ISBN 0 946554 00 5 (set of two volumes)

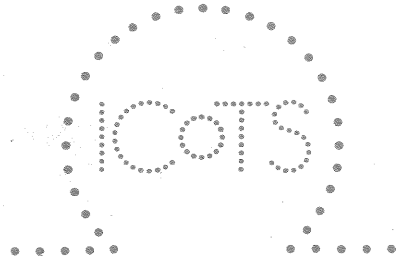
ISBN 0 946554 01 3 (Volume I)

ISBN 0 946554 02 1 (Volume II)

Designed and printed by the University of Sheffield Printing Unit

PROCEEDINGS of the FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS VOLUME I

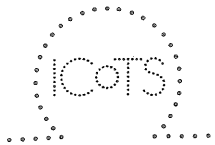
John TAFFE



PROCEEDINGS of the FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS

VOLUME I

University of Sheffield
9 – 13 August 1982



University of Sheffield
9 – 13 August 1982

PROCEEDINGS of the FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS

VOLUME I

Edited by:
Organising Committee of the First International Conference on Teaching Statistics
D R Grey, P Holmes, V Barnett, G M Constable

Published by:
Organising Committee of the First International Conference on Teaching Statistics

D R Grey, P Holmes, V Barnett, G M Constable

© Organising Committee of the First International Conference on Teaching Statistics 1983

ISBN 0 946554 00 5 (set of two volumes)

ISBN 0 946554 01 3 (Volume I)

ISBN 0 946554 02 1 (Volume II)

Printed by the University of Sheffield Printing Unit

Published by Teaching Statistics Trust



Professor Lennart Råde, Chairman of Programme Committee



Professor Vic Barnett, Chairman of Organising Committee

CONTENTS OF VOLUME I

	Page No.
Preface	xi
INTRODUCTORY TALK	
Why Teach Statistics? V Barnett	3
PLENARY TALKS	
Stochastics at the School Level in the Age of the Computer L Råde	19
Optimum Balance between Statistical Theory and Applications in Teaching C Radhakrishna Rao	34
The Relevance of Statistical Training J Gani	50
TOPIC 1: TEACHING STATISTICS IN SCHOOLS (6-11 YEARS AGE GROUP) ORGANISER: PROFESSOR K MATUSITA	
Statistics in the Curriculum for Everybody - how Young Children and how their Teachers react Tamas Varga	71
Teaching of Statistics in Elementary Schools in Japan T Yanagawa	81
Two Examples of Statistics Activities for Elementary School Classes Clare Heidema	91
Using Games to teach some Probability Concepts George W Bright, John G Harvey and Margariete M Wheeler	110
Teaching Statistics in Elementary Schools in Japan H Midzuno	116
TOPIC 2: TEACHING STATISTICS IN SCHOOLS (12-15 YEARS AGE GROUP) ORGANISER: PROFESSOR P L HENNEQUIN	
Why Teach Descriptive Statistics in Secondary Schools? Heinrich Winter	127
The Teaching of Statistics in French Secondary Schools G Dumousseau	144
Teaching Statistics Concepts through Instructional Games George W Bright, John G Harvey and Margariete M Wheeler	158
A Practical Approach to Statistics R Kapadia	169
On the Teaching of Order Statistics K Bognár	179

TOPIC 3: TEACHING STATISTICS IN SCHOOLS (16-18 YEARS AGE GROUP)
ORGANISER: PROFESSOR J SWIFT

Amount and Kind of Statistics to be taught to 16+ year olds
F Eicker 191

Teaching Statistics to 16-18 year olds: Dreams and Reality
T Nemetz 199

Statistics for the 16-18 year olds: a Discussion
A Hanley 208

TOPIC 4: TEACHING STATISTICS TO NON-STATISTICIANS
ORGANISER: DR W G GILCHRIST

We must preach what is practised
A S C Ehrenberg 215

Teaching Statistics to Non-Statisticians - an Application oriented Approach
G Sadasivan 219

Utility through Collaboration: a Problem oriented Approach
D Taylor and G E Skipworth 232

An Open University Service Course
John Bibby 238

The Training of Service Teachers
R W Madsen 250

TOPIC 5: DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL
ORGANISER: MR P HOLMES

Development of Teaching Materials at School Level
A R Hoffer 263

Developing Materials for the Schools Council Project on Statistical Education
P Holmes 269

EDA for 16-19's?
J Bibby 277

TOPIC 6: TRAINING OF TEACHERS IN STATISTICS
ORGANISER: PROFESSOR A P SHULTE

Preparations for College Teaching in Statistics
A K Gupta 283

Teaching Design of Experiments in Indian Universities
V Bhagwandas 287

Clinical Approach to training Statistics Teachers
J Rustagi 298

Discussion of Teacher-Training Papers
A Morley 302

TOPIC 7: USE OF CALCULATORS AND COMPUTERS IN TEACHING
ORGANISER: PROFESSOR L RADE

The Case for using Computers in teaching Statistics
B L Joiner 307

Computer Algorithms for the Classical Distribution Functions
H G Posten 313

Use of hand held Programmable Calculators in Statistics Courses for Nontechnical Students
D R Barr 331

Simulation of Games studied in the Seventeenth and Eighteenth Centuries
J Badrikian 336

TOPIC 8: TEACHING STATISTICS WITH THE HELP OF CASE STUDIES
ORGANISER: PROFESSOR B EPSTEIN

Case Studies for an adequate understanding and interpretation of Results by Statistical Inference
M Borovcnik 341

Is a Scientific Legend a comprehensive Case Study?
V V Shvyrkov and A C Davis 354

TOPIC 9: TRAINING STATISTICAL PRACTITIONERS
ORGANISER: PROFESSOR M BENYAKLEF

Training of Statistical Practitioners
P Osinski 375

Preparing Statisticians for Government Service - a View from the United States
Marie D Eldridge and Katherine K Wallman 389

Programmes d'Enseignement et Couts de Formation
M Benyaklef 403

Les Enseignements Statistiques a l'ENSAE
Ph Tassi et A Trognon 415

CONTENTS OF VOLUME II

Page No.

TOPIC 10: EDUCATION AND TRAINING IN STATISTICS IN DEVELOPING COUNTRIES
ORGANISER: DR M N MURTHY

Statistical Education and Training in Latin America: a General Appraisal
Enrique Cansado 431

Statistical Education and Training in the Developing Countries in Africa
J N Adichie and Biyi Afonja 439

Statistical Education and Training in Developing Countries of Asia and the Pacific
M N Murthy 450

Comments on the above Papers
K T de Graft-Johnson 483

Report on session on Statistical Education and Training in Developing Countries
B P Adhikari 487

TOPIC 11: CO-OPERATION BETWEEN ACADEMIC AND PRACTISING STATISTICIANS
ORGANISER: PROFESSOR W G HUNTER

Co-operation between University and Industry Statisticians in the United States
R D Snee 491

Some Observations from Europe
Lars-Erik Oller 503

Building Bridges between the Academic and Real Worlds - some Observations from South America
C A Peres and P A Morettin 509

Building Bridges between the Academic and Real Worlds
B Jones, B J T Morgan and G B Wetherill 521

Closing the Gap between Statisticians and Statistics Users
Guido E del Pino and Carlos Prado-Campos 527

Co-operation between Academic and Practising Statisticians
A Prat Bartes 536

TOPIC 12: TEACHING SURVEY SAMPLING
ORGANISER: DR D SINGH

The Teaching of Sample Design to Government Statisticians from Developing Countries
Joseph K Garrett 541

Teaching of Survey Sampling
D Singh 552

Teaching of Survey Sampling
P K Bose 566

TOPIC 13: TEACHING DESIGN AND ANALYSIS OF EXPERIMENTS
ORGANISER: DR D A PREECE

The Design and Analysis of Experiments - what is it about?
D A Preece 575

Teaching Experimental Design in the Computer Age
R Mead 591

Teaching Concepts of Design and Analysis using Real and Simulated Experiments
J R Elliott and J C Robinson 603

Teaching Experimental Design: prescribed by a Medical Statistician
Sheila M Gore 627

Contribution to Discussion of the above Papers
B Jones 633

TOPIC 14: TRAINING INDUSTRIAL PRACTITIONERS
ORGANISER: DR R D SNEE

The Polymath Consultant
A A Greenfield 641

Teaching Applied Statistics for and/or in Industry
H C Hamaker 655

Practicals, Case Studies, Projects and Professional Experience: some Methods of developing Skills among Budding Statisticians
G K Kanji 671

Training in Statistical Quality Control
A F Bissell 685

Discussion of Papers by Greenfield, Hamaker and Kanji
A F Bissell 691

Discussion of Papers by Greenfield, Hamaker and Kanji
D G Watts 695

TOPIC 15: PRINCIPLES OF LEARNING PROBABILITY AND STATISTICS
ORGANISER: PROFESSOR E FISCHBEIN

Learning from Experience - Bayes' Theorem: a Model for Stochastic Learning Situations?
H Steinbring and G von Harten 701

Children's Choice Behaviour in Probabilistic Situations
Ruma Falk 714

An Idea of structuring Probability Teaching based on Dienes's Six Stages
Anna Olecka 727

Does the Teaching of Probability improve Probabilistic Intuitions?
E Fischbein and A Gazit 738

Stochastics Teaching based on Common Sense
H-J Bentz 753

A Survey of Probability Concepts in 3000 Pupils aged 11-16 Years
D R Green 766

Misconceptions of Probability, systematic and otherwise; teaching Probability and Statistics so as to overcome some Misconceptions
J M Shaughnessy 784

The Role of Explanation in Teaching Elementary Probability
J L Myers 802

Probabilities: Learning by doing or learning by imitating
A Abele 813

TOPIC 19: TEACHING OF AGRICULTURAL STATISTICS
ORGANISER: DR D SINGH

Training of Personnel for Agricultural Statistics in Jordan
Z Takrouri 831

PREFACE

The first International Conference on Teaching Statistics (ICOTS) took place in the University of Sheffield over the period 9-13 August, 1982. The Conference attracted much interest from all corners of the world. Over 400 full participants from more than 60 countries (accompanied by 100 or so associate participants) spent an active week attending Plenary Sessions, Invited Talks, Contributed Paper Sessions, Workshops and a variety of social events including walks in the Peak District hills and a reception and buffet generously provided by the Lord Mayor of Sheffield on behalf of the City.

The formal programme was very full and covered a wide range of interests in statistical education and training, including the teaching of statistics for all age groups in schools and colleges, the training of practitioners in different applied disciplines, the special problems in developing countries and the difficulties of teaching specific topics. An opening address was given by Professor S. Gregory, Pro-Vice-Chancellor of the University of Sheffield and the closing session was chaired by Mr. H. Ben-Amor, the Director of the UNESCO Office of Statistics.

The Conference was sponsored by the International Statistical Institute and the organisers are grateful to the ISI, and to UNESCO, for financial support. The technical programme and local organisation were arranged by the Programme Committee and the Organising Committee under the respective chairmanships of Professor L. Råde and Professor V. Barnett. The full membership of the two committees was as follows:

Programme Committee		Organising Committee
L. Råde (Chairman)		Professor V. Barnett (Chairman)
E. Cansado	G. E. Noether	G. M. Constable
P. Holmes	J. O. Oyelese	D. R. Grey
B. L. Joiner	B. Penkov	P. Holmes
M. N. Murthy	R. Zielinski	A. Barnett (Conference Secretary)

This Conference arose from the special interests and actions of the International Statistical Institute (ISI): in particular, from the establishment in 1978 of an Education Committee Task Force charged with organising an international conference on the teaching of statistics. The field of statistical education has become a very extensive one and the need for international conferences on the teaching of statistics has been emphasised on several occasions, especially at the many round table meetings on the teaching of statistics arranged by

ISI, and at the international conferences on mathematical education organised by ICMI (International Commission on Mathematics Instruction).

These published Proceedings of the first ICOTS consist of all plenary and invited papers together with contributions from invited discussants, collected together under appropriate topic headings. It is regrettable that the sheer volume of this material has precluded the possibility of including any of the very worthy contributed papers.

It is appropriate to acknowledge the vast effort extended by all members of the Programme and Organising Committees in setting up the Conference and to thank all Topic Organisers for helping to provide an interesting programme. In addition to thanking all authors for their co-operation in submitting papers in suitable form for publication, it is necessary to express particular indebtedness to three people who have given invaluable assistance in bringing this publication into being: Dr. David Grey, who organised the overall preparation of the Proceedings, Mrs. Jenny Cameron, who has undertaken the mammoth task of re-typing all the papers in standard format, and Mr. Bill Seaton, one of the Conference participants, who during an extended period of sabbatical leave in Sheffield, has willingly devoted many hours to the onerous task of proof-reading.

In accordance with views expressed by Conference participants it is hoped that this Conference will be the first of a series of such international conferences on teaching statistics to be held at four-yearly intervals.

V. Barnett
(Chairman, Organising Committee)

Sheffield, December 1982

INTRODUCTORY TALK

WHY TEACH STATISTICS?

V Barnett, University of Sheffield

It might appear a little strange that the opening comments in the First International Conference on Teaching Statistics (ICOTS) should seem to question the very basis of that conference. Such ambiguity of theme is even more perverse from someone who has for more than twenty years been thoroughly immersed in efforts to awaken interest, or change attitudes, to the teaching of statistics at the national and international levels and across the spectrum of school, university and professional work.

For a long time I have been as vocal as anyone in stressing the need for an international conference solely and specifically on statistical education. Of course, the teaching of statistics has featured as a topic of earlier conferences - I would remind you of the invited-participation conference on pre-college statistics teaching at Carbondale, Illinois in 1969 (Råde, 1970). Then there were the round-table meetings held in conjunction with ISI sessions: in Vienna in 1973 (Råde, 1975) and Warsaw 1975 (Breny, 1976) both on school-level statistical education and concerned with general principle more than classroom implementation. A few sections of other ISI conferences have considered specific aspects of teaching or training: e.g. of government statisticians, at the New Delhi session in 1977. Detailed consideration of the teaching of mathematics takes place at the quadriennial meetings of the International Congress on Mathematical Education. Sections at Exeter (1972), Karlsruhe (1976) and Berkeley (1980) were devoted to statistical education, but with very limited coverage in respect of time, topic and level.

The springboard for ICOTS was the New Delhi ISI session of 1977 where a strong initiative developed for a truly international and wide-ranging conference on teaching statistics. The ISI Education Committee was restructured and set up task forces to take positive action on different aspects of statistical education. One of these task forces (under the Chairmanship of Professor L. Råde from Gothenburg, Sweden) was charged with the task of planning ICOTS, and proposed that the ISI should invite the Department of Probability and Statistics at Sheffield University to organise the conference; in acknowledgement of its long-term and wide-spread commitment to the study and development of statistical education. We were very honoured to receive this invitation.

And so we find ourselves together this week to share experiences, to learn new ideas and to hear of individual approaches in the context of a vast range of different national backgrounds (over 50 countries are represented here)

and on a broad canvas of topics and emphases in statistical education. The specialist, and the generalist, is catered for with sessions on all levels of school and college work, development of teaching materials and use of calculators, computers, case studies and so on. We will be considering the special needs of developing countries; the teaching of statistics to non-statisticians as well as statistical practitioners, in industry and in other spheres; the training of teachers, and the teaching of specialised topics such as survey sampling and experimental design. Since I have not been concerned with drawing up the programme I feel that I can congratulate the programme committee on preparing such a veritable feast (might I say orgy) of unremitting delight and enlightenment. I hope it does not lead to indigestion!

This conference is an exciting prospect! I hope it may be only the beginning of a series of international meetings which attempt to bring together the wide assortment of valuable experiences of different individuals and nations in furthering the cause of sound statistical education. Too many separate attempts have been made to 'invent the wheel' in this matter - too many of the invented wheels have been square. Let's begin to work jointly to make a round wheel that we can push along together for mutual benefit. We are approaching a time when every citizen in every country must be given some rudimentary understanding of statistical ideas simply to cope with the need for numeracy in modern society.

The conference comes at an interesting time in the development of statistics and of statistical education. In the United Kingdom we can look back on some distinct landmarks rather interestingly spaced about 25 years apart.

The Royal Statistical Society is to celebrate its 150th anniversary in 1984. About 125 years ago, an energetic and determined woman started a lone crusade for the publication of extensive social statistical data and in particular for the establishment of a professorship in Applied Statistics at Oxford University, ending with a cogent plea on this matter to the famous geneticist Francis Galton almost exactly 100 years ago. She was known at the time as the 'passionate statistician' and never saw the fruits of her plea - in fact, Oxford established its first chair of Statistics in 1948. Some of you may have heard of this lady under a different name : as 'the lady of the lamp' - Florence Nightingale, whose successes in pioneering nursing practice are better remembered than her efforts for statistical education. 1984 also sees the centenary of Karl Pearson's appointment to a Professorship in Applied Mathematics at University College London but it was a further 25 years or so before he was to become the head of the new (indeed first) University Department of (Applied) Statistics. At this same period (about 75 years ago) we were seeing in the U.K. the start of an avalanche of practical statistical applications in such vital areas

as biology, agriculture and industry. It was around 50 years ago that the statistical world began to reverberate to the critical conjunction of Jerzy Neyman and E.S. Pearson. Both were engaged at University College London on their crucial work on the Neyman-Pearson theory of inference and E.S. Pearson took over the headship of the Department of Applied Statistics in 1934 (and soon also found the same hallowed corridors stalked by the famous R.A. Fisher in the role of Galton Professor of Eugenics). Even the presence of statistics at school-level is approaching its quarter-centenary and the single statistic that 80% of secondary schools in England and Wales now teach some statistics is an impressive, if somewhat arbitrary, reflection of changes in attitudes and emphasis to statistics as a subject of study.

Of course, different nations have had different experiences. In Hungary, statistics (or at least probability and combinatorics) have been in and out of the secondary school curriculum for over 130 years, whilst it would appear that Japan (in spite of the valuable contributions its nationals make to statistical theory and practice) has no separate curriculum for statistical education at school-level (some topics are taught in the mathematics lessons) and no independent departments of statistics in the Universities. I look forward to hearing in the later sessions of recent developments in Japan. Statistics does not feature even in the school mathematics courses in many countries, whilst in contrast in Sudan 'basic statistics is in the syllabus of a subject that must be passed by a student wishing to enter [any subject of study in] higher education' (Downham and Gattar, 1982). Unfortunately it is all too easy for pupil and even teacher to ignore the statistics component. My role as Chairman of TOTSAS, the ISI taskforce on teaching statistics at school level, has given me a unique opportunity to witness the enormous range of different attitudes to, and levels of incorporation of, statistics in schools in different countries. We have just produced a book entitled Teaching Statistics in Schools throughout the World (Barnett, 1982). This was published by the ISI only in the last week or so and copies are available at the conference should you wish to explore some of the widely assorted systems and practices in different countries.

But let me return to my main theme. In the light of the various positive remarks I have just made, can there really be any doubt of the value of teaching statistics or of holding a major conference on such a topic. As far as this gathering is concerned the answer must surely be a firm NO : we must teach statistics more comprehensively, more widely, more wisely, more realistically, more everything!

As a secondary-school class-room teacher of statistics you will be very concerned to extend and improve (or even just to introduce) the teaching of the subject. As a teachers-training-college lecturer in statistics you will be anxious to broaden the exposure of student teachers so that their influence can be felt

in the schools. If you are concerned with statistical training in medicine, industry or government you will want to ensure that practitioners can contribute soundly on important and real problems and that their organisations are well aware of what the statistician has to offer. If you are engaged in statistical curriculum development at any level you will be anxious to improve teaching materials and syllabus content and to disseminate the results of your work to a wide audience.

Notice, however, that the emphasis in my remarks is on what you as a statistical educator will want to do, and statistics is seldom a subject which controls its own destiny at any level or in any educational or national environment. It is a very different, and very crucial, matter to establish the need for statistical instruction or for the use of statistical techniques in the wider framework of a school curriculum, in applied subjects at college level or in the operations of a business or government organisation. Why should we teach statistics? - is very much the automatic response to any proposal to do so, and it is a natural and proper response. There is no intrinsic reason why a school or college should offer statistics or why an individual should wish to study it.

I am sure that many of you have had experiences which are not so different from mine in this matter.

As a mathematics undergraduate my head of department (one of the most eminent pure mathematicians of his generation) repeatedly asserted that statistics was an inferior subject not worthy of the attention of anyone with claims to mathematical ability. It was this attitude which was paramount in my decision to become a statistician! Teaching at university level I have had to cope with many mechanical engineers, agriculturalists, physicians (and so on) who saw no good reason to attach importance to service courses in statistics for their students. The grudging acceptance of a 10 lecture course on the whole of statistics (including Markov chains and factor analysis) is very much a feature of applied departments at college level or even in teacher-training. And as far as the 'man on the Clapham omnibus' is concerned the lay attitude to statistics is one of suspicion

'You can prove anything by statistics'

'Lies, damn lies and statistics' etc.

This is a worrying state of affairs, since so few areas of social or personal life can afford to ignore what statistics has to offer - the very reading of a newspaper or watching of a television programme or of advertisements requires some degree of statistical understanding. And of course such an attitude of resistance or even of hostility is bound to be reflected in the way in which the educational system responds to the incorporation of statistics. Some countries

are beginning to see an easing of such suspicion and a growing acceptance of the need for statistical instruction at different levels, whether this is for the production of statistical experts or merely to 'arm the citizen against the media'. But in contrast, some countries have not even reached that level of public awareness of statistics for any resistance to have yet developed. They face the double problem of trying to introduce ideas of statistics at the same time as reassuring their society of its respectability.

Such considerations imply that the question: Why teach statistics? must be a crucial one in any attempt to introduce or extend the teaching of statistics, or in any effort to encourage widespread use of statistical methods. the onus of justification must rest with the statistician; as practitioner, teacher or educationalist. And in my view that justification must be based on one clear principle - that statistics is a practical discipline for understanding the indeterministic world that we live in and for solving the real problems in society from agriculture, through meteorology to zoology - from A to Z! The message is not that statistics might be just another topic of possible study, but that it is central and essential to the education and training of the school-child, the university student, the doctor or even the linguist. The statistician has to be more evangelical in this matter, whether he is a junior teacher in a mathematics department trying to persuade sceptical colleagues to see the place of statistics in the mathematics programme, or whether he is trying to influence central governmental attitudes on the place of statistics in the school curriculum or in society generally.

Thus the main response to the question is 'because we cannot afford not to', and we have to persuade others (including some statisticians and mathematicians) to believe this. But the question has a much more fundamental role for the teacher of statistics. Whether contemplating how you as a teacher will present a particular syllabus, or as an educationalist what recommendations you should make for course design, training programmes or curriculum development, the starting point has to be the same. Before considering how to teach some prescribed syllabus or even what should be the content of a syllabus, we need the clearest possible notion of just why we are teaching statistics at this level, to this group of students etc. Only after we have faced this basic question can we hope to make sense of such matters as the choice of content, teaching method, reinforcement technique, motivation and assessment.

Take a simple example - the teaching of the binomial model and distribution. The approach will depend on many factors. Is it the first introduction of the topic to 14 year olds in a general studies class or to 18 year olds in a mathematics lesson, or is it part of the statistical methods component of a university geography,

or sociology, course; is it in a statistics course for secondary school student-teachers or does it feature in a refresher course on quantitative methods for doctors or for junior managers in industry? The level and method of presentation must be carefully matched to the student group in respect of their maturity and current state of knowledge; it makes no sense to use the binomial distribution as a vehicle for teaching the mathematics of the weak law of large numbers to 14 year olds in a general studies, or even mathematics, lesson. But again, even more important is the question of why such a group, at such a level is being taught such a topic : the binomial distribution.

I can recall my experiences in the mid-sixties as an examiner on a Pure Mathematics with Statistics A-level GCE examination (the public examination of 18 year olds in the U.K.). In various years there would be a question which might typically start

"Define the binomial distribution..."

and then go on to ask about its practical application, or for the solution of a practical or numerical exercise involving the binomial distribution. It was a depressing experience to find candidate after candidate responding:

"The binomial distribution is the expansion of $(p+q)^n$..."

What has such an answer to do with probability and statistics? Obviously such an irrelevant response was reflecting how the candidates were taught the topic (not in all schools, but uncomfortably many) - with no regard for the nature of a probability distribution or for its crucial role in modelling real problems. Instead, it was being presented as a convenient tool to illustrate a mathematical expansion. Why should this attitude be taken? We can only hazard a guess, but my conclusion was that it was not due to a conscious desire to conceal the importance of the binomial distribution as a model - to deliberately manipulate its use to the service of some mathematical exercise. Sadly it was because no-one (curriculum developer, examiner or teacher) had provided a clear enough answer to why the binomial distribution was being taught in the syllabus. Too few teachers at that time could provide the answer from their own education or practical experience - they probably had to keep one step ahead of the class from some equally formal and inadequately motivated textbook. (The dual problems of teacher-training and availability of training aids are vital - I am sure they will figure widely in the discussions at this conference).

But let us stay for a moment with the example of teaching the binomial distribution and see how different purposes in teaching this topic produce quite distinct approaches to suit different needs. We will go through the groups in order.

i) '14 year-olds in a general studies class'. For this group there can be no question of any mathematical treatment. We need to awaken interest through many practical examples (winning or losing a game of tennis, answering 'yes' or 'no' to a question, passing or failing a test, 'rain' or 'fine' tomorrow) in the notion of random variation in a dichotomous situation : that in six games of tennis the number of wins is random (it might be six, or one, depending on who plays who, and on what occasion!). Statements such as '8 out of 10 dogs prefer Dognosh' or 'one in every four children is Chinese' begin to make sense and even primitive notions of estimation develop : 'If 90 people out of a sample of 102 have a radio, perhaps about 90% of the population do so'.

(This sort of awakening of statistical understanding for 11-16 year olds across the curriculum, for all abilities, and tied totally to everyday problems, was the basis of the Schools Council Project on Statistical Education which was directed over 6 years in Sheffield by Peter Holmes. The publications include 27 teachers' and pupils' units, 'Statistics in Your World', and books entitled Statistics in Schools 11-16 : a review (Holmes et al, 1981) and Teaching Statistics 11-16).

ii) '18 year olds in a mathematical lesson' can cope with a more mathematical (algebraic) treatment - with the notion of a distribution and its mean, and an estimator. But again it is crucial that the empirical basis is stressed and illustrated and that the model can be applied to the study of real problems (whether cars on ferries, or bacteria in test tubes). I have no sympathy at this level with probability or statistics being exploited merely to illustrate mathematical concepts.

The most grotesque example I recall of such use was a GCE A-level question along the lines:

A continuous random variable X has probability density function

$$f(x) = a + bx + cx^2 \quad (0<x<2).$$

The mean and standard deviation of X are 8/5 and 1/5, respectively.

Determine a, b and c.

This is purely an exercise in solving three linear equations, from which we obtain a = 3/20, b = 3/20, c = 15/40. Who would want to use such a probability model (especially with f(x) negative for $x < 0.46$).

This is not to deny the importance of mathematics in statistics - it is vital. I merely question the teaching of statistics as if it were mathematics.

But let us move on briefly through the other groups:

In the

iii) 'statistical methods component of a university geography course', the binomial distribution has to be seen to be important to the geographer. Examples for motivation and analysis must be specific.

There are some teachers of statistics who hold the view that it should not be necessary to choose practical examples which are closely related to the interests of the students - they would claim that statistics is so clearly important that whether we are teaching an economics student or an engineering student the same example on the distribution of earthworms will serve equally in both cases. I totally disagree with this view; my experience convinces me that fully specific practical illustrations are vital and must be carefully chosen to add excitement and relevance to the teaching. We can illustrate this for those other groups for whom we contemplated using the binomial model.

In the case of

iv) 'statistical methods for university sociology students' the typical problem is to generate some interest in what students may feel is a rather tedious short extra course. A useful vehicle could be the idea of randomised response or the 'principle of the irrelevant question'. For example, we might describe a survey on some sensitive issue - the use of a drug like marijuana or 'pot'. To ask respondents directly 'do you smoke pot?' invites a dishonest or indignant reply. Instead we build up confidence by playing a game with a randomising device e.g. a box with red and green balls, where just one appears in a window at random. We then ask the respondent to answer a questionnaire such as the following

---

SHAKE THE BOX AND OBSERVE THE COLOUR OF THE BALL

IF RED, ANSWER QUESTION 1; IF GREEN, ANSWER QUESTION 2

REPLY YES OR NO.

Q1 IS YOUR BIRTHDAY IN APRIL?

Q2 DO YOU SMOKE POT?

YES/NO

Apart from the practical advantages of instilling confidence through confidentiality and avoiding legal problems of having specific information about the individual, this provides a most compelling use of the statistics topic; the binomial model. If we know the proportion of red and green balls, and of birthdays in April, we obtain an immediate binomial estimate of the proportion who smoke 'pot'!

Peter Holmes, on his Schools Council Project for 11-16 year olds, tells me that he used this approach most successfully with a class of 12 year olds on the issue of smoking cigarettes during the school lunch break. A quick estimate (incorporating a bit of luck) of precisely 4 smokers led to 4 red faces and a class clamouring to hear more and more about statistics!

For the

v) 'secondary school student-teachers' the challenge is a double one of having the experience to set the topic in any of a range of practical situations (depending on need) and the skill to teach the topic in a compelling way within that situation. All the teacher needs is a degree in statistics, 3 or 4 years practical experience in two or three different fields of application, a postgraduate qualification in education and a charismatic personality! Unfortunately, all the teacher often possesses is uncertainty bred of a poorly presented 10-20 lecture introductory course at College and an inadequate text book!

For the

vi) 'doctors' or 'junior managers' the task is much easier in principle - although skill is still required on the part of the instructor. The audience is already self-motivated. As practitioners they will respond to anything that increases their power to operate. So the binomial distribution needs merely to be offered in this spirit : e.g. as the basis of clinical trials analysis for the doctor or of sampling inspection techniques for the industrialist. This works, as long as you convince the audience that you really do know what you are talking about in practical down-to-earth terms.

So how am I to answer the question : why teach statistics? Essentially in the following way!

We are surrounded with problems - social, economic, scientific, personal. Almost without exception they involve indeterminism, the element of chance. To sovle such problems we need to understand probability and statistical method. We teach statistics to give the individual (depending on need and ability) a level of understanding or expertise in this matter of problem-solving. It might be thought a simple answer, but if whole-heartedly accepted it has profound implications for a wide range of further considerations such as - who should teach statistics, to what groups of pupils or students, for what purpose, on what syllabus, etc. To spell out these implications in detail I would need to write a book, rather than to present an introductory talk.

Consequently, I shall conclude by taking one or two specific topics or themes germane to the different sessions of the conference as a basis for some purely illustrative comments.

A. Teaching Statistics in Schools (12-15)

Society is placing ever growing demands on citizens in respect of their numeracy their ability to assimilate and interpret quantitative information. Such information is becoming increasingly statistical in nature and statistical edcuation for this age range must face up to this need. We should not be aiming to produce precocious statistical experts but should be seeking to instill, across the curriculu and ability range, some basic understanding of statistical reasoning and informal inference. This is best approached through real practical problems from everyday experience. The problems are manifold. Where do we find appropriate teaching material? Who can cope with such an empirical, look-and-learn approach to teaching? How is it to be fitted into the school programme? The Schools Council Project on Statistical Education had just these problems to face in producin its units for pupils and teachers aimed at arming the future citizen to cope with the outpourings of television, newspapers, government, etc. (See Schools Council Project on Statistical Education, 1980a, 1980b; Holmes et al, 1981).

Some groundwork may have been laid in the primary schools (6-11 year olds) in respect of collection, tabulation and presentation of data, and the citizen-arming process will pave the way to the variety of needs at the next level (16-18 year olds - see below). One suggestion taken up in the recent report of the U.K. Committee of Enquiry into the Teaching of Mathematics in Schools (Cockcroft, 1982, Mathematics counts) is an interesting one - that there might be a 'statistics coordinator' in a school who acts as a focus for the statistics teaching across different subjects and ability levels.

B. Teaching Statistics in Schools (16-18)

Here we have a new dimension to face. There are quite different needs for distinct groups of pupils. It is common to think that the mathematically oriented and able are the important ones to cater for. But they represent a minority and are the easiest to deal with. If they have had appropriate earlier motivation, they can be given a more mathematical consolidation prior to further studies at college or unviersity. Their syllabus is relatively straightforward - as is choice of teaching method (still with empirical reinforcement) and there is hope of finding at least some mathematics teachers who can present the material appropriately. One danger for this group is the universal overestimation of how much material can reasonably be presented. Syllabuses, which their designers claim to be of reasonable extent, can often on detailed scrutiny be shown to contain as much as 100% more implicit content than a parallel (and supposedly equal) course in some other subject e.g. physics.

But there are two other larger groups at this age which present a greater challenge and are even less well provided for. There are those in schools and colleges who are either studying purely qualitative or craft/vocational subjects who need a more sophisticated version of the 'citizen-arming' approach proposed for the 12-15 year olds, and, in contrast, those who are studying quantitative or technical subjects (including economics, business studies; even biology and geography) who need to be introduced to relevant statistical method without unnecessary mathematical frills and taught with copious reference to real-life data-based problems. But again, we must ask whether we have teachers capable of teaching in this way, at this level, in the different subjects and whether we have appropriate teaching material? A start on this problem in the U.K. is being made in the new Sheffield-based Leverhulme project for 16-19 year-olds studying business or technological subjects in Colleges of Further Education; or economics, biology, geography, etc. in schools. As with the earlier Schools Council Project, the Leverhulme Project (STEP 16-19) is again being directed by Peter Holmes. A detailed and lengthy pre-project report on the education and training of 16-19 year olds was prepared two to three years ago for the Centre for Statistical Education in Sheffield University (see Barnett et al, 1979).

C. Teaching Statistics to Non-statisticians

The question of who should teach statistics 'service courses' has always been a thorny one - should the biologist teach statistics to his own students or invite the statistician to do so? Relevant illustrative material is vital; so is sound well-constructed statistical argument. At school-level the former consideration (and time-tabling problems) may support the subject specialist; beyond school level I believe it is the statistician who can do the better job provided time is invested in accommodating the special subject realistically. The major policy committee in our university recently concluded 'that all teaching, provided it is directed to the particular needs of its recipients, is best conducted by those who are specialists in the subject'.

D. Training of teachers, development of teaching materials, use of teaching aids

Here we have absolutely central considerations which feature widely in the conference. Any approach to teaching statistics, but particularly the empirical practical style I favour, depends on quality in these three matters. Sadly, no country can feel particularly proud of its efforts in teacher-training or materials-development for statistics, and much investment is needed. There are real problems in teacher-training, not least of which is how to find enough time (in training or in service) to convey the right attitude and experience to potential teachers

of statistics. There is also a fundamental problem : statistics is not the usual linear discipline where we teach by placing one brick of understanding on top of another already firmly placed one. Instead we must persuade teachers to adopt a 'window-gazing' approach - first showing something of the breadth of application of statistics and gradually focussing down on more and more specific aspects as students develop gerater confidence, familiarity and experience. This is no easy task.

As far as teaching-materials are concerned, we need to produce high quality well-reasoned soundly-motivated texts and resource matter, and must not hesitate to discard much inadequate matter which is hampering progress. We must also be alert to an insidious problem : calculators and computers can be important in teaching statistics, but they are merely tools, not teacher nor text!

E. Curriculum Development

The 1960's and 1970's saw an avalanche of activity in curriculum development across all subjects. Money flowed like water (or at least crude oil) and some of it even slopped over statistics. It seems unlikely that we shall return soon to a liberal financial climate in education and yet we have much to do in the development of syllabus and courses (or, more fashionably, programmes) in statistics. I see a major difficulty here. Shortage of financial support for research and development in statistics will certainly restrict the detail in which courses and materials are produced and virtually eliminate vital stages of testing, validation, reappraisal etc.

To put this in perspective we might note that the two books and 27 teaching units of the Schools Council Project on Statistical Education took about 20 man years of effort. Depending on your stand-point it cost about:

1 000	Gold Krugerrands
4 000	Gold Sovereigns
110 000	Dinar
220 000	UK£
400 000	US$
1 000 000	DM
2 500 000	F Francs
6 600 000	Schillings
42 000 000	Pesetas
100 000 000	Yen
500 000 000	Lire

- a lot of money by any standards!

But the reduction of support funds will also encourage the 'statistics on a shoestring brigade'! In my capacity as Chairman of the ISI taskforce TOTSAS I have been involved in many discussions in different countries on how they might introduce a programme of statistics teaching in schools. On three separate occasions the attitude was identical - to call together a small committee representing university statisticians and the national statistical body (and perhaps an odd teacher) and to ask them to meet on 3 or 4 occasions to draw up 'the statistics syllabus'. The result is always the same (we saw it in the U.K. twenty five years ago!) A page of topic headings such as might describe any introductory course at any level, for any group and with any emphasis, which is repeatedly misinterpreted and misapplied until a decade later someone asks 'where did we go wrong?' The error in my view was in not asking and seriously answering the very question with which I started this talk - just why are we wanting to teach statistics in the first place? Let me leave you to think further on this matter. Whether or not you agree with my answer to the question, I hope that it might provide a springboard for discussion and debate in many of our sessions throughout the conference.

References

[1] Barnett, V. (Ed) (1982) Teaching Statistics in Schools throughout the World. International Statistical Institute; Voorburg, The Netherlands.

[2] Barnett, V., Cable, P., Conway, D.A., Harris, J., Holmes, P., Mann, E.J., Morley, S.A., Oldfield, S.J. and Wilkinson, R. (1979) Statistical Education and Training for 16-19 year olds. Centre for Statistical Education, University of Sheffield.

[3] Breny, H. (Ed) (1976) The Teaching of Statistics in Schools. International Statistical Institute; Voorburg, The Netherlands.

[4] Cockcroft, W.H. (Ed) (1982) Mathematics Counts. Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr. W.H. Cockcroft. HMSO, London.

[5] Downham, D.Y. and Gattar, H. (1982) Chapter 12 of Barnett (1982).

[6] Holmes, P., Kapadia, R., Rubra, G.N. and Turner, D. (Ed) (1981) Statistics in Schools 11-16: a review (Schools Council Working Paper 69). Methuen Educational, London.

[7] Råde, L. (Ed) (1970) The Teaching of Probability and Statistics. Almqvist and Wiksell, Stockholm.

[8] Råde, L. (Ed) (1975) Statistics at the School Level. Almqvist and Wiksell, Stockholm.

[9] Schools Council Project on Statistical Education (11-16) (1980a) Teaching Statistics 11-16. Foulsham (for the Schools Council), Slough, U.K.

[10] Schools Council Project on Statistical Education (11-16) (1980b) Teachers and Pupils Notes. Statistics in Your World. Foulsham (for the Schools Council, Slough, U.K.

PLENARY TALKS

STOCHASTICS AT THE SCHOOL LEVEL IN THE AGE OF THE COMPUTER

Lennart Råde, Chalmers University of Technology, The University of Gothenburg

INTRODUCTION

Programmable machines, from programmable calculators and microcomputers to large computer systems, are extremely important in statistical work today. This is one reason why the use of such machines should be considered in stochastics (probability and statistics) courses both at the school and the college level. Another, maybe even more important, reason is the possibility these machines offer for a concrete and experimental approach to the teaching of probability and statistics, especially at the school level. The programmable calculator is especially important in this respect because of its low cost (comparable to that of a book) and its availability. It allows each student to have access to her or his own "pocket computer".

In this paper mainly two examples will be used to demonstrate the flavour of a computer oriented approach to the teaching of stochastics at the school level.

THE TWO TOWERS, AN INTRODUCTORY EXAMPLE

Two towers are built of m and n blocks. A tower is chosen at random and a block is moved from this tower and put on top of the other tower. This procedure is continued until all the blocks have been moved from one tower to the other. We are interested in the total number of times a block has been moved from one tower to the other tower.

This random experiment can be simulated according to the flow charts in Figure 1 and Figure 2. The first flow chart is for a programmable calculator and the second flow chart describes a computer program in BASIC.

With a program according to these flow charts successive trials can be made of the experiment with the two towers.

With such a program the following observations were obtained. Here the values m=3, n=3 were used.

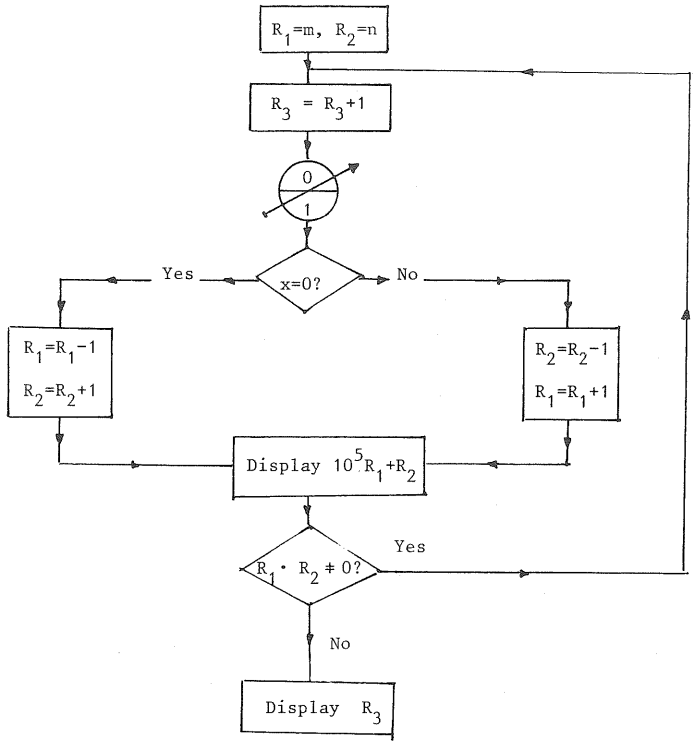


FIGURE 1

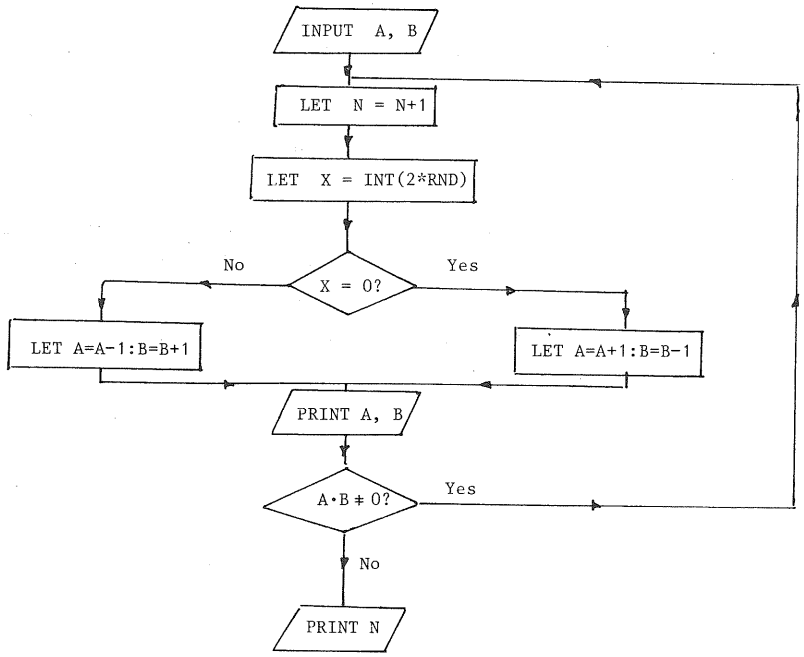


FIGURE 2

7	3	7	3	5	5	9	7	27	33
31	5	19	23	7	17	11	3	5	15
3	7	3	3	5	3	9	29	5	5
5	3	5	5	3	9	3	37	3	3
9	11	13	9	11	31	11	15	5	5
3	3	11	7	3	25	9	15	5	5
5	5	5	7	5	9	3	9	9	7
9	7	5	5	5	19	5	9	17	9
3	3	5	3	25	25	15	5	5	15
3	7	7	5	5	5	19	15	9	3

These data can be summarized by a bar graph or a box plot, see Figure 3.

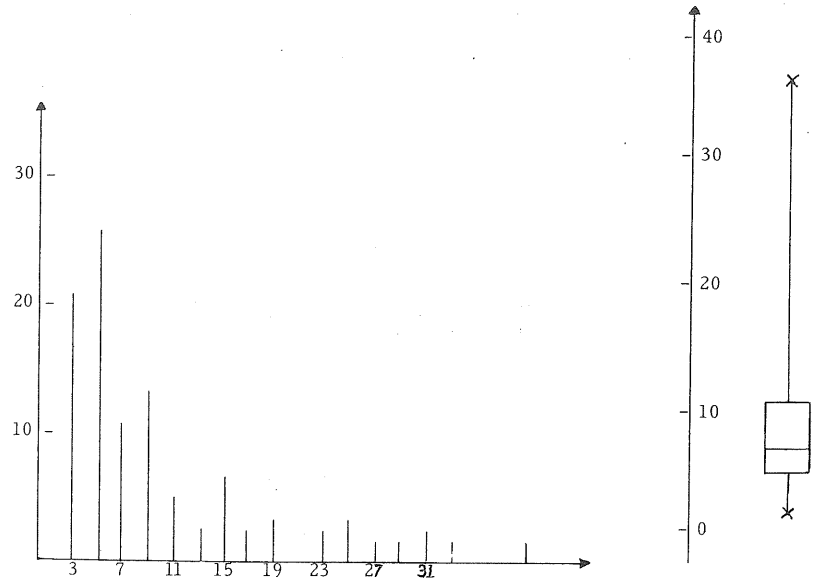


FIGURE 3

Also the mean can be calcualted. It is found to be 9.4. All this provides us with some information about the random variation of the number of moves in this case.

It is easy to modify a program according to the flow charts in Figure 1 and Figure 2 so that it will do a number of trials and calculate the average number of moves in these trials. With such a program the following averages were obtained, each time from 100 trials and with n=m=3.

10.86	8.52	9.84	9.96	9.26

Thus the mean in a large number of trials seems to be in the range 9-10. Running the program with 1000 trials gave the mean 8.838.

More insight could be obtained by doing an investigation with different values of m and n.

The following table gives the results from such an investigation where the mean in 100 trials was obtained for some different values of m and n.

m \ n	1	2	3	4	5	6
1	-	1.83	3.06	3.55	5.16	6.21
2	-	4.06	6.20	7.48	9.77	12.04
3	-	-	8.56	10.25	13.94	17.08
4	-	-	-	15.12	21.79	23.30
5	-	-	-	-	24.62	31.90
6	-	-	-	-	-	32.50

In Figure 4 these results have been graphed. Averages for m≠n have been plotted twice by interchanging m and n.

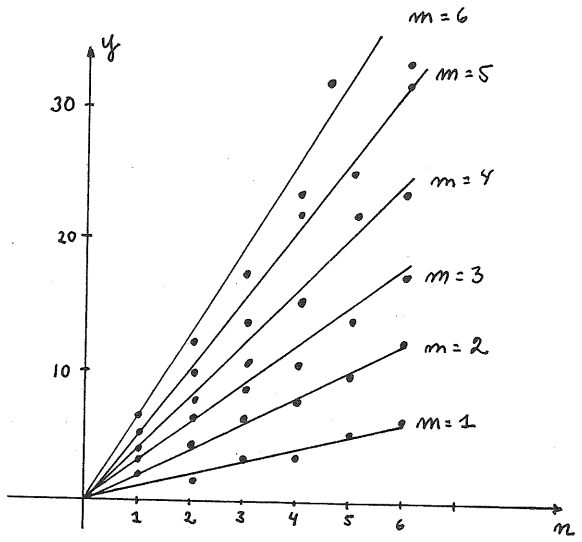


FIGURE 4

These results indicate that the mean in a large number of trials (the expectation) for given value of m is proportional to n. As a matter of fact the results indicate that the expectation is equal to the product m·n. In Figure 4 also the lines m·n have been drawn.

Is it possible to prove theoretically that the average in a large number of trials (the expectation) is m·n? Let us look on the case m=n=3. Then the random experiment can be described by the graph in Figure 5.

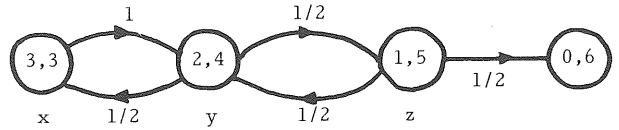


FIGURE 5

Let x, y, z be the expectations from the different states according to Figure 5. It is then easy to derive the following equations.

$$x = 1 + y$$
$$y = 1 + \tfrac{1}{2}x + \tfrac{1}{2}z$$
$$z = 1 + \tfrac{1}{2}y$$

This gives

$$x = 9 = 3.3 \qquad y = 8 = 2.4 \qquad z = 5 = 1.5$$

A similar analysis is possible in the general case.

Let us do one more thing before we leave this introductory example. It is easy to modify a program according to the flow charts in Figures 1 and 2 so that it will do successive trials and calculate the mean of the observations obtained so far after each trial. The results in Figure 6 have been obtained in this way.

The problem which has been discussed here is a classical problem in probability theory. Usually it is called the problem about "the duration of a game". In this context the problem has to do with two players, who make successive trials of a fair game. Initially they have m and n monetary units and after each trial the winner gets one such unit from the other player. After how many trials will one player be ruined?

This problem may look somewhat remote from practical interest. So let us consider a more direct practical application.

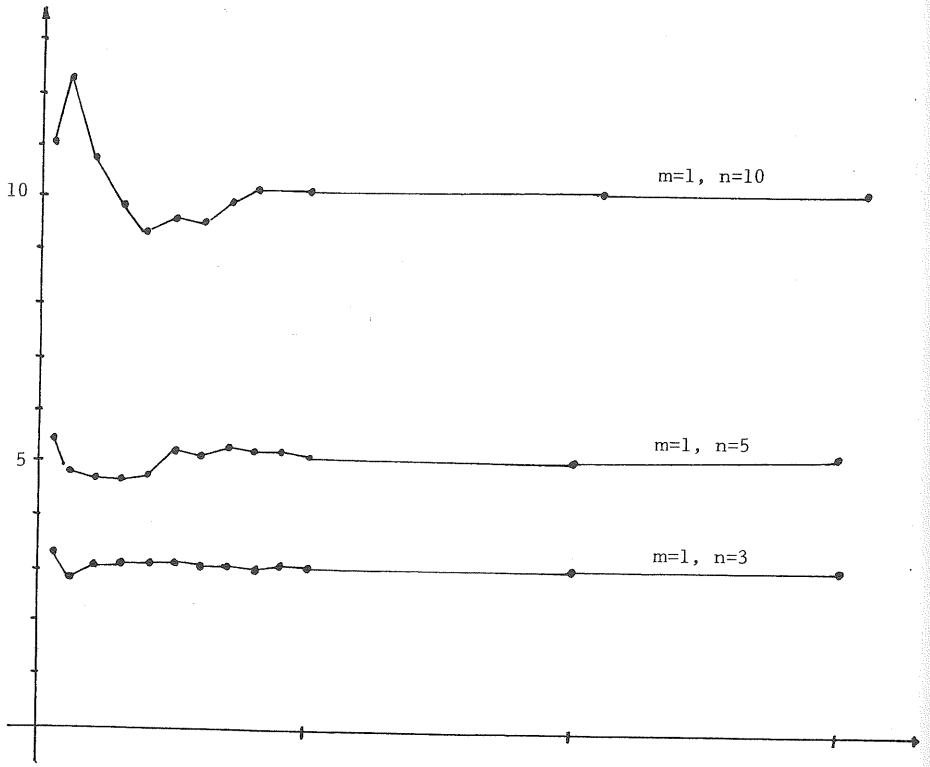


FIGURE 6

RELIABILITY AND COIN TOSSING

Consider the following random experiments dealing with the tossing of coins. First n coins are tossed simultaneously. Those which come up heads are removed and those which show tails are tossed again. The process continues until each coin has come up heads.

We are interested in the number of (group) tosses to accomplish this.

This again looks like a typical problem from a textbook in probability, charming but not dealing with a real life application. The problem, however, has an interesting application concerning the reliability of a system.

Consider a parallel system which consists of n components. The system is subject to shocks which cause the components to fail independent of each other with probability p. Each shock will remove a random number of the components like each toss in the previous problem will remove a random number of coins. We thus have the same problem if we let the coins give the outcome "head" with probability p.

This random experiment can be simulated with programs according to the flow charts in Figure 7.

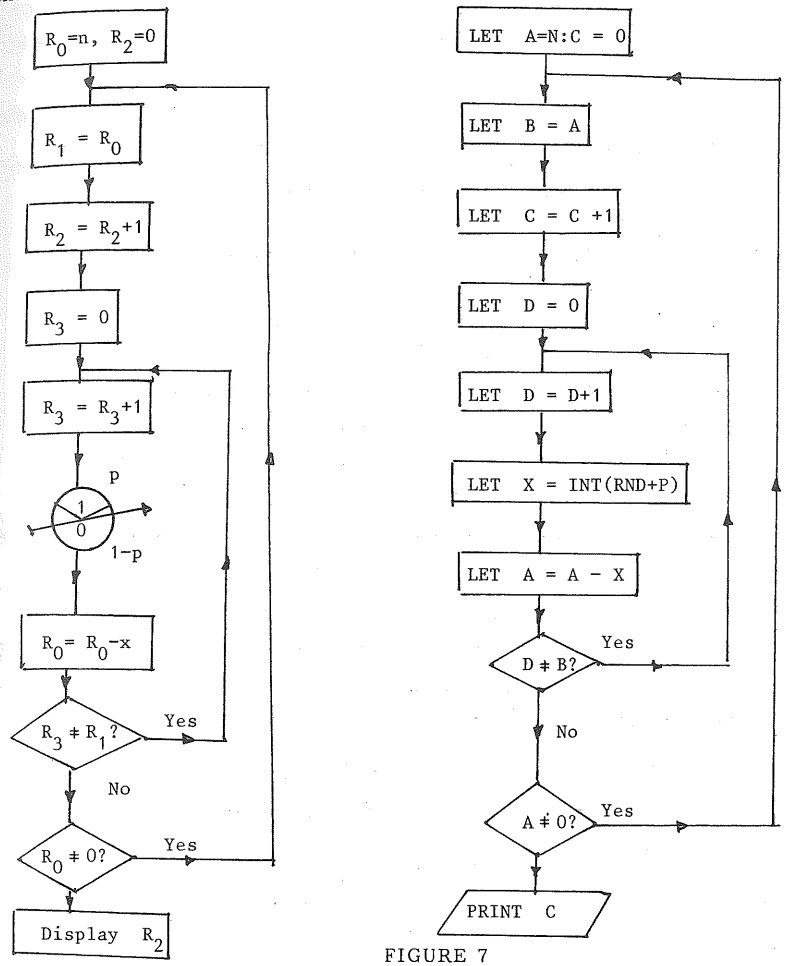


FIGURE 7

Registers and variables are used as follows in these programs.

R_0, A = number of components which have not failed.
R_1, B = number of components which are hit by shock
R_2, C = number of shocks
R_3, D = variable which counts the components in the system when hit by shock

The programs give the length of life of the system in terms of number of shocks until failure of the system.

The following lengths of life have been obtained for two specific cases with a program according to Figure 7.

n=2, p=0.3	n=5, p=0.1
4, 1, 2, 3, 5, 1, 5, 4, 1, 5	15, 8,11,26,13,17,12,21, 3,11
3,16,14, 4, 3, 1, 2, 6, 9, 2	22, 9,48,21,22,16,26,26,18,35
12, 1, 7, 3, 2, 5, 5, 1, 1, 2	21,22, 5,26,46,20,43,20,29,25
3,19, 7, 1, 2, 7, 8,10, 2, 8	16,10, 9,27,17,24,39,28,19,23
9, 3, 6, 4, 5, 4, 2, 6, 4, 8	11,17,22,14,17,12,55,16,31,19
1, 5, 1, 7, 9, 9, 3, 2, 6, 2	16,43,43,21,13,10,14,49,32,20
4, 3, 5, 4, 6, 2, 8, 6, 8, 1	16,16,17,34, 8,15,17,31,15,14
7, 5, 2, 3, 5, 1,19, 8, 1, 1	29,10, 9,40, 7,21,38,26, 5,10
4, 4, 4, 5,13, 9, 2, 3, 2, 9	11,59,28,33,15,17,28,20,16,16
6, 5, 2, 3, 6, 2, 3, 6, 2, 3	40,17,30,14,19,38,12,14,15,15

The 50 first observations in the two cases are shown in the following stem and leaf diagrams.

n=2, p=0.3			n=5, p=0.1		
0*	12311331221321123122 32	(22)	0	83959	(5)
0·	455454655645464	(15)	1	5137216860791747269	(19)
0··	97778898	(8)	2	6121266126009574832	(19)
1*	20	(2)	3	591	(3)
1·	64	(2)	4	863	(3)
1··	9	(1)	5	5	(1)
		(50)			(50)

* 0,1,2,3.

· 4,5,6

·· 7,8,9

The length of life distribution is obviously very different in the two cases.

A program according to the flow charts in Figure 7 can easily be modified to give the average length of life in a number of trials. With such a program the following data have been obtained. In each case 100 trials have been made.

Average length of life in 100 trials						
	p					
n	0.05	0.10	0.15	0.20	0.25	0.30
1	21.67	10.23	6.89	5.13	4.12	3.48
2	29.24	15.20	10.27	7.13	6.39	5.10
3	41.34	16.55	11.81	8.29	6.47	5.60
4	38.75	18.64	13.77	9.58	7.64	6.13
5	44.69	22.59	14.92	10.56	8.57	6.85
6	47.71	24.59	14.83	12.06	8.83	7.39
7	53.54	26.41	16.99	12.37	9.63	7.60
8	52.90	23.96	17.86	13.17	9.66	8.13
9	55.52	25.67	18.17	12.40	10.65	8.94
10	57.94	28.39	19.45	14.10	10.96	8.85

Let us now take a closer look on the case n=2. Then the situation can be described by the diagram in Figure 8.

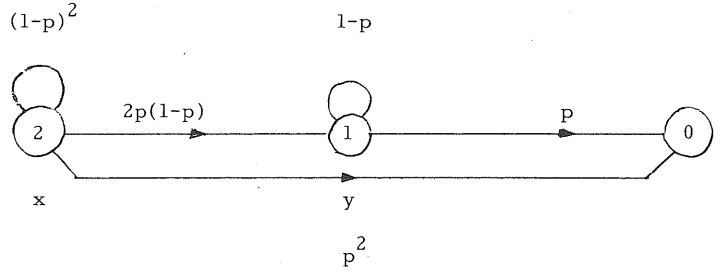


FIGURE 8

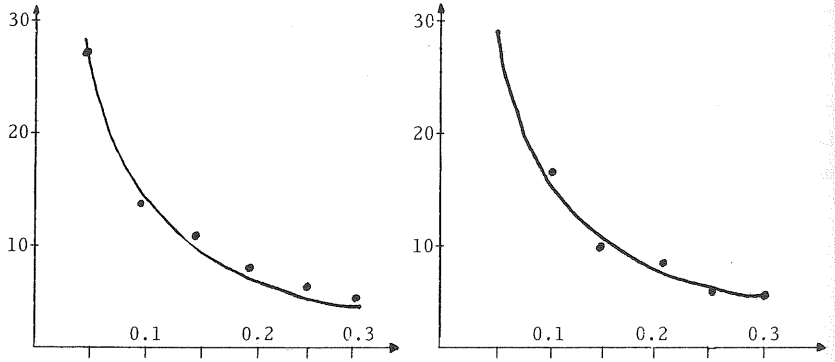
Let x and y be expectations from states according to Figure 8. Then

$$x = 1 + (1-p)\overset{2}{x} + 2p(1-p)y + p^2 \cdot 0$$
$$y = 1 + (1-p)y + p \cdot 0$$

This gives $x = (3-2p)/p(2-p)$, which is the expected length of life of the system for n=2. The following table shows this expectation for p=0.05(0.05)0.30 and also means obtained in 100 and in 1000 trials.

p	Expected system life	Average in 100 trials	Average in 1000 trials
0.05	29.74	29.283	
0.10	14.73	14.07	15.158
0.15	9.73	10.88	9.483
0.20	7.22	7.68	7.231
0.25	5.72	5.81	5.624
0.30	4.71	4.73	4.728

These results are also shown in Figure 9.



Expected system life and average life in 100 trials

Expected system life and average life in 1000 trials

FIGURE 9

It is seen that already 100 trials give a very good estimate of expected system life and that 1000 trials usually (but not always) give a better estimate than 100 trials.

Let us now try to investigate the case n=5 in the same way. The situation is now illustrated by the graph in Figure 10, where, however, we have not indicated the probabilities for transitions between states.

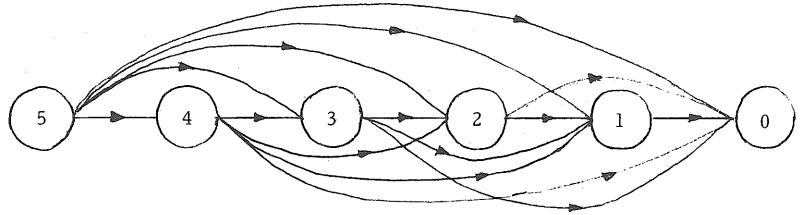
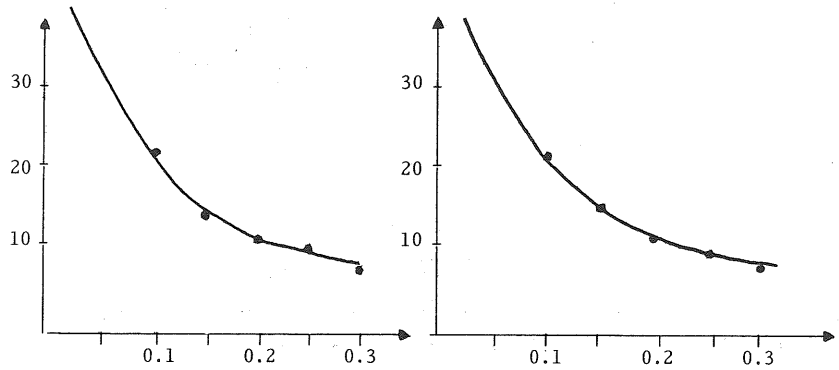


FIGURE 10

It seems to be a very laborious task to try to calculate the expectation in this case. It is, however, easy to simulate the situation with the program used in the previous case but now with n=5 instead of 2. The following data were obtained in such a way.

p	Average in 100 trials	Average in 1000 trials
0.05	42.09	45.138
0.10	22.70	22.546
0.15	14.13	14.582
0.20	10.90	10.915
0.25	8.81	8.451
0.30	6.50	6.558



Average system life in 100 trials

Average system life in 1000 trials

FIGURE 11

These averages are shown in Figure 11. We have in this case not calculated the expectation as a function of the component probability p but we can instead make an estimate by fitting a curve to the observations.

We can also combine a simulation program with a statistical data analysis program. The flow charts in Figure 12 show how programs can be written, which will calculate mean, variance and standard deviation of simulated observations of the system life. The programs can easily be modified also to give the frequency distributions.

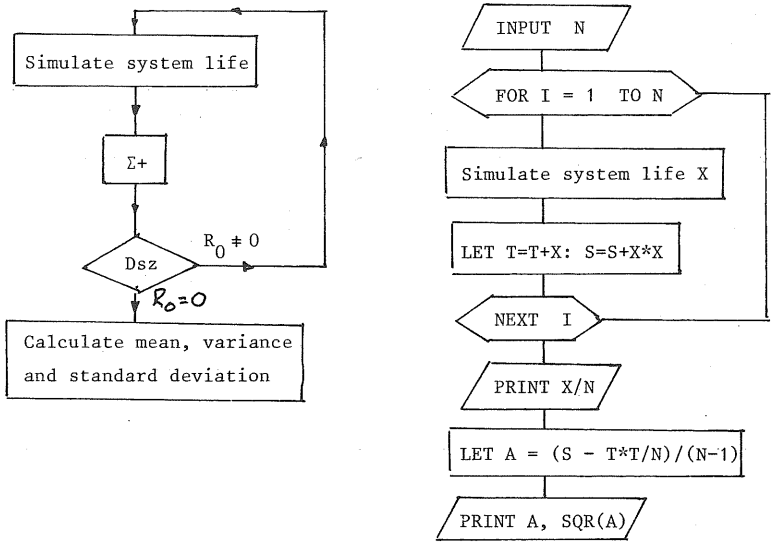


FIGURE 12

The data in the following table have been obtained with a program according to Figure 12. In all cases 100 trials have been made.

n	p	Mean	Variance	Standard deviation
2	0.05	29.7	440	21.0
2	0.10	15.4	108	10.4
2	0.15	9.41	43.6	6.61
2	0.20	7.08	26.1	5.11
2	0.25	5.66	13.9	3.73
2	0.30	4.54	7.87	2.80
5	0.05	42.9	516	22.7
5	0.10	20.8	89.6	9.46
5	0.15	15.4	78.6	8.86
5	0.20	11.0	35.0	5.91
5	0.25	8.88	19.8	4.46
5	0.30	7.13	11.6	3.41

Why is it better to estimate the expected system life with the average in 1000 trials than in 100 trials? This problem can be explored by simulating a sequence of averages and by calculating the variance for the averages. If s^2 is the variance for 100 (say) averages of n observations each, the data obtained will soon indicate that it may be suitable to calculate ns^2. The following data were obtained in this way.

Variance s^2 for 100 and 1000 means of n observations.
In all cases a system with two components and p=0.3 has been used.

100 means			1000 means		
n	s^2	ns^2	n	s^2	ns^2
1	9.39	9.39	1	9.98	9.98
2	4.08	8.16	2	5.02	10.0
3	3.50	10.5	3	3.31	9.92
4	2.67	10.7	4	2.38	9.51
5	1.88	9.40	5	2.02	10.1
6	1.55	9.30	6	1.62	9.72
7	1.42	9.94	7	1.37	9.59
8	1.47	11.81	8	1.31	10.5
9	1.14	10.3	9	1.14	10.2
10	0.981	9.81	10	1.05	10.5

The data strongly indicate that the variance of a mean of n observations is 1/n of the variance for one observation.

It should finally be remarked that it is possible to derive a closed expression for the expectation in the general case with n components or n coins; see J. Kinney [3] and L. Råde [10].

REAL DATA - SIMULATED DATA

The students will in a computer oriented course in stochastics to a large extent work with simulated data. The statistical treatment of such data is of course the same as for real data with the exception that simulated data are often generated within the computer and are not to be given from outside to the computer for statistical data analysis. But the students should also work with real data and should see some examples of how difficult it may be to distinguish between real data and simulated data. This point is discussed in Rao [4] and in Råde [6].

OTHER PROBLEMS

The following is a list of problem areas which provide lots of challenging problems which can be used in stochastics courses at the school level.

Random number generators
Test of random number generators
Simulation of probability distributions
Calculation of probability distributions
Simulation of games of chance
Stochastic geometry
Stochastic number theory
Simulation of Stochastic service systems
Simulation of storage systems
Simulation of statistical estimation procedures

CHARACTERISTICS OF A COMPUTER ORIENTED APPROACH TO STOCHASTICS

The following is a list of some advantages of a computer oriented approach to the teaching of stochastics at the school level. Hopefully these advantages have been demonstrated in the two examples above.

1) A combination of a theoretical and an experimental approach may be used.
2) Important concepts like probabilities, expectation and variance and their properties can be motivated and illustrated by appropriate experiments.
3) Work with data obtained from simulations will provide lots of opportunities for statistical data analysis.
4) Stability and variability in random experiments and statistical estimation procedures can be illustrated.

5) The students will learn about simulation, which is an important method of solving complex probability problems.
6) Simulation of random experiments can be organized as a teamwork of groups of students.

References

[1] Engel, A. (1977) Elementarmathematik von algorithmischen Standpunkt, Ernst Kleff, Stuttgart.

[2] Engel, A. (1982) Statistik auf der Schule: Ideen und Beispiele aus neuerer Zeit, Der Mathematikunterricht, 28: 1.

[3] Kinney, J. (1978) Tossing coins until all are heads, Mathematical Magazine, 51, 1.

[4] Rao, C.R. (1970) A multidisciplinary approach for teaching statistics and probability, The Teaching of Probability and Statistics: (ed. L.Råde, Almqvist & Wiksell, Uppsala.

[5] Råde, L. (1976) Probability, simulation and programmable calculators, Anwendungsorientierte Mathematik in der Sekundarstufe II (eds. (eds. W. Dörfer, R. Fischer), Verlag Johannes Heyn, Klagenfurt.

[6] Råde, L. (1981) Statistics, Simulation and Statistical Data Analysis at the School Level, Stochastik im Schulunterricht, Holder-Pichler-Tempsky, Wien.

[7] Råde, L. (1981) Random digits and the programmable calculator, Teaching Statistics and Probability, NCTM yearbook (ed. A.P.Shulte), Washington.

[8] Råde, L. (1977) Take a Chance with your calculator, Probability problems for programmable calculators, Dilithium Press.

[9] Råde, L. (1981) Algorithms for calculation and simulation of probability distributions, Department of Mathematics, Chalmers University of Technology.

[10] Råde, L. (1976) Reliability systems in random environment, Journal of Applied Probability, 13.

OPTIMUM BALANCE BETWEEN STATISTICAL THEORY AND APPLICATIONS IN TEACHING

C Radhakrishna Rao, University of Pittsburgh

There have been several round table meetings and discussion sessions on teaching of statistics organized by the International Statistical Institute (ISI) and the International Committee of Mathematical Instruction. These meetings and the deliberations of the Education Committee of the ISI at the last fifteen biennial sessions have served a useful purpose in identifying the various problems connected with the education and training of statisticians and bringing to the forefront several issues which need to be discussed in a wider forum. I am glad that the ISI has taken the responsibility in this direction to organize the first international conference on the teaching of statistics. The program committee headed by Lennart Råde deserves our congratulations for working out a comprehensive program and securing the participation of experts from all over the world. We hope that our deliberations will result in specific recommendations and, what is more important, that we will be able to set up a suitable machinery for following up on these recommendations, bringing them to the attention of educational institutions all over the world and helping in their implementation.

The program committee has placed on me a heavy responsibility in asking me to speak at one of the plenary sessions. I could not resist the temptation of accepting the invitation to participate in a conference on a subject of considerable interest to me. As Director of Research and Training at the Indian Statistical Institute for over 35 years, I had the opportunity to develop and institute a wide variety of educational and training programs in statistics at various levels:

* Pre college, undergraduate and graduate courses leading to bachelor's, master's and Ph.D. degrees.
* Applied courses for research scientists working in basic disciplines like biology, psychology, sociology and economics.
* Refresher and advanced courses for statisticians employed in government offices, business, industry and research organizations.
* Short courses for factory workers to help in the implementation of quality control programs.
* Workshops for field workers who gather information by interviewing people or by direct observation.

I shall describe the efforts we have made in formulating these varied types of educational and training programs and in implementing them. I hope my experience will be of some use to others engaged in these endeavours.

WHITHER STATISTICS?

Much has been said about statistics as a discipline, a science and a profession. One has only to read the presidential addresses appearing in the JASA and the JRSS year after year, expressing concern about the future of statistics, lamenting the nonrecognition of statistics as a separate discipline by scientific academies, criticizing the statistical educational programs in the Universities as too theoretical, emphasizing the need to close the "vicious gap" between theoreticians and practitioners of statistics, diversification of statistics into narrow areas of specialization, and so on. They all express the same anxiety. What is the future of statistics? How to develop statistics as a truly viable and socially meaningful science?

Whatever may be the views expressed on the status of statistics as compared to the basic disciplines, it is a fact that statistics has become an indispensable tool in all branches of human endeavour from scientific research and complex decision making to regulation of our daily lives. In his presidential address on "The Expansion of Statistics" to the Royal Statistical Society thirty years ago, R.A. Fisher (1953), described statistical science as

> ..... the peculiar aspect of human progress which gave to the twentieth century its special character; and indeed members of my present audience will know from their personal and professional experience that it is to the statistician that the present age turns for what is most essential in all its important activities. They are the "backroom boys" of every significant enterprise.

The sudden spurt in the applications of statistics which started in the second quarter of this century has led to phenomenal developments in statistical methodology such as design of experiments, sample surveys, quality control and sequential analysis. The large membership of statistical societies all over the world, and the ever increasing number of journals devoted to the theory and applications of statistics ensure us that statistical activity is a growing concern. It is growing not because statistics has reached the self generating stage like mathematics where it can create its own problems but because of its need for investigations in other areas. Unlike in other disciplines we do not study statistics to solve problems in statistics but to solve problems of other disciplines, which gives statistics its special character.

While new problems in statistics arise from its interface with other disciplines, the development of statistical methodology itself requires good knowledge of mathematics. It is interesting to note that in the table given by Box (1980) to illustrate that "interaction with reality can be very fruitful", the names of the investigators mentioned against the practical problems which led to important

techniques in statistics, are those of Gauss, Bernoulli, Laplace, Wald, Fisher, etc., who are all outstanding mathematicians.

Both Fisher and Mahalanobis emphasized that progress in statistics can take place only through new applications while agreeing on the relevance and use of mathematics in developing the statistical tools. I quote from Mahalanobis (1965):

> Pure mathematics is indispensable for supplying rigorous logical foundations and for exploring the limits and refinements of statistical theory but has never given it essentially new concepts or tools. New statistical concepts and method have continually emerged only in dealing with observations in the world of reality.

The level of mathematical education necessary for a practising statistician is, of course, a different issue which has to be carefully examined.

STATISTICS AT THE HIGH SCHOOL LEVEL

Mathematics is essential and has to be taught at a young age as a method of reasoning and a discipline in rigor essential for the growing mind, and as a preparation for the study of basic sciences later in their educational career. That all sciences are in the abstract mathematics is the basic argument for teaching mathematics at an adequate level of sophistication in the high school. But study of mathematics emphasizes only one aspect of reasoning, deductive in character that of arguing from given premises or deriving the consequences from a given set of postulates. This is highly restrictive in nature and allows the students to see only one side of the coin. It is, perhaps, equally important to expose the young minds to reasoning the other way around, inductive in character of arguing from given consequences to the premises, by which new knowledge is created. While deductive reasoning was systematized through the study of mathematics more than 20 centuries ago, inductive reasoning because of its uncertain nature remained uncodified. It is only in the beginning of the present century with the emergence of statistics as a method of drawing "inference on a population from a sample" that serious attempts have been made to provide rules for inductive reasoning through appropriate statistical methodology. Thus, while mathematics provides an abstract framework for describing natural phenomena, statistics provides the concrete methodological tools in searching for a working, if not true, model for observed phenomena.

If we agree that it is important to expose the students at an early age to inductive type of reasoning as well through the study of statistics, then we have to think seriously how statistics should be introduced as a part of the

high school curriculum. Should statistics be introduced as a separate subject to replace or by cutting down the syllabus of some subject that is being currently taught? Should statistics be taught to all students in the high school or only to those who wish to specialize in social and life sciences later. The two papers by Peter Holmes (1981) in the Proceedings of the Gothenburg Symposium on Teaching Statistics held in 1980 contain useful thoughts on these issues. I quote from one of the articles:

> ..... there was a growth in the numerical content of the subjects such as geography, economics and life sciences and representation was made to the examination boards (in the U.K.) that it would be much more useful for students in these courses if, in their mathematical courses they could study probability and statistics instead of theoretical mechanics. In response to this call the examination boards began to make available examinations in pure mathematics and statistics as an alternative to the more traditional pure mathematics and theoretical mechanics.

This is only to give one instance where statistics was introduced by cutting down some topics in the syllabus for mathematics, and statistics was made a compulsory subject of study only for some students. Is this the right thing to do? Some of us may not think so.

What about the contents of the course itself once it is decided to introduce statistics as a separate subject at the high school level? Naturally, one thinks of elementary probability to be taught through exercises involving permutations and combinations, probability distributions through histograms, calculation of median, mode and a few moments by using specified formulae, concepts of correlation and regression and elementary testing of hypotheses. Peter Holmes points out that over a period of years since 1959 when statistics was introduced, there was a growing dissatisfaction with these courses. This should be evident from the high school texts prepared to cover the syllabus in statistics from which the teacher is supposed to teach and from which the students are expected to memorize the formulae for answering the numerical exercises set in the examinations. The students find such a course dull and drab, and it does not serve the purpose for which a case is made for the study of statistics at an early age, viz., to inculcate the spirit of enquiry and to expose the growing minds to concepts of inductive reasoning. Peter Holmes mentions that some changes have been made in the syllabus to include project work to bring the students in contact with live data. This is desirable, although much depends on the teacher, and how the project work is conceived and the role of statistics as a method of quantitative reasoning is emphasized.

What is perhaps ideal is not to treat statistics as a separate subject but reframing the syllabus of the social and life sciences, and in natural sciences as well, injecting statistical concepts and methodology. Experiments of this nature were attempted at the Indian Statistical Institute under the guidance of J.B.S. Haldane who himself designed and taught courses in statistics at the precollege level (see Rao, 1969, 1974). This is not an easy task but worth trying.

We must remember that our aim in teaching statistics at an early age is not to impart statistical skills per se but to expose them to statistical concepts, chance mechanisms in nature and decision making under uncertainty in real situations, with examples chosen from the subjects they are studying. To arouse the interests of the students in statistics, examples given in books such as Mosteller et al (1973) and papers by Rao (1977a,b) may be discussed.

MAKING OF PROFESSIONAL STATISTICIANS

In planning for the education and training of professional statisticians we have to take into account the future demand for statisticians in different areas, the nature of problems they will be called upon to study and the skills they are expected to have. I believe the greatest demand for statisticians will be in industry and business and to some extent in the local, state and federal governments. There will be a limited demand for teachers in statistics in high schools and universities. Apart from some basic common courses in theoretical and applied statistics, different types of courses may be needed for statisticians preparing to work in different areas. In an excellent report on "Preparing statisticians for careers in industry" the ASA committee on training of statisticians in industry (with R.D. Snee as chairman) noted:

> When a statistician leaves school his or her skills and experience should include statistical knowledge, practical problem solving, consulting practice, and ability to communicate orally and in writing with non-statisticians.

This view expressed by the committee is generally applicable to the training of all types of practising statisticians. The ASA committee made a number of recommendations with special reference to the training of statisticians for industry (see ASA, 1980). It would be useful to have similar reports for training in other areas of applications of statistics.

Based on my experience as a teacher and a consultant in statistics, I suggest the following program of studies at the undergraduate and graduate levels for the training of professional statisticians. The courses are divided into three broad categories:

CATEGORY I Basic Theory	CATEGORY II Basic Applied	CATEGORY III Special Topics (Optional)
Probability theory	Design of experiments	Econometrics
Linear models and analysis of variance	Sample surveys	Time Series
Categorical data and chi-square tests	Computer simulation	Biometrics
Statistical inference (a) hypothesis testing (b) estimation (c) decision theory	Cross examination of data	Technometrics Advanced topics in (a) probability (b) inference (c) multivariate analysis

The courses in category I are meant to provide the basic knowledge in statistical concepts and theory and may be taught without relevance to particular applications. These courses should be compulsory for all statisticians. Some of the topics in these courses can be treated at different levels of mathematical sophistication depending on the mathematical equipment of the students. But knowledge of advanced calculus should be sufficient for a rigorous treatment of most of the topics.

The courses in category II will be of great value to practising statisticians in any field, although some techniques are more commonly used than the others in specific areas. The nature of instruction in these courses should be different from those in category I. In subjects like design of experiments and sample surveys, the theory can be built around practical applications. There are excellent books on these subjects which give numerous examples that may be used to provide proper motivation for developing the theory to meet different practical situations (see Des Raj (1968), Kish (1967), Murthy (1967) and Zarcovich (1966) on sampling). These subjects are ideal for introducing students to new concepts like pilot surveys, local control, cost and variance functions and uniformity trials, all of which are useful in designing optimum surveys and experiments. They also provide situations for examining problems like defining a population in terms of a frame and units for sampling, representativeness of a sample, response errors, missing values and repeated surveys, which a practising statistician in any field should be aware of.

In addition to regular instruction in theory illustrated with examples, it will be useful to assign projects to students, which involve designing of experiments and/or sample surveys, collection and analysis of data. One can think of a wide variety of projects which the students can undertake within the limited time and resources available to them while taking these courses. It will be

useful if a guide book on this aspect of training can be prepared listing the titles of possible topics for project work and indicating in some detail how each project can be organized.

The possibility of arranging visits of students to experimental farms and government departments engaged in large scale surveys should be explored. Such visits would give the students an opportunity to see how the theory they are learning is actually put into practice.

In category II, I have also listed two other subjects, Computer simulation and Cross examination of data which deserve great attention. We are all familiar with computer simulation which is an invaluable aid in model building and in solving or providing an insight into problems which are mathematically intractable. Experience in this area and numerical analysis in general will be a great asset to practising statisticians. There is extensive literature on this subject, and we should try to frame a suitable syllabus for a special course on computer simulation for statisticians, giving typical examples.

Cross examination of data is a term suggested by R.A. Fisher for preliminary analysis of primary data. It literally means questioning the data to elicit relevant information before embarking on more sophisticated statistical analysis. What is an appropriate model for given data? Do the observations come from a homogeneous population? Are there outliers, recording errors, contaminating observations and other deficiencies in the data? Are the given data genuine or faked? Is there evidence that the observations are edited or censored in some way? These questions are ignored or not often raised by the users of statistical methods. Textbooks on statistics do not emphasize the importance of raising these questions and eliciting answers from data, which are useful in the choice of appropriate statistical methodology and interpretation of results. These questions are not easy to answer as there are no systematic methods developed to guide us. Cross examination of data and choice of appropriate statistical methodology have remained more in the realm of art and still depend to a large extent on the skill and experience of the statistical practitioner. See for instance the illuminating papers by Fisher (1936) and Haldane (1946) on faking of genetic data, and Mahalanobis (1933) and Majumdar and Rao (1958) on scrutiny of anthropometric data (scrutiny is an earlier term used by Mahalanobis for a preliminary look at the data). However, in recent years, attempts have been made to devise systematic methods for cross examining data, and there is enough material to cover in a separate course on the subject. I suggest the following tentative outline for the syllabus.

Syllabus for cross examination of data

* Graphical methods for display of data.
* Methods for detection of outliers and other deficiencies in data.
* Methods for examining whether data are edited, censored or faked.
* Multivariate techniques such as principal components and canonical analysis.
* Critical reviews of published papers containing statistical analysis of real data.

Books by Gnanadesikan (1977) and Tukey (1970) can provide valuable material for teaching this course. They can be supplemented by making a list of original papers such as those by Fisher, Haldane and Mahalanobis already referred to, for discussion in the class. There is no dearth of published papers containing wrong applications of statistics leading to misleading interpretations, which can be selected for critical reviews by students.

The courses under category III are optional. Those who are preparing to work in industry can take courses on special statistical methods used in industry such as quality control, acceptance sampling, experimental designs for exploring response surfaces and evolutionary operation, operations research, life testing, reliability etc., which may be broadly described as Technometrics. Those who want to work in government offices and business firms can take courses on econometric models, time series, construction of economic and social indicators, which may be broadly described as Econometrics. Those who wish to work in research wings of pharmaceutical companies, hospitals and public health departments may take courses in Biometrics which includes subjects like bio-assays, applied multivariate analysis, mathematical genetics, life tables and so on. Finally, a small number of students may wish to specialize in mathematical statistics to do research on some outstanding problems in statistical theory. They may take courses on advanced topics in probability, statistical inference and multivariate analysis.

The level of special courses in category III should be somewhat high both in its theoretical content and applied flavor. At the Indian Statistical Institute, the students taking the special courses in this category are required to spend some time as apprentices in statistical offices engaged in specified statistical activities as part of their training. Thus, those who are taking courses in technometrics are sent to industrial establishments having well developed statistical divisions. Those who are taking courses in econometrics are sent to government statistical offices, and so on.

The aim of the special courses in category III is to prepare the students to

* undertake on their own responsibility practical investigations in their special area, and
* identify new problems and do research by themselves and/or in collaboration with others for developing new statistical methodology.

THE Ph.D. PROGRAM

I have discussed the making of a professional statistician through appropriate courses of study and training at the Universities. These courses should qualify a student for a Master's degree as a mark of a high level of professional attainment. The Ph.D. program for statisticians, on the other hand, requires different considerations. The standards for such a program differ considerably from university to university within a country and between different countries. The standards for the award of the Ph.D. degree have also changed considerably over the years.

During my first visit to the United States, I asked the head of the statistics department of a well known university as to what kind of program they had in statistics. He replied with a feeling of satisfaction that it was graduate leading to a Ph.D. Then I enquired as to what they were preparing the students for. He said, for the Ph.D. degree in statistics. This was quite in contrast to my own experience while working with R.A. Fisher and at the Indian Statistical Institute. Before I went to Cambridge on an assignment to work on an anthropometric project, I had already done research work in theoretical statistics which could have earned for me a Ph.D. degree in India. However, I thought I should work with Fisher and get a Ph.D. degree from Cambridge. During my first interview with Fisher, I asked him to suggest a problem for my Ph.D. thesis. In his characteristic manner, Fisher said, "The problem must be yours and I might be able to offer you some help in solving it". I was reminded of what Poincaré said emphasizing the ability to see a problem as the major step in research. "The question is what is the question". A Master's program in statistics is not worthwhile if a student who has completed the program is unable to find by himself or herself a new problem for investigation. In Cambridge, I was working on an anthropological project which involved measuring bones and stones and doing a comparative statistical analysis of observations made on different collections of skeletal material from Africa. I was aware that the statistical discriminant function for deciding between two alternatives had its origin in sexing a human mandible. Fisher was glad when I chose discrimination between several populations, which was relevant to my project, as my thesis topic. On my return to India, we set up the tradition at the Indian Statistical Institute of encouraging the Ph.D. students to choose their own problems for writing the thesis. This policy resulted in several highly

original theses on topics, which were, more often, unfamiliar to the supervisors.

I was surprised at what Box (1980) has mentioned:

> The student may receive verbal and nonverbal messages that tell him that the kind of Ph.D. thesis that is motivated by a scientific problem with real data is not likely to be so acceptable as a pretty but narrowly introspective mathematical exercise..... He may suspect that when his publications are assessed, those with any taint of application will be down-weighted, if not entirely discounted.

This is contrary to the very essence of research, and I hope good sense will prevail in departments where such views are held.

REFRESHER COURSES FOR PRACTITIONERS

During the last few years, the ISI has been concerned with the problems of the "growing gap between theoretical and practical statisticians" and "increasing specialization in various subject matter areas of statistics leading to diversification of interests". I have referred to some of these problems in two papers entitled "Growing responsibilities of government statisticians" (Rao, 1975) and "Perspectives in statistics" (Rao, 1979). I believe these problems need a three pronged attack.

First, theoretical statisticians working in universities and research organizations should be encouraged to work from time to time in government statistical offices and industrial establishments. This will provide an opportunity for the academic and practising statisticians to interact and benefit from each other's knowledge and expertise. The visiting theoretical statisticians can study some of the real problems faced by the practitioners and hold seminars and workshops to discuss the possibilities of using new theoretical developments in the analysis and interpretation of data.

Second, the officers working in the government and in the industry should be given an opportunity to take refresher courses on the new developments in statistical theory. These courses should be somewhat different from the academic type given in the university departments. They should be specially designed to meet the needs of practising statisticians in different areas, with emphasis on new applications. The officers taking such courses should be asked to bring data from their offices and analyse in consultation with the teaching staff. At the Indian Statistical Institute, such special courses of three to six months duration are regularly offered to officers deputed from industry and government.

Third, there should be periodical conferences of statisticians working in specific areas to discuss current developments. It will serve a useful purpose if such

conferences are centered round a few problems on which key note papers by both theoreticians and practitioners are invited and discussed.

COURSES FOR RESEARCH SCIENTISTS

At present, the method of providing instruction in statistics to students specializing in other disciplines at the college level is far from satisfactory. In some universities all the courses in statistics are centralized and are offered by the department of statistics, which announces a list of courses ranging from elementary statistics to highly specialized topics. Students from other disciplines can take one or more of these courses given by the faculty of the statistics department. In other universities, each department for a particular discipline formulates its own courses in statistics and has a faculty member to give these courses to their students. Both practices have some advantages and disadvantages.

It must be noted that the courses in statistics for students in other disciplines can have only a limited objective:

* Make them aware of the usefulness of the statistical approach to problem solving.
* Demonstrate errors in measurements, differences between observers and instruments, and scrutiny of data.
* Graphical display of data and summarization of data in terms of quantiles, mean, mode and lower order moments.
* Principles of design of experiments, difficulties in interpretation of historical data.
* Preparation of schedules for recording observations.
* Tests based on mean values and proportions. Concept of the standard error.
* Use of computers in data analysis.

The courses should have a strong applied bias with examples chosen from the fields in which the students are specializing and emphasis on the need for properly defining the measurements, making observations and ensuring comparability of results. The students should be made aware of the literature in their disciplines on investigations which use more advanced statistical methodology for interpretation of results. The need to consult a competent statistician rather than blindly using computer programs in such cases may be emphasized.

The elementary courses offered by the statistics departments do not generally meet the objectives listed above, and further such courses are assigned to inexperienced staff members or teaching assistants who can only teach from cookbooks. There may be a case for specialists in each individual discipline

with a good knowledge of statistics to give such courses to students specializing in that particular discipline.

THE TRAINING OF TEACHERS

The teacher is an important component in all educational and training programs. It is not enough to have an impressive sounding syllabus. More important is the availability of good teachers who have the necessary knowledge and experience to present the different topics of the syllabus in an interesting way and arouse the enthusiasm of the students, especially at the high school level. Along with our deliberations on the objectives and contents of the courses at different levels we should also consider what steps can be taken to ensure a high quality of teaching. I suggest the following for further discussion:

* Preparation of teaching manuals and audio-visual aides along with the syllabus.
* Workshops, short courses and discussion meetings for teachers.
* Encouraging teachers to take up consultation and acquire experience in solving practical problems.

It will be generally useful for teachers and research workers in statistics to get acquainted with the subject matter of a basic discipline in some depth to be able to work by themselves and/or effectively collaborate with workers in the discipline on research projects involving a quantitative approach.

IN CONCLUSION

The object of introducing statistics in the high school curriculum is not to provide them with a kit of statistical tools and a standard set of prescriptions for use in practice. It is not for producing statisticians at lower levels, if at all such a category has a meaning. It is for inculcating a spirit of enquiry and exposing the young and growing minds to inductive reasoning as complementary to deductive reasoning for advancement of knowledge. This is best achieved by not teaching statistics as a separate subject in the high school but by injecting statistical ideas and quantitative approach in the syllabus for social and life sciences.

The curriculum at the undergraduate and graduate levels should be designed to meet the future demand for statisticians at the level of practitioners in special areas of application (such as in government, industry, business, etc.), general consultants, research workers in statistical theory and competent teachers. Suitable syllabi for these purposes are outlined, providing a proper balance between theory and applications.

Refresher courses for practitioners of statistics is important in order to keep them abreast with latest developments in statistical theory and refinements in statistical methodology. This will solve, to some extent, the problem of "transfer of methodology between academic and practising statisticians".

Courses in statistics for research workers in other disciplines should emphasize the quantitative approach in problem solving, and cover the use of standard statistical methodology.

Finally, there is a need to produce teachers of statistics with a strong theoretical background and experience in the application of statistics in at least one basic discipline.

References

[1] American Statistical Association (1980) Preparing statisticians for careers in industry: Report of the ASA committee on training of statisticians for industry. The American Statistician, 34, 65-75.

[2] Box, G.E.P. (1980) The American Statistician, 34, 76-77.

[3] Des Raj (1968) Sampling Theory, Macgraw Hill.

[4] Fisher, R.A. (1936) Has Mendel's work been rediscovered? Annals of Science, 1, 115-137.

[5] Fisher, R.A. (1953) The expansion of statistics. J.Roy.Statist.Soc., A, 116, 1-6.

[6] Gnanadesikan, R. (1977) Methods for Data Analysis of Multivariate Observations, John Wiley.

[7] Haldane, J.B.S. (1946) The faking of genetic results, Eureka, 6, 21-24.

[8] Holmes, Peter (1981) Why teach statistics and what statistics should be taught to pupils aged 11-16; Using project work in teaching statistics to students aged 16-18. Proc.Gothenburg Symposium on Teaching Statistics (Editor, L. Råde).

[9] Kish, L. (1967) Survey Sampling, John Wiley.

[10] Mahalanobis, P.C. (1933) A revision of Risley's anthropometric data relating to the tribes and castes of Bengal. Sankhyā, 1, 76-105.

[11] Mahalanobis, P.C. (1965) Statistics as a key technology. The American Statistician, 19, 43-46.

[12] Majumdar, D.N. and Rao, C. Radhakrishna (1958) Bengal anthropometric survey, 1945 - a statistical study. Sankhyā, 19, 201-408.

[13] Mosteller, F., Kruskal, W.H., Link, R.F., Pieters, R.S. and Rising, G.R. (eds.) (1973) Statistics by Examples, Addison-Wiley.

[14] Murthy, M.N. (1967) Sampling Theory and Methods, Stat. Publishing Soc., Calcutta.

[15] Rao, C. Radhakrishna (1969) A multidisciplinary approach for teaching statistics and probability. Sankhyā, B, 31, 321-340.

[16] Rao, C. Radhakrishna (1971) Data, analysis and statistical thinking. Chapter 17 (pp.383-392) in Economic and Social Development, Vora and Co., India.

[17] Rao, C. Radhakrishna (1974) Teaching of statistics at the secondary level - an interdisciplinary approach. Statistics at the School Level, 121-140, Almqvist and Wiksell.

[18] Rao, C. Radhakrishna (1975) Growing responsibilities of government statisticians. Occasional Papers, No.4, Asian Institute of Statistics, Tokyo.

[19] Rao, C. Radhakrishna (1977a) A natural example of weighted distributions - a classroom exercise. The American Statistician, 31, 24-26.

[20] Rao, C. Radhakrishna (1977b) Statistics in search of truth. Science Today, 27-36.

[21] Rao, C. Radhakrishna (1979) Perspectives in Statistics, Sankhyā B, 41, 129-137.

[22] Tukey, J. (1970) Exploratory Data Analysis, Addison Wesley.

[23] Zarkovich, S.S. (1966) Quality of Statistical Data, FAO, Rome.

APPENDIX

During the last few days, I had the opportunity to listen to various speakers at the conference on topics such as: Why teach statistics in high school? What topics in statistics should be taught in high schools and in colleges? How should statistics be taught at different levels? In this appendix, I shall comment on some of the views expressed by the previous speakers, and explain the basis of the suggestions made in section 3 of my paper on "The Making of Professional Statisticians".

At the high school level, the question of how to teach statistics is extremely important and has been the subject of considerable discussion during the last few years. But at the graduate and postgraduate levels, the question of what to teach is far more important than how to teach. In order to decide on a suitable course of study for qualifying as a professional statistician, we should recognize three main aspects of statistics as a tool in applied research, Data, Methodology and Quantitative thinking, as detailed in the following chart.

Traditionally, the term statistics is used to denote numbers, each of which is a well defined measurement made on an individual or more generally, on unit of a population (real or hypothetical). Each number has a meaning in the sense that it conveys information on the particular unit which provided the measurement. Often, we are not interested in individual units, but would

THREE ASPECTS OF STATISTICS

DATA	METHODOLOGY	QUANTITATIVE THINKING
(1) Observations (2) Mechanism Generating Observations (3) Prior Information about the Population	A. COLLECTION OF DATA (1) Sample Surveys (2) Design of Experiments B. THEORIES OF INFERENCE (1) Estimation (2) Hypothesis Testing (3) Decision Making	(1) Trouble Shooting (2) Formulation of New Problems

like to have some information about the population as a whole, comprising the different units. If all the units could be measured, then any information about the population would be known. If only some units are measured, then the inference on the population would be uncertain. The term statistics, as it is used now, stands for the whole process of making inference under uncertainty, which consists of collecting observations, analysing them, (taking into account the mechanism of generating observations and any prior information available about the population), and interpreting the results. The observations, the mechanism generating them and the prior information constitute the data. The techniques of acquiring data in an efficient way (i.e., to provide maximum information at given cost through well designed experiments and sample surveys) and answering specific questions (like testing given hypotheses and estimating specified parameters) constitute the statistical methodology. Usually, data contain more information than what is required for answering specific questions. The process of extracting such information as a guide to future investigations may be called quantitative thinking. Thus, data analysis has a dual purpose:

DATA ANALYSIS	=	ANSWERING SPECIFIC QUESTIONS	+	RAISING QUESTIONS FOR FURTHER ENQUIRY

Section 3 of the paper "The Making of Professional Statisticians" contains details of compulsory and optional subjects to be taught at graduate and postgraduate levels. Some additional remarks on teaching are as follows:

* There are considerable controversies on theories of inference and methods of data analysis, which should be brought to the attention of the students to make them aware of the limitations of statistical theory, as developed up till now.

* The importance of well documented and accurate data as prerequisite for correct inference should be emphasized.

* There is often more information in data than what can be extracted by methodology designed to answer specific questions. The method of extracting such information has no specific rules and remains as an art depending on the skill and experience of the statistician. Discussion of numerous live examples from applied papers emphasizing the three aspects of statistics in research work would be extremely useful.

This work is sponsored by the Air Force office of Scientific Research under Contract F49629-82-K-001. Reproduction in whole or in part is permitted for any purpose of the United States Government.

THE RELEVANCE OF STATISTICAL TRAINING

J Gani, University of Kentucky

INTRODUCTION

It is in the nature of every professional to regard his calling as indispensable to the common good; practising statisticians and teachers of statistics in schools and universities are no exception to this rule. Yet, in the face of competing claims for the importance of their areas from professionals in agriculture, the arts and humanities, business and administration, the law, the biological and medical sciences, the physical sciences, the social sciences and technology, statisticians may experience an occasional twinge of doubt as to the relevance and standing of their subject.

It is my intention to demonstrate in a very pragmatic way that statistics is a subject of vital importance; it enters substantially into the quantitative content of many (if not all) other professional studies. Statistical information is pervasive: a detailed analysis of The New York Times newspaper for Saturday 22 May 1982 will leave us in little doubt of this fact. The intelligent reading of such a newspaper, or of any other current information material, therefore makes a basic grounding in statistics essential for all citizens. This premise leads very naturally to a discussion of what a common core of statistical training should consist of, and how it might best be imparted to students in schools and colleges throughout the world.

At this stage we are faced with the question of whether basic statistical training can be culture-invariant. Do the citizens of a developed country require the same statistical knowledge as those of a developing nation? Is a grasp of demographic or agricultural statistics, which may be essential to a Nigerian, radically different in kind from an understanding of economic or industrial statistics, which may be more relevant to a Japanese? It may be desirable in theory for all of us to know the same elements of statistics, but in practice our training will depend on the educationl resources available in our different countries. Even if a unviersal course in basic statistics were agreed upon, the different circumstances of each nation would dictate the use of different local examples and illustrations. There would clearly be little point in using stock market statistics as an example of a time series trend in the USSR or illustrating the concept of variability with industrial quality control data in Upper Volta.

The picture that emerges from our study of The New York Times is one of considerable reliance on statistics, in the sense of numerical data, its graphical description, and its comparative evaluation. The estimative and inferential

roles of statistics still remain fairly rudimentary at the layman's level of need. Yet the world-wide spread of scientific education and the growing use of home mini-computers in the developed countries make it increasingly important that a modicum of statistics be taught to every student in our schools and colleges. It has been one of the functions of this International Conference on the Teaching of Statistics to consider how this can best be done.

THE NEW YORK TIMES OF 22 MAY 1982

To establish the importance of statistics in our daily life, I thought it useful to analyze the contents of a newspaper of international standing, with a wide popular circulation. The New York Times, a national US daily, which I purchased in Minneapolis, seemed ideally suited for the purpose. On 22 May 1982 it consisted of 34 pages, 3 of which were advertisements, while 31 contained articles and photographs of the news of the day, together with assorted items of information such as the composition of a US Senate roll-call, tables of financial data, sports results, death notices and weather records. I estimated that the proportion of this information which contained statistical material of some kind was 32.6% of the total available space of 31 pages, or just under one third of the paper; one would have difficulty in denying the relevance of the subject in the face of such evidence. Let us examine some of the material in greater detail.

Financial Reports

The biggest single part of the newspaper relying heavily on statistics was the financial section in the 12 page "Business Day". Statistical material in this totalled 6.25 pages, or 20.2% of the newspaper's information content; it covered:

1. N.Y. Stock Exchange Bond Trading
2. Treasury Bills, Bonds and Notes
3. Money Market Mutual funds
4. Federal Reserve
5. New York Stock Exchange Issues
6. New York Stock Exchange
7. Futures Prices
8. Cash Prices (Foods, Grain and Feeds, Fats and Oils, Metals, Textiles and Fibers, Miscellaneous and Petroleum)
9. Foreign Exchange (Lstg 1 = US $1.80, A$ 1 = US $1.06, Yen 1 = US $0.004213, Sw Fr 1 = US $0.5102)
10. World Gold (US $335.75 per troy oz)
11. Over-the-counter Quotations
12. Trading in Stock Options
13. American Stock Exchange
14. Other US and Foreign Stock Exchanges
15. Corporate Earnings
16. Foreign Stock Index

© New York Times Company, 1982
Reprinted with permission

FIGURE 1

N.Y. Stock Exchange Bond Trading

FRIDAY, MAY 21, 1982

Sales	Current Yield	In $1,000	High	Low	Last	Chge.	Net
A B C D							
AetnLf 81/607	13.5	10	60¼	60¼	60¼	+ ½	
AlaBn 15.35s99	15.4	7	99¾	99½	99½	.....	
AlaBnc 9¼s84	10.5	20	91	90¾	90¾	+ ¼	
AlaP 9¼2000	15.0	13	60	60	60	+ ½	
AlaP 8¼s01	15.3	5	55¾	55¾	55¾	+ ¼	
AlaP 7¾s02	14.4	4	53¾	53¾	53½	.....	
AlaP 8¾s03	15.2	10	58¼	58¼	58¼	+ 1½	
AlaP 9¾s04	15.6	10	62¾	62¾	62¾	+ ¼	
AlaP 10¾s05	16.0	14	68¾	67¾	67¾	+ ¼	
AlaP 10¼s05	15.8	5	65	64¾	64¾	+ ¼	
AlaP 9½s06	15.2	31	59½	59	59½	- 1½	
AlaP 8¾s07	15.3	11	57	57	57	- 1½	
AlaP 9½s08	15.3	11	62	62	62	.....	
AlaP 9¾s09	15.5	15	63¼	63	63¼	+ ¼	
AlaP 12¾s10	16.1	11	79½	78½	78½	.....	
AlaP 15¼s10	16.3	4	93¼	92¾	93½	+ ¾	
AlaP 13¾11	16.9	28	102¾	102	102¾	+ ¾	
AlaP 13¼s09	16.7	5	109	109	109	+ 3	
AlskH 16½99	16.3	36	100¾	100	100	.....	
AlskH 18½01	16.9	15	108½	108½	108½	+ ½	
AlskJ 10¾99	17.1	5	62¾	62¾	62¾	+ ¼	
AlldC 87s	.....	49	50½	50½	50½	+ ½	
AlldC 98s	.....	18	154	154	154	+ ½	
AlldC 2000s	.....	50	12	12	12	.....	
AlsChe 10.35s99	16.4	28	63¼	62¾	63¼	+ ½	
AltaFF 7¼87	10.2	10	71	71	71	.....	
AluCa 9½95	14.5	5	65½	65½	65½	+ ¼	
AMAX 8s86	10.1	2	79¾	79¾	79¾	.....	
AMAX 8½85	9.5	20	91	89¾	89¾	+ 3-2	
AAirl 4½92	11.0	80	41¼	39¾	39¾	- 2¾	
AAirl 5½98	.....	25	68	68	68	+ 1½	
ABrnd 8½85	9.3	10	87¾	87¾	87¾	+ ¼	
ACan 3¾s88	5.7	13	66	66	66	+ ¼	
ACan 9¼84	10.2	10	90¾	90¼	90¼	.....	
ACan 11¾10	15.5	7	76¼	75½	76¼	+ ¼	
ACanN 64s91	cv	4	49¾	49¾	49¾	+ ¾	
AExC 81/2s85	10.1	10	84	84	84	.....	
AExC 8½s86	10.2	44	83	83	83	.....	
AExC 10.15s90	12.5	25	80¾	80¾	80¾	- ¼	
AmGn 11s07	cv	3	102½	102½	102½	.....	
AmMed 9½01	cv	122	89	88	88½	- ½	
AmMot 8s88	cv	7	56	56	56	- ½	
ATT 3¼s84	3.8	38	84¼	84¼	84½	+ ¼	
ATT 4¼s85	5.3	8	82¾	81¾	82¼	+ ½	
ATT 4¾s85	5.4	5	81¾	81¾	81¾	.....	
ATT 7¼s87	4.1	20	70¾	70½	70¾	+ ¼	
ATT 7½s90	6.3	23	62	61	62	+ ¼	
ATT 8¼00	13.0	328	68¼	67¼	67½	- ¼	
ATT 7s01	12.3	127	57¼	56¾	56¾	- ¼	
ATT 7½03	12.5	141	57¼	56¾	56¾	.....	
ATT 8.80s05	13.5	327	65¾	65¼	65¼	+ ¼	
ATT 8¾07	13.3	47	66¾	66¼	66¼	+ ¼	
ATT 10½90	12.3	25	84¼	84	84¼	- ¼	
ATT 13¼91	13.6	266	97¾	96¾	97¼	- ¼	
Amfac 8¼94	cv	5	60¾	60½	60¾	.....	
AMF 8¼s85	10.0	5	86½	86½	86½	- 1½	
Amoco 5¼94	cv	20	61¼	61¼	61¼	+ ¼	
Ancp 9½00	cv	22	91¼	91¼	91¾	- ¼	
Ancp 13¾02	cv	20	94	94	94	.....	
Anhr 9 94s86	11.3	3	87¾	87¾	87½	+ 1¾	
ApolP 11¾s2015.6		1	96	95¾	96	2296-2296	
Arco 8¼s85	8.9	114	94½	94	94	.....	
Arco 8s84	8.6	27	90¾	90½	90¾	.....	
Arco P.22 7.7		15	97	1-1697	97	1-1697	
ArizP 10¾00	14.9	5	71½	71½	71½	.....	
ArizP 11¾99	14.8	12	82	81¾	82	+ ¼	
ArlnKit 5s86	Cv	1	47	47	47	.....	
ArmA 8.79s95	13.4	19	66	66	66	.....	
AshO 4¾93	Cv	10	70	70	70	+ 4¾	
AshO 8.80s00	14.4	20	61	61	61	.....	
AtlRi 9½00	17.4	3	54¾	54¾	54¾	.....	
Astrn 4¾84	5.2	5	84¾	84¾	84¾	+ 1	
Atchsn 4s95rf	8.9	13	45¼	45¼	45¼	.....	
Atchsn 4s95	8.8	10	45¼	45¼	45¼	+ ¼	
ARIch 7s97	10.7	63	68¼	68¼	68¼	+ ¼	
AvcoC 5½93	Cv	19	67¼	67	67¼	.....	
AvcoC 12s98	15.1	51	79¾	79¾	79¾	+ ¼	
AvcoF 7½97	14.6	4	51¼	52¾	52¾	.....	
AvcoF 8¼s84	9.4	3	90¾	90¼	90¼	+ ¼	
AvcoF 9½98	15.1	5	60¾	60¾	60¾	- 2	
Bache 14s00	15.4	23	90¾	90	90¾	.....	
Bally 6s95	Cv	25	101	102	102½	.....	
Bally 12s06	Cv	24	101½	100½	100½	+ ½	
B O 4½95	9.9	4	43¼	43¼	43¼	.....	

U.S. Govt. Bonds	Other Dom. Bonds	Foreign Bonds	Total All Bonds
Day's Sales .....$ —	$825,410,000	$70,000	$25,680,000
Thursday's	630,440,000	70,000	25,510,000
Year to Date.....	6,400,000	62,304,412,000	2,220,608,000
1981	51,501,000	61,812,839,000	51,822,230,000

e—includes international Bank Bonds

BOND ISSUES TRADED

	Issues	Advances	Declines	New Highs	New Lows
May 21	896	283	310	20	7
May 20	862	335	316	34	18
May 19	911	339	376	21	9

Sales	Current Yield	In $1,000	High	Low	Last	Chge.	Net
Chlco 9¼00	cv	39	70¼	68¾	68¾	- ¼	
Chlco 8.45s07	14.1	29	60½	60¼	60½	+ ¼	
Chlco 8½07	14.0	4	58	58	58	+ ½	
Chlco 16.10s98	16.8	12	96	95¾	96	+ ¼	
Chlco 16.95s04	17.1	17	134	93¾	94¾	.....	
Chlco 14¾86	14.4	86	100½	100¼	100¼	- ¼	
Chlco 15.30s86	14.7	30	104	102½	104	+ 2	
Chlco 16.45s86	15.2	90	108½	108¼	108¼	.....	
Chlco 16.45s86	15.9	25	104¼	103½	103½	- ½	
Chylmr 9½97	14.8	5	61	61	61	.....	
Chylmr 9½97	14.7	12	62¼	62¼	62¼	.....	
ClkEq 9¾03	9.6	30	97	95.22	97¾	97¾ + 3-22	
ClevEl 8¾04	9.0	20	91	89¾	89¾	+ ¼	
ClevEl 8¼05	15.0	10	58¾	58¾	58¾	- 1¾	
ClevEl 9¼09	14.8	2	62½	62½	62½	+ 1¾	
ClevEl 9.85s10	15.1	15	65¼	65¼	65¼	.....	
Cole 14s00	14.9	5	93¾	93¾	93¾	+ ¼	
ColuG 9s94	13.0	5	69½	69¼	69½	- ¼	
ColuG 9¼99	14.1	5	70½	70¼	70¼	+ ¼	
ColSO 9¼00 9.3		9	99	15-2299	15-2299	15-32 + 1	
ColSO 16¾91	16.6	1	100	100	100	- 3¾	
Comdis 13s01	cv	62	107	102	107	+ 2	
CmlCr 8¾88	11.5	20	76	75¾	76	+ ¼	
CmlCr 16¼95	14.4	22	98	98	98	.....	
CmwE 8s02	14.8	9	54¾	54	54	.....	
CmwE 8¼05	15.1	10	55	55	55	+ ¼	
CmwE 12¼88	12.4	27	97¾	97½	97¾	- ¼	
CmwE 14¾87	14.9	177	99¾	99¾	99¾	.....	
CmwE 15¾00	15.6	5	98½	98½	98½	+ ½	
CmwE 11½10	15.1	20	77¾	77¼	77¾	+ 1¼	
CmwE 16s91	14.8	1	108	108	108	+ ¼	
CmwE 16¼89	15.9	20	104	102½	102½	.....	
CmwE 17½88	16.2	51	108	108	108	+ ¼	
CmwE 16s90	15.8	91	102	101½	101½	- ¼	
ConEd 5s82	7.4	5	67¾	67¾	67¾	.....	
ConEd 9¾00	14.0	19	68	67	67¾	+ ¼	
ConEd 7.90s01	14.0	13	58	56½	56½	- ¼	
ConEd 7¾02	14.0	10	56	55½	55½	- ¼	
ConEd 7¾03	14.3	5	54¼	54¼	54¼	- ¼	
ConEd 8.60s03	14.3	15	59¾	58¾	58¾	- ¼	
ConEd 9¾04	14.3	11	63¾	63¾	63¾	- 1¼	
ConEd 9¾05	13.9	10	69¾	69¾	69¾	.....	
ConEd 12¾07	15.3	5	80¼	80¼	80¼	+ ¼	
ConEd 9¾08	14.3	5	67¾	67¾	67¾	+ ¼	
ConEd 8¾09	14.3	15	61¾	61¾	61¾	+ ¼	
ConEd 9¾10	14.3	11	63¾	63¾	63¾	- ¼	
ConEd 9¾84	14.3	11	68¾	68¾	68¾	.....	
ConEd 10¾85	14.5	10	71¾	70	71¾	+ ¾	
ConEd 9¾85	11.4	13	83½	83¼	83½	- ¼	
ConEd 9.55s09	14.2	13	68¾	68	68¾	+ ¼	
ConEd 11¾83	12.4	12	96¾	95¾	96¾	.....	
ConEd 12.60s87	13.0	42	97	97	97	.....	
ConAC 6s94	cv	10	101½	101½	101½	+ ½	
ConAC 9¾86	13.8	10	65	64½	65	+ ½	
Cnsl Ppr 9¼01	cv	29	80	79½	79½	- 1½	
Contl 8¾84	8.7	10	100¼	100¼	100¼	.....	
Contl 8¾85	9.2	15	96¾	96¾	96¾	.....	
Contl 8¾86	9.1	10	96	96	96	.....	
Contl 8¾87	9.2	30	95½	95	95½	+ ¼	
Contl 8¾88	9.2	10	96	95¾	96	+ ¼	
Contl 8¾89	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾90	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾91	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾92	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾93	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾94	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾95	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾96	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾97	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾98	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾99	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾00	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾01	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾02	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾03	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾04	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾05	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾06	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾07	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾08	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾09	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾10	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾11	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾12	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾13	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾14	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾15	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾16	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾17	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾18	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾19	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾20	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾21	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾22	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾23	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾24	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾25	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾26	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾27	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾28	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾29	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾30	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾31	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾32	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾33	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾34	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾35	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾36	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾37	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾38	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾39	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾40	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾41	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾42	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾43	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾44	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾45	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾46	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾47	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾48	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾49	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾50	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾51	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾52	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾53	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾54	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾55	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾56	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾57	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾58	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾59	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾60	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾61	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾62	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾63	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾64	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾65	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾66	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾67	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾68	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾69	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾70	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾71	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾72	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾73	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾74	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾75	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾76	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾77	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾78	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾79	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾80	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾81	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾82	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾83	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾84	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾85	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾86	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾87	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾88	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾89	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾90	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾91	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾92	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾93	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾94	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾95	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾96	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾97	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾98	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾99	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾00	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾01	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾02	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾03	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾04	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾05	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾06	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾07	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾08	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾09	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾10	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾11	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾12	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾13	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾14	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾15	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾16	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾17	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾18	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾19	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾20	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾21	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾22	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾23	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾24	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾25	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾26	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾27	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾28	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾29	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾30	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾31	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾32	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾33	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾34	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾35	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾36	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾37	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾38	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾39	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾40	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾41	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾42	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾43	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾44	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾45	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾46	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾47	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾48	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾49	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾50	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾51	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾52	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾53	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾54	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾55	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾56	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾57	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾58	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾59	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾60	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾61	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾62	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾63	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾64	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾65	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾66	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾67	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾68	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾69	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾70	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾71	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾72	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾73	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾74	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾75	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾76	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾77	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾78	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾79	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾80	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾81	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾82	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾83	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾84	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾85	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾86	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾87	9.2	10	96¼	96	96¼	+ ¼	
Contl 8¾88	9

Economic Problems

The second largest section in the newspaper involving statistical information could be classified under the heading of economic problems, both US and international. This took up 4.4% of the paper's information content: the 12 articles in it covered such diverse topics as

17. Senators Prepare to Vote on Budget with $784 Billion
18. Consumer Prices Up Only 0.2%; Energy Costs Off Sharply in April
19. 252 Pieces of Federal Land Picked for Possible Sale, Watt Says
20. Living Costs Up Slightly in the New York Region
21. Furniture Freeze Violator
22. Money Supply Up $2.3 Billion
23. Orders for Durables Dropped 3.9% in April
24. Continental to Buy STSC
25. Harvester Statement
26. Stocks Halt Slide With Dow Up 3.2
27. Dominicans Fear US Sugar Quotas
28. OPEC Rein on Output Extended: Group to keep Oil Price at $34 a Barrel

To give the flavour of these articles, and of their statistical contents let me select two among them for analysis: one of them relates to the US economy (item No.18) and the other to its effect on the Dominican Republic (item No.27).

In article No.18, it was reported that the Consumer Price Index in the US had declined by 0.3% in March 1982 (see Figure 3) for the first time in almost 17 years. Although there had been an increase of 0.2% in the Index in April this continued to reflect the slowing of inflation in the US over the past 7 months. The total increase in the Index over the last 12 months was 6.6%, the lowest since 1978; the expected rate of increase in 1982 was 6-6.5% as against 8.9% in 1981. However, at the same time in 1981, when the Index had increased by 10% over the preceding 12 months, the unemployment rate was 7.3%, whereas the current Index increase of 6.6% was associated with an unemployment rate of 9.4%, threatening to break 10%. A detailed breakdown of costs for medical expenses, housing, food, alcohol, and restaurant meals was given, together with a graph of the Index for February, March and April 1982, comparing its value for the whole of the US with that for New York and Northeastern New Jersey.

The statistical arguments here are somewhat more complex: we are dealing with rates of change in the consumer Price Index, and related changes in unemployment. Readers are now expected to understand the economic terminology, the basic statistical data, and the concept of rate of change which leads to the conclusion that the rate of inflation is gradually slowing down. In addition, they must be aware of the possible random variation of the percentage changes

FIGURE 3

Consumer Prices Up By Only 0.2% in April

Continued From Page 1

the 8.9 percent rate for last year and half the double-digit rates of 1979 and 1980. Treasury Secretary Donald T. Regan said this week that the rate would be in the 6 to 6.5 percent range.

Donald Ratajczak, director of economic forecasting at Georgia State University, said that getting to this level for the year meant an average rise of around five-tenths of 1 percent a month for the rest of the year.

"The news continues to be good," said Larry Speakes, the White House deputy press secretary. "The trend toward lower inflation is continuing. Compared to double-digit inflation a year ago, the trends are in a favorable direction."

Representative Guy Vander Jagt of Michigan, the chairman of the Republican Campaign Committee, said that "the Democrats can stamp their feet and hold their breath until they are blue, but it won't change facts: Republicans are beginning to make things better."

Most economists agree that the decline in inflation is a result of the sharp recession, which has pushed unemployment to a postwar record, the decline in energy prices because of the worldwide oil surplus and the very small increase in food prices.

Last year at this time, the unemployment rate was 7.3 percent and the Consumer Price Index, for the preceding 12 months, was up 10 percent. This month, the 12-month rate is down to 6.6 percent but the unemployment rate is up to 9.4 percent and threatening to break 10 percent.

The rise in the index for April translates to a compound annual rate of 3 percent. The compound annual rate is based on a 0.2470879 rise in the index, which is then rounded off to the figure of two-tenths of 1 percent that was reported today. The numbers are adjusted to remove normal seasonal fluctuations in prices.

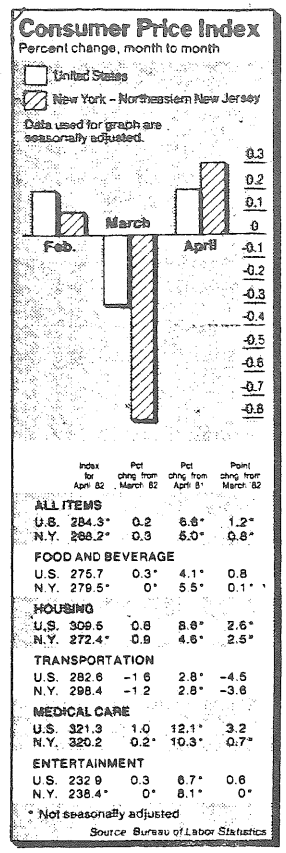
The largest price increases in the month were for medical care, which has resisted the downward trend in inflation. Medical care costs were up 1 percent for the month and 12.1 percent in the last 12 months.

Housing costs and food prices, after declining in March, were up in April. The price of houses, for reasons analysts have trouble explaining, was up 1.2 percent and mortgage interest rates inched up six-tenths of 1 percent.

Housing Can Distort Index

The housing component, because it emphasizes interest rates and house prices, can distort the Consumer Price Index on a monthly basis. The Labor Department's experimental index, which treats housing costs as rent, declined in April by two-tenths of 1 percent. For the last 12 months, this experimental index is up 6 percent, just below the 6.6 percent in the regular index.

Food prices at the grocery store were up three-tenths of 1 percent, following a decline of three-tenths of 1 percent in March. Beef and pork prices rose sharply, pushing the index for meat, poultry, fish and eggs up 1.1 percent. Prices for eggs, fruits and vegetables fell. Prices for alcoholic beverages were up three-tenths of 1 percent and restaurant meals rose four-tenths of 1 percent.



	Index for April '82	Pct chng from March '82	Pct chng from April '81	Point chng from March '82
ALL ITEMS				
U.S.	284.3*	0.2	6.6*	1.2*
N.Y.	288.2*	0.3	5.0*	0.8*
FOOD AND BEVERAGE				
U.S.	275.7	0.3*	4.1*	0.8
N.Y.	279.5*	0*	5.5*	0.1*
HOUSING				
U.S.	309.5	0.8	8.8*	2.6*
N.Y.	272.4*	0.9	4.6*	2.5*
TRANSPORTATION				
U.S.	282.6	-1.6	2.8*	-4.5
N.Y.	298.4	-1.2	2.8*	-3.6
MEDICAL CARE				
U.S.	321.3	1.0	12.1*	3.2
N.Y.	320.2	0.2*	10.3*	0.7*
ENTERTAINMENT				
U.S.	232.9	0.3	8.7*	0.6
N.Y.	238.4*	0*	8.1*	0*

* Not seasonally adjusted

Source Bureau of Labor Statistics

The New York Times, May 22, 1982

© New York Times Company, 1982
Reprinted with permission

in the CPI, which may mean that there is no significant difference in the February/March/April figures. The article also raises the possibility of statistical correlation between unemployment and inflation.

In article No.27 we are launched into the international arena . To help domestic sugar producers the USA is proposing a quota on sugar imports into the country, and the Dominican Republic is to get nearly 18% of this quota. However, this could still cut its present exports of sugar to the USA by one half, causing a decline of $209 million in revenue. Since 100,000 people are employed in the Dominican sugar industry, this could cause further unemployment in a

country with a current unemployment rate of 30%. Because more than 60% of the sugar is produced and exported directly by the government, each pound shipped at a loss creates a larger deficit and hence higher inflation in the country. It costs the Dominican government 17 cents to produce one pound for export, while the private company Gulf and Western Industries can do it for 12 cents. The government, unlike the company, has not been able to modernize its equipment. The present world market price is just over 8 cents a pound; at current market rates, the Dominican Republic would therefore lose between 6-9 cents per pound. The Central Bank estimates that 1.13 million metric tons of sugar will be exported this year. The Dominican government has been hard pressed financially, and has not been able to find the capital to invest in a change of its land usage from sugar to alternative cash crops.

Once again readers need to understand some economic concepts: quotas, government subsidies, inflation, unemployment. They must also appreciate the comparative statistics of state and private production, and the rationale of investment in agricultural and technical improvements. There is, in addition, the possibility of a misuse of statistics; of the 100,000 employed in the sugar industry, it is not made clear how many are employed in government, and how many in private industry.

Social Issues

The third largest section containing statistical information was concerned with social isues, exclusively in the USA. This took up 3.5% of the paper's information content: the 7 articles in it were

29. Disability Cutoffs Assailed by Panel
30. 107 Seized in "Smoke Shop" Raids Arraigned on Charges in New York
31. Drops in Diseases Traced to Policy
32. Is Court the Right Place for Wayward Youths?
33. Lower Rises Seen for Some Rentals
34. For the Devotees of Liverwurst, More Good News
35. Caution on College-Cost Advisory Services

These include discussions on social security, criminological, epidemiological, housing, education and (amusingly enough) gastronomic problems. We shall again select two of these for analysis.

In article No.29, Members of Congress criticized the Social Security Administration for the way in which it was removing people "with callous indifference" from the disability insurance program. The US government pays more than $17 billion a year in disability benefits to 4.5 million disabled workers and their families, the average benefit being $419 a month. In the last 15 months there

has been evidence of a rather hard drive by the federal government to remove many people entitled to benefits from the disability rolls. In the first 13 months of accelerated reviews, 436,000 cases were examined, and 174,000 or 40% of them were ruled ineligible. The article requires an understanding of social science terminology, a grasp of statistical averages, and an appreciation of whether ineligibility is a likely event in such a review, or whether it is significantly different from the previous rejection rate.

Article No.31 was concerned with the New York City policy of allowing only immunized students to enter the public school system; it was claimed that this policy had rid the city of measles, and reduced the incidence of rubella and mumps. In September, when the policy was put into effect 16% of the city's 907,000 public school students had not been immunized; 67,000 had been initially excluded at the end of September, but 47,000 of these were permitted to return in the following days if they produced the required parental consent for immunization. Only 3,163 were still being excluded because they lacked these, or around 0.35% of the total number.

The policy appeared to have been successful. Only 11 cases of measles had been reported in the first five months of 1982, as against 859 in 1979, 1,210 in 1980, and 106 in 1981. The Deputy Mayor had stated that no cases of rubella and only 30 cases of mumps had been reported. But the Director of the Health Department's immunization program reported that there had been 21 cases of rubella in the first 5 months, compared with 196 in the same period in 1979, and 30 of mumps as compared with 67 in the same period in 1975. It was clear from the article that there was a discrepancy in the Deputy Mayor's and the Health Department's reports; notwithstanding these, the program does appear to have had a positive effect.

Here, the reader needs to understand basic ideas of significance; are the results quoted truly indicative that the program has been successful? Or could the results be random? In this case the figures are fairly clear: immunization has indeed proved to be successful.

Political Arithmetic

Political issues involving statistical data, all of them international, occupied 1.7% of the information content of the paper. The 3 articles in this section were:

36. Britain Seizes a Beach Head in Falklands; Reports Downing 17 Planes, Had 5 Ships Hit.
37. State Votes Hurt Mrs. Gandhi
38. In the Namibia Talks, a New Approach.

Article 36 was full of statistics of planes lost, ships sunk, and numbers killed and wounded by the combatants. The statistical data were essentially descriptive, but there were serious discrepancies between the Argentine and the British figures. No advanced statistical concepts were necessary to understand the article.

Perhaps the most interesting of the three articles was item No.38 about the five Western Nations trying to mediate between South Africa and the South West Africa People's Organization (SWAPO) to negotiate an independence plan for Namibia. The article in fact contains no statistical data at all, but raises very important statistical questions. It deals with a most difficult problem of voting formulae, and how these reflect the composition of the electorate. A complex voting formula has been proposed, mixing a proportional representation system with direct election on the basis of defined constituencies. South Africa was insisting on the mixed voting system, in which each ballot would, in effect, be counted twice. This would ensure that right-wing whites supporting the National Party would be represented on the Assembly drafting the constitution of an independent Namibia. Since whites are almost everywhere outnumbered by blacks, only the proportional representation system would achieve this. But this would equally favour SWAPO, whose support is heavily concentrated in the North of the country. Since South Africa's other aim is to see that insurgents fall short of the two thirds vote necessary to control the Assembly, it also favours direct voting to dilute SWAPO's support.

This article brings up some very complex statistical issues, though we can understand their basic principles well enough. We all know that the British system of voluntary voting, with the election going to the candidate with the majority vote, is different from the Australian one, with compulsory voting and the preferential transfer of votes from a losing candidate to one still in the race. The article describes the manipulations leading to the political advantage of one or other party. Of such political arithmetic is the government of countries made up, and it requires quite a sophisticated statistical understanding to grasp its finer points.

Sporting Records

These records occupied 1.7% of the paper and listed the detailed statistics of baseball games played throughout the USA, the tennis scores in the Italian Open at Rome and the German Open at West Berlin, and the seeding of both men (1. Jimmy Connors, USA) and women players (1. Chris Evert-Lloyd, USA) at the French Open at Paris due to be played on Monday 24 May (see Figure 4).

Golf scores for the Atlanta Classic played on Thursday 20 May were also listed as were the soccer results for various US soccer leagues. Such statistics are essentially descriptive, but could certainly serve as predictive data for the betting man.

Weather Information

Lastly, the weather took up 1.1% of the paper, a full 1.8 columns of the back page, over one third of it statistical in type (see Figure 5). On that day the forecast for Washington was "rain likely", temperature 73-61°F and for Minneapolis "mostly cloudy", 66-48°F. The temperatures and weather conditions on the previous day in London at noon were "cloudy", 64°F and in Buenos Aires at 9.00 a.m. "cloudy", 41°F. Once again these statistics were essentially descriptive; the forecasts and synoptic chart, while relying on complex numerical and statistical methods, were not accompanied by technical details.

Let us now summarize our survey. The contents of the New York Times of 22 May 1982, which involved at least some statistical material, can now be broken up into the following categories:

Category	Percentage space of all paper
Financial Reports	20.2%
Economic Problems	4.4%
Social Issues	3.5%
Political Arithmetic	1.7%
Sporting Records	1.7%
Weather Information	1.1%
TOTAL	32.6%

None of these categories involved difficult statistics, though a deeper scrutiny of some of the articles would inevitably lead to statistical complexities. Statistical concepts arose mostly in the financial, economic and social fields, with the recreational (sporting records) and the scientific fields (weather information) forming only a small proportion of the total. One may conclude from this that since all men are social creatures, it is most likely that they will encounter statistics in the economic and social contexts. It is clear that with nearly one third of the newspaper involving statistical information, the subject cannot be easily avoided; statistical training, at least at a basic level, must become an essential ingredient in the balanced education of every citizen.

© New York Times Company, 1982
Reprinted with permission

FIGURE 4

Baseball Records and Results

National League

Based on 65 or more times at bats.

	G	AB	R	H	Pct.
J.Thpsn, Pitt.	33	120	26	44	.367
Moreind, Chi ...	38	142	19	50	.352
Ballor, Mets ...	24	72	10	25	.347
Ru.Jones, S.D.	32	128	29	42	.328
Raines, Mont. ..	33	135	15	44	.326
Stearns, Mets ..	33	124	19	40	.323
Lo.Smith, St.L..	38	156	35	50	.321
Knight, Hou.	39	150	23	48	.320
Guerrero, L.A. .	38	144	23	46	.319
Wilson, Mets	37	157	25	50	.318

HOME RUNS

Kingman, NewYork, 13; Murphy, Atlanta, 12; J. Thompson, Pittsburgh, 11; Horner, Atlanta, 10; Hendrick, St. Louis, 9; Guerrero, Los Angeles, 9.

RUNS BATTED IN

Murphy, Atlanta, 37; Kingman, Mets, 35; Moreland, Chicago, 33; J. Thompson, Pittsburgh, 30; B. Diaz, Philadelphia, 27; Horner, Atlanta, 27.

STOLEN BASES

Moreno, Pittsburgh, 23; Dernier, Philadelphia, 21; Lo. Smith, St. Louis, 18; Wilson, Mets, 15; Raines, Montreal, 12.

RUNS

Lo. Smith, St. Louis, 35; Murphy, Atlanta, 34; Horner, Atlanta, 30; Ru.Jones, San Diego, 29; Dawson, Montreal, 27; Kingman, Mets, 27.

HITS

Moreland, Chicago, 50; Wilson, Mets, 50; Lo. Smith, St. Louis, 50; Knight, Houston, 48; Oester, Cincinnati, 47.

DOUBLES

Garner, Houston, 12; Lo. Smith, St. Louis, 11; Knight, Houston, 11; T. Kennedy, San Diego, 11; B. Diaz, Philadelphia, 10; Parker, Pittsburgh, 10; O. Smith, St. Louis, 10; Horner, Atlanta, 10.

TRIPLES

Wilson, Mets, 3; Rose, Philadelphia, 3; Herr, St. Louis, 3; R. Ramirez, Atlanta, 3; Concepcion, Cincinnati, 3; Oester, Cincinnati, 3; Knight, Houston, 3.

PITCHING (5 Decisions)

Forsch, St. Louis, 5-0, 1.000, 3.86; Sutton, Houston, 6-1, .857, 2.40; Ra. Jones, Mets, 5-2, .714, 3.15; Puleo, Mets, 4-2, .667, 3.72; Garber, Atlanta, 4-2, .667, 1.67; Berenyi, Cincinnati, 4-2, .667, 3.35; Welch, Los Angeles, 4-2, .667, 3.31; Rogers, Montreal, 5-3, .625, 1.66.

STRIKEOUTS

Carlton, Philadelphia, 71; Soto, Cincinnati, 66; Rogers, Montreal, 51; Collar, San Diego, 43; Valnzuela, Los Angeles, 42.

(Last night's games not included.)

LAST NIGHT'S GAMES

Mets at Houston.
Chicago at San Francisco.
Montreal at Cincinnati.
Philadelphia at Atlanta
Pittsburgh at San Diego.
St. Louis at Los Angeles.

EASTERN DIVISION

	W.	L.	Pct.	G.B.
St. Louis	24	15	.615	—
Mets	21	17	.553	2½
Philadelphia	19	18	.514	4
Montreal	16	19	.457	6
Pittsburgh	15	20	.429	7
Chicago	16	22	.421	7½

WESTERN DIVISION

Atlanta	25	13	.658	—
San Diego	19	17	.528	5
Los Angeles	19	20	.487	6½
Houston	18	21	.462	7½
San Francisco	17	22	.436	8½
Cincinnati	16	21	.432	8½

(Last night's games not incl.)

TODAY'S GAMES

Mets at Houston (8:35 P.M.) — Falcone (3-1) vs. Ryan (3-6).
Chicago at San Francisco (4:05 P.M.) — Ripley (1-0) vs. Hammaker (2-1).
Montreal at Cincinnati (7:05 P.M.) — Sanderson (3-3) vs. Soto (3-3).
Philadelphia at Atlanta (2:20 P.M.) — Krukow (3-2) vs. Walk (4-3).
Pittsburgh at San Diego (10:05 P.M.) — Sarmiento (1-0) vs. Eichelberger (3-5).
St. Louis at Los Angeles (10:05 P.M.) — Rincon (2-2) vs. Power (0-1).
(All times Eastern daylight.)

TOMORROW'S GAMES

Mets at Houston (n.).
Chicago at San Francisco (2).
Montreal at Cincinnati.
Philadelphia at Atlanta.
Pittsburgh at San Diego.
St. Louis at Los Angeles.

American League

Based on 65 or more times at bat

	G	AB	R	H	Pct
Bonnell, Tor. ..	31	90	17	34	.422
Harrah, Clev. ..	35	137	22	54	.394
McBride, Clev.	25	80	8	31	.388
Wshgtn, Minn. .	29	91	8	32	.352
E.Murry, Balt..	33	100	14	35	.350
Paciorek, Chi. ..	34	129	15	44	.341
Cooper, Milw. ..	35	141	21	48	.340
Lwnstn, Balt. ...	25	65	15	22	.338
Sundberg, Tex.	30	107	10	36	.336
Almon, Chi.	34	108	18	36	.333

HOME RUNS

Thornton, Cleveland, 11; Hrbek, Minnesota, 10; Roenicke, Baltimore, 9; Harrah, Cleveland, 9; Murphy, Oakland, 8.

RUNS BATTED IN

McRae, KansasCity, 37; Thornton, Cleveland, 36; Luzinski, Chicago, 29; Ogilvie, Milwaukee, 27; Rice, Boston, 26; Baylor, California, 26; Hrbek, Minnesota, 26.

STOLEN BASES

R. Henderson, Oakland, 39; LaFlore, Chicago, 14; Molitor, Milwaukee, 11; Lopes, Oakland, 11; Wathan, Kansas City, 9.

RUNS

R. Henderson, Oakland, 36; Harrah, Cleveland, 32; Thornton, Cleveland, 29; LeFlore, Chicago, 28; Dauer, Baltimore, 27; Evans, Boston, 27; Wathan, Kansas City, 27.

HITS

Harrah, Cleveland, 54; Cooper, Milwaukee, 48; LeFlore, Chicago, 48; Garcia, Toronto, 46; Bochte, Seattle, 45.

DOUBLES

Otis, Kansas City, 14; White, Kansas City, 12; Evans, Boston, 11; McRae, Kansas City, 11; Cowens, Seattle, 11.

TRIPLES

Herndon, Detroit, 5; G.Wright, Texas, 4; Evans, Boston, 3; McBride, Cleveland, 3; Yount, Milwaukee, 3; Upshaw, Toronto, 3; Morrison, Chicago, 3; Cowens, Seattle, 3.

PITCHING (5 Decisions)

Hoyt, Chicago, 8-0, 1.000, 1.53; Guidry, Yankees, 6-1, .857, 2.95; Tudor, Boston, 4-1, .800, 3.26; Renko, California, 4-1, .800, 1.86; Gura, Kansas City, 4-1, .800, 4.17; Beard, Oakland, 4-1, .800, 2.91; Clancy, Toronto, 5-2, .714, 4.21; Zahn, California, 5-2, .714, 2.35. F. Bannister, Seattle, 5-2, .714, 3.19.

STRIKEOUTS

F. Bannister, Seattle, 58; Eckersley, Boston, 47; Perry, Seattle, 47; Guidry, Yankees, 46; Denny, Cleveland, 40.

(Last night's games not included.)

Cubs 8, Dodgers 3

THURSDAY NIGHT

CHICAGO	ab	r	h	bi	LOS ANGELES	ab	r	h	bi
Hendsn lf	3	1	1	1	Sax 2b	5	0	2	0
Woods lf	1	0	0	0	Landrx cf	3	1	1	0
Wills 2b	4	1	1	0	Baker lf	3	0	0	0
Bucknr 1b	4	2	2	1	Guerer rf	3	0	0	1
Morelnd rf	3	1	2	3	Cey 3b	4	0	0	0
Durham cf	3	1	1	2	Garvey 1b	4	0	0	0
H.Cruz lf	1	0	0	0	Scioscia c	3	2	2	0
J.Davis c	3	0	2	1	Russell ss	4	0	0	1
Bowa ss	4	0	0	0	Valnzuel p	1	0	1	1
Sndbrg 3b	4	2	2	0	Johnstn ph	1	0	0	0
Jenkins p	2	0	0	0	Forster p	0	0	0	0
					Orta ph	1	0	0	0
					A Pena p	0	0	0	0
					Mondy ph	0	0	0	0

Total	32	8	11	8	Total			33	3	6	3
Chicago					429	011	000	—	8
Los Angeles					010	000	011	—	3

	IP	H	R	ER	BB	SO
Chicago						
Jenkins W,3-4	9	6	3	1	4	8
Los Angeles						
Valnzuela L,5-4	5	8	7	5	2	7
Forster	2	2	1	0	0	1
Orta	2	1	0	0	2	1

Mets Records

BATTING

	ab	h	hr	rbi	avg.
Hodges	18	8	1	6	.444
Veryzer	26	10	0	1	.385
Ballor	72	25	0	11	.347
Sweet	3	1	0	0	.333
Stearns	124	40	3	10	.323
Wilson	157	50	2	17	.318
Backman	75	23	0	5	.307
Jorgensen	11	3	0	1	.273
Youngblood	89	24	2	11	.270
Brooks	131	34	0	17	.260
Kingman	131	32	13	35	.244
Gardenhire	107	26	0	11	.243
Valentine	68	16	0	8	.235
Foster	130	28	4	16	.215
Raisich	43	8	1	3	.186
Staub	26	3	1	4	.115
Team — AB, 1298, H, 340, HR, 27, RBI, 150 .262					

PITCHING

	w	l	ip	so	era
Falcone	3	1	41	31	3.76
Zachry	3	1	42	16	4.71
Jones	5	2	60	20	3.15
Puleo	4	2	48	19	3.72
Swan	2	2	29	1	.84
Scott	3	2	53	22	6.36
[illegible]	1	2	24	4	3.33
[illegible]	0	2	21	6	7.53
Norman	0	4	26	16	7.77
Frazier	1	0	22	9	2.49
Guidry	6	1	58	46	2.95
Morgan	3	1	32	15	2.78
Rawley	2	1	22	22	2.86
Erickson	4	4	48	18	5.06
John	3	4	56	18	2.59
Righetti	2	3	39	25	5.26
Alexander	0	2	16	5	3.94
[illegible]	0	2	25	30	2.98
Way	0	2	32	24	1.78

(Last night's game not included.)

Yankees Records

BATTING

	ab	h	hr	rbi	avg.
Piniella	59	21	1	9	.356
Hobson	18	6	0	2	.333
Randolph	120	38	2	16	.317
Winfield	126	38	5	25	.302
Mumphrey	102	30	0	11	.294
Griffey	103	28	1	6	.272
Nettles	50	13	1	3	.260
Wynegar	100	25	2	12	.250
Murcer	45	11	1	8	.244
Collins	81	19	0	3	.235
Cerone	86	20	1	8	.233
Balboni	26	6	1	3	.231
Mayberry	66	15	3	8	.227
Smalley	113	25	2	11	.221
Robertson	10	2	1	1	.200
Dent	82	13	0	5	.159
Gamble	58	9	2	6	.155
Foote	30	4	0	1	.133
Team — AB, 1227 H, 308 HR, 19, RBI, 132 .250					

PITCHING

	w	l	ip	so	era
Frazier	1	0	22	9	2.49
Guidry	6	1	58	46	2.95
Morgan	3	1	32	15	2.78
Rawley	2	1	22	22	2.86
Erickson	4	4	48	18	5.06
John	3	4	56	18	2.59
Righetti	2	3	39	25	5.26
Alexander	0	2	16	5	3.94
[illegible]	0	2	25	30	2.98
Way	0	2	32	24	1.78

(Last night's game not included.)

Tigers 11, A's 3

THURSDAY NIGHT

OAKLAND	ab	r	h	bi	DETROIT	ab	r	h	bi
RHndn lf	4	0	0	0	Lemon rf	3	1	1	0
Meyer rf	5	0	1	0	Jones rf	1	0	0	0
Jhnsn 1b	5	0	0	0	Cabell 1b	5	1	2	0
Gross 3b	3	2	1	0	Gibson cf	3	2	1	3
Godwn dh	3	1	3	2	Ivie dh	5	2	2	4
Lopes 2b	3	0	0	0	Herndon lf	5	2	3	1
Newman c	3	0	1	1	L.Parrish c	3	1	1	2
Bosetti cf	4	0	1	0	Brokns 3b	4	1	2	1
Stanly ss	2	0	0	0	Tramml ss	4	0	0	0
Burghs ph	1	0	0	0	De.Johns ss	1	0	0	0
					Whitakr 2b	5	1	3	0

Total	33	3	7	3	Total			39	11	17	11
Oakland					000	200	010	—	3
Detroit					201	500	12x	—	11

E—Meyer. DP—Oakland 1. LOB—Oakland 9, Detroit 11. 2B—Meyer 3B—Herndon 2. HR—Goodwin (2), Ivie 2 (4), L.Parrish (6). SF—Newman.

	IP	H	R	ER	BB	SO
Oakland						
Flowers L,0-1	1	3	4	4	2	1
Owchinko	1	3	4	2	2	1
Klimek	1	1	2	1	1	0
Detroit						
Wilcox W,3-2	9	7	3	3	1	5

Orioles 6, Twins 0

THURSDAY NIGHT

MINNESOTA	ab	r	h	bi	BALTIMORE	ab	r	h	bi
Wshgt ss	4	0	1	0	Bumbry cf	3	0	0	0
Mlls,krn 2b	4	0	1	0	Dauer 2b	4	0	0	0
Vega 1b	4	0	1	0	Lownstn lf	3	0	0	0
Brmsky cf	4	0	0	0	Singtn dh	4	1	1	1
Ward lf	4	0	1	0	Crowly 1b	4	1	2	0
Engler f	3	0	0	0	Dwyer rf	2	2	0	0
Castino 3b	3	0	0	0	Nolan c	4	2	2	1
Gaetti dh	3	0	1	0	Ripken 3b	3	0	2	3
Laudnr c	3	0	0	0	Sakata ss	2	0	0	1

Total	32	0	5	0	Total			29	6	7	6
Minnesota					000	000	000	—	0
Baltimore					030	300	00x	—	6

DP—Minnesota 1. LOB—Minnesota 5, Baltimore 4. 2B—Crowley, Ripken, Nolan. HR—Singleton (3). SB—Bumbry. SF—Sakata.

	IP	H	R	ER	BB	SO
Minnesota						
Redfern L,5-8	8	7	6	6	4	3
Baltimore						
McGregor W,5-1	9	5	0	0	0	5

T—1:58. A—12,996.

N. B. A.

Red Sox 11, Mariners 2

THURSDAY NIGHT

SEATTLE	ab	r	h	bi	BOSTON	ab	r	h	bi
Brown lf	4	0	0	0	Remy 2b	2	2	0	0
Serna 2b	4	0	3	0	Evans rf	5	2	3	0
Bochte 1b	4	0	0	0	Rice lf	4	2	2	5
Cowens rf	4	0	2	0	Ystzmsk dh	4	1	2	2
T.Cruz ss	3	1	1	0	Boggs dh	0	0	0	0
Meier dh	4	0	1	0	Lansfrd 3b	5	2	3	0
Hndsn cf	4	1	1	0	Stapltn 1b	4	1	1	0
Bulling c	3	0	0	0	Gedman c	4	1	1	3
Flrova c	1	0	0	0	Hoffmn ss	2	0	1	1
J.Cruz 2b	3	0	1	2	Valdez ss	1	0	0	0
					R.Miller cf	1	0	0	0
					Nichols cf	3	0	1	0

Total	34	2	9	2	Total			35	11	16	11
Seattle					000	000	200	—	2
Boston					013	313	08x	—	11

DP—Seattle 2, Boston 2. LOB—Seattle 6, Boston 7. 2B—Evans, Rice, Cowens. HR—Ystrzemski (6), Rice (6), Gedman (2).

	IP	H	R	ER	BB	SO
Seattle						
M.Moore L,1-5	3	6	7	7	1	0
Clark	3	6	4	4	5	0
Seattle	2	2	0	0	1	3
Boston						
Tudor W,4-1	9	9	2	2	1	4

T—2:43. A—14,587.

Blue Jays 2, Indians 0

THURSDAY NIGHT

TORONTO	ab	r	h	bi	CLEVELAND	ab	r	h	bi
Garcia 2b	5	0	2	1	A.Bnstr 2b	3	0	0	0
Mulinks 3b	4	0	1	0	Harrah 3b	4	0	2	0
Iorg 3b	0	0	0	0	Hargrv 1b	3	0	0	0
Powell rf	4	0	0	0	Thrntn dh	4	0	2	0
Barfield rf	0	0	0	0	McBridrf	4	0	0	0
Upshaw 1b	4	0	3	0	Hayes lf	4	0	2	0
Revring dh	3	0	0	0	Mannng cf	4	0	0	0
Nrdhgn dh	1	0	0	0	Hassey c	3	0	1	0
Woods lf	3	1	1	0	Craig pr	0	0	0	0
Bonnell lf	1	0	0	0	Dybzkl ss	3	0	1	0
Moseby cf	4	0	0	0	Charbnph	1	0	0	0
Whitt c	3	1	2	1					
Griffin ss	4	0	1	0					

Total	36	2	10	2	Total			33	0	8	0
Toronto					000	000	200	—	2
Cleveland					000	000	000	—	0

E—Garcia. DP—Toronto 3. LOB—Toronto 8, Cleveland 9. 2B—Hayes, Hassey.

	IP	H	R	ER	BB	SO
Toronto						
Clancy W,5-2	9	8	0	0	3	4
Cleveland						
Sorensen L,3-3	7 ⅓	10	2	2	0	2
Glynn	1 ⅔	0	0	0	1	1

T—2:19. A—6,226.

Eastern Conference

BOSTON VS. PHILADELPHIA

May 9 — Boston 121, Phila. 81.
May 12 — Phila. 121, Boston 113.
May 15 — Phila. 99, Boston 97.
May 16 — Phila. 119, Boston 94.
May 19 — Boston 114, Phila. 85.
Last night — at Philadelphia.
*Tomorrow — at Boston, 3:30 P.M.

(Philadelphia leads series, 3-2)

*-If necessary.
All times Eastern daylight.

Western Conference

SAN ANTONIO VS. LOS ANGELES

May 9 — L.A. 128, San Antonio 117.
May 11 — L.A. 110, San Antonio 101.
May 14 — L.A. 118, San Antonio 108.
May 15 — L.A. 128, S.A. 123.
(Los Angeles won series, 4-0)

Tennis

AT ROME

ITALIAN OPEN

Quarterfinals

Elliot Teltscher, U.S, d. Eddie Dibbs, U.S., 6-2, 6-2; Andres Gomez, Ecuador, d. Jose Higueras, Spain, 6-2, 3-6, 6-3; Pablo Arraya, Peru, d. Wojtek Fibak, Poland, 2-6, 6-3, 6-2; Mats Wilander, Sweden, d. Tomas Smid, Czechoslovakia, 6-1, 6-1.

AT WEST BERLIN

GERMAN OPEN

Quarterfinals

Bonnie Gadusek, U.S., d. Hana Mandlikova, Czechoslovakia, 7-5, 6-3; Kathy Rinaldi, U.S., d. Mima Jausovec, Yugoslavia, 4-6, 6-3, 7-5; Bettina Bunge, U.S., d. Duk Hee Lee, South Korea, 6-0, 6-1; Sylvia Hanika, West Germany, d. Andrea Leand, U.S, 6-2, 7-5.

FRENCH OPEN AT PARIS

Seedings for the $400,000 French Open tennis championships which start Monday in Paris:

Men's

1, Jimmy Connors, U.S. 2, Ivan Lendl, Czechoslovakia. 3, Guillermo Vilas, Argentina. 4, Jose-Luis Clerc, Argentina. 5, Vitas Gerulaitis, U.S. 6, Eliot Teltscher, U.S. 7, Peter McNamara, Australia. 8, Yannick Noah, France. 9, Andres Gomez, Ecuador. 10, Balazs Taroczy, Hungary. 11, Brian Gottfried, U.S. 12, Eddie Dibbs, U.S. 13, Jose Higueras, Spain. 14, Steve Denton, U.S. 15, Chip Hooper, U.S. 16, Mel Purcell, U.S.

Women's

1, Chris Evert-Lloyd, U.S. 2, Martina Navratilova, U.S. 3, Tracy Austin, U.S. 4, Andrea Jaeger, U.S. 5, Hana Mandlikova, Czechoslovakia. 6, Sylvia Hanika, West Germany. 7, Mima Jausovec, Yugoslavia. 8, Anne Smith, U.S. 9, Bettina Bunge, U.S. 10, Billie Jean King, U.S.

THURSDAY NATIONAL TENNIS CENTER

LUPUS FOUNDATION TOURNAMENT

Men's Singles

Ricky Meyer d. Edmund Plass, 6-2, 6-1; Tom Jones d. Paul Moss, 7-5, 26, 6-1; Ken Barton d. Scott Seaman, 6-3, 6-4; Chris Garner d. Michael Elnsidler, 6-4, 7-6. Jeff Pabell d. Todd Lazarus, 6-1, 6-2, Eric Fromm d. Vic Casone, 6-1, 6-4, Mike Kazimek d. Larry Hirsch, 6-1, 6-2, John James d. Yaya Doumbia, 6-2, 6-0.

Golf

THURSDAY AT ATLANTA ATLANTA CLASSIC

The Leading Scores

(54 golfers were unable to complete the round because of weather)

Larry Nelson	33—66
Keith Fergus	34—66
Howard Twitty	32—67
Peter Oosterhuis	35—67
Peter Jacobsen	34—68
Hubert Green	34—68
DeWitt Weaver	33—69
Pat McGowan	35—69
Calvin Peete	35—69
Barry Harwell	33—69
Wayne Levi	33—69
Doug Black	34—69
Mark Pfeil	34—69
Jack Newton	36—69
Roger Maltbie	35—70
Mike Nicolette	36—70
Payne Stewart	36—70
Perry Arthur	36—70
Ronnie Black	35—70
Lennie Clements	35—70
Steve Melnyk	35—70
Dan Halldorson	36—70
Rod Curl	35—70
Bill Rogers	35—71
Tommy Valentine	38—71
Gavin Levenson	36—71
Lou Graham	35—71
Tom Shaw	35—72
Charles Krenkel	34—72
Terry Anton	38—72
Pat Lindsey	36—72
Scott Watkins	36—72
Lindy Miller	38—72
Fred Couples	38—72
Bob Tway	34—72
Dane Quigley	37—72
Mark McCumber	36—72
Doug Tewell	35—72
Jim Booros	37—72
*Wayne Player	34—72
George Cadle	36—72
David Edwards	36—72
Andy Bean	36—72
Jeff Mitchell	38—73
Jack Ferenz	39—73
Rod Nuckolls	36—73
Bob Murphy	38—73
Vance Heafner	37—73
Rex Caldwell	39—73
Paul Linger	40—73
Richard Zokol	37—73
Thomas Gray	35—73
Bob Proben	37—73
Mike Sullivan	39—73
Mick Soll	34—73
Ken Green	37—73
John Schroeder	37—74
Leonard Thompson	39—74
Greg Powers	38—74
Ron Commans	37—74
Brian Lindley	36—74
Chuck Thorpe	37—74
Roger Calvin	36—74
Gene George	39—74
Brad Bryant	36—74
Jay Haas	36—74
Tom Jenkins	39—74
Chris Hall	39—74
Barney Thompson	40—74
Mike Donald	36—74
Jim Bertoncino	37—74
Frank Coner	37—74
Bill Britton	37—74
Jeff Thomsen	36—75
Scott Steger	35—75
Larry Ziegler	35—75
Dale Douglass	37—75
Larry Mize	39—75
Grier Jones	38—75
Forrest Fezler	38—75
Buzz Ijiw	39—75
Barry Jaeckel	39—76
Steven Liebler	39—76
Al Geiberger	36—76
Bruce Douglass	40—76
Ben Crenshaw	38—77
Jodie Mudd	38—77
Jim Dent	40—77
Mike Hill	37—77
Jim Simons	39—77
Joe Hager	35—77

College Results

BASEBALL

Seton Hall 17 ..... N.J. Tech 5

TOURNAMENTS

N.A.I.A. AREA 2

Texas Wesleyan 8 ..... William Carey 7

AREA 3

Mayville St 7 ..... Wayne St. 2
Southeastern St 7 ..... Emporia St. 4

AREA 5

Birmingham So. 5 ..... Trevecca Naz. 1

Soccer

North American Soccer League

THURSDAY NIGHT'S MATCH

Toronto 2, Chicago 0.

Eastern Division

	W	L	BP	Pts	Goals For	Agt.
Cosmos	8	2	22	68	23	13
Toronto	7	3	15	55	18	12
Montreal	3	4	8	26	9	8
Chicago	1	7	9	15	9	19

Southern Division

Ft. Lauderdl	7	3	20	60	22	17
Tampa Bay	4	6	12	36	13	17
Jacksonville	3	5	14	32	14	17
Tulsa	2	5	14	32	14	17

Western Division

San Diego	5	3	12	40	14	10
Edmonton	5	3	9	35	12	11
Vancouver	4	5	10	32	12	18
San Jose	4	4	8	30	10	11
Portland	4	3	9	29	13	8
Seattle	2	6	9	19	11	15

TODAY'S MATCHES

San Diego at Montreal (n.).
Tampa Bay at San Jose (n.).
Toronto at Chicago.
Tulsa at Jacksonville (n.).

Six points awarded for a regulation or overtime victory. Four points for a shootout victory. One bonus point for every goal scored with a maximum of three per game. No bonus point is awarded for overtime or shootout goals.

M.I.S.L.

FINAL

NEW YORK VS. ST. LOUIS

May 14 — St. Louis 3, New York 2 (overtime).
May 16 — New York 5, St. Louis 3.
Last night — at St. Louis.
May 23 — at St. Louis, 8:30 P.M.
*May 26 — at New York, 8 P.M.
(Series tied, 1-1)

*If necessary.
All times Eastern daylight.

A.S.L.

LAST NIGHT'S MATCH

Nashville at Georgia.

	W	L	T	GF	GA	BP	Pts
Georgia	2	0	1	7	4	7	19
Nashville	1	1	0	4	3	4	9
Detroit	1	1	0	3	2	3	8
Pennsylvania	1	0	0	3	2	3	8
Carolina	0	1	1	2	4	2	4
Oklahoma City	0	2	0	2	5	2	2
Rochester	0	0	0	0	0	0	0

(Last night's match not included.)

TONIGHT'S MATCH

Oklahoma City at Carolina.

Five points awarded for a victory, two for a tie, and one bonus point for each goal scored up to a maximum of three.

Lacrosse

N.C.A.A. MEN'S CHAMPIONSHIP

SEMIFINALS

Today

North Carolina at Cornell.
Virginia at Johns Hopkins.

Final

May 29 — At Charlottesville, Va.

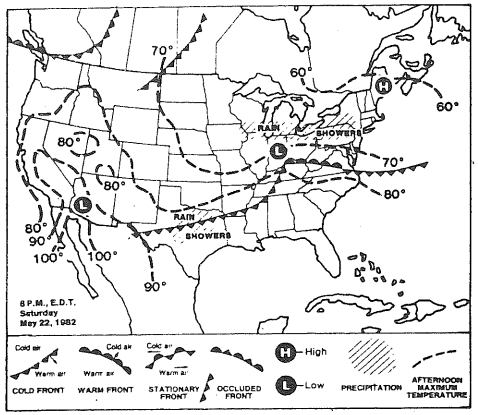
FIGURE 5

The Weather

Meteorology by the National Weather Service

National Forecast

Thunderstorms and showers will be scattered from the Middle Atlantic States and throughout the Tennessee, Ohio and middle Mississippi Valleys. Most of New England, the Gulf States, especially along the coast, and the West will have sunny skies and seasonable temperatures.



New York Area

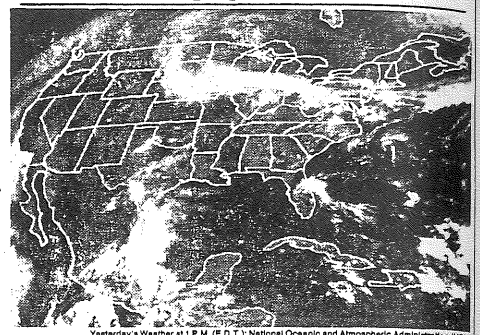
Cloudy skies, cool temperatures and occasional, scattered showers are likely in the New York metropolitan area today as a slow-moving frontal system passes across the region. Easterly winds will maintain cool temperatures for much of the weekend. Tonight and tomorrow, periods of rain are likely. Yesterday was partly sunny, with 72 the high for the day and 67 the low.

Travelers' Forecast

The following forecast for today shows the projected weather conditions and temperature range, from afternoon high to evening low.

Atlanta: Partly cloudy. 87-64.
Boston: Mostly sunny. 63-51.
Chicago: Showers likely. 62-48.
Cleveland: Thunderstorms. 71-54.
Dallas: Partly cloudy. 88-67.
Denver: Mostly sunny. 76-46
Houston: Partly cloudy. 88-71
Kansas City: Mostly cloudy. 69-52.
Las Vegas: Partly cloudy. 92-65.
Los Angeles: Partly cloudy. 74-60.
Miami: Partly cloudy. 84-75.
Milwaukee: Thunderstorms. 58-46.
Minneapolis: Mostly cloudy. 66-48.
New Orleans: Partly cloudy. 85-67.
New York City: Rain likely. 63-51.
Philadelphia: Showers likely. 64-57.
Phoenix: Mostly sunny. 97-69.
Pittsburgh: Thunderstorms. 69-56.
St. Louis: Partly cloudy. 78-54.
San Francisco: Mostly sunny. 66-49.
Seattle: Mostly fair. 71-51.
Toronto: Showers likely. 65-50.
Washington: Rain likely. 73-61.

Satellite Photograph



Yesterday's Weather at 1 P.M. (E.D.T.): National Oceanic and Atmospheric Administration/UPI

A band of frontal clouds with embedded rainshowers is seen reaching from the North Atlantic States across the lower lake region into the Plains States. Thunderstorms can be seen over Florida, while high, patchy cloudiness is over Texas. Much of the remainder of the country has clear skies.

Foreign and U.S. Cities

Abroad

Following are the temperatures and weather conditions in foreign cities yesterday at the local time indicated:

Aberdeen	Noon	55	Cloudy
Amsterdam	1 P.M.	67	Cloudy
Antwerp	2 P.M.	63	Pt.cldy.
Athens	2 P.M.	74	Clear
Auckland	Midnt.	56	Cloudy
Beirut	2 P.M.	72	Clear
Berlin	1 P.M.	64	Cloudy
Bonn	1 P.M.	66	Pt.cldy.
Brussels	1 P.M.	66	Cloudy
Buenos Aires	9 A.M.	41	Cloudy
Cairo	2 P.M.	†	
Casablanca	Noon	73	Pt.cldy.
Copenhagen	1 P.M.	60	Cloudy
Dakar	Noon	71	Cloudy
Dublin	Noon	54	Rain
Geneva	1 P.M.	66	Pt.cldy.
Helsinki	2 P.M.	53	Clear
Hong Kong	8 P.M.	78	Clear
Jerusalem	3 P.M.	73	Clear
Lima	7 A.M.	58	Clear
Lisbon	Noon	79	Cloudy
London	Noon	64	Cloudy
Madrid	1 P.M.	79	Clear
Managua	9 A.M.	75	Clear
Manila	8 P.M.	83	Pt.cldy.
Montreal	2 P.M.	57	Pt.cldy.
Moscow	3 P.M.	54	Cloudy
New Delhi	8 P.M.	92	Clear
Nice	1 P.M.	70	Clear
Oslo	1 P.M.	47	Rain
Paris	1 P.M.	63	Cloudy
Peking	8 P.M.	90	Tstms.
Pretoria	2 P.M.	78	Clear
Rio de Janeiro	9 A.M.	71	Pt.cldy.
Riyadh	3 P.M.	106	Clear
Rome	1 P.M.	72	Clear
Seoul	8 P.M.	67	Clear
Stockholm	1 P.M.	52	Cloudy
Taipei	8 P.M.	89	Clear
Tokyo	8 P.M.	69	Clear
Toronto	2 P.M.	61	Fair
Vienna	1 P.M.	70	Rain
Warsaw	1 P.M.	59	Pt.cldy.
Winnipeg	2 P.M.	66	Pt.cldy.

Following are the highest temperatures the day before yesterday, the lowest temperatures in the 12-hour period ended yesterday at 8 A.M. (E.D.T.) and the conditions yesterday at 8 A.M. (E.D.T.)

Acapulco	92	79	Cloudy
Barbados	86	79	Pt.cldy.
Bermuda	77	68	Pt.cldy.
Freeport	85	68	Pt.cldy.
Guadalajara	86	59	Cloudy
Havana		75	Pt.cldy.
Kingston	89	76	Pt.cldy.
Mérida	99		Pt.cldy.
Mexico City	73	55	Cloudy
Montego Bay	86	73	Cloudy
Monterrey	87	72	Cloudy
Nassau	84	70	Pt.cldy.
St. Thomas		†	

United States

Following are the temperatures and weather conditions through out the nation yesterday at 2 P.M. (E.D.T.):

Albany	†	
Anchorage	†	
Atlanta	†	
Birmingham	†	
Bismarck	59	Pt.cldy.
Boise	/1	Pt.cldy.
Boston	†	
Buffalo	†	
Casper	60	Fair
Charleston, S.C.	82	Pt.cldy.
Chattanooga	†	
Chicago	†	
Cincinnati	†	
Cleveland	†	
Denver	†	
Des Moines	51	Rain
Detroit	66	Cloudy
Fairbanks	†	
Fort Worth	†	
Great Falls	†	
Hatteras	†	
Honolulu	78	Fair
Indianapolis	71	Fog
Jacksonville	†	
Kansas City	68	Pt.cldy.
Key West	†	
Knoxville	†	
Las Vegas	81	Cloudy
Little Rock	†	
Los Angeles	61	Haze
Louisville	80	Tstms.
Macon	†	
Medford, Ore.	68	Fair
Memphis	†	
Miami	†	
Minnpls-St. Paul	63	Windy
Nashville	†	
New Orleans	†	
Oklahoma City	†	
Omaha	†	
Pendleton	†	
Philadelphia	†	
Phoenix	89	Fair
Pittsburgh	†	
Portland, Me.	†	
Portland, Ore.	62	Fair
Rapid City	67	Pt.cldy.
Richmond	†	
St. Louis	†	
Salt Lake City	67	Cloudy
San Antonio	†	
San Diego	†	
San Francisco	†	
San Juan	†	
Sault Ste. Marie	†	
Spokane	66	Cloudy
Tampa	†	
Washington	†	

†Temperatures and conditions unavailable from the National Weather Service.

© New York Times Company, 1982
Reprinted with permission

A CORE OF STATISTICAL TRAINING FOR THE LAYMAN

What are the basic statistical concepts required for an intelligent reading of The New York Times of 22 May 1982? Can we extract from this single newspaper issue, and from a composite of other ordinary information material, a core of statistical topics which would assist the layman to a deeper understanding of our world?

The first point that must be made is that while the terminology may vary in different areas such as finance, economics, social science, politics, and natural science, the statistical principles which are applied to numerical data in each of these fields are identical. They are, in effect, largely context-free. The material we have examined above requires some grounding in the following topics:

1. Descriptive statistics: numerical data and ranges of variation
2. Graphical statistics: graphical representation of data (including time series).
3. Economic statistics: economic indices and their rates of change.
4. Demographic and social statistics.
5. Polls, sample surveys and voting formulae.
6. Concepts of statistical inference.
7. Errors and fallacies in statistical arguments.

For the sake of completeness, one might add for good measure:

8. Statistical calculations: mean, variance and the concept of a distribution.
9. Elements of correlation and regression.
10. Probability models and how to fit them to statistical data.
11. Basic concepts of insurance statistics.

This will undoubtedly seem a very ambitious program, but it is one which is essential to a proper understanding of the information we encounter, and to the conduct of our daily life. If we intend to prepare our students to grapple with the problems of society, we have little choice but to include such a core of statistical training in their schooling and college education. Because of its dependence on numeracy, such a program must inevitably be carried out in conjunction with the students' development of arithmetic and computational skills.

How can this best be done in our schools and colleges? It would be tempting to insist that we set up formal courses in statistics to cover the above topics, as part of every school and university curriculum; we could then demand that every student attend them. Such an attempt would frighten many, particularly those with modest numerate skills, and might well prove counter-productive in the long run. Possibly a lower-keyed approach on multiple fronts would prove more successful. Mathematically oriented courses in statistics should certainly be provided for those who can benefit from them. But one must

simultaneously cater for students of modest numeracy by providing them with a more statistical orientation in business, biology, economics, geography, history, politics and social science courses, as well as all the other fields of study offered in our educational system. In doing so, one would discuss most of the 11 topics previously outlined.

Since statistics enters into the subject matter of most fields in which quantitative data arise, it is important that students should be able to understand its uses, and learn to handle simple statistical problems with some facility. Several statistical text books exist at a variety of levels, both school and college, which endeavour to achieve this aim, and which teachers have found useful as guides. But they are mostly published in the developed countries of the world; it will be one of our challenges to explore whether they can be adapted to the requirements of developing nations, or whether completely new texts will need to be written for these.

Text books may be important in our drive for statistical literacy but skilled teachers will prove crucial to its success. The shortage of such teachers is one of the difficulties experienced in the developed countries over the past two decades, when the role of statistics in education has begun to achieve some recognition. There is no surplus of school and college teachers competent to teach statistics, or to inject a statistical orientation into courses in other fields of study. Not all teachers of mathematics are fully familiar with basic statistical concepts; biologists, geographers and social scientists, among others, are often not sufficiently confident of their statistical skills to discuss statistics problems as they arise in their courses.

Hence, to succeed in educating our youth statistically, we must begin by educating our teachers. In this campaign, the production of suitable text books, the publication of such magazines as The American Statistician, and Teaching Statistics, the organization of Conferences such as ICOTS, and the recommendations of the ISI Committee on Statistical Education and UNESCO all have a valuable role to play. But the most important factor of all will be the teachers' recognition of the need for statistical literacy, and their personal willingness to retrain themselves to respond to the problem. Such retraining may require them to attend appropriate university statistics courses; they may need to convince their Education Departments that statistical material should be included in various student programs of study, and plan to modify syllabuses gradually so that statistics becomes an integral part of other subject courses. If the ISI Committee on Statistical Education can assist such progress in any way, please ask for its help.

We are envisaging an enterprise of the utmost importance: a core training program for world-wide statistical literacy. This is an endeavour linked with the spread

of numeracy, which could well take several generations to bear fruit. To carry it out, we must train our teachers first, so that they can in their turn train our youth at both the school and college levels. Such training will require detailed and careful organization, with different emphases in different countries. In bringing together teachers of statistics of all nations, ICOTS has made a contribution to this long-term education process.

CAN BASIC STATISTICAL TRAINING BE CULTURE-INVARIANT?

We have put forward the suggestion that, for the majority of students, it may be more appropriate to introduce statistical concepts and techniques into the teaching of other subjects, rather than set up a universal basic course in statistics. We must now face the question of whether such training can be culture-invariant. Do the citizens of developed countries require the same knowledge of statistics as those of developing nations? Will the basic topics of statistical training vary depending on whether we teach Nigerians or Japanese, Russians or Upper Voltans? Clearly there are many responses to this question; it can be answered partially by stressing the universality of statistical theory and methods. But another partial answer will point to the differing needs of individual countries, as well as to their educational and economic resources.

If we concentrate purely on the concepts and methods of statistics, as embodied in the 11 topics of Section 3, then we can reply very briefly that basic statistics is independent of culture. Means and variances, significance levels, and the fitting of regression lines are identical for all countries: statistical procedures have the same culture-free validity as the other methods of science. But to the extent that statistical data arise in an economic or social context, rather than a scientific one, then they may well vary from country to country. It is relevant here, to point out that neither the Sunday Times of Nigeria of 6 June 1982, nor the Overseas Hindustani Times of 31 December 1981, published as large a proportion of statistical material as The New York Times. In both of them, approximately 10% of the material involved elementary descriptive statistics. It would appear that in 1982, technically advanced countries were more statistically oriented than agricultural ones, even those with vigorous nascent industries.

Even in developed nations, examples and illustrations of statistical principles may differ substantially. While it might be sensible to discuss variation and trends in economic time series by reference to the stock market in the USA or Britain, such concepts would perhaps be more suitably discussed in terms of annual oil or steel production in the USSR, where the stock market does not exist. In developing nations, statistical discussion must center on topics of vital concern for the particular country. In Egypt, for example, one might select the statistics of fertility and mortality, data on the cotton crop, or the

growth of oil production, while in the Dominican Republic the sugar crop will clearly provide statistical data of direct interest. In short, the discussion and analysis of statistical data must reflect local conditions and interests, if statistics is to become an integral part of every citizen's education. Statistics text books and courses in different countries will show signs of such cultural differences; even if a standard English text is adapted for use in a developing country, its examples may require modification to make the contents more interesting to its new students.

A more fundamental point concerns the differences in economic and educational resources in different countries, and the consequences of these for a statistical training program. There is an order of human priorities which must be kept clear in our minds: it makes little sense to press for statistical sophistication in a country where unemployment is persistently high, where literacy is only partial, and where disease is endemic. Calcutta is the home of the famous Indian Statistical Institute; yet I would hardly rate the statistical education of Calcutta's citizens as a first priority. Like all other skills, statistical expertise subsumes health, an acceptable standard of living, and a reasonable level of literacy and numeracy. We must not, like Dr. Pangloss in Voltaire's "Candide" get carried away with the beauties of philosophy (or statistics) in the midst of famine, carnage and human injustice. In this sense, statistical training cannot be culture-invariant.

Most European countries are considered to be developed. They have well established school and university systems, with a trained body of teachers in mathematics and the natural and social sciences, capable of handling the basic statistical training of students which we have proposed, either immediately, or within a relatively short period. It is no accident that Teaching Statistics was published in Britain, or became immediately popular in Germany, Scandinavia, and other European countries. The ground there is ripe for statistical literacy.

The same is not true, for example, of India or South America. Here, with some notable exceptions, it would be difficult to introduce statistics in schools at this stage; it is only in some of the universities that progress in the spread of statistical education can be made. But if progress can be made in the universities in this generation, and some basic statistical training can become part of every college student's education, then there is hope that such education will filter downward in the next generation, as the standards of living and education rise. Every country must decide for itself how best statistical literacy can be increased, and which educational sector is most likely to be responsive to the growth of statistical education. And in every country, school and university teachers interested in statistics must try to convince their governments of the value of laymen's understanding of the subject.

This postulates that every government is keen to rule over an intelligent and well informed people. One may cynically observe that this is not always so, either in the democratic or in the totalitarian countries. In this sense, the drive for statistical literacy may be only one of the facets of the human struggle for social justice, and for the fulfilment of every living person's potential. These are sentiments nobly enshrined in the United Nations Universal Declaration of Human Rights, but more often honoured in the breach than in the observance.

To summarize our views, one might say that while basic statistics is culture-invariant, the case studies by which statistics will be illustrated may be culturally oriented. It is up to us, as concerned statisticians in every country of the world, to devise the best methods of imparting statistical literacy to our youth, to produce the best texts that we can with relevant local examples, and to incorporate in these the core of statistical principles which we have previously outlined. The ISI Education Committee is vitally interested in this campaign for statistical literacy, and will give the fullest possible assistance in its pursuit.

CONCLUDING REMARKS

Statistics is relevant: in a world where science and technology are rapidly assuming a dominant role, it is impossible to understand the daily newspaper without a smattering of statistical knowledge. To achieve this, we have suggested a basic core of statistical literacy, and a program of teachers' training, followed by the widest possible statistical education of students in schools and universities. We have tried to set this campaign for statistical literacy within the broader context of the struggle for human rights.

And now, we face a further revolution, a revolution not only in our technology, but also in our methods of supplying, storing and using information. While the electronic computer has been with us for over thirty years, it is only recently that mini-computers of great power have entered into our homes in the developed countries. The basic Sinclair computer, which hooks up to a TV set, costs only $150; others like the Apple III, or the TRS 80, or the Commodore PET cost at most a few thousand dollars, and are rapidly decreasing in price.

These mini-computers handle numerical and statistical data with great facility, calculate means and variances, plot histograms, detect trends in time series, carry out standard significance tests, and generally do all that a sophisticated statistician might require of them. The layman may not understand the finer details of the procedures which the computers are programmed to follow, but he can certainly understand and use the results. It is therefore important that the layman be made aware, not only that there are statistical fallacies, but also bad statistical programs. To statistical literacy, we may in future

have to add programming literacy.

These two may gradually become confounded, but for the moment let us meet one problem at a time: we should begin with a campaign for numerical and statistical literacy, to equal in this 20th century that for literacy itself which occurred during the 19th century. When we have made some headway with it, it will be time enough to turn to computer and programming literacy. Life constantly renews itself, and we can rely on it to provide us with new challenges tomorrow, as soon as we have resolved the problems which we face today. Let us concentrate in the 1980s on statistical literacy, and hope that within our own lifetimes we shall see the results of our efforts, and the consequent spread of such literacy throughout the world.

TOPIC 1: TEACHING STATISTICS IN SCHOOLS

(6-11 YEARS AGE GROUP)

ORGANISER: PROFESSOR K MATUSITA

STATISTICS IN THE CURRICULUM FOR EVERYBODY - HOW YOUNG CHILDREN AND HOW THEIR TEACHERS REACT

Tamás Varga, National Institute of Education, Budapest

STATISTICS in the title means more than simple descriptive statistics such as representing children born in a given month or families with a dog or a cat by piled up matchboxes and the like. But it means less than any serious inferential statistics, if only

because

CURRICULUM FOR EVERYBODY in the title is meant to be the one implement in the first four grades of every Hungarian school six to ten year olds, where statistics, in fact, appears at every grade level along with combinatorics and probability, as also in the subsequent fifth to eighth school year, where the implementation is still under way, and also

because

YOUNG CHILDREN are meant to be the said six to ten year olds though most classroom examples will be related to a particular class of third graders in Offenbach,* West Germany, nine year olds on average, with whom I decided to gather some fresh experiences while preparing this talk. As for the TEACHERS, I have in mind mainly those I regularly meet in Hungary during in-service training courses. But the bulk of what I am going to say is about children.

Back to the first word in the title, the reason for going beyond simple descriptive statistics as early as in the first few primary grades is that young children seem to be more responsive to essential, deep questions than to occasional, superficial ones. They are eager to study random phenomena, which they sense to be worthy of their attention. Questions of birth and life and death, accidents and disasters just as well as funny coincidences rouse their interest. So do random phenomena involving actual experiences in their world of plays and games. Their world also includes tales: stories and beings of their imagination, which can be modelled in their games. One more facet of their interest is that they like personal involvement. Pet graphs appeal to them if their own pets are included. In their experiences about random phenomena their involvement is greatly increased by first predicting the outcome then performing the experiment, finally comparing the result with their prediction. In some cases a further step is an attempt to explain the pattern of results by some mini-theory. This carries them further toward probability. During their games they are motivated to devise strategies which increase their chances, and then

* I owe big thanks to the classroom teacher, Mrs. Sommer, for allowing me to take over her class of 14 girls and 11 boys, coming from seven different countries, during eleven times 45 minutes and sometimes longer.

they get some feeling for game theory. The mini-theories demand combinatorial considerations. Yet since the order is

hypothesis ⟶ experience ⟶ theory,

I feel it appropriate to subsume such learning sequences under the term statistics.

The present talk is based on extensive experience, but no controlled experiment, no statistical data, no research like that done by Piaget, Estes or Fischbein.

It is just my private opinion that children usually behave as if they have read their books and knew their roles related to their ages. This is particularly so in the beginning of a sequence of lessons. If, however, the lesson inspire them - and I hope it is really inspiration rather than indoctrination - then they may forget their age and role, and behave at a more mature level. This should, of course, be tested, and possibly in longitudinal research, in order to prove how lasting the effects are.

Now I start to report on certain segments of actual lessons, intermittently adding my remarks.

TWO LANGUAGES FOR PREDICTING LEVELS OF CERTAINTY

Teacher: 'We are going to study probability. It is about things you cannot be sure of. They are in between what is certain and what is impossible. For instance tomorrow will certainly be a Wednesday, which I can show this way (shows a stretched palm with fingers upward). It is impossible that tomorrow it will be Monday (stretched hand, fingers downward). But you cannot be sure whether it will rain or not. You can at most be pretty sure, but not absolutely certain. (Stretched hand, fingers slantly upward). Or you may feel, it is rather unlikely to have rain tomorrow. (Slantly downward). Or maybe you feel so: it is just as likely that it will rain as that it will not. How would you show that? (Some children stretch their fingers forward, others wave them up and down). Discussion follows about the possible interpretations of these two gestures. 'If you wave your hand, then I will think that you change your mind in every moment'.

Examples come next, imaginary and actual events where children express their opinions or predictions by the hand language. Random devices are used: coloured beads to choose of, blue-red counters to throw, etc. Later - maybe in a few days - children learn a written (or rather drawn) language to express their predictions. They do it by stick pictures like these: (Fig.1),



FIG. 1

or these (Fig.2)



it is certain (bound to happen)	there is more than even chance	chances are even	there is less than even chance	it is impossible (cannot happen)

FIG. 2

Two important cases are near the two extremes: (Fig.3)

almost certain, but not quite  almost impossible, but not quite

FIG. 3

WHY THESE LANGUAGES?

Both the gestures and the stick pictures are vehicles of non-verbal communication. The gesture language enables instantaneous, simultaneous, and - last but not least - silent feedback in a class of children. The stick pictures are initially a rough representation of children's probabilistic intuitions. The sticks can be drawn of any height; it is the relative position of the marking that counts.

The use of square grid (graph paper) allows a more refined quantification, and paves the way to representing the distribution of random variables.

Both languages suggest the idea of probability as a quantity between two extremes, while avoiding the sophisticated idea of numbers between 0 and 1 - and at the beginning totally avoiding the use of numbers.

CHILDREN SIMULATE COINS

Every pupil receives a piece of graph paper with a rectangular frame around 8 times 16 square boxes. Their task is to figure out and record the behaviour of a randomly tossed coin. One way of doing this is to fill each of the eight rows of 16 squares from left to right one after the other by two different slant lines like these:



for Head for Tail

The choice of these two symbols results from an attempt to minimize both the time of recording and the difference between the effort to record Head and Tail. The results are found in Table 1 arranged in decreasing number of runs of length 1. (A boy was absent). For easy comparison, expectations in 128 actual tosses are given in the last columns.

TABLE 1

		1 B	2 B	3 B	4 B	5 B	6 B	7 G	8 G	9 B	10 G	11 G	12 B	E.
Runs of length	1	101	73	69	66	60	58	55	54	47	42	41	41	32
	2	12	18	7	6	28	10	35	23	34	9	28	17	16
	3	1	5	2	14	4	4	1	8	3	8	9	9	8
	>3	0	1	7	2	0	7	0	1	1	8	1	6	8
Runs		114	97	85	88	92	79	91	86	85	67	79	73	64
Tosses in runs of length	1	101	73	69	66	60	58	55	54	47	42	41	41	32
	2	24	36	14	12	56	20	70	46	68	18	56	34	32
	3	3	15	6	42	12	12	3	24	9	24	27	27	24
	>3	0	4	39	10	0	38	0	4	4	44	4	26	48
Tosses		128	128	128	128	128	128	128	128	128	128	128	128	128
Heads		64	51	63	67	65	59	61	62	62	72	68	65	64
Tails		64	77	65	61	63	69	67	66	66	56	60	63	64
Difference		0	26	2	6	2	10	6	4	4	16	8	2	≈9

		13 B	14 G	15 G	16 B	17 G	18 G	19 G	20 G	21 G	22 G	23 G	24 G	E.
Runs of length	1	40	39	39	37	34	34	31	29	27	22	17	11	32
	2	16	35	28	6	23	16	30	19	14	7	17	8	16
	3	9	5	11	2	10	15	5	6	13	9	14	13	8
	>3	7	1	0	12	4	4	5	10	8	13	8	14	8
Runs		72	80	78	57	71	69	71	64	62	51	56	46	64
Tosses in runs of length	1	40	39	39	37	34	34	31	29	27	22	17	11	32
	2	32	70	56	12	46	32	60	38	28	14	34	16	32
	3	27	15	33	6	30	45	15	18	39	27	42	39	24
	>3	29	4	0	73	18	17	22	43	34	65	35	62	48
Tosses		128	128	128	128	128	128	128	128	128	128	128	128	128
Heads		60	65	67	56	66	65	69	65	68	63	64	62	64
Tails		68	63	61	72	62	63	59	63	60	65	64	66	64
Difference		8	2	6	16	4	2	10	2	8	2	0	4	≈9

For those who have ever administrated a test like this to children - or, for that matter, to uninitiated adults - it will not be surprising that most of them believe shorter runs to be more frequent than can be expected in actual tosses. The different behaviour of boys (B) and girls (G) is in this small sample quite striking - girls seem to believe in more constancy on the part of a coin than boys, though most of them avoid very long runs, too.

COINS - FAIR OR UNFAIR?

The teacher prepared some loaded coins and dice, along with practically fair ones. The loaded coins, and some unloaded ones too, were rather thick, with real coins on one side, cardboard on the other, the whole of them wrapped, and marked by a 0 and a 1. The whole class assisted while a coin produced these outcomes:

00100 10100 00010 01110 11000 11101

So far there were 17 zeroes (Heads) and only 13 ones (Tails). Yet, in the last five tosses all but one was a Tail. The majority were Tails in another round of five. Is the coin loaded or fair? Or should we go on with the tosses, before taking position?

Children voted in favour of going ahead, doing 30 more tosses. Before continuing, the teacher asked them about the numbers of Tails in 60 tosses they felt would be still acceptable on the part of a fair coin. Would the coin be still rather suspicious if in 60 tosses there were only

20, 22, 24, 25, 26 tails?

Up to 25 all of them said they could not believe that the coin was fair. At 26 some of them said it could be. At this point the teacher tried to make the meaning of the vote more precise:

... 22 23 24 25	26 27 28 ...
In all these cases the coin would be maintained by them to be suspicious.	In such cases they could still accept it as being fair.

Then they went on with the numbers. At 27 the majority voted that from this onward numbers were already acceptable. Some votes were cast to 28 and 29. (There was no mention of numbers greater than 30). Most children presumably had just a very vague idea of what it was all about, but the teacher felt that these vague ideas moved them in the proper direction.

After 30 more tosses the end result was: 24 Tails. The coin was declared by all of them to be suspicious. Cut it! Cut it! - they yelled. The teacher cut the coin, and pulled out two big Hungarian coins from one side.

He was somewhat deceived - not with the children, but with the coin. He expected a more clear cut majority of Heads. At home he found a relative frequency very near to 2/3 in several hundred tosses. There he kept the conditions strictly constant by dropping the coin on the floor from a distance of about 120 centimeters. Was the floor in the school much different? Or did the fact that he now tossed it change the chances? Maybe it did. But even under precisely the same conditions the difference between 24 and 20, only slightly

beyond σ, was hardly enough reason for any great surprise, let alone for modifying his original assumption about the approximate probability. Most children expected still more conformity on the part of the coin in order to accept it to be fair - well within one σ without, of course, knowing anything of the binomial distribution and its S.D. At their level of knowledge they still behaved more intelligently, their probabilistic intuitions worked better, than that of their teacher. Sixty tosses are clearly too few to distinguish between two conflicting assumptions such as 1:2 or 1:1 odds - rejecting any of the two in favour of the other - at an acceptable level of significance.

Soon after this the falsity of a strongly loaded big white cardboard die was quickly detected: after a majority of sixes they cut it and found a lot of plasticine below the one-dotted face.

HOW DIFFICULT IS IT TO THROW A SIX?

The teacher started with the two Freudenthalian questions:

1. Which is easier to throw, a one or a six?
2. To whom is it easier to throw a six, you or me?

Most children were acquainted with a game where players could start moving their pieces only after they threw a six. The memory of their impatience may be one of the reasons why initially many of them believed it was more difficult to throw a six, than - say - a one. But the opinions were not unanimous. Nor were they about the second question. These children were not totally uninitiated.

I will try. You will count the number of throws until the six appears. Who of you believes that this will already happen the first time?

He writes the number 1 on the board, but no name goes below it. Nor below the 2. Some believe that the six will first appear at the third throw the teacher writes their names under the number 3. Most names get below the 4 and the 5, some below the 6, one below the "10 or more".

Let us see! Incidentally the first throw is a six. Tally mark below the 1!. Let us start again: one, two, three. The third is a six. Those who predicted are the champions, they get gold medals. Tally below the 3. So they go on together - some get silver, some bronze medals - then they continue the game in groups of two or three: one of them throws, another writes the tallies after the numbers (this time in a column, on the left side of their worksheets, from 1 to 20 or more) a third may just look and check if everything is OK. After ten minutes: dice back, outcomes are pooled. One in each group shows with his fingers, how many tallies are after the 1's; these numbers are added. Then after the 2's etc. Unbelievable! The 1 got the most tallies, then the 2.

There are good chances of this to be the case anyway - if a few hundred tallies are pooled - before oscillations prevail.

Through this activity children get a feeling of

a) the interplay between law and casualness,
b) a law - geometrical distribution, where mode, median and mean, here 1,4 and 6, are clearly distinct.

This will help them to grasp the meaning of each of these, now confusedly present in their intuitions. They cast their votes to numbers around the median (when does it start to be more likely to have thrown a six than not to have yet thrown it?) or the mean (theoretically: the expected number of throws until a six appears), without realizing, that their winning chances can be maximized by asking another question: the one about the most frequent (theoretically: the most likely) number of throws until the six appears.

FIRST DIGITS

Which is the most frequent letter in German?

Quite a few children voted for E, and the teacher showed on the top of a paper roll he started to unroll that it was in fact the E with 173 appearances in 1,000 letters. The top of the roll was like this:

E	173
N	104
R	81
I	75
S	64

Is the E the most frequent letter at the beginning of words as well?

Some said yes, some voted for D or other letters. Let us see! Open your reading books at random, read a passage chosen at random, do the tallying after the letters you have on your worksheets, alphabetically arranged.

After a few minutes the teacher asked them to dictate - the numbers appeared on the board - and then add the frequencies first of the E's, then of the D's as initial letters. The D's got a clear majority.

Maybe there is some other letter more frequent than the E as initial letter?

They tried the S and the T, also the W and a few others. Then the teacher changed the topic and asked:

What are to numbers as letters are to words?

A lot of them knew and shouted: digits! Which is the most frequent digit? The O, said a boy, and it was a reasonable answer. They discussed that O's very often appear at the end of numbers, though in prices 9's compete with them, or they even prevail.

Which is the most frequent first digit?

The digit 1 was pronounced. A girl started to list products, mainly foods, the unit price of which was between DM 1 and DM 2. Others were not convinced, and started to list further foods, with prices higher than DM 2.

This was the end of the lesson, and the teacher had only time to let children choose from a thick bundle of newspaper clippings mainly advertisements of foods, furniture, trips, sport articles, also radio and TV programmes etc.. He invited them to examine first digits. Some children undertook a different task: to open a telephone directory and do the tallying about the first digits of the house numbers. It was made clear that they were expected to prepare tally lists of this format:

1	
2	1
3	11
.	
.	
.	
9	

Next day most of them came to school with the finding that 1 - and possibly 2 - were more frequent than any other first digit, in quite different situations.

During the lesson some more situations were added. One of them was within mathematics: the multiplication table from 1 x 1 to 10 x 10, previously written on the board. Others were taken from life, e.g. the dates of birth of children, and possible dates of birth in general. Let us forget of the years; their first digits are always 1 (for the time being, and not much back in time). In four of the twelve months the first digit is 1 - the other eight digits have an equal share. In the possible days of the month 1 and 2 have an equal share with eleven possibilities each, 3 has only three possibilities of being a first digit, the other digits have only one.

A similar argument explained the case of TV schedules.

How about house numbers? Is there an upper limit? This was a more complicated situation, it needed some idealisation. Imagine a town where all the streets are of different length. In one of them there is only one house, its number is 1. Call it street one. In another street, call it street two, there are two houses, number 1 and number 2. And so on. Are there in any of these streets an equal number of houses with each of the nine possible first digits?

Yes, in street 9 each of the digits 1 to 9 appears as a "first" digit exactly once. It was harder to detect that the same was true about street 99. It was not important to go on to 999 etc.

Is there any street, where the champion on the first place is the 9? Or the 8? Or, for that matter, the 2?

They realized that none of these could happen. Even the 2 could only climb up to sharing the first prize with the 1, not to be superior.

The argument was not pursued further, it remained at this qualitative level. The invariance under a transformation - such as the areas of countries in square kilometres and in square miles - was kept for a higher level.

FROM PERCEPTION TESTING TO HYPOTHESIS TESTING

Could any of you distinguish these sorts of chewing gum if blindfolded?

Some said they could. One of them chose two gums, one wrapped blue, the other green. She was granted a short learning period, half a minute or so. This the teacher felt necessary, to avoid much change in the probability of hitting upon the correct answer in the course of trials. (Immediate feedback information about the hits had to be given to the subjects, if only because of the others who cheered at hits and sighed at failures). She succeeded in five trials in a row, and this the others accepted as a perfect proof of her ability.

A second volunteer, a boy, succeeded in three trials out of five, but failed in two. Some children wanted to be nice and tended to accept his performance saying that he still more often succeeded than not. The teacher warned them: How can you tell that the 3 to 2 result was more than a little luck? But he understood that tests administrated this way would not be highly motivating. He postponed this kind of activity until a later session. He prepared various test materials where the perceptive difference could gradually be decreased. This would give subjects an initial experience of success in cases where the differences were considerable, and a feedback about their perceptive ability levels which, of course, could not be unlimited. The following types of tests were devised, not all of them used so far:

Length estimation:

Rectangles with slightly different sizes were displayed. The subjects had to tell whether or not the vertical sides were the longer ones (whether or not the rectangles were turned by a right angle while the subject could not see them).

Number estimation:

Cuisenaire rods of the same size (and hence of colour, too: the red ones were used) on two trays, were shown a few more on one of the trays. Subjects had to tell whether it was the one to the right or the left. They had only one second to inspect the trays, and the number of rods was at least 15,

to avoid counting. The rods were not arranged in any regular pattern, and their positions were also changed, not only their numbers.

Weight estimation:
Rods were used again, but in closed boxes, and children had to distinguish which was heavier by lifting them.

Time pitch and taste estimations were also planned; the latter with a small quantity of salt - or of sugar - in one cup, to be distinguished from the other, just water.

From the point of view of teaching statistics, the reactions of children are more interesting than technical details. Batches of five trials were usually administered. Children felt that the following qualification is fair:

No failure in five trials : excellent
One failure in five trials : acceptable

How many in ten trials do you think would be acceptabe? the teacher asked.
The answer was again almost immediate: four.
The teacher tried to make clear that "acceptable" meant hardly by pure chance, and asked them, if they felt that two or less failures in ten trials could be the result of guesswork just as easily as one or none in five. They did.
He decided to postpone confronting them with contrary experimental evidence at a later date.
This was done by letting them shake boxes containing five, ten, or twenty two-coloured counters. (They were very small but relatively thick, to avoid overlapping). Without any calculations these experiments convinced children that their reasoning through proportionality - which they never learned but intuitively felt to be correct - was hardly applicable in situations like these.

Some tentative conclusions:

1. Statistics as a school subject is quite suitable in the primary school.
2. Children need not know much of arithmetic before they are initiated into statistical ideas.
3. Proportionality is not an indispensable prerequisite to understanding statistical - or probabilistic - ideas. It can even be a hindrance, and not only after being conditioned to look at proportionality as a key to most mathematical questions.
4. Learning statistics enhances children's arithmetic understanding and skills.
5. Learning statistics enhances, in particular, their ability of applying arithmetic to life situations.
6. Learning statistics enhances a number of other abilities related to mathematics and its applications, not just arithmetic ones.

TEACHING OF STATISTICS IN ELEMENTARY SCHOOLS IN JAPAN

T Yanagawa, Kyushu University

INTRODUCTION

Mathematics, science and social studies are the three major subjects which have the responsibility for teaching statistics in elementary schools in Japan. The basic and systematic teaching of statistics occurs in the mathematics classes. Students are exposed to various kinds of statistical charts, graphs and tables in science and social studies. They are supposed to study there how to use charts, graphs and tables for communicating information clearly and effectively. although these three subjects are taught by one teacher, textbooks are edited separately and independently. The statistics taught in these subjects and their interrelation are mainly results of a history of trial and error. The present system has promoted the teaching program of traditional topics such as charts, tables and graphs. But, at the same time, it has suppressed the introduction of new ideas into the program; in fact, statistics taught in elementary schools in Japan in the past three decades has changed only a little.

Japan has changed drastically during these three decades. Today we live in a highly developed technological society. Students are exposed to a great deal of numerical information in their daily lives. The importance of probabilistic and statistical ideas has thus increased considerably. It seems unreasonable that the essence of statistical education available in elementary schools today is almost the same as that of three decades ago.

The influence of Japanese statisticians on the teaching of statistics has been insignificant; we still have no research and teaching institute such as a Department of Statistics in Japanese Universities. The following are examples of our efforts and indications of the improvement in the environment and the quality of statistical teaching in Japan:

(i) The Japan Statistical Society established a committee of statistical education about 10 years ago and has delivered a lecture/consultation every year to school teachers, at the site of the annual meeting, hoping to improve the environment and the quality of statistical education.
(ii) For the first time in our history, the name of a statistician can be found among the 5-10 authors of each of the latest high school mathematics textbooks. We hope this will influence the teaching of statistics in elementary schools.
(iii) Teachers colleges have begun to employ statistical scholars in recent years. It is estimated that about 17% of teachers colleges have staff members who have published at least one theoretical or applied paper in statistics.

- 81 -

The purpose of the present paper is to describe how, what and when statistics is taught in elementary schools in Japan and to discuss its advantages and disadvantages with several suggestions for its improvement. In section 2 we will show the objectives and plans for the teaching of statistics in elementary schools established in the guideline of mathematics issued by the Ministry of Education. The guideline is detailed and all textbooks must pass its examination before they are published. In section 3, looking into mathematics textbooks, we will show that textbooks of mathematics are specialized for the systematic and efficient teaching of mathematics. Traditional topics in statistics such as charts, graphs and tables, and their prototypes, are efficiently employed for introducing and developing important ideas in mathematics. We will show in the section how beautifully such topics may be taught in mathematics. It is important to give students experience in collecting and analyzing data. However, the collection of data is neither encouraged in the guideline nor in the textbooks. We need to establish the program regarding the collection of data, coordinating with all subjects taught in elementary schools. Further the other important topics in statistics which cannot be incorporated in the systematic teaching of mathematics such as the role of the mean for central tendency, the dispersion, the idea of sampling and that of probability are poorly handled in most mathematics textbooks. We will look into these problems in section 4 with several suggestions for its improvement.

THE GOALS OF THE GUIDELINE

The objectives and plans for teaching statistics in elementary schools in Japan are established in the guideline of mathematics, issued by the Ministry of Education, as follows:

Objectives

1. To identify the aim of the problem and to organize the data accordingly.
2. To classify, arrange and present data in tabular and graphical forms, and to compute means and ratios.
3. To read and interpret tendencies of data presented in tabular and graphical forms, and summarized in means and ratios.
4. To compare tendencies of the data with subsets of the data, and to explore the idea of inference.
5. To introduce the idea of probability as a measurement of the degree of uncertainty in stochastic events.

Plans in each grade

1st grade :

a. Identify the set having the same numbers of members as a given set.
b. Match a set with its number properly.
c. Classify figures by shape, colour, etc.

2nd grade :

a. Count the total number of figures, classifying them by shape, colour, etc.
b. Read a one dimensional table. Record data in a simple one dimensional table.
c. Read picture graphs showing comparisons and recognize their usefulness.
d. Construct simple picture graphs.

3rd grade :

a. Classify data by date, location etc. Display it in a table. Check missing and overlapping data by computing the total of the table.
b. Read and interpret a simple bar graph.
c. Display data presented in tabular form in a bar graph.

4th grade :

a. Classify and arrange data in a two dimensional table. Check missing and overlapping data by computing totals of the table.
b. Read and interpret a bar graph.
c. Round data to nearest tenth... the nearest hundredth, thousandth, and display it in a bar graph.
d. Read and interpret a broken-line graph (line graph). Display data in a broken-line graph.

5th grade :

a. Construct a table of proportions from a table of real numbers.
b. Display data in a band graph. Display data in a circle graph (pie chart). Read and interpret band and circle graphs.
c. Show what a mean means and use it correctly.

6th grade :

a. Compare histograms and introduce a measure of dispersion (range). Recognize that the mean and range summarize the tendencies of data.

b. Select and construct the most appropriate graph for a given data set among those studied in 6 years in elementary schools.

c. Explore the relationship of the mean and ratio of a sample with those of a population, and explore, informally, the idea of sampling with an emphasis on the representativeness of the samples.

d. Explore empirical probability by observing the outcomes of a simple experiment.

TABLE 1 : School Hours*

	Grade					
	1	2	3	4	5	6
Mathematics	136	175	175	175	175	175
Statistics**	3	4	7	12	6	17
% (math/stat)	2.2	2.3	4.0	6.9	3.4	9.7

* 1 school hour = 45 minutes, ** estimated

Table 1 lists school hours for mathematics and estimated school hours for statistics taught in mathematics. Table 2 summarizes the above plans. To characterize the plan we compared it to that of the U.S.A. Note that, in the U.S.A., each state is responsible for the education of its students; that there is no similar plan to regulate the teaching in all schools in the state as in Japan; and that a very flexible plan is generally established independently by each of the local school units, called school systems. In the table the above plan is compared with that established by the Chapel Hill-Carrboro School Systems (CCSS), North Carolina, which is a unit consisting of 6 elementary schools. Mathematics In Our World is the textbook which is used. The table shows that statistical topics taught in Japan are wide and extensive and that they are allocated to grades systematically, whereas the CCSS plan limits topics and teaches them repeatedly.

TEACHING OF STATISTICS IN MATHEMATICS TEXTBOOKS

A large portion of statistics taught in elementary schools in Japan consists of tables and graphs. They are formally introduced in the grades shown in the previous section. However, the prototypes of tables and graphs are abundantly and effectively utilized for introducing and developing mathematical ideas in earlier grades. This training, coupled with the recent development of colour printing technologies, facilitates students' quick and easy understanding of mathematics. Note that students are also exposed to tables and graphs in

TABLE 2 : Comparison of teaching plans of Japan and Chapel Hill-Carrboro School Systems (CCSS), North Carolina, USA

		Grade					
		1	2	3	4	5	6
1 dimensional table	Japan		0	0			0
	CCSS						
2 dimensional	Japan				0	0	0
	CCSS						
picture graph	Japan		0				
	CCSS	0	0	0	0		
bar graph	Japan			0	0		
	CCSS		0	0	0	0	0
broken-line graph	Japan				0		0
	CCSS			0		0	0
circle graph	Japan					0	0
	CCSS			0			
band graph	Japan					0	0
	CCSS						
histogram	Japan						0
	CCSS						
mean	Japan					0	
	CCSS					0	
dispersion	Japan						0
	CCSS						
sampling	Japan						0
	CCSS						
probability	Japan						0
	CCSS						

Note: CCSS textbooks deal with picture graphs in 5th and 6th grades, too. The case of single symbol representing a many to one ratio is discussed in upper grades. It also deals with mathematical probability of equally likely events in 6th grade. Textbooks in Japan deal with empirical probabilities which covers also unequally likely events.

---

science and social studies prior to being formally introduced in mathematics. Figure 1 displays a part of the results by Miyagawa (1981). She carried out a test in two elementary schools in Kusatsu City (the total number of students in each grade is about 200). The figure compares percentages of students correctly answering questions regarding tables and graphs in the one grade

under and the grade where these topics were formally introduced. Although the correct answer rates depend on questions asked, the figure shows that students have a pretty good knowledge of tables and graphs before they are formally introduced in textbooks in mathematics. The finding gives rise to questions on the utility of the formal introduction (further discussion on the issue is given in the next section).

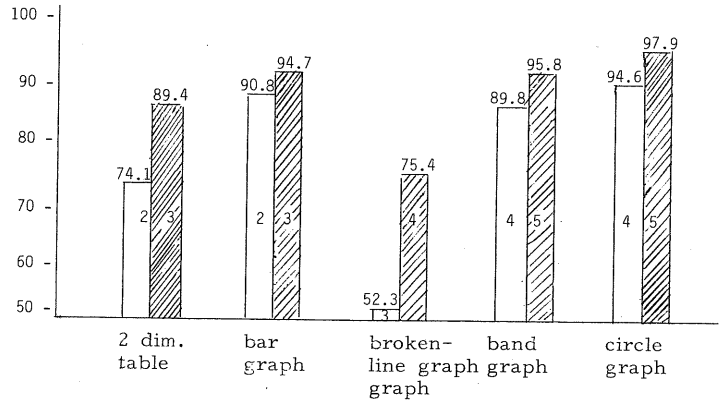
Figure 1

Percentages of students correctly answering questions regarding 2 dimensional table, bar graph, broken-line graph, band graph and circle graph.

▨ : the grade in which the table and graphs are formally introduced.

□ : the grade immediately preceding the above.

Data source: Miyagawa, N. (1981)



A summary of examples showing how tables and graphs, or their prototypes, may be utilized effectively for introducing and developing important ideas in mathematics is shown below.

TEACHING OF STATISTICS IN ELEMENTARY SCHOOLS

We have seen in the previous section how effectively tables and graphs, and their prototypes, can be used for introducing and developing important ideas in mathematics. Similarly tables and graphs are effectively used for communicating information clearly and quickly in science and social studies. We, Japanese statisticians, encourage these teachings. For, so far as reading and interpretation are concerned, tables and graphs are easy and straightforward and, as was indicated in Figure 1, students today have pretty good ability to understand them. What is important and necessary for teaching of statistics is to give students experience in collecting data, summarizing it in tables, displaying

OBJECTIVES (MATHEMATICS)	TABLES, GRAPHS AND THEIR PROTOTYPES
	GRADE: 1
1. To identify the number of items in a group and to match the number names with the number they present. 2. To compare two objects and be able to state which is larger or shorter; greater than, equal to, or less than etc. 3. To classify figures by means of shape or colour.	To use picture graphs. ... number lines. ... coloured tapes. ... picture graph. ... 1 dimensional table.
	GRADE: 2
1. Understanding of addition, subtraction and how they are related. 2. Introduction of multiplication; showing how addition and multiplication are related. 3. To count the total number of figures by clasifying them by shape or colour.	To use number lines and coloured tapes. ... picture graphs. ... 1 dimensional tables.
	GRADE: 3
1. Introduction of the fractional parts (1/2, 1/3, 1/4). 2. Introduction of addition and subtraction of fractional parts. 3. Introduction of decimals: addition and subtraction of decimals.	To use circle and band graphs. ... number lines. ... coloured tapes. ... number lines. ... segments (coloured tapes).
	GRADE: 4
1. Introduction of round numbers: addition and subtraction of round numbers. 2. Understanding the relationship of two dependent variables.	To use statistical data (populations of cities), tables and bar graphs. ... broken-line graphs.
	GRADE: 5
1. Introduction and application of ratio and per cent.	To use band graphs and circle graphs.
	GRADE: 6
1. Introduction of direct proportion and inverse proportion.	To use broken-line graphs

it in graphs and interpreting its tendencies. The first objective of the guideline states "to identify the aim of the problem and to organize the data accordingly". Note that it says "organize" and not "collect". The collection of data is neither encouraged in the guideline nor in the textbooks, since it consumes school hours. However "to have a great reserve" is a recent slogan of education in Japan. We think that it is a good time to advocate incorporating problem solving in textbooks and allocating ample time for its study. In the problem solving we expect students to organize the various topics they have studied and to deepen their knowledge by solving problems taken from their daily lives. We should not overlook the importance of demonstrating to students, especially those in elementary schools, the usefulness of what they have studied at schools in their lives. Problem solving is frequently incorporated in the textbooks of science and social studies. We suggest that teachers should encourage students to apply statistical techniques to find facts in these subjects. Problem solving is very rare in mathematics. We suggest to establish it in mathematics, too. This is not only important for the teaching of statistics but also for the teaching of mathematics.

The other important topics in statistics which cannot be incorporated in the systematic teaching of mathematics are poorly handled in the guideline as well as in mathematics textbooks. In fact, the role of the mean for central tendency, dispersion, the idea of sampling and that of probability are introduced at the last part of the 6th grade, after almost all of the required elementary school mathematics have already been taught. We think it is unreasonable. The importance of these topics in our society has greatly increased. Students are exposed to these concepts in their daily lives and seem to have already the potential to understand them in earlier grades. In order to improve the teaching of these topics, we advocate the establishment of a section concerning mathematics in games in mathematics textbooks. Here, we may introduce the idea of random variation, the role of the mean for central tendency, the idea of dispersion and the idea of probability by discussing students' favourite games. In particular, our idea is to introduce them in the 4th grade, develop them in the 5th grade and then introduce the idea of sampling in the 6th grade.

The establishment of the topics of problem solving and mathematics in games must be, of course, well coordinated with all subjects taught in elementary schools. Further, consideration is necessary to reduce the chances of excessively increasing the students workload by incorporating these programs. For example, regarding statistics, we could reduce the workload of students a little by minimizing the formal introduction of graphs (for example, the formal introduction of band graphs could be eliminated) which are unnecessarily precise and uniform from the point of view of teaching statistics.

While the above suggestions are moderate in theory, it is still considered too radical to carry them out in practice in the present educational climate in Japan. Therefore, as the "next best" policy, we are encouraging teachers to give students experience in collecting and analyzing data in the following way:

(i) The school hours in Table 3 are reserved for special activities in elementary schools in Japan, and they can be used partly for solving problems arising from classroom activities/school lives of students. We are suggesting that teachers encourage students to look into these problems by applying statistical techniques they have studied.

(ii) As homework is regularly given to students during long (summer/winter) school holidays in elementary schools in Japan, we are encouraging teachers to give students problems in collecting data such as recording daily temperature etc., displaying the data in tables and graphs, and interpreting the data's tendencies.

The suggestions depend totally upon the teachers' enthusiasm and their knowledge of statistics. Unlike mathematics, the answer to a question in statistics is not necessarily unique and fixed. This requires that teachers have sufficient abilities in teaching statistics. However, the problem is that not many teachers in Japanese elementary schools have such abilities. As teachers have several workshops on teaching each year, we have strongly encouraged them to establish workshops on teaching statistics. So far as teaching statistics in high schools is concerned, quite a few workshops have been successfully established and progress has been made with the consultation of statistical scholars. However, almost no statistical teaching workshops have been established so far in elementary schools. We find that obtaining the cooperation of teachers in elementary schools is a very difficult task. We hope that the recent developments in our environment shown in the introduction can change it.

TABLE 3 : School Hours for Special Activity

Grade	1	2	3	4	5	6
School hours*	34	35	35	70	70	70

* 1 school hour = 45 minutes

Acknowledgements

The author is indebted to Professor Shiro Yamazoe for his critical discussions during the preparation of this monograph. He also expresses his deepest appreciation of Mr. Carl Yoshizawa for his generous help in refining the English of this manuscript.

References

[1] Chapel Hill-Carrboro School System (1981) Mathematics Objectives.

[2] Ministry of Education (1977) Summary of Guidelines for Elementary Schools, (in Japanese).

[3] Ministry of Education (1978) The Guideline for Elementary Schools; Mathematics, (in Japanese).

[4] Miyagawa, N. (1981) Teaching of tables and graphs in elementary schools and its consideration, B.A. dissertation, School of Education, Shiga University (in Japanese).

[5] Mathematics In Our World, (1981) Addison Wesley.

TWO EXAMPLES OF STATISTICS ACTIVITIES FOR ELEMENTARY SCHOOL CLASSES

Claire Heidema, CEMREL-CSMP, St.Louis

SHUNDA'S NEWSSTAND

The business world offers many opportunities for statisticians, especially in advising decision-making processes. Elementary school age youngsters in their everyday lives also are confronted with decision situations, some with a business flavor, where a rudimentary understanding of statistics may prove useful. However, the pedagogical concerns of making the study of these situations accessible at an early age often prohibits consideration until the problems can be discussed on a high mathematical level. Consider, for example, the classic "Newsboy's Problem" concerning a newspaper seller attempting to maximize profit. Children can appreciate such a problem; indeed they may have paper routes or operate newspaper stands themselves. SHUNDA'S NEWSSTAND provides an excellent example of how an operations research problem involving statistics in a decision-making context can be presented at the elementary school level.

Crucial to the presentation of the problem in terms that 9-11 year olds can understand is the use of stimulating pictures to view the sample data in a variety of ways. The pictures and graphs provide an alternative to technical numerical methods; an alternative that is both aesthetically and pedagogically appealing.

Shunda's Problem

A common problem in the business world is to determine the optimal quantity of items to buy or produce in order to make the most profit. Many factors affect such decisions and no one can be guaranteed of always making the best choice. Still, educated decisions can be made by doing a careful study of sample data and by assuming that past behavior of the consuming public is a good predictor of future behavior. There is always a risk; all one can do is decide what is the most reasonable prediction. SHUNDA'S NEWSSTAND is concerned with such a business world problem as it affects Shunda, a young newspaper seller.

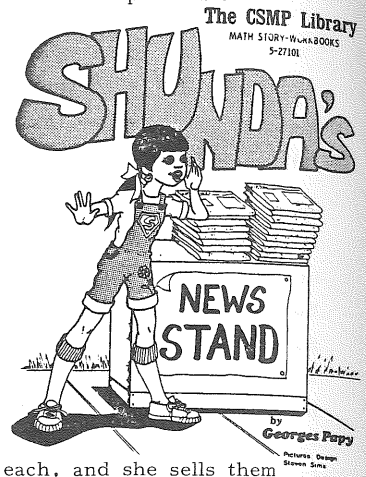
We present the situation in a story-workbook*, that is, as a story told in the pages of a "comic book". Along the way the students respond to questions and solve intermediate problems. They become familiar with a variety of pictorial representations of information and supply the necessary results needed to understand the main problem. Prompted by a series of stimulating pictures for recording

* SHUNDA'S NEWSSTAND is in the CSMP Library, a collection of math story-workbooks providing fanciful excursions into the colorful world of mathematics.

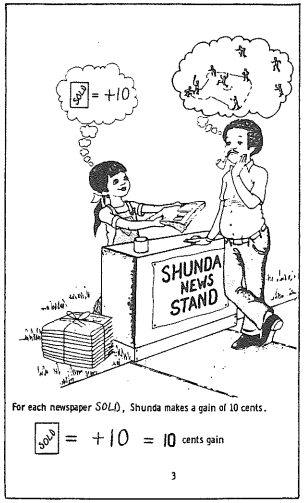
data, the students use the daily demand to determine the most profitable inventory of newspapers.

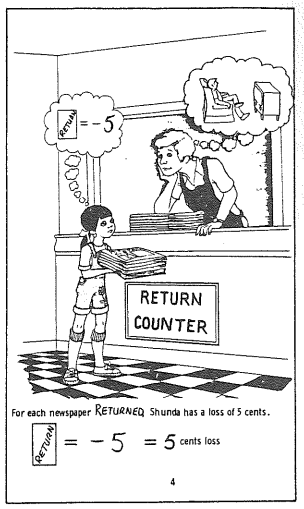
Our story is about Shunda, a newspaper seller, who has a newsstand on the corner of Hamilton Street and Euler Avenue. Each day between 4:00 and 6:00 PM, she sells newspapers to people passing by her stand.

Several pages at the beginning of the booklet* present basic information about Shunda's job.



Shunda buys newspapers from a dealer for 10 cents each, and she sells them to her customers for 20 cents each. Since the number of Shunda's customers varies from day to day, the dealer agrees to buy back the unsold papers for 5 cents each.

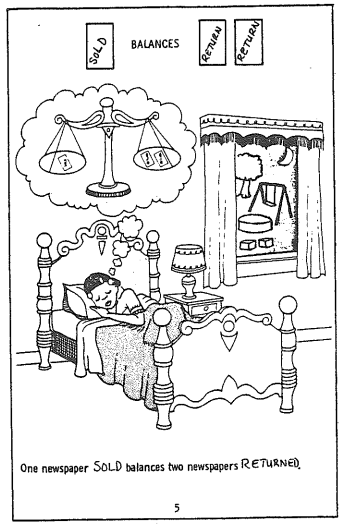


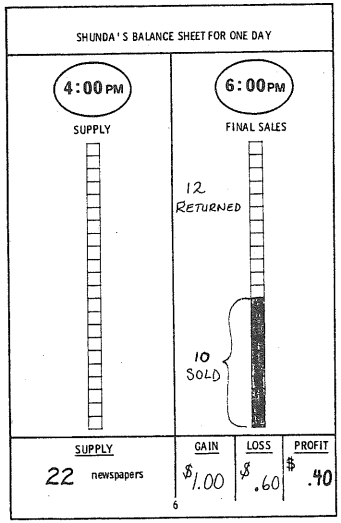


Shunda wants to have a successful business and to make as much money as possible. She devises procedures for keeping records of her business from day to day. The students work through several pages of practice using Shunda's record-keeping procedures. Selected pages from the booklet will suggest the variety of pictures and the nature of the individual work.

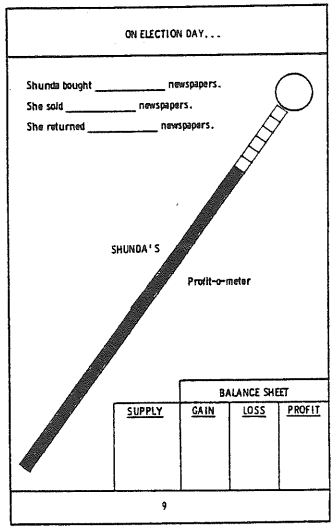
* In this paper, we use actual reproductions (in reduced form) from the booklet, omitting color. Light and dark shading will replace essential distinction of red and blue.

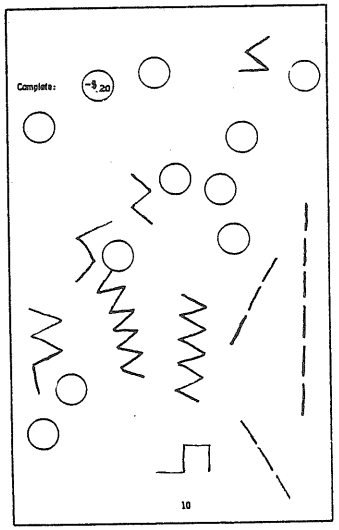
Shunda begins her newspaper business as an apprentice, so may buy differing numbers of papers from day to day depending on what she expects the demand to be.



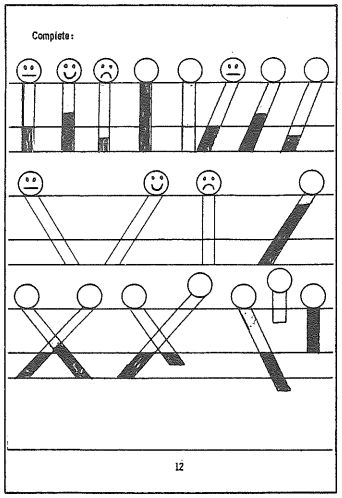


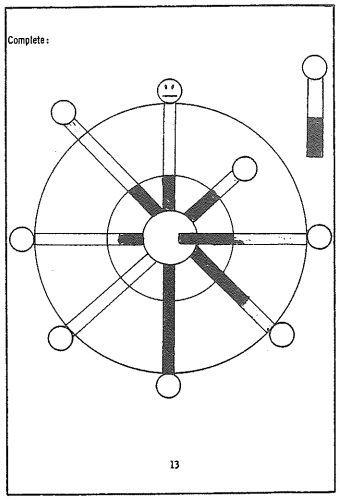
Stylization on page 10 reflects Shunda's whimsical side. Such artistic freedom requires little or no explanation, and it contributes to the students' experience viewing data in a variety of ways.





On days when Shunda has a positive profit, she is happy; when she has a negative profit, she is unhappy; and when she breaks even, she shrugs her shoulders and hopes for a better next day.

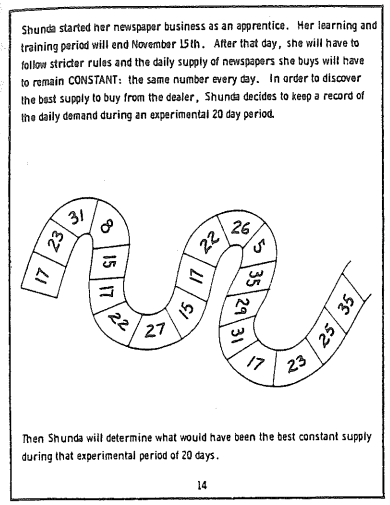


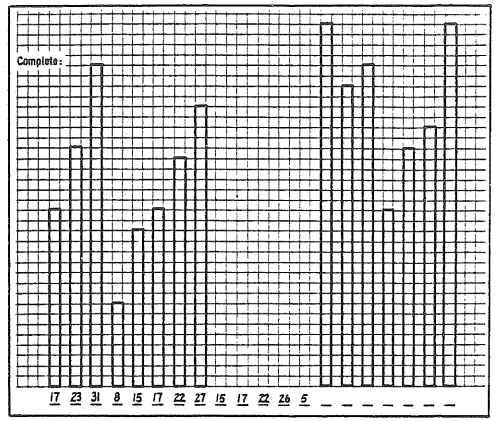


Students must use only the light:dark ratio to determine Shunda's mood or sales on several days. This step suggests the critical role in Shunda's business of a one-third, two-thirds ratio of returns to sales. Later we will see how these pictures contribute to a way of determining a most profitable inventory.

The story continues with new information that presents Shunda's main problem. The training period for Shunda will end shortly. Then she will have to follow stricter rules; the daily supply of newspapers will have to remain constant; the same number every day. Shunda must decide what constant daily supply would be the best to buy from the dealer. How can she make such a decision? In order to determine the best daily supply of newspapers to buy from the dealer, Shunda keeps a record of the daily demand during an experimental twenty-day period.

Students draw a graph of Shunda's sales record.





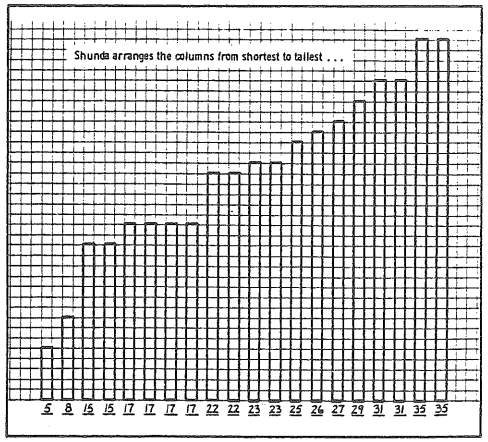
Shunda hopes that if she can determine what would be the best constant supply during that experimental period, then she can expect it also to be the best constant supply in the future.

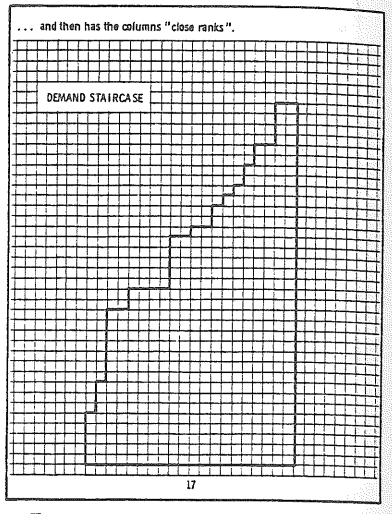
Before proceeding towards a solution, students stop to express their opinions on Shunda's decision. Such a problem interests 10 to 12 year olds; they may be just beginning to earn their own spending money. This discussion can bring out the students' natural curiosity about business matters; it tests their intuition about implications in the statistics for a twenty-day experimental period. Our experience suggests that students will think of a variety of ways of viewing the data; for example,

* Shunda should get only as many papers as she is sure to sell; then she is sure to have a positive profit every day.
* Shunda should get about 17 newspapers every day because she most often sold that number of papers.
* Shunda should get about 22 newspapers every day because half the time she sold more than 22 newspapers and half the time she sold less than 22 newspapers.
* Shunda should get about 20 newspapers because that is halfway between 35 (the most she sold) and 5 (the least she sold).
* Shunda should get 22 newspapers every day because that is the average number she sold daily during the experimental period.

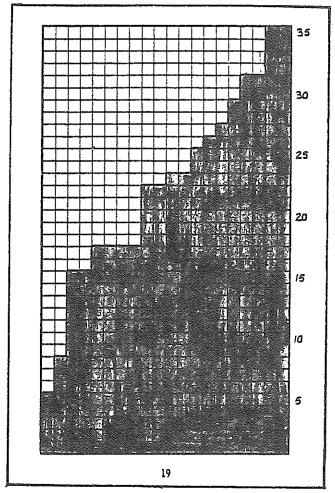
Shunda's Solution

We continue our story with Shunda organizing and studying the demand process as reflected in her records for the experimental twenty-day period.





The largest demand during the twenty-day experimental period was 35. Suppose Shunda would decide on a constant daily supply of 35 newspapers. Page 19 of the booklet shows a record for the twenty-day experimental period with the constant supply of 35 newspapers. We can view this picture as putting Shunda's profit-o-meters for the twenty days side by side and as highlighting the demand staircase.



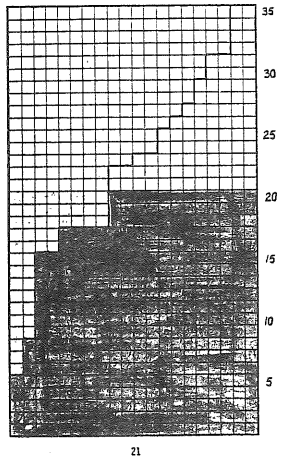
Students complete a balance sheet . . .

BALANCE SHEET for 20 day experimental period			
DAILY SUPPLY	TOTAL GAIN	TOTAL LOSS	TOTAL PROFIT
35	$44.00	$13.00	$31.00

and then question the choice of 35 as a constant daily supply. With such a supply, Shunda would satisfy her customers every day, but is it her most profitable choice? It seems that there is considerable loss (light) in the previous picture.

With the above picture and previous experience of finding daily gain, loss, and profit, students explore "what if" questions in preparation for determining a most profitable constant daily supply.

Suppose Shunda would decide on a constant daily supply of 20 newspapers. Page 21 of the booklet shows a record for the twenty-day experimental period with the constant daily supply of 20 newspapers. Would this choice produce more or less profit than a constant daily supply of 35 newspapers? There is much less loss (light) but also less gain (dark).

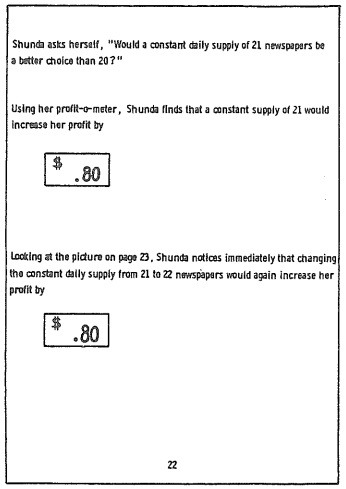


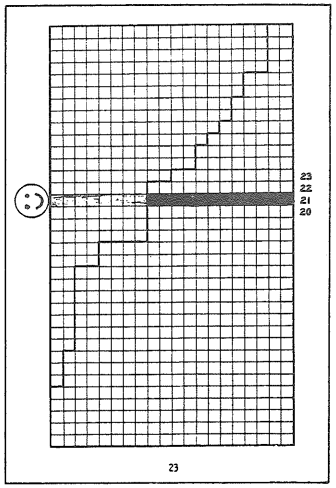
Students complete a balance sheet . . .

BALANCE SHEET for 20 day experimental period			
DAILY SUPPLY	TOTAL GAIN	TOTAL LOSS	TOTAL PROFIT
20	$35.10	$2.45	$32.65

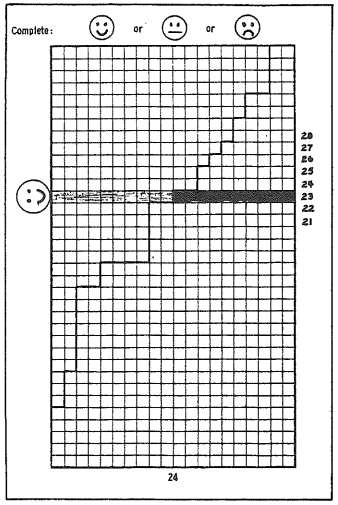
and find that 20 is a better choice than 35 for the constant daily supply. But is it the best possible choice?

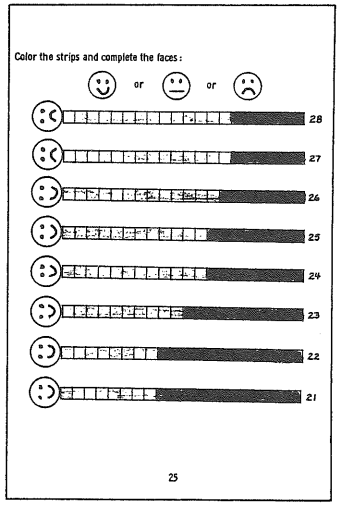
The next two pages of the booklet suggest a way to view the effect of changing the constant daily supply. Therefore, with a constant daily supply of 22, Shunda's total profit would be $34.25 ($32.65 + $80). This procedure demonstrates how the earlier stylizing of Shunda's profit-o-meter contributes appropriate experience for studying the graph that describes the demand process.





Students complete additional profit-o-meters on pages 24 and 25 of the booklet (shown here with solutions) to find that Shunda could continue to increase her profit by increasing the constant daily supply until the supply is 26. Changing the constant daily supply from 26 to 27 would result in less profit as suggested by the frown on the horizontal line for 27.





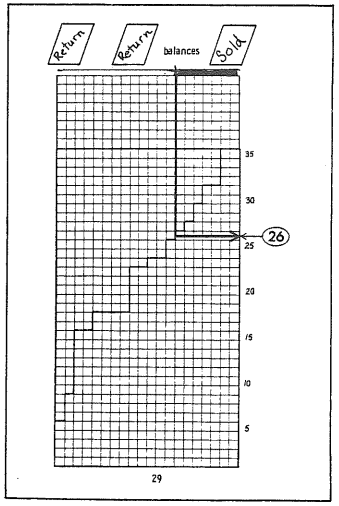
Shunda's Conclusion

During the twenty-day experimental period, the largest profit would have been obtained with a daily supply of 26 newspapers.

BALANCE SHEET for 20 day experimental period			
DAILY SUPPLY	TOTAL GAIN	TOTAL LOSS	TOTAL PROFIT
26	$40.80	$5.60	$35.20

Returning to the early observation that one newspaper sold balances two newspapers returned, Shunda takes another look at the graph describing the demand process for the twenty-day experimental period.





The story-workbook ends with a discussion of Shunda's conclusion. Since 26 papers would have been the best constant supply for the twenty-day experimental period, Shunda decides to adopt it. Of course, she cannot be sure this will remain the most profitable choice. She does not try to forecast future demand; what she does is make the best choice based on the limited evidence of the experimental period. As in most business ventures, she cannot avoid taking some risk.

This example of statistics activities for the elementary school intends to involve students in a real world application of mathematics. There is considerable calculation practice in a meaningful context. There is the opportunity to use bar graphs in a dynamic way to organize data. And, most importantly from a mathematical point of view, there is the experience of using statistics for problem solving.

POPULATION GROWTH

Population statistics appeal to students' natural curiosity about the world around them and the future that lies ahead. The sixth-grade CSMP curriculum includes a series of lessons on population growth. The lessons provide the opportunity to organize, interpolate and analyze real life statistics. More importantly, students use the data to make inferences about the past, present, and future - a rare activity at the elementary level in the study of statistics.

Population Growth Rates

A first lesson introduces students to the concept of population growth rate. Using their intuitive sense of ratio, students develop techniques to compute the growth rates and net population gains of several United States' cities.

The lesson begins with a discussion focused on the question: "What factors affect population growth?" Students conclude that the factors affecting population fall into four categories: births, deaths, immigration, and emigration. They proceed to develop several methods for computing growth rate and net gains. By not providing specific procedures for computing the net growth or the growth rate, the teacher allows students to create their own techniques. The strategies they develop increase their understanding of the concept "population growth rate" - the key for later lessons in this sequence.

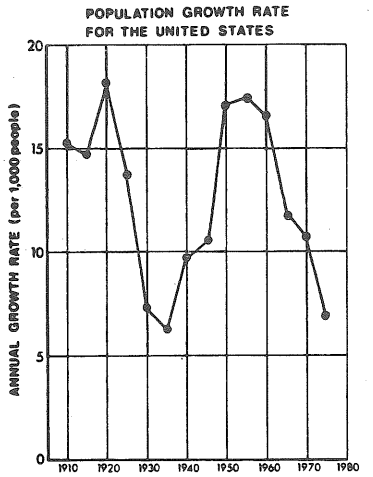
Tables and Graphs

The activities in a second lesson demonstrate that appropriate organization of data facilitates both the interpretation of the data and inferences based upon the data. Students begin by examining a mock newspaper article:

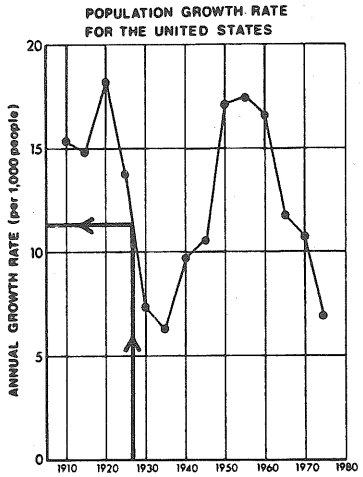
> The United States population growth rate has been declining. Growth rate per 1,000 people was 17.5 in 1955 whereas it was 10.7 in 1970 and 7.0 in 1975. It also declined for a while before World War II; it was 13.8 in 1925 and only 6.3 in 1935. There was a sharp increase in the growth rate after this war increasing from 10.7 in 1945 to 17.1 in 1950. Other yearly growth rates were 15.2 in 1910; 14.9 in 1915; 18.8 in 1920; 7.3 in 1930; 9.8 in 1940; 16.6 in 1960; and 11.8 in 1965.

The teacher highlights the effect of the disorganized nature of the data by asking several questions such as: When was the growth rate the highest? lowest? During what periods was it increasing? declining? The students' difficulty in answering these questions motivates the central theme of the lesson: Are there better ways in which to present this data? After constructing a table and a graph, the class reconsiders the questions they previously found difficult thereby experiencing the benefit of organized data.

Year	Annual Growth Rate (per 1,000 people)
1910	15.2
1915	14.9
1920	18.8
1925	13.8
1930	7.3
1935	6.3
1940	9.8
1945	10.7
1950	17.1
1955	17.5
1960	16.6
1965	11.8
1970	10.7
1975	7.0



Further questions reveal the advantage of the graph over the table. For example, what was the growth rate in the year 1927? Through interpolation, students can estimte a 1927 growth rate of 12 people per 1,000. The graph also shows the large fluctuations in the U.S. population growth rate.



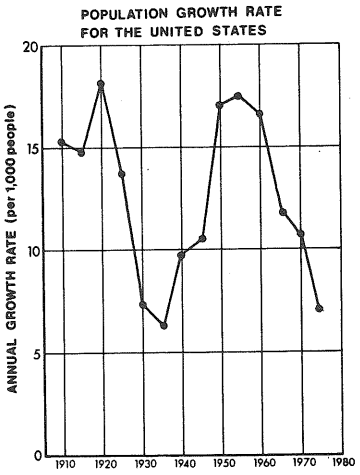
The remainder of the lesson focuses on some historical events that could have caused fluctuations. Students suggest the Great Depression and World War II as possible causes for the low rate of growth from 1930 to 1945. However, changing sociological views, such as people deciding to marry later or electing to have fewer children, seem a more likely cause of the recent low growth rate from 1970 to 1975.

The activity of organizing data, found in the first part of this lesson, is fairly commonplace in the study of statistics. However, the discussion of factors that may have caused the large fluctuations in the population growth rate represents an important shift in the lesson. The students move from simply reading the data to interpreting and making inferences based upon the data. These attempts to interpret data allow students to appreciate the purpose of organizing data into graphs, rather than rotely practising techniques of data organization. Correlating data with known historical and sociological factors provides a framework for later using current data to predict future events.

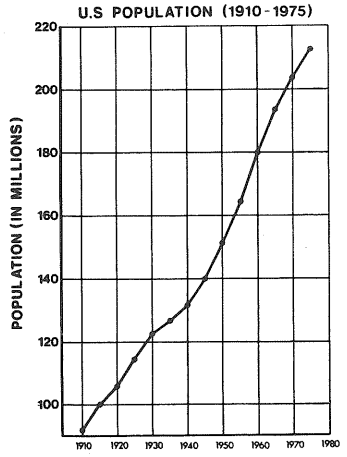
Population Pyramid

There are many ways to graph a set of population data. The choice of a particular graph reflects the feature of the data that the statistician wishes to highlight. In a third lesson, students explore several population graphs. First, the comparison of two distinct graphical representations of population growth during the period 1910 to 1975 allows students to observe the different ways that each graph portrays the same information.

A: Population Growth Rate for the U.S. (1910-1975)



B: U.S. Population (1910-1975)



Since these two graphs display the same census data, they reflect the same trends in different ways. For example, the population is always increasing, shown in graph A by the rate always being positive and in graph B by the total population always rising. Also, the low period in growth from 1930 to 1940 shown in graph A corresponds to a "leveling off" in graph B. Conversely, the sharp rise from 1950 to 1960 in graph B indicates the high growth rate

that is recorded in graph A. This activity encourages students to observe different ways that graphs can display rates of growth. In particular, the graphic comparison demonstrates that a declining (but positive) growth rate and a rising total population can occur simultaneously.

Besides studying the overall growth rate, demographers also analyze the United States population according to various factors: age, race, sex, religion, etc. With this in mind, the following statistic motivates the idea of population distributions.

United States: K-8	
1970:	34,300,000 students
1975:	32,000,000 students

The decline in total elementary school enrolment surprises most students. Who might be concerned about a decline in school enrolment? Teachers are most certainly concerned. Implications of this single statistic include school closings, teachers being laid off, and many other all too familiar problems. An open student discussion of this statistic focuses on such implications.

When first confronted with the above statistic, students might assume that the decline in school enrolment is due to an overall decline in population, contradicting the previous observation that the United States had a positive growth rate from 1910 to 1975. We then present the idea of population distribution by age group.

U.S. Population by Age Group - 1976

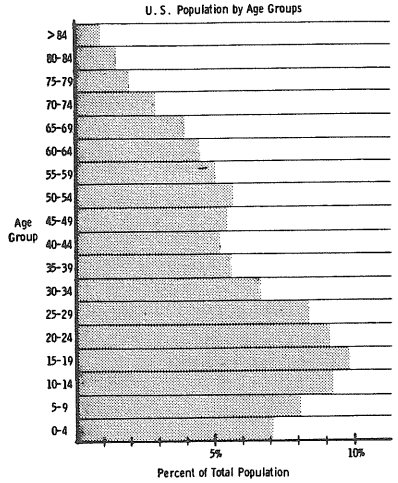
Age	Percent of Total Population
0-4	7.1
5-9	8.1
10-14	9.2
15-19	9.8
20-24	9.1
25-29	8.3
30-34	6.6
35-39	5.5
40-44	5.2

Age	Percent of Total Population
45-49	5.4
50-54	5.6
55-59	5.0
60-64	4.3
65-69	3.9
70-74	2.8
75-79	1.9
80-84	1.3
>84	0.9

By comparing the percent of population in the 5-to-14 age groups (the approximate K-8 enrolment in 1975) with the 10-to-19 age groups (the K-8 enrolment in 1970), students note that the enrolment loss appears due to a decrease in the number of people in that age group. The table reveals how the population in a particular age group can decline even as the total U.S. population continues to increase.

The introduction of population distribution by age suggests other questions of human interest such as: How many Americans remember World War II? World War I? How many Americans are over the age of 65?

After exploring these questions, we conclude the lesson by constructing a population pyramid of the United States population in 1976.



Comparative Population Graphs

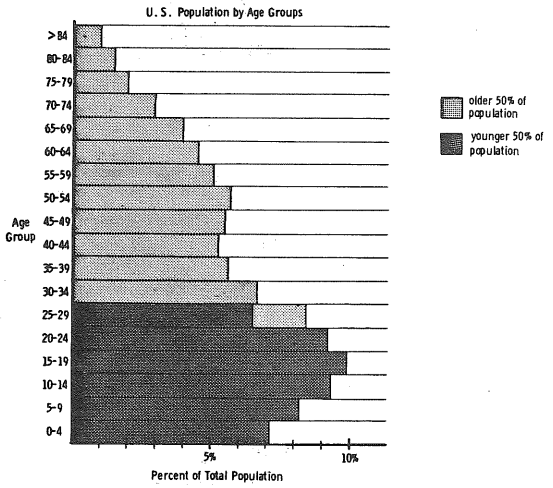
Upon graphing and interpreting a population pyramid for the United States, students naturally assume pyramids for other countries to be "normal". In a fourth lesson, the very dissimilar shapes of both Sweden's and Mexico's population distribution graphs as compared to the United States' graph disagrees with the students' assumption. A goal of this lesson is to interpret the three countries' graphs as a means for conjecturing the political priorities of the countries and what problems may lie in the future for each nation.

The beginning of the lesson returns to a conjecture made in the first lesson. Between the years 1930 and 1945 the United States experienced a low rate of growth. Two historical occurrences were given as possible causes: The Great Depression and World War II. The teacher challenges the students to use the population pyramid to determine which event appears to have had greater impact.

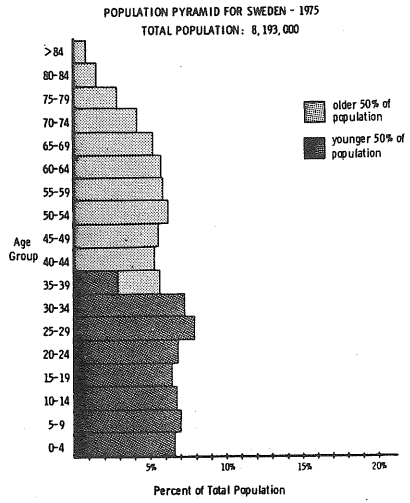
A notable feature of the US population pyramid is the relatively low percent of people in the 40 to 44 age group compared to neighboring age groups. Being of age somewhere between 40 and 44 in 1976 means being born sometime in the years from 1932 to 1936 - the midst of the Depression. Apparently, either people during the Depression chose to have fewer children or their children had a much lower life expectancy. The graph suggests that the Depression was a greater contributing factor to the lower population growth rate than was World War II.

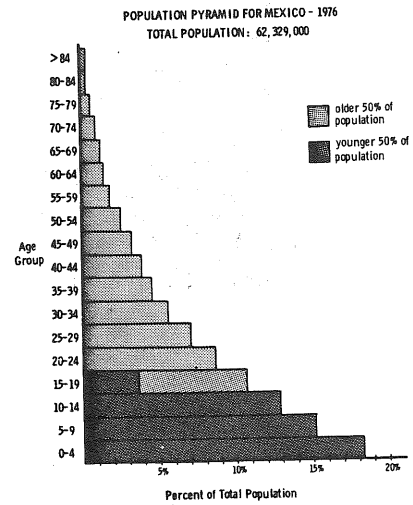
Besides graphs, another simple statistical tool for comparing national populations is averaging. The population pyramid does not allow easy computation of the mean age of the population; however, the median age is appropriate and is easily computed.

The median age of the US population in 1976 was about 28.



After completing this computation, students discuss whether other countries are likely to have similar median ages. Even at first glance, the population pyramids for Mexico and Sweden reveal dramatically different population distributions in different countries.





With Sweden's uniform population distribution, its median age is about 35, higher than that of the United States. In Mexico, a greater portion of the population is in the younger age groups; the country has a low median age of about 17.

What insights into areas of concern for these governments do these population graphs and medians suggest? The remainder of the lesson focuses on a comparison of some political and social issues these countries may soon face, such as:

* United States and Sweden have a large population in the older age groups who need support.
* United States is closing schools while Mexico needs to build more schools.
* Mexico's rapidly increasing population could cause shortages of food, housing, and health facilities.
* Sweden has the largest percent of older people, while Mexico has the smallest.

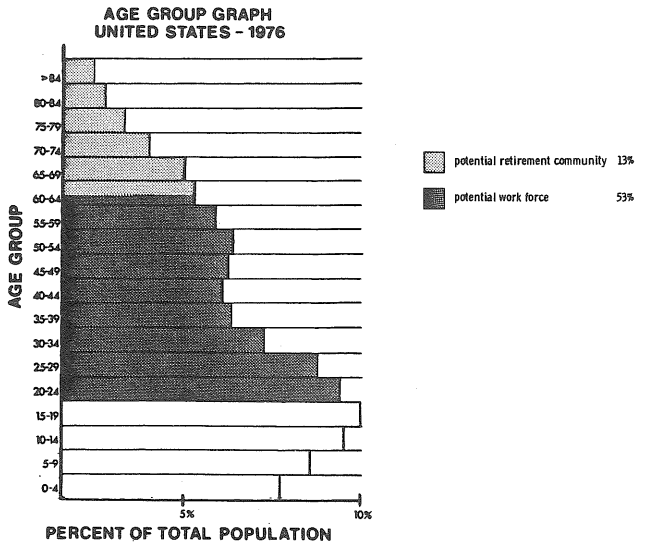
As in previous lessons, the lesson has moved to a discussion about the data. The importance is not so much the statistical tools that the students acquire, but the important realization that these tools enhance the discovery and discussion of events in the world around them.

An extension of the lesson studies the impact of the current population distribution of the United States to a topic of recent controversy: the Social Security System.

There have been several rises in the Social Security tax in recent years, each one promising to set the system on firm ground. As we shall see, the problem with the system may lie in its basic design, therefore calling for more creative measures to rectify the system. The system is designed so that the current work force generates the revenue for the retirees currently on social security. In order to gain only a crude overview and to simplify the analysis, we will contrast the total potential work force (ages 20-62) to the total potential social security receivers (over 62).

The following questions highlight some difficulties in the Social Security System:

* How many people in the potential work force does it take to maintain one person on social security?
* In 1981, how many people entered the work force for every person who entered the retirement system?
* Will the situation get better or worse? in 1980? (Assume life expectancy of 80).



The graph provides answers to these questions. In 1976, approximately four people (4 x 13 is about 53) in the potential work force were needed for every one person on social security (a ratio of 4 to 1). In 1981, however, approximately 5% of the population entered the social security system while only 10% of the population entered the work force (a ratio of 2 to 1); thus the potential work force is not growing as fast as the potential retirees. And as we can see from the graph, the problem will continue to worsen in the future since the population entering the work force is decreasing. The important statistical activity again gives the students an opportunity of not just taking the data at face value but drawing implications from the data in regard to some social issues.

SUMMARY

The series of lessons on population growth exhibits several valuable features for the teaching of statistics, namely:

* topics that interest students at a particular grade level,
* a problem-solving atmosphere, and
* a unifying factor for a variety of mathematical tools and concepts.

Students' curiosity about the world around them begins to emerge in the intermediate grades. Utilizing this curiosity is valuable in teaching, in particular teaching mathematics, effectively. Building on a foundation of interest, the mathematical content is then introduced as a means to resolve questions and to generate a source of new inquiries, thus creating a true problem-solving atmosphere. Students learn new statistical methods as well as gain an immediate appreciation of their application.

References

I. The story of SHUNDA'S NEWSSTAND was written by Georges Papy. Clare Heidema developed the series of lessons in the CSMP curriculum based on this story-workbook and she also wrote the paper on this topic in the book, CSMP: Probability and Statistics.

II. Richard Armstrong developed the series of lessons on population statistics for the CSMP curriculum and Thomas Giambrone wrote the paper on this topic in the book, CSMP: Probability and Statistics.

USING GAMES TO TEACH SOME PROBABILITY CONCEPTS

*George W Bright, Northern Illinois University, John G Harvey, University of Wisconsin and Margariete M Wheeler, Northern Illinois University

Since 1976 the authors have been engaged in a systematic investigation of the cognitive effects of games on mathematics learning. This investigation was initiated and continues for a variety of reasons.

1. Games seem to be important instructional activities, in the sense that teachers can choose to use a game in a particular instructional setting for instruction pertaining to a particular instructional objective.

2. Teachers of students aged 6-11 report (Graber, Rim & Unks, 1977) regular use of games in their classrooms. The types of games used were not surveyed.

3. Students aged 6-11 report (National Assessment of Educational Progress, 1979) not only that mathematics instructional games are used in their classrooms but also that games seem to help them to learn mathematics.

4. Prior to 1976 the limited and very fragmented research on the cognitive effects of mathematics instructional games had produced few substantiated results (Bright, Harvey & Wheeler, 1977).

From the outset, this investigation has focused on identifying effective uses of games so that recommendations for appropriate incorporation of games into school practices can be made. A game is defined by seven criteria:

1. A game is freely engaged in.
2. A game is a challenge against a task or an opponent.
3. A game is governed by a definite set of rules. The rules describe all procedures for playing the game, including the goal(s) sought; in particular, the rules are structured so that once a player's turn comes to an end, the play just completed cannot be retracted or exchanged for another move.
4. Psychologically, a game is an arbitrary situation clearly delimited in time and space from real-life activity.
5. Socially, the events of the game situation are considered in and of themselves to be of minimal importance.
6. A game has a finite state-space (Nilsson, 1971). The exact states reached during play of the game are not known prior to beginning of play.
7. A game ends after a finite number of moves within the state-space.

The first five of these criteria, based upon an earlier discussion of play, games, and gaming situations (Inbar & Stoll, 1971), give formal, psychological, and sociological characteristics of games. The last two criteria were evolved

in order to exclude other, related activities or structures such as play and puzzles.

An instructional game is a game for which a set of instructional objectives has been determined; these instructional objectives may be cognitive or affective ones and are determined by the person(s) planning instruction. The source of ideas for the instructional games used in the research was Developing Mathematical Processes (DMP) (Romberg, Harvey, Moser & Montgomery, 1974, 1975, 1976). DMP is a rich source of mathematics instructional games for elementary school students. These games frequently also seem easily adaptable for other ages and for more sophisticated content.

Before the details of the research are presented, the concept of cognitive effects needs to be defined. In describing a taxonomy of educational objectives, Bloom (1956) stated that the cognitive domain "includes those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills" (p.7). Thus, cognitive effects are effects which describe changes or differences in the attainment of objectives in the cognitive domain. For the study reported here, the cognitive effects are defined as the change in attainment of an objective in the cognitive domain due to a particular instructional treatment; that is, due to the playing of mathematics instructional games.

In DMP one of the grade six (about age 11) lessons is designed to introduce the mathematical concept of fair/unfair by having students play three pairs of related games. (One pair of the games follows the reference list). For each pair, play of the games results in counts of the number of times each of the two players is a winner. By comparing these counts, students decide which game of the pair is more fair. The three pairs of games were expanded to eight pairs (Bright, Harvey & Wheeler, 1981) and used in a series of studies, one of which was with 36 grade six (about age 11) students.

The assumptions underlying this study included the following:

1. The mathematics content was substantial and would not be learned without instruction specifically designed to teach that content.
2. Prior to the start of the study, the students had not received instruction intended to provide mastery of the content.
3. To measure the effectiveness of games on substantial content, which were played in isolation from other instruction on that content, it was not important to have a control group which did not play the game.
4. The concept, fair/unfair, as embodied in the games was at the knowledge cognitive level (Bloom, 1956). Knowledge cognitive level content includes knowledge of specifics or of ways and means of dealing with specifics (pp.65,68).

The study was conducted on ten non-consecutive school days, during five consecutive weeks, in schools in Illinois. On day 1 two pretests were given: the four-card test of formal operations (Adi, Karplus & Lawson, 1980) and a 28-item content test. In each of the 28 items two games were described and students were asked to determine the fairer of the pair. For one 14-item subscale, items were accompanied by data summarizing 50 plays of each of the two games. (These data were generated by computer simulations of the games). The items on the other 14-item subscale were not accompanied by simulated plays. A sample item from the second subscale is provided in Figure 1.

Figure 1

Sample Test Item for Fair/Unfair Game

Which game gives you the best chance of winning or doesn't it make any difference?



A. Game 1 You spin Spinner A once and you win if you get blue.

B. Game 2 You spin Spinner B once and you win if you get green or orange.

C. It doesn't make any difference.

On days 2 through 9 the eight pairs of fair/unfair games were played. A different pair was played for 20 minutes on each of these instructional days.

After playing each pair, students were asked to choose the fairer game, or to say that the games seemed equally fair. On day 10 the content test was given as a posttest.

The means and standard deviations and the comparisons of the pretest and posttest scores are presented in Table 1. In order to determine whether there were sex-related differences, the posttest scores for each subscale and for the total test score were used in an analysis of covariance with the four-card

test and the corresponding pretest score as the covariates and with sex as the classification variable. None of these analyses revealed significant sex-related differences.

Table 1

Pretest - Posttest Comparisons

Instrument	Pretest Mean (SD)	Posttest Mean (SD)	t	p
14 items with data given	5.72 (1.70)	7.33 (1.74)	4.33	.000
14 items without data given	5.31 (1.67)	5.50 (1.88)	0.45	.654
total test	11.03 (2.56)	12.83 (2.66)	3.14	.003

The fair/unfair games seemed to have helped students learn to distinguish fair from unfair situations in the presence of data about the outcomes of playing those games. This effect did not seem to generalize to situations in which students were presented two games without any accompanying data. Perhaps this result is related to the fact that students had not received instruction on probability or on the concept of fair/unfair. Their only source of such knowledge would have been their personal experiences outside the classroom.

It is encouraging that the students seemed to have learned from their interactions with each other about how to interpret the data that they generated by playing the games. This same pattern of results was also noted in a replication of this study with grade eight students (Bright, Harvey & Wheeler, 1982).

One possible follow-up study would be to interview students regarding determination of the fairer game when sample data are not provided. The effects of knowledge about probability might be expected to display itself in the strategies that students use in making this determination. For example, examining students' perceptions about relative areas of spinners or relative distributions of sums of numbers on dice might provide useful information about the ways that students try to build mathematical structures from their personal knowledge bases.

References

[1] Adi, H., Karplus, R. & Lawson, A.E. (1980) Conditional logic abilities on the four-card problem: Assessment of behavioral and reasoning performances. Educational Studies in Mathematics, 11, 479-496.

[2] Bloom, B.S. (Ed.) (1956) Taxonomy of educational objectives. New York: McKay.

[3] Bright, G.W., Harvey, J.G. & Wheeler, M.M. (1977) Cognitive effects of games on mathematics learning. Madison, WI. University of Wisconsin - Madison. (ERIC Document Reproduction Service No. ED 166 007).

[4] Bright, G.W., Harvey, J.G. & Wheeler, M.M. (1981) Fair games, unfair games. In Schulte, A. & Smart, J. (Eds.), Teaching statistics and probability: 1981 yearbook of National Council of Teachers of Mathematics. Reston, VA: NCTM, 49-59.

[5] Bright, G.W., Harvey, J.G. & Wheeler, M.M. (1982) Teaching statistics concepts through instructional games. Paper presented at the First International Conference on the Teaching of Statistics, Sheffield, England, August 1982.

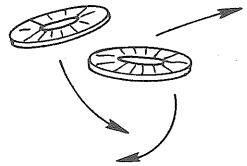
[6] Graeber, A.O., Rim, E. & Unks, N.J. (1977) A survey of classroom practices in mathematics: Reports of first, third, fifth and seventh grade teachers in Delaware, New Jersey and Pennsylvania. Philadelphia, PA: Research for Better Schools, Inc.

[7] Inbar, M. & Stoll, C.S. (1970) Games and Learning. Interchange, 1(2), 53-61.

[8] National Assessment of Educational Progress (1979) The second assessment of mathematics, 1977-78, released exercise set. Denver, CO: The author, May 1979.

[9] Nilsson, N. (1971) Problem solving methods in artificial intelligence. New York: McGraw-Hill.

[10] Romberg, T.A., Harvey, J.G., Moser, J.M. & Montgomery, M.E. (1974, 1975, 1976). Developing mathematical processes. Chicago: Rand McNally.

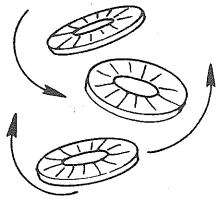


Flipped Chip

Game 1

(2 players)

Each player will need 15 chips. Take one extra chip and put a blank label on each side. Label the sides A and B. Decide which player will be A and which will be B.



Flip the chip. If A comes up, player A places a chip on his or her car. If B comes up, player B places a chip on his or her car.

The winner is the first person to fill all 15 chip spaces. Play two or three games. Then answer these questions:

Does each player have an equal chance at winning?
Does the same person win each time?
Is the game a fair game?

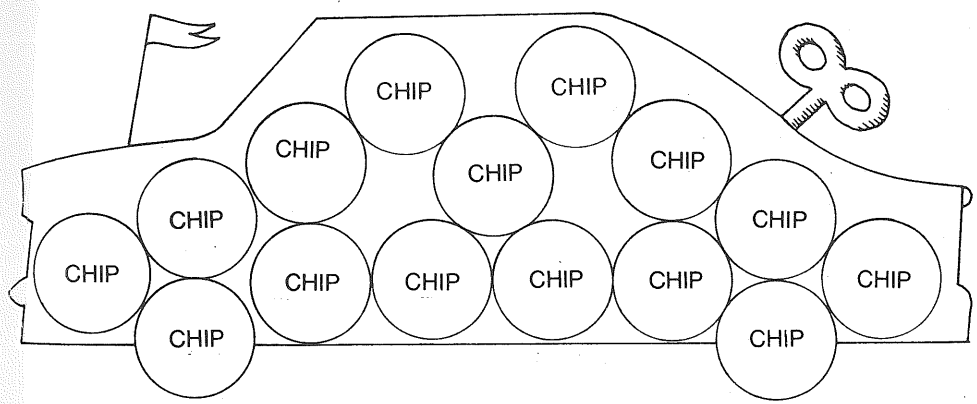
Game 2

(4 players)

Label three chips so that each has an A side and a B side. One player (AAA) puts a chip on his or her car if three A's come up. Another player (ABB) puts a chip on his or her car if exactly two B's come up. The third player (AAB) puts a chip on his or her car if exactly two A's come up. The fourth player (BBB) puts a chip on his or her car if three B's come up.

Flip all three chips at the same time. The winner is the first player to cover his or her car. Play two or three games. Then answer these questions:

Does each player have an equal chance at winning?
Does the same person win each time?
Is the game a fair game?



TEACHING STATISTICS IN ELEMENTARY SCHOOLS IN JAPAN

H Midzuno, Nippon Institute of Technology and Prime Minister's Office, Japan

Before starting a discussion of Statistics Education (SE) at elementary schools, it may be necessary to explain the nature of elementary education in my country.

Nine years of elementary education is compulsory. Parents or custodians of children between 6-14 years of age are obliged to send their children to schools, as stated in the Basic Law of Education. The Law is considered as a supplement to the Constitution, with a special preamble referring to the Constitution.

In the Constitution the basic right to receive education and the basic obligation of parents and custodians to provide education are stated with some related matters but the details are delegated to the Law.

Article 26 of the Constitution states:
"All citizens have the right to receive education equally, according to their ability as specified in the Law. All citizens are obliged to let the children under their custody receive ordinary education as specified in the Law. Compulsory education shall be free".

The preamble to the Basic Law of Education states:
"We have established the Constitution of Japan and revealed our determination that we construct a democratic and cultural nation and contribute to world peace and the welfare of the human kind. The realization of this ideal has to base itself upon the power of education.
We expect to bring up persons who respect the dignity of the individual and seek for truth and peace, as also we have to promote thoroughly education aiming to create a culture which is both universal and unique.
Expressing clearly the objective of education based on the spirit of the Constitution of Japan, we hereby enact this Law, in order to establish the basis of education for the new Japan."

Based on this, the Law determines many points important in education, e.g. objective and so forth, and as the result of these provisions the basic nature and scheme of education have been determined. Article 6 states "Schools referred to in the Law are public in nature and shall be established either by the nation, local public bodies or the legal bodies stipulated in the Law". Objectives are stated and matters pertaining to educational courses are to be determined by the supervisory organs.

An important specification is about the textbooks. Only textbooks approved by the Minister of Education or prepared by the Ministry of Education can

be utilized.

Accordingly the substance is included in the Guideline. This is issued by the Ministry of Education, as the supervisory organ is set to be, for the time being, the Ministry of Education.

Thus the Guideline prepared and the textbooks approved determine the contents of teaching almost entirely.

The Guideline was revised in 1978, and education has been going on since 1980 using thoroughly the new materials.

The Composition of the educational courses is set to be,

National language
Social discipline
Mathematics
Science
Music
Drawing and handiworks
Home Affairs
Sports
Morals
and special activities

In private schools religion may be added or exchanged with morals.

The number of hours and their allocation are also set in the Law.

Teaching of any subject, including statistics, has to find its place somewhere in such a scheme.

It was one of the important issues where to include teaching of statistics. Statistics are used and taught in sciences and social discipline but the basic and systematic, if one may say so, lectures were decided to be given in mathematics.

The goal of the educational course for mathematics states, "To provide the basic knowledge and technical abilities pertaining to quantity and figures and to build up the ability and attitude which enables to understand mathematically, think systematically and reasonably and handle the events occurring commonly." Teaching statistics is also included in this goal.

The goal is further sub-specified with relevant contents for each grade. For such purposes the field or educational course of mathematics is sub-divided into:

A. Numbers and computation
B. Quantities and measurement
C. Figures
and
D. Quantitative relations

and statistics is understood to belong to D.

The present Guideline was revised in 1977 superseding the one issued in 1968. At that time "to provide less tight but complete school life" was one of the principles, and it was suggested, in short, to reduce the contents, limiting only to the basic and those to be repeatedly treated are to be simplified.

Thus, if we compare the two Guidelines, we note the teaching of statistics is also simplified or deferred in general.

In the current Guideline, explicit sub-division of D (Quantitative Relations) into "presentation by formula" and "statistics" has vanished. Statistical subjects start from the third grade while they were from the second in the previous one. The term "probability" was included in the former one, but it is deleted from the present. The parts of the Guideline where statistics is explicitly or implicitly referred to are as follows:
"To arrange in order", "to sort and count", "to express in the form of tables and/or charts", and "to read tables and/or charts" appear here and there but not necessarily in connection with statistics. Thus, such were not included in the following:

Grade 3

Goal No.4

Through arranging the material and using formulae and graphs, improve the ability to express and study quantities and their relationship.
(One of the 4 goals for Grade 3).

Contents

D. Quantitative relationship

(2) To be able to express the data in, and read them from, tables and/or charts.

a. To sort and/or arrange the data in tables on the basis of time and place.

b. To learn how to read and draw bar graphs.

Grade 4

Goal No.4

To improve the ability to express or study quantities and/or their relationship on the basis of formulae and graphs as also to be able to investigate dependency and/or classify and arrange according to the objective.

(one out of 4)

Contents

D. Quantitative relationship

(1) To improve the ability to express or study the relationship between two related variates.

a. For simple cases, to study the relationship through considering the corresponding quantities and/or expressing the set of values in the form of a table.

b. To express the change in line graphs and/or to understand the characteristic nature of the change from such.

(3) To improve the ability to collect, sort and arrange and study the characteristic feature according to the purpose.

a. To study the cases of occurrence for two events.

b. To check the omission and duplication of data.

c. To express the data in bar graphs and line graphs or to study the characteristic feature and trend from graphs.

Grade 5

Goal No.4

... Also, to improve the ability to study statistical data through the use of percentages and pie graphs.

(one out of 4)

Contents

B. Quantity and measurement

(3) Deepen understanding of the observation of the magnitude of quantity and the meaning of the observed values.

...

(b) To understand the meaning of the average and use it.

D. Quantitative relationship

(1) Explain the meaning of percentages and enable the pupil to use them.

(4) To be able, according to the purpose, to sort and arrange the data, express them in pie graphs and bar graphs.

Grade 6

Goal No.4

Along with improving the ability through the understanding of proportion etc. to consider functionally, improve the ability to consider and/or express statistically, e.g. studying the distribution of the data.

Contents

A. Numbers and computation

(2)

e. To understand that numbers can be used to express the degree of occurrence of undeterministic events.

D. Quantitative relationship

(3) To improve further the ability to consider and/or express statistically, - e.g., to study the dispersion of the data for simpler cases.

a. To learn about tables and/or charts expressing the frequency distribution.

b. To learn there are cases where the tendency of the entirety can be known by, for example, rates obtained from partial data.

c. To select adequately tables and/or charts according to the objective and to devise something convenient.

(4) To develop gradually the ability to study the possible cases for occurrence through the orderly arranging of simple events.

The annotated part of the Guideline provides goals as follows:

Middle grades

... to improve the ability to express and consider quantities and their relationships with formulae, tables, charts and the like as also to improve gradually the ability to study the dependency, characteristic feature of, through sorting and arranging, the materials, according to the purpose. ...

Higher grades

... main objectives are ... to improve the ability to consider statistically, e.g. to survey the distribution of data ...

To achieve the goals set the contents of the syllabuses are grouped into the 4 regions of

A. Numbers and computation
B. Quantities and measurement
C. Figures
D. Quantitative relationship

D is a little different from others and may be further grouped into, (a) understanding of function, (b) expressing and/or interpreting formulae, (c) statistical consideration, (d) understanding of aggregation, etc. ...

D is not provided for grades 1 and 2.

c) Statistical consideration

The objectives here are to find out the relationship, tendency and rules existing, through observing as an aggregate, out of events which are governed by randomness, each element of which does not seem to follow any rules.

(1) Characteristic feature of the studies

In order to improve the attitude and ability to consider statistically, let the pupil study a series of processes characterized as follows so that they master the knowledge, techniques and fundamental thoughts involved in them:

a) Clarify the problem and objectives and collect appropriate materials. Then quantify them adequately.

b) Sort the data from the viewpoint based on the objective and facilitate grasping of characteristic features.

c) Interpret what kind of characteristic feature or tendency the aggregate has, through sorted data and/or means and ratios.

d) Infer on the basis of partial materials, what kind of statement can be made with regard to the population. Further engage them in inferring the tendency and/or characteristic feature not covered by the materials, and trying to solve the problems.

In the middle grades b) and c) are highlighted, to provide the basic knowledge and techniques of statistics.

Starting with grade 4 such contents as to collect materials on the basis of the purposes, to consider the general tendency ... are to be given.

(2) Cases for lower grade

As given, for example, "to sort simple cases and express them in tables and/or charts..." cf. (1) for grade 2, it is necessary to give basic experiences of statistical processing of sorting and then facilitating understanding of the characteristic feature of the data and make accustomed with such viewpoints.

Contents for respective grade

Grade 3

Arranging data and graphs

a) Sorting and arranging data

b) Reading and drawing bar graphs

Grade 4

(11) Relationship between two dependent variables

a. Detecting rules

b. Reading and drawing line graphs

(13) Sorting and arranging data and graphs

a. Cases of occurrence of two events

b. Study of characteristic feature and/or tendencies

Grade 5

(11) percentages

(12) Relationship of quantities expressed by a simple formula

(14) Pie graphs, belt graphs

Grade 6

(2)

c) Expressing by a number, the degree of occurrence of undeterministic events

(8) Dispersion of data

a) Tables and/or charts for frequency distribution

b) General tendency

(9) Possible cases of occurrence

Complete sets of textbooks for the educational course in mathematics are to be shown with some explanation in the meeting.

The author feels the basic skeleton of the framework of statistics education at elementary education has been provided, although there remain many other aspects needing clarification, e.g. pertaining to teachers, their qualification, special courses, additional materials to mention just one genre. Naturally to present a comprehensive view of statistics education in elementary education is too gigantic a problem and cannot be done here. Thus,the author intends to present one more aspect only.

The Japan Statistical Society is the professional institution representing the field of statistics with about 1,000 members. Some members are interested in the matter of statistical education in elementary education but they are a very minor group and reports pertaining to such do not appear significantly. Besides JSS, there are many professional societies in the field of statistics, e.g. Mathematical Society of Japan with a section on Mathematical Statistics, Applied Statistics Society, Research Association of Statistical Sciences, etc. The situations are the same and very few or no teachers in primary or lower education have very limited contacts with such as they are organized in other associations and societies.

The Society for Mathematical Education in Japan and a few others are for mathematical education, thus concerned with statistics education, as explained elsewhere.

However, a unique association of statistics education which cannot be missed is the All Japan Council of Statistical Education Research as the association of prefectural Council of Statistical Education Research established in nearly all prefectures with memberships of 1,000-2,000 each.

Each of the prefectural Councils organized the training sessions and the research meetings every year and the national ones are conducted with about 1,500-

2,000 participants. They issue biannually "Statistics Education Research" incorporating reports of those sessions.

They have issued "Manual for Statistical Education", which is utilized extensively. The manual has been quite influential for statistical education in primary education.

The Council issued in 1980 the first collection of the Actual Examples of Statistical Education, which provides good examples of actual statistics education in primary schools.

TOPIC 2: TEACHING STATISTICS IN SCHOOLS
(12-15 YEARS AGE GROUP)

ORGANISER: PROFESSOR P L HENNEQUIN

WHY TEACH DESCRIPTIVE STATISTICS IN SECONDARY SCHOOLS?

Heinrich Winter, Rheinisch-Westfälische Technische Hochschule

"A man standing with his one foot on a hot stove and on ice with the other enjoys - in terms of statistics - pleasantly warm feet. That is the logic of the average". (H. Paulsen - a German doctor - in: DER SPIEGEL, Number 43, 1978, p.217).

In this paper I intend to propose some arguments in favour of Descriptive Statistics (DS) in the school curriculum. I think every pupil has to learn some elementary procedures, skills and concepts concerning gathering, representation, summarization, characterizing, analyzing and interpretation of data. The main aim of this paper is to convince teachers in Secondary Schools both of the importance and teachability of DS; therefore, I will give some practical suggestions too. Let me present my paper in the form of 5 theses; the first 4 may be called descriptive and the 5th prescriptive.

THESIS 1: DS is more substantial than many people assume.

(a) DS is more than computing averages or standard deviations of lists of numbers. Rather, DS concerns describing and modelling reality in which chance plays an important role as to quantities and values of any real entities undergoing variations and fluctuations.

Many pupils fail, when they have to solve a problem of common "deterministic" real life mathematics, even if it is easy with regard to computation. The main difficulty consists in mapping a real life situation onto an appropriate mathematical model. If the real life situation is indeterministic, then its mathematical modelling requires much more mental effort. The reasons are: (1) The regularities or laws to be seen in the situation are often quite hypothetical and may come into conflict with intuitive ideas. (2) The data must often be gathered by the pupils themselves and appear variable and unsteady. (3) The representation and processing of the data are not carried out by simple and fixed rules. (4) The interpretation of the solution requires a feeling for the special factors involved in the problem.

In this respect, DS can be considered as too demanding rather than as too trivial.

Let us have a look at some illustrative examples.

(i) Average temperature of a day

How can we define "the" average temperature of a day? Who is interested in this information? How exact should this average value be?

Pupils might suggest:

$$x_A = \frac{x_{max} + x_{min}}{2}$$

This definition is indeed common (e.g. in the USA); and the average is easy to find using a maximum-minimum thermometer.

But discuss this definition. Construct cases in which this average is misleading, is not misleading.

Assume, the "real" average (as arithmetic mean) is $\bar{x}$. Estimate the difference between x_A and x. It can be shown:

$$|\bar{x} - x_A| \leq \frac{x_{max} - x_{min}}{2} \quad \text{(a little theorem!)}$$

Proof:

$$x_{min} \leq \bar{x} \leq x_{max}$$

$$x_{min} - x_A \leq \bar{x} - x_A \leq x_{max} - x_A$$

$$-\frac{x_{max} - x_{min}}{2} \leq \bar{x} - x_A \leq \frac{x_{max} - x_{min}}{2}$$

$$|\bar{x} - x_A| \leq \frac{x_{max} - x_{min}}{2}$$

Another definition (also used by geographers) of average temperature is:

$$x_B = \frac{x_1 + x_7 + x_{13} + x_{19}}{4} \qquad x_i \text{ means: Temperature at i o'clock}$$

A new problem: Compare $\bar{x}$, x_A, x_B.

Assume the temperature-curve of a day is given on graph paper. How can you find an average temperature from that? In fact: We have actually a problem of calculus! But we can use the temperature-curve for computing an average temperature in different ways. We can easily add similar problems: average temperature of a month, of a year, and so on.

(ii) School report

A teacher (in Germany) has to judge his pupils' achievements in the last half year. He has to find a final mark (from 1 to 6. 1 = very good, ..., 6 = poor). The marks from 6 tests are the basis of the final decision. Here are the test marks of two pupils, A and B.

Test Number	1	2	3	4	5	6
Test marks of pupil A	2	3	3	1	3	6
Test marks of pupil B	4	4	4	4	1	1

Can you propose a fair final report mark for each pupil? Is A better than B? Why? Why not? Under which circumstances is B better than A?

Here are several possibilities of combining the 6 single marks.

Final mark by	Report mark of A	of B	Result
Arithmetic mean	3	3	A as good as B
Median	3	4	A better than B
Modal	3	4	A better than B
Maximum	6	4	B better than A
Minimum	1	1	A as good as B
33 1/3 % value	2	1	B better than A
66 2/3 % value	3	4	A better than B
Mid range	5	3	B more constant than A
Variance	2 1/3	2	
Function in time	oscillates	decreases	B becomes more and more better

Well, what is "really" a fair report mark? Each judgement may be wrong. But we must finally come to a decision. The statistical processing of the data helps us only to describe our decision; we need, however, extra-mathematical arguments to justify our decision.

(iii) Pocket money

Someone may say that the higher the children's pocket money the higher is the income of their parents.

Is that a true statement? How can we test it?

A classroom experiment may be carried out - in fact a delicate matter - perhaps by means of an anonymous interview. Each pupil writes two numbers on a slip of paper:

My pocket money ..... DM
My parents' monthly income..... DM

We collect the slips, mix and numerate them. We note the information in tabular form, for instance (some fictitious figures!):

Nr.	x	y	Nr.	x	y
1	20	2 000	11	20	2 500
2	20	2 300	12	15	2 400
3	15	2 200	13	15	2 000
4	10	1 500	14	10	2 000
5	-	2 400	15	40	2 900
6	15	2 100	16	30	3 200
7	20	1 800	17	10	1 800
8	25	1 500	18	-	1 400
9	30	3 000	19	30	2 100
10	25	3 400	20	30	2 200

x: pocket money in DM; y: monthly income in DM

This table shows the data in a badly arranged way. How can we make it more transparent?

Here is a lucid possibility, the so-called stem-and-leaf display:

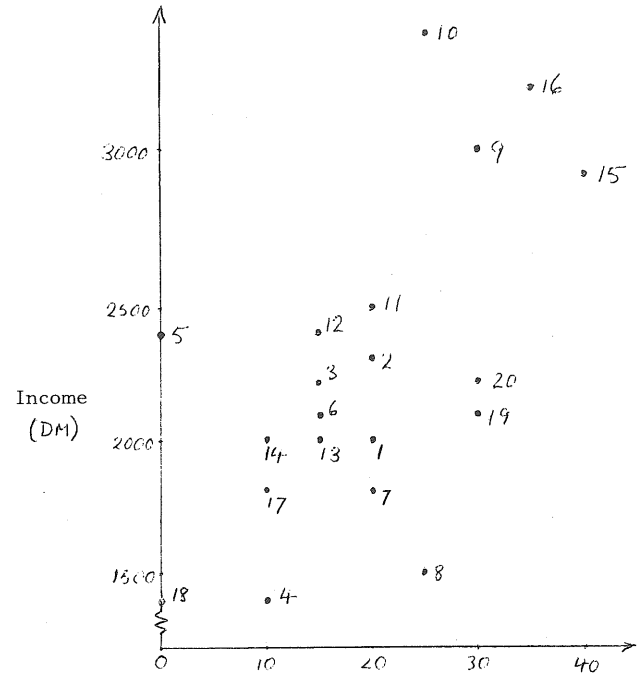
pocket money (DM)		incomes (100 DM)	
0	00	1	45588
1	0005555	2	000112234459
2	000055	3	024
3	0000		
4	0		

Obviously this is a better presentation than the traditional histogram. We have a "numerical sight" of the frequency distribution, too.

We are, however, mainly interested in the relation between the two random variables: pocket money, income. Therefore, we need a simultaneous representation.

A scattergram gives us a first answer about how allowance is correlated with income (in our sample!).

At the first glance the scattergram (next page) presents a confused picture; it seems to be a muddle. But on closer examination we may discover a trend which corresponds to our statement: the higher the pocket money the higher the income. On the other hand, we see many "exceptions", too. Can we calculate the extent of the correlation in some way?



A four fold table might be suggested:

y>2100	3,5,12 d=3	2,9,10,11,15,16,20 b=7
y≤2100	4,7,13,14,17,18 a=6	1,7,8,19 c=4
income y / allowance x	x ≤ 15	x > 15

Pupils may discover a measure of association, such as:

$$k = \frac{(a + b) - (c + d)}{a + b + c + d}$$

In our case, k is (13 - 7) : 20 = 0.3.

Now discuss k.

As is well known, the usual measure of correlation is:

$$f = \frac{ab - cd}{((a + c)(b + d)(a + d)(c + b))^{1/2}}$$

but it is surely impossible to explain it to pupils of this age.

Pupils should be encouraged to find another approach to our correlation problem.

We note from our sample the rank of each pupil according to his allowance and to his parents' income. We obtain this new table:

Number of pupil	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Allowance rank	10	11	6	3	1	7	12	14	16	15	13	8	9	4	20	17	5	2	18	19
Income rank	6	13	11	2	14	9	4	3	18	20	16	15	7	8	17	19	5	1	10	12

If the allowance would be proportional to the income then the lower rows would be identical. But only the 17th pupil has the same rank in both cases. An obvious measure of association might be suggested: the arithmetic mean of the absolute differences between the respective rank numbers.

In our case we obtain 90/20 = 4.5. But this is a bad measure, because it is not standardized.

Let's seek a better measure. Look at the rank pairs $\binom{10}{6}, \binom{11}{13}, \binom{6}{11}$

Compare all combinations of rank pairs; there are 190 cases. We can define:

$$\binom{x_a}{x_i} \text{ is concordant } \binom{y_a}{y_i} \text{ if } \begin{cases} x_a < y_a \text{ and } x_i < y_i \\ \text{or} \\ x_a > y_a \text{ and } x_i > y_i \end{cases}$$

$$\binom{x_a}{x_i} \text{ is discordant } \binom{y_a}{y_i} \text{ if } \begin{cases} x_a < y_a \text{ and } x_i > y_i \\ \text{or} \\ x_a > y_a \text{ and } x_i < y_i \end{cases}$$

For example: $\binom{1}{14}$ is concordant $\binom{15}{20}$; $\binom{1}{14}$ is discordant $\binom{7}{9}$.

Now we classify all 190 possible cases. If P is the number of concordant, Q the number of discordant cases, then

$$g = \frac{P - Q}{P + Q}$$

is a reasonable measure of correlation. Here we find $g = \frac{130-60}{190} \approx 0.368$.

(b) DS as a part of stochastics probably does not comprise as many ideas, problems, theorems as e.g. geometry does; but one should not consider DS merely as a collection of poor tools applied in non-mathematical fields. On the contrary, there are some mathematically interesting concepts and theorems. For instance, we may discover, formulate and prove the following theorems:

Arithmetic mean as a linear concept

(i) $\overline{(x + a)} = \bar{x} + a$

(ii) $\overline{(x \cdot a)} = \bar{x} \cdot a$

(iii) $\bar{x}$ is the centre of gravity of the frequency distribution.

(iv) $V = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2$

(v) Minimum property of the mean $\bar{x}$ with respect to the variance

$$V = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \leq \frac{1}{n} \sum_{i=1}^{n} (x_i - x)^2 \quad \text{for all } x.$$

(vi) Minimum property of the median $\tilde{x}$ with respect to the arithmetic mean of the absolute differences:

$$D = \frac{1}{n} \sum_{i=1} |x_i - \tilde{x}| \leq \frac{1}{n} \sum_{i=1}^{n} |x_i - x| \quad \text{for all } x.$$

and so on.

Proof for (v) without using any methods of calculus:

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - x)^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i^2 - 2x_i x + x^2)$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2x \cdot \frac{1}{n} \sum_{i=1}^{n} x_i + \frac{n}{n} x^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2x\bar{x} + x^2 + \bar{x}^2 - \bar{x}^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \underbrace{(\bar{x} - x)^2}_{\geq 0} - \bar{x}^2$$

$$\geq \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 = V \text{ (iv)}$$

These few statements might have sufficiently supported my thesis 1, namely DS is more substantial than is generally assumed.

THESIS 2: Teaching DS fosters general aims of education in compulsory schools

(a) If school education aims at the development of children's abilities to communicate in their daily and public life, then teaching DS is helpful. In particular, pupils can and must learn to understand, to criticize and to apply numerical or half-numerical arguments as they are widely used in our society.

Among other things, pupils should be enabled

(i) to read the information contained in graphs, diagrams and tables which we find in newspapers, in books, on television, ...,

(ii) to draw instructive graphs,

(iii) to detect misuses of (iconic) representations of data,
(iv) to become aware of the meaning of half-numerical expressions in natural language, like: roughly, approximately, about, round, almost, at least, never, seldom, by chance, ...,
(v) to distinguish between a numerical statement itself and its biased presentation, e.g. "What a disaster: Prices have rocketed up by 10% since last year!" - compared with: "Things are getting better: only 10% price rise since last year!",
(vi) to insist on further or possibly on complete information (for example, upon the statement "Homework of a 7th degree pupil lasts 2 hours a day" it is necessary to ask "Is 2 hours the minimum, maximum, median, modal, the arithmetic mean, or whatever? How was this value found, by a random sample? How large was the sample?"),
(vii) to unmask specious numerical arguments (for example: "In such and such a state house building is increasing greatly: This year 50% more houses were built than in the preceding year." - "Mother's milk is very healthy, for 9 in 10 of 90 years old people were breast-fed in their infancy."

(b) If school education aims at supporting the pupils' abilities to understand their environment, then we should teach DS. There are, around us, a lot of indeterministic processes. Pupils should become aware of, and, at least, basically understand randomness in our daily world. I give some examples: Weather in our country, weight of eggs, height of 7-year-old boys, lifetime of people in our country, quality of an industrial product, service times (at filling stations, telephone boxes, in shops,.....), nominal values and problems of calibration, prices and their ups and downs, peoples' consumer behaviour, illnesses and medicines, events in sport, and so on.

(c) Statistical methods of various levels infiltrate steadily into many jobs, even into the domain of handicraft. Random tests are increasingly used in order to control the quality of mass products, to control the nominal values (e.g. the weight of packets of sugar), to prove the effectiveness of a medicine, estimate the number of future customers of a product, to assess damage caused by hail, and so on. Therefore, DS should be taught, as school education has to prepare children for their future professional life.

THESIS 3: <u>The methods of, and the pursuits in, DS may concern, probably more than other mathematical activites, the pupil's actual existence.</u>

In my view, this thesis presents the prominent reason for teaching DS in compulsory schools.

We often notice a lack of motivation for learning in classrooms, in particular

in Secondary School classrooms, whatever the reasons are for that. Pupils are (openly or covertly) uninterested in most learning objectives. They obviously believe that school learning does not solve any problems of their actual life. I, of course, cannot offer suggestions, which would surely improve this deplorable situation. I think, however, DS gives a good chance to foster motivation for learning, at least for some pupils.

(a) Examples can easily be found in pupil's actual surroundings. There are a lot of subjects attracting his interest: sports, games, leisure time; school life; motor vehicles, traffic; illness and death; election forecast; and so on.

It is important to engage pupils in the choice of the subjects and in the collection of the data. The data should generally not be fictitious but as realistic as possible. The more authentic the problems are the stronger a pupil's interest and activities will be.

(b) Learning DS is learning important aspects of human life. Pupils become aware of life being an uncertain matter. Preferred subjects should be those which stimulate a pupil's thinking about his own existence now and in future.

A pupil can learn to realize that nobody and no institution will be able to guarantee a long life, a healthy life, an economically successful life, a satisfying job, to find a trustworthy partner, and so on. The future is dark and unknown, life is full of risks. A pupil, nevertheless, can learn how to provide for the future. Thus, every pupil should e.g. obtain a basic knowledge of insurance. This matter being rather complicated should be simplified in a responsible way.

Children also become aware of certain aspects of human learning and behaviour. Results of restricted observations are often generalized and decisions are made on grounds of observations. For instance: "A number of times I have bought a book in a particular book shop, and each time the shop assistant was friendly and knowledgeable. So, I think this is an excellent book shop". In the classroom might be discussed: How large does the number of observations have to be?

Can we learn on the basis of only one observation? What actually do we do in generalizing results from limited observations? Is the generalization grounded <u>in</u> the observations? Yes, but isn't it more than that? Do we not superimpose some mental construction on the data: do we not induce an hypothesis? How sure is such an hypothesis? How can we corroborate it? Can we interpret the data by means of the hypothesis? What should we do when new observations do not fit the hypothesis? In particular, children can learn to criticize the naive use of common sense and, above all, the role

of prejudices.

Another aspect of human behaviour should be led to awareness, I mean actions on grounds of intuitively assigned probabilities. E.g., we switch on the light, but the bulb stays dark. Why? We suppose: very probably the bulb filament is broken, or the fuse has gone. We hardly suppose: the power station has been blown up, although this would be possible. Thus, we would change the bulb rather than collect some money for a new power station. Another example: we go to the bus stop, although all the bus drivers may have fallen ill. This is reasonable because the probability of some epidemic among bus drivers is rather small. Obviously, we assign probabilities to the various events of daily life - rough probabilities, but mostly precise enough for common actions; and these probabilities originate from former experiences.

(c) By learning DS pupils experience direct applicability of mathematics. The pleasant feature of DS - unlike mathematical fields - is that pupils are concerned with interesting and significant problems without being obliged to use complicated and technical concepts and methods.

Some concepts of DS can immediately be applied. For instance: The modal of a sample is indeed a simple notion, and is extremely easy to find in the value table or in the frequency distribution of a (random) sample. But how significant are the applications of the modal in real life situations! Politics: absolute, simple, bare, vast majority in parliaments; majority votes; coalitions; plural voting; - Society: pressure by majority opinion; problems of minorities; influence of public opinions; - Checkings: Is the most frequent result always the "true" result? - Children can often directly watch statistical activities in their real world: customs examination; traffic census; traffic control; investigation of air or water or soil, and so on.

(d) Finally DS might offer the chance for one or other pupil to try a new start in learning mathematics at all. The immediate applicability and the small effort of conceptual developments could be a basis for motivating those pupils who have been unsuccessful in mathematics or have hated mathematics or have suffered from anxiety.

Perhaps the teacher attempts a new start and his work becomes more interesting and more successful.

THESIS 4: <u>DS can easily be merged into the canonical syllabus without overloading it.</u>

(a) DS should not appear well separated from classical school topics. On the contrary, DS can be closely interwoven with other common topics of mathematics. In addition, DS projects can be chosen from every other school

subject.

Mathematical topics:

Set theory: sets of statistical units; samples as subsets; frequency distributions as functions; classifying; ranging.
Fractions: relative frequencies; percentages.
Natural numbers: numerations; combinatorics.
Geometry: graphs, Monte-Carlo Methods.
Square equitations: variance; standard deviation.

Other subjects (examples):

Geography: weather; trade; production; soils.
Biology: life spans, growth; sizes.
Physics/Chemistry: measurements.
Road Safety Education: accidents.
Language: texts.
Economics: prices; incomes; production.

A further advantage of DS is that most of the topics can be described only by means of natural language, so that there is no need for endless exercises to make the pupils familiar with elaborated formula notations.

(b) Thus, there is little danger of overloading the canonical syllabus; DS, instead, may enrich aspects of other topics in mathematics.

Let me illustrate this through two examples:

(i) Fractions
Many children consider fractions like $\frac{29}{134}$ as monsters lacking in clarity and practical significance. The DS approach to fractions (≤ 1) makes more sense: $\frac{29}{134}$ should be interpreted as a relative frequency, for instance: Among 134 interviewed persons 29 preferred TV programme A. Is that a big proportion? Do people in general like this programme A? The situation leads to the transformation of $\frac{29}{134}$ into decimals or percent, 0.216 ... 21.6%, and this may be interpreted: About 22 in 100 persons preferred programme A, that is about 1 in 5 persons.

DS also offers further motivation for computing fractions, even multiplying fractions (≤ 1) : If f(a) is the relative frequency of the attribute a, f(b) the relative frequency of the attribute b, and if a is independent of b, then f(a) f(b) is the relative frequency of the combined attribute a and b. Now it also follows if two fractions are both smaller than 1, then the product is smaller than each of the fractions.

(ii) Functions

The notion of function is (at least since Felix Klein) a main and guiding idea, inspiring mathematics throughout.

DS supplies a host of various examples of functions:

- individual x ⟼ interesting attribute of x
- attribute a ⟼ relative frequency of a
- sample S ⟼ mean value of the sample S

and so on.

More important is the fact that DS may bring forth a deeper understanding of what a function essentially is.

For instance, we gather the data on distances and the respective fuel consumptions of a certain car. The driver noted two values each time he filled up his fuel tank. He may have obtained this table:

distances (km)	fuel consumption (l)
325	35.8
340	38.1
309	33.4
182	20.1
405	43.9
265	29.1
352	38.7
315	34.6
340	37.5
352	38.1

It seems clear that there is a strong dependence: the longer the distance, the more petrol was consumed. But is that exactly true? Does the table represent a function in the mathematical sense? Is there only one single value of fuel consumption for a given distance? Obviously not for

340 ⟶ 38.1
340 ⟶ 37.5

It is not unequivocal on the right. If any distance out of our table is given, then we can not always find only one corresponding consumption. Therefore, the table above is not the set of pairs of a function. On the other hand, we are not far from an exact function. We can write down a simple function approximately fitting the data:

f: distance (km) ⟶ fuel consumption (l)
x ⟼ 0.11 x = f(x)

However, is this function the best approximation? We found it by using our eyes! Is there a method to calculate best approximations? Well, we are very close to regression problems.

THESIS 5: In order to benefit by the educational value of DS, mathematical modelling of problems in the pupils' surroundings should be one of the main and fundamental activities.

This is a prescriptive thesis in so far as it contains some suggestions for classroom activities.

Teaching of DS will not produce any relevant educational effect, if we do not explicitly concentrate our efforts on the goals mentioned in the previous theses 1 to 4. Teaching DS will not automatically be successful in the sense I was trying to describe. Of course, mechanical skill training and the repeated solving of small tasks, all of the same type, is totally inadequate.

(a) At the start, one should have a challenging real life situation stimulating the pupils to talk and to think, to ask questions and to discover problems.

The extra-mathematical aspects of the situation must be carefully respected; these aspects should not be misused as merely methodical decorations. It is strongly recommended that pupils carry out projects by interaction with other branches of study.

(b) Pupils must become aware both of the powers and of the limits of statistical procedures. In particular, children should learn that most of the data are blind, useless and unsuitable for drawing any conclusions or for establishing any rules.

(c) At any rate, DS must not be reduced to description in a too strict sense. DS should be regarded as a part of stochastics. Therefore, aspects of probability must be incorporated.

(d) DS should be taught according to the principle of curriculum spiral. The notions and procedures of DS must gradually grow and be related to other mathematical ideas. DS should not be an isolated and separated subject, but should rather be an instructional principle.

To illustrate the 5th thesis let me outline an example:

Life and Death

(i) Possible starting points

- a TV discussion about problems of old age insurance (decreasing working population, increasing number of pensioners)
- a fatal traffic accident in our street
- a newspaper report about infant mortality in developing countries

The classroom discussion could bring up the question:
What is the life span of man?

(ii) First explorations

We seek information.

We look up the obituary notices in a newspaper.

For instance we find in the "Aachener Volkszeitung" (February 9th 1980; authentic figures!):

Age (of dying)	83	70	69	75	95	71	80	77	43	82	85	65	71
Sex	f	f	m	m	f	m	m	f	m	f	m	m	m

What does this table show? Try to concentrate the data. Do you dare to make any speculations?

We find averages:

Arithmetic means	74.3	(total)	Medians	75	(total)
	69.9	(male)		71	(male)
	81.4	(female)		82	(female)

It is worth noticing: Most peoplehad a life span longer than 70 years. On average, women had a longer life than men.

Can we, however, generalize this observation? How reliable is our source?

Further sources of data are:

- cemetery
- municipal administration
- official statistical books

(iii) Discussing a mortality table

The teacher presents the mortality table of the next page. The table might stimulate to discuss and solve problems like these:

- How can we obtain a good survey of these 100 numbers?
  Draw a cartesian graph of the function

  age x $\longmapsto$ living men of age x
  x $\longmapsto$ l(x)

  What strikes you? Describe this "mortality curve".

- Investigate the mortality per year. Where is the minimum, the maximum? What is particularly striking?
  Assume, the mortality per year would be constant!
  We define:
  Probability of death at age x = p(x) = $\frac{l(x) - l(x+1)}{l(x)}$
  This is the probability of dying in the course of the next year if someone is x years old.
  Draw a cartesian graph of the new function:
  x $\longmapsto$ p(x)

General Mortality Table 1970/72 Germany (FRG) Male

x	l_x	x	l_x	x	l_x
0	100 000	35	93 245	70	54 909
1	97 400	36	93 049	71	51 838
2	97 249	37	92 838	72	48 673
3	97 152	38	92 610	73	45 438
4	97 067	39	92 361	74	42 161
5	96 989	40	92 089	75	38 872
6	96 918	41	91 794	76	35 601
7	96 854	42	91 475	77	32 373
8	96 795	43	91 131	78	29 212
9	96 741	44	90 761	79	26 137
10	96 692	45	90 363	80	23 167
11	96 647	46	89 934	81	20 321
12	96 604	47	89 468	82	17 619
13	96 561	48	88 958	83	15 083
14	96 515	49	88 398	84	12 735
15	96 459	50	87 787	85	10 595
16	96 383	51	87 104	86	8 678
17	96 273	52	86 369	87	6 990
18	96 118	53	85 574	88	5 529
19	95 927	54	84 717	89	4 287
20	95 732	55	83 789	90	3 253
21	95 541	56	82 779	91	2 407
22	95 357	57	81 673	92	1 735
23	95 182	58	80 460	93	1 213
24	95 016	59	79 130	94	824
25	94 858	60	77 675	95	539
26	94 705	61	76 087	96	339
27	94 555	62	74 357	97	204
28	94 405	63	72 477	98	117
29	94 253	64	70 440	99	64
30	94 097	65	68 242	100	33
31	93 937	66	65 882		
32	93 773	67	63 361		
33	93 604	68	60 685		
34	93 429	69	57 864		

x: age

l_x: living men of age x

Source: Statistiches Bundesamt, Bevölkerung und Kultur, Kohlhammer, Stuttgart 1976

Describe the probability of surviving: x $\longmapsto$ 1 - p(x)

- A man is 40 years old today. What is the probability that he can celebrate his 65th birthday?
- What is the probability that all 3 new born babies are still alive in 50 years?
- Compute some probabilities concerning yourself.
- Can you find the median and the arithmetic mean of life span by using the table?

It is easy to find the median (roughly 71 years). To find the arithmetic mean, however, demands some calculations. It is useful and not very difficult to display the formula:

$$\bar{x} = 0.5 + \left(\sum_{x=1}^{100} l(x) \right) : 100\,000$$

Here is $\bar{x} = 67.4$. This value is also called "average life expectancy at birth". Find the remaining average expectancy of 15 (30,50,77,.....) years old men.

- It is a quite expensive matter to make a mortality table. How is it made? Who makes it? The teacher should outline how such a table is set up, emphasizing the basic ideas.
- Who might be interested in mortality tables?
  In a simple but responsible way the fundamental ideas of life insurance should be developed and some concrete examples should be considered. We calculate (e.g.) the life insurance premium rate a man has to pay under the following simplified conditions: given his age of insurance beginning, the amount of the policy, the length of the endowment period, the interest rate. The company has to pay the amount either at death or at the end of the endowment period, whichever occurs first. The premium has to be payed at once.

(iv) Comparing further mortality tables.

Excerpts from Mortality Tables from 1871/80 to 1970/72 (Germany):

	MALE						FEMALE					
AGE	1871/80	1910/11	1932/34	1949/51	1960/62	1970/72	1871/80	1910/11	1932/34	1949/51	1960/62	1970/72
0	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000
5	64 871	77 213	89 654	92 880	95 929	96 989	68 126	80 077	91 535	94 270	96 782	97 690
10	62 089	75 984	88 793	92 444	95 620	96 692	65 237	78 816	90 753	93 937	96 579	97 492
15	60 892	75 189	88 244	92 097	95 388	96 459	63 878	77 930	90 270	93 701	96 434	97 349
20	59 287	73 832	87 298	91 466	94 812	95 732	62 324	76 659	89 490	93 295	96 188	97 059
25	56 892	72 130	86 032	90 531	93 948	94 858	60 174	75 043	88 390	92 711	95 884	96 755
30	54 454	70 425	84 715	89 518	93 166	94 097	57 566	73 115	87 139	92 039	95 485	96 429
35	51 815	68 545	83 234	88 428	92 322	93 245	54 685	71 020	85 754	91 221	94 949	95 997
40	48 775	66 227	81 481	87 102	91 218	92 089	51 576	68 659	84 135	90 225	94 184	95 331
45	45 272	63 238	79 285	85 342	89 659	90 363	48 481	66 187	82 211	88 901	93 081	94 308
50	41 228	59 349	76 322	82 648	87 230	87 781	45 245	63 231	79 620	86 991	91 442	92 683
55	36 544	54 290	72 147	78 562	83 221	83 789	41 308	59 350	76 038	84 225	89 063	90 272
60	31 124	47 736	66 293	72 852	76 652	77 675	36 293	54 016	70 984	80 166	85 484	86 903
65	24 802	39 527	58 106	64 999	66 941	68 242	29 703	46 484	63 712	73 875	79 839	81 647
70	17 750	29 905	47 059	54 394	54 461	54 909	21 901	36 448	53 184	63 994	70 840	73 157
75	10 743	19 328	33 479	40 700	39 784	38 872	13 677	24 517	39 132	49 605	57 076	60 033
80	5 035	9 711	19 122	25 106	24 156	23 167	6 570	12 981	23 500	31 787	38 507	42 046
85	1 635	3 297	7 732	11 321	10 861	10 595	2 232	4 794	10 323	15 225	19 500	22 478
90	330	679	1 966	3 175	3 092	3 251	471	1 126	2 868	4 815	6 480	8 016

Source: Statistisches Bundesamt, Bevölkerung und Kultur, Kohlhammer, Stuttgart 1976

What strikes you?

- Investigate the development of infant mortality.
- Investigate sex differences. What do you think about that?
- Find the approximate medians and the arithmetic means for the different times.
- Make inquiries about mortality in several other countries.

References

[1] Barnett, V. (1980) Teaching Statistics in Schools in England and Wales in: Comparative Studies of Mathematics Curricula, Change and Stability, Institut für Didaktik der Mathematik, Materialien und Studien, Band 19, Bielefeld, p.444-463

[2] Dinges, H. (1982) Deskripive Statistik für die Mittelstufe, in: Der Mathematikunterricht, 28, Klett Stuttgart, p.5-27

[3] Engel, A. (1982) Statistik auf der Schule: Ideen und Beispiele aus neuerer Zeit, in: Der Mathematikunterricht, 28, Klett Stuttgart, p.27-85

[4] Ganadesika / Kettering/Tukey (1982) Themen aus der Datenanalyse: Begriffe, Beispiele und Pädogogik, in: Der Mathematikunterricht, 28, Klett Stuttgart, p.28-56

[5] Heller, W.D. (et al) (1979) Beschreibende Statistik, Birkhäuser Basel

[6] Winter, H. (1981) Zur Beschreibenden Statistik in der Sekundarstufe 1, in: Dörfler/Fischer (Ed.): Stochastik im Schulunterricht, Hölder-Pichler-Tempsky-Teubner Wien-Stuttgart, p.279-303

THE TEACHING OF STATISTICS IN FRENCH SECONDARY SCHOOLS

G Dumousseau, IREM de Bordeaux

Statistics is not at present taught to French 11 to 15 year olds. Syllabuses do not require it and text books offer very few activities or exercises bearing on statistics. Very simple statistical sequences may be found in some books, in the sections on proportionality, percentages or reading graphs.

Mathematics masters, who generally are not trained in statistics and who are not encouraged officially to use this subject, practically ignore it. History and geography masters employ more statistics, albeit in simple examples, like reading a chronological sequence in a table, or following the evolution of some economic variable for which the statistical treatment has already been carried out.

There are, however, many teachers who feel it is unreasonable that children should be left solely in the deterministic framework, and essential that they should also have contact, as soon as possible, with numerical data. Over the last decade, the research institutes for the teaching of mathematics (I.R.E.M.) and the association of state school mathematics masters (A.P.M.E.P.) have insisted on the scholastic value of interdisciplinary teaching, with activities related to the child's environment. It has been suggested to devote some of the teaching to the study of themes, and reports on experiments have been published.

All this should favour the development of statistics in secondary schools, but the introduction of new subjects and methods needs teachers with more extensive training, initially and in their careers, than is presently given to them.

The present state of affairs is likely to change, slowly, with the modification of fourth form (15 to 16 year old) syllabuses and with the promotion of educational projects. This may influence the future syllabuses in secondary schools. The educational projects, which are interdisciplinary, may be related to the child's surroundings and may include individual work or work in groups, etc., Various subjects have been tried, particularly regional economic studies, social and cultural enquiries.

Research

Since 1971 groups of teachers in France have worked or are working on the possible development of the teaching of statistics in secondary schools. This research has been organised by the national teaching research institute (I.N.R.P.) by the I.R.E.M.'s and by members of the A.P.M.E.P. Accounts of experiments

have been published (see the bibliographical appendix).

Looking back on all the work done in secondary schools, since 1971, I feel that the experiments carried out can be summed up by the points made in "Mathématiques du quotidien dans les collèges, activités d'analyse des données dans le premier cycle" published in 1979 by the I.N.R.P. - I.R.E.M. research group whose progress was specially followed by P.L. Hennequin from 1972 to 1978.

This research team comprised teachers working with the I.N.R.P. in Paris, or with the I.R.E.M.'s in Bordeaux, Grenoble, Lyons, Orléans, Rennes and Rouen.

Other groups have worked more recently on statistics in secondary education, e.g. at Clermont Ferrand, but their means, methods and conclusions are not different.

The points of view adopted

In secondary schools, the intention was to teach "neither probability theory, nor the conditions of validity of statistical tests. On the other hand, statistics may be taught with a view to analysing data, since that can be done before pupils are acquainted with the theory of probability. Moreover, while many interesting studies have shown that (even very young) children are interested by random phenomena, in the form of games based on the raffle, it seems that this method forgets an important aspect of the statistician's activity, namely not favouring any a priori probability model, but observing data obtained in a natural context".

The data are real and numerical. At this age pupils are not afraid of, and may even "be interested by the juggling of numbers". They are beginning to be proficient in the basic operations, but are curious to use and to get the feel of large amounts of numbers, or of large numbers.

A boat of 10,000 tonnes displacement means nothing to a young pupil who has just learned this number. But after a study of the traffic in a port, or perhaps visiting a port, his reaction will be much more conscious. "Moreover, while the analysis of numerical real data is in itself a statistical activity, it can lead to a probabilistic model. This may occur through the need to predict, or through the repetition of measurements".

The reality of the data is important to these adolescents who are becoming aware of their social and cultural environments. Besides this, the ability to understand a table, to interpret a graph or to use certain properties characteristic of statistical series are necessary training for all pupils, especially those who will not be going on to further education.

Besides these goals, some teaching principles were chosen for this training: Action and communication, at the basis, can be used to develop each child's learning ability, at his own pace, or to incorporate certain statistical or mathematical ideas brought up by the questions which the groups asked themselves about the data.

These general conditions imply that there is no question of stating a theory, but that "the goal is to create situations where the following abilities are required:

- understanding information
- criticising information
- making decisions with knowledge of the chance of error
- accepting discussion and the confrontation of ideas within a group"

"This approach is also an opportunity to teach mathematics a different way. By a continual exchange between the pupil's representation of a situation and the situation itself, a place is found for mathematical structures, at the level of cognitive analysis which may be used."

Scholastic objectives may then be put in three groups:

1) To teach the pupils the general method of the statistician, i.e. unprejudiced examination of the data; their analysis with respect to a question or a problem to be solved; a critical attitude to the data, the pertinence of the question to the methods employed and to the answer obtained. This critical attitude is developed by the comparison of ideas from different groups.

2) To acquire certain habits
The pupils are expected to know

. how to draw up a table of the data
. how to ask questions about the data
. how to take the initiative to answer these questions: choice of method, calculation, graphical representation or other methods of interpretation.
. how to criticise a statistical treatment or an interpretation.

3) To develop mathematical training and activities.

"The actual mathematical side of these goals is of course essential. During the treatment of the data, pupils use their previous knowledge, may build new tools for the analysis of the data and may learn new ideas specially introduced for the occasion".

Amongst other mathematical concepts and procedures which may be envisaged are:

- classifications, mappings
- work on $\mathbb{N}$, Z, $\mathbb{D}$, $\mathbb{Q}$
- mean and spread

- proportionality
- reading and drawing graphs
- an approach to probabilities

Thus, without adding to the mathematical syllabuses, we have a concrete "approach to many important parts of the syllabuses".

Methodology of most of the published research

There was not any laboratory research. It was decided to work on observations of groups of children in their classes.

The I.N.R.P. group writes, "Observation is a means of obtaining an objective opinion of teaching method. Instead of just saying "That worked", or "They didn't understand a thing", by careful osbervation we may collect sufficient information to explain the childrens' reaction to a teaching method. In our view, observation is an illustration and not a proof, since the point is not to prove a general truth, but to describe a particular truth.

The goal is thus to find out what was wanted to be done, what was wanted to be observed and what was done and what actually happened".

These observations allow us to infer how the children grasped the situation.

The shortage of psychologists in the research teams forced the teachers to learn, without the supervision of psychologists, the methods of objective osbervation and rigorously to respect instructions, particularly in order to reconstruct the chronological order of the work, the expected and unexpected behaviour.

Research on 11 to 13 year olds done by the I.N.R.P. and the I.R.E.M. involved seven teams from Bordeaux, Grenoble, Lyons, Orléans, Paris, Rennes and Rouen.

Each team comprised mathematics masters and psychologists. There was freedom of choice of the subject.

Pupils worked in groups of three or four. Teachers set the work in the least restrictive way possible, such as:

"With all these data what kind of enquiries can you imagine"

or "Find a subject that interests you and write a questionnaire to be given to your classmates".

The observers, one per group, whenever possible, noted the childrens' discussions, their behaviour and the teacher's action. The team's report was a synthesis of these observations. A synthesis of the team's work was presented in the group's report.

All the teams produced a report on the subjects they studied (see biographical

appendix).

Teaching situations produced in the classes

Once more, we shall use the I.N.R.P. - I.R.E.M. research as an illustration, but the results are true of what happened in other research groups.

135 ordinary urban and rural classes were involved in the I.N.R.P. - I.R.E.M. research.

"Topics studied by the teams were of general or local interest (variation of the population of a county, economic activity of a port) but were adaptable to any area, or were related to the childrens' immediate environment (surveys of hobbies, comics, television programmes, etc.,...). Topics were chosen with a view to the mathematical and statistical ideas they would involve and with a view to involving the children, e.g. by collecting data and building a table, children became more aware of the difficulty of carrying out an enquiry or a survey."

Method used when the data was supplied

Teachers distributed tables of data, generally large ones, and after a few explanations of how to read them, gave the least restrictive instructions possible for their use.

Pupils knew that the work could last several classes (sometimes a dozen hours). A period spent looking for questions about the data was followed by a class synthesis where ideas could be shared. Each group then chose a subject on which it was to work and to prepare a report.

The next stage involved work in groups. Time spent on this second part varied widely with the subject (2 hours, sometimes 4 hours). Setting out was sometimes slow, since the groups had to organise themselves, define their goals, choose secretaries, chairmen and the sharing out of the work.

Work in groups must benefit all the class and may help other groups, so at the end of this stage a class synthesis was held.

Chairmen of each group explained the aims and the work accomplished up to date.

The teacher must decide when to hold the synthesis. This is sometimes difficult to decide.

When much work is involved the above synthesis is an opportunity to review the subject, to foresee possible extensions, to decide on work to be corrected, completed or re-oriented and to reveal similarities between the work of different groups.

The third stage was often sufficient to finish the work in each group before the final synthesis which gave all the pupils the feeling of having participated in a collective activity... which they should be able to explain.

Some long subjects may need more than three sessions.

Pupils had to explain their work and to justify their actions in a way beneficial to all.

The teacher's role is above all to animate the sessions. When several or all the pupils are in difficulty, the teacher can organise information sessions, i.e. classes, specially on certain necessary mathematical tools.

Enquiries

The main difference with the above was in the first stage.

This involved several phases: choosing a subject, writing the questionnaire, preparing and carrying out the enquiry, examining the results obtained.

These phases are done or not by the pupils, depending on the teacher's intention when it was decided to work on enquiries.

Some enquiries involved a complete population of pupils, while others were based on samples and on the whole population in order to stimulate reflection on the idea of a sample.

The data

The type of data in each report is given over.

Behaviour observed

Owing to the amount of elementary work involved, to certain aims, especially the use of communication, and to the possibility of different rates of work, group work was preferred to individual work.

The advantages of group work were pointed out by all the research teams: mutual aid, spreading the work, which allows everyone to participate, the useful discipline implied by communication.

When there was argument between the leaders of a group, the teacher had a delicate part to play in solving problems of relations, or in making some groups efficient.

Pupils' interest was stimualted by the (relative) freedom in the choice of subject and the methods employed. They decided, in their own language, their topics. This freedom led them to accept responsibility, since choice implies a commitment, insisted on by the teacher, to produce a piece of work.

Report Number	Subject	Source
1	Weather (Bordeaux, Brest, Strasbourg)	Meteorological offices in Bordeaux, Brest and Strasbourg
2	Weather (Brest, Chateauroux, Orléans, Strasbourg)	Meteorological offices in Brest, Chateauroux, Orléans, Strasbourg
3	Water consumption in Grenoble	Grenoble Water Board
4	Electrical consumption in the South East from 1971 to 1975	South-East Electricity Board
5	Electrical consumption in the South-East on the third Wednesday in December*	Lyons and Marseilles Electricity Board
6	Numbers of pupils	Ministry of Education
7	Sporting performances	Files kept by the sports masters
8	Population of the Gironde and the Dordogne	l'I.N.S.E.E. (National Statistical office)
9	Football	Publications of the Fédération Française de Football
10	Car registration in the Rhône	Prefecture Lyons
11	Car registration in the E.E.C.	Motor trade paper
12	Loire river	Service hydrologique Centralisateur d'Orléans
13	Hobbies - holidays	Enquiry done by the pupils
14	Measurements of the length of a room	Done by the pupils
15	Port of Bordeaux	Bordeaux port authority
16	Port of Rouen	Rouen port authority
17	Sport results	Files kept by the sports masters
18	Surveys	Enquiries done by the pupils
19	Heights of 4th formers in a school in Orleans	enquiry done by the teachers
20	Heights and weights of 4th formers in a school in Valence	Enquiry by the pupils
21	Heights, weights and sports results	Enquiry by the pupils
22	Heights and weights from birth to 18 years	Collected by science masters from the health records
23	Television	Questionnaire sent to the pupils

* N.B: This is the last Wednesday off school before Christmas.

Number		Number of Tables	Dimensions of the largest tables	Amount of numbers given
1	Weather (Bordeaux, Brest, Strasbourg)	11	13 x 12	1,100
2	Weather (Brest, Châteauroux, Orléans, Strasbourg)	5	20 x 12	1,007
3	Water consumption in Grenoble	3	31 x 12	1,750
4	Electrical consumption in the South East from 1971 to 1975	10	31 x 6	1,795
5	Electrical consumption in the South East on the third Wednesday in December	3	24 x 6	204
6	Number of pupils	12	24 x 6	591
7	Sporting performances	9	25 x 2	250
8	Population of the Gironde and the Dordogne	4 2	15 x 9	435 270
9	Football	1	20 x 20	380
10	Car registration in the Rhône	2	8 x 12	188
11	Car registration in the E.E.C.	4	34 x 8	943
12	Loire river	10	31 x 12	3,653
13	Hobbies - holidays	0	0	0
14	Measurements of the length of a room	2	100	60
15	Port of Bordeaux	26	38 x 8	3,982
16	Port of Rouen	12	50 x 6	1,970
17	Sport results	2	34 x 5	260
18	Surveys	3	14 x 12	288
19	Height of 4th formers in a school in Orleans	1	2 x 8	16
20	Heights and weights of 4th formers in a school in Valence	8	34 x 4	800
21	Heights, weights and sports results	5	32 x 7	784
22	Heights and weights from birth to 18 years	8	39 x 4	1,070
23	Television	6	35 x 10	1,500

Most themes favoured interdisciplinary collaboration, which was indeed done in some cases. This much improved the atmosphere and everybody's interest. Questions asked by the groups required the teacher to inquire for more information from teachers of other subjects.

There were pupils, however, good scholastic ones in particular, who were not interested in their work, especially at the beginning: "But that's not mathematics!".

Lower school children, especially the 11-12 year olds, had much imagination and curiosity, so that ideas were numerous, much more so in fact than with the 16-18 year olds.

Mathematical ideas used or introduced

Sets. Cartesian product, relations, mappings, operations in P(E), partitions. Generally speaking, the set ideas were not often explicitly used by the pupils, even when they were known to them. This may be because this kind of work is often felt by the children to be cut off from any mathematical context. The opposite attitude was also seen. Some pupils wanted at all costs to find or introduce mathematical ideas.

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{D}$, $\mathbb{Q}$. Sums - ordering. Many questions were asked about comparisons, variations, and calculating differences.

Numerical calculations

Calculations were often repetitive, so in order to avoid boredom, pupils could use calculators, sometimes programmable ones.

When, in some teams, the repetitive numerical work had to be done again, to develop certain questions, results were obtained with a microcomputer set at the pupils' disposal (weather study in Bordeaux, when after the work on temperatures the pupils wanted to undertake a study of rainfall).

Graphical representation

"Experiments show that while most pupils have at some time of their education seen graphs, only rarely have they had to make them... Pupils often have seen and read graphs in natural science and geography classes, and in these subjects only a few classes have made graphs".

Pupils' experience in this area generally varied widely. Tests were carried out on 679 pupils to clarify this matter. The analyses of these tests point out:

"11 to 12 year olds seem to have some knowledge of the word graph. Nonetheless, this knowledge is often vague. More often than not, the work done involved labelling points...".

Even 13 year olds were rarely given the choice of representation, or even, the graduation of graphs.

The most common spontaneous graph was of the "temperature" kind, especially when time was involved.

Yet one is surprised by the number of pupils who, confronted by a problem with two variables, do not think of a graph to picture or to understand better the variations. This was also osberved among 16-18 year olds.

Difficulties in this kind of work, requiring research by the children and help from the teacher, were due to the choice of graduations (regular or not, no graduation), to the choice of origin, and to the data which sometimes were too large.

But these difficultites, the failures, the reports to the class, classmates' criticisms gradually led to more efficient graphs, richer in information. These personal contributions and successive corrections are very improving, on condition that the pupil is allowed to create these transformations himself. (See Appendix 2) "Pupils who draw a graph "to see" do not always feel the need to deduce consequences from it, or to interpret graphs which become an end in themselves". Many could profit by training in the use of graphs as a tool for analysing a question.

Mean - Mode - Median

11 to 12 year olds often know the word average without knowing what it means, in which case some pupils make its calculation an end in itself.

They often refute that an average can contain decimal places. "There can't be a car cut in two" (report; car registration).

It appears that for most pupils the mean is a number in the sequence, the most representative number, and not a real number the central characteristic of the sequence, i.e. a very abstract number in fact, produced by a formal definition.

This remains true of 12 to 15 year olds. Moreover, it seems necessary to familiarise the pupils with the average.

Mode and mean are sometimes mixed up. The median is quite frequently used. "What if we took the middle pupil" (report; sports results).

The mode and above all the median are much more concrete measures than the mean.

The mean certainly is difficult for children, e.g. the mean of the means, cases where the same number appears several times, or where null values are forgotten, or mixing up comparison of the means and comparison of the size of the populations.

Ratio - Frequency - Proportionality

These ideas occur spontaneously in statistical calculations whenever comparisons are required.

Yet, "reasoning on ratios and a fortiori the use of percentages is far from easy to 11-15 year-olds... as no doubt to many grown ups. This kind of reasoning is considered by Piaget and Inhelder as one of the characteristics of formal reasoning which begins at 11 to 12 years". Complete mastery of these ideas is achieved only slowly and at different personal rates.

"All the evidence shows that frequency and percentage are difficult ideas. They are assimilated only slowly, through contact with many rich situations where the pupil can ask himself questions. Current exercises, which put too much accent on the calculation, do not favour mastery of the idea". It appears that for 11 to 13 year olds calculation of frequencies as approximate decimal fractions is sufficient.

Spread

Certain themes required groups to compare variations and to deduce their consequences, e.g. the work on weather.

The reports say little on the small interest given by the children to spread and in particular to its representation by a number.

The average of a set of numbers seems natural enough, but bothering with the average deviation from the average and with a number to represent it require model building and abstraction beyond most 11 to 14 year olds. The idea of distance, moreover, is not properly mastered.

Some themes led to drawing scatter diagrams (report on heights, weights and sporting ability). Groups took account of the cloud of points associated with a sequence of two variables by drawing a line with as many points above it as below. (This idea was often suggested by the teacher). Pupils then drew two lines parallel to the first, enclosing the cloud of points and even rejecting some far flung ones. Nearly all the points lay in this band (92%, report on sporting ability).

The talk given to the class by this group had a large impact, a sign that pupils are favourable to some elementary modelling of real phenomena.

Correlation

Study of statistics in secondary schools can lead to consciousness of correlations between variables. Three of the themes used by the I.N.R.P. - I.R.E.M. research gave rise to studies correlation:

In the study "Heights-weights-armspread" pupils spontaneously sought out

"Are the tallest the heaviest?"

"Do height and armspread favour putting the shot?"

In the study of sporting ability, we found the question, put by the pupils,

"Is running related to the long jump? to the high jump?"

Finally, in the study of the port of Bordeaux, the variation with time of the number of boats coming in and their laden weight led pupils to study the relation between the two and to seek why the first was decreasing and the second increasing.

The methods used were graphical (comparison of several graphs, study of scatter diagrams) or tables of spreads.

Awakening of awareness to random phenomena and prediction

In these experiments, the childrens' interest was always very strong for any prediction which they were told they would be able to check.

"But extending in time the results obtained from a descriptive study was not always spontaneously suggested". At this age, we can try only to make the children sensitive to the problem, in view of a later introduction of probabilities.

Some mathematical tools (absolute value, average, frequency, spread, elementary combinatorial analysis,...) are not sufficiently mastered, besides which setting up even an elementary model requires formalisation which the children often find difficult.

They are sensitive to the gaming aspect of prediction, and ready to propose ideas. What was done was considered to be useful. Pupils, aided by the teacher, became aware that very often measurement of some quantity is associated with a random variable, that any future experiment involves risk, but that randomness generally ensures that the results lie in some interval [m, M] determined by analysis of the results already obtained.

Some examples are:

1) Sampling

Sampling was approached through prediction. One pupil proposed: "If I question only 6 out of 35 pupils, will I have everyone's opinion?".

Study of a sample and then of the entire population was strongly motivating.

2) Treatment of the data

"How many French made cars will be registered in 1975?"
(report on car registrations)

"Between what two numbers does the length of the class lie?"
(report on the measurement of a length)

"Can a pupil's speed be used to predict his long jump?"
(study of sporting ability)

Methods used called for the mean, the extent of a sequence and graphical methods (lines illustrating tendencies, scatter diagrams).

During the discussions in the syntheses, many pupils were aware that prediction is the use of a model (a very simple one of course at this stage), a formalisation using elements of the statistical study: mean, percentage, tendency lines, the interval [m, M].

CONCLUSION

We have discussed here research done by the I.N.R.P.-I.R.E.M. from 1973 to 1978, but other more recent work exists also, e.g. Clermont-Ferrand 1980-81, Rennes 1978-79, suggestions for adult teaching adults in Lilles, but which could be adapted to secondary school use, research on graphs (interdisciplinary, not just statistical) done by an I.N.R.P.-I.R.E.M. group.

All the conclusions corroborate what we have said above, particularly:

Secondary school children are not at all put off by large amounts of information. Indeed, it is a situation resembling closely their everyday experience.

They are strongly motivated by real data, which they may have collected themselves, or which deals with their social, economic or cultural environment.

They are ready to undertake long projects with many calculations, provided they are clear of the aims and if they have some equipment at their disposal as well: calculators or microcomputers.

They are able to and interested in working on random phenomena, in foreseeing possible errors, in predicting. It would thus be a shame to restrict them to purely deterministic teaching whereas they live in a world where statistics and probability play a large part. They will often have to make decisions. Our duty is to give them the tools to make these decisions conscious, calculated and involving a minimum of risk.

References

[1] Mathématiques du quotidien dans les collèges (Activités d'analyse des donnés dans le premier cycle). ed.: l'Institut Notional de recherche Pédagogique.

[2] Dossiers de la recherche I.N.R.P.-I.R.E.M.

Nos.	TITRE DU DOSSIER	EDITE PAR
1	Climatologie (Bordeaux, Brest,Strasbourg)	I.R.E.M. de Bordeaux
2	Climatologie (Brest, Châteauroux, Orléans, Strasbourg)	I.R.E.M. d'Orléans
3	Consommation d'eau à Grenoble	I.R.E.M. de Grenoble
4	Consommation d'électricité dans le Sud-Est de 1971 à 1975	I.R.E.M. de Lyon
5	Consommation d'électricité le 3e mercredi du mois de décembre dans le Sud-Est	I.R.E.M. de Lyon
6	Effectifs scolaires	I.N.R.P. Paris
7	Étude de performances sportives	I.R.E.M. de Bordeaux
8	Évolution de la population en Gironde et en Dordogne	I.R.E.M. de Bordeaux
9	Football	I.N.R.P. Paris
10	Immatriculation de voitures dans le département du Rhône	I.R.E.M. de Lyon
11	Immatriculation de voitures dans les pays de la C.E.E.	I.R.E.M. de Rennes
12	Loire	I.R.E.M. d'Orléans
13	Loisirs - Vacances	I.R.E.M. d'Orléans
14	Mesures de la longueur d'une salle	I.R.E.M. de Bordeaux
15	Port de Bordeaux	I.R.E.M. de Bordeaux
16	Port de Rouen	I.R.E.M. de Rouen
17	Résultats sportifs	I.R.E.M. de Grenoble
18	Sondages	I.R.E.M.Paris-Nord
19	Taille d'élèves de 5e d'un C.E.S. d'Orléans	I.R.E.M. d'Orléans
20	Taille et poids d'eleves de 5e d'un C.E.S. de Valence	I.R.E.M. de Grenoble
21	Taille, poids, envergure et résultats sportifs	I.R.E.M. de Rennes
22	Tailels et poids: croissance de la naissance à dix-huit ans	I.R.E.M. d'Orléans
23	Télévision	I.R.E.M. de Rouen

- Autres dossiers

[3] Etude d'un thème statistique en classe de 5ème-6ème : étude de la vente de journaux à St. Brieux (I.R.E.M. de Rennes)

[4] Dossiers statistiques (cotes et débits du barrage de Rochebut (Allier)) I.R.E.M. de Clermont

[5] Extraits du travail de recherche spontanée en 6ème. I.R.E.M. de Clermont

[6] Activités mathématiques. Centre Université Economie d'éducation permanente (C.U.F.E.P.) Lille

[7] Matemactives. I.R.E.M. de Lille

[8] Publications de deux groupes de stagiaires (en préparation). I.R.E.M. de Montpellier

TEACHING STATISTICS CONCEPTS THROUGH INSTRUCTIONAL GAMES

George W Bright, Northern Illinois University, John G Harvey, University of Wisconsin and Margariete M Wheeler, Northern Illinois University

Since 1976 the authors have been engaged in a systematic investigation of the cognitive effects of mathematics instructional games. From the outset the research has focused on identifying effective uses of games in order that recommendations for appropriate incorporation of games into mathematics instruction can be made.

A game is defined by seven criteria:

1. A game is freely engaged in.
2. A game is a challenge against a task or an opponent.
3. A game is governed by a definite set of rules. The rules describe all procedures for playing the game, including the goal(s) sought; in particular, the rules are structured so that once a player's turn comes to an end, the play just completed cannot be retracted or exchanged for another move.
4. Psychologically, a game is an arbitrary situation clearly delimited in time and space from real-life activity.
5. Socially, the events of the game situation are considered in and of themselves to be of minimal importance.
6. A game has a finite state-space (Nilsson, 1971). The exact states reached during play of the game are not known prior to the beginning of play.
7. A game ends after a finite number of moves within the state-space.

The first five of these criteria, based upon a discussion of play, games, and gaming situations (Inbar and Stoll, 1971), give formal, psychological, and sociological characteristics of games. The last two criteria were evolved in order to exclude other, related activities or structures such as play and puzzles.

An instructional game is a game for which a set of instructional objectives has been determined; these instructional objectives may be cognitive or affective and are determined by the teacher(s) planning instruction. The source of ideas for the statistical instructional games used in the research was Developing Mathematical Processes (DMP) (Romberg, Harvey, Moser and Montgomery, 1974, 1975, 1976). DMP is a rich source of mathematics instructional games for elementary school students. These games frequently also seem easily adaptable for other ages and for more sophisticated content.

The notion of cognitive effects derives from the work of Bloom (1956). In describing a taxonomy of educational objectives, Bloom (1956) states that the

cognitive domain "includes those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills" (p.7). Thus, cognitive effects are effects which describe changes or differences in the attainment of objectives in the cognitive domain. For the studies reported here, the cognitive effects are defined as the change in attainment of an objective in the cognitive domain due to a particular instructional treatment; that is, due to the playing of mathematics instructional games.

During 1979-82 a series of studies has been conducted to examine the effectiveness of mathematics instructional games in each of several combinations of instructional level and cognitive level. Three instructional levels have been identified: post-instructional level, co-instructional level, and pre-instructional level. An instructional activity is at the post-instructional level if it is used after instruction designed to provide students with mastery of the content; an activity is at the co-instructional level if it is used in conjunction with instruction designed to produce mastery; and, an activity is at the pre-instructional level if it is used prior to such instruction. Hence, instructional level is dependent on the prior instruction provided to a particular group of students.

Cognitive level is defined in terms of Bloom's 1956 taxonomy of six levels: knowledge, comprehension, application, analysis, synthesis, and evaluation. Of these six levels, evaluation essentially never appears in U.S.A. school mathematics curricula; synthesis appears only very rarely. Because the research conducted was limited by school curricula in the U.S.A. and was not designed to change those curricula, the evaluation and synthesis cognitive levels were omitted from consideration.

The remaining four cognitive levels and the three instructional levels yield twelve combinations; within each of these combinations it is possible to study the effectiveness of games as teaching tools. However, one of these combinations, namely pre-instructional/analysis, does not seem appropriate for the use of games. In this combination, students would be expected to use content at the analysis level prior to instruction designed to produce mastery of the content. While some students might succeed at this activity because of well-developed problem solving skills, most students should not be expected to use content at this high cognitive level prior to instruction. Hence, this combination was also omitted from further consideration.

For each of the remaining eleven combinations, studies at two different grade levels were conducted during 1979-82 with a variety of mathematics content. Of these twenty-two studies, four dealt with probability and statistics concepts with students aged 12 to 15 years. The distribution of these four studies among the eleven combinations was as follows:

1. pre-instructional/knowledge: one study
2. co-instructional/comprehension: one study
3. pre-instructional/application: two studies

Three of these studies are reported in this paper; the fourth, one of the pre-instructional/application studies, is in progress.

Several assumptions underlie these studies:

1. The mathematics content was substantial and would not be learned without instruction specifically designed to teach that content. Hence, control groups were not consistently used in these studies. Specifically, in the pre-instructional studies it was not necessary to have a control group in order to measure the effectiveness of a game. In the co-instructional study it was of course vital to have such a control group, since regular instruction was also occurring.
2. At both the pre-instructional and co-instructional levels, prior to the start of the study the students had not received instruction designed to produce mastery of the content.
3. Games could be found or modified which would provide instruction of the mathematics content at the appropriate cognitive level.

GENERAL RESEARCH DESIGN

The same basic design was used in all of the studies. Each study was based on the instructional day; that is, any school day during which instruction on the particular content of the game was given and/or the game was used. Games were played for eight instructional days. In pre-instructional studies these eight days were spread across four consecutive weeks; that is, the games were played twice per week for 20 minutes on each day of play. For the co-instructional study the eight days augmented standard instruction, so the game playing days were unevenly spaced over a three-week period. Again, on each of the eight days, the games were played for 20 mintues per day. In each study, during the week prior to the game playing pretests were given, and during the week after the end of game playing, posttests were given. Cognitive effects were measured by comparison of pretest and posttest scores in the pre-instructional studies and by analysis of covariance on the posttest scores with pretest scores as the covariate in the co-instructional study. As noted earlier, in the pre-instructional studies control groups were not used. In the co-instructional study, however, a sample of six students from each classroom was randomly selected, subject to a sex distribution approximately equal to that of the whole class, to serve as a control group. Control subjects received the regular instruction provided by the teacher, but they played

Mastermind (Invicta, 1972) instead of the particular statistics game being studied.

Also, in each study the four-card test of formal operations (Adi, Karplus and Lawson, 1980) is given as a pretest. Students aged 12 to 15 may be in transition between concrete and formal operations, and the particular level of skill that they have in dealing with material requiring logical reasoning may affect their success with that material. Hence, it was decided to measure formal operations so that content scores could be adjusted for formal operations skill.

PRE-INSTRUCTIONAL STUDIES

Two studies were conducted at the pre-instructional level. One dealt with knowledge of fair/unfair, and one dealt with the application of probabilities in decision making.

For the fair/unfair study eight pairs of games (Bright, Harvey and Wheeler, 1981) were used. These games were adapted from Developing Mathematical Processes. The content of these games is at the knowledge cognitive level. Subjects were 36 students in grade eight (about age 13).

A 28-item content test, consisting of two 14-item subscales, was given as both a pretest and a posttest. In each of the 28 items two games were described and students were asked to determine the fairer of the pair of games. For each item on one subscale students were given data summarizing 50 plays of each of the two games. (These data were generated by computer simulations of the games). The items on the other subscale were not accompanied by these simulated plays. A sample item is provided in Bright, Harvey and Wheeler (1982).

The means and standard deviations and the comparisons of the pretest and posttest scores are presented in Table 1. Analysis of variance of the subscale without simulation data given, with sex as the classification variable, yielded a significant sex-related effect ($F = 6.62$, $p < .015$) favoring the males. Analyses of covariance, with the four-card test and the pretests as covariates, yielded no significant effects.

TABLE 1

Pretest - Posttest Comparisons for Fair/Unfair Games at Grade 8

Scale	Pretest Mean (SD)	Posttest Mean (SD)	t	p
14 items with data given	6.50(2.27)	7.56(2.30)	3.83	.001
14 items without data given	6.44(1.78)	5.81(2.12)	-2.05	.048
Total test	12.94(3.22)	13.37(3.91)	0.93	.357

The games seemed to cause an improvement in performance at the knowledge cognitive level when data summarizing plays of the games were available. The marginal negative effect on the scores of the subscale without data is inconsistent with the results of a similar study at grade six (Bright, Harvey and Wheeler, 1982) and in the larger context of this project seems to be a spurious result. It is encouraging that in this setting students seemed able to learn to interpret data through the playing of the fair/unfair games.

It should be noted that in an earlier study (Bright, Harvey and Wheeler, 1980b) the same games were used with 82 grade seven students (about age 12). Although not identified at the time, the instructional level was pre-instructional. Two types of game-playing groups were formed: heterogeneous groups and homogeneous groups. Based on pretest scores students were grouped into four strata of equal size. Heterogeneous groups of size four were formed by randomly choosing one student from each stratum. Homogeneous groups of size three or four were formed by randomly choosing students from the same or adjacent strata. Thirteen heterogeneous and twelve homogeneous groups were created. Students who did not complete the posttests were dropped from the analyses. The pretest and the posttest were like the subscale reported earlier without any simulation data presented with the items.

The analyses of the data are presented in Table 2. Analysis of covariance with the pretest as the covariate was conducted on the posttest scores with the treatment and sex as classification variables. None of the F-values were significant at the .05 level; that is, neither treatment nor sex differentiated performance level.

TABLE 2

Pretest - Posttest Comparisons for Fair/Unfair Games at Grade 7

Grouping	N	Pretest Mean (SD)	Posttest Mean (SD)	t	p
Heterogeneous	49	16.0(5.1)	17.7(4.6)	3.01	.004
Homogeneous	33	15.7(4.9)	16.9(4.6)	2.82	.008

The second pre-instructional study dealt with application of probabilities in making decisions. Number Golf, the game used in this study, was based on a game described by Bright (1980). In this game a goal number is designated and players try to reach this goal number by adding or subtracting the sum of two numbers generated by dice rolls to their cumulative scores. The content of Number Golf is at the application cognitive level. Subjects were 74 students in grade seven (about age 12).

For this study three tests were used. One test was given only as a pretest; it was a 10-item test in which students were told that a pair of regular dice had been thrown 36 times; students were asked to predict how many times each of 10 given sums would appear. Students were given two points for each answer that matched exactly the predicted number of occurrences of each sum and one point for each answer that differed from the predicted number by only one. Two tests were given both as pretests and as posttests. The first was a 20-item test; in each item students were asked to choose the most likely of three sums generated by the toss of a pair of dice. Ten of these items dealt with standard dice; ten, with non-standard dice. The second was a 42-item test; for each item a goal number, a given score, and a number designated as the sum of a throw of two dice were given. Students were asked to choose the better of two possible moves, generated by adding or subtracting the given sum to the given score, in order to reach the goal number in as few moves as possible. Students could also indicate that the two moves were equally good. A sample item is given in Figure 1. Twenty-one of the items dealt with standard dice; 21, with non-standard dice.

FIGURE 1

Sample Test Item for Number Golf

Goal Number: 12
Given Score: 9
Dice Roll: 5

Which move would be better?

a. Subtract 5 from the Given Score to get a total of 4.
b. Add 5 to the Given Score to get a total of 14.
c. It doesn't matter which move is made.

The means, standard deviations, and comparisons of pretest and posttest scores are presented in Table 3. Analysis of covariance on the most likely sum/regular dice test with the four-card test, the number of occurrences test, and the pretest of most likely sum/regular dice as covariates and with sex as the classification variable yielded a significant sex effect ($F = 6.85$, $p < .011$) favoring the males. Analyses of covariance on the other three posttests with analogous sets of covariates and sex as the classification variable yielded no significant effects.

TABLE 3

Pretest - Posttest Comparisons for Number Golf Games at Grade 7

Scale	Pretest Mean (SD)	Posttest Mean (SD)	t	p
Number of occurrences[a]	6.94(3.95)	not given		
Most likely sum[b]				
regular dice	2.61(2.30)	4.42(2.12)	6.80	.000
non-standard dice	1.99(1.70)	3.38(2.09)	4.67	.000
Choosing move[c]				
regular dice	6.92(2.29)	7.27(2.94)	1.19	.240
non-standard dice	6.10(2.10)	6.00(2.50)	-0.31	.754

a. 10 items

b. 10 items in each subscale

c. 21 items in each subscale

The most likely sum test is not at the application cognitive level but is rather at the comprehension cognitive level for students who have not yet received instruction on probability. Hence, Number Golf was not effective at producing improvement in application cognitive level performance among the seventh graders who participated, though performance at a lower cognitive level did improve. A replication of the study with grade nine (about age 14) students has just been completed. Data should be analyzed by mid-summer 1982.

CO-INSTRUCTIONAL STUDY

One co-instructional study dealt with the concept, average. This study was at the comprehension cognitive level. The game, Average Hands, was taken from Developing Mathematical Processes. In the game students are dealt 12 cards from a deck of number cards. Players must arrange their cards into three subsets: one with three cards, one with four cards, and one with five cards. After all of the arrangements are completed, players compare the averages for each size of subset. The player with the highest average scores 3, 4 or 5 points, depending on the size of the subsets being compared. Variations of this game were developed by changing the scoring rule to reward the smallest average in a comparison and by varying the numbers occurring in the number decks. Subjects were 102 grade seven students (about age 12) with 80 students in the experimental (game playing) group and 22 in the control group.

A variety of tests were used. Five speed tests were given only as pretests: (a) a test of all 100 multiplications basic facts, (b) a test of 12 addition items, with four to eight two or three-digit addends in each item, (c) a test of 18

long division items, with one and two-digit divisors, (d) a test of 8 rounding items, with each item being a mixed number to be rounded to the nearest whole number, and (e) a test of 7 multiple-choice estimation of number items, each of which was a "word" problem with five approximate answers given. Two tests were given as pretests and posttests. One was a test of 20 averaging items, each of which was a set of four to eight one-, two-, or three-digit numbers for which an average was to be computed. The other test contained 20 items; for each item a set of nine numbers, in the range 1 to 99, was given, and students were asked to choose two disjoint subsets of three or more numbers each, so that the averages of the subsets were within specified ranges. A sample item is provided in Figure 2. The sets of nine numbers were randomly generated and the ranges of the averages sought were chosen so that each item had one or more solutions.

FIGURE 2

Sample Test Item for Average Hands

In each of the questions which follows you are given a list of nine numbers. Using the numbers in each list, you should make two sets of three or more numbers so that the average of the numbers in Set One is between the numbers given for that set, and the average of the numbers in Set Two is between the numbers given for that set. Do not use any number in the list more than once. You do not have to use all the numbers in the list. Usually there will be several correct ways to choose the two sets of numbers.

Numbers: 1,2,3,4,5,6,7,8,9

Set One: _____ average between 2 and 4

Set Two: _____ average between 6 and 8

The means and standard deviations of the tests are given in Table 4. Analysis of co-variance on the posttest of averaging, with the four-card test, the addition test, the division test, and the pretest of averaging as co-variates and with the treatment group as the classification variable yielded a treatment effect ($F = 5.1$, $p < .030$) favoring the experimental group. Analysis of covariance on the posttest of averaging with the same four co-variates and with both treatment group and sex as classification variables yielded a treatment effect ($F = 5.3$, $p < .025$) but did not yield either a significant sex effect or a significant treatment by sex interaction. Analysis of covariance on the posttest of choosing sets, with the four-card test, the rounding test, the estimation test, and the pretest of choosing sets as co-variates yielded no significant effects for treatment, sex, or the interaction.

TABLE 4

Means (Standard Deviations) for Average Hands Games at Grade 7

Scale	Experimental Group		Control Group	
	Pretest Mean (SD)	Posttest Mean (SD)	Pretest Mean (SD)	Posttest Mean (SD)
multiplication	99.10 (3.1)		99.62 (0.8)	
addition	10.60 (1.7)		10.71 (1.3)	
division	16.58 (1.9)		16.71 (1.3)	
rounding	6.21 (2.2)		5.62 (2.6)	
estimation	3.95 (1.4)		3.24 (1.6)	
averaging	7.03 (3.6)	9.75 (3.7)	6.36 (3.1)	8.27 (3.0)
choosing sets	6.58 (4.5)	12.46 (5.7)	5.46 (3.9)	11.36 (6.2)

The game seemed to enhance regular instruction on computing averages. This behavior, however, is marginally at the comprehension cognitive level. A reasonable argument could be made that for age 12 students this behavior is only at the knowledge level, since the algorithm for computing averages involves subskills that are presumably well-known to these students. Separating a set of numbers into subsets with averages within specified bounds is certainly at the comprehension cognitive level, but the game did not seem to cause any improvement in the performance of that skill beyond that provided by the regular instruction. It is also quite likely that the teacher did not expect mastery of that skill.

CONCLUSIONS

The three studies have the common thread that the students involved had not previously been exposed to instruction designed to produce mastery of the particular statistical concept or skill needed to play the particular game. All of the games seemed effective at improving performance, either alone or in addition to the regular classroom instruction, at the knowledge cognitive level. The games did not seem effective at the comprehension or application cognitive levels. In part, this may be explained by the fact that students started the experimental period with essentially no knowledge, or at least very little formal knowledge, about the content. Hence, students may have had to build all of the cognitive structures required to absorb the content within the relative brief treatment period of approximately three hours spread across eight instructional days. This may be an unreasonable expectation.

The implication of this observation for teaching is that one game, or one variety of games, probably should not be expected to carry the brunt of initial teaching of concepts like those encountered in probability and statistics. Mathematics

instructional games typically focus on a limited range of concepts or skills within a limited range of cognitive levels. Initial teaching of a mathematics concept may require broad coverage of many aspects of that concept. One type of game should not be expected to provide such broad coverage. Also, for probability and statistics in particular, basic concepts may be contaminated by students' inaccurate perceptions of real-world experiences. It seems important to help students not only identify their inaccurate perceptions but also give up their confidence in those perceptions. Perhaps highlighting the important attributes of concepts prior to using games, would effectively assist understanding of probability and statistics concepts.

Previous research (Bright, Harvey and Wheeler, 1979, 1980a) had indicated that after a concept has been taught to mastery, games are effective materials for the practice and reinforcement of performance. Performance at comprehension and application cognitive levels may depend heavily on well-developed skills at the knowledge cognitive level. For games used at the pre-instructional level this background of knowledge cognitive level performance may not be strong enough to support the higher level work. It may be unrealistic to expect an eight-day treatment to provide enough time to build both the lower level skills and the higher level skills.

Future research should examine the kinds of skills and concepts that particular students develop while playing a game at the pre-instructional level as well as the different effects that may accrue with more rounds of play of such a game. Dennis, Muiznieks and Stewart (1979) indicated in their examination of the games literature that no studies examined the effects of varying the number of rounds of play of a game. Especially in pre-instructional settings where development of skill would be expected to be slow, such studies could prove of great value in the development of guidelines for the use of games.

References

[1] Adi, H., Karplus, R. and Lawson, A.E. (1980) Conditional logic abilities on the four-card problem: Assessment of behavioral and reasoning performances. Educational Studies in Mathematics, 11, 479-496.

[2] Bloom, B.S. (Ed) (1956) Taxonomy of educational objectives. New York: McKay.

[3] Bright, G.W. (1980) Game moves as they relate to strategy and knowledge. Journal of Experimental Education, 48, 204-209.

[4] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1979) Using games to retrain skills with basic multiplication facts. Journal for Research in Mathematics Education, 10, 103-110.

[5] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1980a) Using games to maintain multiplication basic facts. Journal for Research in Mathematics Education, 11, 379-385

[6] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1980b) Achievement grouping with mathematics concepts and skills games. Journal of Educational Research, 73, 265-269.

[7] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1981) Fair games, unfair games. In: A. Schulte and J. Smart (Eds) Teaching statistics and probability, 1981 yearbook of National Council of Teachers of Mathematics. Reston, VA: NCTM, 49-59.

[8] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1982) Using games to teach some probability concepts. Paper presented at the First International Conference on the Teaching of Statistics, Sheffield, England.

[9] Dennis, J.R., Muiznieks, V.J. and Stewart, J.T. (1979) Instructional games and the computer-using teacher. Urbana, IL: University of Illinois (ERIC Document Reproduction Service No. ED 183 189).

[10] Inbar, M. and Stoll, C.S. (1970) Games and learning, Interchange, 1(2), 53-61.

[11] Invicta Plastics Ltd. (1972) Mastermind - Regular, Oadby, Leicester, England: The author.

[12] Nilsson, N. (1971) Problem solving methods in artificial intelligence. New York: McGraw-Hill.

[13] Romberg, T.A., Harvey, J.G., Moser, J.M. and Montgomery, M.E. (1974, 1975, 1976) Developing mathematical processes. Chicago: Rand McNally.

A PRACTICAL APPROACH TO STATISTICS

R Kapadia, Polytechnic of the South Bank

INTRODUCTION

Open any newspaper or magazine and one is likely to find some statistical information. Probabilistic words like chance, odds or likely are part of everyday language. Data is collected pertaining to many and various aspects of life. Decisions made by governments and industries rely heavily on published figures. Many subject areas have become increasingly quantitative. It is clear that pupils are growing up into a world becoming more dominated by statistical ideas and it is up to education to take up the challenge by providing suitable courses for all children. This paper considers the needs of ordinary children who will complete their education at age 16 or below. An attainable scheme of work, which takes into account the statistical awareness required by adults both in society and in their employment, is presented.

The first part provides a perspective showing how the influence of the subject has grown. From this the main aims are advocated and a specimen syllabus linked to applications is outlined. Suitable teaching strategies are described, illustrated by a specific example. References to suitable teaching materials are also given. Throughout, the international context is borne in mind, for statistics is an essential part in any education system.

HISTORICAL PERSPECTIVE

In a sense statistics began when man started counting and keeping records. This is a rather naive view in that, though records have been kept for many centuries, they were rarely interpreted, analysed, or more critically, used to make inferences. This had to await the emergence of probability, which was a slow process.

For man held a deterministic view of the world over a long period of time. The search for pattern and order was guided by a casual point of view. Science used empirical methods which were firmly based on cause and effect. If an experiment was repeated in the same conditions, it was assumed that the same results would ensue. This deterministic viewpoint is now being superseded especially at the microscopic level, by theories based on probability.

Probability, as a mathematical discipline is thought to have originated from a (characteristically) chance communication about gambling by Chevalier de Mere to Pascal. The study of inferential statistics began later through the work of Graunt and Petty in the seventeenth century on medical records and

life insurance.

In teaching probability and statistics, we must not forget the long time the subjects took to develop and the radical change of philosophy they entail. For, according to Kendall (1970) "the late emergence of probability calculus was due to some more fundamental factor. The very notion of chance itself, the idea of natural law, the possibility that a proposition may be true and false, in fixed relative proportions, all such concepts are nowadays so much part of our common routines of thought that perhaps we forget that they were not so to our ancestors. It is in basic attitudes towards the phenomenal world, in religious and moral teaching and barriers, that I incline to seek for an explanation of the delay".

In a perceptive analysis presented at a colloquium on the history of statistics Hacking (1981) asserts that, though it seems only to provide information, statistics is part of the technology of power in a modern state. For social facts have simply become facts which are statistical in character so that 'many of the modern categories by which we think of people and their activities were put in their place by an attempt to collect numerical data'.

It was during the nineteenth century it became accepted that 'real chance exists and is part of the underlying structure of the world'. Hacking cites the following events which led to this:

(a) the theory of probability was developed,
(b) the avalanche of numbers revealed an astonishing regularity in the statistics of crime, suicide, epidemics and mathematicians attempted to analyse such phenomena,
(c) the normal curve was postulated as a universal model for human, social and biological traits,
(d) social reformers thought that one could recognise the 'boundary conditions' under which a population was governed by statistical laws,
(e) statistical mechanics and thermodynamics heralded the intrusion of chance into physics,
(f) Darwin and Galton introduced probabilistic notions into biology.

Hacking goes on to assert that 'almost no domain of human enquiry is left untouched by the events that I call the avalanche of numbers, the erosion of determinism and the taming of chance'. Paradoxically, in a quantitative universe, exactness has become impossible.

Thus, it is clear that statistics has become a vitally important subject. It affects the way we think about ourselves by determining classifications. Even birth and death, the most individual of events, are governed and described

by probabilistic laws. Thus it is vital (in the literal sense too!) that children be taught statistics in school.

MODERN APPLICATIONS

One could write many pages in describing the many and varied recent applications of statistics; to give some idea of the scope just a few examples are described, according to the classification in Tanur (1972). Four major subject areas, where statistical reasoning is being increasingly applied, are considered; these are medicine, science, politics and everyday life.

Statistics has played a vital part in the progress of medical science. Substantial tests preceded the introduction of the Salk poliomyelitis vaccine, which has led to the virtual eradication of this terrible disease. A variety of drugs have been produced to combat a wide range of illnesses. Their development has required careful statistical screening. Probability has played a crucial role in many sciences. Quantum physics is firmly based on probabilistic notions, as enshrined in the famous Heisenberg Uncertainty Principle. Even life is now thought to have been caused by a chance mutation.

One finds evidence of statistics in politics constantly. Opinion polls are the popularly known manifestation yet published figures and trend analysis are used to help in making more informed decisions on a wide variety of issues.

This overlaps with the many applications of statistics in the social world - communication, leisure, national data to name just a few areas. Thus statistics affects each citizen. All major countries conduct a national census to build up a basic picture; it is important that pupils understand the rationale behind this massive exercise in data collection and the safeguards enforced to ensure confidentiality. Prices are continually rising yet people do not understand how the retail price index works. [A recent survey, published in the Cockcroft Report (1982) showed that over 50% of adults do not understand the rate of inflation]. Bar charts, pie diagrams and pictograms are often used in literature; there are also striking examples of abuses which children need to be aware of. Manipulation of figures can make a substantial difference in pay negotiations.

It is clear that statistics plays a vital role in everyday life for everyone, both at home and at work. This is the fundamental reason why the subject needs to be taught to all pupils.

SCHEME OF WORK

Having made the case for teaching statistics to all pupils one must consider the aims of a course to meet the general needs. Two aims are proposed,

following the work of the Schools Council Project on Statistical Education. The author was one of the team who undertook the research for this project. The aims represent the two aspects of statistics, which is defined to be the science of numerical data. Pupils need to become aware of range of applications to which statistical methods can be applied.

(1) Children should become aware of, and appreciate the role of statistics in society. That is, they should know about the many and varied fields in which statistical ideas are used, including the place of statistical thinking in other academic subjects.

(2) Children should become aware of, and appreciate, the scope of statistics. That is, they should know the sort of questions that an intelligent use of statistics can answer, and understand the power and limitations of statistical thought.

This attitude to statistics teaching implies that pupils should deal with statistical ideas in a variety of practical situations in order to experience the ubiquity of the subject. These examples need to be chosen from different parts of the school curriculum and also from important areas of everyday life or future work. Thus the two aims are pursued concurrently so that children are introduced to the techniques and methods of statistics via realistic examples. Clearly this is a very different philosophy to the traditional, rather arid approach where statistical techniques are divorced from real data. The modern approach of teaching within context fits much more naturally with the movement towards the introduction of problem-solving in the classroom.

The statistical techniques to be studied can be divided into five categories: data collection; data tabulation and representation; data reduction; probability; inference (note the emphasis on data). These categories need to be studied in the following broad areas of application: mathematics; humanities; sciences; social sciences; other. The five areas have been deliberately chosen to fit in, as far as possible with the normal school curriculum. Nevertheless the last area is by no means the least important: it includes important applications of statistics such as medicine, management, quality control and everyday life which do not fit the neat compartments of the school curriculum.

Thus teaching materials need to be devised according to a two way classification, whereby units of work develop concepts within specific areas.

For example one could choose an application from the sciences, say the plight of blue whales. This can be used to introduce the concept of sampling and show some of the problems inherent in obtaining data. Appropriate measures of location can be applied, perhaps even a probability distribution. One can then interpret the results and make relevant inferences.

This example shows how the approach described above can be achieved in practice. It starts with a problem (How many whales are there?) and develops the necessary statistical techniques (sampling, simple inference) to help to solve it (as far as possible).

Clearly care needs to be taken in the choice of the problem and the development of statistical techniques. One needs to consider the background of the children both in terms of their cultural outlook and statistical knowledge. Some examples may be appropriate in certain countries but not in others. A selection of materials adopting this approach needs to be developed. A start has been made [7].

To ensure the breadth and scope of statistics is properly covered one needs a syllabus and list of core applications. This is presented below but the above proviso needs to be borne in mind. In some countries one would need to make appropriate adaptations. Also what is presented below is intended to be taught to all pupils, particularly those who will complete their education at age sixteen or earlier. The needs of those who will continue their study later will be much wider [8 p.71].

Syllabus

This covers the broad areas: the depth of treatment would vary according to the pupils and choice of applications.

(a) Uses of a census
(b) Types of data
(c) Sampling
(d) Obtaining data
(e) Tabulation of data
(f) Pictorial representation of data
(g) Measures of location
(h) Measures of dispersion
(i) Index numbers
(j) Correlation
(k) Intuitive understanding of probability
(l) Reading and interpreting tables or diagrams
(m) Comparing data
(n) Detecting trends

A more detailed list of objectives may be found elsewhere [8, p.68].

Core Applications

Below is a list of the important applications of statistics which all children should meet.

(a) Census of Population
(b) Retail Price Index
(c) Opinion polls
(d) Health - smoking, vaccination
(e) Population growth
(f) Measures used in society e.g. housing, crime, leisure
(g) Conservation and pollution
(h) Advertising
(i) Errors and accuracy
(j) Weather prediction

It is important to remember that the syllabus and applications should be considered jointly. A technique is important in statistics in terms of its applications. Thus each of the areas listed in the syllabus will be met within one of the applications; indeed some areas will occur in several applications. This will clearly be beneficial for pupils will see statistics as an interesting, live subject relevant in a variety of disciplines. Clearly one has to be careful not to introduce too many new ideas simultaneously; reinforcement is also necessary. This interlinking of techniques and applications will make statistics an exciting and meaningful area of study with direct relevance to both everyday life and future work.

TEACHING APPROACHES

Having presented a framework one must consider what teaching approaches are suitable. Three basic principles are important [8 p.41].

1. Concepts and techniques will be developed in a practical context.
2. Techniques will not necessarily be completely developed at the time they first arise. Many of the ideas first introduced in the early years will continue to arise in later years.
3. A full theoretical justification of all the topics is neither necessary nor desirable. Some items will only be touched upon in particular problems; other concepts will be covered only through experiments and will not be justified theoretically.

A mixture of teaching strategies needs to be used: exposition, discussion, investigation, project work, simulation, case studies and practical work. The starting point should always be a problem as statistics is used to solve problems. After that one should proceed using appropriate strategies to help children find their own solutions to the problem.

Exposition can play a role in this, indeed this has always been a fundamental ingredient of work in the classroom. But there needs to be a dialogue whereby teachers should take account of and respond to the answers which pupils give, even if these answers are wrong or unexpected. Indeed statistics is the subject of uncertainty, making judgements in this situation is crucial. This can lead to worthwhile discussion and increased awareness for both teacher and pupil.

It is also useful for pupils to discuss ideas themselves, perhaps in small groups. This can be particularly valuable in a brain-storming session if a problem seems particularly intractable. Discussion can also clear up misunderstandings or help to clarify ideas.

The idea of investigation is fundamental. Investigations can be quite substantial and lead into individual project work, but this need not be the case. Indeed, perhaps more commonly, an investigation need be neither lengthy nor difficult, perhaps arising in response to a pupils' question such as "what would happen if...". Very often the question may be resolved quite quickly, sometimes it will need to be discussed at another time. The essential requirement is that pupils should be encouraged to think in this way and that the teacher should utilise such opportunities. There should be willingness even to follow some false trails.

Simulations can play a key role in presenting relatively difficult ideas in an intuitive way. Sometimes a full theoretical analysis may be rather difficult or even impossible but a simulation can throw light on the underlying situation and help pupils build up statistical intuition. This is particularly important in tackling probabilistic notions some of which can be quite tricky. The key idea is to simulate the uncertainty inherent in a particular situation and hence develop a deeper understanding. The advent of microcomputers will revolutionise the place of simulations as a large number of trials can be carried out effectively, reducing the drudgery. More details about simulations and Monte Carlo methods can be found in the NCTM yearbook [5].

Practical work is essential in putting over the main ideas behind data collection. This is a far wider and diverse area than in the traditional approach which pays only lip service to data collection before moving into often meaningless calculations. For, if one ignores data, there is little purpose in performing calculations of the mean, median or standard deviation. Indeed it is only by looking at data carefully that one can decide which measures are important. One also needs to be aware of the problems inherent in data collection - inaccuracies, omissions, bias. Children can only learn about this by trying to collect data in practical situations. This can be done by a small census, a questionnaire or a survey; sometimes one might use published data but one needs to be aware of appropriate sources. Occasionally one might only be able to take

a sample, in which case care over inferences is necessary. In fact sampling is widely used in society and children need to be aware of the power and limitations of such techniques, preferably by taking samples themselves. Thus practical work lies at the heart of statistics, particularly in building up a firm background.

To facilitate effective teaching some equipment and resources are necessary but the essentials are relatively inexpensive [Holmes, 1981]. Micro-computers are likely to play an important role in statistics teaching but they would probably be bought as a resource for the whole school; besides their price is falling rapidly.

To summarize, it is important that all these techniques are used in the teaching of statistics. A problem-solving approach is essential as a starting point after which a variety of strategies should be used as appropriate. An example is presented below to give an idea of what is envisaged.

Retail Price Index

The retail price index is one of the major measures of the national economy now that inflation seems here to stay. This affects everyone yet few people understand how it works.

A suitable starting point for children is their weekly pocket-money. One might ask whether pocket-money this year has the same purchasing power as last year. This will lead into an interesting discussion of price rises of items such as sweets, crisps or drinks. Pupils can then try to construct their own individual little index using just four items by comparing present prices with those of the previous year. There is scope here for discussing accuracy of data and choice of items for the index. A pie chart or some other form of representation can be drawn. One can then go on to the idea of weights; clearly one buys more of certain items than others. Each pupil can choose his/her own items and weights to calculate a personal index number. This allows comparisons to be made within the class and helps to emphasize that the retail price index cannot reflect the effect of price rises on any individual but is a combined average. An appropriate average can be calculated from the class results.

After this it is appropriate to look at grocery prices in more detail, perhaps including a survey in local shops. Each pupil can find out what his/her family buys each week to build up a sort of Family Expenditure Survey. Thus one can work out a food index in a way which mirrors what is actually done in practice.

Finally one can broaden the index to include other commodities. The ideas of a clothing index and the forms that could take for families of different sizes

can be considered. This leads naturally into an explanation of how the Retail Price Index is actually worked out, using the relevant published figures.

This teaching outline shows how various strategies can be used within a problem-solving framework. There are opportunities for exposition by the teacher, discussion between pupils, simulation, investigation and practical work. Techniques covered include problems of obtaining data, data tabulation and representation, applying measures of location and dispersion, index numbers, and comparing data. Clearly pupils will not become proficient in all these areas simultaneously but, by using this approach, they will meet these techniques in an appropriate context and, hopefully, learn how to apply the techniques in other situations.

Extensive evaluation of this approach has been conducted by the author as part of the research on the Schools Council Project on Statistical Education. These ideas were developed in a set of 27 units of work which have been carefully tested in a wide range of schools with teachers from varying backgrounds, teaching classes of different abilities. In general the materials produced were successfully taught to pupils of most abilities. More details can be found in Kapadia (1980). The materials have now been published [7].

CONCLUSIONS

Statistics is a hard subject to teach properly. By its very nature it is uncertain. In conducting experiments or surveys and making inferences one cannot predict what might happen. The statistics teacher needs to be particularly alert. Statistics also involves open-ended questions which have no 'right' answer. This is why a problem-solving approach linking techniques to applications is vital.

Probability also poses problems, not least because of the counter-intuitive nature of some of its results (do you believe the birthday paradox?). Another aspect which is hard to put across is the way that probability acts as a model in many situations. The fundamental point of modelling a supposedly unique event, like death, with probability is a difficult idea to accept. One knows it works (else insurance companies would have become bankrupt long ago) yet there is still a feeling of unease. It is by exposure to such ideas through simulation and investigation that children will learn to come to terms with these concepts. To facilitate understanding one should start with a frequentist approach estimating probabilities from experimental data; this emphasizes the link with statistics. It is also important to develop a subjective approach as well as using a priori considerations. These ideas are only presented briefly as this paper is mainly concerned with statistics. Yet the theoretical basis afforded by probability must not be ignored.

This paper has presented the outline of a course suitable to provide all children with a necessary background in statistics. The historical perspective and modern applications provide the rationale. The aims of the course must be to link objectives and applications within a problem-solving framework. A variety of teaching approaches is necessary as illustrated by the example of the retail price index. The general overall aim is to achieve a level of statistical numeracy so that all pupils learn to appreciate the powers and limitations of statistics.

References

[1] Cockcroft, W. (1982) Mathematics Counts, HMSO.

[2] Hacking, I. (1981) How should we do the history of statistics? I. & C. No.8, pp.15-26.

[3] Holmes, P. et al (1981) Statistics in Schools 11-16: A Review, Methuen.

[4] Kapadia, R. (1980) Developments in statistical education. Educ.Stud.Maths., 11, pp.443-461.

[5] Shulte, A. (ed) (1981) Teaching Statistics and Probability, N.C.T.M.

[6] Pearson, E. and Kendall, M. (eds) (1970) Studies in the history of statistics and probability, Griffin.

[7] School Council Project on Statistical Education (1981) Statistics in Your World (27 pupil booklets, 27 teacher booklets), Foulsham.

[8] School Council Project on Statistical Education (1981) Teaching Statistics 11-16, Foulsham.

[9] Tanur, J. (ed) (1972) Statistics: A guide to the unknown, Holden-Day.

ON THE TEACHING OF ORDER STATISTICS

K Bognár, Eötvös Loránd University

To deal with order statistics is very fruitful from several aspects. Besides getting the pupils acquainted with different statistical notions (median, range, etc.), we have an opportunity to show how to distinguish between independence and dependence and the different levels of dependence (correlation). The variety of combinatorial methods used in connection with order statistics is also an argument in favour of teaching this topic.

The program to be presented is for pupils of 13-15 years who learned some stochastics (relative frequency, probability, mean etc.) within the new mathematical curriculum.

The problems were usually connected with games, which has several advantages. The collection of data (which is essential in the teaching of statistics) is not so monotonous if an experiment is performed for deciding who is the winner, whose tip is good. The games are generally such that pupils may choose from different strategies and this gives them some means of "defending" themselves against chance.

The games were played in the classroom with coins, balls, dice and pearls but, later on, pupils used tables of random numbers. By the aid of the usual tables of random digits they prepared tables of 0-1 sequences in which the relative frequency of zero is not $\frac{1}{2}$.

GAMES WITH ZERO-ONE SEQUENCES

One may disagree with treating 0-1 sequences in connection with order statistics, since many facts become trivial in this case but - in our opinion - 0-1 sequences help to understand notions just because of their simple character.

Instead of tossing a fair coin (or unfair coin), we draw (with replacement) from a box containing a "zero" ball and a "one" ball (or some "zero" balls and some "one" balls). Thus we generate 0-1 sequences, say, of length n. The probability of drawing a zero will be denoted by p_o.

The digits generated will first be written as they come out (unordered sample) and then they will be ordered nondecreasingly (ordered sample).

GAME 1:

n will be fixed as 3, 4, 5; $p_o = 0.5$. The player writes two 0-1 sequences of length n: a tip for the unordered sample and one for the ordered sample. Under each digit in their tips they write a nonnegative integer (stake), the sum of these integers in each tip should be 2n. Also they pay 2n toplak to the bank. After that a 0-1 sequence is generated, and for each hit the players get the amount appearing under the corresponding digit in their tip.

The game can be played either by the whole class together or in pairs; in the latter case much more random sequences will be obtained. Pupils prepare the following table:

	unordered sample								ordered sample						
tip	1	1	1	1	1			pure gain −1	0	0	1	1	1		pure gain 5
stakes	2	2	2	2	2				10	0	0	0	0		
generated sequence	1	0	0	1	0			number of hits 2	0	0	0	1	1		number of hits 4
hits	X			X					X	X		X	X		
tip...															

Questions to be discussed:

1. Are there preferable strategies in case of unordered and in case of ordered samples?
2. Which n ($3 \leq n \leq 10$) do you prefer?
3. Which p_o do you prefer?

GAME 2a:

First let $p_o = 0.5$. Before generating 0-1 sequences of length 5 pupils write (secretly) their tip concerning the last zero in the ordered sample. (If they think there will be no zero in the sequence, they write O). Those whose tip is correct get 5 toplak; those who err by one get 4, those who were mistaken by 2 get 1 toplak; the others get nothing.

GAME 2b:

The same game, but now the 0-1 sequences are generated with the aid of a box with 2 "zeros" and 1 "one" ($p_o = \frac{2}{3}$).

To the question "How many toplak are you ready to give for a game?", most of the answers were 3 (which is not bad, since the game is profitable up to 3.79 when the player's tip is 3).

The pupils determined the distribution of $X_i^{(5)}$ (where $X_i^{(n)}$ denotes the i-th digit of the ordered sample of length n) and compared it with the empirical distribution (see Table 1, $p_o = \frac{2}{3}$).

i	$P(X_i^{(5)} = O)$	Empirical (200 sequences)
1	0.996	0.99
2	0.955	0.96
3	0.790	0.77
4	0.461	0.46
5	0.132	0.12

TABLE 1

Afterwards, it was easy to discuss the fact that it is a bit easier to find out the ordered sample in case $p_o = \frac{2}{3}$ than in case $p_o = \frac{1}{2}$.

GAME 2c:

A special case of the Polya model is considered. At the beginning we have two balls in a box, numbered 0 and 1 again. A ball is drawn at random from the box, its number is written down, then (before the next draw) it is replaced and also an additional ball of the same type is put into the box. This process is repeated five times. Thus we get as a result a Polya-type 0-1 sequence. The tip concerns again the last zero in the ordered sample.

Pupils can see after a short time that these 0-1 sequences behave differently from the sequences generated by a simple replacement. The empirical distribution of the last zero (see Table 2) was surprising to them.

i	0	1	2	3	4	5
Empirical (150 sequences)	0.15	0.21	0.15	0.17	0.15	0.17

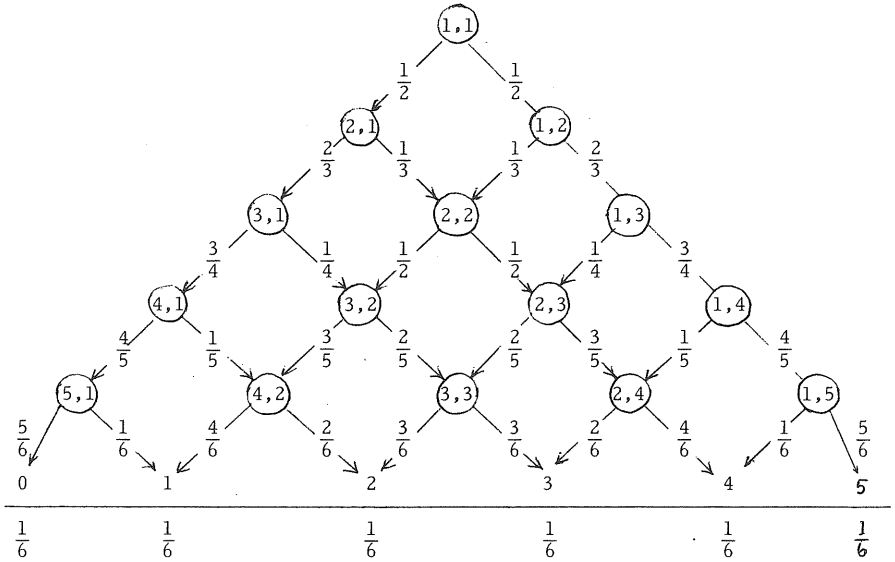
TABLE 2

Observing always say, the second digits in the unordered samples, pupils could see that nearly half of them were zero (as if they were drawn with replacement from a box with $p_o = \frac{1}{2}$). Further, pupils could see that the connections between first and second; first and third; first and fifth digits in the ordered samples became looser and looser.

GAME 3:

Player A tells an integer n, this will be the length of the 0-1 sequence to be generated ($p_o = \frac{1}{2}$). Then B tells k ($1 \leq k \leq n$) which means that A should find out the k-th digit of the ordered sample. A gets (n-k) toplak from B if A's tip is good, and A pays 2(n-k) toplak (or k(n-k) toplak in another version) in case of a mistake. The average gains of A (by different strategies) can be seen in Table 3 (or in Table 4 in case k(n-k) toplak). The pupils did not compute these, but they found out good strategies. Pupils were asked to draw a scheme and indicate their gains by jetons of two colours (positive and negative), from which everybody could see which strategies were good for A and B. They soon discovered the difference between n=2s and n=2s+1.

This game was played also with Polya-type sequences. Pupils evaluated the exact distribution of the number of zeros with the help of the following graph (zero to the left, one to the right).



GAME 3a:

A version of Game 3 is the following. Player A tells an integer n, then B tells i, k ($1 \leq i < k \leq n$), which means that A should find out the i-th and k-th digits of the ordered sample of a random 0-1 sequence of length n. A gets 3 toplak from B for a correct tip and pays to B 2 toplak in case of any mistake.

To find out how to play well, the pupils evaluated the joint distribution of the i-th and k-th digit of the ordered sample.

B \ A	2	3	4	5	6	7	8	9	10
1	0.3	1.3	2.4	3.6	4.8	5.9	6.9	7.5	9.0
2	0	-0.5	0.1	1.3	2.7	4.1	5.4	6.6	7.7
3		0	0.1	-1.0	-0.1	1.3	2.8	4.4	5.9
4			0	0.4	-0.1	-1.5	-0.4	1.2	2.9
5				0	0.7	0.6	-0.3	-2.0	-0.7
6					0	0.8	1.1	0.7	-0.5
7						0	0.9	1.5	1.5
8							0	0.9	1.7
9								0	1.0
10									0

TABLE 3

B \ A	2	3	4	5	6	7	8	9	10
1	0.5	1.5	2.6	3.8	4.8	5.9	6.9	8.0	9.0
2	0	-0.5	0.1	1.3	2.7	4.1	5.4	6.6	7.7
3		0	-0.3	-2.0	-1.1	0.4	2.1	3.8	5.5
4			0	0.1	-1.4	-4.5	-3.3	-1.3	0.8
5				0	0.3	-0.7	-3.5	-8.0	-6.3
6					0	0.6	-0.02	-2.3	-6.6
7						0	0.7	0.6	-1.1
8							0	0.8	1.0
9								0	0.9
10									0

TABLE 4

GAME 4:

Each of two players A and B generates a 0-1 sequence ($p_o = 0.5$) until at least one of them obtains a "run" 11 (i.e. the digit 1 appears twice one after the other).

1. If both A and B obtained their first run within 10 trials, namely k_A trials needed for A and k_B for B, then

a) in case $k_A = k_B$, the game is finished and nobody pays

b) in case $k_A < k_B$, either B pays $k_B - k_A$ toplak to A, or B begins to generate a new sequence (in which case A may also choose to begin a new sequence) and the rules are applied for the last sequence of each player.

2. If A obtained while B did not obtain a "run" within 10 trials, then B either stops and pays $10-k_A$ to A or generates a new sequence until a run appears and pays the length of this new sequence to A.

3. If neither A nor B obtained a run within 10 trials, each of them should generate a new sequence and they pay according to 1.

Denote by P_i the probability that the number of drawings needed until a run 11 appears is just i. Table 5 contains these probabilities for a few i's (evaluated by the pupils).

i	2	3	4	5	6	7	8	9	10
P_i	0.250	0.125	0.125	0.094	0.078	0.063	0.051	0.041	0.03

TABLE 5

GAMES WITH ORDERED SAMPLES FROM UNIFORM DISTRIBUTIONS

To get nearer to more realistic situations we treat ordered samples from the uniform distribution on $0,1,\ldots,r$ $(r > 1)$.

GAME 5:

A fair die (numbered from 0 to 5) will be thrown in this game 10 times. Before that, everybody writes 10 integers (as tips for the results), and orders The generated sequence is also ordered, and is written under the tip sequence. The differences between the corresponding digits of the tip and the generated sequence are computed, and +1, -1 or 0 is written in the corresponding place, according as the difference is positive, negative or zero. Thus everybody has a sequence of elements +1, -1 and 0 (error sequence). If a "random walk" (starting from the origin) with this error sequence remains in the interval [-3, 3], the player gets 2 toplak; if in [-2, 2] then 5 toplak, and if in [-1, 1], then 10 toplak.

Pupils listed the maximal distance from the original (M) and also the sum of the digits in the error sequence (S) (Table 6).

Tip	1	1	2	2	3	3	4	4	6	6			
Ordered Sample	1	1	1	2	3	4	4	5	6	6	M	S	Gain
Error Sequence	0	0	+1	0	0	-1	0	-1	0	0	1	-1	10

TABLE 6

GAME 6:

The bank has a box containing x balls numbered from 0 to x. The player should find out x by drawing with replacement from the box 3, 4 or 5 times. The player should pay the bank 1 toplak for each drawing and gets from the bank x, x-1 or x-2 toplak respectively if he hits x, errs by one or two respectively; otherwise he gets nothing (x may be between 3 and 10).

The average gains of the player (when estimating x by the maximal value of the numbers drawn) can be seen in Table 7.

Average gains of the player in Game 6

Number of draws / x	3	4	5
3	-0.6	-1.4	-2.3
4	0.1	-0.6	-1.4
5	0.7	0.1	-0.6
6	1.3	0.8	0.1
7	1.7	1.3	0.8
8	2.0	1.8	1.4
9	2.4	2.3	1.9
10	2.7	2.6	2.4

TABLE 7

GAME 7:

A box contains r balls numbered from 0 to r. From this box n balls are drawn with replacement. The range of the sample (the difference between the largest and the smallest digit) is computed. Players write a tip for the range before generating the sequence and get 5, 4 or 1 toplak from the bank according as the absolute value of the difference between the tip and the actual range is 0, 1 or 2 (otherwise they get nothing).

Pupils evaluated the exact distribution of the range for r=5, n=3; for r=5, n=5; for r=9, n=5 and decided which is more profitable and which sum is worth paying to the bank for a game.

For r=3 (die) they examined the connections between digits of different distances. For n=3 they had a lot of empirical data from which they evaluated the conditional expectation of the i-th digit given the value of the first digit. Table 8 contains their results.

Value of the first digit	Conditional expectation of the			
	second digit		third digit	
	Theoretical	Empirical	Theoretical	Empirical
0	1.81	1.78	3.63	3.72
1	2.48	2.53	3.95	4.10
2	3.14	3.26	4.27	4.30
3	3.79	3.77	4.58	4.56
4	4.43	4.24	4.86	4.88
5	5	5	5	5

TABLE 8

They offered measures for the connection with the aid of this table. They offered games in which players buy information concerning the value of a certain digit.

GAME 8:

Player A puts (in secret) some white and some red balls (at least one of each colour) into a box, so that the total number of balls in the box is 10. Another box is prepared by A with the same content, both given to B. Then A tells B that there are either k_1 or k_2 red balls (among 10) in each box (the same number in both). B should find out which in fact is the case. B draws with replacement simultaneously from both boxes till a red ball comes at least from one of the boxes. If the last balls both happen to be red, then B is obliged to stop. Otherwise B may stop. In both cases of stopping B makes a decision (between k_1 and k_2). B pays one toplak to A for each (double) drawing. Further, B gets (pays) $\frac{10}{|k_1-k_2|}$ from (to) A if the decision is correct (incorrect). If B does not stop, he may go on drawing from the other box till a red ball appears there also (but no longer). The sums to be payed are as above.

Denote by $X_1(k)$ and $X_2(k)$ the number of drawings needed till a red ball appears from the first and second box respectively (each containing k red balls). Further denote $q = 1-\frac{k}{10}$.

Table 9 contains some values of the probability distribution of minimum $(X_1(k), X_2(k))$. $(P(\min(X_1,X_2) = i) = q^{2(i-1)}-q^{2i})$.

i \ q	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.99	0.96	0.91	0.84	0.75	0.64	0.51	0.36	0.19
2	0.01	0.04	0.08	0.13	0.19	0.23	0.25	0.23	0.15
3	0.00	0.00	0.01	0.02	0.05	0.08	0.12	0.15	0.12
4	0.00	0.00	0.00	0.00	0.01	0.03	0.06	0.09	0.10
5	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.06	0.08
...									
	$P(X_1(k) = X_2(k) = i)$								
1	0.81	0.64	0.49	0.36	0.25	0.16	0.09	0.04	0.01
2	0.01	0.03	0.04	0.06	0.06	0.06	0.04	0.03	0.01
3	0.00	0.00	0.00	0.01	0.02	0.02	0.02	0.02	0.01
...									

TABLE 9

References

[1] Engel, A. (1976) Wahrscheinlichkeitsrechnung und Statistik, I-II. Klett: Stuttgart.

[2] Rényi, A. (1969) On the teaching of probability, Rouse Ball Conference.

[3] Rényi, A. (1970) Foundations of probability. Holden-Day, Inc., San Francisco, Cambridge, London, Amsterdam.

[4] Vincze, I. Introduction into the theory of order statistics... (Manuscript).

[5] Bognár, K. (1980) Random fluctuation through games. "Stochastik im Schulunterricht". Hölder-Pichler-Tempsky, Wien, B.G. Teubner, Stuttgart.

[6] Bognár, K., Nemetz, T. and Tusnády, G. (1980) Ismerkedés a véletlennel. Tankönyvkiadó, Budapest (In Hungarian).

TOPIC 3: TEACHING STATISTICS IN SCHOOLS

(16-18 YEARS AGE GROUP)

ORGANISER: PROFESSOR J SWIFT

AMOUNT AND KIND OF STATISTICS TO BE TAUGHT TO 16+ YEAR OLDS

F Eicker, University of Dortmund

PRESENT STATUS OF STATISTICS AT SCHOOLS. FUTURE STEPS.

The lectures at this conference have shown once again a broad consensus among statisticians on arguments in favour of the school subject statistics and also on general issues of how to teach it: e.g. the importance of education in numeracy in today's world, the teaching should start at a young age, realistic data of interest to the pupil exist abundantly and can be used at school. But also quite a few details of the teaching methods and contents seem to find widespread approval: e.g. certain proposals on how to introduce and to apply specific concepts or methods, how to tie probability to statistics, to let children collect their own data, etc.

Our future work on teaching statistics and at conferences such as this is likely to concentrate on balancing the new subject of statistics with the other school subjects, on further details of instruction and particularly on sharing internationally the teaching experience already available. This procedure may lead rather quickly to a sufficiently large collection of tried out teaching topics from which the (modest) content of a statistics syllabus may be chosen to suit the needs of the particular groups of students at hand. Many problems are similar internationally as again this conference has shown. New textbooks may then be developed like 'Statistics in Your World' for the 11-16 year olds [5] and existing textbooks may be adapted. When good teaching material exists which represents well enough the statistical activities this will provide the basis for a new discussion of the nature and of the place of statistics at schools. This certainly is necessary in many countries, e.g. in Germany.

Since I am not a school teacher I cannot contribute anything regarding classroom experience. But since teaching 16-19 year olds and teaching statistics to non-statisticians at college have much in common, let me describe some personal experiences and ideas, and let me propose a few methodological details and sources of applications useful at schools.

But let me first, though, recall some characteristics of statistical work in general and some of its implementations for school teaching. Let us reflect in particular how it differs from mathematics. All this of course is basic to our efforts. It is not new what I say here but it seems necessary in many countries, particularly in mine, to draw the attention to the very nature of our subject in order to present the field of statistics appropriately in our teaching and to guard against unforeseen developments such as they occurred e.g. in set theory at schools.

THE NATURE OF STATISTICAL REASONING

Let me illustrate by an example some relevant aspects of the nature of statistical reasoning. Part of the example seems also directly suitable for school instruction: let us look at regression analysis which is probably the most widely used and also a most typical statistical technique (for a typical recent book see [1]). Its general aim is to try to represent observed values of a variable y sufficiently closely as a certain function of the values of another observed variable x usually called regressor or predictor. If nothing is known about the causal relationship between both (e.g. between height x and weight y of a growing plant) one would of course better look for a more general interrelation $g(x,y) \approx o$ between both. (The case of two or more regressors is frequent in practice but needs not be taken up at schools nor for our present purpose).

The closeness of fit of f depends on the purpose of the investigation. Often neither the type of the function f nor the degree of closeness of the fit are known in advance but have to be tentatively or successively determined. Typically this is a problem of data analysis or of descriptive statistics. The representation can be considered a mathematical model $y \approx f(x)$ in case the causal relation between x and y is clear (as e.g. between the weight of y of a plant and the amount x of water or a fertilizer). Then also the name predictor for x becomes evident. The simplest case here is a linear function f of only one regressor: $y = a + bx + e$ which can always be assumed in case both the x and the y-values are clumped in small neighbourhoods and the global function f is smooth on the x-neighbourhood.

Any relevant prior information from the empirical object of study (which has nothing to do with Bayes' prior distribution) regarding the structure of the function f, the kind and number of regressors, the errors e etc. should be exploited in this whole process. The previous knowledge largely belongs to the empirical sphere of the problem but clearly must be investigated and understood by the statistician. This part of his responsibility thus has the nature of an inductive investigation typical for quantitative empirical science. If no or only insufficient prior knowledge is available it may be substituted by structures obtained from exploratory data analysis as in the above regression example indicated.

Once these steps have led to a mathematical model the derivation of its mathematical properties of interest has the nature of purely deductive reasoning and thus is part of mathematics, in particular often of probability theory and mathematical statistics. But the reinterpretation of deduced results into the empirical context then again leads out of the domain of mathematics.

Thus statistical analysis comprises both induction and deduction simultaneously and consciously or not, and regression is very well suited to exemplify this according to the very brief description above. Regression is also very well suited under didactical aspects since there are very simple as well as very complicated situations and an abundance of interesting examples.

An important conclusion we derive from this and many similarly suited examples of statistical analyses is that statistics can properly be subsumed neither under mathematics nor under any empirical field alone. If we could reach general agreement on this it would answer many futile disputes once and for all. Moreover, the relevant course of the school curriculum should appropriately be named just statistics or at least 'mathematics and statistics' instead of mathematics alone.

It must be emphasized that the empirical work includes all relevant technicalities of the often very complicated data gathering processes. These should be subject to statistical requirements or control whenever possible (e.g. in controlled experiments). Consequently the statistician must cooperate most closely with the empirical worker who as a rule does the actual data collection.

The above account suggests that statistics can be defined as a methodology to acquire quantitative knowledge under uncertainty. (The definition occasionally made as decision making under uncertainty is too narrow). Its duality of approaches leads to calling it an 'art and science'. Remarks to this end can be found in many forewords of statistical textbooks, in the International Encyclopaedia of Statistics [3] under the article on Statistics by W. Kruskal and elsewhere.

It should still be mentioned that often only higher precision of the results requires the use of probability models and tools in statistical analysis, unless prior information demands a specific probabilistic model. Since in many cases and in particular at school high precision is not required, purely data analytic methods constitute an important branch of statistics and of teaching it. They are often mathematically fairly easy, so they can be treated conveniently at school. Since in hardly any cases where precision is needed the probabilistic model assumptions are acceptable beyond any doubt, data analysis is still an indispensable tool in establishing or checking these assumptions (e.g. checking residual plots by eye). Thus statistics instruction at school restricted to descriptive data analysis may often be completely sufficient. Part of it can be taught at a very young age, other parts can be quite demanding mathematically (details and examples below).

Another general comment is that no misleading illusion about the potentials or the scope of statistical methods should be awakened (e.g. if the data base

is shaky). Some of the limitations can be inferred from the above said. On the other hand it should not be reaffirmed statistically what is evident beyond any doubt simply by looking at the data.

In summary then regarding teaching statistics at school it should be clear from the above that

- the didactical and organizational problems are quite specific and should be well contemplated;
- the limitation to the mathematical aspects or foundations only is no substitute for a proper statistics course and would undoubtedly lead to a quick end to the episode of statistics instruction at schools;
- a balanced presentation of statistics can provide lots of excellent motivation for any age and ability age group of pupils. It thus can foster objective and mathematical thinking for all.
- overselling statistics at school should be avoided as carefully as underselling it.

TARGET GROUPS

A discussion of the amount and the kind of statistics to be taught at school in any one country must differentiate among the various target groups of students to be instructed. A possible statistics programme for those leaving school at 16 will look quite different from that for students up to 18 or 19 years. And again for the 16+ group there will be different programmes for those concentrating on vocational or on general education courses. We shall subsequently not discuss specific requirements for special types of schools or course programmes but rather concentrate on general education courses. To a large degree, although this varies temporally and geographically, these students will not study at universities later on, and again, of those who do go on to universities, only a fraction will ever use statistics professionally or even study mathematics, science or statistics. Hence it certainly is not justifiable to adjust statistics programmes to the specific needs of some minority groups. The pedagogically much more demanding task and the greatest need rather arises from the group of students that later on uses statistics peripherally in business and industry, in the biosciences and health, socio sciences, psychology and other areas. Often enough those taking jobs in these areas may actually get their only statistics education at school. But an equally important challenge lies in the above mentioned citizen awareness type of education in an age when many published statements use (and abuse) statistical concepts. Still another need exists in the other school subjects taught concurrently. In summary then it should be clear that we must concentrate our discussion on the modest needs of an average user (or consumer) in the above sense and from this there arises a didactical task

quite specific for the field as should have become clear by delineating the very nature of statistical reasoning.

Since a statistics syllabus in most countries has not yet been rigidly introduced, much of this article concerns both the pre-16 and post-16 age group. Where appropriate a dividing line is pointed out.

SELECTION OF STATISTICAL SCHOOL TOPICS

The case of regression mentioned earlier is but one example from the wide range of statistical subjects. We have to choose very few representative topics for school, perhaps one or two: regression may be a possible one because of its flexibility. Obviously there is no optimal choice. Nor do we want to train a mini-statistician. Rather it would be a serious question what a minimum amount of teaching material is that still makes sense. But we are not trying to answer it. Certainly the whole task is very much eased if elementary concepts and examples are already taught early during schooling. This possibility certainly exists and is well tried out (cf. the recent school book series 'Statistics in Your World'). A separate session as you know is devoted to this question. Let me now loosely list and discuss some themes which may be candidates for school. Of course the list is incomplete and I may have forgotten even some very important topics. In the case of each and every subsequently suggested method this method should be strictly tied into a real life motivating problem. This is usually a time-consuming job to teach.

To begin with the children should collect their own data a number of times. If this is done at a higher age (e.g. to take interviews) the students may in addition describe facts that influence or determine a probability model for the sampling process. Elementary data analysis can be done early. Location and scale can be discussed and investigated via median or mean, resp. range and variance. These measures must not just be calculated and then left but they must be interpreted in the empirical context. For higher grades more complicated methods of exploratory data analysis may well be up to the mathematical abilities, e.g. regressions, fitting curves by least squares, smoothing, trends, seasonal variations, even lagged variables, forecasting, orthogonal regressions and the question of causality, possibly also rank correlation. Even a limitation only to examples from descriptive statistics might seem reasonable. It would avoid the lengthy and delicate introduction of probabilistic inference, model building etc. which would almost require a separate course if done properly. But such a restriction should not rule out what might be called 'intuitive inference' (G. Noether). Descriptive statistics must always even at the earliest stages provoke the question 'What do these figures tell me?'. It is almost needless to say that they can only be taught by way of real life examples. It is anyway

the natural preparation for probabilistic inference: data analysis often gives hints for particular probability models and it is used to check their validity and the inferences based on them. We said above that this is the way one proceeds in statistical practice. The nature and value of inference statements can then be better understood with their help. There is a lot of material illustrating this from statistical practice and much can be used in school.

Only a few of the simplest examples of inferential statistics should be explained. These examples are often excellent ways of showing the value of 'means' and similar functionals. In particular discrete, finite valued random variables and distributions suffice. But the real meaning and validity of (simple) stochastic models should be deeply and critically discussed and understood (if discussed at all). Elementary non-parametric methods are superbly appropriate in this connection, as has been advocated for a long time e.g. by G. Noether. Point estimation and confidence interval estimation should be preferred to testing.

Probabilistic tools should be kept to the utmost minimum and should only be introduced if their bearing on statistical applications is clearly and immediately visible to the students. Actually it has often been advocated by teachers and statisticians that at the school level all that is needed is an intuitive understanding of the frequency aspect of probability. Much can be referred to tables or simulation experiments. But if probability is really used it should always be the aim to ensure that the student really understands the leading ideas and that he knows what they mean in the real world and in a statistical statement, e.g. the concepts of randomness, stochastic independence, probability as a measure of uncertainty etc.

It should be understood that this latter probability concept is actually often applied in everyday life to such non-frequentist applications as the probability of a horse winning a race, a particular football team winning, of passing a particular examination. But still this is not a reason for including a substantial course on probability theory for any age group, e.g. the impression must strictly be suppressed that the concept of an integral (or even improper integral) is a necessary prerequisite for handling simple statistical problems. On the other hand statistics and its application offer a wealth of opportunities to select and illustrate any number and any level of mathematical concepts and techniques that are possibly desired in general school instruction. It is thus in perfect agreement with a strong and lasting strand in formulating the goals of mathematics education in general as expressed e.g. in 'Studies in mathematics education' [4] (cf. in particular vol.2, p.159ff): it consists of stressing its relation to and bearing on many needs of real life problems.

By contrast a course on probability alone (or one with just a brief formal appendix on a topic of theoretical statistics) would be much more in the strand of abstract mathematics teaching. It would then also be in competition with other purely mathematical topics. Such a course would not incorporate the motivation by statistical applications. It must therefore be expected to have the fate of a brief transitional fashion in advanced mathematical school instruction like many other themes in the past and in the present. It would be fully in the tradition of most mathematics instruction. It also does not make a difference if such a course carries the new name 'stochastics' (as in Germany) which by definition excludes statistics in a comprehensive sense.

Continuing the survey of possible school topics it should be added that when the pupils actually take their sample also the defining questions and the practical problems connected with it should be pointed out to them, e.g. the question what is a household or what errors can occur when taking an interview. Experimental design may be mentioned, no matter whether the issue is spelled out in detail. Although usually the statistician does not himself collect the data he must know and often control everything relevant to his inference about this process. Hence also the students should hear about these steps.

Another good source of examples in sampling, random sampling or stratified, is the (rough) estimation of finite population parameters from such sampling. The probability model in this case is quite easy to obtain by the way.

Some statistical topics that actually occur in geography, economics, biology, science as taught today to the post-16 group are, according to findings of the Schools Council Project at Sheffield [2], the design and use of questionnaires (and interviews), reliability of data, index numbers, averages, trend analysis. Some, or all (and many more) of these basic concepts are vital, and they can be taught in statistics courses. Their importance as well as their presentation obviously need not or cannot be compared with traditional mathematical teaching topics. It is our didactical challenge to incorporate such new concepts into school instruction, to prove that this is possible and then press for an adaptation of the curricula.

Good sources of real life examples and often a sufficient description of methodology can typically be found e.g. in good textbooks on introductory applied statistics for biology, medical and economics students; they exist in many languages.

It also seems obvious today to use electronic computers in the actual statistical calculations. There are many encouraging experiences in this direction. For detail refer to proper sessions of this conference.

Whatever is going to happen, a self-critical and unambitious attitude is appropriate with regard to the total amount of teaching material as well as the degree of

sophistication. A few enlightening and really well understood encounters with statistics within an empirical context in the pupil's entire school career seem completely sufficient and at the same time all that can realistically be hoped for. For instance in the 16+ age group a practical case or two of sampling and estimation of the population parameters from such sampling might fully suffice. On the other hand time and again students should be reminded of the existence of statistical investigations and methods applied to a wide variety of cases so that they can see the flexibility. But as much as an attempt at a systematic coverage we must avoid the mechanical reproduction of some cookbook recipes. Considering the long list of possibilities, demands and restrictions there is much hard work still to be done on practical teaching beyond the consensus on general issues.

One of these acute tasks is teacher training. There are again special sessions on this, so let us remark just one or two things. Obviously the difference between statistical and mathematical work must be effectively reflected in teacher training, possibly also in (new) university courses. The work and scope of the average teacher, i.e. of their majority should be taken into account. The development of realistic training programmes as well as of curricula will require patience and time. It is not easy to design the appropriate statistics training programme for teachers. For instance courses taught other than by mathematics departments may be appropriate and suitable for teacher students. But this problem must be realized and tackled or the experiment 'statistics in schools' will become another shortlived and disappointing episode in the history of school mathematics.

Acknowledgements

For critical comments and additions on a preliminary version of this paper the author is much indebted to G. Fillbrunn, P. Holmes, H. Kütting, T. Nemetz, G. Noether and J. Swift.

References

[1] Draper, N.R. and Smith, H. (1981) Applied regression analysis. 2nd ed. Wiley.

[2] Holmes, P. et al (1981) Statistics in Schools 11-16: A Review, Methuen Educational.

[3] Kruskal, W.H. (1978) Statistics. Internat.Encycl. of Stat., The Free Press, Div. of McMillan Co., New York.

[4] Morris, R, (ed) (1981) Studies in mathematics education, 2 vols., UNESCO, 7 Place de Fonteroy, 75700, Paris, France.

[5] Schools Council Project on Statistical Education (1981) Statistics in Your World, School book series and teacher's guide. Foulsham Educational, Slough,UK.

TEACHING STATISTICS TO 16-18 YEAR OLDS: DREAMS AND REALITY

T Nemetz, Mathematical Institute of the Academy, Budapest

S U M M A R Y

Many statisticians are convinced that statistics, as an independent and compulsory subject, should be included in secondary-level general education. There are different (sometimes quite ambitious) proposals for the subject matter. Our real possibilities seem, however, not to match them. In this paper we discuss some aspects of this confrontation. Our conclusion is, that to a certain extent, statistics can be dealt with within mathematics. Examples are given to show how one can embed statistical notions into traditionally taught mathematics. Still much effort has to be taken to improve conditions for a widespread introduction of an independent statistics course.

DREAMS

Likely most statisticians have had adventures into the area of "what should and could be taught" at school level. It is even more likely that all of us can turn up with beautiful ideas and programs after just short consideration. In what follows we try to summarize the typical first-thought-conceptions. This is mainly composed of (a bit biased) quotations (without references), mixed with my own first notions.

The starting point of my dream is a perfect harmony among the three main disciplines of the area called stochastics: the descriptive, the mathematical statistics and the probability theory. The best men in these areas are eager to cooperate in finding the best possible answer to the question: what, at which level and how should it be taught. The educational authorities, as well as the teachers of all subjects are cooperative, too.

Since there is hardly any historical or didactical reason for separating these three branches, their teaching goes parallel, following a kind of spiral structure.

The goal of teaching statistics at schools is making the children familiar with statistical thinking with emphasis on influencing their thinking rather than providing them with recipes for specific situations. The subject matter aims at the "average" (in many senses) pupils and is assembled to provide them with a proper understanding of the surrounding environment and to satisfy their curiosities about the "behaviour" of a random media, which is vital for intellectual growth. There are many illustrations for the wide applicability of stochastic methods with proper examples from various fields. Such examples are also incorporated into the teaching of the basic disciplines like geography,

biology, physics, economics etc.

An important aspect is that the pupils start getting acquainted with the world of randomness at a time when they are not yet infected with the notion of determinism, i.e., from the very beginning of the primary school. Statistical and probabilistic ideas are spread through the whole curriculum, from grade 1 onwards, allowing a gradual discovery of the "regularities in the uncertain". At the last stage a small, compact course is offered, containing formal, correct definitions and some theorems with proofs. Of course, this is an independent subject, and not a (possibly hidden) part of other subjects.

As to the goals and guidelines the dreams of many of us go more or less along the same lines. As to the details, however, the "first-thought-conceptions" vary from one extreme to the other. We bring just a few such extremes:

- Descriptive statistics is too specific to be taught. It suffices to mention just a few characteristics like the average, in connection with arithmetic.
- Descriptive statistics plays the most important role in everyday life, therefore its detailed teaching is of vital importance.
- Mathematical statistics is to be handled just to the extent of helping the teaching of probability (if included in the subject matter at all).
- I have encountered the proposal to include in the curriculum even the theoretical background of the method of chi-square tests.
- The children do not gain anything from the over-idealized/ theoreticized area of the probability theory - no need for it.
- A real understanding of the stochastic ideas supposes the knowledge of the central limit theorems, therefore one has to insist on its teaching.

Luckily, most dreams are much more realistic. These include the dreams of those who had taken the trouble to work for a concrete teaching project. Hopefully this conference helps us to collect a short description of their last dreams.

REALITY

Below we list some of the main difficulties in realizing not only our nice dream, but even more realistic conceptions. Of course, our experiences in this respect are strongly influenced by the particular Hungarian conditions, where we had, till last year, a very rigid school system (see [3,4]).

Let us start with a mild disappointment: there seems to be a complete disharmony among descriptive statisticians, mathematical statisticians, researchers in

probability theory and the ones who apply it. This disharmony leads to contrasting (if not conflicting) recommendations toward the educational authorities, who have by now accepted the idea of introducing statistics/stochastics in the curricula of the pre-university schools. Introducing statistics means a struggle for the authorities from the point of assigning time to its teaching. At present, this can be done to the detriment of some well-established subject only. And we have to admit that good teaching of statistics uses much time, especially at secondary level, and cannot be done in just a few lessons. And the authorities face the most serious problem: the lack of teachers with proper education in statistics.

Round table conferences on the teaching of statistics during ISI meetings show that the "teacher problem" is the biggest obstacle in introducing statistics to schools, all the world over. The initial training of teachers seldom contains even passing contact with statistics. Secondary level mathematics teachers have had some training in probability, but usually in a rather formal way, which does not help them to understand statistical ideas. Re-training teachers seems to be very difficult. Besides financial problems, the administration of re-training also makes difficulties, and the effectiveness thereof is questionable. We completely share the Polish experiences (W. Szlenk [2], pp.50-54): "... it is hardly possible for an adult to assimilate the elements... in just a two-week course... it is a very hard question how to teach an over 40-years old teacher. In my experience only a few of them could get used to notions like probability and independent events".

Some people have the opinion that this problem can be overcome by relying on teachers of other disciplines like economics, biology, social sciences etc. (see [2], pp.17,18). In Hungary this is no solution. Economics is taught in a few highly specialized secondary schools, and we are concerned with the education of "average" students in general secondary schools. Teachers of other subjects are lacking even the formal acquaintance with probability of the mathematics teachers. We can see some attempts in biology: among the 4500 test-questions assembled as candidate questions to the final examination in biology, there are 3 of statistical nature, one of which is given in the Appendix, Example 1, for judgement.

The examination-oriented teaching raises another difficulty. We have an "almost uniform" final examination in secondary schools in the main disciplines, also entrance examinations to universities. Due to this uniform nature, no one can expect a test-question about some statistical problem in any subject while it is taught just in part of the schools. This prevents a possible gradual introduction, since teachers stress and drill usually the more probable examination themes, neglecting simply what is not important in this respect.

One may count on the natural curiosity of the pupils. If we do, we must find situations apt to motivate the children, and to put life into the theory. This seems to be a hard problem, since the allotted time is usually too short. Realistic examples are awfully complex, the simple ones are mainly artificial. We should find statistical populations, which can be easily sampled, inspected in a short time, and have time and space invariant stochastic character, providing a sound ground for the "average teacher". There are not so many such populations.

Field trips and case studies do not seem appropriate for our purposes, however desired they are. They can hardly be embedded into the framework of a unified centrally governed school system. We are furthermore convinced that only enthusiastic and knowledgeable teachers can treat them successfully.

Finally, one more observable problem: 16 year-old children can make statistical inference faster and more easily than the 18 year-old ones, if none of them had earlier statistical education. This fact contradicts the attempts to postpone the teaching of statistics to the last grade.

WHAT CAN WE DO NOW?

Under the present circumstances teaching of statistics should be accomplished within mathematics as part of the compulsory education with a possible facultative independent statistics course at the final grade of the secondary school. Topics in statistics and probability are to be dealt with from the very beginning.

For the time being we are forced to follow "The less the more principle", considering the present state of the teachers' knowledge. Treat just that part of statistics which harmonizes with traditional mathematics. A large part of descriptive statistics satisfies this condition. There are other problems, although usually neglected in introductory courses, which are common exercises in mathematics, but help to interpret statistical terms, e.g. the minimum-problems leading to the arithmetic mean and to the median in a statistical population are straightforward exercises in traditional mathematics, at the same time they clearly show one aspect of why they are considered as averages and why they are linked to the variance and to the standard deviation.

I consider it important to clarify the role of the pair mean-variance in representing statistical data, i.e. Tschebyseff-inequality for statistical data. No mathematics teacher objects to its teaching if properly handled (for a possible formulation, see Example 2 in the Appendix). This gives a bit of insight into the magic 3-sigma-rule, one of the most frequent cook-book recipes.

Teachers do not object to statistical inference problems if they can be solved on a common-sense basis, without sophisticated methods. One widely used example in this respect is breaking a simple substitutional cipher. This highlights an important difference between statistics and mathematics: statistical work may include some shrewd non-mathematical guesswork. A short cryptotext shows the difficulties with a small sample. A number of similar problems based on written languages can be considered equally well.

Many statisticians insist on dealing with tests of simple versus simple hypotheses. Some of them find that these must be preceded by providing the children with a sound background in probability. If we want to develop a proper attitude to statistical problems, this is not necessarily the case. There are simply formulated test problems which can be successfully attacked by "ad hoc" procedures proposed by the class, and the goodness of the decision rules can be experimentally investigated. Such a problem is stated in Example 3 in the Appendix.

The teaching of probability should concentrate on the notion of the random variable, in a possible strong connection to some statistical problem. At secondary level, it is difficult to design and perform a consistent and didactically acceptable course where probability is defined as a function on sets. Therefore we are against both axiomatic and set-theory approach, meeting the experiences in the Netherlands (see [2], p.53). The way, we propose, looks like this:

- Make a prediction of the distribution of a simple random variable
- Collect data
- Compare the observed data with the prediction at an intuitive level

This concept is illustrated by Example 4 in the Appendix.

WAYS TO SOLVE THE "TEACHER PROBLEM"

The 4th Round Table Conference on the Teaching of Statistics of ISI has worked out a recommendation in this respect. Let us recall a part of it ([2], p.18): "Teachers should be assembled in seminars for a few days - a week, say - and work together on a set of lessons already experimented to analyze, criticize, modify and adapt them... in those seminars, some of the time could be used to familiarize them with some real statistical work".

The loosely stated duration of such seminars becomes more concrete in a work of V. Barnett [1], who claims: "Courses of about 40 hours with plenty of practical work and student participation surely represent the minimum contact necessary to begin to bring teachers to a position where confidence and ability may develop".

It should be added that the best way to encourage them to teach the subject is to provide them with very detailed manuals, containing an almost minute-to-minute description of the lessons, and listing a wide variety of possible traps with solutions.

CONCLUDING REMARKS

Despite the great difficulties, elements of statistics can be included in the compulsory curricula of general educating secondary schools. For the time being we have to be self-restraining regarding the subject matter of instruction. The chances of its gradual extension are fair. It is vital, in this respect, to test carefully planned projects by a number of "average teachers" and not by enthusiasts. Teachers, participating in trials of the projects, should collect and formulate their experiences and problems. These, together with appropriate comments, and a detailed description of the project, should be widely circulated. An ambitious curricula, or a compulsory, independent statistics course may be introduced just after having re-trained a considerable number of the teachers.

References

[1] Barnett, V. (1980) Statistical Education in Schools in England and Wales. Technical Report for ISI TOTSAS, University of Sheffield.

[2] Breny, H. Ed. (1976) The Teaching of Statistics in Schools. Proceedings of the 4th International Statistical Institute Round Table Conference on the Teaching of Statistics.

[3] Nemetz, T. (1979a) Pre-University Stochastics Teaching in Hungary. Technical Report prepared for ISI TOTSAS.

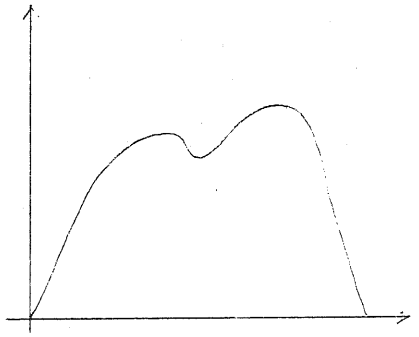
[4] Nemetz, T. (1979b) Stochastics in the Hungarian Schools: Pre-Univeristy level. In: Convegno l'insegnamento pre-universitaro della statistica con particulare riferimento alla scuola secondaria superiore. Bressanone, 19-21 settembre 1979.

A P P E N D I X

Example 1: A prospective examination-test question in biology.

The figure shows the distribution of a feature. Choose the most realistic statement from the following one!

A) There is no feature with such a distribution
B) The figure shows the typical distribution of a quantitative feature
C) It shows the typical distribution of a qualitative feature
D) It shows the distribution of a quantitative feature based on a very small sample
E) It shows the distribution of a qualitative feature based on a very small sample



Example 2: The proof of Tschebyseff inequality does not frighten teachers of mathematics.

Claim: Let $x_1, x_2, \ldots, x_n$ be a sample of non-negative numbers, $\bar{x}$ be the arithmetic mean, and $A > \bar{x}$ an arbitrary number. Then there are at most $\frac{n\bar{x}}{A}$ sample elements which are not smaller than A.

Proof: $n \cdot \bar{x} = x_1 + x_2 + \ldots + x_n =$

$= \{$the sum of x_i's which are less than $A\} +$

$+ \{$the sum of x_i's which are not smaller than $A\}$

$\geq \{$the sum of x_i's which are not smaller than $A\}$

$\geq A \cdot \{$the number of x_i's which are not smaller than $A\}$, etc.

Interpretation: The proportion of elements in a sample of non-negative numbers which do not exceed a given number A, where $A >$ sample mean $\bar{x}$, is at least $1 - \frac{\bar{x}}{A}$.

Now, if we have a sample $y_1, y_2, \ldots, y_n$ containing arbitrary real numbers, and the arithmetic mean $\bar{y}$ and the variance S^2 are known, then we can apply our result to the sample $(y_1-\bar{y})^2, (y_2-\bar{y})^2, \ldots, (y_n-\bar{y})^2$, $\quad A=\lambda^2 S^2$.

Since the arithmetic mean of this new population is S^2, we get:

The proportion of the elements y_i in the sample $y_1, y_2, \ldots, y_n$, for which $(y_i-\bar{y})^2 \leq \lambda^2 S^2$, is at least $1 - \frac{S^2}{\lambda^2 S^2} = 1 - \frac{1}{\lambda^2}$.

Interpretation:

If we know only the mean $\bar{y}$ and variance S^2 of a sample $y, y_2, \ldots y_n$ of reals, we can be sure that at least $(1 - \frac{1}{\lambda^2})n$ elements belong to the interval $[\bar{y}-\lambda S, \bar{y} + \lambda S]$

Here, one can take concrete values for λ, e.g. $\lambda = 3$.

Example 3: A decision problem.

A statistical population contains pairs of type A and type B objects. Both types of objects have 4 characteristics. These characteristics can be measured in a given order.

A sample of 10 elements was taken and is given here. The first column gives the label assigned to the given pair. Another person did not pay attention if he was measuring a type A or a type B object. In his records, therefore, the ordered quadruples may have got interchanged. Your task is to guess if it happened or not. Write AB into the last column if you think that they are not interchanged, and write BA otherwise!

No.	TYPE A				TYPE B			
7	11	7	1	1	15	2	5	1
8	14	1	11	1	21	3	27	9
9	29	6	1	1	9	20	18	3
10	19	1	1	1	14	5	6	48
11	2	12	1	1	1	1	1	1
12	14	1	1	1	30	1	1	40
13	2	20	1	7	1	1	1	64
14	8	10	1	1	26	7	2	27
15	6	72	10	1	19	11	5	3
16	3	1	1	8	16	1	8	17

No.	1. Data Set				2. Data Set				?
17	6	26	32	17	32	24	1	3	
18	12	1	1	2	19	4	4	7	
19	1	18	9	16	19	1	1	3	
20	2	35	5	5	1	1	1	31	
21	9	12	3	29	21	98	6	1	
22	8	33	22	47	1	1	1	3	
23	8	9	10	1	5	1	3	1	
24	8	26	11	6	34	1	15	3	
25	18	4	25	7	30	1	1	16	
26	6	12	17	3	15	7	2	10	
27	7	23	4	3	4	3	5	3	

For the reader's information: type A data was obtained in the following way. We have taken the book [2], and opened it at the page given in the first column. Then we have counted the letters and spaces on that page till the first letter *T*, to get the first "measurement". Starting from here we have counted again till the first *H* to get the second one, then till the next *E*, finally, till the next SPACE. The key-word STAR was used to get type B data. Finally, the interchanges were made on a coin-tossing basis.

Example 4: The problem of two coins.

Concerning the number of heads when tossing two coins there are usually two models suggested in the classroom:

$$A = \begin{Bmatrix} 0, & 1, & 2 \\ 1/3, & 1/3, & 1/3 \end{Bmatrix}; \qquad B = \begin{Bmatrix} 0, & 1, & 2 \\ 1/4, & 1/2, & 1/4 \end{Bmatrix}$$

The question, which is the "better", is dealt with via experiments. It is not boring to collect some 100-120 observations. A typical result;

$$E = \begin{Bmatrix} 0, & 1, & 2 \\ \frac{28}{120}, & \frac{58}{120}, & \frac{34}{120} \end{Bmatrix}$$

I was surprised to learn, and this is why this example was given, that children of 16 years of age frequently come up with the following decision-rule:

The probabilities in models A and B, the relative frequencies in E can be viewed as points in the 3-dimensional space. We will say A is better than B, if E is nearer to A than to B, and vice versa.

STATISTICS FOR THE 16-18 YEAR OLDS: A DISCUSSION

A Hanley, Northampton

There seems to be agreement amongst statistical educators that appropriate statistics courses should be provided for all 16-18 year old students. Professors Eicker and Nemetz have described what form this statistics curriculum could take. They have also given reasons to justify the introduction of statistical education which could be presented to the education authorities who, in many countries, decide centrally what the components of the school curriculum shall be. A major task is then to persuade these authorities that statistics merits a place in the curriculum even if the time allocated to other studies has to be adjusted to accommodate it. When this task of persuasion is accomplished schools will be obliged to introduce statistics into their curricula.

In the United Kingdom the task has a different focus. There is no centrally prescribed curriculum so that schools are free to choose and plan their own. Adjacent schools could be providing different curricula, each chosen by the staffs of the schools to realise their personal views on what constitutes the best course for students. In practice there is not the variety there might be owing to the influence exerted by the public examinations. These examinations are so highly regarded by parents, employers, teachers and hence by students that courses not leading to a public examination are regarded as of very low status and value. Schools therefore tend to offer packages of courses based on public examination syllabuses which leads to a fair degree of uniformity in the range of subjects on offer whilst variety of course content arises from the selection of a particular subject syllabus.

Changes in a school's mathematics curriculum usually follow an initiative by the Head of the Mathematics Department who may be influenced by a school colleague, by a local authority adviser, by an inservice course, by a professional association or by a national "fashion". Suggestions about desirable curricular changes are regularly made in reports published by the Government following the studies of Committees established to examine particular aspects of the educational system or following investigations carried out by Her Majesty's Inspectors of Schools. Such suggestions are frequently not acted upon. For example, the Crowther Report [1], published in 1959, had this to say about "numeracy", a word invented by the Crowther Committee and used often at this Conference.

> Literacy has long been important, and its value is as great as ever. Just as by "literacy", in this context, we mean much more than its dictionary sense of the ability to read and write, so by "numeracy" we mean more than mere ability to manipulate the rule of three. When we say that a scientist is "illiterate", we mean that he is not well enough read to be able to communicate effectively with those who have a literary education. When we say that a historian or a linguist is "innumerate" we mean that he cannot even begin to understand what scientists and mathematicians are talking about. The aim of a good Sixth Form* should be to send out into the world men and women who are both literate and numerate. It is perhaps possible to distinguish two different aspects of numeracy that should cocnern the Sixth Former. On the one hand is an understanding of the scientific approach to the study of phenomena - observation, hypothesis, experiment, verification. On the other hand, there is the need in the modern world to think quantitatively, to realise how far our problems are problems of degree even when they appear as problems of kind. Statistical ignorance and statistical fallacies are quite as widespread and quite as dangerous as the logical fallacies which come under the heading of illiteracy. The man who is innumerate is cut off from understanding some of the relatively new ways in which the human mind is now most busily at work ..... However able a boy may be, he cannot reach a Sixth Form level of numeracy except on the foundation of a Fifth Form* level; but, if his numeracy has stopped short at the usual fifth form level, he is in danger of relapsing into innumeracy. It is now one of the most important tasks of the Sixth Form to ensure that no boy or girl leaves school as innumerate as most have done in the past, and as far too many do even today.

Despite this exhortation, few schools over 20 years later have introduced the types of Sixth Form course envisaged by Crowther. Indeed, a prime concern of this Conference is that such courses should be developed for the 16-18 year olds.

---

* The Sixth Form consists of the 16 to 18 year old students in secondary education. The Fifth Form is the group of 15 to 16 year old students.

One of the reasons why advocated changes do not take place is that central government does not usually designate any portion of the grants to local authorities specifically to curriculum development. The latest report [2] has, inside the front cover, the following "disclaimer" (reminiscent of the Government Health Warning appearing on packets of cigarettes sold in the UK) "No assumption can be made about government commitment to the provision of additional resources as a result of this report." Allocation of financial resources to implement curriculum development requiring them is therefore a matter of local decision.

Whether innovation is directed centrally or chosen by the individual school the support that will be demanded from those already involved in statistical education will be substantially the same. Most teachers are not prepared by initial training to teach statistics and even mathematics specialists may have little statistical background. Inservice courses could be provided systematically to train teachers over a suitable period of time. (In the UK there is no commitment to compulsory inservice training that would ensure teachers were exposed to this). Suitable teaching materials are scarce but they are slowly being produced. Here there are the Schools Council's Project on Statistical Education booklets, a few Open University courses, some BBC schools' broadcasts and associated materials, and soon there will be materials coming from the Leverhulme Statistical Education Project 16-19. There is a taxing exercise in front of all of us to produce all that is going to be demanded to introduce sound statistical education into the curriculum of 16-18 year olds.

A recent survey of mathematics teaching in schools in England and Wales [3], investigated the mathematical needs of adults in their work and in everyday life. It was found that many adults were very apprehensive about being asked to face simple mathematical tasks - indeed, a good proportion of those approached refused to co-operate when they knew the topic was mathematics. Many of those questioned could not cope with percentages or simple averages nor read a timetable correctly. They ascribed their lack of confidence and understanding to a variety of reasons but a major one was the irrelevance to them of the work they had been asked to do at school. Is there a danger that courses in statistics might be subject to the same criticism? If so, what should we do to avoid this pitfall?

Lastly, we are a gathering of enthusiasts. Many of the teachers who may be required to tackle new courses will have no better, perhaps worse, statistical backgrounds than the students they teach. How can they be given the practical experiences of relevant statistics themselves so that they may teach their students with the conviction and insight based upon the confidence of personal

experience? A number of projects have tried to produce resource materials which are "teacher-proofed" in the hope that teachers cannot misuse them. Can we adopt the positive attitude of designing materials that will be "teacher-improving"?

References

[1] HMSO (1959) 15 to 18: A report of the Central Advisory Council for Education (England)

[2] HMSO (1982) Education 5 to 9: an illustrative survey of 80 first schools in England.

[3] HMSO (1982) Mathematics Counts: Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr. W.H. Cockcroft.

TOPIC 4: TEACHING STATISTICS TO NON-STATISTICIANS

ORGANISER: DR W G GILCHRIST

WE MUST PREACH WHAT IS PRACTISED

A S C Ehrenberg, London Business School

SUMMARY

Much of what we teach in statistics is seldom used in practice.
Much of what students need, we do not teach.

In teaching non-specialists, we need to cover the procedures they will come across in their other studies and subsequent work. Since many courses and introductory texts do not do so, students are often disillusioned with them and regard such courses as irrelevant.

I now describe some topics which we often teach unnecessarily (or with the wrong emphasis) and some which we mostly do not teach at all even though they are often needed. They range from deep matters like statistical inference or causation to more hum-drum (but important) ones like means, medians and modes.

FREQUENCY DISTRIBUTIONS OR TABLES

We spend much time teaching frequency distributions, both in dealing with observed data (grouped and ungrouped readings, histograms, etc.) and in telling students about theoretical models like the Normal, Poisson, or Binomial distributions. However, students seldom come across frequency distributions in their readings or practical work. This also applies to teachers - in my experience, when asked when they last used or saw a frequency distribution, teachers generally refer to some teaching occasion rather than to anything practical.

What students see instead is aggregated data: averages, totals, percentages and so on. In particular, they do so in the form of structured tables, where each entry is some such summary statistic. However, we hardly ever teach students how to analyse and interpret tables.

I am not suggesting that frequency distributions should not be taught at all. They are usually the basis of the aggregated statistics people do see. They also occur in dealing with deviations from a model, and they arise in random sampling (at least hypothetically). But if we want to avoid students being puzzled at what they are being taught, we ought to explain to them the real role of frequency distributions: why they do not often see them explicitly and why they nonetheless ought to know something about them.

MEANS, MEDIANS AND MODES

We spend a good deal of time teaching the calculation of means, medians and modes. Yet the latter two measures are virtually never mentioned again in the rest of the course or textbook, nor do they occur much in students' other reading.

Medians and modes are helpful concepts in understanding observed data. But they are seldom used in analysis, e.g. in comparing different sets of data or in model-building. This is not a value judgement, but a matter of fact. It ought to be reflected in our teaching.

THE MEAN DEVIATION AND THE STANDARD DEVIATION

Many introductory texts do not mention the mean deviation at all. Others denigrate it. Yet faced with a range of numbers like the deviations from a fitted model - e.g. 3, -1, 4, 0, 1, -3, 3, 0, etc. - we generally note that they are roughly 2. We form a mean deviation either in our head, or more precisely on paper, $15/8 \doteqdot 2$.

The standard deviation has well-known advantages, being both easier to compute from raw data, especially for large n, and also vastly more amenable to theoretical developments. But for the beginner the standard deviation only stops being a rather mysterious abstraction if we lead into it as a development from the mean deviation.

Both as a teaching aid and as something which most people often use in practice, the mean deviation is sadly neglected.

RANDOM SAMPLING VERSUS OBSERVATIONAL DATA

Random sampling in general and randomised experiments in particular are often treated in introductory texts as if they were the only ways of obtaining proper data.

But randomised experiments are only just over 50 years old. In practice, observational studies predominate, with all their additional problems. Our teaching should reflect this.

Further the vast majority of data are not collected by any method of explicit random selection. Each set of data is in effect a small ad hoc population. Yet such data are not necessarily bad. Measuring the whole population may be wasteful, but it is not wrong.

What is needed is the notion of generalisation. But this is hardly ever discussed. I have not seen the word "generalisation" or anything like it in the index of

a statistical textbook, let alone in the body of the text. Yet the idea of a result which is known to hold under a wide range of different conditions is basic to ordinary science and technology.

BEST-FIT EQUATIONS VERSUS LAWLIKE RELATIONSHIPS

Virtually all our teaching of the relationships between two or more variables is in terms of finding the best-fitting equation for a single set of data. But this is not what students do (or should) come across in their ordinary studies - a result which was only established once and which has never been repeated.

In contrast, fitting an equation to two or more different populations (or samples therefrom) is not discussed in our literature. The derivation and interpretation of the lawlike relationships of science does not seem to be regarded as a fit occupation for a statistician.

PROBABILITY VERSUS CAUSATION

Probability figures large in many statistics courses and texts. Some even start by saying that statistics is about probabilistic data.

Yet having been taught the laws of probability (usually in terms of games of chance with equally likely outcomes - a highly improbable concept that never occurs elsewhere in life or science), students find that neither they nor their introductory text make any further use of probabilities. (The basic notion of random sampling does not require a detailed mastery of the probability calculus).

In contrast, the process of trying to move towards causal interpretations of observed data is seldom discussed in the usual texts.

EXPLORATORY ANALYSIS VERSUS REPETITIVE STATISTICS

Most situations treated in traditional statistics are <u>exploratory</u>. This was so even prior to John Tukey's emphasis on EDA.

Each new set of data is to be analysed as if there were no prior knowledge of other similar data. This for example underlies our usual notions of inference ("Is this correlation significant?") and that of fitting a regression line ("How does y vary with x in this data?").

But in practice, most of the data professional people come across are <u>repetitive</u>. They have seen similar data before, usually week-in week-out. They therefore have much prior knowledge and understanding of the phenomena (except perhaps in the case of statistical consultants).

DISCUSSION

In this paper I have briefly outlined a number of points where what is taught (i.e. what is in our textbooks) differs greatly from what students actually need.

I believe that we must change what we teach, to cover the kinds of statistics that students come across in their main subject-areas and their practical work. This seems a more important topic than whether courses should be taught by statisticians or subject-matter specialists (to which there can hardly be a single general answer).

The re-orientation I am suggesting will also overcome statistical teachers' continuing lack of practical teaching materials. They often report that they cannot extract practical illustrations or case-histories of using "Technique X" from statistical practitioners. But as Professor Gerald Goodhardt has pointed out, the reason is simply that although Technique X is in our textbooks and is taught in our courses, it is practically never used. That is why there are no practical examples.

In adapting our teaching, some temporary compromises will be needed. There are two reasons. First, individual teachers will not always find it easy to give up instantly much of what they have been teaching for many years. Secondly, we may have to go on covering items that are required by syllabuses or examination systems (which often take longer to change). Nonetheless, we can work at making our texts and teaching materials more relevant.

TEACHING STATISTICS TO NON-STATISTICIANS - AN APPLICATION ORIENTED APPROACH

G Sadasivan, Indian Agricultural Statistics Research Institute

INTRODUCTION

Teaching statistics to non-statisticians is more difficult than teaching the subject to statisticians. Some methods of teaching statistics to statisticians are given in Kempthorne (1980), Mosteller (1980), Noether (1980) and Tukey (1980). The role of statistics in the teaching of agricultural scientists is given in Sadasivan (1966). In the author's opinion the classroom type of teaching is much better for such personnel than the seminar type of teaching. This is specially so if the performance of the trainees is to be assessed by any type of test subsequently. It may be noted that the statistical topics to be taught to different categories of non-statisticians are different. The methods of teaching these topics should be oriented towards the methods of assessment of the trainees. Teaching with detailed mathematical derivation would be troublesome to the non-statisticians. For instance, for introducing statistical quality control in textile industry the mill owners are to be instructed in very simple terms how statistics can be used to benefit production of textiles of acceptable quality level. This is a seminar type of teaching non-statisticians. This type of teaching is of relevance in Government Departments as well, where the boss does not know anything in conventional statistics, but has to take decisions using statistical data in his day to day activities. The agricultural scientists in India are also a group of non-statisticians who know next to nothing in statistics and it is the job of the statisticians to explain, say, the benefit of using a complicated experimental design for such an audience. For instance, the present generation of agricultural scientists in India did not want to use anything more than randomised blocks or split-plots for their field experiments. In many cases they used to lay out their experiments without using any statistical considerations and later on approached a statistician to squeeze out as much information as possible from their experiments. Although this gives rise to new problems in statistics, it is not a happy situation. One way to prevent such mishaps is to impart appropriate statistical training to such groups of non-statisticians. But a word of caution is necessary here. The training of such personnel should not be highly mathematically oriented lest they totally dispense with the services of statisticians.

It is clear from the exposition above that there are different categories of non-statisticians as (1) ordinary scientists (2) agricultural scientists (3) engineers and technologists (4) business scientists (5) management scientists (6) administrators (7) farmers (8) other laymen and (9) the general public. The teaching

method need not be the same for each of these categories because the educational level especially in mathematics is different among these categories. Hence the teaching method has to be oriented towards the needs of the particular subgroup we are considering. We have also to consider whether there can be a general method of teaching applicable to all non-statisticians.

Some of the organisations involved in teaching statistics to non-statisticians in India are the Indian Statistical Institute, Indian Agricultural Statistics Research Institute, Agricultural Universities, Indian Institutes of Technology and regular Universities. Teaching non-statisticians in these organisations calls for inter-disciplinary collaboration. As far as collaboration between agricultural scientists and statisticians is concerned, the scientists will always have an edge over the statistician.

Teaching experimental results to farmers also falls in the class of teaching statistics to non-statisticians. Since all experimental results are expressed statistically, usually only the interpretation of statistical analysis of the results is given to the farmers. They are taught with the help of diagrams, charts, tables, etc. and with the help of projectors by the extension specialists. The data are derived from comprehensive experiments in the cultivator fields or from the results of the experiments in research stations. The teaching involves results of precise and careful data analysis. If the literate farmers are given the rudiments of the concepts in statistics, they will be able to appreciate the experimental results much better. To make this programme effective the agricultural extension workers are to be trained adequately in statistics.

There are many statistical training courses for non-statisticians at the Indian Agricultural Statistics Research Institute, New Delhi. In 1958 we framed two courses for the benefit of agricultural scientists. These were later reorganised into six courses for experimental statistics. Then the necessary pre-requisite courses in Mathematics were framed especially for students in Agricultural Physics, Agricultural Engineering, Genetics, Agricultural Economics etc. It may be noted that the courses formulated for each of these disciplines contain the appropriate topic in statistics useful to them. The courses for Genetics students include Biometrical Genetics, Population Genetics and Biometrical Methods, courses for Entomologists include Tagging Models, Probit Analysis, Bioassay Techniques, courses for Mycologists include Contageous, Negative Binomial, Log-normal and other types of distributions apart from Elementary Statistical Methods. Similarly the courses for Agronomists include Advanced Design of Experiments, courses for Extension include Nonparametrics and Psychometric methods, courses for Economics include Econometrics. Other disciplines in Agriculture use elementary statistical methods. Similarly there are certificate

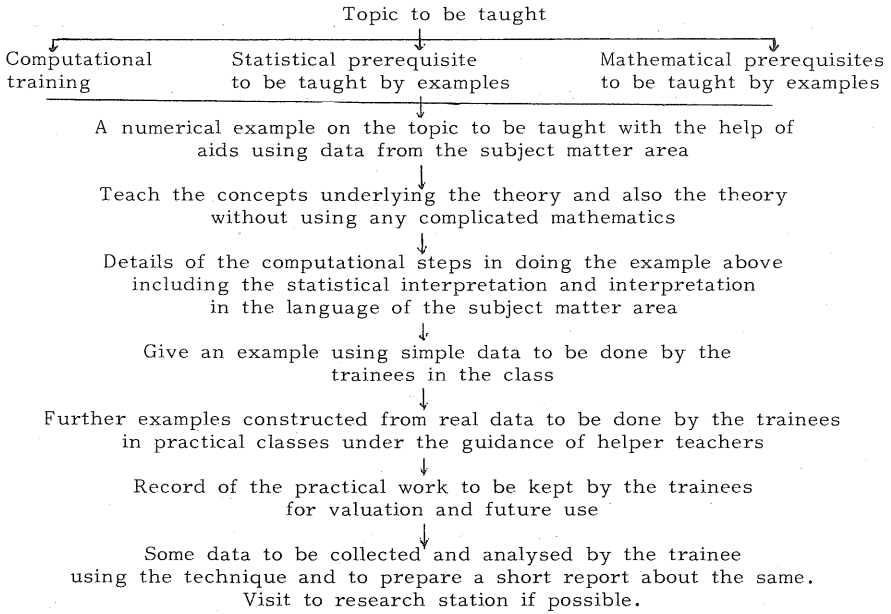
courses in the Institute for training nominees from different states in India and abroad who had no formal training in statistics to enable them to perform their day to day statistical duties. These trainees belong to a wide spectrum, most of them matriculates or degree holders in Agriculture or Animal Sciences but by virtue of their jobs are doing statistical work in their parent departments. Some of the trainees have deficiencies in languages. There is arrangement to train them in English before their initiation to statistics courses. Apart from these there are short term courses in the Institute which impart training in statistics to Directors of Agricultural Institutes, FAO Trainees, personnel of the All India Services such as Indian Administrative Service, Indian Forest Service, Agricultural Research Service, etc. The method used by the author for training these personnel along with some of the other possible methods are detailed in this article. Apart from these, the author is associated with the Food Scientists in India who use a number of non-parametric methods in their training programme in statistics. It may be noted here that the training imparted to Indian Administrative Service and other probationers is of a general nature which helps them to understand the working of the different sections and organisations dealing with statistics under the Indian Government. Further I.A.S.R.I. had a programme to train non-statisticians to professional statisticians. Under this programme a graduate in Agriculture can be trained for the M.Sc. degree in Agricultural Statistics. The first year of the course is used for teaching supplementary mathematics. The second and third year are used for training in statistics including a research project.

Now a word about teaching non-statisticians in Universities. The methods of teaching used below can also be used in this case. Usually, for teaching in Universities real data are not used. The students are not given any project work.

Advisory work to Agricultural Scientists for conducting their research is a type of training of the non-statisticians in the art of data collection. This involves training in planning and analysis of their field and laboratory experiments in such a way that data on different characters can be collected and analysed. In these cases knowledge of programming is an advantage to the non-statisticians.

THE MAIN TEACHING METHOD

Here we discuss an application oriented method of teaching statistics to non-statisticians. It is described by chart as follows:-

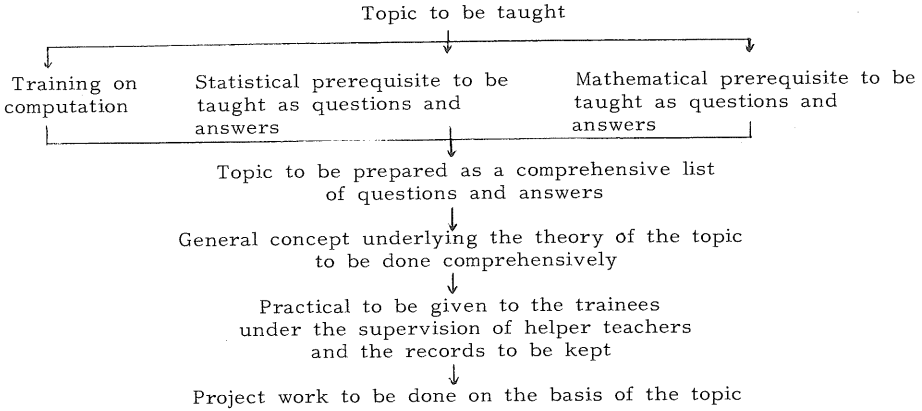


The method can as well be described as follows: First decide about the topic to be taught. Then teach the required pre-requisite Mathematics and Statistics as supporting courses in a simple way by doing simple exercises. Then construct a numerical example from the subject matter area for the topic we decide to teach and interpret the statistical procedure using the numerical example. If charts or other teaching aids are helpful they can be used for this teaching. Proceeding from this we can state the general statistical concepts underlying the procedure along with the necessary formulae, stating clearly what the symbols in the formulae stand for. Then the computational steps involved in doing the practicals relating to the procedure are to be given in such a way that the results can be calculated easily by the substitution of the unknowns in the formulae from the numerical. A statistical interpretation of the results obtained should be given next. This interpretation is further translated in the language of the subject matter area for the convenience of the non-statisticians. The computations in the above should be such that they can easily be performed with the hand or with the help of a facit calculator. In cases where the non-statistician knows programming the computations are made using an electronic computer. But this procedure will help the scientist to comprehend fully the implications underlying the statistical method only if he knows programming. Now give a small example to the trainee as classwork. Then good practicals based on the procedure are to be constructed from the subject matter area

and given to the trainee. These computations are demonstrated by helper teachers. This will help the non-statistician to keep a good record of the work done by him for future use in all his scientific endeavour. If possible, some data should be collected, analysed by the trainee from field or laboratory or any other source using the technique under study and a short report written. Visits to research stations practising the methodology will further help the non-statistician to understand the method.

OTHER TEACHING METHODS

The teaching methods should be oriented towards the method of assessment. The method given above would be useful if the questions are comprehensive in nature. But the recent trend in assessment method is to ask objective type questions. For this method of assessment a teaching method designated Method II is diagrammatically as follows:-



In words Method II can be described as follows. First the prerequisites to the topic can be taught in the form of questions and answers. The next step is to prepare a comprehensive list of questions, the answers to which would cover the topic to be taught. Next the general concepts underlying the topic can be taught followed by a list of questions for practical work supervised by helper teachers. The rest is as in Method I. In this method of teaching much expertise is required for proper drafting of questions and their possible answers. The questions can be oriented towards practical things as well as theory. In teaching non-statisticians, the quantum of practical questions should be more than that for theory.

In the following Method III yet another technique of teaching non-statisticians is discussed. Here the topic to be taught is assigned to a trainee. Under

the supervision of the statistician the student is to prepare the teaching material and teach the topic to the trainees. The statistician should be present in the class to resolve any difficulties while teaching. The order and technique of presentation can be either as in Method I or Method II. It may be noted that Method III facilitates better discussion.

REQUIREMENTS TO FACILITATE BETTER TEACHING

In the following we discuss some requirements other than those given earlier for better teaching. The courses for training of non-statisticians should be introduced from elementary levels to advanced levels. For training such personnel teaching aids are very helpful. They include, graphs, diagrams, projectors, calculators, television programmes, computerised courses, practical classwork, class notes, statistical and mathematical tables, boards, chalk, library and other equipment. Each category of non-statistician requires a special area in statistics. His training should be oriented towards this need. This specialised training is very useful for developmental research in general. There are special statistical publications which are of interest to non-statisticians. These must be listed and brought to their notice. Non-statisticians should be encouraged to collaborate in statistical research projects applicable in their areas. This will ensure better applicability for the methods evolved from such projects. Non-statisticians can participate in statistical research seminars which are of relevance to their specialised areas. Teaching of non-statisticians requires inter-disciplinary collaboration. If non-statisticians develop a supercilious attitude such collaboration is difficult to achieve. In training non-statisticians there should be specific reasons for including statistics as a part of the curriculum. Every scientist should be taught data gathering, principles of design and survey planning. In teaching non-statisticians case studies can be utilised. The needs of course contents for different types of non-statisticians are different. Some of these were briefly discussed in the introduction.

EXAMPLES OF TEACHING METHODS

Example I: Let us consider the example of teaching the binomial distribution and some of its applications to agriculture.

Step 1 - teach the following prerequisite mathematics and statistics to the trainees: proportions and percentage; square roots; theory of indices; frequency distributions - discrete and continuous; central tendency and dispersion, calculation of mean and variance from grouped data; probability - addition and multiplication rules; mathematical model building; sampling for proportion from finite and infinite populations; estimates and their properties. Train the participants in computational techniques as well. Then teach a numerical example like the following:-

Step 2 - in an experiment on wheat we desire to find the proportion of wheat plants affected by wheat rust along with its error. Divide the area into micro-plots of the same shape and size, each plot containing n = 10 plants. Select N = 100 such plots at random from the tract which has in all N^1 = 1,000 such plots. The distribution of the number of affected plants per plot is as follows:

Number of affected plants per plot (x_i)	0	1	2	3	4	5	6	7	8	9	10	Total
Number of plots (f_i)	5	7	12	15	15	20	15	6	3	1	1	100

Assume that the probability p, for incidence of wheat rust is constant over the area. The problem is to estimate this probability along with the appropriate error assuming that the binomial distribution gives a close fit to the data. First assume that the N = 100 sub-plots come from an infinite population. If the binomial gives a close fit to the data, the mean of the data $\bar{x} = np = 10p$, the mean of the binomial. Thus an estimate of p is $\hat{p} = \bar{x}/10 = 4.11/10 = .411$. A biased estimate of the error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/10} = .156$. An unbiased estimate of the error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/9} = .164$. Since the above data gives the exact distribution of $\hat{p}$, the exact error of $\hat{p}$ works out directly from data as

$$\sqrt{\frac{1}{99}\left[\sum_{i=1}^{11} p_i^2 f_i - \left(\frac{\Sigma p_i f_i}{100}\right)^2\right]} = .212 \text{ where } p = x_i/n.$$

These values are different since they have been obtained by different theories. Again, considering that the sample is coming from a finite population of N^1 = 1,000 plots, using Cochran (1953, p.34) an unbiased estimate of the variance of $\hat{p}$ is

$$\frac{(N^1 - N)}{N^1 N(N-1)}\left[\Sigma p_i^2 f_i - \frac{(\Sigma p_i f_i)^2}{N}\right] = 0.000407$$

giving an estimated standard error of .0202. If we use the concept that the sample we have is a simple random sample of 1,000 plants from 10,000 plants,

$$\vartheta(\hat{p}) = \frac{(N^1-N)}{(nN-1)N^1}\,\hat{p}(1-\hat{p}) = 0.000218$$

and hence estimated standard error of $\hat{p}$ = .0148. This smaller value of the error should be expected since the procedure assumed here is quite different from the one used.

Step 3 now we interpret the theory behind the numerical example as follows. We have assumed that the binomial distribution gives a good description of the data. Thus there is a constant probability $\hat{p}$ of a plant getting affected by rust in the tract. Hence assuming sampling from infinite population, the

expected frequencies of samples with 0,1,2,...10 affected are given by the successive terms in the expansion of 100 $(\hat{q} + \hat{p})^{10}$ where $\hat{q} = 1-\hat{p}$. The mean of this distribution is 10 $\hat{p}$. The estimated standard error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/10}$. An unbiased estimator of standard error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/9}$. From sampling theory for finite populations, since every sampling unit contains the same number n of plants an estimate of the variance of $\hat{p}$ is

$$\vartheta(\hat{p}) = \frac{N^1-N}{N^1N} \cdot \frac{1}{(N-1)} \sum_{i=1}^{n} (p_i-\hat{p})^2 f_i.$$

If the plants are taken as the ultimate units and if we are drawing n N elements out of nN^1 elements at random, another estimate of variance is given by

$$\vartheta(\hat{p}) = \frac{(N^1-N)}{(nN-1)N^1} \hat{p}(1-\hat{p}).$$

Step 4 - The different formulae for computation can easily be derived from the above. Using these in the numerical example, the mean proportion of incidence of the disease in the tract is .411 with the most appropriate estimate of error .202. Hence the recommendation to the extension scientists is as follows. At a particular time of crop growth under some suitable conditions rust disease can occur in the wheat crop. Knowing the extent of incidence appropriate spraying can be made to save the plants from disease. The recommendation to farmers can be as follows. Find the proportion of plants in the tract affected by the disease. If it is above danger mark as given by the mycologist resort to spraying. Thus the interpretation of the results goes outside the ambit of statistics.

Other steps in teaching Method I can easily be followed. Method II can also be used for teaching this example.

Example 2: The data used in this example are adopted from Davidson and Beaver (1977) in preference testing for packed food mix labelled 1,2,3,4. they are given below.

Pair (i,j)	Number of repetitions (n_{ij})	Number of preferences for i $a_{i \cdot ij}$	Number of preferences for j $a_{j \cdot ij}$
-	-	-	-
(12)	85	37	48
(13)	85	47	38
(13)	83	52	31
(14)	84	49	35
(24)	82	48	34
(34)	85	45	40

The problem is to teach the non-statistician the generalisation of the Bradley - Terry Model by Dykstra (1960) using the above data.

First teach the computational methods and also the following prerequisites in a simple way.

Mathematics: Matrix algebra; inversion of matrices; logarithms, solutions of linear equations; differentiation; maxima and minima, Lagrangean multipliers.

Statistics: probability and likelihood; maximum likelihood; concept of parameter-free procedures; hypothesis testing; chi-square distribution; efficiency concept in non-parametrics.

These topics can be taught by providing prerequisite courses or independently.

Step 1 - We teach by Method I. The problem here is to obtain the ratings for the four packed food mixes using Dykstra's model. Under the model the estimates p_i (i=1,2,3,4) are obtained from the normal equations set up by the method of maximum likelihood by an iterative procedure. We programmed the iteration and the solutions converged to $p_1 = .26398837$; $p_2 = .32149677$; $p_3 = .21121396$; $p_4 = .20330090$.

Since the number of repetitions are large in the experiment an asymptotic test can be used to test whether p_i's differ significantly among themselves. The test statistics $-2\ell n\lambda$ in this case has been computed as 11.1028 with three degrees of freedom and is hence significant showing that the food mixes differ significantly among themselves. The statistic to test the appropriateness of the model to the data above has been computed as 1.0519 which is distributed as a chi-square with 3 degrees of freedom and is hence not significant showing that the model is quite appropriate. The 3 x 3 asymptotic variance-covariance matrix for testing any contrast among p_i's is:

$$\begin{bmatrix} .00107 & .00028 & .00054 \\ .00028 & .00109 & .00026 \\ .00054 & .00026 & .00133 \end{bmatrix}$$

Step 2 - Now we discuss the theory behind this method. Under Dykstra's model it is assumed that t treatments are tried in t(t-1)/2 pairs, each pair being repeated an unequal number of times. It is postulated that corresponding to the t treatments $T_1, T_2, \ldots T_t$ there exist parameters π_i for T_i (i = 1,2,....t) such that $\pi_i > 0$ and $\sum_{i=1}^{t} \pi_i = 1$. The parameters are further defined with the probability statement that, if X_i generally denotes an observation on a sample of T_i, $P(X_i > X_j) = \pi_i/(\pi_i + \pi_j)$ in the comparison of T_i with T_j, $(i \neq j)$. Under these formulations, the general likelihood under the model becomes

$$L(\pi_i) = \prod_i \pi_i^{a_i} \prod_{i<j} (\pi_i + \pi_j)^{-n_{ij}} \quad \ldots\ldots (5.1)$$

where $a_i = \sum_{j \neq i} a_{ij}$ and $a_{ij} = 2n_{ij} - \sum_{k=1}^{n_{ij}} v_{ijk}$

$a_{ij} + a_{ji} = n_{ij}$ where $v_{ijk} = 1$ if $x_i > x_j$ and $v_{ijk} = 2$ if $x_j > x_i$ in the kth repetition of the pair (ij). Next we test the null hypothesis $H_o: \pi_i = \frac{1}{t}, (i=1,2,\ldots t)$ against the alternative hypothesis H_1: no π_i is assumed equal to any π_j $(i \neq j)$. Then differentiating the natural logarithm of (5.1) with respect to π_i using appropriate Lagrange multipliers and denoting the maximum likelihood estimates by p_i we obtain $(a_i/p_i) - \sum_{j \neq i} [n_{ij}/(p_i+p_j)] = o$ $(i,j = 1,2,\ldots t)$ with the usual restriction $\sum_1 p_i = 1$. For estimating p_i the above equation can be put in the form:

$$p_i = a_i / [\sum_{j \neq i} n_{ij}/(p_i+p_j)] \ldots\ldots\ldots\ldots (5.2)$$

Reasonable first estimates may be obtained by assuming that the p_i's are not too different from each other so that we may assume $p_i = p'$ for all $j \neq i$. Then

$$p' = (1-p_i)/(t-1) \quad \text{and we obtain}$$
$$p_i = a_i / [(t-1) \sum_{j \neq i} n_{ij} - (t-2)a_i] \ldots\ldots\ldots\ldots\ldots (5.3)$$

(5.3) can be used to get first estimates of p_i. These estimates can be substituted into the right-hand side of (5-2) to obtain second estimates. The procedure is iterated till two successive estimates are equal. The number of iterations required under this procedure is small provided the p_i's are not widely diverging. For testing H_o against H_1 the likelihood ratio depends on the statistic,

$B_1 = \sum_{i<j} n_{ij} \log (p_i + p_j) - \sum_1 a_i \log p_i$. Then, if λ is the likelihood ratio $-2\ell n \lambda = 2 (\sum_{i<j} n_{ij}) \ell n 2 - 2B_1 \ell n 10$, is distributed in the limit as a χ^2 with (t-1) degrees of freedom when all n_{ij}'s become large. For testing appropriateness of the model we can use the statistic

$$-2\ell n \lambda = 2\ell n 10 \, (\sum_{i \neq j} a_{ij} \log a_{ij} - \Sigma n_{ij} \log n_{ij} + B_1) \ldots\ldots (5.4)$$

which is distributed in the limit as a chi-square with (t-1) (t-2)/2 degrees of freedom when all n_{ij}'s become large. (5.4) is derived under the assumption

$$P(X_i > X_j) = \bar{\pi}_{ij}$$

and the statistic tests the hypothesis

$$H_o: \bar{\pi}_{ij} = \pi_i/(\pi_i+\pi_j) \quad (i \neq j, i,j = 1,2,\ldots t)$$

against the alternative

$$H_a: \bar{\pi}_{ij} \neq \pi_i/(\pi_i+\pi_j)$$

for some i and j. The precision of the estimated P_i's can be computed by the formulae:

$$\lambda^o_{ii} = (t^2/4) \sum_{j \neq i} n_{ij} \quad (i=1,2,\ldots\ldots t)$$
$$\lambda^o_{ij} = -(t^2/4)\, n_{ij} \quad (i \neq j;\ i,j = 1,2,\ldots\ldots t) \qquad \ldots\ldots\ldots (5.5)$$

The superscript here indictes the null case when $\pi_i = \frac{1}{t}$ $(i=1,2\ldots t)$. Then the asymptotic variance-covariance matrix of the treatment ratings is given by the inverse matrix $[\lambda^o_{ij}]^{-1} = [\sigma^o_{ij}]$ $(i,j = 1,2,\ldots\ldots,(t-1))\ldots(5.6)$.

Efficiency in such experiments can be studied with the help of pairwise precision which gives the sample size N_{ij} which would be required in a two sample test in order to obtain $p_{ij} = est\pi_{ij} = p_i/(p_i+p_j)$ with the same precision as results from a paired comparison analysis. For the pair (i,j) it is given by

$$N_{ij} = 4/t^2 (\sigma_{ii}^o + \sigma_{jj}^o - 2\sigma_{ij}^o) \ldots\ldots (5.7)$$

Step 3 - Compute the first estimates of p_i using (5.3). Substitute the estimates on the R.H.S. of (5.2) to get the subsequent estimates. The iteration is continued till two successive estimates coincide. To test whether the p_i's differ among themselves significantly, compute

$$-2\ell n \lambda = 2(\sum_{i<j} n_{ij}) \ell n_2 - 2B_1 \ell n 10$$

where

$$B_1 = \sum_{i<j} n_{ij} \log (p_i+p_j) - \Sigma a_i \log p_i.$$

Test for significance of this using χ^2 table. For testing the appropriateness of the model compute the statistic (5.4) and test using chi-square tables. Compute the asymptotic variance-covariance matrix of p_i's using (5.5) and (5.6) and pairwise precisions by (5.7). The computations in the numerical example were shown in Step 1. The main statistical interpretation is that the food mixes differ significantly and that the model is appropriate. Interpretation in the subject matter area is that T2 is the more preferred food mix followed by T1, although they do not differ significantly. Comparison among the other foox mixes can be made in a similar manner. Pairwise precision for the pair (1,2) is 156 thereby meaning that in this experiment with 85 repetitions for this pair, the precision achieved is as good as in a single pair experiment with 156 repetitions. Thus paired comparison experiments improve the precision of treatment comparisons due to hidden replications. These results can be given to the food industry for practice.

In the next steps practicals can be given to the trainees and the record of class work kept. The trainees can do an experiment in sensory evaluation, analyse the data and write a report. This makes the training very comprehensive.

SOME REFLECTIONS ON TEACHING NON-STATISTICIANS

In this section we will refer to some dubious points. Should construction of designs be included in the curricula for scientists? The answer to this is possibly 'no'. For teaching collection and compilation of agricultural statistics are the teaching methods discussed in this article necessary? This also has a negative answer. This is not so for teaching research techniques in agricultural sciences and also for teaching statistics minor in universities. Does the training of non-statisticians to some extent bridge the gap between academic statisticians and practising statisticians? The answer to this is possibly 'yes' because some of the practising statisticians are non-statisticians. Should all the non-statisticians trained be assessed? In my opinion all need not be. There is also the fundamental question whether statistics should be taught to non-statisticians at all. The answer is that they should not be taught statistics to the detriment of statisticians.

CONCLUSIONS

Thus the method of teaching statistics to non-statisticians must be application-oriented. They should be trained as to how to present interpretative results. Examples must be given to demonstrate what can be inferred and what cannot be inferred from real data. The teaching method can be the same in each specialised area except in areas as collection of agricultural statistics where the contents are purely descriptive. A course in computer programming is a must for most types of non-statisticians. Field work involving independent designing of data collection and analysis is also a must for most non-statisticians.

For assessing agricultural scientists we have been using the credit-course system of America. The trainees are given quizzes, mid-terms and a final examination according to this system. Maximum weight is given for the final examination in the four point grading system adopted here. Further, there are comprehensive examinations and vivas in the relevant subjects as well as for the thesis work of non-statisticians. Similar methods are adopted for grading non-statisticians in Agricultural Universities and Indian Institutes of Technology. For Junior, Senior and in-service training courses of the IASRI, we adopt the external examination system of the Indian Universities including a "viva voce" examination. Students in these courses are awarded classes. For traineees of All India Services, the courses given are seminar type with neither assessment nor grading. For students joining on short term projects, the project reports are assessed. Over the years we have developed some specialised curricula for the non-statisticians. Teaching materials for such trainees have also been developed. The results of the consolidated work have been published as 'Statistical Methods for Agricultural Workers' and

'Handbooks of Practicals' apart from a number of applicational papers which appear from time to time in the Journal of the Indian Society of Agricultural Statistics.

References

[1] Cochran, W.G. (1953) Sampling Techniques. John Wiley & Sons Inc.

[2] Kempthorne, O. (1980) The Teaching of Statistics. The American Statistician. 34, No.1, pp.17-21.

[3] Mosteller, F. (1980) Classroom and Platform Performance. The American Statistician. 34, No.1, pp.11-16.

[4] Noether, G.E. (1980) The Role of Non-parametrics in Introductory Statistics Courses. The American Statistician, 34, No.1, pp.22-23.

[5] Sadasivan, G. (1966) Statistics and Agricultural Education. Journal of University Education, Delhi University, 5, No.1. pp.19-24.

[6] Tukey, J.W. (1980) We need both expository and confirmatory. The American Statistician, 34, No.1, pp.23-25.

[7] Davidson, R.R. and Beaver, R.J. (1977) On extending the Bradley-Terry model to incorporate within-pair order effects. Biometrics, 33, pp.693-702.

[8] Dykstra, O. Jr. (1960) Rank analysis of incomplete block designs; a method of paired comparisons employing unequal repetitions. Biometrics, 16, pp.176-188.

UTILITY THROUGH COLLABORATION: A PROBLEM ORIENTED APPROACH

D Taylor and G E Skipworth, Teesside Polytechnic

The teaching of statistics to non-statisticians is a well established practice. Indeed such is the success with which the argument for the relevancy of statistics has been put that the practice has been extended to include undergraduate and postgraduate courses not just for engineers, nor only geographers and sociologists, but even such inveterately literary characters as historians! While the latter might be considered an extreme case, there is a common problem to be found in all courses of this nature. Put simply, it consists of convincing students of the relevance and utility of statistics to their particular discipline which in turn requires the conversion of plausible theory into credible practice. The theoretical justification receives considerable attention, especially from the plethora of validating bodies to be found in certain sectors of higher education, but too little time has been given to the more mundane, but no less important, task of actually teaching a course in a manner that successfully converts theory into practice.

The most common practice found in the teaching of such courses is that of the 'invited alien'. Given the absence of staff with the necessary statistical knowledge in the host department (a situation which may be changing slowly), it is not surprising to find that the 'visiting statistician' is called in to do the job. All too often, however, he is left to work in isolation. Lacking the detailed knowledge of the subject that is the prime interest of the students, he becomes an outsider, a once (or twice) weekly performer with a display of statistical pyrotechnics that, depending upon the nature of the audience, arouses either admiration for their elegant irrelevance or deep scepticism for their inhuman manipulation of real facts! This difficulty is compounded by the lack of time commonly devoted to such courses. Despite the importance attached to statistics courses by host departments, there are always more important indigenous courses that crowd the timetable and restrict the visiting statistician to little more than a token appearance. In these circumstances it is not surprising to find that the course becomes a whistle stop tour of some of the better known statistical techniques. If it's week three it must be correlation! Faced with these practical difficulties, noble aspirations quickly turn to unsatisfactory practice. Unsatisfactory for both staff and student as a series of techniques, thinly disguised with examples from the host department's discipline but with limited relevance to the day-to-day work of the students, are presented in a manner that has more in common with forced feeding than genuine scholarship.

Although a general dissatisfaction with the practice of servicing work was an important element in the formulation of the particular course that is the main subject of this paper, there were also more specific considerations that were to the forefront in the construction and evaluation of the course. Indeed, we would argue that the emphasis must be on specific operations rather than on general crusades. It was clear from preliminary soundings that, while the staff mounting the M.A. in Local History were convinced of the usefulness of statistics to their area of study and hence of the value of an introductory course for the students, many prospective students would have no knowledge beyond basic arithmetical skills and, moreover, would exhibit apprehension if not antipathy to any course that involved statistics. In these circumstances it was felt that an introductory statistics course could only succeed if it were clearly demonstrated to the students that it was an integral part of the whole M.A. programme and relevant to work that was taking place elsewhere in the programme. To this end we sought to show that statistical skills were both a necessary and natural part of the historian's 'tool-bag' and that these skills could be applied to problems introduced elsewhere in the students' studies. Utility through collaboration became the motto of the course, which was not only conceived but also put into practice as a joint venture. Statistician and historian would be seen working together and in a way that was mutually beneficial.

It is now appropriate to consider in more detail the Introduction to Statistics course in the M.A. in Local History offered at Teesside Polytechnic. The first point to stress is that each session lasts approximately three hours. Although this is in part the outcome of the fact that the M.A. is a part-time evening only course, a deliberate decision was taken to utilise the maximum time period so that a self-contained and integrated topic could be presented in any one session. The second point to emphasise is that, with the exception of the introductory meeting, each session is devoted to a specific topic and, moreover, a topic that relates directly to other courses in the M.A. programme. Thus, for example, one session is given over to a consideration of the use of measures of association to throw light on the determinants of nineteenth century voting behaviour, a problem that is of major importance on another, essentially historical, course that is devoted to a consideration of regional patterns of political action.

However, to illustrate the approach used on the course we will consider another example in greater detail. The question of changing living standards is one that has generated much argument among historians. One part of this debate has centred on the measurement of real wages as a measure of material well-being. The technique most commonly, and all too often uncritically, used is that of index numbers. This, then, provides the general problem area

for a session. The session is divided into three roughly equal parts, but throughout every session both members of staff - statistician and historian are present and will contribute, though the balance will vary from one sub-period to another. The opening sub-period is led by the historian who firstly outlines the general problem, the range of opinions on the subject and, given the nature of the M.A., the relevance of the local study to the wider issue.

As part of this preview the case for using statistical techniques is also presented. In this instance particular emphasis is placed on the need to generalise, the level of generalisation appropriate to the question under review and, particularly, the limitations of the concept of real wages in itself and in relation to the wider issue of living standards. His second, and equally important, task is to provide basic data and a critical commentary thereon. This last point is doubly important. From the historian's point of view the critical scrutiny of material is a central part of his practice. In this case particular attention is paid to the difficulties of assembling data relating to working class expenditure and consumption patterns that underlay any attempt to construct a cost of living index. Budgetary information for the late nineteenth century is limited. It is not possible to give details of year to year changes, nor is there specifically local material to work on. With regard to local prices the position is brighter. The local press as well as the minutes of the local Board of Guardians provide valuable material but even here there is the problem of using adjusted institutional prices to be recognised. The need to recognise the imperfections of the data, however, is no less relevant to the statistician. Although applied statisticians, rightly, emphasise the practical nature of their work, there is a tendency in teaching the subject to have perfect examples. Understandable though this is, there are certain dangers inherent in such an approach - most notably oversimplification. However, in this course all concerned are face-to-face with imperfect data from the outset. Moreover, the question is constantly posed: is the data of sufficient substance and quality to justify further analysis? Clearly, with the raising of such a question the expertise of the statistician is brought in to the discussion.

The second part of the session, predicated on the assumption that the reply to the above question will be, however heavily qualified, in the affirmative, is led by the statistician. In this he sets out the basic principles underlying, in this case, Index Numbers. This is done via a combination of commonsense and basic arithmetic rather than detailed mathematical proofs. Further he makes clear any requirements of the data demanded by a particular technique (something, clearly, that will vary from technique to technique). Finally, he draws attention to the limitations of the technique itself. All three aspects deserve further comment. First, the presentation of the technique is seen

to be a middle way between one in which mathematical proof is seen to be a prerequisite, inappropriate given the background of the students on this particular course, and one in which students are given a 'recipe book' of formulae, pre-digested rules to be applied without thinking, inappropriate to a course at graduate or undergraduate level. Second, the emphasis on the data requirements from the statistician's viewpoint is felt to be a most necessary part given the propensity of many non-statisticians, and often those most enthusiastic about quantification, to apply statistical techniques in any circumstances. Finally, the consideration of the imperfections of statistical techniques is seen as a necessary antidote to the common belief among non-statisticians that statistics offer certitude in both method and result. As a result of this sceptical approach the earlier question asked of the data is supplemented by another: does the technique provide a satisfactory approach to the problem in hand? This brings the historian back into the discussion and underlines the interplay between the two disciplines.

The third and final sub-period involves a consideration and interpretation of the results obtained from the application of the technique considered in part two to the data introduced in part one. The students are not expected to work out the results for themselves but are presented with a combination of handouts and overhead projections. In the evaluation both statistician and historian contribute to an assessment of both the statistical and the historical significance of the results. It is also at this point that the students can make (and indeed are expected to make) a contribution to the general discussion. It will be readily appreciated that this part of the session, and the related follow-up essay, is both the culmination of what has gone before and the justification for the whole exercise.

Having outlined the approach adopted in the course it is necessary to see to what extent the theory has stood up in practice. So far there have been two intakes of students which provide the empirical basis of our evaluation. It was clear from the outset that the major problem would be one of weighting both within topics in any given year and between groups from year to year. The first intake, a dozen in size, was relatively homogenous in terms of ability, sympathetic to the aims of the course, and well motivated. One or two, significantly non-historians by first-degree, had an elementary knowledge of basic statistics but the majority could be considered, and certainly considered themselves to be effectively innumerate, at the outset at least! The first running of the course revealed two weaknesses. First, the balance between the three sub-periods was difficult to maintain. In particular the third sub-period tended to be squeezed in terms of time and this was further compounded by the fact that students were responding immediately and in some cases to historical problems that they were meeting for the first time. To some extent this was masked

by the enthusiasm of the students; in another direction, it was also offset by the follow-up essay work, though it should be made clear that the essay work was based on only one of the topics, chosen by the student from the range offered on the course. The second weakness, and considered to be only minor in the light of the first running of the course, was the lack of any opportunity for the students to 'dirty their hands' albeit on a small scale. Interestingly, this shortcoming was regretted by the students as well as the staff. In the light of this assessment it was decided for the second cycle to retain the same basic format with two modifications: (i) the deletion of one topic and (ii) the introduction of substantial handouts relating to sub-period two. In these ways it was hoped to allow more time for discussion and thereby meeting our original aspirations.

The second intake, twenty in number, exhibited different characteristics. The range of ability in both general and particular terms was much greater than with the first intake. Equally significant, while there was little difference in terms of overall motivation, there was an element not so much of scepticism but of cynicism about the objectives of the course. While this was overcome for the most part, there remained the two weaknesses identified in the first cycle but intensified by the first characteristic of this group. Superficially, the problem of internal balance was solved in that more time was made available for discussion but a more serious problem emerged, namely bringing those at the lower end of the ability range up to the necessary level of understanding to participate fully and satisfactorily in discussion, and indeed to produce the required standard of written work. This led us to re-consider what had been seen as the secondary weakness from the first cycle, that is the lack of any opportunity for students to 'dirty their hands'. It now seems clear that such an opportunity would be beneficial not only to strong students but also to those weaker students who had struggled to grasp some of the basic concepts introduced in the course. Thus, the introduction of partly worked examples, offering the students the chance of applying by imitation the techniques introduced in the session, should give a firmer grasp of the subject to all students, but particularly to the less able and less sure. Finally, the introduction of a 'dirty hands' element will necessitate a carry-over from one week to the next. While this obviously destroys the unity of each session it is felt that this will be more than compensated for by the firmer grasp of the subject matter and by the greater opportunity for reflection on the results.

The emphasis in the preceding paragraphs has been on the shortcomings of the course in its two presentations. However, it would be wrong to give the impression that the approach has been unsuccessful in overall terms. Given the backgrounds of the students coming onto the course and the limited time

at our disposal the results, as reflected in essay marks and examination performances, have been encouraging. Students with little or no formal knowledge of statistics at the start of the course have finished with not just a grasp of basic techniques but also an appreciation of how those techniques can be applied in a proper manner which will further their work as practising historians. Even in other, less successful cases it is by no means clear that the shortcomings are confined to the statistics course. There are, however, other less formal but no less important measures of success. First, student response - almost unanimously in the case of the first intake, and a clear majority in the second - was positive and encouraging, with many wishing to take the subject further. Second, several students, in the course of their M.A. dissertations, utilised both the specific knowledge gained from the course and the general expertise of the staff. Finally, and no whit less important, the staff response has been favourable! Both authors have found the course not simply novel (though in their direct experience it can certainly claim novelty) but both stimulating and satisfying. This is not to say that there have not been problems and periods of heart-searching but it is a measure of the overall success of the course that these very shortcomings have acted as a stimulus to further refinements and improvements to what is seen as a basically sound way of teaching statistics to non-statisticians. No longer need the statistician be an 'invited alien': genuine collaboration can bring mutual benefits to both staff and students and to the disciplines involved in such an integrated offering.

AN OPEN UNIVERSITY SERVICE COURSE

John Bibby, The Open University

INTRODUCTION

This paper describes an Open University course entitled:

MDST 242: Statistics in Society

This course will be available to Open University students for up to 10 years starting in February 1983. However, in addition to being anecdotal and specific, the paper also raises general issues and questions concerning the teaching of statistical methods to nonspecialists.

DEFINING THE CURRICULUM

In his introductory speech to this conference, Vic Barnett has already raised the question

WHY TEACH STATISTICS?

He has, indeed, also provided tentative answers. However, a full answer will depend upon our attitude to a further crucial question, namely:

WHAT IS STATISTICS?

This can be answered in several different ways. The conference has discussed various competing definitions of statistics, and not all teachers can afford the luxury of questioning the curriculum - this is often defined by our examining body, head of department, departmental tradition, or other institutional constraints. In such situations one may easily unquestioningly accept or rationalise the status quo, and avoid returning to first principles.

However, many teachers do have considerable latitude in determining their syllabus. Some few are prepared to stretch institutional constraints to their limit. This is certainly the case at The Open University - possibly excessively so. We have the power (and problems) associated with rethinking the curriculum from first principles. This power must of course be exercised responsibly, although it is not always easy to know what such responsibility implies, when traditions and technology in our subject are changing as rapidly as they are today.

CONSTRAINTS

Nevertheless, way back in 1979 when planning for the course MDST 242 started in earnest, we did face certain constraints. Paramount among these were the following:

(a) We were concerned with an individualised learning system - students would study alone, at homes scattered throughout the United Kingdom, with minimal face-to-face tuition. The teaching material had to respond to this.

(b) We were aiming at a course which was non-mathematical. "Statistical Literacy", "Education for Everyman" and "Citizen-Arming" are watchwords which describe the approach we were aiming at. Moreover, we wanted to attract students from all six Open University Faculties. Therefore our applications could not be too discipline-specific.

(c) Our resources were finite. That is not to say they were inconsiderable - we could count on 15-20 person-years of academic person-power, and also other Open University resources such as text, audio-tape, television programmes and computing. (However, we decided not to include any computing, for various reasons including obsolescence-phobia).

(d) The most important constraint however was student time. Our course should take around 160 hours of study time in total. This may be compared with a lecture course of 40-50 hours (assuming 2-3 hours of back-up work per hour of lecturing).

(e) Moreover, we were conscious (and still are) that this could be an extremely influential venture (for better or worse), for statistics teaching across far greater horizons than just the Open University. We have 100 or so tutors scattered across the country, and many of these also teach at other colleges and universities. Other institutions, we hope, will consider using our work. Also, in the ten years of presentation (1983-1992) our course would reach up to 20 thousand students - a figure which is several orders of magnitude bigger than many other statistical ventures. Thus we sought a reasonable compromise between the requirements of tradition and those of good statistical practice.

FIGURE 1: Tensions to be resolved

(1)	Statistics as an "exact science"	⟷	Statistics as a "social product"
(2)	a few "case studies"	⟷	many "examples"
(3)	descriptive/ exploratory statistics (EDA)	⟷	inference

TENSIONS

In the course of preparing our material, many decisions have had to be made. However, three key areas may be identified as encapsulating dilemmas which many teachers must face. These are summarised as the "tensions" shown in Figure 1.

(a) First is the tension between statistics as an "exact science" (objective, rigorous, culture-free, technique-oriented) and statistics as a "social product" (produced as the outcome of human responses to a wide variety of conflict-laden situations).

In resolving this tension we resolved that we did wish to teach technique (this was not to be simply a "statistical appreciation" course), but we also wished to sensitize our students to the technical and non-technical limitations of technique.

(b) The second tension was between presenting a few "case studies" on the one hand or many different examples on the other hand. The first course of action would allow us to go into material in some depth. However, it might unwittingly provide hostages to fortune, in view of the wide variety of subject-backgrounds from which our students would be drawn. Also, the second option could be too disconnected.

Therefore we resolved to focus on a few "topic-areas", from each of which a variety of applications would be selected. The topic-areas should involve students as citizens (producers, consumers, parents, etc.) rather than as adherents of a particular discipline. Thus we focussed on the following four main topic-areas:

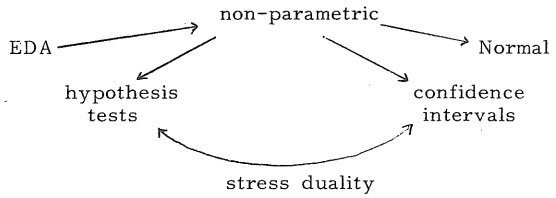
Economics
Education
Health and Medicine
Energy

Moreover we sought - somewhat optimistically as it turned out - to teach only techniques which "arose naturally" out of these topic-areas.

(c) The third tension referred to in Figure 1 relates to the balance between exploratory and data-reductive techniques on the one hand (à la Tukey/ Ehrenberg) and classical inference on the other hand. This tension is certainly extremely topical now. However, we believe that EDA is no transient bandwagon. It reflects a long-standing unease concerning the role of probability-based methods. The use of these methods often carries a certain lack of conviction. Moreover, the philosophy and mathematics on which they are based can be obscure and are often positively intimidating.

However, they survive faute de mieux, and clearly are relevant in certain well-defined situations such as sampling and randomised experiments.

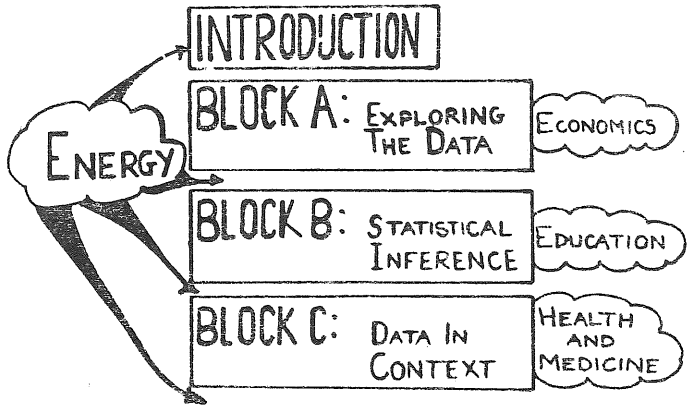
FIGURE 2: Our approach linking EDA with conventional inference



EDA is more generally applicable, but is less precise in the answers it gives. Thus, we sought a balance between exploration and inference. The balance we found is summarised in Figure 2: the course starts with EDA, and then approaches inferential statistics via a limited range of non-parametric methods.

We start with the sign test for the median, and its associated confidence interval. This has the advantage of mathematical simplicity - the pattern of Pascal's triangle is crucial here - which allows us to concentrate on teaching statistical points such as the duality between hypothesis testing and confidence intervals. The Normal distribution arises via the large-sample properties of primarily non-parametric methods - we are currently planning a television programme on this.

FIGURE 3: Structure of MDST 242



OUTLINE OF COURSE

Figure 3 outlines the course structure. Further details of structure and content appear in a Course Description available from the author. However, the main features are the following:

The course has 3 main Blocks, each consisting of 4 units. These may be summarised as follows:

Block A covers EDA, focussing on topics from economics (prices, wages etc.). This block also covers line-fitting using Tukey's resistant method, price indices and survey methods.

Block B introduces inference, with educational examples providing the vehicle. We focus on the limited menu of techniques shown in Figure 4.

Block C consolidates the techniques of exploration and inference. It also discusses notions of experimentation (drug-testing) and public policy-making (smoking).

FIGURE 4: "Menu" showing main inferential techniques

MDST 242

M E N U



Hors d'oeuvre

Triangle de Pascal

Starter

Sign test/confidence interval for median

Main Course

Morceau de Mann-Whitney
French fried Wilcoxons
Crême a la Chisquared

Dessert

z (No t)

Notes:

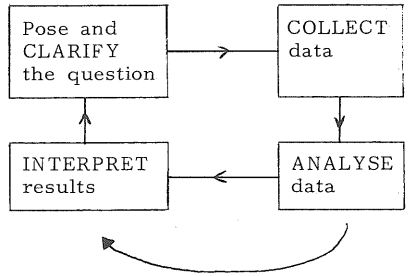
(1) Each confidence interval is served beside along and between the associated hypothesis toast, together with a strong gravy of duality.

(2) No one-tailed beasties.

(3) 5% significance only

Interleaving the 3 main blocks are 4 further units on energy statistics, using especially some of the work of the Energy Research Group at the Open University. These units are primarily "stock-taking" units : the student watches 2 television programmes, submits assignments, draws breath, and proceeds to the next block.

FIGURE 5: The model-building loop



STATISTICS AS MODEL-BUILDING

The "model-building loop" shown in Figure 5 is fundamental to the portrait of statistics which this course will paint. The loop has many directions of relevance:

* it ties in with the "Data = Fit + Residual" (DFR) equation of exploratory data analysis
* it illuminates the non-statistical problems associated with hypothesis-formulation
* it emphasises the iterative genesis and tentative status of statistical knowledge, thus supporting the notion of "statistics as a social product". (The debate in the sociology of science between Popper and Kuhn has many interesting parallels here).

The model-building aspect of statistics has scarcely been emphasised at this conference. This is not surprising, because the model-building viewpoint can query the very rationale of the conference's existence. It raises the question of whether statistics should be separated from mathematics at all. Applied mathematics properly defined is little more than deterministic model-building. By contrast, statistics is model-building done properly i.e. stochastically. This is a viewpoint which MDST 242 aims to emphasise.

However, while many courses concentrate on the COLLECT and ANALYSE boxes of the model-building diagram, our course (and in particular the television

programmes) go to great lengths to illustrate the INTERPRET and CLARIFY boxes. Thus in TV3 (which was shown at the conference), fully half the programme is concerned with setting up the rationale for a line-fitting strategy in one particular context (the installation of cavity wall insulation). By contrast, the technique of line-fitting is correspondingly underplayed. (We do, however, find audio tape a very suitable medium for the teaching of techniques: it allows the student to stop and replay, self-assessment exercises can be interspersed; also technical details can be given a suitably low profile).

THINGS WE DID NOT INCLUDE (BUT MIGHT HAVE DONE)

All honest teachers learn from experience. Speaking personally, the experience of preparing MDST 242 has taught me many things. Predominant among these has been the joy and excitement of working with a group of lively-minded colleagues who are primarily concerned with applications, rather than theory. Despite occasional frustrations, preparing this course has been fun : and fun is a factor which should not be underplayed in any statistics course. (I only hope that the students enjoy the course as much as I have done; and also that they learn something from it!).

However, if given the opportunity to prepare the course again, there are several alternative possibilities which I would want, at the very least, to consider seriously.

(a) Statistics as a Social Product

This concept was referred to briefly in the section "Tensions". However, there remains an insidious force within our subject which pushes relentlessly towards technique. The tendency is enhanced by the fact that "statistics as a social product" remains an amorphous and ill-formulated concept : this is seen as weakness in a world where precision is a sign of strength. Nevertheless, I would wish that we had stood firmer against the tide of technique; although I am unsure whether a firmer resolution would actually increase our chance of success.

(b) Inequality as a Topic

Our choice of "Education" as a topic-area may have been misplaced, although undoubtedly it will prove to be popular with Open University students, many of whom are teachers. "Inequality" may have been a better choice - indeed inequality arises naturally in economics, in education, and also in health. However, a more explicit recognition of inequality as a theme may have been preferable. (Indeed, inequality and poverty are world-pervasive issues - a point which has not been lost on us concerning the world-wide marketability of our course, especially in developing countries).

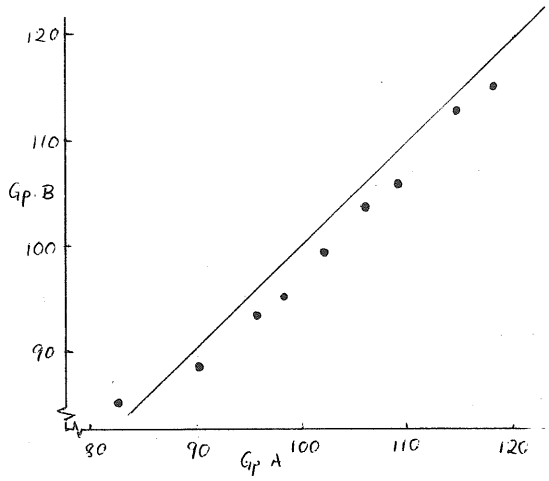
(c) Median Polishing

This technique, introduced by Tukey, is a non-parametric counterpart of analysis of variance. However, it is much, much simpler. The analysis of a 5x3 table can even be used to introduce the notion of median. Our early drafts proposed this, but unfortunately median polishing was subsequently winkled out by pressures of time and space, as well as by the feeling that the technique is as yet insufficiently used and understood.

FIGURE 6: Back-to-back stemplot of test scores achieved by members of two groups

GROUP A		GROUP B
	7	5
32	8	58
950	9	358
9621	10	34
974	11	35
	12	6

FIGURE 7: Quantile plot based on Figure 6



(d) Quantile plots

This method of comparing two batches is informal, informative, and relatively easy to use. As an example, the data on test scores presented by Professor Noether at the conference leads to the back-to-back stemplot shown in Figure 6, and to the quantile plot shown in Figure 7. If the two batches were identical, then each point would lie exactly on the 45° line. All other patterns are to be related to this bench mark or "FIT". Common patterns are the following:

* shift of 45° line : indicates difference in level
* tilt of 45° : indicates difference in scale
* "tailing off" : indicates something about the tails, outliers perhaps.

The plot shown in Figure 7 seems to indicate a difference in level between the two batches (Batch A greater than Batch B), with a slight reversal in the left-hand tail. However, there are problems associated with the definition of quantiles in small discrete batches, and further research may be necessary on the real-life properties of quantile plots before the technique is suitable for introduction into elementary statistics courses.

(e) Use of stemplots for easing calculations

Apart from its use in calculating order statistics, the elementary grouping provided by stemplots also provides convenient "short cuts" for calculating several standard statistics, in particular the mean, standard deviation and Mann-Whitney U statistic. The short cuts are similar to those provided by the "use of working mean" methods and are particularly useful when calculators are not available.

FIGURE 8: Use of stemplot for calculating the mean

s_i		f_i	$s_i f_i$	t_i
7	5	1	7	5
8	5 8	2	16	13
9	3 5 8	3	27	16
10	3 4	2	20	7
11	3 5	2	22	8
12	6	1	12	6
			104	55

Note: f_i represents the frequency at each level of the stem, and t_i represents the total of the leafs. (For further details see text).

(i) Calculating the mean

The mean is easily calculated from the stemplot shown in Figure 8, without using a calculator. The stemplot decomposition of a value x_i is

$$x_i = 10s_i + \ell_i \qquad (*)$$

where s_i is the "stem value" of x_i and ℓ_i is its "leaf value".

In Figure 8, for instance, the stem value of 75 is 7, and the leaf value is 5.

The rest of the calculation is now evident. We have

$$\Sigma x_i = 10 \Sigma s_i + \Sigma \ell_i.$$

Now if f_i represents the frequency at level i of the stem, and if t_i represents the total of the leaves at this level, then

$$\Sigma s_i = \Sigma s_i f_i$$

and

$$\Sigma \ell_i = \Sigma t_i.$$

(Some liberties have been taken with notion here, in the interests of clarity and simplicity).

Therefore

$$\Sigma x_i = 10 \Sigma s_i f_i + \Sigma t_i.$$

The righthand side involves fewer and simpler additions than the lefthand side. The data in Figure 8 gives

$$\Sigma x_i = (10 \times 104) + 55 = 1095,$$

a value which can easily be checked using conventional methods. The mean is now easily calculated.

(ii) Calculating the variance

The stemplot decomposition described above leads to several alternative methods of calculating the variance. The most obvious of these is the following:

From (*) we deduce that

$$x_i^2 = 100s_i^2 + 20s_i\ell_i + \ell_i^2.$$

Therefore $\Sigma x_i^2 = 100 \Sigma f_i s_i^2 + 20 \Sigma s_i t_i + \Sigma \ell_i^2$.

FIGURE 9: Calculation of variance based on Figure 8

s_i	f_i	t_i	$f_i s_i^2$	$s_i t_i$	$\Sigma \ell_i^2$
7	1	5	49	35	25
8	2	13	128	104	89
9	3	16	243	144	98
10	2	7	200	70	25
11	2	8	242	88	34
12	1	6	144	72	36
			1006	513	307

A worked example is shown in Figure 9. The first three columns are taken from Figure 8. Using the above formula gives

$$\Sigma x_i^2 = (100 \times 1006) + (20 \times 513) + 307 = \underline{111,167}.$$

The final term, involving ℓ^2, is relatively unimportant in this calculation, and may even be omitted without affecting the result unduly.

(iii) Calculating the Mann-Whitney statistic

The Mann-Whitney statistic U is used for comparing two groups. It calculates for each member of one group the number of members of the other group which are less than it; U is then the sum of these numbers.

The stemplot provides an easy way of calculating U, and even easier ways of calculating upper and lower bounds. The method is illustrated in Figure 10.

FIGURE 10: Calculation of U based on Figure 6

Stem value	Group A cum.freq. (d)	Group B freq. (f)	fd	fd^{-1}	Corrections
7	-	1	-	-	-
8	2	2	4	-	4
9	5	3	15	6	4h
10	9	2	18	10	4
11	12	2	24	18	1
12	12	1	12	12	-
			73	46	13h

Note: "h" means "half"

The approximation with the clearest rationale is shown in the column headed fd^{-1}. Here each frequency from Group B is multiplied by the number of observations in Group A which are at a lower value on the stem (i.e. higher up the stemplot, because low values are at the top of the stemplot). The total of this column represents a lower bound on (i.e. is less than) the required value U.

An upper bound can be calculated as the total of the column headed fd. Alternatively the values fd^{-1} can be "corrected" by looking at the individual leaves at each level on the stem, one at a time. By referring back to Figure 6 the corrections shown in Figure 10 may be derived. Adding in these corrections we find the value of U is 59h or $59\frac{1}{2}$. ("h" is a convenient shorthand for "half". It is simpler, quicker to write, and less dominating than "$\frac{1}{2}$").

A very close approximation to the value of U can be obtained by averaging the totals obtained in the columns headed fd and fd^{-1} (7 and 46 in this case).

CONCLUSION

Any course which is innovative is bound to have weaknesses as well as strengths. We hope to monitor feedback from students and tutors on our course, in order to improve its presentation in subsequent years.

However, it is also to be hoped that some ideas may be found here which merit application by other teachers from outside The Open University. The author would be very interested to hear of such applications, and of the results.

THE TRAINING OF SERVICE TEACHERS

R W Madsen, University of Missouri

Those of us who are attending this International Conference on Teaching Statistics no doubt have as great or greater interest in teaching than do many of our colleagues. Certainly as a group we have many years of teaching experience among us and many of us, over the years, have developed good teaching skills and techniques. I believe that most of us realize that we can still improve our teaching and that is one reason that we have come to this conference: so that we can learn some teaching ideas that are new (at least to us).

How have we learned our own teaching skills? Most of us have learned them from the "best teacher". At least we can say that we have if the cliche that "experience is the best teacher" is true! It is a well-known fact that, at least in American colleges and universities, most professors are highly trained in the subject matter of their chosen discipline but receive little or no formal training in how to teach that subject matter. These professors (including most of us) have learned teaching by experience and from informal observation of our own professors when we were students. This is quite a contrast to the training received by elementary and secondary school teachers who are required to take 40 or more semester hours of course work in areas such as Oral Communication, Psychology of learning, Psychological Development of the Adolescent, Education of Exception Student, etc. Many of us have in fact learned over the years to become reasonably good teachers, but it has taken more time than it might have had we received some formal training.

If it is true that formal training in teaching would be beneficial for individuals who have completed Ph.D. programs and who are attempting to teach at a college level, then such must be equally true for individuals who are only beginning their graduate education. In many American universities introductory level courses in many disciplines are taught at least in part by graduate teaching assistants. It is the purpose of this workshop to discuss together how these teaching assistants might be trained to teach.

Most of the teaching done by graduate teaching assistants is in service courses. Bancroft (1972) described service courses as "courses which are provided by one or more departments for students majoring in a different department or departments". Much has been written on the mode of teaching service courses. Fiscal restrictions generally dictate one of two models: large lectures taught to 300-500+ students by an experienced instructor (Sherr, 1981) with

perhaps recitation-discussion sections of 30-50 students conducted by teaching assistants or smaller lectures taught to 30-50 students by teaching assistants. (Most of us would prefer to have smaller classes to allow for student interaction and to have these classes taught by experienced faculty but that is, unfortunately, not a fiscally admissible option!). The teaching of service courses, whether by teaching assistants or regular faculty, is often more difficult than teaching courses to our own majors because many of the students in the class are "required" to take the course and are not taking it because they want to. This difficulty makes it even more important to have trained mature instructors. (See Federer, 1980).

Havign discussed the need for the training of teachers of service courses, especially when these teachers are graduate teaching assistants, we might address the question of who should be involved in this training. The following quotation may give some idea: "Can one blind man be guide to another? Will they not both fall into the ditch? A pupil is not superior to his teacher; but everyone, when his training is complete, will reach his teacher's level..." (Luke 6: 39-40 NEB). Using this as a guide, we should have our best and most interested teachers lead the training of new teachers.

We have discussed the "why" and "who" of the training of teachers and now turn our attention to the question of "when" and "how". To this end we will first discuss what has been done in the Statistics Department at the University of Missouri-Columbia. We will also freely admit that this program of training, which has evolved over a period of six years, has not nearly achieved perfection. Even if it were a perfect program for us, it would not be a perfect program for you. It is our desire to interact with each other so that we might each find ways of making our training programs more effective. What follows is an outline of topics we have discussed with new teaching assistants and abbreviated versions of handouts given to them. In many cases questions are raised on the handouts, but the answers are left for the teaching assistants to fill in based on discussions within the group. We should also add that we have a small department (about 10 full time faculty) so we have had as few as three and only as many as eight in a group at a time. The amount of time devoted to each session is 30-50 minutes.

Session I: Keys to Effective Teaching

This session (as well as Sessions II and III) are conducted prior to the first day of class. We want the teaching assistant to think about their new role as "teacher" rather than "student".

Handout I

1. Understand what it means to "teach".
   (a) To show how to do something.
   (b) To provide with knowledge and insight.
2. Know the material to be taught.
   (a) Know what will be taught in the entire course.
   (b) Know answers to problems assigned.
3. Know your students.
   (a) Know the majors so that you can use applications of interest.
   (b) Know as many of the students as possible by name.
4. You must teach so that the students can learn.
   (a) You must do your part and encourage the student to do his part.
   (b) Motivate the student by showing the need for and applicability of probability and statistics.
   (c) Students can be encouraged to learn, but cannot be forced to learn.
   (d) Try to view your teaching from your student's point of view.
5. Involve the students in the learning process.
   (a) Ask them questions.
   (b) Allow them to ask questions.
   (c) Have them participate in generation of data and examples.

Session II: First Day Tips

After more than 14 years of teaching I still get nervous on the first day of class. In addition to new faces, the first day activities are quite different than on other days. Students may be in the wrong class, have the wrong text, etc. It is helpful for the new teachers to know explicitly things that they should or shouldn't do.

Handout II

1. Expect to be somewhat nervous, especially if this is your first day of teaching.
2. Locate the room in which you will teach in advance.
3. Begin and end this and all classes on time (ideally within 1 minute of the scheduled time.
4. Give information to the students.
   (a) Class name, number, and section.
   (b) Prerequisites.
   (c) Your name, office location, phone.

5. Get information from the students.
   (a) Name (to check against class roll).
   (b) When prerequisite was taken.
6. Give information about course mechanics.
   (a) Homework policy.
   (b) Exam schedule.
   (c) Grading
7. Refer any out-of-the-ordinary question to a faculty advisor.
8. Be prepared to give a regular lecture on the first day.

Session III: Lecture Preparation

One of the best articles I have seen written on the teaching of statistics is by Mosteller (1980). He writes "... I plan to bring out a few ideas that I have found useful in teaching. Although no two teachers are alike, I have often found devices used by others adaptable to my own style... I shall speak as if to new teachers eager for my advice". It is in this spirit that I present my ideas on lecture preparation to new teaching assistants.

Handout III

1. Prepare in advance.
   (a) Do not wait until the hour before class.
   (b) Prepare about 1/3 to 2/3's of the next lecture in addition to preparation for the present one.
   (c) News articles, examples from the literature, etc. should be filed when found for future use.
2. Follow a schedule/outline.
   (a) Use departmental guidelines.
   (b) Construct your own.
3. For each main concept use a three step method.
   (a) Motivate
      (i) Describe a realistic problem which needs to be solved. Problem used should be of interest to your students.
      (ii) Clearly state basic assumptions including simplifying assumptions.
      (iii) Leave the problem unsolved at this point.
   (b) Present Theory
      (i) Make transition from the particular motivating problem to a more general setting.
      (ii) Use notation consistent with the text. If the notation is at variance with standard notation, make note of the fact.
      (iii) Use appropriate mathematical level. You should neither over-simplify nor be overly rigorous.

(c) Illustrate the Theory

(i) Solve the problem raised in part 3(a). Identify parameters. Explain where and why simplification was necessary.

(ii) Illustrate mechanics. Work out mechanics in some cases and/or give students an opportunity to try on their own. (Be careful to make good use of class time on this. Calculations should take only 1 to 2 minutes or less).

(iii) Try to provide insight by using a word problem. Explain how you check conditions to know that the theory applies.

4. Prepare notes of adequate detail.

(a) Write enough so that you can follow your own work.

(b) Work out numerical answers in advance.

5. BE ORIGINAL.

(a) Do not use examples directly from the text (except in special circumstances)

(b) Modify your own notes from semester to semester.

Session IV: Lecture Delivery

On the assumption that it is not possible to assimilate a great deal of information in a short period of time, after the first three sessions we meet once a week for a teaching seminar. One of the important topics to cover early is that of lecture delivery. Mosteller (1980) also makes some pertinent comments on delivery.

Handout IV

1. Spoken Communication

(a) Speak to the class. Make eye contact with each person.

(b) Speak clearly. If you have an accent, be aware of it and make extra efforts to overcome it.

(c) Speak loudly enough so that you can easily be heard in the far corner of the room.

(d) Avoid a monotone voice.

2. Written Communication

(a) Write legibly. Print if your writing tends to be sloppy. Also write large enough so that writing can be read from the back of the room.

(b) Be careful about location. Do not write too low on the board or too far from the middle.

(c) Write down all important information. Most students will write in their notes only that material which is written on the board.

(d) Make use of graphs and diagrams. Use colored chalk as appropriate.

3. Nonverbal Communication

(a) Your own enthusiasm about teaching and about the subject matter is important. It has a strong effect on student attitude toward the course.

(b) How you say things, e.g. your tone of voice and attitude, may be as important as what you say to the student.

Session V: Exam Construction and Grading

There are two schools of thought on the administration of examinations for multisection courses. One is to have a common departmental exam for all sections and the other is to have different exams made up and administered by the individual teachers. In our department we do some of both, so we want to give our new teachers some help in knowing how to construct exams. Since good multiple choice exams are harder to construct (although easier to grade) we have concetnrated on the written exam format.

Handout V

1. First steps

(a) Determine in advance the number of exams and the relative weights of the exams.

(b) Announce the date of the exams at least one week in advance.

(c) Tell the students what material they are responsible to know.

2. Characteristics of "good" and "bad" exams.

(a) As perceived by the students.

(b) As perceived by the instructor.

3. Exam construction

(a) Make a list of the various topics and concepts to be covered.

(b) Write the exam remembering the following:

(i) Give explicit instructions.

(ii) Cover as many topics as possible within time constraints.

(iii) Use some of the following types of questions.

1. Definitions.
2. Routine mechanical problems.
3. Application or word problems.

(iv) Questions should have the property of

1. Being legible.
2. Having different parts of questions independent.
3. Being unambiguous.
4. Giving point values.
5. Having appropriate mathematical level.

4. Proctoring of exams.

(a) Discourage cheating by setting up appropriate physical conditions.

(b) Discourage cheating by your physical presence.

5. Grading exams.
   (a) Promptly.
   (b) Give partial credit consistently.
   (c) Show where errors are made.

6. Learning from the exams.
   (a) You should make note of which questions are too hard or too easy.
   (b) Go over the exam questions, especially the difficult ones, so that the students can learn from their mistakes.

Session VI: Involving Students in the Learning Process

One of the chief advantages of a small number of students in a class is that there is opportunity for teacher-student interaction. If we do not make use of these opportunities we are, I believe, doing the students a disservice. Mosteller (1980) states "Gathering relevant data in class and then analyzing them with the help of the class gets the students involved because we are analyzing their data and they can contribute to the analysis". Hansen (1980) makes a similar statement: "Students are more interested and attentive when they are involved in the class. In applied statistics class[es] which discuss data analysis, a physical demonstration is one good way in involving the students".

Handout VI

1. Allow and encourage students to ask questions:
   (a) On homework or previous lectures.
   (b) On the current lecture (at appropriate points in time).

2. Ask the students questions.
   (a) Some questions may be rhetorical.
   (b) For non-rhetorical questions allow the students time to think.
   (c) Have students work out simple computational practice problems.

3. Have students participate in generation of data and conduct simple experiments.
   (a) Let individual students perform simple experiments. (Coins, dice, cards, etc.).
   (b) Let the entire class participate in some experiments and then summarize the results. This approach can readily be used in studying binomial, negative binomial, Poisson, or uniform distributions as well as for the Central Limit Theorem and random sampling.
   (c) Use the class as a source of data.
      (i) Can be useful for descriptive statistics, study of proportions (binomial), distribution of random variable, etc.
      (ii) Has positive effect on class attitude in that people (frequently) like to talk about themselves and are curious about others in the class. (E.g. question about family size, state in which they were born, birthday problem, etc.).

      (iii) Can be negative if you ask the wrong question. (Don't ask about low GPA's, divorced or alcoholic parents, or other potentially embarrassing questions).

Session VII. Use of Teaching Aids

The topic of teaching aids or teaching materials receives much attention. Several sessions at this conference have been devoted to the topic, the American Statistician has a section entitled "Accent on Teaching Materials", and the journal Teaching Statistics often contains articles on teaching aids. New teachers of statistics often need to be encouraged to use teaching aids and should be told how to incorporate them into a lecture.

Handout VII

1. What are teaching aids?

2. What are some purposes of teaching aids?
   (a) Gain and hold attention.
   (b) Add interest.
   (c) Increase retention of knowledge through association.
   (d) Show applicability to real world situations.

3. What are sources of teaching aids?
   (a) Commercially prepared aids.
   (b) Adult or children's games.
   (c) Newspapers, magazines.
   (d) Homemade materials.

4. Types of teaching aids.
   (a) Visual aids
      (i) Dice, cards, coins, marbles, etc.
      (ii) Colored chalk.
      (iii) Cartoons.
      (iv) Packages of products, examples of defectives.
   (b) Audio aids. Tape recordings of sports, speeches, news, weather, etc.
   (c) Printed aids.
      (i) Newspaper and magazine articles. (Separate handout on these).
      (ii) Mail advertisements.
      (iii) Books - especially from courses students will take (or have taken).
   (d) Projected aids.
      (i) Overhead projector with transparencies.
      (ii) Slides, movies.

5. When should teaching aids be used.
   (a) Whenever appropriate, but most frequently as part of introductory motivating example.

(b) As often as possible. Most people err in using too few rather than too many teaching aids.

Session VIII. Teaching Aids - Newspapers and Magazines

One of the most convincing ways of showing students the applicability of statistics to everyday life is to use common everyday sources. Among the most common everyday sources are newspapers and magazines. In view of the importance of these items as teaching aids, a separate session is devoted to them.

Handout VIII

1. How can suitable material be found?
   (a) Purposefully.
   (b) Serendipitously.
2. Where should materials be kept?
   (a) Specific files can be made.
   (b) Clippings can be placed in the text at the appropriate place.
   (c) They should not be left in a newspaper or magazine to be cut out "later".
3. Can material be "re-used"?
   (a) Not in the case of dated or easily replaced materials.
   (b) Possibly in the case of "classic" examples.
4. For what subject matter will these aids be most useful?
   (a) Probability: precipitation probabilities; odds for sports events:
   (b) Misuses, abuses of statistics: Proper use of mean vs median; incorrect or misleading labels on graphs (especially in advertisements)
   (c) Descriptive Statistics: graphs, tables, histograms, pie charts, etc. (especially abundant in news magazines).
   (d) Estimation (point estimtes, interval estimates, question of sample size): Use proportiosn given in Gallop Poll etc. (Not strictly random, but probably representative). Also point out difficulties with "man-in-the-street" polls, etc.
   (e) Hypothesis testing (especially about proportions): Same as (d).

Session IX. Student Evaluations

One of the most controversial topics among educators is that of the use and value of student evaluations. Rencher (1979) states that "Widespread disagreement among faculty members about the merits of student ratings would seem to point up the fact that the concept of teaching effectiveness is a complex phenomenon to evaluate". I believe that part of the reason for the controversy is the fact that administrators sometimes use these evaluations, to greater or lesser degrees, in evaluating teaching. Our purpose is not to discuss the appropriateness of such use. I do believe, however, that individuals who desire to improve their teaching can make use of student evaluations for self-improvement. I encourage our teaching assistants to view them primarily in this light.

Handout IX

1. Discussion of validity of student evaluations.
2. Discussion of factors which might affect responses.
   (a) Time of class.
   (b) Required or elective course.
   (c) Student grades.
3. A look at the evaluation form.
4. Use of evaluations.
   (a) By faculty advisor.
   (b) By teacher.
      (i) Trimmed mean approach.
      (ii) Mechanics.
      (iii) Attitudes.
      (iv) Language or accent problems.

References

[1] Bancroft, T.A. (1972) On Teaching of Service Courses in Statistics, The American Statistician, 26, 14-16.

[2] Federer, Walter T. (1980) Pressures For and Against Effective Teaching, Proc.Sec. on Statistical Education, American Statistical Association, 28-30.

[3] Hansen, James L. (1980) Using Physical Demonstrations When Teaching Data Analysis, Proc.Sec. on Statistical Education, American Statistical Association, 61-62.

[4] Mosteller, Frederick (1980) Classroom and Platform Performance, The American Statistician, 34, 11-17.

[5] Rencher, Alvin C. (1979) An Analysis of Student Evaluation of Instruction and Relationships Among Salary, Chairman's Ratings, and Student Ratings of Faculty, Proc.Sec. on Statistical Education, American Statistical Association, 1-5.

[6] Sherr, Lawrence A. (1981) Teaching the Introductory Course to Large Classes, Proc.Sec. on Statistical Education, American Statistical Association, 8-12.

TOPIC 5: DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL

ORGANISER: MR P HOLMES

DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL

A R Hoffer, University of Oregon

It is clear to all of those who are attending this conference that statistics should be an important component in the education of all people and that statistics education should begin with young children and be carried on continuously throughout the years of formal education. However, a rather large percentage of teachers, especially at the school level, have not themselves experienced an adequate statistics education and hence lack knowledge about statistical topics per se as well as what types of statistical activities would be appropriate for their students and how to integrate such activities into their course of study. It is necessary to attack this cycle of ignorance from several directions: (1) the initial training of teachers should have an adequate statistical component; (2) there should be a variety of inservice programs for practising teachers at colleges, within school systems, and for teacher self-study; (3) curriculum developers and directors should more enthusiastically incorporate adequate statistical activities in their programs and courses of study; (4) there should be more active research into questions of statistical didactics; and (5) there should be developed more effective statistical lessons and learning sequences for students. Concerning the item of research most of the available studies deal with probability rather than other aspects of statistics.

There is a nice analogy between statistics learning and computer science learning. It is insufficient just to talk about all the potential uses of computers, and it is not enough to use only pre-prepared programs. Students need to write their own programs as well as to pose and solve their own problems with computers. Likewise one needs to "do" statistics. Carrying the analogy further, there are people, including middle school teachers, who are afraid of computers just as there are those who are afraid of statistics, in so far as they understand the topic. Hence to improve statistics education at the school level we must work to help teachers gain a positive attitude toward the subject as well as to increase their knowledge.

There are two more obstacles of which one must be aware as we develop statistics teaching materials. First, we must realize that statistics makes demands for sophisticated reasoning processes to which most people are unaccustomed. The usual activities at the school level deal with arithmetic, algebraic and geometric computations. Hence it is a fairly simple matter to instruct students to distinguish between different types of averages or compute means and so forth. It is more difficult to develop skills in collecting and organizing data,

analyzing probabilities, testing hypotheses, and more generally thinking critically and being comfortable with uncertainty. A second obstacle deals with the heavy demands that are placed on elementary school teachers. They are usually responsible for covering many subjects besides mathematics and statistics. (As we note below this can be turned into an advantage). They must find time and energy to devote, for example, to the daily situations of students with varying abilities, to the demands of parents and school administrators, to the results of research, new learning theories and teaching strategies, and to the new teaching materials in the various disciplines. It is folly to expect that simply by developing more teaching materials at the school level will we be assured that more statistics learning will take place. The materials must be of the type that are readily acceptable by teachers. Indeed a proof of this situation may be evidenced by the availability over the past twenty years of numerous statistics teaching references with little or no change in school programs.

Statistics is a dynamic endeavor, forever changing. By their very nature statistical worksheets and textbooks with pre-prepared data are outmoded by the time they are printed. The real import of statistics learning rests in the relevance to learners as to how statistics is used to affect their lives and how they can use statistics to more fully understand phenomena. This points out a need then for at least two types of teaching materials: one type deals with the development of background knowledge and basic statistical skills while the second deals with engendering in the learner the abilities to recognize in ill defined situations how statistics can be used to obtain more information and to help in making decisions.

In the school year 1976-1977 we brought together at the University of Oregon a team of people to develop a resource on statistics for teachers of middle school children (ages 10-13 years). The project writing staff consisted of two middle school teachers, a high school teacher, a statistician, a mathematics educator and a mathematician. The resource, entitled Statistics and Information Organization [see Hoffer, 1977] contains background material for the teachers on statistical content and didactics along with an extensive collection of suggested classroom activities. Although the staff collaborated on all aspects of the writing, for specific sections each person had responsibilities which included an analysis of existing materials and publications, outlines, and writing of initial drafts. The resource consists of the following components:

Content for Teachers
Didactics
Teaching Emphases
Classroom Materials
Glossary
Annotated Bibliography
Selected Answers

The Content for Teachers section provides teachers with background information on statistics and probability. The subsections deal with gathering data, tables, graphs, scatter diagrams, averages, deviations, sampling, experimental probability, probability with models, counting techniques, probability with counting, and inferential statistics. There is also an appendix with 22 computer programs. There are numerous references here to other parts of the resource especially the classroom materials that serve as examples of the ideas as well as providing the teachers with connections to potential work for their students.

The Didactics section provides teachers with information on learning theories, teaching techniques, diagnosis and evaluations, goals and objectives. There are separate articles on components of instruction, classroom management, and statistics and probability learning. The latter includes a survey of the available research studies. Once again these articles contain several references to other components in the resource, especially the classroom materials.

The Teaching Emphases section contains a collection of processes, approaches, and aids that are emphasized throughout the resource. The subsections deal with critical thinking, decision making, problem solving, models and simulations, calculators and computers, and laboratory approaches. As the name implies these are aspects that should be emphasized as we teach statistics in addition to specific subject matter topics. There are cross references in the teaching emphases sections to the content, didactics and classroom materials sections of the resource.

The Classroom Materials section deals with concepts, skills and applications and takes up over half of the resource. The sixteen subsections here cover many of the usual topics including gathering data, scatter diagrams, sampling and monte carlo methods. The areas of application deal with business and commerce, environment, health and medicine, people and culture, and recreation. Before each subsection there are teacher commentaries which give a rationale for teaching the topic; suggest alternate ways to introduce or develop topics; suggest ways to involve students; highlight the classroom pages; and give more ideas that interrelate with the content, didactics and teaching emphases sections. The classroom materials are of two basic types: one intended for

direct student use and the other as suggested activities directed by the teachers. In both cases there are short lessons as well as sequences of integrated lessons. Each subsection contains worksheets that teachers can duplicate for class use; activity cards that may be laminated; transparency masters; games for small group or class use; teacher directed activities; teacher ideas that can be used to develop other activities; and teacher pages that give teachers mathematical background for specific student pages.

To round out the resource there is a Glossary that gives informal definitions and an Annotated Bibliography that provides teachers with sources for additional ideas and references.

As noted above, the various components of the resource are related. The classroom materials are keyed to each other, to the teaching emphases, to the commentaries, and to the didactics and content sections by using "teacher talk" in non-reproducible blue print on the student pages. The commentaries refer to specific classroom pages and often a classroom page is shown reduced in size next to the discussion of the page. The commentaries relate the various teaching emphases to the statistical topic of the sub-section and to the content and didactics sections. Each teaching emphases includes a rationale and examples from the classroom materials. The concepts and skills sections contain references to the applications sections to point out the potential uses of statistics. The didactics and content sections also refer to specific classroom pages and show a classroom page reduced in size to illustrate a topic of the sections.

The resource can be used by individual teachers, teams of teachers within a school or district, and by inservice leaders for classes and workshops. Teachers can use the resource

- To learn more about statistics. The content section has many ideas and useful background information for teachers who want to review or learn concepts in statistics and probability. There are several exercises included to help teachers improve their knowledge of the subject.
- To gain insight into difficulties in learning statistics. For example, some students find it difficult to use a systematic approach in listing permutations. Difficulties such as these are discussed in the didactics sections.
- To obtain information about curriculum trends and research in mathematics education. Many of the current trends in middle school mathematics are discussed in the teaching emphases sections and keyed on the classroom pages. The didactics section contains accessible synopses of the available research related to teaching statistics and probability.
- As a source of ready-to-use activities to supplement and vary the curriculum. Worksheets can be duplicated to provide a copy for each student; activity cards can be duplicated and laminated for repeated use; transparencies

can be made from a page to form the basis of a teacher-led discussion; or a page can be the focal point of a bulletin board or display.

- As a springboard for developing activities, sequences or lessons, or curriculum. Teachers can use the classroom pages as examples which they can alter to fit the interests of their classes. From the pages of teacher ideas and from the suggestions in the commentaries and didactics papers teachers can develop other activities and sequences of lessons.
- As a source of different ways to teach a topic. A variety of teaching approaches are included in resource. Concrete models are used in many of the classroom activities and attention is given to concept formation in the commentaries and didactics sections.
- To provide students with materials according to their interests. For example, there are classroom materials for students who are interested in applied problems, students who want to be challenged, and students who need experience with concrete activities in experimental probability.
- To improve attitudes. Many activities can be used to help students have a feeling of success and accomplishment. Some students who do not succeed in arithmetic can be successful in gathering data and using the self-checking pages. The many visual and hands-on materials can be used to capture student interest.
- To increase problem solving abilities and build basic skills. Many classroom pages provide problem solving activities and the problem solving teaching emphasis contains several hints and background information.
- To gain access to other sources. The annotated bibliography describes the features of various available references for teaching and learning statistics. Many of the sources are cited in the classroom pages, and in the other sections of the resource.

The resource was made available to teachers in the United States and Canada in 1978. The reaction was varied. At one extreme teams of middle school teachers reorganized their entire curriculum by integrating the resource materials with their traditional program. At the other extreme teachers ignored the resource entirely. It was more common for teachers to phase in the resource gradually by first trying a few classroom pages, then organizing a sequence of lessons, then using one of the teaching emphases as a theme for a unit of study, and finally developing materials of their own for their students.

To some extent the resource has influenced teachers, supervisors, curriculum directors and authors of articles and textbooks just as they influenced the developers of the resource, and yet the resource is far from a panacea. It offers to the more creative, more energetic teachers splendid opportunities to enrich the learning experiences that they offer to children. In many ways

the resource is too flexible especially for those teachers who only use lessons that are completely prepared for them.

While statistics is a subject in its own right, it is also a fiber that interrelates all other school topics in the social sciences, the sciences, the arts, mathematics and especially in the societal demands on people. The aspects of statistics as a truly interdisciplinary subject should be emphasized as new teaching materials are developed. Since teachers at the school level usually have the responsibility to teach several subjects, the lower grades are a natural place to begin a rich program of statistics edcuation. However, the materials that are developed should take the format of sequences of lessons that use statistics in analyzing questions in other subject areas. This connection with familiar and standard topics will make more palatable the inclusion of statistics or any other subject that may be new to the teachers.

In light of the many excellent sources that are already available [see Shulte, 1981], the materials that need to be developed now should be part of a thorough curriculum revision with a strong statistics component beginning at the school level and continuing throughout the secondary school [see, for example, Usiskin, 1979] as well as a similar integration of statistics in teacher education programs.

References

[1] Hoffer, A. (director) (1977) Statistics and Information Organization. Palo Alto, CA, Creative Publication.

[2] Shulte, A. (editor) (1981) Teaching Statistics and Probability. Reston, VA, National Council of Teachers of Mathematics.

[3] Usiskin, Z. (1979) Algebra Through Applications. Reston, VA, National Council of Teachers of Mathematics.

DEVELOPING MATERIALS FOR THE SCHOOLS COUNCIL PROJECT ON STATISTICAL EDUCATION

P Holmes, University of Sheffield

The Schools Council Project on Statistical Education (1975-1981), working in England and Wales and based at the University of Sheffield, was the first national project in the world to look specifically at the statistical education needs of all pupils in the 11-16 age range and to develop teaching materials to meet these needs. Although it was working at the same time as groups in France and in Hungary there was no connection between these projects at a time when they could have influenced each other [1].

The brief of the Schools Council Project was:

1 to assess the present situation in statistical education as regards content, level, motivation and teachers' attitudes, and to relate these to the position of statistics outside schools;

2 to survey the needs of teachers, whether they are teaching statistics as a specialist subject or working in a related field; implications for the needs of both initial and in-service training needs are to be considered;

3 to devise detailed proposals for the implementation of the teaching ideas;

4 to produce teaching materials such as notes for teachers, descriptions of experiments, work cards, sets of examples and audio-visual aids.

No curriculum development project works in a vacuum. All have to work in a particular country, taking into account the needs of society, the nature of the subject and its place in the curriculum, the background and the needs of the teachers and the needs of the pupils. In this talk I will take each of these four areas in turn and show how they affected the teaching materials developed by the Schools Council Project on Statistical Education. At the end of the talk I shall indicate the final result by giving some details from some of the units [5].

SOCIETY

Britain is a well developed society in which statistics plays a major part. There is a major Government Statistical Service whose findings are widely publicised. There are regular censuses of the population and all of the population has to fill in forms for statistical purposes at one time or another. Opinion polls are carried out frequently on many topics and many of their conclusions made public in the national press and television. People are expected to be

able to understand data presented to them in popular articles, in advertising and in the financial and commerce columns of newspapers. They have to use statistics in their own places of work to make decisions. A further account of this background to the work in the project is given in [4]. Clearly the teaching material developed by the project has to reflect this national picture.

At the time the project was working there was a growing national concern about levels of numeracy amongst young people. This concern showed itself in the setting up of a national committee under the chairmanship of Dr. W.H. Cockcroft and the committee's report has recently been published [2]. In it there are a number of references to statistics. In particular it states that "such textbooks (in statistics) as are available concentrate on theory rather than practice; it seems clear that there is a need for the provision of further teaching materials which will emphasise a practical approach to the teaching of statistics" (para. 780). The Schools Council Project was concerned to produce just this type of teaching material. Another indication of the way the project and the enquiry were thinking on the same lines, and both reflecting the needs of society, is shown in para. 781. Here the committee quote approvingly the evidence submitted to it by the project that "Statistical numeracy requires a feel for numbers, an appreciation of appropriate levels of accuracy, the making of sensible estimates, a common sense approach to the use of data in supporting an argument, the awareness of the variety of interpretations of figures and a judicious understanding of widely used concepts such as means and percentages. All these are part of everyday living, "Good statistics teaching can encourage pupils to think in these ways". Again, obviously, the project was trying to develop materials that would satisfy these criteria with an emphasis on discussing accuracy of data, being involved with collecting data so as to be more aware of the problems involved and on drawing practical inferences from the results. It is shown in the concern the project has for choosing everyday examples (Sampling the Census [6], Equal Pay [7], Smoking and Health [8]), emphasising real data and the real implications of the inferences that can be drawn, with less emphasis on introducing techniques for their own sake. The project allowed and encouraged the use of calculators and, for example, provided partly drawn graphs so that more time could be spent on discussing the meaning and implications of the figures.

National needs require that most of our pupils have a general understanding of statistics. They need to know what can be done with statistics, they also need to be able to identify misleading uses of statistics. For this latter purpose the project developed a special unit on misleading data [3]. Other pupils will need to go further and develop skills in a proper use of statistics. For them it was important to devise materials that would lay the foundations of

the nature of the subject - getting accurate data, allowing for variability and uncertainty, drawing inferences etc. - so that later formal work would fit into a well established schema in the pupil's mind.

Curriculum development takes place against the background of what a country expects from its education system. In England and Wales over the last few years the trend of opinion has been more towards the pragmatic, utilitarian and useful rather than the academic subject for its own sake. One of the biggest influences on the school curriculum in England and Wales are the national examinations set for 16 year olds by a few examination boards. These examinations are subject oriented, there are examinations specifically in statistics, but they are taken by only a minority of students. There is some statistics as part of many mathematics courses but, as is stated in the Cockcroft report (para. 776) "When statistics is taught within secondary mathematics courses too much emphasis is very often placed on the application of statistical techniques, rather than on discussion of the results or ordering and examining the data and on the inferences which should be drawn in the light of the context in which the data have been collected. The work can therefore become dry and technique-oriented and fail to show the power and nature of statistics". The project decided not to draw up a syllabus in statistics for one of the national examinations. This would have reached only a minority of pupils and would also have been very time consuming. It may not even have been possible within the three years for which the project was initially funded. Instead it was decided that we should try to influence courses as they currently existed, especially mathematics since this was a subject taken by nearly all pupils. The project therefore conceived of its work as a coherent series of booklets that could be used individually within different courses. The aims were to improve present content and practice, to develop what was desirable rather than what was directly assessable under examination conditions, to encourage more practical work and indirectly to change the prevailing national view of statistics particularly of mathematics teachers, mathematics examiners and those responsible for drawing up the national syllabuses.

The accepted pattern for curriculum change also affects the way teaching material is developed. In some countries the curriculum is controlled centrally; in England and Wales the final decision rests with the schools (though clearly there are regional and national pressures for conformity). This means that we had to involve schools and teachers with the development work from a very early stage. This is very good for the project team. Teachers have to be convinced that the material is useful, sensible and will work with their pupils. At the same time they make practical suggestions as to what will satisfy these criteria. By working with teachers in this way and letting them advise on, test and evaluate the trial material there is less likelihood of the material being

seen as something unwanted imposed from the outside. It helps the material match up with current classroom practice. It also builds up a band of informed teachers sympathetic to the aim of the project.

THE SUBJECT AND ITS PLACE IN THE CURRICULUM

A start was made on indicating the current place of statistics in the school curriculum in the previous section. Part of the early work of the project was to investigate this position more closely. Statistics is not well established as a separate subject in the main school curriculum. On the other hand there is a widespread and growing use of statistics being made in many other subjects across the school curriculum. It also has an established part in the teaching of mathematics. This is more fully documented in the project's survey [2]. This widespread use poses several interdisciplinary problems of its own. In many cases teachers of different subjects do not know what their colleagues are doing. There are also different approaches to and requirements of statistics by teachers of different subjects. The mathematics teachers are more likely to concentrate on the mathematical aspect of the subject and seek for a mathematical understanding of the techniques. Their colleagues in other subjects may well see statistics more as a tool and not be so concerned with developing understanding. In some cases it was clear that the level of difficulty of the concept being used was not appreciated. The implications of this for the project's teaching materials were that we tried to choose topics for the booklets that could be used in different subject areas of the school curriculum and that would also encourage some interdisciplinary co-operation. Booklets developed the techniques in a context to help students, and their teachers, appreciate the why and when of using statistics. Statistics is clearly a subject that is widely applied and used in practice. The developed materials should reflect this fact.

This brings us to the impact of the nature of the subject on the materials developed. As the Cockcroft report says, quoting evidence from the project, "Statistics is not just a set of techniques, it is an attitude of mind in approaching data. In particular it acknowledges the fact of uncertainty and variability in data and data collection. It enables people to make decisions in the face of this uncertainty". All this was reflected in the project's aims.

1. Children should become aware of and appreciate the role of statistics in society. That is they should know about the many and varied fields in which statistical ideas are used, including the place of statistical thinking in other academic subjects.

2. Children should become aware of, and appreciate, the scope of statistics. That is, they should know the sort of questions that an intelligent use of statistics can answer, and understand the power and limitations of statistical thought.

These two aims can be seen as wanting children to know about the breadth of statistical applications and the depth of statistical thought. Clearly statistics introduces some very difficult ideas. These include indeterminism, uncertainty, randomness, a sample to represent a population, making estimates from samples, drawing inferences and so on. These ideas take some time to mature and occasionally run counter to the pupil's previous experiences at school. This is particularly true of the ideas of indeterminism and inference in the mathematics lessons where most of the answers are very deterministic and the emphasis has been much more on the certainty of logical deduction. Because of this the project's materials spend much time on practical work, on experiencing the problems of getting the data, of seeing not only the different outcomes that might happen but also that some are more likely than others. There is work on simulation, sampling and discussion of possible implications well before any such ideas are formalised. The practical nature of statistics encouraged the project to produce materials which taught techniques in a context and not to isolate them. This helps the pupil to decide when to use the technique and to see more clearly how it is used. Because ideas take time to mature they are introduced in an elementary way in the early units, they then occur again in later units and so the ideas are gradually reinforced. A full summary of which ideas occur in which units is given in the teachers handbook [4].

The relationship between probability and statistics also had an impact on the way the materials are being developed. At a later stage probability models, particularly probability distribution, can be seen as models for different measures in a population and hence we can quantify our inferences drawn from samples. This is far too sophisticated a concept for eleven year olds, yet probability ideas are important and take a long while to develop. Thus the early project units develop probability and statistics mainly in separate units - though there is one early unit [10] which starts to bring the two ideas together through the concepts of fairness and randomness. The early units in probability emphasise uncertainty, using games and simulations to see that decisions will sometimes be right and sometimes wrong but to get a feel that some decisions will be better than others. The early units in statistics concentrate on getting accurate data, seeing what it tells you, summarising and presenting it in different ways and drawing informal inferences from these summaries. Only in the later units do the two strands start to come together and more formal work in estimation and significance testing etc. is left to later.

THE TEACHERS

In the last resort any teaching material developed has to be accepted by, and usable by, the teachers. The initial survey carried out by the project showed that even amongst the mathematics teachers more than one-third had had to study their statistics on their own, not as part of their initial training. This proportion was even higher for teachers who had been in the profession for some time. Even for those teachers who had had some training in statistics it was usually only a small part of their background, it tended to be rather academic and did not greatly help them identify and appreciate what was suitable for their pupils. Very few had ever used statistics in a practical context. Very few indeed had had any training in the teaching of statistics, such work playing no part in by far the large majority of initial teacher training courses. There was a general feeling that statistics was a 'difficult' subject and so a lack of confidence when it came to teaching. This meant that in developing its materials the project had to think carefully how it would also help the teachers. One result was the production of teachers' notes for each of the units and a teachers handbook [4] to give background and added support.

In order to build up confidence the project team not only regularly visited the trial schools but also arranged local and national meetings for the teachers to share their problems and to become more involved with the project. Because most teachers felt their lack of competence it was not possible to persuade them to write the initial drafts, though they did give useful advice, and this writing was done by the project team. Their willingness to help with testing, evaluating and making positive comments certainly improved the quality of the material, and some of them wrote the revised version of the units prior to final editing.

Teaching styles differ from teacher to teacher and from school to school. The project worked with teachers whose natural styles varied from the stand up and talk method to the organising of completely individualised learning schemes. The teaching material had to be developed in such a way as to match up with as large a range of these styles as possible - subject to the nature of the subject allowing such a style to be sensible. The project's materials therefore include, in the pupil booklet, all the information that a pupil needs so that it is possible to use the teaching style which encourages pupils to work together with less teacher involvement. Completely individualised learning was not found appropriate to an approach which required practical experimental work, collecting data and discussing the results. Teachers used to a more formal presentation of their lessons can still work in this way but, in order to encourage greater pupil participation, there are places in the material where discussion of the results is required or there are open-ended questions where it is inappropriate

for the teacher to tell the pupils "the answer". In practice it was found that even the low ability pupil had something to offer to such discussions either from practical experience or by imaginative comment.

THE PUPILS

The most important consideration of all curriculum development of teaching materials is the pupils themselves. In particular the materials must take into account the age, ability, background, present and future needs of the pupils. The Schools Council Project was developing materials for pupils of all abilities from the ages of 11 to 16. It could not assume that any earlier work had been done by the pupils on probability and statistics, though some pupils had in fact met some elementary ideas on collection and representation of data in their earlier years at school. Between the ages of 11 and 16 pupils mature greatly and their interests change and broaden. The project's material reflects this by choosing topics more from home and school in the early units where elementary ideas of probability and statistics are being developed and moves to topics of more general, social, topical and personal interest in the later units.

Because pupils of all abilities are envisaged then each unit has to include something that all can do. This might be the easier parts; it is also the sections on open questions and discussion which can be answered at many levels. Most units also incorporate extra material for the more able and also have different stopping points for those who can't work as quickly.

The present and future needs of the pupils may be for use in other subjects, in their everyday life, as people who use statistics at work or (in a minority of cases) as future specialists. The effect of these considerations on the teaching material produced has been considered above.

In this paper I have tried to show some of the many things anyone developing new teaching materials has to take into account. I have indicated how they affected the teaching material of the Schools Council Project on Statistical Education and the approach taken by the project team. I hope that this will help any of you who have to develop teaching material in your own countries to identify the areas of concern and produce relevant material for your pupils.

References

[1] Barnett, V. (1982) Teaching Statistics in Schools Throughout the World, International Statistical Institute, Voorburg.

[2] Committee of Inquiry into the Teaching of Mathematics in Schools (1982) Mathematics counts: Report of the committee of Inquiry... under the chairmanship of Dr. W.H. Cockcroft, HMSO, London.

[3] Holmes, P., Kapadia, R., Rubra, G.N. and Turner, D.E. (Ed.) (1981) Statistics in Schools 11-16: A Review, Methuen, London.

[4] Schools Council Project on Statistical Education (1980) Teaching Statistics 11-16, Foulsham, Slough.

[5] Schools Council Project on Statistical Education (1980), Statistics in Your World, (17 Pupil Booklets, 27 Teachers' Booklets) Foulsham, Slough.

[6] Schools Council Project on Statistical Education (1980) Sampling the Census, Foulsham, Slough.

[7] Schools Council Project on Statistical Education (1980) Equal Pay, Foulsham, Slough.

[8] Schools Council Project on Statistical Education (1980) Smoking and Health, Foulsham, Slough.

[9] Schools Council Project on Statistical Education (1980) Phoney Figures, Foulsham, Slough.

[10] Schools Council Project on Statistical Education (1980) Being Fair to Ernie, Foulsham, Slough.

EDA FOR 16-19's?

J Bibby, The Open University

ABSTRACT

This paper asks various questions concerning the relevance of Exploratory Data Analysis (EDA) ideas for teaching 16-19 year old's, and reports on a workshop which considered these questions.

INTRODUCTION

Four key questions can be posed which are pertinent to the question of preparing EDA materials for use at school level:

(a) are EDA ideas relevant at this level?
(b) if so, then which ideas and techniques should be taught?
(c) which data sources are sufficiently motivating for this age group?
(d) how should or could such material be prepared?

Before attempting to answer these questions it is important to be clear what EDA does and does not do.

STEMPLOTS

This simple technique provides an overview of the distribution of values in a batch. In particular, it provides an impressionistic overview of

- the general location of the batch (the "main body" of data)
- its spread, and in particular whether one batch is more, or less, spread out than another
- the location of peaks or modes
- the existence of tails, and the relative skewness or symmetry of the batch
- the existence of outliers.

However, this overview can be misleading if care is not taken to experiment with the effect of different grouping intervals.

In addition, the process of drawing the stemplot pushes the analyst's nose "close to the data", and should bring to light

- patterns related to the order of the data, e.g. auto-correlation
- any clearly aberrant observations.

In short, drawing the stemplot often provides an informal check for errors.

FIGURE 1: Measurements made by 4 individuals when attempting to bisect 12 lines by eye (Source: Kanji, 1982).

	A	B	C	D
1	3.04	3.32	3.26	3.11
2	3.16	3.23	3.14	3.10
3	3.10	3.28	3.20	3.15
4	3.01	3.35	3.18	3.16
5	3.12	3.26	3.23	3.12
6	3.07	3.20	3.27	2.98
7	3.13	3.24	3.19	3.04
8	3.07	3.32	3.25	3.04
9	3.16	3.19	3.22	3.26
10	3.18	3.26	3.19	3.20
11	3.19	3.41	3.18	3.14
12	3.27	3.03	3.22	3.13

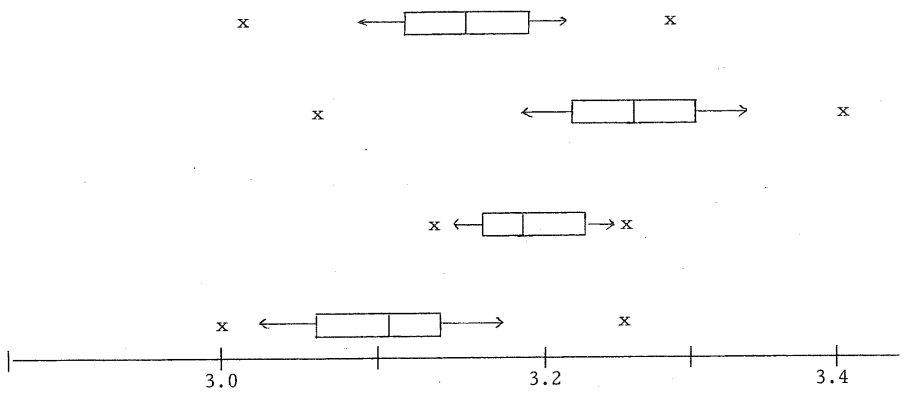


FIGURE 2: Decile boxplots based on Figure 1
(arrowheads denote deciles)

BOXPLOTS

Decile boxplots provide an indication of the distribution of data values in a batch. For example, the data in Figure 1 can be shown more graphically in Figure 2. (For further details of decile boxplots see The Open University 1983, Unit A2).

GENERAL POINTS CONCERNING EDA

There are of course many other EDA methods, e.g.

- line-fitting
- median polishing
- transforming for symmetry/linearity/additivity

All these are relatively simple, but may take time to trickle down to the school level.

The pro's and con's of EDA may be summarised as follows:

The Pro's of EDA

- EDA emphasises pattern, which is an intuitive "fuzzy" action (unlike probability)
- EDA calculations are generally simple; they use counts, not sums : hence there is no need in EDA for the intimidating panoply of sigma signs which are required by conventional methods
- EDA methods are generally applicable as they do not require a probability model; therefore they apply also where the data does not arise from a sample, where the sample design is complex, or where there has been a complete census
- EDA encourages emphasis on residuals, the DFR equation (Data = Fit + Residual), the notion of statistical analysis as an interative procedure, and the model-building loop (see An Open University Service Course, Figure 5, by John Bibby, this volume).

The Con's of EDA

However, EDA does have certain disadvantages, including the following:

- many EDA procedures are difficult to define precisely and algorithmically (e.g. deciding whether to 'squeeze' or 'stretch' a stemplot)
- terminology in the area is far from standard
- there may be problems in finding suitable teachers (it is important that statistics should be teachable even by bad teachers!)
- EDA is dangerously empiricist - it risks encouraging the notion that knowledge somehow "arises out of the data"; it downplays prior knowledge and the role of theory.

RAPPORTAGE FROM WORKSHOPS

Workshops were set up to consider questions (a) to (d) outlined in the Introduction. They are reported as follows:

There seems to be some place for EDA ideas at this level, although an immediate problem is the unfamiliarity of teaching staff with EDA ideas and techniques. There may be a need for training courses in this area. Also, certain EDA material is more suitable for younger children.

The question of suitable data sources was reckoned to be a key problem. These could be self-generated, as in the experimental task arranged by Gopal Kanji (1982) in which participants were required to bisect by eye each of a set of straight lines. Errors could be related to:-

(i) right/left handedness
(ii) sequence number on page
(iii) position on page (left/right)

Data relating to pupils' families was also reckoned to be relevant, as were large data sets such as those collected by Governmental surveys.

The preparation of EDA materials obviously had to involve classroom teachers. However, it also required some central initiative, such as could be provided by the Leverhulme Project directed by Peter Holmes at Sheffield University.

References

[1] Kanji, G. (1982) Developing materials for 16-19 year old students. (Paper prepared for ICOTS (Mineo, 5pp)).

[2] The Open University (1983) MDST242: Statistics in Society (16 vols.) Open University Press.

TOPIC 6: TRAINING OF TEACHERS IN STATISTICS

ORGANISER: PROFESSOR A P SHULTE

PREPARATIONS FOR COLLEGE TEACHING IN STATISTICS

A K Gupta, Bowling Green State University

INTRODUCTION

In recent years there has been considerable and growing interest in the use of statistics throughout many sectors of our society. The quantitative nature of this discipline has created opportunities for scientists to participate in the new technology of the twentieth century (Rao, 1973), characterized by its unique methodology, akin to what is termed "the scientific method" in the philosophy of science. This increased demand for statistical knowledge (see Amstat News, 1981) can be met adequately if the universities will produce qualified college teachers of statistics. Here we are concerned with the training of trainers. The subject is broad and many statisticians have written about the teaching of statistics and the need for reform, beginning, perhaps, with Wishart (1939 and 1948) and including Yates and Healy (1964), Rao (1972), Hogg (1972), Kempthorne (1974), Daniels (1975), Gupta (1982) among others.

Let us first look at the profile of a college teacher of statistics. A college teacher of statistics is required, with rare exceptions, to have a doctorate. He enters a college in which a pattern already exists, and in the beginning his main concern is to teach already existing elementary courses in mathematics and courses in statistics. He works in an academic atmosphere and usually works alone on his research problems. Being a specialist, he is also engaged in consulting, writing referee's reports for journals and reviewing. He is also called upon to contribute his time and energy to the work of the committees in which much important activity of an academic community is centered. Such are the demands of a college teacher of statistics. In this paper some of the basic knowledge and skills in statistics that are needed in order to prepare oneself to teach statistics effectively are discussed. Besides statistics, competencies in mathematics, probability, computational techniques, history and a cognate area are identified. The importance of participation in undergraduate teaching, statistical consulting, and continued professional development is also discussed.

REQUIRED COMPETENCIES

The prospective college teacher, as a graduate student, has the desire to learn and the ability to learn, but must also have the desire to teach and the ability to impart knowledge in a logical way. But teaching of statistics calls for expertise and hence in-depth understanding of the subject is the best preparation. The rationale here is that college teaching is a skill

which can be acquired by receiving instructions. But what kind of instruction should a prospective college teacher of statistics pursue? In the light of the profile of a college teacher of statistics described above we list below the competencies which should be acquired in order to do the job effectively.

(i) Mathematics. A sophisticated training in relevant mathematics is a prerequisite to the advanced study of statistics. The competency should be gained in (a) advanced calculus (including special functions and transformations), (b) linear algebra (including vectors, matrices and eigenvalues), (c) differential equations and (d) real analysis.

(ii) Probability. Since all branches of statistics use probabilistic arguments, competency in the following topics should be acquired: (a) basic concepts (especially conditional probability), (b) stochastic processes (including time-series analysis), (c) asymptotics and (d) distribution theory.

(iii) Statistics. There are many different branches of statistics and statistics is still proliferating. The main point here is that graduate students be exposed to different points of view. Courses should emphasize modern statistical theory and its applications and include (a) statistical inference (estimation and testing), (b) nonparametric analysis, (c) Bayesian analysis, (d) sequential analysis, (e) decision theory, (f) multivariate analysis and linear models (including regression), (g) data analysis, (h) sampling techniques and design of experiments.

(iv) Computing Techniques. Computers are used as an aid in teaching statistics, analyzing data and simplifying calculations. Hence knowledge of at least one programming language (e.g. Fortran, Basic, APL) and familiarity with most useful statistical packages (e.g. BMDP, SAS, SPSS, MINITAB, etc.) is imperative. Courses should include (a) programming, (b) data processing and (c) simulation techniques.

(v) Cognate Studies. Statistics is a dynamic subject which pervades so many fields including biological, medical, physical and social sciences. In fact a great many problems that statisticians solve come from these fields. Competency, maybe as a minor, in at least one such allied field should be acquired.

(vi) History. The history of statistics is a fascinating one and every teacher should know it. Knowledge of history gives the teacher confidence and makes teaching lively. Every statistician should be familiar with the works of Galton, K. Pearson, Fisher, E. Pearson, Neyman and other pioneers.

PARTICIPATION IN UNDERGRADUATE TEACHING AND STATISTICAL CONSULTING

So far we have been concerned with the coursework training of prospective college teachers of statistics. However, teaching also belongs to the activities people learn by doing. It has been suggested (MAA, 1979) that teaching assistantship should be teaching apprenticeship during which prospective college teachers (then graduate students) get training in the skills that will make them effective in the classroom. Next to the competencies to be acquired through coursework, participation in the supervised undergraduate teaching is the best preparation for a college teacher.

Another mode of preparing by doing is the participation in statistical consulting. This is usually achieved by working in a statistics laboratory or consulting center at a unviersity. By working on real problems, a student learns the art of asking questions, reads relevant material on the topic and gains experience in the art of communication so important for a statistician and college teacher.

IN-SERVICE TRAINING

A college teacher of statistics needs to keep up with the changes in the subject matter, advances in instructional technology and new societal demands. He needs skills of curriculum planning. Competency in these areas can be achieved through in-service training and continued professional development programs. In this context the concept of useful half-life is pertinent. For example, an engineer is thought of as losing half of his effectiveness each 8 years. Hence universities and professional societies should create the mechanism for training and refresher programs for college teachers of statistics. The American Statisical Association and the American Mathematical Society do provide short courses in relevant new topics in statistics (and mathematics), some of which are available in the form of films and tapes. It will be very useful to establish some sort of clearing house to publicize the availability of such audio and video materials so that college teachers of statistics throughout the world could benefit.

CONCLUDING REMARKS

The universities in the United States are doing an adequate job of preparing college teachers in statistics. However, improvements are necessary. There can hardly be a single international approach to the education of college teachers in statistics, though one should expect common international standards of excellence in statistical education. Effective college teaching may be independent of the subject but a regional approach is very important in pre-

paring for college teaching in statistics due to factors such as: resources, language, etc. However, international collaboration among nations in preparing college teachers in statistics could be of great value. The International Latin American Training Center in Statistics involving 15 countries is an example of such a collaboration. In this regard an International Directory of Academic Statisticians is proposed to be made available. Such a directory should include, in addition to their specialty areas and research interests, information about their linguistic capabilities. Such a reference source, I hope, will be useful in drawing from the world pool of trained statisticians, to help prepare college teachers in statistics for regions where they may be needed.

References

[1] Daniels, H.E. (1975) Statistics in universities - a personal view (with discussion). J.Roy.Statist.Soc., Series A, 138, 1-17.

[2] Gupta, A.K. (1982) Random thoughts on developing statistics in the thrid third world. Trabajos de Estadistica.

[3] Hogg, R.V. (1972) On statistical education. American Statistician, 26, 8-26.

[4] Kempthorne, O. (1974) The teaching of statistics at the M.S. level. An essay prepared for the inaugural ceremonies for the Center of Statistics and Computation, Chapingo, Mexico.

[5] Rao, C.R. (1972) A multidisciplinary approach for teaching statistics and probability. Chapter 13, pp.251-272 in The Teaching of Probability and Statistics. Edited by L. Råde.

[6] Rao, C.R. (1973) Linear Statistical Inference and its Applications. Wiley, New York.

[7] "Supply and Demand for Statisticians". (1981) Amstat News, April, No.74, 5.

[8] Training Programs for Teaching Assistants in Mathematics, (1979) Mathematical Association of America.

[9] Wishart, J. (1939) some aspects of the teaching of statistics. J.Roy.Statist.Soc., Series A, 102, 532-564.

[10] Wishart, J. (1948) The teaching of statistics. J.Roy.Statist.Soc., Series A, 109, 532-564.

[11] Yates, F. and Healy, M.J.R. (1964) How should we reform the teaching of statistics? J.Roy.Statist.Soc., Series A, 127, 199-210.

TEACHING DESIGN OF EXPERIMENTS IN INDIAN UNIVERSITIES

V Bhagwandas, University of Indore

A GLIMPSE OF INDIA

India is a vast country, the Indian population being 15.53% of the World's population (reference year 1980-81) and the second largest country after China (21.72%). This country is divided into 22 states and 9 union territories inhabited by 339.9 million males and 318.2 million females with percent decadal growth rate 1971-81 as 24.64%. Increase in population density from 177 persons per square kilometer in 1971 to 221 persons per square kilometer in 1981.

Literacy in the Census of India is defined as follows:

The person who can read, write and understand a language is defined as literate whereas the person who can only read a language but cannot write it is supposed to be illiterate. Children below the age of five years are deemed to be illiterate. On the basis of working paper I (1981) Census of India (which includes children below five for 1981 data; due to absence of this information separately till the publication of this paper), the literacy rate is as follows:

	Persons	Male	Female	Male/Female Ratio
1961	24.02	34.44	12.95	941
1971	29.45	39.45	18.69	930
1981	36.17	46.74	24.88	935

*Little more cross section for 1961 and 1971 (1981 data is not available) for the age-group 15-35 i.e. adult group has the following data on literacy rate.

Literacy position of the population in the age group 15-35 by sex and Urban/ Rural areas in 1961 and 1971 as follows:

		% of Illiterate			% of Literate		
Year	Area	Persons	Male	Female	Persons	Male	Female
1961	Total	67.60	53.17	82.45	32.40	46.83	17.55
	Urban	39.28	26.98	54.41	60.72	73.02	45.59
	Rural	74.61	60.37	88.63	25.39	39.63	11.37
1971	Total	57.97	43.29	73.24	42.03	56.71	26.76
	Urban	31.40	21.47	43.33	68.60	78.53	56.67
	Rural	65.75	50.27	81.20	34.25	49.73	18.80

*Source: Literacy Statistics at a Glance, Directorate of Adult Education

The definition of urban area adopted after the 1951 Census is as follows:

(1) All places with a municipality, corporation, Cantonment board or notified town area committee etc.

(2) All other places which satisfy the following criteria:

i) a minimum population of 5,000

ii) at least 75% of the male working population engaged in non-agricultural pursuits

iii) a density of population of at least 400 persons per square kilometer

We give Number of Institutions for 1978-79 and 1979-80 and enrolment by stages/classes for the year 1978-79 as Table 1 and Table 2 as given by Planning, Monitoring and Statistics Division, Ministry of Education and Culture, Government of India "Studies in Educational Statistics No.1 - 1981".

Under the Constitution of India, education is a primary responsibility of the state and states have compulsory education for children between the ages 6 to 11 and that explains the large primary schools in the country. High School/Higher Secondary and Intermediate or college education is imparted both by private as well as government schools and colleges. Unfortunately, the quality of education imparted at government schools/colleges is poorer as compared to private and a commoner cannot afford higher fees at private institutions. Just to give an idea of pass percentage in High School and Higher Secondary School, after looking into the results of various Boards of Secondary Education of States in 1977 we give the following table:

Table 1

NUMBER OF INSTITUTIONS

Type of Institutions	1978-79	1979-80
(a) Pre-Primary Schools	9,172	10,426
(b) Primary Schools	474,993	478,249
(c) Middle Schools	110,902	114,720
(d) High Schools	36,190	37,419
(e) Higher Secondary Schools (Old Pattern)	3,121	3,127
(f) Higher Secondary Schools (New Pattern)	4,531	5,044
(g) Inter/Junior Colleges	3,017	3,197
(h) Arts, Science & Commerce Colleges	3,261	3,336
(i) Universities	106	108

Table 2

ENROLMENT BY STAGES/CLASSES (1978-79)

(INDIA)

	Boys	Girls	Total
Ph.D./D.Sc./D.Phil.	16,376	5,817	22,193
M.A.	106,417	64,350	170,767
M.Sc.	48,529	22,458	70,987
M.Com.	31,719	2,001	33,720
B.A./B.A.(Hons)	626,231	360,494	986,725
B.Sc./B.Sc.(Hons)	381,346	136,069	517,415
B.Com./B.Com.(Hons)	412,801	58,538	471,339
B.E./B.Sc.(Eng)	94,501	3,246	97,747
B.Ed./B.T.	43,761	32,199	75,960
M.B.B.S.	53,617	16,717	70,334
Inter Arts	190,617	65,055	255,672
Inter Science	158,455	35,023	193,478
Inter Commerce	39,124	2,201	41,325
Pre-Degree/Pre-University (two-year course)	186,739	100,365	287,104
Pre-University (one-year course)	57,143	22,215	79,358
Higher Secondary (New Scheme XI and XII Classes)	807,670	386,284	1,193,951
Higher Secondary (Old Scheme IX, X and XI Classes)	945,313	303,765	1,249,078
High School (IX and X Classes)	4,854,358	2,066,915	6,921,273
Middle (VI-VIII Classes)	12,548,337	6,152,894	18,701,231
Primary (I-V Classes)	43,662,003	27,278,383	70,940,386
Pre-Primary Stage	488,961	415,170	904,131
Teacher Training Schools	59,411	45,713	105,124
Polytechnics	116,907	10,845	127,752
Technical and Industrial Schools	163,185	37,701	200,886

	Number Appeared			Number Passed			Percentage		
	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
High School (Annual)	2,549,794	983,064	3,532,858	1,141,566	427,898	1,569,464	44.8	43.5	44.4
High School (Suppl.)	555,722	146,548	702,270	163,663	39,352	203,015	29.5	26.9	28.9
Higher Secondary (Annual)	416,218	144,775	560,993	196,478	91,150	287,626	47.2	63.0	51.3
Higher Secondary (Suppl.)	44,697	21,946	66,647	17,792	9,342	27,134	39.8	42.6	40.7
Intermediate Jr./Pre-university (Annual)	387,041	106,987	494,028	229,248	68,793	298,041	59.2	64.3	60.3
Intermediate Jr./Pre-University (Suppl.)	53,232	6,400	59,632	27,518	3,003	30,521	51.7	46.9	51.2

The pass percentage rises up as we go from high school level to university level due to students feeling more responsible; the education is costlier and not many, therefore, can afford higher education.

UNIVERSITIES IN INDIA

The first universities to be established were Calcutta, Bombay and Madras in the year 1857, the next Punjab University in 1882 and then Allahabad in 1887. For quite some time these five universities constituted the university system in India, and there were 19 universities before independence in 1947. By 1960, 45 universities were established and by 1970 the total number had risen to 83; today the number is 137, out of which 93 are traditional universities, 21 are agricultural universities and 10 belong to the category of the institutions of national importance and 12 are known as deemed to be universities, and one not covered was established in 1981, named as Jagannath Sanskrit Vishvavidyalaya.

In India, colleges were established first and universities came later. When the universities came to be established the colleges ceased to be independent units and became affiliated units of the universities. In fact the affiliating system is to be found only in India. On the recommendation of the Sadler Commission some of the unitary and residential universities, like Lucknow, Dacca, Waltair and Delhi were established but just after independence and due to a tremendous surge in enrolment most of the expansion took place along the lines of the establishment of a large number of colleges/universities which were affiliating in character and became the foci of growth and their number rose from 14 in 1947 to 74 as of today.

In recent years, research facilities have grown a great deal in range and variety and with new universities, departments of postgraduate education and research are increasing and new postgraduate colleges too try to equip themselves with research facilities. There had been a reduction in the funds available with the U.G.C. for research but other agencies like CSIR, ICAR, ICMR, Department of Atomic Energy, Department of Science and Technology, etc. do support researches in Science and Technology. But unfortunately, in India there is no rationale behind distribution of these funds and some universities and national institutes are favoured much more as compared to others.

Out of 94 traditional universities if we exclude 10 universities like Sanskrit Universities, Technology Universities, Ayurved Universities, then 84 can be classified according to type as follows:

S.No.	Type	Number
1.	Teaching and Affiliating	42
2.	Affiliating Teaching and Residential	10
3.	Affiliating cum Teaching	9
4.	Affiliating	6
5.	Teaching and Residential	5
6.	Unitary, Teaching and Residential	4
7.	Unitary and Teaching	2
8.	Federal	2
9.	Teaching	1
10.	Unitary and Residential	1
11.	Teaching cum Residential	1
12.	Unitary and Affiliating	1
	Total:	84

STATISTICS IN UNIVERSITIES

(a) 13 out of 21 Agricultural Universities have Statistics departments, 3 being departments of Agricultural Statistics. The largest of these departments is the Department of Agricultural Economics and Statistics in Punjab rao Krishi Vishwavidyalaya, Akola, having 2 Professors, 4 Readers and 17 Assistant Professors; the second largest is that of Punjab Agricultural University.

(b) 10 institutions of National importance include 5 Institutes of Technology at Bombay, Delhi, Kanpur, Kharagpur and Madras and only I.I.T. at Bombay and Kanpur have regular postgraduate courses in Statistics. One Institute, Dakshina Bharat Hindi Prachar Sabha is for the 'Hindi' language and 3 Medical Institutes, out of which the All India Institute of Medical Sciences at Delhi and the Postgraduate Institute of Medical Education and Research at Chandigarh take the services of Bio-Statisticians. Tenth is the internationally known institute viz. Indian Statistical Institute, meant for Statistics and Statisticians.

(c) Out of 12 institutions deemed to be universities the Indian Agricultural Research Institute at Delhi is the only institution having a large department of Agricultural Statistics doing useful work in teaching and research in Statistics.

(d) The following table gives the number of universities imparting postgraduate (at Master's level) instruction in Statistics.

S.No.	Type	Number of Universities having P.G. in Statistics
1.	Teaching and Affiliating	21
2.	Affiliating Teaching and Residential	7
3.	Affiliating cum Teaching	6
4.	Affiliating	3
5.	Teaching and Residential	3
6.	Unitary Teaching and Residential	4*
7.	Unitary and Teaching	Nil
8.	Federal	1
9.	Teaching	Nil
10.	Unitary and Residential	Nil
11.	Teaching cum Residential	1
12.	Unitary and Affiliating	Nil
	Total:	46

* All the four universities which are unitary, teaching and residential have a postgraduate department of teaching in Statistics. Half of the universities in India have yet to open doors for master's level instructions in Statistics.

TEACHING AND RESEARCH IN DESIGN OF EXPERIMENTS IN INDIAN UNIVERSITIES

The importance of Statistical Designs in Statistics was established by Federer & Federer [1] in 1973, while making a case for a separate research journal in Design of Experiments and allied branches. Out of the articles published in 18 journals between 1968 and 1971, 15.6% of articles and 15.2% of page coverage was only in the area of Experimental Design, Treatment Design Measuring and Measurement, Sampling Survey Design, Sequential Design, Modelling Design and others. Now if we drop the topics like Sample Survey Design and Sequential Design, we can see that percentage still does not fall below ten.

The main centres in India for research in Design of Experiments have been the Indian Statistical Institute Calcutta, Calcutta University, Bombay University and the Indian Agricultural Research Institute, Delhi and the main sources of inspiration have been people like Professor C.R. Rao in Calcutta, Professor S.S. Shrikhande and Professor M.C. Chakraborty in Bombay and Professor M.N. Das in I.A.R.I., New Delhi.

The questionnaire was sent to all the 46 universities' departments and institutes to find out the status of these departments with respect to teaching and research in Design of Experiments. Only 25 departments and institutes responded - the other 21 departments did not feel it necessary to respond. The findings are as follows:

(1) Most of the universities have Design of Experiments as a compulsory subject forming half the paper along with Multivariate Analysis, Matrix Theory or Sampling Theory except in Madras, where Design of Experiment is also offered at M.Phil. level and at optional level. Four universities, namely Calcutta,

Punjabi University Patiala, Gauhati and Indore offer a compulsory as well as an optional paper in Designs. At optional level, one treats the subject at advanced level. Construction of Designs is taught in nine universities.

(2) Topics covered mostly in these courses on design include: Theory of Linear Estimation, Gauss Markoff's Theorem, Linear Models, General Theory of Analysis of Experimental Designs, Randomised Block, Latin Squares, Graeco-Latin Squares, Cross-over Designs, Incomplete Block Designs; Balanced and Partially Balanced, Youden Square, Split Plot Experiments, Symmetrical and Unsymmetrical Factorial Experiments, Missing Plot Techniques and Response Surface Analysis.

On the construction side, the knowledge of EG(n,s), PG(n,s) complete set of m.o.l.s and their connection with PG(2,s) EG(2,s), Orthogonal Arrays, Hadamard Matrices, Tactical Configurations, Duals of Incomplete Block Designs, the application of Graph Theory in Construction of Designs, Construction of various GD designs, Triangular Designs, L_r association Scheme etc. is usually covered in a good course.

(3) Books followed in order of preference are given below. The numbers in brackets indicate the number of departments using the books.

(1) Oscar Kempthorne	"Design & Analysis of Experiments" (21)
(2) Cochran & Cox	"Experimental Designs" (16)
(3) Chakraborty, M.C.	"Mathematics of Designs and Analysis of Experiments (13)
(4) Raghavrao, D.	"Construction & Combinatorial Problems in Design of Experiments" (9)
(5) Das, M.N. & Giri, N.C.	"Experimental Designs" (8)
(6) Federer, W.T.	"Experimental Designs" (7)
(7) John, P.W.M.	"Statistical Design and Analysis of Experiments" (6)
(8) Mann, H.B.	"Analysis & Design of Experiments" (4)
(9) Scheffe, H.	"Analysis of Variance" (3)
(10) Rao, C.R.	"Linear Statistical Inference" (2)
(11) Jambunathan	"Linear Estimation" (2)
(12) Cox, D.R.	"Planning of Experiments" (2)
(13) Fedorov	"Theory of Optimal Experiments" Academic Press
(14) Fisher	"The Design of Experiments"
(15) Pearce, S.C.	"Field Experimentation with Fruit Trees and Other Perennial Plants"
(16) Montgomery, D.C.	"Design & Analysis of Experiments"
(17) Anderson, R.L. & Bancroft, T.A.	"Statistical Theory in Research"
(18) Panse, V.G. & Sukhatme, P.V.	"Statistical Method for Agricultural Workers"

(19) Ogawa, J.	"Statistical Theory of the Analysis of Experimental Designs"
(20) Ryser, H.J.	"Combinatorial Mathematics"
(21) Hall, Marshall	"Combinatorial Theory"
(22) Bose, R.C., Clatworthy, W.H. & Shrikhande, S.S.	"Tables of PBIBD with two associate classes"

The book numbers (20) and (21) are recommended in Nagpur University for half paper on Construction of Design due to the fact that Professor Shrikhande works in Nagpur as Honorary Professor after retirement from Bombay.

(4) Research publications in Design of Experiments during the last five years by various university departments:

1. Indore Univeristy (13) one thesis submitted
2. Madras Univeristy (10) and 3 technical reports
3. Calcutta University (9)
4. Punjabi University (9) two theses in Design
5. University of Rajasthan (8) two theses in Design
6. Lucknow University (6)
7. Sardar Patel University (6)
8. Annamalai University (5)
9. Utkal University (2)
10. Osmania University (1)
11. Bombay University (1)

Total (70) research papers + 3 technical reports.

(5) Research Project work completed by departments in Design of Experiment:

Calcutta University

(i) On PBIB association scheme
(ii) On response surface designs
(iii) On asymmetrical factorial
(iv) Balanced factorials
(v) Estimation of variance components

Rajasthan University

(i) Construction and analysis of Multi-dimensional designs.

Lucknow University

(i) Optimality of designs in Regression Analysis

Punjabi University

(i) Construction of incomplete block designs and of designs for Bio-assays
(ii) PBIB designs, their uses in genetical experiments and their construction
(iii) Incomplete block designs and their construction

Bombay University

(i) Generalised group divisible designs and their properties and generalised Kronecker Product of Designs

Indore University

(i) Construction of Boss Block Designs
(ii) Construction of Semi-regular and Regular GD Designs

(6) Some of the suggestions to improve the existing courses and research facilities are as follows:

(i) Distinguished scholars (!) should be invited for talks (usuallly people forget about the talk after delivery, if not basically interested in the topic of the talk).

(ii) Encouragement must be provided by the government for writing monographs on Design of Experiments.

(iii) In India, we have the tradition of a Summer-Institute organised for four to six weeks on an all-India/regional basis on specialised topics in various subjects and it is suggested that a Summer-Institute be organised in Design of Experiments at all-India level.

(iv) One of the suggestions includes separate books for analysis and construction respectively on design of experiments.

(a) It is true that the only book available on construction of design is due to Raghav Rao and a good book supplementing the content of Raghav Rao is welcome. Das and Giri's book is good for its general use but a book on Analysis with rigorous treatment is also needed in the subject of design of experiments.

(b) Most of the universities in India are very poorly equipped with respect to libraries. Even if some journals are subscribed, it is no wonder if some of its issues are missing due to non-subscription (yearly), the reason being irregular and insufficient flow of grant from supporting agencies like the University Grants Commission. In fact, at least one university in a State must have all issues of the most important journals of the subject. Various universities of the State may be allowed different subjects in order to avoid any confrontation.

(c) The interaction between various academicians in the same area of the subject is minimum at present in India and any effort like regional seminars/discussions etc. including participation in a Summer Institute is a necessity.

(d) Education, Agricultural Scientists, Psychologists etc. are consumers of the products of design of experiments and interaction with them is going to help in evolving new methods in design which are lacking in the country. Consumers listed above use available techniques instead of asking a design of experiment person to modify the technique to suit them. In fact in the process, sometimes they tailor out even a good piece of data.

Reference

[1] Federer, W.T. and Federer, A.S. (1973) "A Study of Statistical Design Publications from 1968 through 1971". The American Statistician, Vol.27, pp.160-163.

CLINICAL APPROACH TO TRAINING STATISTICS TEACHERS

J. Rustagi, The Ohio State University

INTRODUCTION

Statistics is a relatively new field of academic and scientific endeavour. With the advent of new technology, especially the computer, the art of data analysis and statistical computing has attained a new perspective. With the emerging field of computer graphics, statistical sciences is bound to receive new and innovative techniques for solutions to many problems. In spite of substantial theoretical and experimental developments in statistics, the teaching of statistics has not received adequate attention by the profession. It is only recently that attempts have been made by the national and international statistical associations to address the issues involved. The latest such effort was made at The Ohio State University in organizing a conference in Teaching of Statistics and Statistical Consulting in 1980. The proceedings of this conference were published under the editorship of Rustagi and Wolfe (1982).

One of the important issues not addressed at that time was the training of teachers of statistics for schools and colleges. In this paper, we propose to provide a framework to enrich the training of statistics teachers at the college level. Recently the need of such training has been felt especially in the United States, after the Mathematical Association for America proposed courses in Statistics at the undergraduate level through its Committee on Undergraduate Programs in Mathematics. Such courses are being taught often by persons not really trained in the statistical sciences. We propose a "clinical" approach for training statistics teachers so as to provide "bedside" teaching of statistics, so often used in medical sciences. Such a training is especially suited to persons who have an understanding of the mathematical basis of the techniques and need to develop an appreciation of the nature of the statistical phenomenon.

STATISTICS AS AN IMPORTANT DISCIPLINE IN SCHOOLS AND COLLEGES

The applications of statistical techniques are made in a large number of scientific disciplines. The problems of basic design of experimentation, analysis of experimental results, acceptance or rejection of scientific hypotheses are extremely important to scientists. In the non-experimental sciences such as economics, the numerical information is utilized in developing models. As an important branch of knowledge, statistics is being required in the curricula of many undergraduate and graduate degree programs. In high schools an

understanding of probability and statistics is required of the students and statistics courses have been introduced at high school levels in many countries. Several mathematical techniques can be motivated through statistics. Suggestions to teach elementary mathematical concepts through statistics have been made by several mathematics teachers (for reference see Anslander et al. (1971)).

At present a usual unit in high school mathematical curriculum in American high schools includes concepts in probability and statistics. In some school systems separate courses in statistics have been introduced. The television course on probability and statistics in the early sixties by F. Mosteller and further developments in the high school curriculum (Mosteller, 1971), popularized the introduction of statistics in high schools in the United States and other countries.

The requirement that students take at least one course in statistics with prerequisites of calculus has been introduced in many undergraduate business curricula. A large number of universities and colleges provide an introductory course to various non-mathematical majors with only the minimum prerequisites. Such service courses are increasing at a fast pace and teachers without substantial training in statistics are being called upon to teach them. The Committee of Undergraduate Programs in Mathematics of the Mathematical Association of America (1972) has developed course contents for statistics courses and this has created a demand for statistics teachers in undergraduate colleges. Very few universities in the United States have independent Departments of Statistics and hence this responsibility falls on teachers of mathematics, psychology, economics etc. who have a side interest in statistics. In such cases, not only is it important to provide material for teaching but also to provide techniques for teaching.

Undergraduate programs in biological, social, behavioural and engineering sciences require statistics courses as prerequisites to their courses in their own disciplines. Increasing demand is being made for courses in statistics as a result of the introduction of interdisciplinary programs such as majors in the mathematical sciences which include statistics, mathematics and computer science. There is a growing demand for statistics teachers in schools and colleges at every level. A recent survey at The Ohio State University disclosed that more than two hundred faculty members were teaching, consulting or applying statistics as a major activity of their academic endeavour even though the Department of Statistics has only a faculty of fifteen (Report of Self Study Committee, 1981).

LACK OF TRAINED TEACHERS

In the recent past there has been an increase in the creation of new Departments of Statistics programs. Several new programs in American universities have been introduced for training master and doctoral level statisticians. Many undergraduate programs have teachers without formal training and practical experience in statistics. Such teachers provide, at times, only a formal treatment of the subject. Many times, the required courses are taught to the "unwilling" student by teachers having no enthusiasm and no experience with real application. They hardly transmit any stimulation and motivation, thus depriving a large majority of students of the real potentials of statistical sciences. Many potential scientists, educators and future leaders do not receive rich heritage of the exciting field of statistics. Lack of proper training in statistics at the college level not only affects the field in which it has direct application but prohibits the students from being exposed to one of the most important modern sciences.

To perform the task of teaching statistical concepts effectively, a basic understanding of statistical philosophy and practice is needed. The competent teacher should be able to provide the feeling of the real world problems and the application of statistics to these problems. The statistics teachers should be trained to impart this feeling to their pupils in an effective manner and with enthusiasm.

CLINICAL APPROACH

In order to provide statistics teachers with the subject matter of statistics, basic concepts in the theory of probability, statistical inference, design of experiments, etc. should be conveyed. Facility of utilizing well known statistical packages for numerical computations should be developed. Understanding of the problems of data analysis and development of statistical theory for this purpose can be made through the problem-oriented approach. This approach to training statistics teachers requires that the individual be associated with an actual problem in data analysis. The enthusiasm and insight of the experimenter can be easily transmitted to the trainee if he/she is associated in a meaningful way with the experiment. The purpose of this association is that the student is guided towards the appreciation of all aspects of experimentation, development of model, formulation of hypotheses, data collection, data analysis and the final disposition of the problem.

Another way of providing such a training is through internship programs in some industrial or business organization where the student is exposed to real problems. Practical difficulty of arranging such internships and the

limited contributions of such assignments to the training of statistics teachers, has made the method of little use. However, the flow of practical problems in a university setting is quite large and can provide sufficiently different problems of experimentation to give a broad survey of statistical concepts. In our experience at The Ohio State University, Statistics Laboratory, the range of problems in one quarter of 10 weeks, may see a problem in Home Economics, Anthropology, Educational Research, Engineering, Agriculture, Medicine, Music, Literature and so on. Many of these problems provide the student with a wide variety of applications of statistics to various fields.

The experimenter discusses the problem with a statistical consultant in the presence of the student. The problem is analyzed and statistical questions are formulated. The complete process of numerical evaluation of the results and final disposition of the problem is conducted with the experimenter in the presence of the student. A written report is prepared often and is discussed by all the parties.

The importance of having a statistical laboratory cannot be overemphasized. The laboratory provides a continuous flow of actual problems and gives the student teacher the experience needed. The techniques of teaching statistical consulting are also developed in a similar fashion. A series of papers on this matter is given in Rustagi and Wolfe (1982).

References

[1] Anslander, L., Avenoso, F.J., Dyer, E., Chiefetz, P., Gewirtz, A., Hunte, B.E., Quintas, L.V. and Rauch, H.E. (1971) Mathematics Through Statistics, Mathematical Dimensions, Ill., Chicago, Ill.

[2] Mosteller, F. (1971) The Joint American Statistical Association - National Council of Teachers of Mathematics Committee on the Curriculum in Statistics and Probability, Review of the International Statistical Institute, 39, 340-342.

[3] Panel on Statistics, The Introductory Statistics Courses, Committee on Undergraduate Program in Mathematics, Mathematical Association of America, Berkeley, California, 1972.

[4] Polya, George (1963) On learning, teaching, and learning teaching. The American Mathematical Monthly, 70: 6, 605.

[5] Report of the Self Study Committee, Department of Statistics (1981) The Ohio State University, Columbus, Ohio.

[6] Rustagi, J.S. and Wolfe, Douglas A. (Editors) (1982) Teaching of Statistics and Statistical Consulting, Academic Press, New York.

[7] Situmeang, R.P. (1976) Motivational supplements to lecture discussion in introductory college statistics courses, Ph.D. dissertation, The Ohio State University, Columbus, Ohio.

[8] Tanur, Judith et al. (Editors) (1972) Statistics: A Guide to the Unknown Holden-Day Incorporated, San Francisco.

DISCUSSION OF TEACHER-TRAINING PAPERS

A Morley, Trent Polytechnic

The papers from Professor Gupta and Professor Rustagi are concerned with the training of teachers of statistics at college level. They are particularly concerned with the kind of statistical training experienced by the prospective college teacher, especially in applied statistics, and this is clearly very important. However, they make limited comment on more direct pedagogic training. It may be useful for me to comment on this aspect of training for teachers at both college and school level.

Professor Gupta envisages the new college teacher taking part in established undergraduate courses with a set text. Typically in the U.K. he will be faced with much more responsibility for the presentation; often just given the syllabus. For example, there are no suitable texts for many services courses.

To help with taking this responsibility we have recently instituted required training for new teachers in state institutions for higher and further education, organised on a regional basis. In the first year the newcomer has a reduced teaching load, a three-day orientation course before teaching begins, tutorial supervision by an experienced member of his own department which includes observation of lectures, and a two-week course at the end of the year. In the second year he will attend for one day a week throughout the year a course organised by the Staff Development Officer.

As a supervisor the kind of things with which I am trying to give assistance are:

(i) Conceptual and task analysis of a topic.
Student background-levels of symbolism and justification to be used.

(ii) The shaping and presentation of a lecture; posing and answering a problem - key questions - aids.

(iii) Planning of workshop and seminar activities.

(iv) The overall sequencing of a syllabus. Time to be spent on various topics. The use of texts, handouts, tables, calculators and computers.

(v) Setting and marking of coursework, tests and end of year examinations.

At school level in the U.K. we have to distinguish sharply between training teachers for so-called 'A' level courses in statistics for a selective group (approximately top 15%) of 16-18 year olds and for teaching elementary probability and statistics within integrated mathematics courses for 11-16 year olds.

The number of 16-18 year old pupils taking statistics has grown enormously in recent years. The courses are ambitious, externally examined nationally and assume a concurrent study of the calculus. Those training to teach at this level will mainly be mathematics graduates with a variable knowledge of statistics, doing a one-year course of teacher-training (including 12 out of 36 weeks on teaching practice in schools). There is relatively little time to spend specifically on the teaching of statistics, typically 6-10 hours. The tutor is limited to considering some of the texts available (e.g. Hodge and Seed), the syllabus and type of examination of the different examination boards, including those where projects have a part and study of some key issues - I look at hypothesis-testing, the use of microcomputer graphics, and the Statlab data-bank. There is a great need for in-service courses for this group of teachers.

A similar small amount of time is available for considering statistics for 11-16 year olds and we have to rely on transfer from general experience in the teaching of mathematics. I focus on a few units from the 'Statistics in Your World' series looking at the notion of key questions, identifying skills and concepts, how to keep up-to-date with the data and the teachers' book which discusses the justification for statistics in the curriculum at this level.

TOPIC 7: USE OF CALCULATORS AND COMPUTERS IN TEACHING

ORGANISER: PROFESSOR L RÅDE

THE CASE FOR USING COMPUTERS IN TEACHING STATISTICS

B L Joiner, University of Wisconsin

ABSTRACT

The teaching of basic statistics courses can be greatly enhanced with the use of widely available statistical software. Pocket calculators can relieve some of the computational drudgery of statistics but are of little help with data plotting, one of the key steps in the statistical analysis process. By using computers in basic statistics courses, we can present students with real data sets and problems, and teach them how to approach and analyze data. In this paper, the basic steps in the statistical analysis process are listed, and the important role that computers can play in some of these steps is emphasized.

INTRODUCTION

What can be conveniently calculated and plotted has always played a major role in how statistics was done and in how it was taught. Until recently the best tools available to practitioners and students were mechanical calculators that could only add, subtract, multiply, divide and perhaps take a square root. These were used by many to supplement hand calculations.

In the past decade two important developments have offered new possibilities. These are pocket calculators and widely available easy-to-use computer software. Relatively inexpensive pocket calculators can now do higher level calculations such as compute standard deviations, fit straight lines and even certain types of non-linear functions, compute percent points and integrals of statistical distributions and simulate random data. Using these effectively, as is done, for example, by Koopmans (1981), can relieve much of the computational drudgery traditionally associated with basic statistics courses. However, a key weakness in these devices is their inability to speed up substantially the plotting of data.

Tukey (e.g., 1977) has introduced new graphical displays, many of which can be done by hand relatively easily. Combining these new displays with the use of pocket calculators does open exciting new possibilities for low-budget teaching, as Koopmans (1981) has demonstrated.

In the places where students can have access to computers, however, an even better approach is possible. This approach adds on easy-to-use computer languages such as Minitab (Ryan, Joiner, Ryan, 1976, 1982) to the tools traditionally available. The cost of this approach is decreasing rapidly and, in the United States, may now be as low as $.25 per half-hour session. If

a student had 20 half-hour sessions in a course, this would amount to about $5 per student, or roughly one-third to one-fourth the cost of a text book.

THE STATISTICAL ANALYSIS PROCESS

"I know how to do a t-test, I just don't know when to do one."

If our teaching is to do any good, our students must learn to judge when various statistical procedures are appropriate, what assumptions are critical, how to check for violation of assumptions, how to judge whether calculations have been done at least approximately correctly and how to interpret results. In other words, they must learn the basics of how to do statistics, not just do calculations.

The statistical analysis process has many steps, most of which need to be done in loops several times in any analysis. These steps include the following:

(1) Understanding something about the context of the data
(2) Understanding the goals of the analysis
(3) Checking the data values for reasonableness
(4) Assessing whether or not the data might shed light on the stated goals of the analysis
(5) Understanding something about the sources of variation or uncertainty in the data
(6) Making a variety of plots that increase understanding of the data
(7) Perhaps transforming or re-expressing the data in a more appropriate metric
(8) Selecting a computational procedure
(9) Carrying out the calculations
(10) Checking the calculations for reasonableness
(11) Interpreting the results

It is my view that the goal of a basic statistics course should be to help students understand as much as possible of the statistical analysis process. They will need to understand this process before they can begin to use statistics, of course. But, just as importantly, they will need to understand it if they are to be intelligent consumers of statistical results. Only by understanding some of what should be done, will they have any hope of being able to judge whether a given analysis has been done appropriately or not.

The concepts involved in this analysis process are apparently not very easy to most students. This is especially true if they become ensnarled in a mass of computational details before they begin to get an understanding of the process as a whole.

Computers and calculators can help some, but they do not provide a complete

solution by any means. Note in particular that pocket calculators can only help in steps (7) and (9). In fact, even computers can only help in steps (3), (6), (7) and (9). The rest must come from hard thinking. The best that any computational and plotting aid can do is to free the analyst's/student's mind from the need to worry about the details of the calculations or plots.

The important decisions of what to do are left to the analyst/student. But if computers have freed the teacher from the need to teach computational details, class attention can be focused instead on the more important concepts of the statistical analysis process.

POSSIBILITIES VERSUS TRADITION

In the past most attention in basic statistics courses has been focused on either calculations or on the properties of basic statistical formulas under the narrow assumptions of simple random sampling from a Gaussian distribution. It is now possible to change this emphasis and bring our teaching much more in line with how good statistical analysis is actually done.

Let us now consider the possibilities the computer affords us at the various steps in the analysis process.

(1) Understanding the context of the data. By using computers, more realistic data sets can be considered. They need not be made up, preselected or modified to make it easy for the calculations to be carried our nor to satisfy the narrow assumptions of random samples from Gaussian distributions. A richer variety of larger data sets in areas of interest to the students can be used.

(2) Understanding the goals of the analysis. Along with more realistic data sets comes a much broader class of possible goals. Thus the importance of choosing statistical procedures that are appropriate to the particular goals of the analysis can be emphasized.

(3) Checking the data. Almost every real data set has errors of the non-statistical sort. Dials are misread, impossible answers are recorded and so on. Learning how to look for these trouble spots with histograms, plots and frequency tables is made relatively easy by computers. It is also easy to dramatize the importance of such checking.

(4) Might data shed light on stated goals? More realistic problems can encourage students to think about how hard it really is to gather good data on many questions of interest. This is very important to their future roles as consumers of statistical analysis either in their everyday thinking or in their jobs.

(5) Sources of uncertainty. This again relates to the need to go beyond the assumption of a simple random sample. In many cases the data are, in no sense, any sort of random sample. We can, at best, use the results of standard procedures for guidance (see, e.g. Joiner, 1982; Tukey, 1982). In other cases more complex variance structures are present. Understanding these structures, and their impact on standard procedures is quite important.

(6) Making a variety of plots. Good plots are often very important to the statistical analysis process and in some cases even obviate the need for a more formal analysis. It is here that computers really excel. Plotting by hand a set of 100 observations is a very tedious, error prone activity. Few students or teachers would make more than a handful of such plots by hand in a course. Yet good practice often calls for a number of such plots in each analysis. Data must be plotted in a variety of ways; residuals from various analyses must be plotted in a variety of ways; subsets of data must be plotted; and so on. (See, for example, Joiner, 1981a).

(7) Transformations of data. Data are not necessarily recorded in the units most appropriate for their analysis and much can often be gained from a change of units: gasoline mileage may be better analyzed in gallons per mile than in miles per gallon; acidity in pH rather than in hydrogen ion concentration; toxicity in deaths per hour rather than in hours to death; stock prices in earnings per cost of stock rather than price earnings ratio; and so on. Computers make it easy to re-express data then replot and recompute. (See, e.g. Joiner, 1981b). And, they make these processes seem quite easy and natural.

(8) Selecting a computational procedure. Computers are of little help here, except that by making it easy to do a variety of calculations, they may make it more likely that a procedure will be chosen based on its appropriateness rather than on its ease of calculation. One minor problem is that computers make it almost too easy to do more analyses than are justified. Students should be warned against the trap of doing analysis after analysis until they find one that gives "statistically significant" results.

(9) Carrying out the calculations. Computers greatly facilitate this process in two ways. They obviously make the numerical calculations quite easy, but in addition, their ability to store a copy of the carefully checked data frees the student/analyst from the need to re-enter the data. They also make it easy to loop back and begin again at step (6) by plotting the data in new and more effective ways, and possibly retransforming the data, for another analysis.

(10) Checking calculations for reasonablness. After getting an answer it is important to reflect on whether that answer is reasonable or not given what else we know about the problem and the data. Computers are presently of little help here except in providing more time for this activity. One way they might help, although this seems to be not much exploited, is in providing drill exercises in which students are presented with a list of numbers or a histogram then asked to guess some characteristic of the data such as their standard deviation.

(11) Interpreting the results. Again computers are of little help except in freeing up time to concentrate on this important activity and in making it easy to do further numerical or graphical processing of results. For example, they make it easy to compute and plot predicted values and residuals.

SUMMARY

Computers can be a big help to us in teaching statistics. They can free us of the need to focus on computational details and narrow, unrealistic assumptions. To take full advantage of this opportunity we will need to free our minds of what we have here-to-fore taken for granted must be taught. We must do fresh thinking and study how good statistical design and analysis are actually done in practice - then see how we can best teach this process to our students. If we do this, I think our students will find statistics a much more interesting and relevant subject and will become better users of statistical procedures and better consumers of statistical analyses. We will also have more fun teaching it.

Acknowledgements

I am very grateful to Alison K. Pollack who made a number of suggestions which significantly improved this report.

References

[1] Joiner, Brian L. (1981a) "Lurking Variables: Some Examples". The American Statistician, 35, 227-233.

[2] Joiner, Brian L. (1981b) "How To Use Minitab Effectively in Teaching Statistics". Proc. Gothenburg Symposium on Teaching Statistics, Lennart Råde (ed.) 16-30.

[3] Joiner, Brian L. (1982) "Practicing Statistics or What They Forgot to Say in the Classroom", Teaching of Statistics and Statistical Consulting, Jagdish, S. Rustagi and Douglas A. Wolfe (eds.) Academic Press, New York, 327-342, 391-393.

[4] Koopmans, Lambert H. (1981) An Introduction to Contemporary Statistics, Duxbury Press, Boston, MA.

[5] Ryan, Thomas A. Jr., Brian L. Joiner and Barbara F. Ryan (1976) Minitab Student Handbook, Duxbury Press, Boston, MA.

[6] Ryan, Thomas A. Jr., Brian L. Joiner and Barbara F. Ryan (1982) Minitab Reference Manual, Minitab Project, University Park, PA.

[7] Tukey, John W. (1977) Exploratory Data Analysis, Addison-Wesley, Reading, MA.

[8] Tukey, John W. (1982) "Discussion of Papers by Calvin, Joiner and van Belle", Teaching of Statistics and Statistical Consulting, Jagdish S. Rustagi and Douglas A. Wolfe (eds.) Academic Press, New York, 379-389.

COMPUTER ALGORITHMS FOR THE CLASSICAL DISTRIBUTION FUNCTIONS

H O Posten, University of Connecticut

INTRODUCTION

Access to a collection of computer programs for the classical distribution functions, particularly if these can be readily used interactively by any student, is an important teaching tool for teachers of probability and statistics. The availability of such a collection allows the teacher and student to concentrate on the concepts and development of probability and statistics with little if any emphasis on the details of determining probabilities from tables.

By using the algorithms to be presented, any teacher with beginning level programming ability (or with the assistance of a colleague or student with this ability) can provide a full set of programs for evaluating binomial, Poisson, geometric, negative binomial, hypergeometric, normal, t, F, beta, gamma, and chi-square probabilities. In most cases, the actual algorithm will be reasonably simple and, for the continuous distributions, general enough to allow a user to increase accuracy by specifying the number of terms or steps to be used. Also, more than one algorithm for each distribution will be provided to allow a user to check the quality of one algorithm with another.

ALGORITHMS FOR DISCRETE DISTRIBUTIONS

Recursive Algorithm

A single recursive algorithm is very effective for evaluating probabilities for any of the discrete distributions previously mentioned. This algorithm is based on a common property of each of these distributions, that

$$f(x) = c(x-1)f(x-1) \qquad (1)$$

for all possible values of x appropriate to each distribution, where $f(x) = P(X=x)$ is the probability function and $c(x-1)$ is a simple multiplicative factor. The distribution function $F(x) = P(X \leq x)$ is evaluated by accumulating values of $f(x)$. For each distribution, the appropriate details are provided in Table 1. The geometric distribution is not provided in Table 1 since it is a special case of the negative binomial distribution with $r = 1$.

From Table 1 and formula (1), the general recursion algorithm is given in Table 2. This algorithm calculates the first probability $f(0)$ and, from this, each successive probability $f(k)$ until the probability for the desired value of x is reached. The individual probabilities $f(k)$ are accumulated at each

step to provide the cumulative probability. This algorithm, as listed in Table 2, provides only the values of f(x) and F(x) but can easily be modified to efficiently produce a complete table of probabilities. To do this, it is necessary to output the values of f(k) and F(k) for each value of k in steps 2 and 3 and to continue step 3 until the largest value of k is reached for the given distribution. For the two infinite range distributions, there is

TABLE 1: Elements of Recursive Algorithm for Discrete Distributions

distn.	f(x)	range of x	c(x-1)	f(0)
binomial	$\binom{n}{x}p^x q^{n-x}$	x = 0,1,...,n	$\frac{p(n-x+1)}{qx}$	q^n
Poisson	$\frac{e^{-m}m^x}{x!}$	x = 0,1,2,...	$\frac{m}{x}$	e^{-m}
negative binomial	$\frac{(r+x-1)!}{(r-1)!x!}p^r q^x$	x = 0,1,2,...	$\frac{(r+x-1)q}{x}$	p^r
hyper-geometric	$\frac{\binom{M1}{x}\binom{M2}{n-x}}{\binom{M1+M2}{n}}$	usually x = 0,1,...,n	$\frac{(M1-x+1)(n-x+1)}{x(M2-n+x)}$	$\frac{\binom{M2}{n}}{\binom{M1+M2}{n}}$

no largest value for k and the probability calculations must be arbitrarily terminated at some point to avoid continuing the calculations forever. A convenient strategy, in this case, is to stop the calculations when essentially all of the probability has been calculated. That is, prespecify an allowable RESIDUAL

TABLE 2: Recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Probabilities

STEP 1:	Input value of x and the distribution parameters (e.g., n and p for the binomial).
STEP 2:	Evaluate f(0) and F(0) = f(0).
STEP 3:	Evaluate f(k) = c(k-1)f(k-1) and F(k) = F(k-1) + f(k) for k = 1,2,...,x.
STEP 4:	Output x, f(x), F(x).

amount of probability and terminate the calculations when the accumulated probability exceeds 1-RESIDUAL. By specifying RESIDUAL to be quite small, all but a negligible amount of probability will be provided in the resulting table. Table 3 provides the modified algorithm for producing a table of probabilities for any of the listed discrete distributions.

TABLE 3: Recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Tables

STEP 1:	Input distribution parameters (and RESIDUAL for an infinite range distribution).
STEP 2:	Evaluate f(0) and F(0) and output 0, f(0), F(0).
STEP 3:	Evaluate f(k) = c(k-1)f(k-1) and F(k) = F(k-1) + f(k) and output k, f(k), F(k) for k = 1,2,3,...,d, where d is the largest possible value for k.
STEP 3*	(replaces STEP 3 when distribution range is infinite): Evaluate f(k) = c(k-1)f(k-1) and F(k) = F(k-1) + f(k) and output k, f(k), F(k) for k = 1,2,... until F(k) > 1-RESIDUAL.

Except for hypergeometric distributions, the calculation of f(0) for the algorithms in Tables 2 and 3 is usually quite simple using the forms given in Table 1. The calculation of f(0) for the hypergeometric distribution, however, presents some difficulties since f(0) is the ratio of two combinatorials each involving factorials. A special approach to the evaluation of f(0) is therefore needed for the hypergeometric distribution. It is easy to show for this distribution that f(0) may be written as

$$f(0) = \prod_{j=1}^{n} \frac{M2-j+1}{M1+M2-j+1}, \qquad (2)$$

and a simple program loop can be used to evaluate f(0) as a product of ratios.

The foregoing indicates that it is not difficult to write a recursive program to evaluate any of the indicated distributions with the use of (1), even for the hypergeometric distribution. For the latter distribution, this holds provided the range is x = 0,1,2,...,n. Occasionally, one runs into a hypergeometric distribution for which the sample size is larger than one or both of M1 and M2. In this case, at least one end of the range is restricted. If a general hypergeometric program is required to handle this case, as well as the usual case, the programming is a bit more difficult. It must be taken into account that the general range for the hypergeometric random variable is x = c,c+1, ...,d, where c = max(0,n-M2) and d = min(n,M1). The lowest possible value of x may, therefore, be zero or larger than zero. In the latter case, the algorithms of Tables 2 and 3 must be modified so that f(0) is replaced in STEP 2 by f(c) when c > 0, and in STEP 3, k should start at c+1. It is not difficult to show that (when n > M2) f(c) may be written as

$$f(c) = \prod_{j=1}^{M2} \frac{n-j+1}{M1+M2-j+1} \qquad (3)$$

The form (3) makes it easy to evaluate f(c) as a product of ratios. Also, when producing a table for the hypergeometric distribution, the calculations in STEP 3 (Table 3) should be terminated when k reaches the maximum value

min(n,M1). For this much generality, some additional programming effort is required. The result is worth the effort, however, since any hypergeometric problem can be handled without tables.

In general, the foregoing recursive method accurately evaluates probabilities for any of the listed discrete distributions, even when the range of x is large. It is fairly easy, using this algorithm, to achieve at least ten decimal accuracy by carrying enough digits in the calculations (for example, 16 decimals in double precision calculations on an IBM 360/370 System).

A Warning and a Modification

On reflection, the foregoing recursive algorithm is seen to have one flaw. Should f(0) have a value smaller than the smallest magnitude allowed by the computer, underflow will occur and f(0) will be read as zero by the computer. In this event, all subsequent probabilities f(k) will be calculated as zero since each is the product of a multiplier and the previously calculated probability. (Such an error can be avoided by incorporating into the program at STEP 2 of the algorithm a statement that stops the program when f(0) is read as a zero and informs the user to use another method). In this case, for the finite range random variables one can reverse the recursive algorithm starting at the highest possible value of the random variable and accumulate probabilities backwards instead of forwards. It is possible that both the first and last probabilities can be read as zero because both are too small, but many times this backwards approach can be useful in sidestepping the problem of f(0) being too small.

In practice, the problem of f(0) being too small does not seem to arise for the range of hypergeometric parameter values usually met. Also, for the negative binomial distribution, it is unlikely that this problem would occur for the range of p and r usually met by teachers and students. For this to occur, p would have to be particularly small and/or r would have to be unusually large. The problem does occur for the binomial and Poisson distributions (for the Poisson, m > 180 and, for the binomial with p = .50, n > 260 will yield an f(0) that is too small for an IBM 360/370 System). Consequently, it is wise to provide a program termination when such a problem occurs and then use a backup program. An effective backup procedure for such cases is a modification of the foregoing recursive algorithm which uses a natural logarithm and an exponential operation at each step. Specifically, from (1),

$$\ell n\ f(x) = \ell n\ c(x-1) + \ell n\ f(x-1) \qquad (4)$$

for all appropriate values of x for each of the listed distributions. The modification evaluates $\ell n\ f(0)$ ($\ell n\ f(0) = n\ \ell n\ q$ for the binomial and $\ell n\ f(0) = -m$

for the Poisson distribution) and subsequently $\ell n\ f(k)$ from this using (4) repeatedly. The values of f(0) and f(k) are obtained by exponentiating these natural log values. This avoids the problem of f(0) being too small for the computer for any of the four discrete distributions in Table 1. The appropriate algorithms are provided in Tables 4 and 5.

From a teaching viewpoint, the author prefers to use the recursive algorithm in its simplest form with a terminating statement that informs the user to use another method when f(0) is read as zero. In this form, the algorithm is very simple and can readily be explained to anyone with algebra background, including high school students. As a backup method, for the binomial and Poisson, programs using the log-recursive algorithm are quite effective, ordinarily achieving at least 10 decimal accuracy when enough digits are carried throughout the calculations. Such a backup program is unnecessary for the hypergeometric distribution and probably for the negative binomial, though either can be developed from Tables 4 and 5.

TABLE 4: Log-recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Probabilities

STEP 1:	Input value of x and the distribution parameters.
STEP 2:	Evaluate $\ell n\ f(0)$ (= $n\ \ell n\ q$ for binomial and -m for Poisson), $f(0) = \exp(\ell n\ f(0))$, and $F(0) = f(0)$.
STEP 3:	Evaluate $\ell n\ f(k) = \ell n\ c(k-1) + \ell n\ f(k-1)$, $f(k) = \exp(\ell n\ f(k))$, and $F(k) = F(k-1) + f(k)$ for $k = 1,2,\ldots,x$.
STEP 4:	Output x, f(x), F(x).

TABLE 5: Log-recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Tables

STEP 1:	Input distribution parameter values (and RESIDUAL for an infinite range distribution).
STEP 2:	Evaluate $\ell n\ f(0)$, $f(0) = \exp(\ell n\ f(0))$, $F(0) = f(0)$ and output 0, f(0), F(0).
STEP 3:	Evaluate $\ell n\ f(k) = \ell n\ c(k-1) + \ell n\ f(k-1)$, $f(k) = \exp(\ell n\ f(k))$, and $F(k) = F(k-1) + f(k)$ and output k, f(k), F(k) for $k = 1,2,\ldots,d$, where d is the largest possible value for k.
STEP 3*	(replaces STEP 3 when distribution has an infinite range): Evaluate $\ell n\ f(k) = \ell n\ c(k-1) + \ell n\ f(k-1)$, $f(k) = \exp(\ell n\ f(k))$, $F(k) = F(k-1) + f(k)$ and output k, f(k), F(k) for $k = 1,2,\ldots$ until $F(k) > 1-\text{RESIDUAL}$.

Direct Term by Term Evaluation of Individual Probabilities

A relatively simple algorithm for any of the distributions in Table 1 evaluates the probability f(x) by evaluating each factor and multiplying these together. Such individual probabilities can then be added for cumulative probabilities.

This approach is not usually practical because the combinatorials involved in each of the f(x) can become too large for the computer. This approach, however, becomes practical if logarithms are used and one has an accurate algorithm for evaluating the logarithm of a factorial. The natural logarithm of a factorial can be evaluated using $\ell n\ x! = \ell n\ \Gamma(x+1)$, where $\Gamma(y)$ is the gamma function. An excellent algorithm for evaluating $\ell n\ \Gamma(y)$ will be provided later on in the section headed 'Algorithms for Evaluating the t- and F-Distribution Functions'. The general algorithm is simple and will be illustrated using the binomial distribution. For this distribution,

$$\ell n\ f(k) = \ell n\ \Gamma(n+1) - \ell n\ \Gamma(k+1) - \ell n\ \Gamma(n-k+1) + k\ \ell n\ p + (n-k)\ \ell n\ q. \tag{5}$$

The value of $f(k) = \exp(\ell n\ f(k))$ is obtained from (5) for all required values of $k = 0,1,\ldots,x$ and accumulated to provide the cumulative probability $F(x)$. Corresponding algorithms for the other three discrete distributions are obtained similarly. This method is slower than the two recursive methods, but it is effective and can achieve 10 or more decimal place accuracy (carrying 16 decimals in the computer). For the evaluation of a single probability, f(x), it is relatively fast and most serviceable. For cumulative probabilities it is useful as an independent algorithm for checking the quality of the two recursion algorithms. For teaching purposes, the latter are better since they are faster and do not involve factorials, thereby allowing easier explanation.

Other Methods for Discrete Distributions

A fourth effective method for evaluating a cumulative binomial probability is bsed on the identity

$$F(x;p,n) = \sum_{k=0}^{x} \binom{n}{x} p^x q^{n-x} = 1 - I_p(x+1,n-x)\ , \tag{6}$$

where

$$I_y(a,b) = \Gamma(a+b)\ [\Gamma(a)\Gamma(b)]^{-1} \int_0^y t^{a-1}(1-t)^{b-1}dt \tag{7}$$

is the cumulative beta distribution function. The identity (6) is given by various authors including Greenwood and Hartley (1962) and Kennedy and Gentle (1980). With an accurate algorithm for evaluating the beta distribution function, (6) provides a good algorithm for evaluating the cumulative binomial distributive. Three good algorithms will be provided in the section 'Algorithms for the Beta Distribution Function' for evaluating the beta distribution function. Using these beta algorithms, very good accuracy (ten or more decimals, if enough digits are carried in the calculations) can often be achieved. In fact, using the three methods for evaluating the cumulative binomial probability (recursive, log-recursive, and binomial-beta) often leads to 10 or more decimal place agreement between the three methods (using double precision on an IBM

360/370 computer).

A third effective method for evaluating a Poisson cumulative probability is based on the identity

$$F(x;m) = \sum_{k=0}^{x} \frac{e^{-m} m^k}{k!} = 1 - G_{2m}(x+1,2) \tag{8}$$

given by Greenwood and Hartley (1962) and by Kennedy and Gentle (1980) where

$$G_y(a,b) = [\Gamma(a)b^a]^{-1} \int_0^y t^{a-1} e^{-t/b}\ dt \tag{9}$$

is the gamma distribution function. The quality of this Poisson algorithm will depend upon the quality of the gamma algorithm. This, however, will present no problem since three excellent gamma algorithms will be provided in the section 'Algorithms for the Gamma Distribution Function'.

ALGORITHMS FOR CONTINUOUS DISTRIBUTIONS

Algorithms for the Normal Distribution

Probably the most effective method for evaluating a single standard normal probability is a continued fraction expansion. Such an approach is given by Bargmann (1969) as

$$\Phi(z) = 1 - \phi(z)CF \tag{10}$$

where $\Phi(z)$ is the standard normal distribution function, $\phi(z)$ is the corresponding density function,

$$CF = \frac{z}{1-}\ \frac{z^2}{3+}\ \frac{2z^2}{5-}\ \frac{3z^2}{7+}\ \frac{4z^2}{9-}\ \ldots \quad \text{for } 0 < z \leq 2.5 \tag{11}$$

and

$$CF = \frac{1}{z+}\ \frac{1}{z+}\ \frac{2}{z+}\ \frac{3}{z+}\ \frac{4}{z+}\ \ldots \quad \text{for } z > 2.5 \tag{12}$$

The expansions (11) and (12) are continued fraction expansions which are truncated (i.e., terminated at some step) for evaluation and are then rational functions. For example, (11) is

$$CF = \cfrac{z}{1-\cfrac{z^2}{3+2z^2-\cfrac{3z^2}{7+4z^2}}}$$

if the truncation is made at the indicated step. It is often convenient to evaluate continued fractions recursively. In this direction, (11) may be evaluated by the steps in Table 6 and (12) by the steps in Table 7. Tables 6 and 7 provide a quick and easily programmed method for evaluating CF. The cumulative

normal probability, $\Phi(z)$, is then readily calculated by (10).

TABLE 6: Recursive Evaluation of Normal Continued Fraction ($0 < z \leq 2.5$)

STEP 1:	Evaluate for prespecified N, $W(N) = 2N-1+N(-1)^N z^2$.
STEP 2:	For $K = N-1, N-2, \ldots, 1$, evaluate $W(K)=2K-1+K(-1)^K z^2/W(K+1)$.
STEP 3:	Evaluate $CF = z/W(1)$.

TABLE 7: Recursive Evaluation of Normal Continued Fraction ($z > 2.5$)

STEP 1:	Evaluate for prespecified N, $W(N) = z + N/z$.
STEP 2:	For $K = N-1, N-2, \ldots, 1$, evaluate $W(K) = z + K/W(K+1)$.
STEP 3:	Evaluate $CF = 1/W(1)$.

A careful study by one of my students, P. Hormel, for $z = 0.1$ to 10.0 by .10 showed that less than 40 erms in the continued fraction were necessary to provide at least 13 decimal accuracy for all z in the range $0 < z < 10$, using double precision calculations on an IBM 360/370 computer.

A second algorithm which also provides very good accuracy is the series expansion (13) provided by Abramowitz and Stegun (1964).

$$\Phi(z) = \frac{1}{2} + \phi(z) \sum_{k=0}^{\infty} T_k, \tag{13}$$

and

$$T_k = \frac{z^{2k+1}}{1(3)(5)\ldots(2k+1)}. \tag{14}$$

The normal distribution function, $\Phi(z)$, is evaluated with a truncation of the sum in (13), which itself can be evaluated using the recursive relation

$$T_0 = z, \; T_k = \frac{z^2}{2k+1} T_{k-1}, \; k = 1,2,\ldots . \tag{15}$$

The specific algorithm is given in Table 8. because of teh recursive element, this algorithm is reasonably fast. It is also quite accurate. The study by P. Hormel showed that 100 terms will provide at least 12 decimal accuracy for $0 < z < 9.3$ and at least nine decimal accuracy for values of z up to $z = 10$.

A third very effective method for evaluating the normal distribution function is the simple method of numerical integration known as Simpson's method. This method divides the range of integration into N intervals of length 2H and then approximates the area over each interval by a simple quadratic approximation to the normal density function. The specific algorithm for evaluating a cumulative standard normal probability, from z = L to z = U, is

$$P(L < z < U) = \sum_{k=1}^{N} A_k \; , \tag{16}$$

where

$$A_k = (H/3)[\phi(z_{2k}) + 4\phi(z_{2k-1}) + \phi(z_{2k-2})]$$

is the approximate area over the range z_{2k-2} to z_{2k} and $z_m = L+mH$. The approximations A_k are then accumulated. This method is surprisingly effective for the normal distribution, even with relatively large values of H, but the most effective value of H (using double precision on an IBM 360/370) is about H = .005. With this value of H, accuracy of at least 10 decimal places is obtained for all values $0 < z < 10$.

TABLE 8: Series Algorithm for the Normal Distribution Function

STEP 1:	Input z and N = number of terms to be used in the series in (13).
STEP 2:	Evaluate $T_0 = z$, $SUM = T_0$.
STEP 3:	For $k = 1,2,\ldots, N$, evaluate $T_k = [z^2/(2k+1)]T_{k-1}$ and $SUM = SUM + T_k$.
STEP 4:	Evaluate $\Phi(z) = (1/2) + \phi(z)$ SUM and output z, $\Phi(z)$.

Of the three methods, the most effective for evaluating a single normal probability is the continued fraction algorithm because of its high accuracy when using only a small number of terms in the continued fraction. However, the Simpson method algorithm is the most effective for producing a table of standard normal probabilities, sicne the area, A_k, over each of the N intervals is directly calculated and accumulated. The evaluation of each A_k is fast, and it is a simple matter to output accumulated areas as each A_k is calculated

to provide a table. Students find it interesting to produce their own normal tables which are better than those in any book. The programming challenge is to print out the entire table on one or two regular computer pages (11 by 14 inches). A table of cumulative standard normal probabilities for z = 0.00 to 7.00, by 0.01, can be organized in 70 lines with 10 probabilities in each row, each provided to nine decimal places.

Algorithms for Evaluating the t- and F-Distribution Functions

The distribution functions for the t- and F-variates are quite simple to evaluate, provided an algorithm is available for the beta distribution function $I_y(a,b)$ defined by (7). For the t-variate with v degrees of freedom, the identity

$$P(t_v < t) = (1/2)[1 + I_x(1/2,v/2)], \text{ for } t > 0, \tag{17}$$

with

$$x = t^2/(v+t^2)$$

provides the algorithm. For $F(v_1,v_2)$, the F-variate with v_1 and v_2 degrees of freedom, the identity

$$P(F(v_1,v_2) < F) = I_x(v_1/2,v_2/2) \tag{18}$$

with

$$x = (v_1/v_2)F/[1 + (v_1/v_2)F]$$

yields the algorithm. The identities (17) and (18) are given by Johnson and Kotz (1970).

Algorithms for the Beta Distribution Function

An effective way to evaluate a beta distribution function is to use a continued fraction expansion. One such algorithm, due to Muller (1931) and given by Abramowitz and Stegun (1964), Bargmann and Ghosh (1963), Boardman and Kopitzke (1975) and Kennedy and Gentle (1980), is

$$I_x(a,b) = \frac{\Gamma(a+b)}{\Gamma(a+1)\Gamma(b)} x^a(1-x)^{b-1}CF \tag{19}$$

where

$$CF = \frac{1}{1-}\frac{c_1}{1+}\frac{d_1}{1-}\frac{c_2}{1+}\frac{d_2}{1-}\cdots \tag{20}$$

is a continued fraction with coefficients

$$c_k = \frac{(b-k)(a+k-1)x}{(a+2k-2)(a+2k-1)(1-x)}, \quad d_k = \frac{k(a+b+k-1)x}{(a+2k)(a+2k-1)(1-x)}. \tag{21}$$

One problem with this method is the requirement to evaluate the gamma functions in (19). This can be handled by using an efficient algorithm for evaluating the logarithm of a gamma function. An excellent procedure is provided in the section 'Algorithm for the Natural Logarithm of the Gamma Function'. For the evaluation of (20), it is effective to use a recursive form, as given in Table 9.

TABLE 9: Continued Fraction Algorithm for the Beta Distribution Function

STEP 1:	Input parameter values a and b and the values of x and n = the number of recursive steps to be used (the actual number of steps used in the continued fraction will then be 2n).
STEP 2:	Evaluate $S_n = c_n/(1+d_n)$, using (21) to evaluate c_n and d_n.
STEP 3:	For k = n-1,n-2,...,1, evaluate $S_k = c_k/[1+d_k/(1-S_{k+1})]$, using (21) to evaluate c_k and d_k.
STEP 4:	Evaluate CF = $1/(1-S_1)$ and LA = ℓn $\Gamma(a+b)$ - ℓn $\Gamma(a+1)$ - ℓn $\Gamma(b)$.
STEP 5:	Evaluate LB = LA + a ℓn x + (b-1) ℓn(1-x) + ℓn CF and $I_x(a,b)$ = exp(LB).

Bargmann and Ghosh indicate that this method is effective over a wide range of values of a and b, providing accuracy of at least 10^{-8} using 80 terms in the continued fraction (n = 40 in Table 9). However, Kennedy and Gentle recommend the use of this method for x < (a-1)/(a+b-2) and the use of

$$I_x(a,b) = 1 - I_{1-x}(b,a) \tag{22}$$

for larger x. The present author's experience with this algorithm indicates that often 10 or more decimal place accuracy is achieved using significantly less than 80 continued fraction terms, even above the range for x given by

Kennedy and Gentle. Also, a study by one of my students, J. Leitz, over 91 beta distributions with $a \geq b$ and $0.1 \leq a,b \leq 500$ and over 15 x-values in the probability range 0.0001 to 0.9999 for each distribution, supports this conclusion. Leitz's study, though, suggests that this continued fraction algorithm may not be very effective for the case when a and b are both less than unity, as also suggested by Bargmann and Ghosh. An exception to this is the case $a = b = 0.5$, when the method is quite accurate (10 or more decimals).

Another effective algorithm for evaluating the beta distribution function is based on a Fourier series Chebyshev polynomial expansion (Woods and Posten (1977)). Computationally, the algorithm is effective because a recursive strategy is used with the expansion

$$I_x(a,b) = 1 - (\theta/\pi) - S \tag{23}$$

where

$$\theta = \cos^{-1}(2x-1) \quad \text{and} \quad S = \sum_{j=1}^{\infty} b_j \sin(j\theta) , \tag{24}$$

where the b_j coefficients will be defined. The series S is evaluated by truncating

$$S_n = \sum_{j=1}^{n} b_j \sin(j\theta) \tag{25}$$

at some convenient value for n. The coefficients are evaluated recursively by

$$b_1 = (2/\pi)(a-b)(a+b)^{-1} \tag{26}$$

$$b_2 = (1/\pi)(a+b)^{-1}(a+b+1)^{-1}[2(a-b)^2-(a+b)(a+b-1)]$$

$$b_{j+2} = (a+b+j+1)^{-1}[2(a-b)\frac{j+1}{j+2} b_{j+1} + (j+1-a-b)\frac{j}{j+2} b_j)],$$

$$j = 1,2,\ldots,n-2.$$

S_n is then evaluated recursively and rapidly (without use of any trigonometric functions) by

$$U_{n+1} = U_{n+2} = 0 \tag{27}$$

$$U_k = b_k + (4x-2)U_{k+1} - U_{k+2} , \quad k = n,n-1,\ldots,1$$

$$S_n = 2U_1 (x-x^2)^{1/2} .$$

The algorithm is given in Table 10.

Studies by several students (F. Rockhold, A. Tupek and T. Ashbee) and myself indicate that this algorithm is effective over a broad range of values of a, b and x, achieving eight or more places if enough terms are used. Usually 200 terms provide at least 6-8 decimal accuracy and 500 terms, at least 8-10 decimal accuracy. Because of the recursive approach these terms are computed rapidly. The method is not as accurate as indicated if the distri-

TABLE 10: Fourier Series Chebyshev Polynomial Algorithm for the Beta Distribution Function

STEP 1:	Input values of a, b, x, and n = number of terms to be used in the series (25)
STEP 2:	Evaluate b_k, $k = 1,2,\ldots,n$, recursively from (26).
STEP 3:	Evaluate U_1 recursively from (27), and S_n from (27).
STEP 4:	Evaluate $I_x(a,b) = 1 - (\theta/\pi) - S_n$.

bution shape is such that the probability is strongly compressed towards either endpoint (e.g., a = 100, b = 2). In this case it appears that one can usually expect four or five decimal accuracy with 200 terms. Additionally, this method is quite weak when either parameter is very near zero (a or $b < 0.2$) and is not very strong when either parameter is small ($0.2 \leq a$ or $b \leq 1.4$), though in the latter case one can anticipate several decimal place accuracy. This last weakness is an important consideration since, as indicated in the section 'Algorithms for Evaluating the t- and F-Distribution Functions', F-probabilities are evaluated using a beta algorithm. The beta parameters are each one-half of the corresponding degrees of freedom of the F. Thus, for example, a = 1/2 and a = 1 are important parameter values since they correspond to F-variates with one and two degrees of freedom. Fortunately, the general weakness for small parameter values does not exist when a or b = 1/2 (unless the other parameter value is a different low value, ≤ 1.4). The opposite is, in fact, true. When a or b = 1/2, the method has very high accuracy. On the other hand, when a or b = 1, the method is weak often providing as little as three decimal accuracy. In this case (and other cases), one can use an identity which evaluates the probability after first incrementing the low parameter. The following forms, provided by Abramowitz and Stegun (1964), increment the first parameter by one each time they are used:

$$I_x(a,b) = [\Gamma(a+b)/\Gamma(a+1)\Gamma(b)]x^a(1-x)^{b-1} + I_x(a+1,b-1) \tag{28}$$

$$I_x(a,b) = (1/x)[I_x(a+1,b) - (1-x)I_x(a+1,b-1)] . \tag{29}$$

Similar identities exist for incrementing the second parameter but (22) permits (28) or (29) to be used for this purpose.

A second continued fraction expansion can be used as a third method for evaluating the beta distribution function. This expansion, from Aroian (1941, 1959), is also given by Abramowitz and Stegun (1964), Bouver and Bargmann (1975), Boardman and Kopitzke (1975) and by Kennedy and Gentle (1980).

$$I_x(a,b) = \frac{\Gamma(a+b)x^a(1-x)^b}{\Gamma(a+1)\Gamma(b)} \left[\frac{1}{1-}\frac{c_1}{1+}\frac{d_1}{1-}\frac{c_2}{1+}\frac{d_2}{1-}\cdots\right] \quad (30)$$

with

$$c_k = \frac{(a+k)(a+b+k)x}{(a+2k)(a+2k+1)} \quad \text{and} \quad d_k = \frac{k(b-k)x}{(a+2k-1)(a+2k)} . \quad (31)$$

The algorithm, in recursive form, is identical to that in Table 9 using (30) and (31) in place of (19), (20) and (21) and noting that the exponent on (1-x) is increased by one in (30) as compared to (19) (i.e., b replaces b-1 in STEP 5). Bouver and Bargmann indicate that this continued fraction expansion provides a superior algorithm for evaluating the beta distribution function over an extensive range of values for a and b. They refer to the range $10^{-8} < a,b < 70{,}000$ and specifically state that the method is efficient even for extremely J-shaped and U-shaped distributions. They recommend its use for $x < a/(a+b)$ (i.e., for x < mean), and the use of the complement form (22) for larger x.

Algorithms for the Gamma Distribution Function

Several algorithms for evaluating the gamma distribution function are quite effective over a wide range of parameter values. These algorithms can also be used to evaluate the chi-square distribution function, since the latter is a special case of the gamma distribution function

$$F(x;a,b) = [1/\Gamma(a)b^a] \int_0^x t^{a-1}e^{-t/b}\, dt , \quad (32)$$

with b = 2 and the degrees of freedom for the chi-square = 2a. The evaluation of (32) can be reduced to evaluating a gamma distribution function with b = 1,

$$I_x(a) = [1/\Gamma(a)] \int_0^x t^{a-1}e^{-t}\, dt, \quad (33)$$

since $F(x;a,b) = I_{x/b}(a)$. The algorithms to be presented will be for the form (33).

A very good continued fraction expansion for the gamma distribution, given by Khovanskii (1963), is

$$I_x(a) = [x^a e^{-x}/\Gamma(a)]CFK , \quad (34)$$

where

$$CFK = \frac{1}{a-}\frac{ax}{1+a+x-}\frac{(1+a)x}{2+a+x-}\cdots\frac{(n-1+a)x}{n+a+x-}\cdots \quad (35)$$

is a continued fraction expansion which can be evaluated recursively using CFK = 1/(a-T(1)), where

$$T(n) = \frac{(n+a-1)x}{n+a+x}, \quad T(k) = \frac{(k+a-1)x}{k+a+x-T(k+1)}, \quad k = n-1,n-2,\ldots,1 . \quad (36)$$

The recursion algorithm is given in Table 11. Two of my students, K. Sullivan and M. El-Bolkiny, provided studies of this algorithm for 41 values of a in the range 0.05 to 200.25 over 18 x-values ranging in probability from 0.0001 to 0.999 or higher (except for a < 1, the lowest actual x-value was about x = 0.001). They found the Khovanskii method to be very effective, achieving at least 12 decimal place accuracy with at most 100 continued fraction terms (except for a > 100, up to 200 terms were required for this accuracy for some x-values).

A second continued fraction algorithm for the gamma distribution function is given by Boardman and Kopitzke (1975). They actually provide two continued fractions, one for lower x-values and the other for upper x-values. The two forms are

$$I_x(a) = [1/\Gamma(a+1)]x^a e^{-x}/CFL , \quad (37)$$

where

$$CFL = 1 - \frac{x}{a+1+}\frac{x}{a+2-}\frac{(a+1)x}{a+3+}\frac{2x}{a+4-}\cdots\frac{(a+n-1)x}{a+2n-1+}\frac{nx}{a+2n-}\cdots \quad (38)$$

and

$$I_x(a) = 1 - x^a e^{-x}[\Gamma(a)]^{-1}/CFU , \quad (39)$$

where

$$CFU = 1 + \frac{1-a}{x+}\frac{1}{1+}\frac{2-a}{x+}\frac{2}{1+}\cdots\frac{n-a}{x+}\frac{n}{1+}\cdots . \quad (40)$$

CFL and CFU are continued fraction expansions to be used for lower and upper values of x, respectively. Boardman and Kopitzke suggest using (37) when x < a and (39) otherwise. A recursive algorithm for evaluating CFL is

$$\begin{aligned} T(n) &= x(1-n-a)/[a+2n-1+nx/(a+2n)] \\ T(k) &= x(1-k-a)/[a+2k-1+kx/(a+2k+T(k+1))], \quad k = n-1,\ldots,2, \\ CFL &= 1-x/[a+1+x/(a+2+T(2))], \end{aligned} \quad (41)$$

and for CFU,

$$\begin{aligned} T(n) &= (n-a)/(x+n) \\ T(k) &= (k-a)/[x+k/(1+T(k+1)], \quad k = n-1,\ldots,2, \\ CFU &= 1 +(1-a)/[x+1/(1+T(2))]. \end{aligned} \quad (42)$$

The recursive algorithm is provided by (37) and (41) for the lower expansion, and by (39) and (42) for the upper expansion. The pattern of the algorithm for each expansion is similar to that of the algorithm of Table 11.

Two of my students, E. Cook and G. Karbowski, provided studies of this algorithm over a set of 27 a-values in the range $0.1 \le a \le 200.25$ and for seven values of x in the probability range 0.001 to 0.999 or higher (except for a < 1,

TABLE 11: Continued Fraction Algorithm for the Gamma Distribution Function

STEP 1: Input values of a, x and n = number of terms to be used in the continued fraction expansion.

STEP 2: Evaluation T(n) and T(k), k = n-1,...,1, from (36). Evaluate CFK = 1/(a-T(1)).

STEP 3: Evaluate L = a ℓn x - x - ℓn Γ(a) + ℓn CFK and $I_x(a) = \exp(L)$.

the lowest actual x was about x = 0.001). They found that the lower continued fraction was superior to the upper for the range suggested by Boardman and Kopitzke (and slightly beyond), except for a < 1 where CFL was found to be superior for x-values up to a probability level of about 0.999. Their recommendation was to use CFL for x-values up to a probability of 0.999 when a < 1 and up to a probability of 0.7 for $1 \leq a \leq 200.25$, and to use CFU elsewhere in this a-range. They found that this strategy provided accuracy of at least 12 decimals using at most 200 terms in the continued fraction. An exception to this was the largest a-value (a = 200.25) which required, for one x-value, 400 terms to achieve 12 decimal accuracy.

A third extremely effective algorithm for the gamma distribution function is the series method of Khamis (1965), also provided by Battacharjee (1970),

$$I_x(a) = \frac{x^a e^{-x}}{\Gamma(a)} \sum_{k=0}^{\infty} \frac{x^k}{a(a+1)\ldots(a+k)} , \qquad (43)$$

for which the terms of the series can be evaluated recursively. If T_k represents the k-th term in the sum, $T_0 = 1/a$ and $T_k = (x/(a+k))T_{k-1}$ for k = 1,2,... . A study by M. El-Bolkiny over 41 a-values from 0.05 to 200.25 and 18 x-values in the range 0.0001 to 0.999 (except the lowest x was x = 0.001 for a < 1) showed this series algorithm to be surprisingly strong for all a and x in the range studied. At least 12 decimal accuracy was achieved in all cases using only 10 terms in the series. As a consequence, this algorithm seems superior to the previous two algorithms because of its efficiency and accuracy, at least over the range indicated. The appropriate algorithm is given in Table 12.

TABLE 12: Series Algorithm for the Gamma Distribution Function

STEP 1: Input values of a, x and n = number of terms to be used in the series expansion.

STEP 2: Evaluate $T_0 = 1/a$ and SUM = T_0.

STEP 3: For k = 1,2,...,n, evaluate $T_k = [x/(a+k)]T_{k-1}$ and SUM = SUM + T_k.

STEP 4: Evaluate L = a ℓn x - x - ℓn Γ(a) + ℓn SUM and $I_x(a) = \exp(L)$.

Algorithm for the Natural Logarithm of the Gamma Function

The best algorithm for evaluating ℓn Γ(x) is probably the Euler-MacLauren expansion provided by Abramowitz and Stegun (1964).

$$\ell n\, \Gamma(x) = (x-1/2)\ell n\, x - x + (1/2)\ell n(2\pi) + \sum_{i=1}^{4} T_i \qquad (44)$$

with $T_1 = 1/12x$, $T_2 = -1/360x^3$, $T_3 = 1/1260x^5$, $T_4 = -1/1680x^7$.

This Euler-MacLauren expansion, truncated at four terms, provides an extremely accurate method for evaluating ℓn Γ(x) with the error being less than 10^{-14} if $x \geq 18$ and enough digits are carried in the calculations. When x < 18, it is customary to increment the argument of the gamma function and use

$$\ell n\, \Gamma(x) = \ell n\, \Gamma(x+k) - \sum_{i=0}^{k-1} \ell n(x+1) \qquad (45)$$

where k is the integer which, when added to x, produces a value between 18 and 19. A program is easily written using (44) and (45).

CONCLUDING REMARK

The algorithms presented in this paper provide a useful array of methods for evaluating the classical distribution functions. The accuracy comments included were based on comments from other authors and, to a large degree, on personal experience. In the latter case, they should be understood to apply to the use of an IBM 360/370 computer system using double precision calculations.

References

[1] Abramowitz, M. amd Stegun, I.A. (1964) Handbook of Mathematical Functions. National Bureau of Standards, Applied Math. Series. Washington, D.C.: U.S. Govt. Printing Office.

[2] Aroian, L.A. (1941) Continued Fractions for the Incomplete Beta Function. Ann.Math.Statist., 12, 212-223; Correction (1959) 30, 1265.

[3] Bargmann, R.E. and Ghosh, S.P. (1963) Statistical Distribution Programs for a Computer Language. IBM Watson Research Center Report RC-1094.

[4] Bargmann, R.E. (1969) Principles and Design of Statistical Computer Languages, Chapter 8 in Progress in Operations Research, Vol.3, New York: John Wiley and Sons, Inc.

[5] Bhattacharjee, G.P. (1970) Algorithm AS 32: The Incomplete Gamma Integral. Applied Statistics, 19, 285-287.

[6] Boardman, T.J., and Kopitzke, R.W. (1975) Probability and Table Values for Statistical Distributions. Proc.Statist.Comput.Section. Washington, D.C.: American Statistical Association.

[7] Bouver, H. and Bargmann, R. (1975) Computer Algorithm and Modules for the Evaluation of Statistical Distribution Functions. Tech. Report 110, Dept. of Statistics, University of Georgia, Athens.

[8] Greenwood, J.A. and Hartley, H.O. (1962) Guide to Tables in Mathematical Statistics. Princeton, N.J.: Princeton University Press.

[9] Johnson, N. and Kotz, S. (1970) Continuous Univariate Distributions, Vol.2, Boston: Houghton Mifflin Co.

[10] Khamis, S.H. (1965) Tables of the Incomplete Gamma Function Ratio. Darmstadt, Germany: Justus von Liebig verlag.

[11] Kennedy, W.J. and Gentle, J.E. (1980) Statistical Computing. New York: Marcel Dekker, Inc.

[12] Khovanskii, A.N. (1963) The Application of Continued Fractions and their Generalizations in Approximation Theory. Gronigen, Netherlands: P. Noordhoff.

[13] Muller, J.H. (1931) On the Application of Continued Fractions to the Evaluation of Certain Integrals, with special Reference to the Incomplete Beta Function. Biometrika, 22, 284-297.

[14] Woods, J.D. and Posten, H.O. (1977) The Use of Fourier Series in the Evaluation of Probability Distribution Functions. Commun.Statist. B6, (3), 201-219.

USE OF HANDHELD PROGRAMMABLE CALCULATORS IN STATISTICS COURSES FOR NONTECHNICAL STUDENTS

D R Barr, Office of Naval Research, London Branch
(on leave from the Naval Postgraduate School, Monterey, California.

INTRODUCTION

The usefulness of stochastic models and statistical inferences in a wide range of disciplines has gradually led to the inclusion of probability and statistics courses in many academic curricula. Unfortunately, the required course in statistics is often the most disliked and feared course in a student's curriculum. This is especially the case for "nontechnical" students who generally have limited mathematical capabilities. Instructors of statistics courses for such students often feel they must devote a significant proportion of their instruction to improving the students' algebraic and computational skills. Clearly, allocating course time to algebra and arithmetic cuts into the time that can be spent on statistics.

It is reasonable, therefore, to examine the goals of the statistics courses for such nearly mathematically illiterate students. For many curricula, such examination reveals that the students need to gain an appreciation of statistical reasoning and applications, but it is not absolutely necessary that they be able to carry out the details of the associated computations. In their professional careers, these people will generally need to be informed consumers of statistics; for the most part they will never be skilled producers of statistics. But in order to learn about statistics, isn't it necessary for the student to "do" statistics? And if so, how can such a statistics course avoid the drudgery (from the student's point of view) of the algebra and calculation?

One possibility for teaching an applied statistics course without spending time on mathematical prerequisites is to use a computational device which allows the use of "canned" programs for carrying out the calculations. The particular device used for this purpose can take many forms, ranging from a programmable calculator to a microcomputer or a mini or mainframe computer. For planners of courses relying on such devices, there is an unfortunate tendency to get bogged down in considerations of the specific device to be used. The general approaches that can be used in teaching these courses is rather invariant with respect to the particular computational device used, as long as it is possible to use stored programs of sufficient sophistication. I shall return to this point at the conclusion of this paper.

In what follows, I shall report on a program we have been using for Management Science students at the Naval Postgraduate School in Monterey, California.

In this program, Texas Instruments T159 calculators are issued to all students entering the Management Science (MS) curriculum, and these personal calculators are used by the students throughout their six quarter program leading to a Master's degree in Management Science. These students are military officers who received Bachelor's degrees in a wide variety of subject areas several years before entering the Master's degree program at the Naval Postgraduate School. These students are given a calculus "refresher" during their first quarter (a few of the students apparently haven't had a calculus course before). This is a calculator based course, using pre-recorded programs for computation of limits, derivatives, integrals and so on. An introduction to use of the T159 calculator occupies the first week or so of this course. The two quarter sequence in probability and statistics follows the calculus refresher course; the role of the programmable calculator in this sequence is described below.

THE COURSE

The two quarter course in probability and statistics for the Management Science curriculum is in most respects a "standard" course for such students. The textbook currently used is Wannacott and Wannacott, "Introduction to Statistics" (Wiley). A written supplement was prepared which served as an interface between the text and the calculator. This supplement contained worked examples of problems as well as listings of the programs distributed to the students, and instructions for using the programs. The programs themselves were distributed in the form of pre-recorded magnetic strips. As each topic was encountered in the text, the corresponding supplement material and appropriate programs were discussed in the class and use of the programs was demonstrated by the instructor. Emphasis was placed on the program input requirements (statistical assumptions, data) and interpretation of the program outputs.

This approach allowed adherence to the standard course outline, while emphasizing the idea of making the students into informed consumers of statistics. Several rather unique consequences were observed in this use of computation devices as an integral course resource:

(i) Students are nearly always able to calculate the correct numerical answer to standard problems. When incorrect numerical answers did occur, they were due to erroneous data entry or, more seriously, inappropriate choice of program by the student. This tends to make a student's performance on a given problem type a Bernoulli variable, without much chance for assignment of partial credit. This led us to conduct the grading for the course along the lines of the "mastery concept". Indeed, a list of specific objectives were given to the students for each chapter of the course. In this way, the students knew exactly what they were expected to be able to do. Any student who "mastered" the approximate

200 objectives would receive an A in the course.

(ii) Students usually found calculations of solutions to problems in the text to be quite simple, and they often requested supplementary exercises. This placed something of a burden on the instructors.

(iii) There were occasional mechanical failures of the devices. I recall once, during an examination, there was a very brief power failure that caused the calculator memories to be cleared. There were occasional problems with individual calculators in getting a program card to read, having insufficient battery charge, and so on. Such failures are disturbing to the student because they can affect his (her) grade and yet are not the student's fault. I learned to bring one or two spare calculators into exams, which I could lend to students in case of such failures.

(iv) Occasionally, students just didn't like using the calculator. Such students attempted to solve the homework and examination problems by traditional hand calculation. While this is in some ways a good thing, it takes time; students insisting on hand calculation tended to receive lower course grades because they didn't have time to finish exams.

(v) For any course which includes discussion of programs and demonstration of their use in solving statistical problems, it is necessary that the students all have the same model of calculator. As mentioned above, we were able to arrange to have T159's distributed to all students in our course. This was done with a pool of calculators owned and maintained by the school. Students checked calculators out of this pool. An alternative would be to require that students purchased a specified calculator as a precondition for taking the course. This idea is becoming more feasible as the cost of sophisticated calculators falls. I recall that I was required to purchase a rather fancy slide rule when I first entered an engineering curriculum in the late 1950's. That slide rule cost a little over $30, which was roughly the cost of a modest bicycle at the time. The T159, for example, is available at about $170, which is roughly the cost of a modest bicycle today.

In addition to providing easy calculation of numerical results for exercises in the "traditional" text, the use of calculators allowed a number of non-traditional statistical computations to be performed by the students. Examples are: (a) simulations (both as a way of examining such ideas as confidence intervals, and as an analysis tool in itself), (b) robustness (for example, seeing how the posterior probabilities are affected by variations in the prior probabilities in an application of Bayes' formula), (c) significance testing (wherein the level of significance of the observed test statistic is calculated), and (d) applications of transformations (although transformations such as coding

to simplify calculations aren't necessary with the calculator).

Programs that we prepared for use by the students included:

CDF and inverse CDF for normal, exponential, uniform/random samples for normal, exponential, uniform/random calculation of various sample attributes such as mean, median, range, mean absolute deviation for grouped and ungrouped samples,

significance tests and CI's for μ and σ^2 in normal samples;

significance tests and CI's for the exponential parameter;

significance tests about Bernoulli parameters;

various two sample tests of significance and CI's;

one and two-way analysis of variance (balanced design);

Scheffé contrasts

regression for one and two variables, including CI's for coefficient parameters, CI for a regression value, Prediction intervals, tests of significance for parameters;

Bayes' formula;

simulation of samples for certain discrete distributions;

expected values for functions of discrete random variables.

RESULTS

Generally, the students were enthusiastic about this approach. The calculator based statistics course became a popular course in the Management Science curriculum, in sharp contrast with earlier non-calculator approaches. A few students from each class actually began to choose Master's thesis projects based on the use of statistical analyses. Instructors of some follow-on courses (such as introduction to operations research and economics) were forced to revise their traditional approaches because the students had by that time developed considerable skill in preparing and using programs for their calculators. After an initial experimentation period, the Management Science Faculty endorsed this approach, and considerable efforts are being made to adapt virtually the entire Management Science curriculum to this approach.

There continues to be some resistance to these changes by some faculty members, however. A common criticism which I believe has some merit is that this approach makes the student dependent on a "mysterious black box". "What would one of these students be able to do if he or she were separated from the calculator?", I have been asked. An attempt to answer this brings us back to the idea of trying to decide what the students should be able to do as a result of completing their statistics course. I believe the calculator approach leaves more time for discussing the applications of statistics than did the traditional approach. I think that as a result they are better able to understand and utilize the

results of statistical studies performed by statisticians in their organizations. And I can counter with the challenge, "What were they able to do after completing the traditional course?".

A second criticism I sometimes hear is, "Why don't you just run the analyses off on the computer and hand copies of the computer print-outs to the students?". We have experimented with such an approach, but without much success. First, it seems difficult to sustain motivation of non-technical students for such an approach. Second, there is merit, I believe, in having students handle data. This seems to give them a more accurate feeling that data contain information, that these are the basis of their inferences (rather than preconception or opinion), and that obtaining and handling good, clean data is not a trivial task. I have found the students seem to almost enjoy working with the data. Thus, it appears to me that looking at data sets and entering data into programs is, at least for small data sets, an asset in this approach. Some other potential criticisms to this approach were cited above. These include cost, mechanical reliability, and grading procedure issues.

CONCLUSIONS

The availability of low priced computing resources opens the possibility of restructuring traditional statistics courses. At the "low end", service statistics courses for non-technical students can take advantage of the capabilities of hand held programmable calculators. Such calculators have advantages (over use of computer terminals, for example) in terms of cost and portability. We are now at a point where students might be expected to purchase their own calculators for use in the classroom and at home. Using pre-recorded programs, a statistics course requiring virtually no algebra or computational skills can be designed for non-technical students. There seems to be some merit in having such students "handle" data, however. In such a course, emphasis is on the foundations and assumptions for statistical inferences, and in interpretation of the results of statistical analyses.

SIMULATION OF GAMES STUDIED IN THE SEVENTEENTH AND EIGHTEENTH CENTURIES

J Badrikian, IUT de Clermont-Ferrand

In the 17th and 18th centuries, most mathematicians took random situations from their social environment as subjects for their studies : particularly, games of chance which were played at that time.

We can see how probabilistic thinking developed from these props, by intuition. What Pascal called "biais". A reasoning which was, moreover, presented as "an Art".

To begin with, we suggest a statistical approach to these problems.

One can really understand a game if one often plays or watches others playing. The rules are then assimilated, the prognostics are given and tactics proposed. Practice will lead to familiarization with a random situation.

This is where the computer comes in : the games are simulated and the pupil can follow their development on the screen. The choice of representation (graphic, pictorial) is important. One must, at the same time, give all information on the development of the game in an energetic way (just giving simple score results would be very boring) with precision and clarity.

With the "Jeu de RECONTRE", n cards, numbered 1 to n are drawn successively from a bag. The player who withdraws a card bearing the same number as the order of the draw, at least once in the game, wins.

Representation on screen for 9 cards numbered 1 to 9

1	2	3	4	5	6	7	8	9
7	1	6	4	9	2	3	8	5

Explanation

← Order of draws

The 9 numbers are written down as they are drawn.

← "Invariants" 4 and 8 flicker: number (respectively 8) has been drawn on the fourth (respectively 8th) draw.

Simulation of a game is obviously quicker than the real game itself. The computer will therefore allow one to save time. As many games as one likes can also be played in a short time.

Secondly, the computer's function is to provide statistics. All results are memorized, the frequency calculations are done automatically and given at the end of the operation. The operator can do as many runs as he wishes. He can also develop sequences to any length.

Statistical results reply to questions already resolved elsewhere by mathematical reasoning and this underlines the bond between statistics and probabilities.

Here, the computer is only an aid to the pupil (he saves time, and he is discharged from doing frequency calculations) and it also helps him to maintain active interest in several ways:

- over and above the choice of the number of runs he can:
  - change the parametric values, for example, if two partners play with one probability of winning (0.5 for each player) he can modify this "chance" to his liking and examine the consequences on the different stages of the game;
  - moreover, having statisical replies, the pupil may wish to know how to resolve the problem by deduction.

Now would be the time to do a study, using probability calculus. And one may then need to use the computer again for the actual calculation of the established formulas.

It was Jacques Bernoulli who drew up the tables giving the winning probability for each stage of the game called "JEU DE PAUME". It took him a certain time to do this. A simple program can give them in a few seconds.

Today, I would like to present four simulations, realized on the screen of an APPLE II micro-computer which is connected to a printed output - which will eventually provide numerical results (frequencies, tables).

1. Simulation of the "Jeu de Paume"

This game was the forerunner of tennis. The "problem" is the same - although more complicated - as Pascal's game which the French call "Probleme des Partis" and the English call "Problem of Points".

The simulation is made for one chosen value of the parameter:

- number of games
- number of sets
- possibility of winning for each player.

Initial state (scores already reached) and the number of repetitions done from this state are left to choice.

The user can then modify the possibility of winning each game. Extra time is envisaged for three or more players.

On the other hand, after a certain number of repetitions we give the observer no possibility of following the game but give him just the final result and the frequencies.

2. Simulation of the "Jeu de Rencontre"

This game was studied by MONTMORT and was given the name of GAME OF THIRTEEN, because it consisted of 13 cards, numbered 1 to 13. It was taken up and discussed by DE MOIVRE, EULER, LAPLACE...

One could be interested in one or several "invariants".

3. Game of "The Three Urns" (Daniel BERNOULLI)

Three urns, each one possessing n balls of a given colour are used. We remove one ball from each urn. The ball from urn 1 is put into urn 2, the ball from urn 2 into urn 3 and the ball from urn 3 into urn 1.

From the development of the contents of the urns (which we see carried out on the screen) one can ask oneself numerous questions:

- contents of urns,
- time lapse for return to initial state,
- state of balance, ...

4. Game of Lottery

A lottery consists of n numbers. A draw consists of taking r numbers. After x draws we estimate the chance of only m numbers already given. (This is a question we often ask in France concerning our "JEU DE LOTO").

This problem was also studied by numerous mathematicians (EULER for example) and generalized the problem of throwing the dice by MOIVRE.

On the other hand, it illustrates the situation of drawing without replacement (the r numbers drawn) and draws with replacement (the x repetitions).

TOPIC 8: TEACHING STATISTICS WITH THE HELP OF CASE STUDIES

ORGANISER: PROFESSOR B EPSTEIN

CASE STUDIES FOR AN ADEQUATE UNDERSTANDING AND INTERPRETATION OF RESULTS BY STATISTICAL INFERENCE

M Borovcnik, University of Klagenfurt

NEED FOR NEW TEACHING STRATEGIES IN STATISTICS - PRELIMINARY CONCEPTIONS OF OUR UNDERSTANDING OF STATISTICS

In teaching probability theory some efforts have been made to support the understanding of the term probability as such. The foundation is laid by Kolmogorov's axioms (or some reduced system of axioms) but the intuitive way of teaching is by means of (Laplacean equiprobability and) relative frequencies. One very important aspect of relative frequencies is the empirical law of large numbers. For visualizing this empirical fact the idea of simulation has been (and still is) of great importance in teaching probability.(*)

The character of probability changes, if the field of application is widened to judgement of probability models itself. This is why one cannot rely on what has been developed within (a more or less mathematically oriented) probability theory. This bears the danger of an unfeasible transfer of a somewhat objective character of probability to statistics.

From the point of view of application it is statistical methods that count. Statistical ideas have penetrated modern sciences and ways of thinking. Fallacies, errors, and manipulation are widespread and characterize a helplessness in connection with statistical methods and results.

I want to characterize two ideal positions to tackle this problem in education at all levels:

(I) Probabilistic considerations dominate teaching stochastics, the mathematical level aimed at is too high, the connection to applications is abandoned, statistical problems are of minor importance.

This position has its origin in the following attitude: It is often thought desirable to work out the foundations, and to reason with logical preciseness, to trace back anything to the foundations. For stochastics this means that giving a complete representation of the theory of probability is considered the best way of teaching it. But theoretical connections are endless.

The mathematics of probability theory has many disadvantages:

- it precludes many people from occupying themselves with stochastics
- it obstructs the possibility of advancing to the relevant part of it, i.e. to statistics

(*) Cf. e.g. Heitele who apostrophizes it as a fundamental idea of probability

- it obscures ideas and concepts behind statistical reasoning
- it neglects primary stochastic intuitions, it fails to weaken false conceptions, it is not suitable to build the basis for stochastic schemes

(II) The other extreme position is that of teaching cookery recipes. Probability is too difficult, therefore it is reduced to a minimum, the remainder is an unrelated chain of methods trained by means of sterile examples - unconnected senseless mathematics.

Both of those extreme positions mislead us to apply statistical methods routinely (or else: to exclude this possibility of tackling problem situations) without adequately grasping the problem situation, without interpreting the results critically.

My thesis is the following:
Both positions are not apt to establish a picture of the real interrelations, they tend to give a biased objective conception of the possibilities of statistics.

There is a trend to dodge difficult mathematics, the old idea of simulation is extended to statistical questions in order to give statistics more weight and a better chance for being understood (Cf. e.g. Schrage). Decision rules are evaluated empirically, the whole test (estimation) problem is simulated under the conditions of a specified model and theoretical terms like risk of an error of type I or II can be witnessed in their long-run consequences. This performance should enable us to lay the basis for an operative understanding of central ideas and terms.

One of the central ideas is the following: comparing, evaluating, and grading a single value of a (test) variable within all possible values on the basis of an hypothesis.

This can be realized by theoretical reasoning (deriving the distribution of the test statistic) or by creating the basis for the comparison by simulation of the test situation (and empirically deriving this distribution).(*)

The concept of simulation may also be used to advantage in estimation problems: Visualizing the long-run consequences of models and decisions (i.e. estimations) can give a better insight into notions and terms like unbiasedness, variance of estimation functions or confidence coefficient of interval estimates.

However, there is some danger of performing an unfeasible transfer of the probabilistic side to statistics:

(*) An idea going back to W.S. Gosset (in deriving the t-distribution he filled the gap with simulation studies) which was rediscovered for didactics only recently.

(1) One is tempted to consider the simulated data as real data, to interpret empirical meaningfulness into them and the terms which are visualized in their manifestations, to overlook that such simulated data are hypothetical data as they are gained on a basis of hypotheses. This is even complicated by the relation between the theoretical law of large numbers and the empirical one which has quite a different quality.

(2) There is the danger that one cannot clarify the conditions under which this simulation is done. In the probabilistic standard situation one is confronted with iid (independent identically distributed) realizations of a variable - transferred to the statistical standard situation, this would mean that one has iid-test-situations. There seems to be a very different quality of adequateness of iid-assumptions in some urn (lottery) situation and in the statistical test situation where one can hardly think of the realized situation as one of an iid-sequence of test situations (this fact should be worked out by case studies). Simulation of test situations alone does not meet the problem of teaching statistics in an intellectually fair way; it seems to oversimplify the problem.

(3) Difficulties in learning statistics are reduced to technical ones. It is agreed that there are special ways of reasoning in statistics: one has to grasp the various difficult technical terms as e.g. of the test theory. There seem to be no problems beyond that. This is by no means the case. A technical view on statistics increases the objective character of it and its results inadequately.

I agree that an operative understanding of central statistical terms and concepts is valuable - just as urn situations (and the simulation of them) are of value to visualize the probabilistic mechanism. However the field of real application is limited. The built-up understanding of the terms is biased thus impeding full understanding of the meaning, possibilities, and limitations of statistical reasoning.

Computers will soon aid us in calculating the necessary results. What we have to support by teaching statistics is to confirm our competence in facing seemingly objective results by means of statistics, to demonstrate the limitations and subjectivity of any statistical argument. It does not merely suffice to teach statistics, the conception of it which we facilitate is relevant.

- What are possible interpretations of statistically aided results? Difficulties with an adequate understanding are to be faced.
- What is the role of probability theory in understanding statistics?

Besides the details of the way of teaching statistics one has to avoid a biased understanding of methods and terms. We do believe that knowledge about

statistical methods - meta-knowledge - will facilitate a better understanding. Furthermore, we think that case studies will be suitable to teach statistics as we intend to. We shall discuss some "meta"-case-studies to support our opinion and simultaneously give a picture of some aspects of our understanding of statistics.

Statistical problem situations are open problem situations. There is a much better chance to make this clear by case studies than by "conventional" teaching strategies:

One has to face problems like:

- judging information as to whether it is relevant to the problem at hand, whether it has its origin in qualified knowledge of the subject or is only something like a prejudice,
- deciding whether it is desirable to give an "overall"-treatment of the problem (e.g. the whole distribution in question) with no specified questions, combined with a low level of so-called background knowledge which may be false, in order to ask for details (concerning a parameter),
- deciding to include or to exclude various hypotheses, as it is often done only implicitly, by choosing a test statistic (e.g. χ^2 test or Kolmogorov-Smirnow-test),
- deciding which hypothesis is to be formulated as null hypothesis,
- judging random variability that is possible at a specified (given) sample size: significance in the technical-statistical sense does not mean the same as being of practical importance, failure to prove significance may be due to a small size sample only,
- arbitrariness in choice of technical procedures like different approximation rules, different groupings (χ^2-test) etc.,
- knowing and judging the peculiarities, possibilities, and limitations of special methods.

Knowledge of the problem and the field of application, information and competence to evaluate this information is most decisive in structuring a problem situation. By case studies we want to support our understanding of statisics:

As far as subjective components play a part (consciously or not) in the process of structuring a problem situation, in the choice of a test statistic, the results are of the same subjective character in spite of the mathematical methods used and the numerical exactness of the result.

The necessary competence is to be seen in two different modes:

(I) Knowledge in the field of application and the problem at hand

(II) Knowledge in statistics, in the peculiarities and limitations of the various statistical methods proper.

It is far more a convention that statistical methods are regarded as a useful tool in handling problems in various fields of application and in empirically oriented sciences. We have tried to outline why we think "conventional" teaching statistics gives a biased picture of this "art of reasoning", we have given some aspects of statistics that we consider especially important, now we are to proceed to discuss some "meta"-case-studies.

DISCUSSION OF THE CASE STUDIES

Case study 1:(*)
Number of boys in completed(**) families with four children.

Boys	0	1	2	3	4
Nr.of fam.	1	17	45	31	6

It should be investigated whether the data are consistent with the binomial model or if there are significant deviations from it (at the 5% - level).

We shall demonstrate various possibilities to tackle the problem:

(a) Information from population statistics is used to equate p = 0.51 - χ^2-test gives a prob value of 8%.

(b) χ^2-test with $\hat{p}$ estimated from the sample gives a prob value of 14%.

Neither χ^2-test indicates significant discrepancies between data and binomial hypothesis.

(c) 95%-confidence interval for the parameter p yields (0.508; 0.612). This result may be interpreted as follows:
- sample size too small, no judgement of a detail (of the distribution) seems possible
- information used in case (a) does not seem to be adequate, the percentage of boys may not equal that for the whole population.

(d) Variance test: $\frac{\text{empirical variance}}{\text{var.of spec.binom.model}}$ indicates a significant deviation - prob value 3%.

(*) No details of calculations will be given, only those aspects of results will be quoted which are needed to understand the line of discussion.

(**) i.e. families where no further children are to be expected.

(e) Confidence interval for standard deviation (equivalent to variance test): the theoretical standard deviation is in the upper rejection region of the connected test - the empirical variability being too small for binomial model.

In short, the results are as follows:

- χ^2-test fails to reject binomial hypothesis,
- confidence interval for the percentage of boys indicates: sample size being too small to judge a detail of distribution (local question),
- confidence interval for the standard deviation yields rejection of binomial hypothesis: empirical variability being significantly too small for binomial distribution.

This is an artificial and very academic case study, yet serving our purposes: to show that statistical problems are by no means straightforward. The different results do not cause logical difficulties but are somewhat disturbing if one wants to confer statistical results to reality. In the various ways of solution, different problems have been tackled.

χ^2-test tackles a global question: Is the number of boys binomially distributed or not? Any alternative is included in the consideration, the χ^2-test-statistic reacting to any deviation from the (binomial) model. Background knowledge is used only in the sense that one has a random sample. Other hypotheses are not referred to in case (b). In case (a) additional information on the percentage of boys is used but this does not seem to be justified in the light of case (c). -Using information one can make mistakes.

Confidence intervals for central tendency or for variability deal with local questions: a detail of the distribution in question is discussed. Treating local questions is not generally better than treating global ones: What is a possible interpretation of "no significant results" in cases (c) and (e)? Surely not that there are no significant deviations from binomial model, but only that there are no significant deviations concerning central tendency or variability. Only special alternatives are included in the consideration, for the test statistic is sensitive only in deviations concerning these details.

Using information to formulate local questions may be misleading, for using this information may be unjustified or simply incorrect. In test situations there is always the problem of judging information as to whether it is relevant to put concrete questions in order to get more information out of data - namely concrete answers:

In complete families of four children the variability of the number of boys is reduced in comparison to the binomial model, for those families with four children having only boys or only girls tend to drop out of the population

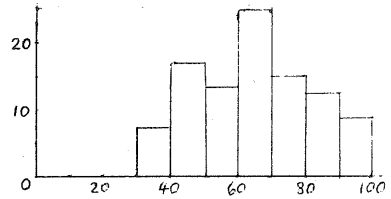
by getting another child.

One cannot escape the problem of local and global questions by arguing that χ^2-test is simply inferior in detecting deviations from a model because it uses only a small part of information contained in the sample. Another example could be given to show that this is not true. It depends on the view of the problem what alternatives are to be considered or to be excluded. There are questions in which any alternative has to be taken into consideration.

Case study 2:

Customers in a newly opened shopping city are interviewed at random. Among other questions they are asked to judge their contentment with the facilities of the new shopping city on a scale from 0 to 100.

The result of a sample of 80 persons is depicted in the histogram.



For further considerations it might be of value to replace this frequency distribution by a theoretical distribution. Is a normal distribution appropriate to describe the distribution?

(a) Test for normality: χ^2 yields a prob value of 34% - normality seems to be quite appropriate.

(b) Test for uniformity: χ^2 yields a prob value of 8% - one cannot exclude uniformity (at the 5% - level).

(c) Let us consider increasing the sample size to 94, the percentages being unaffected.
Normality: p. v. 3%; uniformity: p. v. 4%.

By varying the hypothesis being tested or by varying the sample size one can see: the choice of the null hypothesis is very critical - one must judge the sample size for relevance. The results are very sensitive to these "parameters".

Case study 3:

A firm is interested in improving prevention of accidents. For that reason one wants to investigate (among other things) if there is a significant deviation from a uniform distribution of risk over the days of the week.

Mon	Tue	Wed	Thu	Fri
26	17	22	24	31

(a) χ^2-test for uniformity yields a prob value of 50%. The data cannot at all be regarded as unusual - no rejection of uniformity hypothesis.

(b) Separate 95% -confidence intervals for the percentage of accidents occurring on each day of the week have a width of approximately 15 percentage points, e.g. for Monday (0.14: 0.29). The separately calculated confidence intervals are very wide
- sample size seems to allow no conclusion other than its being too small (simultaneous confidence interval would be even wider).

(c) Supposing there is past information that accidents on Mondays and Fridays are more frequent. Test of $p = 0.4$ against the alternative $p > 0.4$ would just reject the uniformity hypothesis relative to Mondays and Fridays (at the 5% -level).

In cases (a) and (b) the method used was varied - again a choice being open between local and global aspects of the problem. If it were of practical importance then one would decide to get more data.

In case (c) another aspect is considered: The problem of whether one can use information available in order to put more precise questions to get more information out of data. What if this information is just a cursory impression, or even just a prejudice? What if it is extracted from the same data by means of which it should be tested? Judging this information for relevance to one's problem is clearly subjective. The same subjective feature lies in the variability in case study 2 and in the choice of the method in cases (a) and (b).

STRUCTURING THE PROBLEM SITUATIONS - CONSEQUENCES FOR THE CONCEPTION OF STATISTICS

(i) Choice of test statistic

By varying the choice of methods, choice of null hypothesis and information about the problem at hand we have seen that in any statistical (test) situation there is a large spectrum of possibilities. By choosing a test statistic we implicitly or explicitly discern two classes of deviations:

- important deviations, the test statistic being sensitive to them,
- deviations of no importance, the test statistic being insensitive to corresponding alternatives.

(ii) Local and global questions

There are very different aspects connected with these opposite ways of questioning. Does it really help if χ^2-test leads to rejection of the binomial model (case study 1)? What if one cannot reject the binomial hypothesis by the variance test?

One has to balance advantages and disadvantages of all the possible results and their interpretation under rival ways of tackling the problem.

(iii) Information

Justified use of information may lead to concrete (local) questions and in consequence to concrete answers which do yield much more information about the problem at hand (e.g. direction of deviations). On the other hand use of information means exploiting background knowledge which may be false.

One has to balance advantages and disadvantages of using information: Is it better to mark out special alternatives to get more information out of the data thus risking making an error or is it better to avoid this risk and include any alternative?

(iv) Relevance of sample size

Statistical significance is not equivalent to practical importance. All results are relative to sample size, so there is a need for judging the magnitude of random deviations to be tolerated at a specified sample size by direct methods.

- In case of no rejection one has to ensure that this is not due to a small sample size, a sample size at which random variation is so large that even considerable discrepancies between data and model have only a small chance of being detected.
- In case of rejection one has to evaluate magnitude and relevance of the deviations by other (direct) methods than just by saying that e.g. it exceeds the 95% -fractile of the test distribution.

Such judgements and the coherent acts of choice - including, excluding alternatives in the consideration; balancing advantages and disadvantages of local and global questions; evaluating information; judging sample size for relevance etc. - should not be made arbitrarily or simply implicitly without deliberating over their consequences.

Expert knowledge in the field of application, the problem situation at hand as well as the peculiarities of statistical methods in question should form a basis for structuring the problem finally leading to the model situation of mathematical statistics or, if simplified, the corresponding simulation situation.

Of course this structuring process is markedly subjective, many problems being coherent in it: Isn't even using information in order to structure one's problem manipulation? It depends on one's point of view how information is evaluated.

One consequence is that results gained by statistical methods are inseparably connected to the whole context of the process of problem structuring and

all its subjective components. Without communicating these context conditions statistical results are nothing but bare bones: void of any useful interpretation, or: capable of any wanted interpretation.

MEANING OF STATISTICAL RESULTS FOR REALITY

By varying the conditions the case studies make clear that statistical methods are very doubtful if their results are to be interpreted adequately. Regarding the complexity of problems in connection with these judgements something like a return to a "weakened form of determinism" could be desired: To have only one test statistic to check discrepancies between data and model would give the results at least a normative character. The whole process of structuring with all its subjective components would be eliminated. The case studies, however, show that there are interesting questions which could not be answered in this way. Thus, let us forget this idea.

What do statistical results really mean in the light of the discussion of the case studies?

What does it mean that the test statistic XX has realized a value of c, the significance of the result being beyond the 5% -level? Other test statistics would have yielded a different significance level.

It is a usual convention that risk of a type I error is regarded as property of the test procedure in question and not of the tested hypothesis. The testprocedure will reject a correct null hypothesis only in 100α% of performances of the test - α being also called the size of the procedure.

Besides the general difficulties in transferring a meaning of this long-run property to a singular test situation there are many aspects to be considered:

- One includes only error of type I, errors of type II depend on the alternatives determined by the structuring process.
- How can risks of a type I error be compared with each other in different test situations? One does not always test correct hypotheses, the tested hypotheses not always being of the same importance to reality. For example, simulating test situations to give an impression of the corresponding notions misleads someone to interpret empirical meaningfulness into such terms.
- A frequency interpretation of the size of a test is restricted to "clinical" iid-conditions thus being a model conception - the case studies with their structuring process indicate that this could only be an hypothetical (urn) conception for the problem at hand.

- The process of problem structuring makes use of information and other subjective components - the resulting discrepancy between the outgoing problem situation and the model situation being unique in any case. The distinct contents of background knowledge used in the structuring process makes statisical results uncomparable because they lose their sense without the whole context.
- Comparing and judging a realized value of a test statistic with a theoretical distribution is clearly a model comparison. Simulating this process cannot wipe out these theoretical components. Besides it may lead to false impressions because the iid-conditions are clearly satisfied in the urn (simulation) situation but by no means in the real problem situation.

All this means that the singular test situation is compared to a sequence of independent identical test situations. There is one singular realization of the test situation - with all the subjective components of the structuring process - the real problem situation being by no means identical with the model situation in question. Nevertheless the model situation is thought to be identically repeatable and the result of this hypothetical process is used to judge the initial real problem.

The discussion of the case studies stresses that statistical results can only be interpreted hypothetically. The concept of iid-realizations of a random experiment - e.g. playing roulette - is uncritically transferred to statistical reasoning. One loses sight of the fact that it is the process of problem structuring that leads to the "clinical" standard situation. One "creates reality" without reflecting this.

My thesis is the following:
The possibility of interpretation in the clinical iid-situation of probability theory is extended to statistical reasoning by an unfeasible transfer leading to a false objective conception of statistics itself.

The notion of probability seems to have changed its character from objective to subjective via the extended field of application from probability to statistical reasoning. It is desirable to investigate whether it would be of advantage to understand the notion of probability right from the beginning the way the subjectivists do. Statistical methods do not gain objectivity by the fact that many people acknowledge their merits and adhere to the usual mode of interpretation - statistical results being valuable only in connection with a balanced view of the problem situations which ought to be determined by expert knowledge in the field of application and statistical methods - they are not to be separated from the whole problem context and its structuring.

One always has to point out the discrepancy between the initial problem situation and the model situation. Statistical methods cannot replace knowledge of the problem at hand - statistics per se cannot give insight - it can only help finding one's way through.

CONCLUDING REMARKS

The discussion of the case studies and the continuing discussion of the problems of adequate interpretation of statistical results should provide

- a picture of our understanding of this "art of reasoning",
- sufficient support to show that case studies(*) offer an excellent opportunity to facilitate a better insight into the problems of mathematizing real situations and interpreting the results adequately.

In this light statistical problems appear to be open problems.

Several - subjective - choices and judgements have to be made to establish the model situation of mathematical statistics:

(i) Choice of the test statistic
(ii) Choice between local and global questions
(iii) Judgement of information for relevance
(iv) Judgement of relevance of sample size

Case studies should give a chance to see that statistical results can only be interpreted hypothetically - their relative simplicity not only being artificial but also being a basis to facilitate an understanding of this feature of statistics. One cannot interpret statistical results adequately separated from the whole problem context including its structuring. The shortcut "the h_o will not be rejected at the 5% -level", where you can take any h_o and where the problems are exchangeable, is misleading and very academic.

Another step is understanding statistical concepts and the meaning of statistical results can be achieved by doing statistics in an exemplary way, i.e. by means of project studies. We have not tried to tackle the problem of project studies simply because we lack experience. Nevertheless we think project studies should permit the insight that personal purposes and intentions do affect this process of problem structuring and thereby final results if this is reflected in a specific phase. Maybe the problems of real application become too powerful so that we think case studies could facilitate this conception of statistics more easily.

(*) For further case studies especially dealing with limitations and possibilities of statistical methods see [1] - where, for example, a parallel between X^2-test for independence and the Pearson correlation coefficient is outlined.

On the whole, case studies (and exemplary project studies) can permit the perception that not only knowledge of statistics (in the mathematical-technical sense) but especially expert knowledge in the field of application, of the problem at hand, and some "meta"-knowledge about possibilities and limitations of statistical methods in question will and should play a decisive role to compete in situations with uncertainty. The intuitive way to teach statistics should be by means of case studies. Case studies should have the same place in teaching statistics as simulation (*) has in elementary probability.

(*) In connection with demonstrating the consequences of certain probability models - not in the sense of Monte-Carlo studies.

References

[1] Borovcnik, M. (1981) Eine Analyse des χ^2-Unabhängigkeitstests - Möglichkeiten und Grenzen, Schriftenreihe Didaktik der Mathematik, 3, p.35-48.

[2] Heitele, D. (1976) Didaktische Ansätze zum Stochastikunterricht in Grundschule und Förderstufe, phil. diss., Dortmund, 1976.

[3] Schrage, G. (1981) Entscheiden und Begründen - Leitlinien für den Statistikunterricht, Schriftenreihe Didaktik der Mathematik, 3, p.86-101.

IS A SCIENTIFIC LEGEND A COMPREHENSIVE CASE STUDY?

V V Shvyrkov and A C Davis, Rider College, Lawrenceville

We were standing in the back room of historic Bainbridge House in Princeton, New Jersey contemplating this saying of the late Albert Einstein:

"There was this huge world out there, independent of us human beings and standing before us like a great, eternal riddle, at least partly accessible to our inspection and thought. The contemplation of that world beckoned like a liberation".

A child next to us asked his father, "What's Einstein? Sounds like a mosnter". Father: "Some people think so. He was an inventor". Child: "What did he invent, some kind of telephone?".

We are reminded by this nine-year old and his father of our students taking required courses in business statistics. They seem to encounter much difficulty compared with other business courses they take. It is possible that much of the difficulty arises because students perceive statistics as a bewildering array of arcane techniques, bristling with hocus-pocus formulas, the eternal mystery of degrees of freedom, brain-teasing word problems, and a Rubik's cube model of the real world, with data grouped and scrambled beyond hope of reconstruction.

As a result of students perceiving statistics as an overabundance of seemingly unconnected methods and problems, statistics, rather than being viewed as a powerful research tool, is viewed simplistically as the application of a collection of many methods to hundreds of disjoint problems. At best, such an approach to the study of statistics is unrealistic. On the surface, students are being taught how to use statistics practically and with discretion. But their perceived satisfaction of knowledge exceeds reality, so that their little knowledge is a dangerous thing. Unfortunately, the simplified approach to teaching statistics presupposes without verification that all underlying mathematical and methodological assumptions are met. Finally, typical case studies are designed to illustrate only a limited number of concepts.

The authors maintain that much of the trouble we have mentioned could be eliminated if statistics were presented as a scientific investigation of a single comprehensive real-life case study, which we conditionally call a "scientific legend".

I. COMPREHENSIVE CASE STUDY ... PLUS

We have developed a method of teaching Introductory Statistics I and II at Rider College using this case-study approach. Why the name "scientific legend"? Well, a legend is a story more real than reality (distilling the essentials into a compact "case") and our legend is this essence pressed into the use of science! A story teaches reality by distillation, but is neither a small intricate puzzle nor an overwhelming mass of confusion.

Existing case studies on business statistics are based either on samples or on a population data set. Case studies based on samples are designed to present statistics as a complex of disjoint problems. On the other hand, case studies based on a population data set are designed to provide students with unified statistical problems rather than theory. Both types of case studies do not consider statistics as a science, nor do they teach students how to verify the underlying assumptions under which statistical methods are valid.

A scientific legend (SL) is a comprehensive case study based on a population data set. Our particular scientific legend consists of real-life data (150 family budgets) constructed from average values taken from the 1977 Statistical Abstract of the United States. Each budget contains values for nine family characteristics (variables): husband's education; husband's job position; husband's height, wife's height; son's age; family income; number of rooms; expenditures on housing; and, annual bread consumption.

Our broad goal in constructing the legend was to create a vehicle that could be used to present statistics as a scientific investigation of a connected set of real-life economic problems as opposed to a collection of disjoint questions. Thus, in our classes a single set of data is used throughout the semester to illustrate all principal problems and methods usually covered in a business statistics course.

At the beginning of the course, the legend is used to explain the theory of probability. This material is crucial as it forms the foundation for all statistical theory to be covered later on. With the help of the deductive method the legend is used to explain laws of generating samples from populations (in this case, the population consists of the 150 family budgets).

With this background knowledge, the legend is used as the basis for considering generalized statistical measures. Here, each student works with a unique subsample of observations (32 or more budgets) and studies the remainder of the topics. The scientific legend is designed to present statistics as a science through the following stages:

1. Analysis of statistical data set quality.
2. Verification of underlying assumptions.
3. Application of statistical methods.

Audio-visual aids and computer-assistance enable presentation of a scientific legend and its concept. Let us consider the main three conceptual stages of the scientific legend:

1. Success of any statistical research requires coordination between theory (the theory of probability) and practice (data analysis). The theory of probability is based on the foundational assumption that a population is homogeneous (e.g., an honest coin, an unloaded cubical die). Coordination between theory and practice requires the same foundational assumption for the statistical data analysis. Therefore, we can state that statistical data analysis should be based on a homogeneous statistical population.

Quality in economic research depends upon the generation of one-dimensional and n-dimensional homogeneous statistical distributions. These distributions can be considered as a bridge between the theory of probability and statistical analysis in economics. Our student source population is, by design, homogeneous.

2. The verification of underlying assumptions is the second important stage in the use of a scientific legend. On the basis of SL we verify the following underlying assumptions of regression models:

a. Independence between factors.
b. Homoscedasticity of conditional variance.
c. Normality of the residual variable.

For the solution of all these problems the Chi-Square test statistic and the K-S test statistic are employed. Classical theory provides a weak inference because it is based on a single sample subject to errors. We suggest that a stronger inference be constructed on the basis of multiple samples.

3. The scientific legend is designed as a vehicle to teach business statistics on the basis of a "realist" approach. The realist approach follows these rules:

a. Difficulties in statistics should be detected and explained, not avoided.
b. Practical explanation should be made first, and abstract theory second.
c. One should teach "Business Statistics" through practical exercises, not through riddles.
d. In business statistics, linking of terms, definitions, and the concept of data quality (and quality, not necessarily quantity!) is very important, remembering Richard Hamming's dictum, "The purpose of computing is insight, not numbers".

e. Statistics should be taught on the basis of continuity among topics. We offer the following logical sequence in our courses:

Foundations of statistics
Generalized measures
Statistical analysis of regularities
Statistical inferences
Time series
Index numbers

II. AUDIO-VISUAL AIDS

The method of the scientific legend can be practically taught only with the help of audio-visual aids and computer assistance. The purpose in constructing the method was to provide a vehicle that can be used to present statistics as a scientific investigation of a connected set of economic problems (in contrast to a collection of disjoint problems). Through this method, students, working with unique subsamples, analyze a connected set of ungrouped data. Emphasis is put on the verification of statistical data quality before the application of statistical methods.

The traditional chalk-and-blackboard teaching technique, adequate for the customary simplified examples, is insufficient when one deals with ungrouped data. Therefore the realist approach requires transparencies for the instructor and a special student manual for the class. Since student problems have individual answers, computer assistance is required for the generation of homework and exam answers. A standard text provides theoretical material to reinforce the lectures, and videotapes or films provide a backup and review, but not a substitute for the lectures.

Statistics through transparencies is characterized by many illustrations in pictures, graphs, tables, and formulas. Transparencies have been elaborated with didactic attention to the following points:

a. Principle of the still: one idea in simple form with few words
b. Window method (progressive disclosure)
c. Link (sequence) method
d. Method of repetition

Statistics films are mainly based on materials prepared for transparencies. Designed primarily for students missing lectures or wishing to review, they are characterized by fewer words, many illustrations, and a quicker speed of presentation. Our experience has shown that the speed of presentation of materials in a film can be as much as four times that of an ordinary lecture. Since the films will be housed in the college library, they are easily accessible

to students, and since they are all based on data contained in the "scientific legend", continuity in material coverage is more easily effected.

The method of the scientific legend lends itself to the development of interactive teaching and instant feedback on the correctness of homework through the use of terminal oriented mainframes or minicomputers. Students are therefore able to pace themselves, to gain reassurance, and to avoid series of errors due to interproblem dependencies.

Proposed Benefits

Sicne unique subsamples (consisting of 32 or more family budgets) can be drawn from the population, each student works with a unique set of data throughout the term for homework and, if the instructor desires, examination purposes. Although comprehensive case studies (or so-called "practice sets") have been successfully used in other business courses (e.g. accounting), the authors are unaware of any attempts to use this approach in the teaching of statistics. That the legend can be used to generate individualized comprehensive case studies is almost certainly novel. The use of adjusted real-life, rather than purely fabricated, data is designed both to enhance student interest and involvement and to present statistics as a powerful yet practical research tool.

Since the measurement level of the family budget data with which the students are working varies (nominal, ordinal, interval, ratio), it is easier to explain the circumstances under which particular statistical methods are appropriate or inappropriate. The goal of avoiding indiscriminate application of techniques without prior verification of mathematical assumptions is therefore continaully stressed. In our view, this constitutes a more realistic and beneficial teaching approach.

The use of individual samples from a common population generates the additional benefit that consciousness of variability in sample data has already been raised when the subject of sampling distributions comes up. The probabilistic nature of statistical inference is apparent in contrast to the "proofs" of other scientific and logical systems. The use of computers by students at a success and sophistication level far above that experienced in beginning programming courses should help to reinforce positive experiences in the use of these tools.

III. COMPUTER IMPLEMENTATION

Computer-assisted instruction has not taken over the world; certainly not in statistics. Perhaps it is that the existence of computer-aided instructional methods has preceded their essence (Michelson, 1970). We may, just as we need new wineskins to contain new wine, need innovative theory to contain

new methods. The method of the scientific legend is an attempt to provide just such a theory most suitable for implementation on computer media.

"Experimental physics is the strongest argument for scientific realism. The current fad in philosophy is due to doing theory, not experimentation. Nothing but experiment is so lacking in science.... A good experimenter will bring forth phenomena that have not even been seen before." (Hacking, 1982). In statistical education, data and examples are particularly important in the impartation of insight. We only need continually to fight the tendency to lose the insight among all the numbers!

The SL method is so regular that we are able to implement it rather straightforwardly on the computer. [We are now using a Digital Equipment 2060 and writing in an extended Fortran]. Not that any system of this magnitude is trivial or small to implement, but its feasibility has been heartening. We are taking a structured approach to the programming, keeping computational, printing, data management, and interactive functions as separated as possible. Each student at the beginning of the course is assigned an id. number and a password. The id. numbers are the keys to the table of random numbers and the password is used to prevent one student from tampering with another's results or from seeing another's grades.

Several methods of handling the generation of samples might be possible. We could try to do this using published random number tables, and devise a method for each student to select suitable numbers within the range of family numbers in the population. Alternately, we could use the computer to generate random numbers. This is the method we are using. At present, we are taking advantage of a prime-modulus generator supplied with our Fortran function package, generating numbers in floating point form, and by a linear transformation mapping them onto integers in the range of 1 to 150, matching the population. We have tested the output of this generator and function for uniformity of the distribution and are satisfied that it provides numbers suitable for these educational purposes. Before converting the programs to another computer system, we will want to research a portable number generator that will provide numbers that do not vary with hardware and method of arithmetic. The new IEEE standards on computer arithmetic may eventually help in this matter. Of course, these "random" numbers must not be random with each computer system or execution, or the student id.s would be untraceable and meaningless. Since they are generated by a deterministic algorithm, they are repeatable.

We are preserving on disk whole sets of random numbers sufficient for several years' of students. These are indexed with the id. key number given to each student at the beginning of the term. Then the random numbers are used in the computer just as the student uses them: to select subsamples from the

population, which is kept on another disk file.

At present, we have printed answers to all homework and examination problems. These can be used to check the students' final answers. For some chapters we also have generated detailed numerical solutions for each student which are useful in diagnosing where the student has gone astray in his work. We intend to finish these solutions and implement a further refinement in the next year, namely, a program to check a student's homework interactively at a terminal.

The interactive method will work as follows. Students log-in to a special account with execute-only privileges, giving their random number key id. and password. The defaults to this special account take over from there. The student is asked for which problem he wishes to check his answers. Answers are checked, telling the student only whether he is correct within a) a high degree of accuracy or b) within an acceptable degree of accuracy or c) not correct. The accuracy required is generally plus or minus a few percent of the correct value, but depends upon the particular problem and the sensitivity of the calculation. Students are not given the correct answers if wrong, but are sent back to try again until they have mastered the problem. The relative accuracy and the number of times attempted to get within the required accuracy is recorded in the grade table data area for this student. An average score is kept on a running basis according to a pre-announced formula for each student and is available at each session for his perusal. This feature allows an instant measure of progress, a satisfaction we find students crave. Periodically, the instructor may obtain summary printouts on student progress and can call attention to those students who have not progressed in a timely fashion. These reports also serve as backup documents on the grading in the outside event of disk failure or security compromise of the system. In this process virtually all grading labor has been automated, freeing the instructor for more profitable activities such as working individually with students.

A further level of computer implementation would involve computer-assisted instruction (CAI) wherein the student could be walked through the solutions to problems he is not able to get on his own. This would involve a far greater investment in terminal time than the checking of pre-worked answers, as well as a greater level of program sophistication. At present we plan to leave this function to the instructor who can walk through problems with the help of detailed printouts that can be prepared either ahead of time or on demand for each student.

Program Structures Required (see Fig.1)

The following programming structures are required:

1. An executive routine which handles interaction with the students at the terminals and calls the appropriate calculation subroutines.

2. A set of subroutines to implement each of the statistical methods to be taught (e.g. mean, dispersion, regression analysis, analysis of variance, etc.). These routines have interface with the data fields of the legend or legends.

3. A table containing data for the particular legend being employed.

4. A routine to generate tables of random numbers keyed to student number (once for the term).

5. A table of grades and a record of interaction with each student logging-in.

6. A record of statistical information about the operational experience of the system for later analysis (and, we hope, future papers!).

7. Routines for instructors' retrieval of grades and data as well as prepared student materials such as sets of random numbers, detailed problem solutions, exam blanks, etc.

Data and Program Security and Implications for Automatic Grading (see Fig.2)

A number of security problems arise when students are left alone to work on the computer. Our problem is how to minimize these exposures in a practical system. The primary security comes from the nature of the legend method itself. Since each student has individual assignments, no student can copy another's answers. For homework, this is probably sufficient, since the burden of doing assignments for more than one student is too great for anyone. For exams and quizzes other problems arise. One student could pay or coax another student to take the exam for him if it were to be directly entered into the computer without proctoring. The computer cannot see; any identification a student might be asked to provide at log-in could be transferred to his proxy. The following method is suggested to preclude this possibility while maintaining a high degree of automation.

An individual pair of identical answer sheets (possibly with carbon paper) are prepared by the computer system. Exams are administered by the instructor in the usual manner. One copy of the answer sheet is collected at the end of the period and the other is kept by the student for entry at the terminal as in the manner of the homework. After a sufficient time, the computer prints for the instructor copies of the answer sheet filled-in with the students' responses.

These can be spot-checked or candled with the retained copies for differences. the computer will add up scores as usual; any exceptions from the verification (retained) copies can be entered by the instructor. If an optical or mark-sense scanner is available, the student participation in entering answers is not necessary, as the exams can be entered directly in machine-readable form.

Of further consideration is the matter of tampering with the system. Data files (including student grades) and system programs must be protected from alteration by enterprising or malevolent students. Several measures can be taken. First, the programs can be stored in load module form only, which is virtually indecipherable; it will only execute and cannot be read. Second, data file names can be camouflaged and kept secret. The references to them from the programs will not be decipherable in the load modules. As a further security measure, the grade data can be encoded so that even if students were to stumble upon the files, they would be unprintable. Also, they would be unalterable without destruction. Additionally, the protection of the operating system should be used to isolate the vital files from all users other than the programs of the legend system with a "need to know". Finally, programs and data segments are kept in separate accounts, separately protected and with camouflaged account names. A further hazard can be created by the student who wishes to abort the program in mid-session or by system crashes. We therefore find it necessary to trap terminal interrupts (control-C on the DEC-20) so as not to allow the student to abort suddenly, and we must keep a record on disk of the partial progress of each student.

A Review of Our Experiences to Date

One of the ironies of the CAI experience is that computers can provide a level of "personal" attention to the student that faculty are unable to do for lack of time. This is particularly important in light of the large sections now encountered in business schools and other departments teaching "practical" subjects. Further, the workbook and the computer allow the fast students to proceed at their own pace and slow students to take their time. Some of the homework deadline pressure is relieved, since assignments do not have to be due on particular days, but students can enter their results any time computer terminals are working and available, which in most hours of the day and night. The homework is graded by formulas for accuracy and with the speed and patience that only a computer possesses. Tests and quizzes can be given with the frequency that an instructor desires, knowing both that grading tedium has been relieved and that all does not rest upon the exams, but that homework is also a reliable measure of student performance. As in all courses, students want frequent evaluation of their progress. This is particularly so in a grade-conscious generation. The use of computer terminals makes this feasible.

We believe from our experiences to date that this will have the further effect of keeping up morale and interest in the course.

From our experience of teaching by the method of the scientific legend in the last year, we offer these observations about what we have done right and what we wish to alter for next year: First, what is right: Students work hard and gain much experience in practical statistical calculations. Many students find an interest in this method where they had not from conventional approaches. For the teacher it is a relaxing and self-cueing method. (The authors do not find they need to bring any notes to class). Nevertheless, some students say they are addicted to chalk dust (author Davis personally is allergic to it!) and they cannot learn if there is not enough of it flying during a lecture. One student told another, "This is the most bizarre way of learning I've ever seen". However, we have found that after a few weeks the conservatism and the shock of new procedures wears off. (Our methods if observed resemble more the corporate and continuing-education methods that are now prevalent than they do the traditional academic classroom atmosphere).

At the beginning of the course term we would emphasize more attention to detail. We have found many students are sloppy in their habits and we feel one important lesson in a beginning statistics course is emphasis upon accuracy and form. Round-off is another subject that remains a mystery to students unless specifically stressed. Topics of sensitivity of the various methods to errors and the effect of roundoff, e.g., in regression analysis, will be emphasized more in the future. Students will be encouraged to develop a sense of judgement in these matters and to present final results with an appropriate degree of accuracy. Until a complete text is developed with this method, students will be encouraged not to miss class, or in that rare event to review by way of the videotapes. Since this is a unified method, students can lose depth by falling seriously behind. There is always a temptation in this method to work by cookbook, but serious errors in judgement will result thereby.

In order to break what can become after a long time "slide hypnosis" and to afford more opportunity for student interaction, we plan more frequent short quizzes and student-led seminars on the material. We wish to emphasize the importance of obtaining a good personal calculator with scientific functions, and possibly with programmability and elementary statistical operations built-in. We believe many students have been slowed down unnecessarily by tedium in using crude four-function calculators. The college is obtaining a graphics software package and a full-color plotter on its computer next year and we hope to connect this so that students can get plots of their individual sample distributions. We believe this will be another feature to increase student interest

in the material.

CONCLUSION

Our motivation in developing a "scientific legend" has been to encourage students to view statistics as a unified set of procedures by which to scientifically test probabilistic observations about the real world. The existence in the last few years of certain technology on the market, viz., terminal-oriented computers and inexpensive hand-held calculators has made the contemplation of this approach possible. We are attempting to motivate them to work with real-world, ungrouped data, and have been developing a computer package designed solely for instructional purposes.

References

[1] Anderson, Donald T. and Rivers, Richard A. (1980) Practice Problem in Financial Accounting: Saluki United Vending Service, Inc. Houston, TX: Dame Publications.

[2] Cooley, William W. (1969) "Computer Assisted Instruction in Statistics". Statistical Computation, Roy C. Milton and John A. Nelder, eds., New York: Academic Press.

[3] Forsyte, Alan B. and Freed, Jas R., (1979) "An Evaluation of Computer-Aided Instruction in an Introductory Biostatistics Course". J. Computer-Based Instruction, February, 5, 57-62.

[4] Gallant, A. Ronald, (1973) "Some Arguments Against the Use of Statistical Packages in Teaching Statistical Methods". Proc. Computer Science and Statistics: Seventh Annual Symposium on the Interface, October 18-19, pp.223-225. [Plus, other papers this volume].

[5] Gross, Beatrice and Gross, Ronald. (1980) A Review of Innovative Approaches to College Teaching. Sarasota FL: American Accounting Association.

[6] Hacking, Ian (1982) (Stanford University Department of Philosophy), "Experimentation and Scientific Realism", Princeton University Philosophy Seminar, January 29th.

[7] Hobbs, Gerald R., Jr. "Some Advantages of Individually Generated Data Sets in Statistical Methodology Courses". - in forthcoming volume by Academic Press, pp.465-469.

[8] Michelson, S. (1970) "The Existential Reality of Educational Production Functions". Washington, DC: Office of Education.

[9] Shvyrkov, V.V. (1982) Scientific Legend in Applied Probability and Statistics, (Vol.I) Lawrenceville, NJ: G. Throwkoff Press (Student workbook).

[10] Shvyrkov, V.V. (1982) Scientific Legend in Statistical Science, (Vol.II) Lawrenceville, NJ: G. Throwkoff Press (Student workbook).

[11] Tobias, S. (1969) "Effect of Creativity, Response Mode, and Subject Matter Familiarity on Achievement from Programmed Instruction". J. Educational Psychology, 60, 453-460.

[For further references on computer-aided instruction, see our paper (A.C. Davis and V.V. Shvyrkov, "Computer Assistance in Teaching Statistics Using Method of the Scientific Legend") Proc. 1981 Joint Statistical Meetings, (August 10-13, 1981, Detroit MI.)].

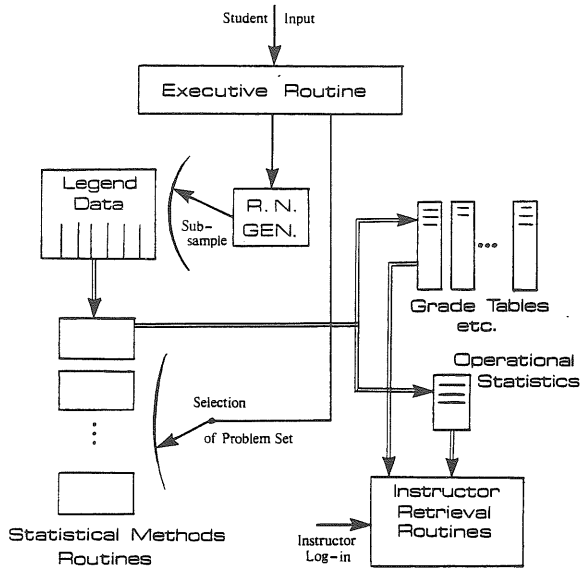
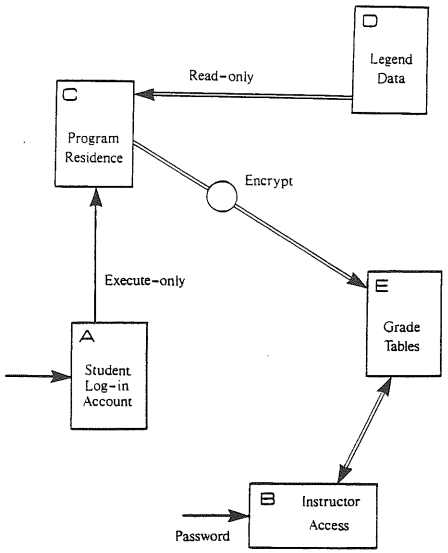


Fig. 1 Program Structures



Squares Represent File Protection Areas

Fig. 2 Program & Data Security

#	HUSBAND EDUCATION	HUSBAND JOB POSITION	HUSBAND HEIGHT IN INCHES	WIFE'S HEIGHT IN INCHES	SON'S AGE (IN YEARS)	INCOME $1000S	NUMBER OF ROOMS	EXPENDITURE ON HOUSING $1000S	CONSUM. OF BREAD (IN POUNDS)
1	H	S	71.0	71.0	12.9	33.0	6	6.83	315.0
2	E	N	67.0	59.0	7.4	5.0	2	1.12	302.5
3	E	S	69.7	69.0	15.9	33.0	6	6.50	325.0
4	E	N	64.3	62.1	9.2	8.8	2	2.86	305.8
5	C	S	71.2	69.5	12.9		7	7.50	325.0
6	H	N	66.1			15.3	5		
140	E	N	70.9	65.4	11.8	14.1	4	3.16	321.0
141	E	N	70.1	65.2	11.7	13.8	3	3.42	319.5
142	H	N	66.9	66.5	9.3	15.2	3	3.47	281.7
143	H	N	71.4	63.7	8.0	10.8	4	2.62	297.7
144	E	N	69.4	65.1	11.6	13.5	3	3.31	318.7
145	E	N	68.2	69.0	13.8	19.5	3	4.50	302.5
146	E	N	69.7	67.2	12.9	17.0	4	3.91	329.0
147	H	S	72.3	67.8	12.6	23.7	3	4.90	285.0
148	E	N	69.9	63.2	10.3	10.8	4	2.37	313.7
149	C	S	70.6	63.4	9.4	13.3	3	3.25	305.0
150	H	N	70.4	63.4	8.0	10.5	4	2.12	296.8

EXPLANATION OF SYMBOLS

E ELEMENTARY SCHOOL EDUCATION
H HIGH SCHOOL EDUCATION
C COLLEGE EDUCATION

N NON-SUPERVISORY POSITION
S SUPERVISORY POSITION

Fig.3 The Source Population

```
 EXECUTE CHAP2
LINK:   Loading
[LNKXCT CHAP2 execution]

 PROGRAM TO PRINT ANSWERS TO PROBLEMS IN CHAPTER  2

ON DISK ARE LEGENDS A THROUGH E, EACH CONSISTING OF
160 STUDENT CASES OF 32 RANDOM NUMBERS EACH

TYPE IN LEGEND LETTER (LIMITS: A TO Z) B
TYPE IN STARTING AND FINISHING STUDENT NUMBERS,
SEPARATED BY A BLANK  17 17
TYPE IN STARTING AND FINISHING PROBLEM NUMBERS,
SEPARATED BY A BLANK  18 21
FOR REPORT IN STUDENT-MAJOR ORDER, TYPE STUDENT ELSE TYPE PROBLEM
FOR PRINTED OUTPUT, TYPE THE WORD print AND PRESS RETURN;
FOR TYPED OUTPUT PLACE PAPER ON LAST LINE OF PAGE AND PRESS RETURN.
TO QUEUE OUTPUT TO DISK FILE FOR30.DAT (E.G., TO PRINT LATER
ON SPECIAL FORMS), TYPE THE WORD disk AND PRESS RETURN.

PROBLEM NUMBER 2.18

LEGEND B

STUDENT NUMBER: 17

| E\J |   B̄   |   B   |  TOT  |
|  Ā  |  .835 |  .165 | 1.000 |
|  A  |  .271 |  .729 | 1.000 |
|     |       |       |       |

PROBLEM NUMBER 2.19

LEGEND B

STUDENT NUMBER: 17

| E\J |   B̄   |   B   |       |
|  Ā  |  .893 |  .380 |       |
|  A  |  .107 |  .620 |       |
| TOT | 1.000 | 1.000 |       |

PROBLEM NUMBER 2.20     LEGEND B     STUDENT NUMBER: 17

| STAGES | P(B̄/Ā) | P(B/Ā) | P(B̄/A) | P(B/A) |
|    1   | 0.792 | 0.219 | 0.208 | 0.781 |
|    2   | 0.835 | 0.271 | 0.165 | 0.729 |
|    3   | 0.846 | 0.288 | 0.154 | 0.712 |
|    4   | 0.849 | 0.293 | 0.151 | 0.707 |
|    5   | 0.850 | 0.294 | 0.150 | 0.706 |
|    6   | 0.850 | 0.294 | 0.150 | 0.706 |

PROBLEM NUMBER 2.21     LEGEND B     STUDENT NUMBER: 17

| STAGES | P(Ā/B̄) | P(A/B̄) | P(Ā/B) | P(A/B) |
|    0   | 0.950 | 0.050 | 0.583 | 0.417 |
|    1   | 0.907 | 0.093 | 0.419 | 0.581 |
|    2   | 0.893 | 0.107 | 0.380 | 0.620 |
|    3   | 0.888 | 0.112 | 0.369 | 0.631 |
|    4   | 0.887 | 0.113 | 0.366 | 0.634 |
|    5   | 0.887 | 0.113 | 0.365 | 0.635 |
|    6   | 0.886 | 0.114 | 0.365 | 0.635 |

CPU time 1.25   Elapsed time 54.06
@
```

Fig.4 Instructors' Retrieval Routines


```
@EXECUTE CHAP9
LINK:   Loading
[LNKXCT CHAP9 execution]

 PROGRAM TO PRINT ANSWERS TO PROBLEMS IN CHAPTER  9

ON DISK ARE LEGENDS A THROUGH E, EACH CONSISTING OF
160 STUDENT CASES OF 32 RANDOM NUMBERS EACH

TYPE IN LEGEND LETTER (LIMITS: A TO Z) A
TYPE IN STARTING AND FINISHING STUDENT NUMBERS,
SEPARATED BY A BLANK  134 134
TYPE IN STARTING AND FINISHING PROBLEM NUMBERS,
SEPARATED BY A BLANK  1 4
FOR REPORT IN STUDENT-MAJOR ORDER, TYPE STUDENT ELSE TYPE PROBLEM
FOR PRINTED OUTPUT, TYPE THE WORD print AND PRESS RETURN;
FOR TYPED OUTPUT PLACE PAPER ON LAST LINE OF PAGE AND PRESS RETURN.
TO QUEUE OUTPUT TO DISK FILE FOR30.DAT (E.G., TO PRINT LATER
ON SPECIAL FORMS), TYPE THE WORD disk AND PRESS RETURN.


PROBLEM NUMBER 1

 x     y'     f

 2   2.965    2

 3   2.403    4

 4   3.470    5

 5   5.454    5

PROBLEM NUMBER 2

SUM x =   61.   SUM y = 60.16   SUM xy =   246.44   SUM x SQ = 249.

y(x) =  -0.202 +  1.039x

PROBLEM 3          ELASTICITY=  1.054

PROBLEM # 4    STUDENT # A134

 1  -----
 2  0.375
 3  0.250
 4  0.300
 5  0.300
 6  -----
 7  -----
CPU time 1.14   Elapsed time 23.16
```

Fig.5 Another Answer Display

STAT 1- Midterm #1

Date:_____ Grade:_____

NAME: PRINTED _____ LEGEND: A STUDENT NUMBER: 15

#	1	2	3	4	5	6
PROBLEMS!	2.1	2.2	2.3	2.4	2.5	2.6
\REQUIRED!						
IRNS \	a1	$P(a \cap b)$	$P(a \cup b)$	$P(a \mid b)$	$P(b \mid a)$	$P(b)$
1						
2						

Random Number Set 1 ONLY:

#	IProb.&IReq'd	Results
7	2.7	$P(A \cap B)$
8	2.8	$P(A \mid B)$
9	2.9	$P(B \mid A)$
10	2.10	$P(A \cap B)$
11	2.11	$P(A \cap \bar{B})$
12	2.12	$P(\bar{b} \mid a)$
13	2.13	$P(\bar{b} \mid \bar{a})$
14	2.14	$P(\bar{a} \mid b)$
15	2.15	$P(\bar{a} \mid \bar{b})$
16	2.22	$\bar{X}=$
17	2.23	$S=$
18	2.24	$P(s=1)$
19	2.25	$\bar{X}=$
20	2.26	$S=$

STAT 1- Midterm #1

Date:_____ Grade:_____

NAME: PRINTED _____ LEGEND: A STUDENT NUMBER: 15

#	1	2	3	4	5	6
PROBLEMS!	2.1	2.2	2.3	2.4	2.5	2.6
\REQUIRED!						
IRNS \	a1	$P(a \cap b)$	$P(a \cup b)$	$P(a \mid b)$	$P(b \mid a)$	$P(b)$
1	21	.656	.875	.600	.955	.200	.800	.688
2	15	.625	.906	.727	.952	.250	.750	.656

Random Number Set 1 ONLY:

#	IProb.&IReq'd	Results
7	2.7	$P(A \cap B)$	.313
8	2.8	$P(A \mid B)$	.294
9	2.9	$P(B \mid A)$	.333
10	2.10	$P(A \cap B)$	.220
11	2.11	$P(A \cap \bar{B})$	.656
12	2.12	$P(\bar{b} \mid a)$	.200
13	2.13	$P(\bar{b} \mid \bar{a})$	.222
14	2.14	$P(\bar{a} \mid b)$	.600
15	2.15	$P(\bar{a} \mid \bar{b})$	.955
16	2.22	$\bar{X}=$	.156
17	2.23	$S=$	.363
18	2.24	$P(s=1)$	.334
19	2.25	$\bar{X}=$	.312
20	2.26	$S=$	.513

Fig.6 An Exam Blank and an Individual Answer Key (reduced in size)

LEGEND A STUDENT NUMBER: 3

RANDOM NUMBER SET 1:

145	85	5	106	110	81	50	51	79	112	87	58	147	44	138	62
129	65	66	140	143	113	53	28	115	56	52	122	1	69	76	150

RANDOM NUMBER SET 2:

46	131	22	149	65	50	79	58	104	6	35	21	47	94	52	42
34	110	2	72	122	150	146	18	31	108	106	83	117	40	136	111

LEGEND A STUDENT NUMBER: 4

RANDOM NUMBER SET 1:

85	32	5	22	99	28	63	108	128	145	107	19	97	2	12	59
26	29	84	56	100	43	90	106	149	73	69	61	82	36	1	136

RANDOM NUMBER SET 2:

46	50	51	34	91	108	126	130	96	88	19	81	77	15	62	127
112	20	145	63	57	125	144	133	70	142	8	71	68	131	114	101

LEGEND A STUDENT NUMBER: 5

RANDOM NUMBER SET 1:

113	78	15	123	59	35	38	62	103	53	147	30	36	37	74	125
61	133	34	101	141	107	1	64	90	110	142	82	48	26	3	106

RANDOM NUMBER SET 2:

40	29	86	14	74	117	132	139	65	134	30	35	39	114	10	70
43	60	44	84	106	27	103	82	147	150	112	142	8	76	61	56

Fig. 7 Random Numbers

TOPIC 9: TRAINING STATISTICAL PRACTITIONERS

ORGANISER: PROFESSOR M BENYAKLEF

TRAINING OF STATISTICAL PRACTITIONERS

P Osinski, Munich Centre for Advanced Training in Applied Statistics for Developing Countries of the Carl Duisberg-Gesellschaft (CDG)

INTRODUCTION

The training of statistical practitioners is an important issue for both developed and developing countries. However, this paper will only deal with the training of statisticians from developing countries. It is felt that statistical offices in developed countries have the means and the possibilities to train their own staff, whenever necessary, to cope with new tasks or to introduce new methods of data collection or analysis. Newcomers to the statistical service are always trained "on the job" and most heads of offices would agree that it takes some time to transform a learned statistician into a statistical practitioner.

In most developing countries this problem is of even greater importance. In recent years a number of specialized schools and institutes have been created in LDC's where young people can obtain a degree in statistics after 2-3 years of study. Statistical theory and mathematical methods are important subjects in the curriculum of these schools and it is not at all certain that the diploma obtained will enable the young statistician to cope with the day-to-day work when joining a statistical office. Hence, there must be some additional practice-oriented training to supplement the theoretical knowledge acquired at statistical schools or universities.

While the issue of "academic" and "practical" training remains controversial the fact is that the main efforts of international organizations and some LDC Governments are directed towards "academic" training leading to a degree in statistics. The Statistical Training Programme for Africa (STPA) fostered by the Economic Commission for Africa (ECA) is mainly meant to support statistical schools in Africa. Relatively little attention is paid to the holder of a diploma of statistics once his studies are terminated. Occasional workshops on specific items of statistics organized internationally indicate the need for further training but they are completely insufficient to solve the problem.

The main concern of this paper is to outline some ideas on how statisticians after having terminated their basic statistical education can most effectively be trained (a) to cope with the tasks they are confronted with and (b) to keep abreast of new developments in statistics. The paper draws heavily on the experience the Munich Centre has gained since 1973 in organizing training courses on specific subjects of statistics. Reference is also made to the author's paper "Follow-up courses", presented at the Fourth Meeting of Directors of International Statistical Training Centers - Rabat, Morocco, May 23-26, 1979.

CHARACTERISTICS OF TRAINING STATISTICAL PRACTITIONERS

The need for improving the quality of statistical personnel is a recognized and long-standing aim of practically all LDC statistical offices. This stems mainly from the request for more and better statistics made by both national and international users. In order to cope with the demand put upon them heads of offices have undertaken considerable efforts in this respect which are starting to pay off increasingly. It should, however, be noted that the training of statisticians does not always figure prominently on the priority list of LDC Governments.

The methods available for improving the skills of statisticians can be summarized in the following way: On the job training (formal or informal), training outside the office in specialized institutions or training centres and a combination of both.

Training on the job is the method widely used in developed and developing countries. It applies to both newcomers to the office and to personnel already employed. Informal training consists of a person being introduced to his job by his superior or supervisor. There is no defined training programme and the trainee learns by actually doing certain jobs which hopefully become progressively difficult. The training is strictly individual and its success depends largely on the activity and the willingness of the supervisor to transfer his knowledge. An advantage of this type of training is that the trainee learns only methods relevant to executing his job. However, there is also a certain danger that the training may become one-sided, especially in offices where the various tasks are clearly separated. This would call for a rotation programme within the office in order to avoid over-specialization of personnel.

Formal on the job training requires a much higher degree of organization. A programme has to be developed, the categories of people to be trained have to be defined and serious thought has to be given to the length of the courses envisaged and the availability of qualified teaching staff. Participation in training courses - although normally on a voluntary basis - should have priority over other office work to make sure that people are actually released from their duties without difficulties. The simplest form of this kind of training are courses organized from time to time on an irregular basis. A more sophisticated form used in the Federal Statistical Office of Germany consists of breaking down a rather voluminous training programme into several parts which staff members aiming at a post of responsibility are supposed to go through within a period of two years. The advantage of formal training is that a great number of young people can participate receiving all the same instruction. On the other hand course participants may feel that some of

the topics of the programme are not very relevant to the work they are presently doing. This underlines the importance of a well-conceived and practice-oriented training programme.

While the conditions for executing a staff training programme exist in developed countries this is not the case in most of the LDC statistical offices. There is a lack of experience with such programmes, qualified teaching staff is in short supply and the material prerequisites such as lecture notes, background literature etc. are missing.

Under these circumstances many of the LDC statistical offices have recourse to training their staff in specialized institutions or training centers, in order to upgrade their knowledge in certain fields of applied statistics. Short-term courses are usually organized by the regional commissions of the U.N. or by institutions created in certain donor countries such as France, Germany, Netherlands, UK, United States and Japan. Most of the courses are held in the developed countries concerned, but more recently there is a tendency to organize them also in developing countries preferably on a regional basis.

What kind of possibilities does the director of an LDC statistical office have to upgrade the knowledge of his personnel in fields like agricultural or industrial statistics or national accounts? Courses in these topics are not readily available in developing countries nor are they in developed countries. What remains is to attempt a sort of "on the job training", however difficult it may be, and to use outside institutions for further training whenever the occasion arises.

In this context it should be mentioned that the upgrading of the level of statistical academic education is also widely operated by LDC statistical offices. In general the staff member is released from his duties for a period of 2-3 years with the obligation to return to the office after the termination of his studies and to work there for a certain period of time. If not doing so, staff members have to pay back the expenses incurred by the Government for their studies. The observance of this rule, however, is difficult to enforce.

PURPOSE OF TRAINING

A statistician participating in a course of short duration on a specific type of statistics is likely to be interested primarily in items having direct relevance to his job. He rightly expects that his efforts will enable him to have a better insight into the problems he is confronted with in order to do a better job. The first and overriding consideration when organizing such courses, therefore, is that they are practice-oriented.

This is easier said than done. Most people will admit that the borderline between "theory" and "practice" is artificial and that a certain amount of theory has its place in any course programme, however practical it may be. On the other hand, a practical course can never be expected to cover all facets of reality. Many statisticians participating in a course expect just this to happen. The ideal for them would be to go home with a sort of recipe book listing all the steps to be undertaken in a statisical survey and giving detailed advice on how to deal with all possible problems. The trouble is that local situations in LDC's are extremely variable and do not lend themselves to ready-made solutions.

The aim of training should, therefore, be to upgrade the specialized subject-matter knowledge of the statistician enabling him to take decisions in an independent and well conceived manner whenever problems arise. To do this he needs a thorough knowledge of the concepts and definitions used in his special field of statistics and some background information why certain concepts or methods have been given preference as compared to other possible solutions. The exchange of experience among statisticians is also very important in order to learn from each other and to adopt methods of data collection, treatment, publication and analysis which already have been used successfully.

SCOPE OF TRAINING

When organizing training courses it is of vital importance that the topic selected is well defined. In principal one can choose either a methodological approach dealing with selected statistical methods applicable to all fields of applied statistics or a sectoral approach concentrating on specific subject-matter such as agricultural statistics, industrial statistics etc. The latter method has the advantage that it corresponds to the organizational set-up of most statistical offices in LDC's. Typical items to be treated are data collection methods (censuses, sample surveys), questionnaire design, data processing with special reference to EDP, data analysis and publication. While the topics are treated in more general terms when using the methodological approach, teaching is geared to the specific subject when using the sectoral approach. The inter-relationship of the subject-matter statistics under consideration with other statistics should also be shown in order to rouse the awareness of the statistician that the statistics he is doing are an integrated part of an overall framework and hence need coordination with related statistics.

When dealing with statisticians from LDC's it should be kept in mind that many statistical offices are still badly equipped with both manpower and material. Therefore, it makes no sense to confront statisticians from these countries with the most sophisticated methods of data analysis or publication as there

is no chance that they will be able to apply them in the near future. This does not mean that very new and advanced methods should be completely excluded from the lectures, but main emphasis should undoubtedly be on methods which can be readily applied.

Another question is whether related subjects should also be included in a training programme. When dealing with economic statistics some selected lectures on the functioning of the economy would certainly be useful. The study of population statistics may also be supplemented, for instance by lectures on sociological aspects of vital statistics. It is somewhat difficult to decide how far one should go in including related subjects in a statistics course. It is clear, however, that a profound knowledge of the subject area a statistician is studying (agriculture, industry, prices etc.) will permit him to produce better statistics and to analyse the results more thoroughly.

An interesting attempt to statistically analyse living conditions and the economic decision making process in African rural societies is undertaken by a group called AMIRA created by the "service de cooperation" of the French statistical service INSEE. This inter-disciplinary group includes inter alia statisticians, economists, sociologists, anthropologists. A similar approach is followed by the Development Research Center of the World Bank and its study of living standards measurement. This integrated approach to statistics has its merits with regard to well-defined research projects at a village or rural level and may lead to a better understanding of household statistics at a regional level. The Munich Centre's experience with presenting some of the findings of the group to our African students has been somewhat disappointing in so far as most of them regarded the situations described as "old-fashioned" and due to disappear in the course of modernisation of the society.

VARIOUS APPROACHES TO TRAINING

The purpose and the scope of training being fixed it has to be decided which method to use in order to motivate the statisticians as much as possible. Various approaches have been outlined by Mr. Waast in a meeting called "journées pédagogiques" organized in Paris by the "Centre Europeen de Formation des Statisticiens Economistes" (CESD) from the 7th to the 11th of September 1981. About 60 persons from statistical schools, training centres and organizations fostering the training of statistics participated. Mr. Waast especially referred to four approaches:

- courses on special statistical topics
- courses on statistical methodology
- critical evaluation of surveys
- practical execution of a survey

Courses on special statistical topics are still the most commonly used method of training statisticians. Since 1973 the Munich Centre has organized two courses of 4 months' duration annually, one of which is in English and one in French. The subjects dealt with so far are national accounts, foreign trade statistics, industry and handicraft statistics, demographic statistics, agricultural statistics, household statistics, social statistics. Around 500 statisticians from African and Asian countries having a university degree or a similar training in statistics have participated in the courses.

In their most primitive form courses on special statistical topics have mainly two disadvantages: they concentrate exclusively on the subject-matter under discussion without showing the interrelationship with other statistics and they are in the form of "ex cathedra" lectures leaving little room for discussion and exercises. Attention should, therefore, be given to exercises being included in the course programme and showing the use to be made of the data collected and their place within a coordinated framework of statistics. The lecturer should be instructed to permit questioning at any moment in order to initiate a dialogue with the statisticians attending. This will give them the opportunity to discuss the actual problems of their daily work, to exchange their experiences and to ask for potential solutions. As far as possible the lecturer should aim at using a fair amount of the time available for practical exercises in order to check that the points made were correctly understood. If these items are observed, courses on special statistical topics are a useful tool for training statisticians.

It should, however, be mentioned that much depends on the lecturers available. Experience in statistical work in LDC's is of vital importance in order to give advice on how to adapt conventional concepts to local conditions. The preparation of lecture notes and written exercises beforehand is also of great value.

Courses on statistical methodology have quite a different starting point. A topic of general economic interest such as an agricultural development plan, or the fight against inflation, may be chosen. First the problem has to be analyzed, then the need for data relevant to the subject has to be determined, and finally the best way of obtaining the data needed has to be elaborated. The course may include a discussion on the validity of existing statistical instruments (national accounts, multi-purpose household surveys etc.), data

collection techniques (sample surveys, interviews etc.), their application in a given social environment and the need for developing an information system.

The approach is a rather theoretical one. It is interdisciplinary in the sense that it may embrace all kinds of statistics. Formal lectures are kept to a minimum and emphasis is placed on the comments and proposals made by the participants. There are a number of disadvantages. One, certainly, is that the group should be restricted to keep the discussion within reasonable limits. More important still is the fact that the participants need a rather advanced technical knowledge of the problems under discussion in order to be able to cooperate fully. This approach, therefore, is more suited to rather short seminars for already experienced statisticians.

A third way of training statisticians consists of the critical evaluation of surveys already undertaken. The exercise includes a study of the preparatory documents (objectives of the survey, methodology) and the survey documents (questionnaires, instructions to interviewers, classifications etc.). The method has the advantage of being based on documents used for real work. In order to stimulate the interest of the participants it would be desirable to investigate a series of surveys using different methodologies rather than a single one. It may, however, be difficult to bring together the necessary background material.

The study of documents may be beneficial to students, especially the questionnaires used and the way they are designed. The analysis of other documents - mostly rather lengthy ones - is often not very stimulating and it is difficult to capture the attention of the course participants by going through them. The course tends, therefore, to turn into a series of lectures on the documents presented by the responsible lecturer rather than an analysis made by the students. A distribution of the documents to the participants may, however, be of great value to further reference.

The last approach mentioned concerns the practical execution of a survey. There is no doubt that a small group of statisticians working together in a team to deal with all aspects of a survey is a good way of making them acquainted with the difficulties of applying statistical methods to real situations. The method educates to be imaginative, to compromise and to find solutions to unforeseeable problems. On the other hand it is time-consuming and rather costly. In order to be of some general value the number of cases included in the survey should be sufficiently large and the sampling frame correctly conceived.

While the method has certain advantages it should be noted that its application very often encounters difficulties. A certain infrastructure is indispensable and the social environment must lend itself to surveys. The training of LDC statisticians in European training institutes by applying this method is generally hampered by restrictive regulations concerning surveys, a certain unwillingness to respond to inquiries on a voluntary basis and a growing concern about the confidentiality of individual data. However, the organization of surveys representing undoubtedly a very important aspect of work in a course on applied statistics does not dispense with studying more "theoretical" items such as concepts and definitions or analytical methods.

A useful tool for training statisticians is case studies. The US Bureau of the Census, through its International Statistical Programme Center (ISPC) has developed a set of materials, i.e. "Atlantida: A Case Study in Household Sample Surveys". It is a rather complex study dealing with all phases of a survey programme and emphasizing the practical aspects rather than the theoretical ones of household surveys. More recently the French statistical office INSEE has developed some case studies. The mythical country "Baramu" represents a case study centered on employment, education and health problems. The intention is to enlarge the study to comprise also agricultural statistics, foreign trade and industrial statistics. Another case study "Kangare" shows how to build up national accounts from given data. The case study "Bernard" intends to provide an introduction to business accounting. The French "Centre de Perfectionnement Economique et Financier" (CEPEF) mainly uses two case studies namely "La Camesie" dealing with annual forecasts and "Tercera" centering on the planification of rural development.(*) All case studies try to approach reality as far as possible. Their educational value depends very much on the ability of the teaching staff to present the studies in an interesting way, to guide the discussions of working groups and to give advice on the various problems as the students go along. The teachers should, therefore, be very qualified having also a rather profound experience in LDC's.

A reproach often made to courses on special statistical topics like, for instance, population statistics is that these courses tend to compartmentalize the teaching of statistics. However, the danger of doing this appears to be even greater when evaluating or executing a survey. The latter exercises do not necessarily require a discussion on the use to be made of the data and the interrelationship with other statistics. In all methods used there is the danger that, for instance case studies or evaluations turn into "ex cathedra" lectures rather than being the result of efforts undertaken by the students.

(*) For more detailed information cf. Compte-rendu des journées pédagogiques des écoles de statistique africaines, Paris, 7-11 September 1981. "Centre Européen de Formation des statisticiens économistes des pays en voie de développement", 3 Av. Pierre Larousse, 92240 Malakoff, France.

Whatever the method of training, one should always keep in mind that a statistician participating in a training course is mainly interested in topics which are useful to the work he has to do in his office. The concentration on a well-defined subject, therefore, still appears the best way to stimulate his cooperation. He will also be receptive to learning which use can be made of the statistics he produces and the interrelationship with other statistics. A combination of lectures, discussions and case studies appears to be the best way of conducting a "practical" training course, whereby a lot of room should be devoted to problems brought forward by the students themselves. Whenever the teaching can be combined with the execution of a real survey this opportunity should be used, as the experience in the field is of greater value than any theoretical case study.

MATERIAL AIDS TO TRAINING

The conventional blackboard and the overhed projector are still the most commonly used aids to training. The direct contact with the lecturer, the possibilities of asking questions and of opening a dialogue on unclear points are valuable assets for the participants. On the other hand the response of the audience enables the lecturer to adapt the progress of his lecture to the understanding of the students. It may be that these advantages had a somewhat retardatory influence on the use of video tapes. It may, however, be expected that they will play a more important role in the future due to technical progress making their use relatively cheaper. The availability of suitable programmes is another important precondition.

The preparation of lecture notes and the distribution of an annotated agenda before the lecture is also of great importance to the success of training. The course participants should know at any moment which point of the lecture is discussed and how it relates to what has been said beforehand and to what is still going to follow. Lecture notes should summarize the content of the lecture in concise form. They are very important as reference material for further study and as aid-memoire for the participants, especially if they come from LDC's where relevant literature or experienced discussion partners may not be available.

Microcomputers have recently become very powerful and extremely cheap, so that even smaller and financially poor statistical offices can afford them. The use of microcomputers should, therefore, become a standing feature in any course programme for statisticians. The purpose should not be to convert them into programmers or computer experts, but rather to familiarize them with the use of electronic equipment. This will hopefully contribute to a better understanding between statisticians and EDP technicians and facilitate their

cooperation. In this sense it is not important that the statistician knows a series of commonly used computer languages. Basic training should be centered on supplementing the statistical courses by practical exercises on the computer. The course participant should be given the opportunity to manipulate the microcomputer as often as possible and to learn its use by the "trial and error" method. This will enable him to open an informed dialogue with the EDP technician in case the job on hand requires the use of a big computer.

In this context some general remarks on the use of computers may be in order. At the moment each office or institution of some reputation will try to have its own computer - and the bigger the better. The accounting work has then, of course, to be adapted to the requirements of the computer which in many cases leads to substantial organizational changes and sometimes to rather significant additional paperwork. The progress in technology increasing the capacity of microcomputers enormously now makes it possible to use these devices for jobs of rather substantial dimension. For the time being producers are still more concerned with producing sophisticated hardware leaving the software aspects somewhat aside. But things are changing quickly and more and more software for microcomputers becomes available. Therefore, it is quite legitimate to expect that a series of microcomputers, each one well adapted to specific jobs, may be able to replace a big computer in the foreseeable future. This is especially important to LDC's statistical offices where, for a number of reasons, the use of big computers encounters difficulties.

TEACHING STAFF

The teaching staff for training statistical practitoners could consist of permanent staff members of the institute doing the training, part-time lecturers or a combination of both. All depends on whether the training is continuous or sporadic and whether the training is supply-oriented (repetition of a fixed set of subjects) or demand-oriented (changing subjects on request).

Supply-oriented training institutes organizing courses on a regular basis tend to have permanent teaching staff. As the subjects dealt with are more or less the same in each period each staff member has his responsibilities within an overall training programme. The advantage is that coordination among the staff members is relatively easy. On the other hand the system is costly and budgetary difficulties are likely to arise in connection with requests for new staff. It is therefore more suited to education in statistics rather than to training practitioners.

If the training is demand-oriented, i.e. the subjects of courses are constantly changing it is preferable to employ part-time lecturers. Best suited are experienced statisticians from statistical offices or institutes as well as academic people with experience in practical statistical work. This will ensure that the training has a certain bias towards practical issues which is rather desirable when training practitioners. Experienced statisticians serving as part-time lecturers usually have a profound knowledge of the subject-matter, they know the latest developments in their field and are able to advise on practical issues. The use of part-time lecturers ensures a certain flexibility and is financially less burdensome. On the other hand it is not always easy to put together the "right" team for a course as good statisticians tend to be rather busy. Coordination is also difficult and lecture notes are often a problem.

When training statisticians from LDC's a knowledge of local conditions is indispensable. Commonly used concepts in developed countries are not always transferable to LDC's and advice on problems brought forward by the course participants will hardly be possible without LDC experience.

COST OF TRAINING

The cost of training is an important consideration in any training programme. In general the opinion prevails that training courses organized in LDC's are cheaper than those held in developed countries. This may be true to a certain extent, but expectations in this respect are often higher than real savings. All depends on whether the organizer of a course is able to ensure the co-operation of a local school, institute or university which is prepared to provide the infrastructure needed (classrooms, teaching aids, accommodation, sometimes even part-time lecturers). If this is the case courses in LDC's can indeed be organized at relatively low costs. However, if course participants have to be lodged in hotels, classrooms etc. hired at normal rates, UN per diem rates paid to participants and teachers flown in from abroad, the saving made on lower expenses for airplane tickets for the students may be largely offset.

Most statistical training institutes are subsidized in one way or the other. The expenditures for housing, administrative and teaching staff, materials etc. are often paid from Government aid funds. Costs vary widely depending on the organization of the centres and the permanent staff employed. It is therefore difficult to establish any general rules. However, it can be said that training costs are enormously high if the organizer intends to recuperate all costs by tuition fees.

At present the direct costs for a statistician participating in a 4 months training course at the Munich Centre are approximately as follows:

-	air economy ticket	US $ 2,000.--
-	clothing allowance	US $ 250.--
-	books	US $ 100.--
-	fellowships ($ 500 per month)	US $ 2,000.--
		US $ 4,350.--

EXPERIENCE OF THE MUNICH CENTRE

Since its inception in 1973 the Munich Centre for Advanced Training in Applied Statistics for Developing Countries has organized the following courses:

- 2 four months courses annually in Munich one of which is in English and one in French
- Seminars of 3 weeks duration in Munich for directors of statistical offices or their deputies
- 4 to 6 weeks courses in developing countries

The 4 months courses are meant for statisticians from LDC's having an academic degree in statistics or an equivalent education and some years of experience in practical statistical work. In spite of efforts to keep the group as homogeneous as possible we are not always successful even though attention is paid to arriving at the same educational background. The professional experience of a statistician is decisive for his performance and much depends on the office he is working in. The maximum number of course participants is 23 coming from about 18 different African or Asian countries. Hence, a certain variety of qualification is unavoidable.

More recently a certain change has been noticeable in the composition of the participants. In the early 1970's most participants had received their basic statistical education from European universities. Nowadays practically all participants studied in their home countries. While this may not necessarily affect their theoretical statistical knowledge, a certain decline in language proficiency is to be noted. Hence, it is advisable to request a language test before admission.

Up to now the course programme was always centered on a specific topic of statistics such as industrial statistics, agricultural statistics etc. The concentration on one topic does not mean that it is looked at in isolation. The course duration of four months permits going through the whole process of statistical surveys, from data collection through publication and analysis. Special attention is paid to the use of data, their interrelationship with other statistics and their usefulness for national accounting and planning purposes.

In 1983 a new approach will be tried by dealing with the topic "collection and analysis of data - methodological aspects". It is mainly meant for statisticians having the function of preparing surveys or analysing data from a variety of sources.

The Munich Centre works with part-time lecturers. They are experienced statisticians who are able to teach in English or French thus avoiding interpretation. Most of them have extensive experience in LDC's. Course participants from LDC's have a preference for lecturers from their own countries. Whenever possible the Munich Centre invites LDC lecturers in spite of the sometimes burdensome additional travel cost. The main problem with part-time lecturers lies with the coordination of their interventions and the preparation of lecture notes which sometimes never arrive due to other obligations of the lecturers.

The teaching is informal. Lecturers can be interrupted at any moment for comments or questions. The cooperation of the course participants in communicating their experience is essential to the success of a course and should be encouraged. Practical exercises are highly appreciated by the participants and should form a substantial part of the programme. Lectures in the morning and discussions and practical work in the afternoon may be a workable alternative approach.

Seminars for the directors of statistical offices or their deputies were organized on an irregular basis. The topics were "the organization of a statistical service" and "the setting and implementing of statistical priorities". The duration was three weeks. This is definitely too long and should be shortened to two weeks as it is difficult for the directors to be absent from their office for longer periods. In preparing seminars of this kind the difficulty lies with finding suitable topics of interest to the directors without being too technical. The various items on the agenda should be well prepared by background papers distributed beforehand. Care should be taken to ensure the cooperation of internationally well-known statisticians to prepare the papers and to serve as discussion leaders as this will have a positive impact on participation. The greater part of the seminar should consist of discussions and exchange of experience. A report reflecting the findings should be prepared by the participants.

Courses in developing countries have been carried out by the Munich Centre since 1980. The topics were "Household statistics" (Cameroon), "EDP for statisticians" (Thailand), "Agricultural statistics" (Ivory Coast), and "1983 World Census of industrial statistics" (Ethiopia). In each case the cooperation of an already existing institute or organization could be ensured, i.e. the

"Institut Statistique de la planification et de l'economie appliquee (ISPEA)" in Yaounde, the "Asian Institute of Technology" in Bangkok, the "Ecole statistique d'Abidjan" and the ECA in Addis Ababa. The provision of classrooms, secretarial help, typing and reproduction facilities, accommodation therefore did not represent any special problems. In all cases it was also possible to employ local part-time lecturers. This ideal solution works, however, only during the summer vacation when the regular students of the schools are absent. At other periods the schools work at full capacity and cannot accommodate additional training courses.

LDC's unconditionally support training courses organized in their countries. Administrative procedures concerning nominations appear to be easier and wider publicity is normally given to the courses. It is advisable that the organizer limits participation to a regional level. This ensures a certain homogeneity of the participants and diminishes travel cost. Another important advantage is that practical work can sometimes be organized by conducting a real survey.

PREPARING STATISTICIANS FOR GOVERNMENT SERVICE - A VIEW FROM THE UNITED STATES

Marie D Eldridge, United States Department of Education
and Katherine K Wallman, Council of Professional Associations on Federal Statistics

INTRODUCTION

In 1979 the American Statistical Association (ASA) Section on Statistical Education formed a Committee on Training of Statisticians for Government. The establishment of this Committee was an outgrowth of work by a predecessor group (Committee on Training of Industrial Statisticians 1980) sponsored by the ASA. Under the leadership of the authors of this paper, the Committee on Training of Statisticians for Government completed its assignment and presented its findings on education and training for federal statisticians at the 1981 annual meeting of the American Statistical Association. The full report (Committee on Training of Statisticians for Government 1982) was designed to assist colleges and universities in developing their statistics curricula, to serve as a counselling tool for prospective federal government statisticians as they select their training courses, to provide guidance to managers who recruit and supervise statisticians as they select employees and plan for their career development, and to offer recommendations on qualification standards for federal government statisticians. Material in the present paper is drawn largely from the Report of the Committee on Training of Statisticians for Government; the paper focuses on preservice and in-service education and training for government statisticians, and on recommendations to colleges and universities and the national government.

THE FEDERAL STATISTICIAN: A PROFILE

Perhaps the largest single employer of statisticians in the United States is the federal government. As of October 31, 1978, the federal government employed over 3,900 statisticians, as well as 3,240 statistical assistants and numerous other employees performing statistical duties but classified in different job series (U.S. Civil Service Commission). While the federal government has an extensive need for statisticians with highly specialized capabilities, the majority of the government's statisticians are generalists involved with data collection activities. Even the Bureau of the Census, with its major emphasis on demographic analysis and research in statistical methodology and statistical data processing techniques, has the collection of general-purpose statistical data as its primary purpose.

* Contribution of the U.S. Government, not subject to copyright in the United States.

The Survey Statistician

Although survey statisticians as a group comprise the major segment of federally employed statisticians, the responsibilities of individual employees differ substantially. A junior or midlevel statistician working on a complex survey for an agency such as the Bureau of the Census may perform duties solely related to frame maintenance, sample selection, survey estimation, data adjustment, quality control, or data analysis. In the Department of Defense, by contrast, the entire survey process from design to analysis is likely to be done by one or two individuals.

The size of the agency responsible for a survey is also a key factor in determining how and by whom the survey is designed and the data are collected. In an agency with few statisticians, the survey design is frequently done by an outside contractor and the actual data collection is virtually always done on contract. In these cases, the survey statistician's responsibilities are largely concerned with writing requests for proposals, evaluating contract proposals, and monitoring the performance of contractors. In most of the large general-purpose statistical agencies (Bureau of the Census, Bureau of Labor Statistics, the Department of Agriculture's Statistical Reporting Service, and the National Center for Health Statistics), the design is virtually always done by an internal unit; the data collection is then done by field units (enumerators) across the country.

The nature and extent of the federal statistician's duties are affected by his or her professional level, as well as by the size of the employing agency and the complexity of the survey. At the junior professional level, much of the statistician's time is devoted to on-the-job training in all aspects of his or her office's program. In addition, the junior statistician frequently becomes an expert in some specialty - either in some analytical technique (such as data adjustment), in some procedure (such as contracts or publications), or in some topic (such as an industry).

At the middle professional level, the survey statistician is expected to be able to handle non-routine technical problems. Responsibilities include developing procedures and schedules for data collection and writing operating manuals and procedures. Technical expertise is used to help define concepts or to identify problems and to develop possible solutions. In the smaller agencies, the midlevel statistician may have full responsibility for the entire survey process from planning the design to writing the final report.

The senior level survey statistician's time is generally divided among supervision, administration, and program planning. The beginning supervisor generally is responsible for a unit comprising three to ten people (statisticians, subject

matter specialists, statistical assistants, and secretaries) and administers one or more surveys or programs. In addition to planning the unit's activities, the senior statistician participates in planning the office's overall program and in discussing budget requirements with budget officials. Representing the agency in dealing with personnel at the highest levels in other agencies and in private firms and the training of agency personnel are usually the senior statistician's responsibility.

Like the senior statistician in charge of data collection or validation and the production of estimates, the senior statistical analyst spends a great deal of time in program planning, supervision and administration, but unlike the senior survey statistician he or she may devote a large portion of time to applied research. This research will be either program-oriented research - that is, analysis of the output from a particular statistical program - or topic-oriented research, using a wide variety of data sources and a wider variety of approaches and techniques to analyze a particular topic, such as a comparison of cost-of-living indices, techniques for analysis of quality change, or a study of the economic benefits of education.

Demographers, Biostatisticians and Other Specialists

While the majority of U.S. federal statisticians are involved in general data collection activities, the government also employs many statisticians whose jobs are highly focused in other areas. Such specialists include demographers, biostatisticians, engineering statisticians (i.e., specialists in such areas as quality control and reliability), researchers in statistical methodology, and an increasing number of specialists in computer software for statistical applications.

Demographers in the federal service work on various programs and problems including public health, manpower analysis, family planning programs, immigration and emigration policy, population forecasting, estimation and projection, and other areas. The principal federal employer of biostatisticians is the National Institutes of Health, which is the medical research arm of the Public Health Service. Biostatisticians work in lesser numbers for other Public Health Service agencies, serving principally as epidemiologists. In the National Center for Health Services Research, biostatisticians focus on analyses of health services data, in the Food and Drug Administration on the evaluation of drug safety, and in the National Center for Health Statistics on the analysis of vital and health statistics.

Statisticians with specialized education in assurance sciences and experimental design are also in demand in the government. Quality control, reliability,

availability, and maintainability are all aspects of assurance science used by Department of Defense statisticians in the course of designing, procuring, storing, and using weapons systems and related material.

Scattered throughout the government are statisticians whose primary function is statistical consulting with scientific researchers. Areas of application include the physical and engineering sciences, agricultural sciences, environmental and geophysical sciences, and the biological and medical sciences. These statisticians work in close collaboration with scientific specialists to develop empirical models of phenomena and statistical models for measurement or observation processes.

The federal government also employs a sizeable number of mathematical statisticians in the area of methodological research. Typical areas of research include sample design and estimation, the theory and methods for survey evaluation, data adjustment, the improvement of data quality and survey operation through efficient quality control procedures, the assurance of respondent confidentiality in published data through mathematical disclosure analysis, the computerization of maps, computer editing, the establishment of production standards for survey personnel, and statistical modelling.

A final example of the specialized uses of statisticians in the federal government involves computer programmers. In several agencies, individuals responsible for programming complex sample designs and the associated estimates and variance calculations are classified as statisticians. Such individuals are also regularly called upon to apply a myriad of standard statistical software packages to survey or other statistical data and to develop software for special purposes.

While the job requirements of the specialized mathematical statistician are frequently much more technical in nature than those of the survey statistician described earlier, the level of responsibility of the junior and midlevel mathematical statistician is similar to that of a survey statistician employed at a corresponding grade. The primary difference in the work of the two types of statisticians occurs at the senior level, where the mathematical statistician is still commonly found in a nonsupervisory advisor position with highly specific technical demands and requirements. Almost all survey statisticians at this level, on the other hand, are in management or supervisory positions.

EDUCATIONAL NEEDS OF FEDERAL STATISTICIANS

Because of the heavy emphasis on data collection in the U.S. federal statistical system, it is virtually imperative that the effective federal statistician be well trained in survey techniques. (Exceptions to this rule are those highly

specialized statisticians providing statistical assistance in a scientific research community). Beyond the earliest entry levels - where an individual would likely not need statistical training beyond the fundamentals of statistical theory, inferential statistics, and correlation/regression techniques - knowledge of survey methods is vital to most government statisticians throughout their careers. In most agencies, the lack of a solid background in applied statistical methods and sample survey procedures would hamper the effectiveness of the statistician by the time the mid-professional level is reached. To be fully prepared for the job at this level, the statistician should have had training in the areas of advanced statistical theory, applied sample survey methodology, computer programming, and technical writing.

The advanced statistical theory desired at this level would include mathematical statistics, theory of probability, matrix theory, and sampling theory. The first three courses are standard at most institutions of higher education, usually forming part of the required curriculum for a degree in statistics. Unfortuantely, sampling theory - probably the most important of the four for many government statisticians - is still not offered in some institutions with statistics curricula and is not required in many others. In fact, in some institutions the student receives almost no applied training until after the master's level. In addition, the quality and appropriateness of the sampling theory courses vary widely from institution to institution. Frequently, the courses are taught by persons who themselves have had no experience in designing a survey or specifying a sampling plan. Such courses - which are usually limited to the derivation of endless mathematical formulas for variances of alternative designs - are of marginal value to government statisticians. A considerably more useful sampling theory course would be one that is presented in the context of how one makes decisions among alternative designs, considering both cost and precision. Unfortunately, few U.S. colleges and universities offer sampling theory courses of this type. The problem is so severe that some agencies find it necessary to encourage newly hired statisticians to take an adequate sampling course.

Of particular importance to the junior professional survey statistician is education in applied sample survey methods. Unfortunately, adequate training opportunities in this area are even rarer than are adequate sampling theory courses. Most graduate statistical curricula are so theoretically based that the elementary principles of applied survey methods are not covered at all. Other than in many social science statistics courses, the principles of questionnaire design and the analysis and interpretation of survey data are rarely touched. And with the exception of a possible digression in a sampling theory course, students very rarely delve into the vital area of the estimation

and control of nonsampling errors in survey data. Missing also is adequate training in imputation and other incomplete data techniques. Yet these areas represent the everyday working tools of the federal survey statistician at the junior and middle professional levels.

The final areas of technical education needed by the junior professional statistician are computer programming (particularly FORTRAN and the more common statistical software packages such as SPSS, SAS and BMD) and technical writing. Virtually all statistical curricula now include heavy reliance on the computer as a statistical tool, but technical writing is usually forgotten. As Barbara Bailar stated in her paper describing government training needs:

> The writing of statistical reports seems to be a problem because of poor writing ability. Graduates may never [have] had to write reports, especially if [their] training was primarily in mathematics. Frequently one is told that one reason graduates majored in mathematics was to avoid writing, which they say they dislike. Unfortunately, it is of little use to be able to analyze data beautifully but not be able to communicate the results of that analysis.

The need for writing skills cannot be overemphasized. As statisticians move into managerial positions, communication - both written and oral - becomes a vital ingredient of a successful performance. As a manager, the statistician must frequently communicate important, but highly complex, technical details to senior officials with no technical training. Inability to communicate effectively in such circumstances could lead to important decisions being made on an incorrect or incomplete understanding of available information.

A possible solution to this problem would be for colleges and universities to require their mathematics and statistics majors to take an applied course in which historical data are analyzed (or applied to a mathematical model, etc.) and a technical report of the study is required. This would give the student at least a nominal exposure to technical writing and applied techniques to counter in part the current imbalance of theory over practice.

By the time a federal survey statistician rises to the top of the mid-professional level, he or she is likely to be responsible for an entire survey or a segment of a major survey. It is at this time that consulting and analytical skills become essential as well. Statisticians at this level should obtain additional technical education in the areas of statistical analysis methods such as non-parametric methods, stochastic processes, time series analysis, and multivariate analysis techniques. In addition, in-service training in introductory principles of supervision is essential, as statisticians at this level are frequently team

leaders, if not formally supervisors.

Finally, senior level general or survey statisticians must have a broad background in supervisory education. While their earlier technical education will still be extremely important to them, the job demands at this level turn increasingly to management.

In 1979, the Statistical Policy Coordination Committee (SPCC), established to advise and assist the President in the improvement, development, and coordination of federal statistical services, appointed a Task Force on Civil Service Reform to make recommendations on policies and activities which would contribute to ensuring the integrity and improving the performance of the federal statistical system. One of the areas considered by the Task Force was the opportunity which executive development programs offer for strengthening the skills and effectiveness of senior statistical personnel in the federal government. The 1981 Report of the SPCC Task Force on Civil Service Reform concluded, among other things, that:

1. Senior statisticians tend to develop their experience in the technical aspects of their positions, and often have been relatively isolated from broader policy and management activities such as interface with the legislative and budget processes, and

2. In many instances, senior statistical personnel in the federal service have spent most, if not all, of their professional careers within a single agency or narrow field of specialization.

As a consequence, the Task Force recommended that executive development programs for senior statisticians in the federal service focus on the acquisition of managerial skills and provide an opportunity for participants to gain a broader understanding of the policies, processes, and programs of the federal government. Types of development needed to improve the government statisticians' effectiveness as senior managers included, for example, education in communication skills, personnel management, and coordination with a wide variety of outside groups and agencies.

Acquisition of the managerial skills needed by the senior government statisticians can be undertaken through various means ranging from formal classroom instruction to informal briefings or discussion groups. Among the options suggested by the Task Force were formal academic training, assignments to another part of the agency (such as the budget or legislative division) or to a different agency for a period of two to three months, intensive (two

to three week) seminars, short (one to five day) workshops, information visits or briefings, attendance at management or other policy formulation meetings, and independent study or discussion groups. The specific mix of executive development activities the statistician should pursue will depend, of course, on the individual's background and interest in assuming managerial responsibilities in the government statistical system.

A summary of the desired education for government statisticians is contained in the following table. (For additional views on this topic, see papers published by Fellegi, Malinvaud and Murthy 1978). In some specialized areas, additional courses in a particular subject may be required, or certain statistical analysis courses may be needed earlier in the employee's career.

IN-SERVICE EDUCATION AND TRAINING OPPORTUNITIES FOR FEDERAL STATISTICIANS

To meet the specialized requirements for education and training of federal statisticians, and to fill existing gaps in their educational experience, the U.S. statistical agencies have relied heavily on in-service training. Apart from the natural "on-the-job" training that derives from association with other statisticians, most federal statistical agencies make a wide variety of education and training opportunities available to their employees.

The most frequent method of in-service training used is agency support of the statistical training offered by local universities. Although many universities now offer degree programs at night to attract full-time workers - particularly in areas with large concentrations of federal workers - there are still few direct arrangements between those institutions and the federal statistical agencies. Many universities educate their doctoral students only for academic careers. Fortunately, some recent activities seem to be changing this picture.

One such program, called the Statistical Survey Institute, was established jointly by the American Statistical Association and the Statistical Reporting Service (SRS) of the Department of Agriculture in 1976. The objective of this program is to enhance graduate education and research in statistical theory by tying instruction in these areas to data collection and analysis at the SRS. A second program is a joint program of the American Statistical Association and the Bureau of the Census, partly funded by the National Science Foundation. The objective is to develop a program of research applied to the social sciences and to supplement the experience of senior social scientists and the education of graduate students in statistics, economics, demography, and related areas by exposing them to the real statistical problems encountered in a large data collection agency.

Suggested Technical Education and Training for Federal Statisticians

Level	Courses*
Entry level	Theory of Statistics Statistical Inference Correlation and Regression Methods Elementary Scaling Principles Graphical Presentation of Data
Junior Professional	Sampling Theory Questionnaire Design Principles Mathematical Statistics Theory of Probability Analysis and Interpretation of Survey Data Estimation and Control of Nonsampling Errors in Survey Data Imputation/Incomplete Data Techniques Matrix Algebra Fortran/Cobol, SAS/SPSS/BMD, Data Base Management Language Technical Writing Basic Government Contract Procedures
Mid-level Professional	Consulting Principles and Practice Nonparametric Methods Stochastic Processes Time Series Analysis Multivariate Analysis Techniques Advanced Scaling Principles
Senior Professional/Manager	Introduction to Supervision Personnel Management, Labour Relations, Equal Employment Opportunity and Related Administrative Areas Effective Communications for Managers Government Budget, Legislative, Regulatory and Procurement Procedures Relationships with Other Governmental Levels, Public Interest Groups, the Media, and the Private Sector

* List is not intended to be in priority or educational sequence order.

Some agencies have developed long-term education agreements with university departments. Under these programs, selected employees are enrolled in the regular graduate program of the institution. In most cases a master's degree is earned by the trainee, although the actual receipt of a degree is secondary (from the federal government's point of view) to the value of the training received.

A major segment of the in-service training that is available to federal statisticians is provided by the agencies themselves. Perhaps the best known of such programs is that of the U.S. Department of Agriculture (USDA) Graduate School. Two of the courses offered through this program - Advanced Agricultural Statistics and Sample Survey Methods - were specifically designed and developed to fill training voids encountered by agricultural statisticians at the entry level.

RECOMMENDATIONS

The preceding sections of this paper describe the environment in which the U.S. government statistician works, the education and training needed in order to work effectively in that environment, and in-service training opportunities for federal statisticians. In the federal government, the well-educated statistician will be one who has been exposed to a balance of theory and practical applications. The statistician who can bring appropriate statistical theory to bear on various methodological and practical problems, and translate this knowledge into terms understandable by the program manager and the policy-maker, will have the greatest impact in terms of bringing statistical expertise to bear on government programs and policies. While there are isolated needs for pure theorists, statistical applications remain the primary tool of the government statistician. For this reason, tools such as computer programming, applied sampling theory, data systems design, incomplete data procedures, technical writing and technical consulting are of primary importance.

In view of its findings with respect to the nature of government statistical work in the United States, the Committee on Training of Statisticians for Government recommended that:

1. Institutions with baccalaureate and master's level programs in statistics should offer a one- or two-semester course in sampling theory. While such a course should be heavily steeped in theory, it should also stress the realities of choosing a design from among several alternative ones, considering both cost and precision. In addition, the course should be taught by a person who has had actual experience in designing a reasonably complex survey or sampling plan. If there is not such a person on its faculty, the institution should explore various means of gaining assistance in

recruiting competent professionals. The institution should, at a minimum, include several lectures given by such a person in its sampling theory course.

2. Institutions with baccalaureate and master's programs in statistics should offer a one- or two-semester course in statistical applications, if they do not already do so. Such a course should include education in design of experiments, quality control and reliability methods, and other specialized statistical procedures, stressing their application to actual data.

3. Baccalaureate and master's programs in statistics should require education in missing/incomplete data techniques. These techniques could be covered in a variety of courses, such as design of experiments, time series analysis, and multivariate analysis. While most graduates have had courses in inferential or analytical techniques, they have had little or no experience in applying these techniques to problems that actually occur. It is rare that one is presented with a data base that is complete and without error. Imputation and validation techniques in an applied setting should be stressed in such a course. Further recommendations on methods for achieving this objective are contained in the report "Preparing Statisticians for Careers in Industry".

4. Institutions should require all majors in statistics to have demonstrated capabilities in the areas of computer programming and statistical software packages. A statistician without such capabilities is not prepared to assume most statistical positions in the labor market - either in government or in private industry.

5. All statistical majors (undergraduate or graduate) should be required to take a one-semester course in technical writing. The most frequent complaint that is levelled against US government statisticians is that they cannot write adequately. Years of hard work on very difficult projects can go for naught if the results and implications cannot be conveyed clearly to senior officials with no technical background.

6. All programs in statistics (undergraduate or graduate) should include at least one course in which the student is required to consult with subject-matter specialists. As with technical writing, too often excellent talent is not properly utilized because the

individual has weak communication skills. The statistical major must be able to translate the subject-matter specialist's needs into statistical specifications, conduct the work, and then translate the results and implications back into the specialist's terms.

It should be noted that many of the foregoing suggestions are being implemented in various education programs. By the same token, it is recognized that not every university department could or should offer programs designed specifically for prospective government statisticians.

To suggest that the responsibility for better preparing the statistician for career opportunities in government statistical service is solely the responsibility of the university clearly would be inappropriate. Improved communication and co-operation between the institutions responsible for training and the prospective employers (in this case, the government agencies) must be achieved. It is hoped that the recommendations in this paper will serve as a useful step in promoting an ongoing exchange between academic and government officials.

The need to introduce the university student to "real-life" problems encountered by government statisticians has been a recurring theme. Methods to accomplish this objective were a focus of the American Statistical Association Conference on Transfer of Methodology between Academic and Government Statisticians held in 1978 (see conference report). Although the scope of the ASA conference was considerably broader than the education of statisticians, the conference's recommendations to government statistical offices are consonant with findings of the Committee on Training of Statisticians for Government and were presented as recommendations to the federal government agencies:

1. Government agencies should foster opportunities for collaboration with academic statisticians through such procedures as the employment of consultants, temporary exchanges (see Savage, 1978) of personnel with academic institutions, and study leaves for senior and junior staff members.

2. Government statistical offices should develop materials to describe vexing problems with which they are faced. Government statistical offices should also develop data packages that simulate practical governmental applications, such as case studies. These materials can be used as training aids in universities.

3. Government statistical offices should develop procedures for increasing access of university researchers to government statistics so that appropriate research can be undertaken.

While the report of the Committee on Training of Statisticians for Government, and those that have been prepared earlier, are intended to communicate needs between and among the several involved institutions, their full value will not be realized without an organized follow-through effort. The Committee on Training of Statisticians for Government endorsed the recommendations of its predecessor Committee on Training of Statisticians for Industry, that is, that "The statistical societies and the universities should create programs to further address these issues," and of the Conference on Transfer of Methodology Between Academic and Government Statisticians, that is, that "national associations should seek to encourage the development of data sources which can be used in the university curricula to give experience in practical statistical education" and that "national associations should develop a compilation of existing exchange opportunities and arrange for distribution throughout the profession." The consonance of findings from the independent assessments of these groups suggests that implementation of recommendations by the academic institutions, the employers, and the associations is appropriate.

References

[1] Bailar, Barbara A. (1978) Training Statisticians for the Government, United States Department of Commerce, Bureau of the Census.

[2] Committee on Training of Statisticians for Government (1982) Preparing Statisticians for Careers in the Federal Government: Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Government, The American Statistician, 36, 69-81.

[3] Committee on Training of Statisticians for Industry (1980) Preparing Statisticians for Careers in Industry: Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Industry, The American Statistician, 34, 65-75.

[4] Fellegi, Ivan (1978) Some Thoughts on Training Needed for Government Statistical Work, in: Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 43-52.

[5] Malinvaud, Edmond (1978) University Teaching for Future Official Statisticians, in: Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 53-67.

[6] Murthy, M.N. (1978) Statistical Education, Training, Research and Their Utilization, in: Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 68-93.

[7] Savage, I. Richard (1978) Exchanges of Statisticians, in: Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 116-129.

[8] Task Force on Civil Service Reform, Statistical Policy Coordination Committee (1981) Executive Development for Statisticians Under the Civil Service Reform Act, United States Department of Commerce, Office of Federal Statistical Policy and Standards.

[9] U.S. Civil Service Commission (1978) Occupations of Federal White-Collar Workers.

PROGRAMMES D'ENSEIGNEMENT ET COUTS DE FORMATION

M Benyaklef, INSEA, Rabat

MISE AU POINT D'UN PROGRAMME D'ENSEIGNEMENT

La mise au point de programmes d'enseignement n'est banale que dans le cas où il s'agit d'apprendre à l'étudiant comment utiliser un ensemble de techniques routinières en vue d'exécuter une tâche assez simplement définissable (certains "recyclages" dans le cadre professionnel tombent dans ce cas).

Dès qu'il s'agit d'objectif complexe, comme la formation d'un "Statisticien-planificateur", les choses ne vont pas d'elles-mêmes.

D'abord une tentative de description du "profil" du cadre à former s'impose. Ensuite un corpus de connaissances et de compétences sera dégagé puis une confection sous forme d'unités pédagogiques en résultera. Enfin un système de "Feedback" permettra des réajustements et des adaptations périodiques des programmes. Ce dernier point est fondamental puisqu'aucun profil ni moyen d'y parvenir ne sont immuables.

Objectif de la formation

La première étape est la plus délicate. Il y a lieu de faire une description des tâches que le lauréat du système devra pouvoir exécuter. Cette description sera aussi détaillée que possible. Par example "calculer des indices statistiques" sera précisé selon le cas par exécuter les calculs des indices classiques" ou par "confection et calculs d'indices de prix".

Toujours dans cette même étape, il y a lieu de décrire l'organisation du travail en définissant le rôle de l'individu au sein du groupe : Comment l'apport de l'individu s'intégre aux autres pour produire les résultats recherchés.

La brièveté de ce que nous en disons n'est pas en rapport avec la complexité de cette première étape.

Corpus des connaissances requises

La deuxième étape consiste à passer en revue les outils nécessaires à l'exécution des tâches voulues. Ceci amène de manière récurrente à évoquer des connaissances prérequises à tel ou tel concept, jusqu'à réduction à un seuil minimal. Ce dernier est le "niveau d'entrée". Deux points sont à noter : Primo, l'algorithme qui dégagera les techniques spécifiquement orientées vers les besoins exprimés ne devra pas nous faire oublier la "formation générale" qui donne une dimension humaine et une ouverture indispensables; Secondo, entraîner quelqu'un à

apprendre et rechercher par lui-même est un investissement précieux pour garantir que sa formation ne s'arrêtera pas avec l'Ecole.

L'exemple de la formation d'un statisticien est édifiant à ce sujet. En effet, étant appelé à collaborer aux questions les plus diverses, la connaissance statistique seule peut être un handicap pour le statisticien.

A titre d'illustration, supposons qu'il s'agisse de l'estimation statistique dans les modèles économétriques. On dégagera d'abord deux composantes : les outils statistiques et les rouages économiques. La première composante se ramènera essentiellement à la connaissance des modèles linéaires et de ceux qui s'y rattachent. L'inférence statistique nécessitée exigera un ensemble de techniques sur la méthode des moindres carrés. Il s'en suit qu'une certaine étude des espaces vectoriels et du calcul matriciel constitue un préalable. Le niveau d'entreé se réduit alors à un bagage modeste. Est-ce tout? Il est clair qu'une autre dimension est souhaitable pour l'étudiant, à savoir celle de l'appréciation de la portée d'un modèle quel qu'il soit, de la vulnérabilité des conclusions face aux hypothèses avancées (linéarité, normalité de la distribution des erreurs, etc...). D'ailleurs la seconde composante, si elle se base sur certaines notions qu'il est indispensable d'acquérir d'abord, il n'en reste pas moins que l'habilité à "la fouille intellectuelle" est une qualité à cultiver.

Toujours dans cette deuxième étape, il s'agit de composer un programme d'enseignement à partir du corpus ainsi défini : arrêter une liste de cours ou d'activités pédagogiques ou de recherche, leur durée et leur agencement ainsi que les méthodes d'enseignement et d'entraînement. D'autre part la question du contrôle des connaissances se pose : est-ce indispensable? Si oui, sous quelle forme est-il le plus efficace? Il n'existe pas de réponse valable pour tous les types d'activité; Cependant une certaine forme d'examen est utile à moins que l'individu ne soit hautement motivé. Il est en effet bon de faire précéder le vrai examen, à savoir la performance dans la profession, par des tests qui ont le mérite d'être significatifs malgré quelques défauts inhérents à tout contrôle.

Feedback et adaptation du système

La troisième étape consiste à mettre au point un système d'évaluation du processus de formation. A cette fin, on peut concevoir un ensemble de questionnaires périodiques à l'intention de chacune des parties suivantes :

a) les étudiants,
b) les enseignants,
c) les employeurs,
d) les lauréats du système.

Cette dernière catégorie devrait émettre des avis assez objectifs pourvu qu'elle porte suffisamment d'intérêt à l'institution.

Les étudiants sont préoccupés d'abord par la traversée du système et n'ont pas en général la vision globale pour parler de réadaptation de programmes. Toutefois, leurs avis sont utiles pour certains aspects de la question. D'autres aspects sont couverts par les avis des enseignants, mais ces derniers ont la tendance naturelle à mettre l'accent théorique sur la part qui revient à leur spécialité. L'équilibre peut être trouvé en contrebalançant avec l'avis des employeurs.

Une confrontation minutieusement préparée des parties en présence, devrait conduire aux évaluations les plus profitables.

Il y a lieu de signaler un aspect important des conclusions à tirer de l'ensemble des avis : un temps de battement est nécessaire avant toute modification du système, le module à retenir devra être au moins égal à la période de formation d'un lauréat. Trop de changements, trop fréquemment rendent méconnaissable le système ce qui détruit les fruits de la chaîne d'adaptation.

Remarque: Quels que soient les programmes dégagés et la qualité de la formation dispensée, il est nécessaire de donner l'occasion aux lauréats de mettre à jour leurs connaissances par l'organisation de séminaires de courte durée sur des sujets spécifiques. Certaines techniques évoluent vite et se prêtent naturellement à une amélioration constante. Si le lauréat ne se maintient pas au courant de cette évolution, il risque d'appliquer des méthodes insuffisante et dépassées.

Les thèmes des séminaires à offrir apparaîtront lors du dépouillement des questionnaires adressés aux lauréats, aux employeurs et aux enseignants.

Ce survol en termes généraux de la problématique de la mise au point de programmes d'enseignement cache d'innombrables questions de détails qui surgissent dans la pratique d'un cas concret.

CAS DE L'I.N.S.E.A.

Une des missions de l'INSEA est la formation des Cadres dans les domaines statistico-économique (nous ne parlerons pas dans cet article du département informatique). L'objectif est de disposer de personnes capables de maîtriser les techniques liées à la planification, en vue du développement national.

La production statistique étant un support fondamental à la planification, la qualité des données et leur manipulation exigent un personnel bien formé.

Les cycles d'études à l'INSEA couvrent trois niveaux de formation correspondants à 3 étapes de l'activité statistique : la collecte et la mise en ordre des données, leurs analyses et leurs synthèses, enfin leur dépassement et la conception de nouvelles approches plus spécialisées et plus fines.

D'après la terminologie adoptée à l'INSEA, le premier niveau est celui des Adjoints Techniques, le second est celui des Ingénieurs d'Application, le troisieme est celui des Ingenieurs Statisticiens-Economistes.

Premiere niveau de formation

Un programme de formation au premier niveau devrait préparer l'étudiant à la pratique statistique concréte. Outre des cours sur l'élaboration statistique et sur la statistique descriptive, il est bon d'inclure des spécificités propres à la Statistique Agricole, Industrielle, Commerciale, etc....

D'autre part, le calcul numérique, le traitement informatique des données ainsi que les techniques de représentation graphique reçoivent beaucoup d'attention.

Afin d'élargir l'horizon de l'étudiant, celui-ci est exposé aux idées générales en Economie, en Socio-Démographie, etc... Un esprit de rigueur et de discipline intellectuelle est maintenu chez lui par des compléments de mathématiques.

Enfin, on ne perd pas de vue que la "connaissance du milieu" où le lauréat exercera constitue une partie intégrante du bagage de l'étudiant.

Ingénieurs d'Application

L'Ingénieur d'Application joue un rôle fondamental dans l'activité statistique, sa formation en trois ans est telle que les deux premières années sont consacrées à l'apprentissage théorique classique quant à la troisième année, elle est essentiellement orientée vers les applications avec une certaine latitude dans les choix de l'étudiant. Les programmes d'enseignement actuels viennent du souci de donner "certaines dimensions" aux lauréats du cycle : l'étudiant est exposé à diverses techniques scientifiques avec une bonne part de démonstration et de rigueur afin de garantir dans les applications, un respect des hypothèses avancées et une évaluation correcte de la portée des résultats utilisés. Cette dimension technique ne s'acquière pas au détriment d'une certaine dimension humaine. En effet, puisque notre statisticien est appelé à travailler étroitement avec des hommes parfois peu versés à la Statistique et puisque le chiffre n'a de vrai sens que dans un environnement bien compris, une part non négligeable de l'enseignement est consacrée à l'Economie spécifiquement marocaine, à la Sociologie, à l'Organisation administrative, etc...

Revenons aux cours proprement dits : l'introduction d'un cours de Statistique-Mathématique (estimations et tests) est précédée par des notions assez développées en statistique descriptive et en calcul des probabilités. D'autre part l'enseignement des mathématiques (Algébre linéaire et Analyse classique) a le double but de fournir les outils indispensables à la statistique et de consolider la formation d'un esprit scientifique. Par ailleurs, l'économie est attaquée d'abord par son aspect non quantitatif (économie politique) puis l'analyse économique (micro-économie, macroéconomie, économétrie) vient orienter l'esprit dans la direction quantitativiste.

Un cours sur les mathématiques de l'optimisation est destiné à unifier des questions rencontrees par l'étudiant dans diverses disciplines. D'autre part l'importance du traitement informatique des données appelle une initiation à l'ordinateur si bien qu'à la sortie du cycle d'études l'étudiant posséde les rudiments indispensables. La dernière année d'étude laisse une grande place à la planification économique et à d'autres disciplines satellites. Le dernier trimestre est consacré à des séminaires au choix allant de l'épistémologie..., aux Finances Publiques.

Le goût à la recherche bibliographique est favorisé chez l'étudiant par sa collaboration à l'étude de thèmes d'un certain intérêt. Par ailleurs, un stage pratique de quelques semaines permet à l'étudiant d'avoir un contact vivant avec le secteur professionnel.

Statisticiens-Economistes

Nous abordons enfin le troisième cycle d'études, celui des Ingénieurs Statisticiens-Economistes.

Les techniques statistiques sont de plus en plus utilisées dans divers domaines, ceci pose le problème de savoir s'il est plus efficace de spécialiser le statisticien dans un secteur donné ou d'enseigner les outils statistiques appropriés aux chercheurs qui en ressentent le besoin (par exemple est-il plus profitable d'apprendre l'économie à un statisticien ou la statistique à un économiste en vue d'en faire des économètres?).

La réponse à la question générale n'est pas unique, elle dépend du type de discipline concerné.

Au niveau du 3e cycle, il est possible d'ouvrir la porte à des universitaires de formations scientifiques diverses pour leur faire un enseignement statistico-économique selon un système souple adapté a différents besoins.

La première année est consacrée à l'homogénéisation des niveaux et à l'étude systématique de la Statistique, de l'Analyse Economique et des matières alliées. A partir de la deuxième année, l'étudiant peut se spécialiser dans le domaine

de son choix (Analyse des données, Théorie et Pratique des Sondages, Econométrie, Planification, Statistique et Economie des Ressources Humaines, etc...).

Un travail de recherche personnelle est demandé à l'étudiant pour s'assurer de son habilité à mettre en oeuvre les connaissances qu'il a acquises.

COÛT PAR ÉTUDIANT

La détermination du coût de formation est intéressante pour au moins deux raisons : toute production a un coût qu'il est important de connaître (la "production de lauréats Statisticiens" ne fait pas exception), et toute dépense est à confronter à d'autres alternatives d'allocation des ressources.

Le calcul du coût de formation d'un Statisticien constitue un exercice compliqué. Il en est de même lorsqu'on veut déterminer la "valeur" de toute personne ayant acquis une certaine qualification professionnelle.

Va-t-on comptabiliser les seuls frais nécessités par l'acquisition de cette qualification ou va-t-on inclure les dépenses encourues avant l'accés de l'individu au Centre de Formation? (dont par exemple le prix de l'acquisition d'un enseignement primaire et secondaire) pourquoi pas remonter même plus en arrière? Le "manque à gagner" qui se dégage du fait que l'individu se consacre aux études plutôt qu'au travail rémunéré n'est-il pas à prendre en considération?

L'autonomie nationale que l'on dérive d'une formation sur place, quel qu'en soit le coût, n'a-t-elle pas une valeur à integrer dans la fonction d'utilité du lauréat?

Il devient donc clair qu'un calcul du coût de formation a besoin d'être canalisé, il doit être entrepris avec des hypothèses de simplification et doit viser un objectif précis. Nous allons nous restreindre donc à la formulation suivante :

(1) A combien se chiffrent l'édification et le fonctionnement du Centre de Formation?
(2) A quelle charge supplémentaire le budget de l'Etat sera-t-il soumis annuellement lorsque l'on décide d'ériger un Centre de Formation?
(3) Le lauréat du système à mettre en place revient-il à combien dans ces conditions?

Une fois que la réponse à la question (1) est donnée, il suffit de retrancher l'aide extérieure éventuelle et d'ajouter ou retrancher les frais de bourse ou les droits d'inscription (ou recettes diverses) selon le cas pour obtenir la réponse à la 2e question. La réponse à la 3e question n'est pas destinée, selon notre optique, à calculer les droits dont l'étudiant devra s'acquitter pour suivre ses études. Elle ne tient pas compte non plus des frais d'hébergement, de nourriture, etc... Le chiffre obtenu traduit plutôt l'effort financier

nécessaire à la production d'un lauréat, étant entendu qu'il ne s'agit là que de la valeur ajoutée par le Centre de Formation.

Voyons comment nous pourrions répondre à cette question précise :

On pourra d'abord comptabiliser le coût de l'étude du projet d'implantation du centre. Soit I_1 l'amortissement annuel de ce coût. Les calculs étant faits sur la base d'un taux et d'une période d'amortissement correspondants aux conditions qui prévalent à l'endroit où le Centre sera implanté.

Soit I_2 le loyer annuel ou l'amortissement des coûts de construction des bâtiments qui abritent le Centre.

Soit I_3 le loyer annuel ou l'amortissement des frais de premier établissement, du mobilier et de l'équipement divers (par exemple location ou achat d'ordinateur, frais d'installation et d'abonnement au téléphone, etc...).

La somme $I_1 + I_2 + I_3$ sera désignée par I. Elle représente la charge constante annuelle de l'Investissement initial correspondant à la conception, à la construction et à l'équipement du Centre de Formation. Cette approche suppose qu'il n'y a pas de problème de financement initial, le crédit étant supposé disponible.

Tout projet solidement étudié et documenté et dont l'utilité est manifeste ne saurait rencontrer de difficultés de financement. En fait le plus souvent, des bâtiments existent déjà pour abriter provisoirement le Centre naissant et le problème se trouve momentanément résolu.

Nous passons aux charges de la composante humaine du Centre de Formation, à savoir les Administrateurs et les Enseignants. L'étude ayant dégagé des besoins précis en hommes, qui peuvent varier d'année en année, soit A_t l'ensemble des salaires distribués au personnel administratif à l'année t. Le coût A_t sera en général réparti uniformément sur l'ensemble des étudiants du Centre à moins qu'il y ait des raisons pour qu'il en soit autrement.

Par exemple, si le Centre reçoit plusieurs catégories d'étudiants, certains d'entre-eux pourraient requérir des services administratifs particuliers. Un ajustement devrait être fait en conséquence.

D'autre part, le personnel pourrait administrer la recherche parallélement à la formation, il ne serait pas alors correct d'imputer les salaires aux besoins de l'enseignement seul. Un coefficient qui estime la part consacrée à la formation (et la part consacrée à la recherche) devrait être introduit. D'une manière générale pour ne pas avoir à nous répéter lorsque nous passerons aux charges d'enseignants, du matériel, etc..., (c'était déjà le cas pour l'investissement initial I), nous supposerons que si un centre a une double mission de formation et de recherche ou si plus généralement diverses missions lui incombent, alors un coefficient relatant la part de la formation dans l'ensemble

des activités du Centre devra être fixé. Ce coefficient sera appliqué à chaque composante des coûts engendrés par l'édification et le fonctionnement du Centre dans sa globalité. Pour ce qui est des recettes éventuelles, leur imputation est claire. En fait chaque fois que la dissociation de la partie - "Formation" peut être faite, l'imputation devient immédiate.

Pour ce qui est de la rémunération des enseignants, l'enseignement s'effectue d'abord par un certain nombre d'heures passées avec les étudiants. Que l'enseignant soit rattaché de manière permanente au Centre, ou qu'il vienne simplement enseigner à titre de "vacataire" ou "consultant", il est possible de déterminer le coût horaire de l'enseignement dispensé étant entendu qu'il est tenu compte des rémunérations éventuelles des services annexes.

Une dernière composante du coût de production est celle de la consommation des matériels (non inclus dans I) et des frais généraux de fonctionnement (à l'exception de ceux afférents aux ressources humaines). Soit F_t la charge annuelle de fonctionnement.

Nous avons adopté l'année comme base du calcul mais il est clair que la méthode reste valable pour déterminer les charges d'un séminaire ou d'une session de recyclage qui ne s'étale que sur quelques mois ou quelques semaines. L'indice t devra être regardé plutôt comme unité de mesure quelconque du temps.

Considérons alors un programme de formation s'étalant sur une période t. Le coût de ce programme est :

$P_t = I + A_t + E_t + F_t$ où I, A_t et F_t sont définis ci-dessus et E_t représente la rémunération des enseignements inscrits dans le programme.

Désignons par N_t le nombre d'étudiants ayant suivi le programme de formation au cours de la période. Le coût par étudiant est donc :

$$C_t = \frac{P_t}{N_t}$$

Si un contrôle des connaissances est institué à la fin de la période de formation et si un nombre n_t seulement ($n_t < N_t$) est déclaré "lauréat", le coût par lauréat est donc :

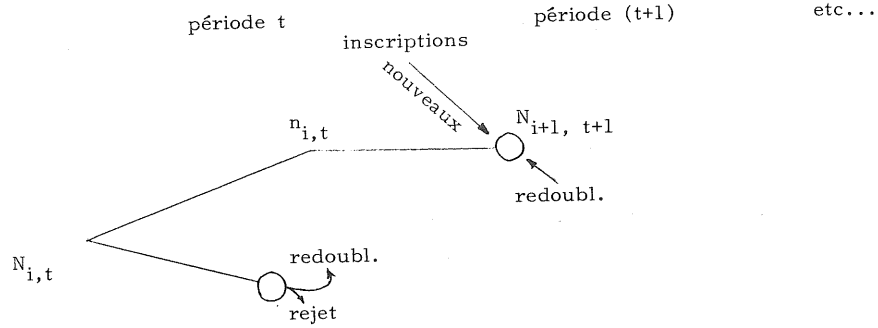
$c_t = \frac{P_t}{n_t}$. Le rapport des coûts $\frac{C_t}{c_t} = \frac{n_t}{N_t}$ exprime le rendement du programme.

FILTRAGE ET RENDEMENT DU SYSTÈME

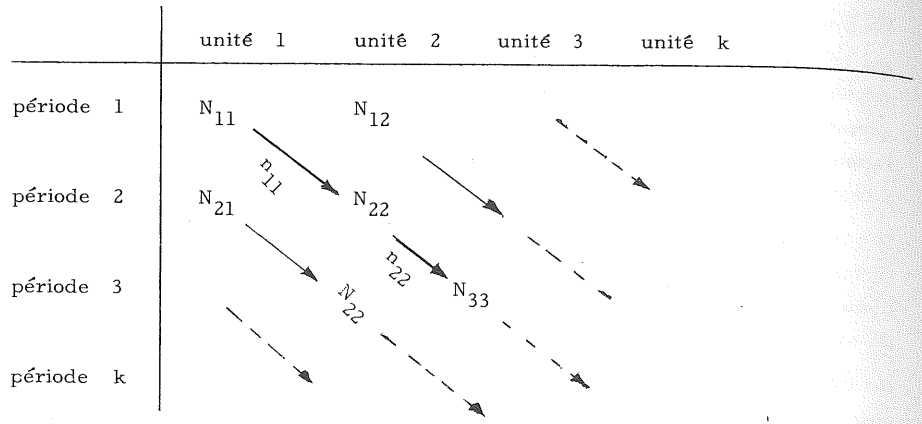
Un programme de formation couvre en général un corpus de connaissances qui se subdivise en unités successives telles que l'inscription à chaque unité requière la réussite à l'unité précédente (exemple des examens trimestriels, semestriels ou annuels). Le contrôle des connaissances ainsi que d'autres facteurs font que le nombre n de lauréats du système est en général inférieur a l'effectif initial N d'étudiants.

Il y a un processus de filtrage à travers le système. A chaque étape, un certain nombre d'étudiants se trouve rejeté du programme prévu par suite d'abandon, d'échec ou de redoublement. Soit N_i l'effectif initial du i^{em} chainon du système (i^{em} unité) et soit n_i l'effectif final ayant réussi cette unité.

On utilisera un deuxième indice pour désigner la période où ces effectifs sont observés et l'on usera des notations $N_{i,t}$ et $n_{i,t}$. En général on a $N_{i,t} > n_{i,t}$ et $N_{i,t} > n_{i-1,t-1}$. La différence $R_{i,t} = N_{i,t} - n_{i,t}$ est constituée par les abandons, exclusions et redoublements au cours et à la fin de la période t. Quant à la différence $\varepsilon_{i,t} = N_{i,t} - n_{i-1,\,t-1}$ elle est constituée par les redoublements de l'unité i et par les inscriptions éventuelles d'étudiants extérieurs au système et le rejoignant en cours de route. Voici un schéma :



Le schéma des transitions est le suivant



Le rendement du système se dégage à partir des taux de succès. Le vecteur de ces taux étant $(s_{11}, s_{22}, \ldots, s_{kk})$ où $s_{it} = \frac{n_{it}}{N_{it}}$.

Remarquer que les taux utilisés correspondent à une cohorte ou promotion donneé. C'est-à-dire que l'on suit l'évolution d'un groupe d'étudiants inscrits à une année donnée que l'on prend pour origine.

Le rendement peut être appréhendé par le produit $\Pi = s_{11} \times s_{22} \times \ldots s_{kk}$ qui n'est pas très différent du rapport des lauréats au nombre d'inscrits à l'année initiale soit : $\tau = \frac{n_{kk}}{N_{11}} = (\tau = \Pi \text{ si } \varepsilon_{ii} = 0)$

Si l'on veut saisir l'importance des redoublements et des rejets du système, on peut déterminer la durée moyenne $\bar{k}$ nécessaire pour produire un lauréat. La durée normale étant de k, soit τ la proportion des étudiants qui terminent leur cycle de formation après k périodes, soit τ_1 la proportion de ceux qui terminent après (k+2) périodes etc..., jusqu'à τ_m. m étant le nombre maximal de périodes permis en supplément de la durée normale du cycle.

Notons par $\tau^* = 1 - \tau - \sum_{\alpha=1}^{m} \tau_\alpha$ la proportion de ceux qui quittent le système sans jamais obtenir leur diplôme. Nous proposons pour durée moyenne :

$\bar{k} = \tau.k + \sum_{\alpha=1}^{m} \tau_\alpha.(k+\alpha) + \tau^*. 2k$ où nous avons choisi arbitrairement le double de la durée du cycle pour tenir compte des abandons et rejets.

On peut également appréhender le rendement de la manière suivante : Soit P_i le coût du programme couvert dans l'unité i, tel qu'il a été calculé plus haut et soit n le nombre de lauréats du système comprenant k unités.

Le coût moyen par lauréat est $\bar{P} = \frac{P_1 + P_2 + \ldots P_k}{n}$. Ce coût moyen par lauréat pourrait être également saisi en cumulant les coûts moyens par étudiant

dans chaque unité

Soit $\bar{\bar{P}} = \frac{P_1}{N_{11}} + \frac{P_2}{N_{22}} + \ldots + \frac{P_k}{N_{kk}}$

Le rendement du système pourrait être traduit par le rapport de ces deux coûts moyens :

$C = \frac{\bar{\bar{P}}}{\bar{P}}$. Le coefficient $C < 1$. Il est égal à 1 dans le cas où il n'y a pas de déperditions dans le système.

En conclusion, nous pouvons appréhender le rendement du système par l'un des paramètres suivants :

1) $\Pi = s_{11} \cdot s_{22} \cdot \ldots \cdot s_{kk}$

2) $\tau = \frac{n_{kk}}{N}$

3) $\bar{k} = \tau.k + \sum_{\alpha=1}^{m} \tau_\alpha.(k+\alpha) + \tau^*. 2k$

4) $C = \frac{\bar{\bar{P}}}{\bar{P}}$.

A titre d'exemple, calculons ces paramètres pour le cas du cycle de formation des Ingénieurs d'Application de la Statistique. Ce cycle s'étale sur 3 années ($k = 3$). Pour la cohorte qui a commencé ses études en 1975, les chiffres sont comme suit :

N_{11}	= 92	, n_{11}	= 39	S_{11}	= 42 %	, P_1	= 1 004 244	Dh
N_{22}	= 40	, n_{22}	= 32	s_{22}	= 80 %	, P_2	= 1 165 963	Dh
N_{33}	= 32	, n_{33}	= 28	s_{33}	= 88 %	, P_3	= 566 857	Dh

D'autre part $m = 1$ et $\tau_1 = 0,26$. Les calculs donnent :

1) $\Pi = 0,29$

2) $\tau = 0,30$

3) $\bar{k} = (0,30) \times 3 + (0,26) \times 4 + (0,44) \times 6 = 4,58$

4) $c = \frac{\bar{\bar{P}}}{\bar{P}} \equiv \frac{57\ 778}{97\ 752} = 0,59$

L'interprétation est que le taux de succès global (pour traverser toutes les étapes du cycle) est de 30 %. La durée moyenne passée par l'étudiant pour sortir lauréat du cycle est près de 4 ans et demi au lieu des 3 années réellement prévues.

Enfin le rapport des coûts moyens entre le système sans déperditions et le système tel qu'il est de 0,59. Dans un certain sens, prés de 40 % des dépenses sont à "fonds perdus".

CAPACITÉ IDÉALE ET COÛTS THÉORIQUES

Les moyens aussi bien humains que matériels dont on dispose ne sont pas nécessairement utilisés à l'optimum. Il n'est pas rare de constater dans certaines institutions qu'une salle de classe d'une capacité de 40 n'abrite en fait qu'une dizaine d'étudiants ou bien qu'un spécialiste pouvant encadrer une vingtaine d'auditeurs ne se retrouve en fait qu'avec un faible nombre de disciples. Cette situation résulte parfois de difficultés de recrutement des stagiaires, de l'inadéquation des locaux, de la nature très spécialisée ou avancée de certains enseignements, etc...

La raison peut être tout simplement due à la nature "discontinue" qui lie le nombre d'étudiants aux besoins en salles de classe et en enseignants. En d'autres termes, et à titre d'exemple, on peut utiliser la même salle et le même professeur jusqu'à 20 étudiants mais pour 6 étudiants supplémentaires inscrits dans le même cours, le dédoublement des moyens devient nécessaire, d'où un "Saut" de la fonction coût.

Dans de telles situations, il est naturel de rechercher un taux d'encadrement humain et un taux d'utilisation du matériel qui dans un certain sens atteignent l'optimum. Parmi les contraintes à imposer dans la recherche de cette capacité idéale, il convient de fixer des effectifs qui conduisent à un taux de succès maximal. L'objectif est de garantir l'assimilation des connaissances par les étudiants dans les conditions les meilleures mais sans gaspillage. La résolution de ce programme fait intervenir de nombreux paramètres et dépend largement des spécificités et des besoins de chaque Pays, chaque institution et chaque cycle de formation. La diversité des conditions qui peuvent prévaloir nous recommande de ne pas tenter une résolution dans le cas général.

Si nous nous plaçons à l'optimum, le coût théorique moyen par lauréat sera saisi par un nombre unique à savoir :

$$\bar{P}' = \frac{P'_1 + P'_2 + \ldots + P'_k}{n'}$$

où les P'_i sont ceux dont il s'agit plus haut mais avec des effectifs n'_{ik} au lieu de n_{ik} ($n' = n'_{kk}$) qui correspondent à un schéma de transition et à des taux de succès contrôlés.

Le rapport des coûts $\frac{\bar{P}'}{\bar{P}}$ traduit le taux d'efficacité du programme et reflete l'amélioration possible des coûts de formation.

LES ENSEIGNEMENTS STATISTIQUES A L'ENSAE

Ph Tassi et A Trognon, ENSAE, Malakoff Cedex

A - QU'EST QUE L'E.N.S.A.E.?

"L'Ecole Nationale de la Statistique et de l'Administration Economique assure la formation de spécialistes en statistique et en économie appelés à prendre place parmi les cadres de l'Administration, des organismes publics et semi-publics et des entreprises. Elle assure en particulier la formation des fonctionnaires du corps des Administrateurs et du corps des Attachés de l'Institut National de la Statistique et des Etudes Economiques (I.N.S.E.E.).

Elle contribue à la diffusion des connaissances scientifiques et techniques relatives à la méthode statistique, à l'analyse et à la prévision économiques, à l'économétrie, au traitement de l'information, à la démographie et à la sociologie quantitative. Elle concourt à la recherche dans les mêmes domaines".

C'est de cette manière que l'article premier du décret régissant son fonctionnement définit l'E.N.S.A.E.. On y décèle immédiatement la multiplicité des tâches qui lui sont assignées, ce qui se traduit dans les faits par une relative complexité d'organisation que nous allons tenter de résumer.

Premièrement, l'E.N.S.A.E. ne forme pas que des fonctionnaires, la majorité de ses anciens élèves a des responsabilités dans les entreprises françaises ou internationales. Deuxièmement, elle dispense deux niveaux de formation: l'E.N.S.A.E., c'est deux écoles en une. Plus précisément, elle abrite deux divisions qui se distinguent par le recrutement et par le contenu des enseignements.

La Division des Statisticiens Economistes et des Administrateurs (S.E.A.) accueille des élèves qui, ayant reçu une formation scientifique générale d'un niveau élevé sont aptes à assimiler un enseignement fondamental en statistique et en économie et acquérir la pratique des techniques les plus avancées. La division S.E.A. recrute par concours au niveau des classes préparatoires aux Grandes Ecoles (Bac + 2) et, par admission sur titres, d'anciens élèves des Grandes Ecoles scientifiques ou de l'université. Les élèves de cette division passent 2 ou 3 ans à l'Ecole selon leur niveau en mathématique et en économie; ils sont de l'ordre de 250 à l'Ecole, 80 à 90 par promotion.

La Division des Cadres de Gestion Statistique et des Attachés (C.G.S.A.) reçoit des élèves qui, ayant déjà effectué des études supérieures, suivront un enseignement et acquerront la pratique leur permettant de diriger l'exécution de travaux statistiques, de participer à la conception de grandes enquêtes

et de collaborer à la préparation de programmes économiques. La division C.G.S.A. recrute par concours au niveau des classes préparatoires et de la deuxième année de Sciences Economiques. Les élèves de cette division effectuent tous deux ans d'études a l'Ecole, ils sont de l'ordre de 160, soit environ 80 par promotion.

D'autre part, depuis 1963, est associée à l'E.N.S.A.E. une école spécialement créée par les communautés européennes afin de former les cadres de la statistique dans les pays francophones en voie de développement : le Centre de formation des Statisticiens Economistes des pays en voie de Développement (C.E.S.D.) Les élèves de ce centre suivent, en plus d'un enseignement spécifique, une part très importante des cours de la division S.E.A.

B - L'ENSEIGNEMENT STATISTIQUE A L'E.N.S.A.E.

Ce cadre rapidement dressé, nous pouvons aborder à présent notre propos principal : l'enseignement de la statistique à l' E.N.S.A.E. Dans cette note rapide, nous n'avons pas dressé un catalogue complet de tous les cours à caractère statistique. Nous avons, au contraire, choisi à titre d'exemple, trois enseignements de la division S.E.A. dont les statuts et les méthodes pédagogiques sont relativement différents : les enseignements de statistique mathématique, d'économétrie et de statistique appliquée.

La statistique mathématique est une discipline fondamentale dont le contenu est assez bien stabilisé.

L'économétrie, par opposition est un domaine en constante évolution dans lequel il faut distinguer un ensemble de connaissances minimales, à acquérir par tous les élèves et les approfondissements s'adressant aux élèves particulièrement intéressés par ce domaine.

Enfin, la statistique appliquée, pour laquelle on ne peut parler ni de niveau minimal à faire connaître, ni de développement constant, procède de la volonté de mettre à l'épreuve de la réalité des données les concepts statistiques et économétriques fondamentaux.

Avant de décrire en détail ces trois enseignements, résumons l'esprit général dans lequel se déroule la scolarité des élèves de la division S.E.A.

En raison des formations antérieures diverses de nos étudiants, la première année dispense un enseignement à dominante mathématique pour ceux des élèves provenant d'une filière à caractère économique et un enseignement principalement économique pour ceux des élèves ayant eu une formation en mathématique.

Tous les élèves de première année acquièrent les connaissances indispensables en calcul et théorie des probabilités. Commencent aussi à un niveau élémentaire et essentiellement pratique des enseignements d'analyse des données et de statistique descriptive.

Année charnière, la deuxième année a pour fonction principale l'enseignement des notions de base de statistique mathématique et d'économie théorique.

Enfin, en troisième année peuvent être mises en application les notions, assimilées lors des années précédentes, soit dans des cours théoriques tel que l'économétrie soit dans des groupes de travail (cf. plus loin) au sein desquels des travaux appliqués sont réalisés, soit encore dans des cours à option.

S'il fallait résumer de manière très synthétique, en quelques mots simples, l'esprit général de chacune des trois années nous pourrions dire :

Première année : rattrapage, mise au niveau.
Deuxième année : acquisition des connaissances théoriques indispensables.
Troisième année : utilisation, mise en oeuvre, application.

LA STATISTIQUE MATHEMATIQUE

Organisation matérielle

Cette matière, enseignée en deuxième année, est constituée par 45 séances(*) de cours magistraux à l'ensemble de la promotion(**) à raison d'une à deux séances par semaine et 34 séances de travaux dirigés (TD). Ces derniers se font par groupe de quinze élèves et ils ont pour vocation de faire résoudre des exercices d'assimilation du cours. Certains TD peuvent nécessiter quelques travaux informatiques ou encore des interprétations de résultats sur listings. D'autres TD, plus volumineux, donnent lieu à notation. Ces TD dits "doubles" sont préparés par deux élèves et présentés oralement aux autres élèves du groupe.

Le contrôle classique des connaissances comprend deux compositions écrites de l'ordre de trois heures, et une interrogation orale.

Contenu de l'enseignement

Destiné à servir de fond à la culture statistique des étudiants et de base aux cours de troisième année où la théorie statistique est un instrument, le cours de statistique mathématique développe les divers éléments de la statistique paramétrique.

(*) Les séances durent 1 heure 10
(**) De l'ordre de 90 élèves

En introduction, cet enseignement définit avec précision la notion de problème statistique en utilisant la théorie de la décision; il rappelle ou présente ensuite un certain nombre d'outils statistiques généraux (exhaustivité, totalité, information, identification etc...) qui seront utilisés dans le cadre des théories de l'estimation et de tests; enfin la dernière partie de ce cours souligne comment les résultats obtenus dans les parties précédentes s'appliquent au cas du modèle linéaire.

Ce cours bénéficie d'un support écrit cohérent d'un niveau d'abstraction élevé (*). Dans son exposé oral le professeur adopte dans ces conditions un ton moins théorique en mettant l'accent sur des exemples pédagogiques. Il se trouve ainsi allégé des démonstrations mathematiques rigoureuses que l'étudiant peut assimiler par lui-même par référence au cours écrit.

Ainsi dans le cadre de la partie consacrée au modèle linéaire le formalisme adopté le rapproche de celui utilisé en économétrie permettant ainsi une liaison plus facile entre le cours de statistique mathématique de 2ème année et celui d'économétrie de 3ème année.

Une expérience pédagogique passée

Au cours de son histoire, l'enseignement de la statistique mathématique à l'E.N.S.A.E. n'a pas toujours correspondu au schéma que nous venons d'indiquer. Ainsi a été réalisée l'éxpérience suivante : en deuxième année avait lieu un cours magistral de théorie des probabilités (il a lieu aujourd'hui en première année) tandis que l'enseignement de la statistique mathématique se déroulait en groupe de travaux dirigés où se mêlaient les exercices et les éléments de cours s'appuyant sur des notes écrites distribuées à l'avance.

Une telle formule présente certains avantages pédagogiques. Elle permet, en particulier, de mieux accorder le niveau du cours à celui des étudiants. Mais, précisément, cette diversité entre les groupes risque d'être néfaste aux enseignements ultérieurs qui s'appuient sur un acquis moyen certain d'une promotion; ce niveau de connaissance peut ne pas être parfaitement défini et très variable d'un groupe d'élèves à un autre. En outre, une telle formule repose fortement sur les notes écrites servant de support aux diverses séances, ce qui entraîne nécessairement une forte rigidité. C'est d'ailleurs essentiellement en raison de la difficulté à la faire évoluer que cette formule a été abandonnée.

(*) Le support de cours est le livre de A. MONFORT : "Cours de Statistique Mathématique", Economica, 1982, France.

LES ENSEIGNEMENTS D'ECONOMETRIE

Organisation

Les enseignements des méthodes statistiques de l'économétrie se déroulent durant l'année terminale, la troisième année. Ils s'articulent autour de deux cours. Le premier, partie intégrante et primordiale du tronc commun, enseigné durant le premier trimestre de l'année scolaire, s'étend sur 22 séances de cours et 10 séances de travaux dirigés. Ce cours dit "Econométrie I" dispense les connaissances fondamentales à l'ensemble de la promotion. Il est sanctionné par un examen écrit.

Le second cours, optionnel, appelé "Econométrie II", est d'un volume de trente séances de cours magistraux et s'adresse aux étudiants les plus motivés par l'économétrie théorique. Ce cours rassemble à peu près le quart d'une promotion. Le contrôle des connaissances repose sur des travaux personnels des étudiants.

Le contenu des enseignements

a) Econométrie I

Ce cours a pour but principal d'apprendre aux étudiants comment les méthodes statistiques qui leur ont été enseignées peuvent être utilisées et adaptées pour la spécification et l'estimation des modèles économiques. L'économétrie est une matière en développement continu. Ce cours ne vise donc pas à couvrir l'ensemble des méthodes actuellement disponibles mais plutôt à dispenser une formation de base solide. Il comporte une réflexion critique sur les hypothèses utilisées et les modalités pratiques d'application des techniques étudiées.

Le cours se divise en trois parties. La première est consacrée à une révision approfondie du modèle linéaire et de la méthode des moindres carrés ordinaires, dont les étudiants connaissent déjà bien les principales propriétés théoriques, en insistant sur les problèmes spécifiquement économétriques que soulève leur mise en oeuvre: multicolinéarité, utilisation de variables indicatrices, estimation sous contraintes. Une seconde partie traite du modèle linéaire général et de la méthode des moindres carrés généralisés; notamment dans les cas suivants: hétéroscédasticité, autocorrélation des perturbations, régressions empilées, modèles à double indice. Enfin, dans une troisième partie sont abordées diverses classes de modèles qui ne sont pas réductibles à des modèles linéaires et pour lesquels des méthodes d'estimation particulières ont été mises au point: modèles autorégressifs, modèles dans lesquels la liste des variables explicatives est incomplète, modèle non linéaire, modèles à équations simultanées, modèles à erreurs sur les variables.

Le cours oral est complété par des notes écrites et par des travaux dirigés. Au cours de ceux-ci les étudiants sont amenés à faire passer sur ordinateur les programmes économétriques classiques, à interpréter les résultats d'études économétriques diverses et à approfondir certains points techniques du cours.

Econométrie II

L'Econométrie, nous l'avons déjà souligné, est une matière mouvante sans cesse en renouvellement. La définition du minimum de connaissance que devrait avoir chacun de nos étudiants est délicate. Aussi offrons-nous, aux étudiants qui le désirent une option d'approfondissement portant actuellement sur trois grands thèmes en développement:

- les modèles à équations simultanées
- les modèles pour données individuelles temporelles (panels)
- les modèles pour variables qualitatives.

Chacun de ces trois thèmes est enseigné par un ou des chercheurs de la spécialité. Le niveau élevé de cet enseignement peut être atteint grâce à des supports de cours très détaillés et avancés, sur lesquels peuvent s'appuyer les exposés oraux en ne dégageant que l'esprit principal sans s'attarder sur les techniques de démonstration. Certains de ces supports ne sont pas encore figés et évoluent au gré de l'avancement de la recherche dans le domaine.

Bien que cet enseignement ne perde jamais de vue les modèles ou les données économiques pour lesquelles les techniques d'estimation et de tests sont élaborés, l'aspect théorique de cet enseignement parait relativement marqué. Cependant, et dans la mesure du possible, on tente d'associer à chaque procédure nouvelle d'estimation ou de test un programme informatique ou quelques éléments d'algorithmique permettant à chacun d'élaborer son propre programme. L'étudiant dispose dans ces conditions des moyens pour mettre en oeuvre en vraie grandeur, dans le cadre par exemple d'un groupe de travail, les procédures qu'il vient d'apprendre.

Afin de faire manipuler par les étudiants les instruments statistiques utilisés dans ce cours avancé, nous avons tenté cette année une expérience de contrôle des connaissances sur devoirs faits à domicile. Ces devoirs comportaient certaines questions ouvertes où l'étudiant pouvait mettre en valeur des qualités de recherche et d'analyse d'un problème économétrique théorique.

LA STATISTIQUE APPLIQUEE

Structure des enseignements appliqués

Il n'existe pas, dans l'ensemble des enseignements de l'E.N.S.A.E., un et un seul cours d'application de la statistique au sein duquel serait fait le passage obligatoire entre les théories statistiques étudiées et leur mise en oeuvre sur données réelles. La pratique statistique est présente tout au long des trois années d'enseignement: en première année, les élèves réalisent, dans le cadre du contrôle des connaissances du cours d'Analyse des Données, des mémoires en vraie grandeur; en deuxième année, existent, depuis 1977, des groupes dits de "Statistique Appliquée", au sein desquels les étudiants sont confrontés à des problèmes sur données réelles, dont la résolution fait appel aux techniques usuelles de la statistique descriptive; en troisième année, enfin, l'aspect pratique intervient par la présentation d'exemples d'application des méthodes statistiques par des contrôles des connaissances sous forme de mémoire, par la conception et l'exploitation d'une enquête dans le cadre du cours optionnel de Sociologie appliquée, et surtout, par les groupes de travail. Conçus de façon analogue aux groupes de Statistique Appliquée de 2ème année, cette forme pédagogique permet d'aborder des questions proches de la recherche, donnant lieu à des travaux théoriques et appliqués où la plupart des cours de l'E.N.S.A.E. peuvent être utilisés.

Nous allons détailler de façon plus approfondie les enseignements de "Statistique Appliquée" et les Groupes de Travail.

Les groupes de "Statistique Appliquée"

La finalité des Groupes de statistique appliquée est, comme cela a été indiqué ci-dessus, de permettre aux étudiants d'aborder des problèmes concrets dont les éléments ne sont pas formalisés.

Cet enseignement commence par huit cours magistraux devant l'ensemble de la promotion, dans lesquels sont abordés les éléments de théorie des indices, de désaisonnalisation et de lissage des séries temporelles, au niveau de la statistique descriptive. Puis les étudiants se répartissent par groupe de 4, chaque groupe choisissant un thème d'étude descriptive parmi plus d'une vingtaine de sujets proposés. Sous la direction d'un animateur connaissant les données et travaillant dans le domaine concerné, les étudiants vont se réunir au rythme de 2 heures tous les quinze jours, pour faire le point de l'avancement de leurs travaux, tout au long de l'année (en moyenne, 10 séances de 2 heures). A la fin de l'année scolaire, un mémoire doit être rédigé, et soutenu oralement.

Les questionnaires d'évaluation de cette forme pédagogique laissent apparaître un intérêt unanime de la part des étudiants, qui découvrent ainsi l'importance des phases antérieures à toute étude formalisée: mise en forme des données, création et structuration de fichiers informatiques, "éclairage" des données par la statistique descriptive.

Il faut également mentionner que la proximité de l'I.N.S.E.E. favorise ce type de travaux, non seulement pour la définition des sujets et la disponibilité des données, mais encore par l'appui de sa structure informatique (terminaux, écrans, logiciels).

A titre d'exemple, nous présentons en annexe 1 la liste des thèmes retenus pour l'année scolaire 1981-1982. Nous l'avons dit, cette forme d'initiation à la description des données est apparue en 1977 dans les enseignements de l'E.N.S.A.E. Jusqu'à cette date, l'apprentissage de la statistique descriptive donnait lieu soit à un cours magistral - voué au désintérêt du public en raison de son manque de support mathématique - soit à des travaux dirigés portant sur les sujets fondamentaux que sont la désaisonnalisation, les indices, la concentration ...

La forme pédagogique des travaux dirigés ne s'est pas révélée adaptée à la matière. La pratique statistique ne s'enseigne pas, mais se découvre et c'est pourquoi la formule des petits groupes autonomes semble la seule possible.

Les groupes de travail

Les groupes de travail constituent, en troisième année, l'analogue des groupes de Statistique Appliquée de 2ème année. L'organisation de l'enseignement est très proche: parmi la quarantaine de thèmes proposés, les étudiants par groupes d'effectif moyen égal à quatre, choisissent un sujet d'études dans des domaines divers: statistique, économie, économétrie, informatique. Chaque groupe se réunit avec le (ou les) animateur(s), deux heures par semaine, et doit en outre, consacrer au moins une demi-journée de travail hebdomadaire à cette activité.

Le travail du groupe est concrétisé par la rédaction d'un mémoire, qui peut parfois donner lieu à publication.

Le fait que le nombre de thèmes proposés soit passé de 6 en 1969, (dans le cadre d'une option), à 43 en 1981, (cette activité étant devenue obligatoire), montre le succès d'un tel enseignement. Ayant dépassé l'analyse descriptive du cours de "Statistique Appliquée", il nécessite la conception et le test de modèles théoriques, et fournit ainsi une première approche d'une direction d'études, qui est souvent l'activité professionnelle d'un ancien élève de la

division S.E.A.

La liste des thèmes des groupes de travail proposés en 1981-1982 est fournie en annexe 2.

* * *

L'évolution rapide de la technologie informatique, de son implantation dans les entreprises et, partant, son impact sur la méthodologie statistique, doit être également l'une des préoccupations majeures d'une école formant des statisticiens pour lesquels l'informatique est un outil indispensable.

Dans ce domaine l'E.N.S.A.E. parait privilégiée par la proximité de l'I.N.S.E.E. L'accessibilité des données, des logiciels (SAS, TSP, BIBLOS) et du matériel (les élèves utilisent un IBM 3033) est une opportunité pour les enseignements à caractère appliqué. Cependant ces facilités ne correspondent pas nécessairement à l'environnement informatique que risqueront de côtoyer nos anciens élèves. Il convient dans l'avenir d'adapter nos enseignements appliqués aux micro-ordinateurs, qui formeront très vraisemblablement l'outil futur du statisticien.

On peut cependant penser que l'apprentissage des logiciels généraux de statistique ne sera pas perdu à terme; l'évolution des tailles de mémoires aidant, certains grands logiciels, par exemple TSP, sont ou seront utilisables sur certains micro-ordinateurs.

La vocation de l'E.N.S.A.E. dépasse largement la déjà très difficile formation des cadres des instituts de statistiques. Les étudiants seront confrontés, dans leur vie professionnelle, à des problèmes très divers, que la plus grande variété des cours à option ne peut évidemment couvrir. Il paraît donc important de fournir à nos élèves un niveau élevé en statistique théorique. Cette culture statistique devrait leur permettre d'être "adaptables" et d'assimiler facilement les connaissances complémentaires dont ils auraient éventuellement besoin par la suite, soit par des lectures personnelles, soit par l'intermédiaire d'actions de formation auxquelles l'E.N.S.A.E. peut participer. Néanmoins, par essence, la statistique est inséparable de l'application. Une école telle que l'E.N.S.A.E. se doit d'être attentive en permanence aux besoins de l'extérieur, savoir quelles sont les méthodes nécessaires aux statisticiens d'entreprises et d'administration: c'est dans cet esprit que se sont développés les stages des étudiants, obligatoires en fin de première année. Sans infléchir véritablement l'enseignement statistique fondamental, de telles informations peuvent orienter le choix des thèmes proposés dans les enseignements de statistique appliquée et les groupes de travail.

Ce double souci de rigueur méthodologique et d'ouverture sur la pratique professionnelle est la préoccupation constante de la direction de l'Ecole et c'est dans cet esprit que celle-ci met actuellement en place une cellule de recherche. Cette cellule, accueillant des jeunes anciens élèves encadrés par des professionnels compétents, devrait être un lieu de rencontre et d'animation intellectuelle qui sera précieux pour l'évolution des enseignements.

A N N E X E 1

THEMES DE STATISTIQUE APPLIQUEE

(Année 1981-1982)

1. Analyse du trafic SNCF (grandes lignes); répartition entre la 1ère classe et la 2ème classe - Impact du TGV.
2. Enquête sur les communications écrites et téléphoniques au sein d'une entreprise.
3. Prévisions de ventes.
4. Utilisation du programme CENSUS pour la gestion.
5. La fécondité en France, au niveau départemental, depuis la 2ème guerre mondiale.
6. Analyse d'exportations et d'importations.
7. Modélisation de la diffusion du téléphone en France.
8. Evolution du marché du chauffage.
9. Performances d'un système informatique.
10. La consommation des ménages aux îles du Cap-Vert.
11. Lien entre activité féminine et fécondité.
12. Caractéristiques des automobilistes en infraction.
13. Influence des traitements de la vigne sur la qualité d'une eau de vie.
14. Une typologie des fraudeurs à l'I.R.P.P.
15. Le chômage régional.
16. Le marché de l'horlogerie.
17. Le rôle de la distribution dans la reconquête du marché intérieur.
18. L'effet de la réduction du temps de travail sur l'emploi dans le secteur du Commerce.
19. Les loisirs en région parisienne.
20. La mobilité pendant le week-end.
21. Prévision de températures journalières.
22. Etude de la validité des tests de personnalité.
23. L'équipement des ménages en biens durables.

A N N E X E 2

GROUPES DE TRAVAIL PROPOSES EN 1981-1982

1. Estimation d'un modèle macroéconomique d'équilibre avec rationnements.
2. La régionalisation: disparités inter-régionales d'épargne et d'investissement.
3. Etude de l'aire d'attraction théorique des espaces verts urbains et péri-urbains.
4. Stratégies de développement et chocs extérieurs.
5. Autonomie du développement économique et ouverture sur l'extérieur dans les pays de l'A.S.E.A.N.
6. Investissement et insertion internationale des économies caractérisées par des salaires réels imparfaitement flexibles.
7. Initiation à l'assembleur I.B.M.
8. Contraintes extérieures au développement: une analyse macroéconomique de déséquilibre.
9. Les méthodes modernes d'analyse des données.
10. Modèles et simulations sur les comptes d'entreprises.
11. La diversification des entreprises industrielles.
12. Analyse des données du projet de comparaisons internationales en Afrique.
13. Pratique de la méthode de Box et Jenkins et développement des programmes.
14. Contrôle monétaire et structures bancaires: évolutions et réformes en cours.
15. Le comportement d'offre.
16. La fécondité différentielle des familles à partir d'un modèle de répartition selon le nombre d'enfants.
17. Rentabilité et stratégie des entreprises françaises.
18. Emploi et productivité: analyse sectorielle.
19. L'analyse statistique et économique de plans d'investissement prévisionnels et de leur validation.
20. Relations entre modèle annuel et modèle trimestriel: application à l'emploi et au cycle prix-salaire.
21. Analyse discriminante.
22. Planification spatiale et marché foncier: appropriation publique des sols.
23. Politique industrielle et préférences nationales des structures.
24. Encombrement du temps et consommation des ménages.

25. Stratégie de développement des pays africains et nouvelle division internationale du travail.
26. Application de la méthode de Box et Jenkins à l'étude prévisionnelle court terme des principaux marchés automobiles européens.
27. Influence du codage et de la distance en analyse des données.
28. Etude de la structure des représentations de l'univers professionnel.
29. Evolution comparée des inégalités sociales en Chine et en Inde.
30. Séries chronologiques de transport de voyageurs.
31. L'endettement international et ses conséquences sur le système financier international.
32. Taux de change et politique économique en France sur longue période (1920-1979).
33. Evaluation des besoins de stationnement auprès des gares de la banlieue parisienne.
34. Spéculation privée et inflation.
35. Application de l'économétrie sur variables qualitatives à une enquête de conjoncture.
36. Dynamiques de long terme et spécialisations des grandes économies occidentales.
37. Incidences régionales des réformes de la politique agricole commune.
38. La conception d'un système d'information.
39. Typologie des entreprises exportatrices dans l'industrie française.
40. Chômage et crise de l'emploi dans les années 1930.
41. Econométrie de la dynamique des revenus et de la consommation.
42. L'équipement des ménages en biens durables.

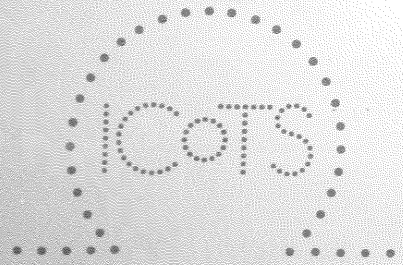
ISBN 0 946554 00 5 (set of two volumes)

ISBN 0 946554 01 3 (Volume I)

ISBN 0 946554 02 1 (Volume II)

Designed and printed by the University of Sheffield Printing Unit

John Taffe



PROCEEDINGS of the FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS

VOLUME II

University of Sheffield
9 – 13 August 1982



University of Sheffield
9 – 13 August 1982

PROCEEDINGS of the FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS

VOLUME II

Edited by:
Organising Committee of the First International Conference on Teaching Statistics

D R Grey, P Holmes, V Barnett, G M Constable

Published by:
Organising Committee of the First International Conference on Teaching Statistics

D R Grey, P Holmes, V Barnett, G M Constable

© Organising Committee of the First International Conference on Teaching Statistics 1983

ISBN 0 946554 00 5 (set of two volumes)

ISBN 0 946554 01 3 (Volume I)

ISBN 0 946554 02 1 (Volume II)

Printed by the University of Sheffield Printing Unit

Published by Teaching Statistics Trust



"It was hard work . . .



. . . but not all the time".

CONTENTS OF VOLUME I

	Page No.
Preface	xi
INTRODUCTORY TALK	
Why Teach Statistics? V Barnett	3
PLENARY TALKS	
Stochastics at the School Level in the Age of the Computer L Råde	19
Optimum Balance between Statistical Theory and Applications in Teaching C Radhakrishna Rao	34
The Relevance of Statistical Training J Gani	50
TOPIC 1: TEACHING STATISTICS IN SCHOOLS (6-11 YEARS AGE GROUP) ORGANISER: PROFESSOR K MATUSITA	
Statistics in the Curriculum for Everybody - how Young Children and how their Teachers react Tamas Varga	71
Teaching of Statistics in Elementary Schools in Japan T Yanagawa	81
Two Examples of Statistics Activities for Elementary School Classes Clare Heidema	91
Using Games to teach some Probability Concepts George W Bright, John G Harvey and Margariete M Wheeler	110
Teaching Statistics in Elementary Schools in Japan H Midzuno	116
TOPIC 2: TEACHING STATISTICS IN SCHOOLS (12-15 YEARS AGE GROUP) ORGANISER: PROFESSOR P L HENNEQUIN	
Why Teach Descriptive Statistics in Secondary Schools? Heinrich Winter	127
The Teaching of Statistics in French Secondary Schools G Dumousseau	144
Teaching Statistics Concepts through Instructional Games George W Bright, John G Harvey and Margariete M Wheeler	158
A Practical Approach to Statistics R Kapadia	169
On the Teaching of Order Statistics K Bognár	179

TOPIC 3: TEACHING STATISTICS IN SCHOOLS (16-18 YEARS AGE GROUP)
ORGANISER: PROFESSOR J SWIFT

Amount and Kind of Statistics to be taught to 16+ year olds
F Eicker 191

Teaching Statistics to 16-18 year olds: Dreams and Reality
T Nemetz 199

Statistics for the 16-18 year olds: a Discussion
A Hanley 208

TOPIC 4: TEACHING STATISTICS TO NON-STATISTICIANS
ORGANISER: DR W G GILCHRIST

We must preach what is practised
A S C Ehrenberg 215

Teaching Statistics to Non-Statisticians - an Application oriented Approach
G Sadasivan 219

Utility through Collaboration: a Problem oriented Approach
D Taylor and G E Skipworth 232

An Open University Service Course
John Bibby 238

The Training of Service Teachers
R W Madsen 250

TOPIC 5: DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL
ORGANISER: MR P HOLMES

Development of Teaching Materials at School Level
A R Hoffer 263

Developing Materials for the Schools Council Project on Statistical Education
P Holmes 269

EDA for 16-19's?
J Bibby 277

TOPIC 6: TRAINING OF TEACHERS IN STATISTICS
ORGANISER: PROFESSOR A P SHULTE

Preparations for College Teaching in Statistics
A K Gupta 283

Teaching Design of Experiments in Indian Universities
V Bhagwandas 287

Clinical Approach to training Statistics Teachers
J Rustagi 298

Discussion of Teacher-Training Papers
A Morley 302

TOPIC 7: USE OF CALCULATORS AND COMPUTERS IN TEACHING
ORGANISER: PROFESSOR L RADE

The Case for using Computers in teaching Statistics
B L Joiner 307

Computer Algorithms for the Classical Distribution Functions
H O Posten 313

Use of hand held Programmable Calculators in Statistics Courses for Nontechnical Students
D R Barr 331

Simulation of Games studied in the Seventeenth and Eighteenth Centuries
J Badrikian 336

TOPIC 8: TEACHING STATISTICS WITH THE HELP OF CASE STUDIES
ORGANISER: PROFESSOR B EPSTEIN

Case Studies for an adequate understanding and interpretation of Results by Statistical Inference
M Borovcnik 341

Is a Scientific Legend a comprehensive Case Study?
V V Shvyrkov and A C Davis 354

TOPIC 9: TRAINING STATISTICAL PRACTITIONERS
ORGANISER: PROFESSOR M BENYAKLEF

Training of Statistical Practitioners
P Osinski 375

Preparing Statisticians for Government Service - a View from the United States
Marie D Eldridge and Katherine K Wallman 389

Programmes d'Enseignement et Couts de Formation
M Benyaklef 403

Les Enseignements Statistiques a l'ENSAE
Ph Tassi et A Trognon 415

CONTENTS OF VOLUME II

Page No.

TOPIC 10: EDUCATION AND TRAINING IN STATISTICS IN DEVELOPING COUNTRIES
ORGANISER: DR M N MURTHY

Statistical Education and Training in Latin America: a General Appraisal
Enrique Cansado 431

Statistical Education and Training in the Developing Countries in Africa
J N Adichie and Biyi Afonja 439

Statistical Education and Training in Developing Countries of Asia and the Pacific
M N Murthy 450

Comments on the above Papers
K T de Graft-Johnson 483

Report on session on Statistical Education and Training in Developing Countries
B P Adhikari 487

TOPIC 11: CO-OPERATION BETWEEN ACADEMIC AND PRACTISING STATISTICIANS
ORGANISER: PROFESSOR W G HUNTER

Co-operation between University and Industry Statisticians in the United States
R D Snee 491

Some Observations from Europe
Lars-Erik Öller 503

Building Bridges between the Academic and Real Worlds - some Observations from South America
C A Peres and P A Morettin 509

Building Bridges between the Academic and Real Worlds
B Jones, B J T Morgan and G B Wetherill 521

Closing the Gap between Statisticians and Statistics Users
Guido E del Pino and Carlos Prado-Campos 527

Co-operation between Academic and Practising Statisticians
A Prat Bartés 536

TOPIC 12: TEACHING SURVEY SAMPLING
ORGANISER: DR D SINGH

The Teaching of Sample Design to Government Statisticians from Developing Countries
Joseph K Garrett 541

Teaching of Survey Sampling
D Singh 552

Teaching of Survey Sampling
P K Bose 566

TOPIC 13: TEACHING DESIGN AND ANALYSIS OF EXPERIMENTS
ORGANISER: DR D A PREECE

The Design and Analysis of Experiments - what is it about?
D A Preece 575

Teaching Experimental Design in the Computer Age
R Mead 591

Teaching Concepts of Design and Analysis using Real and Simulated Experiments
J R Elliott and J C Robinson 603

Teaching Experimental Design: prescribed by a Medical Statistician
Sheila M Gore 627

Contribution to Discussion of the above Papers
B Jones 633

TOPIC 14: TRAINING INDUSTRIAL PRACTITIONERS
ORGANISER: DR R D SNEE

The Polymath Consultant
A A Greenfield 641

Teaching Applied Statistics for and/or in Industry
H C Hamaker 655

Practicals, Case Studies, Projects and Professional Experience: some Methods of developing Skills among Budding Statisticians
G K Kanji 671

Training in Statistical Quality Control
A F Bissell 685

Discussion of Papers by Greenfield, Hamaker and Kanji
A F Bissell 691

Discussion of Papers by Greenfield, Hamaker and Kanji
D G Watts 695

TOPIC 15: PRINCIPLES OF LEARNING PROBABILITY AND STATISTICS
ORGANISER: PROFESSOR E FISCHBEIN

Learning from Experience - Bayes' Theorem: a Model for Stochastic Learning Situations?
H Steinbring and G von Harten 701

Children's Choice Behaviour in Probabilistic Situations
Ruma Falk 714

An Idea of structuring Probability Teaching based on Dienes's Six Stages
Anna Olecka 727

Does the Teaching of Probability improve Probabilistic Intuitions?
E Fischbein and A Gazit 738

Stochastics Teaching based on Common Sense
H-J Bentz 753

A Survey of Probability Concepts in 3000 Pupils aged 11-16 Years
D R Green 766

Misconceptions of Probability, systematic and otherwise; teaching Probability and Statistics so as to overcome some Misconceptions
J M Shaughnessy 784

The Role of Explanation in Teaching Elementary Probability
J L Myers 802

Probabilities: Learning by doing or learning by imitating
A Abele 813

TOPIC 19: TEACHING OF AGRICULTURAL STATISTICS
ORGANISER: DR D SINGH

Training of Personnel for Agricultural Statistics in Jordan
Z Takrouri 831

PREFACE

The first International Conference on Teaching Statistics (ICOTS) took place in the University of Sheffield over the period 9-13 August, 1982. The Conference attracted much interest from all corners of the world. Over 400 full participants from more than 60 countries (accompanied by 100 or so associate participants) spent an active week attending Plenary Sessions, Invited Talks, Contributed Paper Sessions, Workshops and a variety of social events including walks in the Peak District hills and a reception and buffet generously provided by the Lord Mayor of Sheffield on behalf of the City.

The formal programme was very full and covered a wide range of interests in statistical education and training, including the teaching of statistics for all age groups in schools and colleges, the training of practitioners in different applied disciplines, the special problems in developing countries and the difficulties of teaching specific topics. An opening address was given by Professor S. Gregory, Pro-Vice-Chancellor of the University of Sheffield and the closing session was chaired by Mr. H. Ben-Amor, the Director of the UNESCO Office of Statistics.

The Conference was sponsored by the International Statistical Institute and the organisers are grateful to the ISI, and to UNESCO, for financial support. The technical programme and local organisation were arranged by the Programme Committee and the Organising Committee under the respective chairmanships of Professor L. Råde and Professor V. Barnett. The full membership of the two committees was as follows:

Programme Committee

L. Råde (Chairman)
E. Cansado
P. Holmes
B. L. Joiner
M. N. Murthy
G. E. Noether
J. O. Oyelese
B. Penkov
R. Zielinski

Organising Committee

Professor V. Barnett (Chairman)
G. M. Constable
D. R. Grey
P. Holmes
A. Barnett (Conference Secretary)

This Conference arose from the special interests and actions of the International Statistical Institute (ISI): in particular, from the establishment in 1978 of an Education Committee Task Force charged with organising an international conference on the teaching of statistics. The field of statistical education has become a very extensive one and the need for international conferences on the teaching of statistics has been emphasised on several occasions, especially at the many round table meetings on the teaching of statistics arranged by

ISI, and at the international conferences on mathematical education organised by ICMI (International Commission on Mathematics Instruction).

These published Proceedings of the first ICOTS consist of all plenary and invited papers together with contributions from invited discussants, collected together under appropriate topic headings. It is regrettable that the sheer volume of this material has precluded the possibility of including any of the very worthy contributed papers.

It is appropriate to acknowledge the vast effort extended by all members of the Programme and Organising Committees in setting up the Conference and to thank all Topic Organisers for helping to provide an interesting programme. In addition to thanking all authors for their co-operation in submitting papers in suitable form for publication, it is necessary to express particular indebtedness to three people who have given invaluable assistance in bringing this publication into being: Dr. David Grey, who organised the overall preparation of the Proceedings, Mrs. Jenny Cameron, who has undertaken the mammoth task of re-typing all the papers in standard format, and Mr. Bill Seaton, one of the Conference participants, who during an extended period of sabbatical leave in Sheffield, has willingly devoted many hours to the onerous task of proof-reading.

In accordance with views expressed by Conference participants it is hoped that this Conference will be the first of a series of such international conferences on teaching statistics to be held at four-yearly intervals.

V. Barnett
(Chairman, Organising Committee)

Sheffield, December 1982

TOPIC 10: EDUCATION AND TRAINING IN STATISTICS IN DEVELOPING COUNTRIES

ORGANISER: DR M N MURTHY

STATISTICAL EDUCATION AND TRAINING IN LATIN AMERICA: A GENERAL APPRAISAL

Enrique Cansado, CIENES, Santiago

INTRODUCTION

The problems that the teaching of Statistics generates are diverse and complex. They emerge either from the "formal" teaching (primary, secondary, and university), which is offered under the responsibility of the "academic statisticians", or from the training of governmental statisticians, who have responsibility for operation and supervision of technical and professional activities.

Both problems received rather early attention in Latin America. Because of its historical relevance we must point out the settlement of a career in Statistics in the University of Litoral (now University of Rosario) in 1948, in the city of Rosario, in Argentina. This pioneer effort was responsible for the training of a numerous set of professionals now active either in Argentina or in several Latin American countries. Another important early training program was developed by the Instituto Brasileiro de Estadística (IBGE), in Brazil. Both projects actively continue their development in our day. Some efforts have been accomplished in recent years to evaluate the development and evolution of the training in Statistics in Latin America. Among the most important ones, according to our judgement, must be mentioned "The Technical Meetings on the Evolution of the Statistical Profession in the Western Hemisphere", that took place in Santo Domingo, Dominican Republic (November, 1977) sponsored by the Inter American Statistical Institute (IASI), the "Latin American Seminar on the Teaching of Statistics" (SELENES), held in Santiago, Chile (November, 1977), organized by the Inter American Statistical Training Center (CIENES) and "Working Group on the Statistical Profession in the Western Hemisphere" gathered also at CIENES headquarters (December, 1980), under the sponsorship of IASI.

In the Session of the International Statistical Institute (ISI), held last year in Buenos Aires, Argentina, a "Round Table on the Teaching of Statistics in Latin America" was conducted and its proceedings will be published in a short period from now. At this event most of the relevant statisticians of the Region participated and they presented important contributions on the subject.

PRESENT SITUATION

The teaching of Statistics in Latin America presents today a situation highly superior to the one prevalent twenty or thirty years ago but the perspective of improvement is not as clear as it was then.

The countries of the Region are strongly diverse, as it is well known, in what concerns their sizes and general degree of development. The teaching of Statistics presents, as it is natural, different achievements and difficulties in these countries either in what concerns the academic area or in that related to the governmental one.

During a long period of time, and as a consequence of the Inter American programmes conducted with efficiency and continuity, by the Organization of American States (OAS) and the Inter American Statistical Training Center (CIENES), there was available for Latin America a set of training activities that complemented the efforts performed by the national institutions - either academic or governmental - related with teaching, research, extension, and application of statistical methodologies. A relevant part of this work was performed through the Inter American Center for Training in Economic and Financial Statistics (CIEF) from 1953 to 1961, and mainly by the Inter American Statistical Training Center (CIENES), that started in 1962 and is still active today, both located in Santiago, Chile. More than five thousand Latin American Statisticians have participated in the courses developed by these Centers in its headquarters or offered by their professors in the different countries.

This important and large amount of work has unfortunately been limited by the serious financial difficulties that OAS has suffered for some years.

As a consequence of this situation the countries of the Region, as far as the teaching of Statistics is concerned, have a better situation than the one they had in the past, but they will depend more and more on their own resources and initiative to develop the different programs in statistical training required in all of them.

In what concerns the formal education at primary and secondary level, no complete information is available and the little existent leads to the generalized impression that in the Region only a few incipient efforts have been performed as isolated experiments.

According to the available information it can be mentioned that thirteen of the twenty-eight state members of OAS have one or more university careers in Statistics; namely, Argentina, Bolivia, Brazil, Colombia, Costa Rica, Chile, Ecuador, México, Panamá, Paraguay, Perú, Dominican Republic, and Venezuela. In seven of these countries there exists, at the present time, one or more academic university programs at a graduate level (M.Sc. or Diploma) with strong differences in orientation, degree of consolidation, and number of participants. These seven countries are Argentina, Brazil, Colombia, Chile, México, Perú, and Venezuela. Only Argentina, Brazil, and México have programs at doctorate level founded and carried on by local universities.

All the countries recognize the importance of statistics teaching and the need for professionals in this field but they face serious difficulties in establishing such programs of education, perhaps the most important one being the lack of professionals with an adequate formation to offer teaching at a good level. This fact largely conditions the quality of several of the already existing programs. The few graduate programs existent in the Region make them unable to fulfill the present needs.

A characteristic of university teaching in this part of the world, and not extremely different from the one in the rest of the countries, is that in the formal teaching a heavy weight is given to statistical theory and a very light one to applications and actual data analysis; this fact, in many cases, does not enable the participants of these careers to work in offices of the national statistical system unless they make some complementary training.

Research in Statistics, either theoretical or applied, has not reached a significant level in Latin American universities. Research in statistical theory is conducted only in a few isolated places. On the other hand the interest for applied research is becoming more widely recognized. In spite of it, applied research faces at present many problems and the quality of what is being done varies very much among the different countries as well as among universities within the countries.

The countries of the Region are not different from the rest of the world in the fact that the university systems or other institutions of academic training have not provided or do not provide, at present, adequate training that would allow their participants to work in the production of basic governmental statistics. This reason has had as a consequence that several National Statistical offices - generally the heads of the National Statistical Services - have created their own Training Units in order to improve, at a basic and operational level, the present formation of their staff. In very few cases a good cooperation between academic and governmental statisticians exists; the gap between these two groups is far wider than the one existing in more developed countries.

As a very positive effort to solve this problem we should mention the recent creation in México, by a law dated December 1980, of the Instituto Nacional de Investigación y Capacitación en Estadística y Geografía (National Institute for Research and Training in Statistics and Geography), "with the goal of conducting research in order to improve the procedures, methods, and techniques in the field of statistics and geography as well as to offer training to the personnel in general and technical staff to satisfy the needs that an adequate integration and development of the national statistical systems requires".

In 1962 there was installed in Chile the Inter American Statistical Training Center, a dependent of the OAS, which has been offering training to professionals

of the Region (either academic or governmental). Its contribution has been a significant one to this part of the world. In the academic area, the task has been reached through the Course in Mathematical Statistics, a graduate program at M.Sc. level, which has been completed by more than three hundred participants since it started. It was, for many years, the unique program of the Region offering a M.Sc. degree and still now is the course that produces the largest number of graduates at M.Sc. level per year in Latin America.

Several Latin American participants have completed doctoral programmes in American or European Universities. It is interesting to point out that almost half of them had previously obtained an M.Sc. degree at CIENES. The other two courses that CIENES regularly offered in its headquarters were dedicated to governmental statisticians. One was addressed to "producers" of statistics and the other one to "users" or "analysts" of statistics. More than one thousand four hundred Latin American governmental statisticians have profited from these courses. This significant amount of training activities has not a counterpart in the countries of the region.

CIENES has also been in charge of the Multinational Project of Educational Statistics since it was established in 1971. Every year courses on this field have been offered, and about two hundred students have participated in them.

More than one thousand five hundred students have participated in other specialized courses offered by CIENES.

The fields in which these courses were offered cover a very wide range of subjects: Agricultural Statistics, Trade Statistics, Statistics of Tourism, Demographic Statistics, etc.

CIENES has also offered courses in different countries of the region. Sometimes it has also participated in already established training programs sending one or several of its professors. About three thousand Latin American statisticians have participated in these courses, some of them academic ones but the major number in the governmental area.

CIENES continues offering at present courses for all the Latin American countries in its headquarters. Unfortunately the number of scholarships that OAS grants is extremely reduced. In this last year CIENES has dedicated a serious effort to support its Program of Assistance to the Institutionalization of the Teaching of Statistics (PRAIENES), that has as its goal, to offer help to the national institutions in the different countries of the Region, which have regular problems in the teaching of statistics, in order to strengthen such programs or to create new ones.

SOME COMMENTS ON THE PRESENT SITUATION

There is little correlation between the degree in which the importance of Statistics is accepted in an early stage of a process of education and how much has been implemented in this direction at primary and secondary level. So it seems quite commendable to make efforts to strengthen the teaching in this field, when it exists, or to establish it when it is not already offered. In order to accomplish this, well qualified professors must be available and, even before that, well elaborated programmes with a clear enumeration of well selected topics must be prepared. The intuitive aspects of Statistics and its possibilities of application should be clearly pointed out in order to emphasize the difference between statistics and classic Mathematics Problems. Problems like data production, inference, and processes of taking decisions should be presented. Given the complexity and wideness of the problems that the teaching of statistics at primary and secondary levels raises a close collaboration between statisticians and experts in education will be necessary as well as the participation of all national and international institutions involved in the problem. At the ISI Session held last December in Argentina, a round table on the problem of the teaching of statistics in Latin America was held.

It was pointed out there that the difficulties involved in keeping the university programs at an adequate operational level in almost all the countries was caused by economic reasons. The need of an evaluation of the labor market for the graduates of careers in Statistics was also discussed.

It is strongly expected that this Conference will be able to produce some recommendations on the problem.

In December 1980 a Working Group on the situation of the Statistical Profession in the Continent met at CIENES. As a consequence of this meeting very valuable recommendations on the teaching of statistics at University level were made.

These recommendations follow:

1. It is absolutely necessary to continue and to expand educational and training programs in statistics at the graduate level. This would have a twofold effect: It would increase the amount and the quality of research being carried out, and it would have a multiplier effect as well through the use of these professional statisticians as teachers. It should be noted that the quality of instructions improves if the teachers do research work.

2. To diminish the problem of financing research, advisory services, and manpower training for these activities, the groups of researchers should prepare projects in areas of primary interest to government agencies or industry. This point is particularly relevant for researchers working in the academic field.

3. In the field of international technical assistance, and especially where advisory and research services are involved, agencies sponsoring such services must increase their financial support. As a counterpart for this, the organizations that carry out this assistance must propose projects that are sufficiently feasible and significant to attract the interest and support of the governments of the region.

4. Given the great scarcity of publications in the Latin American area, communications among the various groups of researchers is difficult. Efforts should be made to have a regular periodic publication that would essentially benefit research in the region.

5. The more frequent and regular scheduling of national and international meetings would also be desirable. Meetings should be designed to favor the exchange of data on experience gained by researchers, and an interchange of information between researchers and professional users of results derived from research.

6. Research and advisory services should be institutionalized and be mandatory in both governmental and academic organizations. These activities should also be upgraded to be significantly rewarding to the professional at the economic and social level.

7. An active exchange among already existing research groups should be encouraged through periodic visits, exchange of professors, and other similar activities.

A good graduate program in Statistics at M.Sc. level would, as a first step, provide professors with an adequate amount of knowledge to be able to teach in university careers in Statistics. This M.Sc. program can be followed inside the same country, in another country of the Region (in most cases with the advantage of receiving the teaching in the own language of the participant) or in a university outside the Region.

There are very few doctoral programs in Statistics in the Region. Some of them have a serious prestige, but it could be said that at that level almost all the training is followed by Latin American candidates at American and European Universities. In several Inter-American meetings the first ideas of a project on a degree at doctoral level jointly sponsored and granted by some universities of Argentina, Brazil, and México and coordinated by CIENES has been discussed. However desirable the project may be, it seems not to be an easy one and its implementation cannot be thought about in the near future.

If the problems that arise in the process of improving the formal teaching of Statistics are varied, diverse, and difficult, the ones that face the improvement

in the capacity of governmental statisticians in Latin America are even more complex. Given the deep gap that separates (in the Region, even more than in other parts of the world) academic and governmental statisticians, the training of these last ones is not offered by the universities but through international centres. Examples of these last ones are CIENES and, with the same location, the Latin American Center for Demography (CELADE) depending on the United Nations. The National Statistical Offices of the Region are more and more establishing programs of Training to satisfy immediate objectives, but only at a basic or operative level.

SOME STATISTICAL FACTS

On the occasion of the 43rd Session of the International Statistical Institute, Dr. Raúl Mentz and Mrs. Nora M. Jarma de Cortés presented an invited paper [5], which will be the main source of some specific information about the situation of statistics as a profession in the Region.

The paper of Dr. Mentz and Mrs. Jarma de Cortés defines a quite enlightening line of research and CIENES will dedicate some efforts to it in the future. Dr. Mentz and Mrs. Jarma de Cortés belong to the National University of Tucumán, in Argentina. The information derives from a sequence of surveys implemented since 1979. The original frame of 27 centers was complemented with the list of institutions invited to the Latin American Seminar on the Teaching of Statistics (SELENES) gathered at CIENES in 1978, leaving the final number of consulted centers at 106. The answer came only from 29 Latin American Centers, which belong to Argentina (6), Brazil (11), Colombia (3), Chile (4), Guatemala (1), México (3), and Venezuela (1).

Table I of the above mentioned paper is reproduced on the next page.

The survey conducted by Professors Mentz and Jarma de Cortés is a serious effort but the non-response to the questionnaire is reflected in several absences. To give a few examples: There are several groups in Chile which are not recorded; they are located in Santiago, Valparaiso, Iquique and Antofagasta. Two strong Colombian groups are also absent (National University of Bogota and del Valle University in Cali). In the Caracas area are also several groups. This comment must not be considered a non-constructive criticism, it is just made in order to express that it is CIENES intention to present at the ISI 44th Session, which will be held in Madrid in 1983, an inventory of the Latin American statistical situation as complete as possible. The short period of time that elapsed between the ISI 43rd Session and the moment at which this paper is written made it impossible to conduct the required survey.

References

[1] Documentos presentados en la Reunión Técnica sobre la Evolución de la Profesión Estadística en el Hemisferio Occidental (Noviembre, 1977), Instituto Interamericano de Estadística (IASI), Washington, D.C., U.S.A.

[2] Seminario Latinoamericano sobre Enseñanza de Estadística (Noviembre, 1978), Centro Interamericano de Enseñanza de Estadística (CIENES), Santiago, Chile.

[3] Adiestramiento Estadístico en la Región Americana (Agosto, 1979), Centro Interamericano de Enseñanza de Estadística (CIENES), Santiago, Chile.

[4] Informe de la Reunión del Grupo de Trabajo para examinar la Situación de la Profesión Estadística en el Hemisferio Occidental (Diciembre, 1980), Instituto Interamericano de Estadística (IASI), Washington, D.C., U.S.A.

[5] Mentz, Raúl Pedro and Iarna, Nora M. de Cortés (1981) Encuesta a Centros Estadísticos Latinoamericanos (Survey on Latin American Statistical Centers). Invited paper 43rd Session of International Statistical Institute, Buenos Aires, Argentina.

Latin American Statistical Centers Survey

Teaching and Research Personnel classified by their highest degree

Countries	Centers	Total	(1)	(2)	(3)	(4)	(5)	(6)	(7
Argentina	National University of Buenos Aires	14	1	1	-	-	12	-	-
	National University of Córdoba	18	1	1	1	-	13	2	-
	National University of Comahue	10	-	-	-	1	7	2	-
	National University of Rosario	72	-	3	-	-	64	4	1
	National University of Tucuman	18	4	1	1	1	9	2	-
	National Institute for Agricultural Research (INTA) Castelar National Center	9	1	-	1	1	6	-	-
Brazil	University of Brasilia	33	9	1	-	19	4	-	-
	University of Sao Paulo, Department of Statistics,	41	12	2	-	17	10	-	-
	University of Sao Paulo, School of Medicine	9	-	6	-	2	-	-	1
	University of Sao Paulo, Agricultural School "Luiz de Queiroz"	11	-	9	-	2	-	-	-
	University of Sao Paulo, Mathematics at Sao Carlos	5	2	-	-	3	-	-	-
	State University of Campinas	26	4	2	2	17	1	-	-
	Federal University of Pernambuco	29	-	1	-	12	16	-	-
	Federal University of Rio de Janeiro (a)	25	9	5	1	7	3	-	-
	Federal University of Sao Carlos	14	3	1	-	3	7	-	-
	Federal University of Rio Grande do Sul	23	-	5	3	7	8	-	-
	Brazilian National Institute for Agricultural (EMBRAPA), Brasilia (a)	27	4	12	1	10	-	-	-
Colombia	School of Engineering, University of the Andes, Bogota	12	5	-	2	3	2	-	-
	Corporación Centro Regional de Población (Corporation of the Regional Center of Population Bogota	12	-	5	4	2	-	-	1
	SER Institute, Bogota	27	4	1	3	3	14	2	-
Chile	Austral University, Valdivia	7	-	-	1	3	3	-	-
	"Federico Santa María" Technical University	15	2	-	2	8	3	-	-
	Latin American Center for Demography (CELADE) (b)	nd	nd	nd	nd	nd	nd	nd	nd
	Inter American Statistical Training Center (CIENES)	11	3	-	3	1	4	-	-
Guatemala	Institute of Nutrition for Central America and Panama	5	2	-	-	2	1	-	-
Mexico	Autonomous University of Mexico, Institute o for Research in Applied Mathematics and Systems, (IIMAS) Mexico D.F.	13	7	-	-	2	4	-	-
	School of Agriculture, Chapingo (a)	24	10	-	-	14	-	-	-
	National Institute for Agricultural Research (INIA) Mexico D.F.	26	1	-	1	6	18	-	-
Venezuela	University of the Andes (a)	13	6	-	4	2	1	-	-
	TOTAL	549	90	56	30	148	210	12	-

The code corresponding to this table follows:

(1) Doctoral degree in centres of the United States, Europe, etc.
(2) Doctoral degree in Latin American Centres
(3) M.Sc. Degree in centres of the United States or Europe
(4) M.Sc. Degree in Latin American Centres
(5) Professional or B.Sc.
(6) Student
(7) Not specified
(a) Date from 1979-80 wave of the survey
(b) No data available about their current staff.

STATISTICAL EDUCATION AND TRAINING IN THE DEVELOPING COUNTRIES IN AFRICA

J N Adichie and Biyi Afonja
University of Nigeria, Nsukka and University of Ibadan

INTRODUCTION

It is now widely recognised, if still grudgingly by some academicians and persons in older professions, that statistics is both an academic discipline and a profession. In Africa as indeed in every other continent of the world, the teaching of statistics may be conveniently put into two broad categories: the teaching of statistics as an academic discipline in schools and colleges and the training in statistics as a profession. It would therefore be convenient to treat the topic of this paper in two parts, namely: statistical education corresponding to teaching statistics as an academic discipline in institutions and statistical training for a profession.

STATISTICAL EDUCATION IN AFRICA

Until recently, statistics was not considered pure or important enough to be admitted as an academic discipline in its own right. This phenomenon was not peculiar to developing countries, indeed it was even more pronounced in the developed countries of the world. Because the school and college curricula of most African countries were faithfully modelled after the colonial patterns, the type of statistics taught in schools then was made a part of the mathematics syllabus. In schools, the course consisted mainly of elementary techniques in the calculation of means and variances, binomial coefficients and Poisson probabilities. Feeble attempts made to introduce into the school syllabus practical and useful, yet elementary statistical topics like sampling procedures, methods of censuses and surveys, statistical reasoning, were quickly killed by purist mathematicians who would not compromise the elegance and beauty of mathematics by admitting such topics that are deemed to lack mathematical rigour. In colleges and universities, where it was taught at all, statistics was treated as part of Applied Mathematics. This was the state of affairs in most institutions in Africa in the sixties.

Teaching Statistics in Schools

In the early seventies, however, pressure was brought to bear on the schools to cultivate statistics as a distinct school subject. This was the case mainly with the schools in Anglophone countries of West Africa. Two principal factors were responsible for the pressure; the first was the crying need to have a recruitment base for the statistical assistants who would fill the many vacant positions in the National Statistics Offices of these countries. It must be

remembered that about this time, many African countries had gained independence from their erstwhile colonial masters and were therefore very mindful of consolidating their economy in this post-independence period. It was therefore important to strengthen the national statistical offices, by recruiting not only the professional and middle level statistical manpower but also the junior statistical personnel. A secondary school graduate with the General Certificate of Education (GCE) in Statistics and Mathematics was a most fitting candidate for these junior positions of statistical assistants.

The second factor which is related to the first was that the National and Regional Examining Bodies, e.g. the West African Examinations Council (WAEC), agreed to examine statistics in the school and the GCE 'O' level examinations. Appropriate syllabuses had to be worked out, discussed and approved. The WAEC syllabus for the GCE 'O' level and West African School Certificate (WASC) statistics contains the following topics among others:

> collection of data; sources of data, planning of censuses and surveys, random sampling, design of questionnaires, tabulation and presentation of data; diagrams and charts, measures of location and dispersion for a set of numbers and for grouped data. Calculations of simple probabilities simple uses of the binomial and normal distributions. Weighted averages: index numbers. Simple regression and correlation.

This syllabus is run by all the English speaking West African Schools that offer statistics in the GCE 'O' level examination conducted by WAEC. At the moment these are schools in Nigeria, Ghana, Sierra Leone, Liberia and the Gambia. A similar syllabus also exists for the East African Examining Bodies.

Comparative figures showing the number of Nigerian candidates offering statistics and those offering other subjects like Mathematics, Economics and English Language in the GCE 'O' level conducted by the West African Examinations Council are given in the following table.

Nigerian Candidates Entry for GCE 'O' Level Examination

Subject	Year			
	1978	1979	1980	1982
Statistics	1,335	1,486	1,465	1,864
Mathematics	119,486	142,909	176,221	316,979
Economics	104,096	128,704	159,841	291,268
English Language	123,583	148,317	174,259	316,979

Source: West African Examinations Council

This state of affairs is indicative of the position in the entire continent South of the Sahara. Our experience even shows that Nigeria's embrace with school statistics is much warmer than that of other countries.

Facilities for Statistical Teaching in Schools: One of the visible constraints in the development of school statistics in Africa is the lack of qualified teachers. It is discovered that even in the schools that offer statistics in the GCE 'O' level examinations, only a few teachers can really teach the subject. One typical example is found in a Nigerian School where 240 pupils (i.e. 6 classes of 40 each) are taught the subject for two years, by only one statistics graduate teacher! Obviously, most of the time, the students fend for themselves reading whatever is available in the nature of books. And these books are either not there or very scanty indeed.

Lack of suitable text books for school statistics is one of the commonest problems threatening the teaching of school statistics. But this problem has a special dimension for Africa, in the sense that most of the statistics text books used in institutions in Africa are written and published outside Africa, mostly by non-Africans. When there is inadequacy of books in the countries of the authors, its effect on the countries on the receiving end can be imagined. Right now only one or two books specifically written by Africans are near enough to what is needed.

Prospects of School Statistics: It should be noted that with all the arguments in favour of introducing statistics as a school subject, some reservation is still expressed in some quarters. The principal argument is that strengthening the mathematics base of students in schools should be preferred to adding statistics as a subject which itself would need mathematics for further development. The result is that even with the introduction of statistics as a separate examination subject some statistics continues to be taught as part of school mathematics. It is, however, hoped that by the time recommendations of Vic Barnett's ISI (International Statistical Institute) sponsored Task Force on the teaching of statistics at school levels (TOTSAS) are implemented, many more countries in Africa will have firmly established the teaching of statistics in their schools.

Teaching Statistics in Higher Institutions in Africa

The gaining of political independence and the subsequent expansion of their economies have made it necessary for many African countries to seek to strengthen their government statistical services with respect to professional and middle level manpower. A few higher institutions were encouraged to start undergraduate and in some cases, sub-degree courses in statistics either as a separate subject or as a part of mathematics or economics.

The first International "Survey of Statistics Teaching in developing Countries" carried out by UNESCO in collaboration with the ISI in 1971/72 reported that only twelve University-level institutions in the following ten African countries had programmes leading to a qualification of up to a first degree level in statistics as at 1972:

> Ethiopia (Cairo and Alexandria), Ethiopia (Addis Ababa), Ghana (Legon), Ivory Coast (Ecole de Statistique d'Abidjan), Morocco (Institut Nationale de Statistique et d'Economic Appliquee, Rabat), Nigeria (Nsukka and Ibadan), Sudan (Khartoum), Tanzania (Dar es Salam), Uganda (Makerere), Zaire (Lumumbashi).

Among all these African institutions, there were in the 1971/72 academic session, about 230 students in their final year of the first degree programme or equivalent, and over 300 postgraduate students mainly in Cairo and Khartoum pursuing post-graduate diploma and M.Sc. programmes, with a total of 135 full time equivalent teachers.

Apart from the above mentioned institutions that offer degree level courses in statistics, there are other institutions that offer programmes, notably in mathematics and social sciences that contain substantial and often compulsory statistical components. Many institutions also offer service courses in statistics to students of social sciences, engineering, medicine and agriculture. The survey report recorded 33 such institutions in Africa as at the 1971/72 session.

As at 1981, the number of higher institutions in Africa offering degree-level courses in statistics has tremendously increased. For example, in Nigeria alone, the number of universities offering degree programmes in statistics has risen from two in 1972 to ten in 1981. These are at Benin, Calabar, Ibadan, Ife, Ilorin, Lagos, Maiduguri, Nsukka, Port Harcourt and Zaria.

Besides the increase in the number of university level institutions in each country, there is also a noticeable increase in the number of countries that now have institutions offering degree level courses in statistics. Universities in Botswana, Lesotho and Swaziland, and special institutions in the Cameroun, Algeria and Rwanda are recent examples.

Programme in Universities: The programmes in these universities particularly in Nigeria are offered in various forms. For example, there are departments of statistics in Nsukka, Ibadan and Ilorin, while in the other universities the statistics programme is offered in the department of mathematics.

In addition, Ife and Benin have some statistics courses located in demography/social statistics and economics departments respectively.

By and large, the B.Sc. statistics programmes of most of the universities, particularly those in the Anglophone countries, appear similar to the standard B.Sc. statistics programme of either the British or the American/Continental University. There is enough of mathematics - calculus and linear algebra - in the syllabus to help the students understand the basic probability theory required for a thorough grasp of statistical methods.

A typical syllabus for a B.Sc. (Statistics) would have courses built around the following topics:

> Probability and distribution theory, statistical inference, sample surveys, design of experiments, demography, actuarial statistics, statistical methods; non-parametric methods, multivariate analysis, econometrics, and sequential analysis, economic statistics, time series analysis, sources of official statistics, operations research and practical project.

The balance between theory and practice (Applied Statistics, an elusive term) varies among the Universities depending on the needs of the country, the orientation of the lecturers and the facilities available.

In some of the institutions, it appears that efforts have been made to modify the 'imported' programme to suit local conditions. Even then, as in the developed countries the standard B.Sc. statistics programme continues to be criticised as being inadequate for preparing graduates for careers in government. Recent discussion at the 9th Conference of Commonwealth Statisticians held in Nairobi in December, 1980 seems to suggest that this criticism is not limited to developing countries only. The B.Stat. programme run parallel with the B.Sc. at Makerere University is obviously another attempt to meet these criticisms.

A few universities: 3 in Nigeria, 2 in Egypt, one in Sudan, one in Ghana and one in Uganda now offer postgraduate programmes in statistics leading to Master and Doctoral degrees.

Programme in Colleges of Technology and Education: Apart from the universities per se, there are also particularly in Nigeria, many polytechnics, colleges of technology, and colleges of education that include statistics in their programmes. Prominent among this group of higher institutions in Nigeria are: The Polytechnics in Ibadan, and Kaduna, the Institute of Management and Technology in Enugu; the Colleges of Education at Owerri, Awka and Uyo.

The standard of statistics education offered in these colleges is indeed post-secondary if slightly sub-degree. The polytechnics and colleges of technology award Ordinary National Diploma (OND) in Statistics, while a few go up to the Higher National Diploma (HND) after four years of study. The qualifications

are very close to the B.Sc. level, but are usually practical in nature. Particular mention should be made of the programme of Colleges of Education. The graduates of these Colleges of Education are awarded a Certificate in Education with emphasis on the teaching subject studied. Those of them that study statistics (with mathematics usually) get enough of statistics to be able to teach the subject in secondary schools. Some colleges of education now offer additional university type degrees.

Francophone Centres: Institutions offering statistics in the French speaking African countries would for the purpose of this paper be classified as Training Institutions different from the standard University. We shall consider them in more detail in the next section.

TRAINING STATISTICAL STAFF

So far, the discussion has dealt mainly with the teaching of statistics as an academic discipline. Except at the lowest (school) level, teaching statistics for general education may also be seen in the light of preparing staff for professional (statistical) duties.

To meet the needs of governments, special training institutions have been set up particularly in the Francophone countries.

Before examining the training of this form, it is relevant to mention the type and nature of work expected to be carried out by statisticians and their subordinate staff. Broadly, the work falls into three categories:

(a) data collection, processing, analysis and dissemination carried out in government and quasi-government offices.
(b) research of the type carried out in formal research centres or similar organisations.
(c) academic, mainly teaching with possibly some research.

Corresponding to these three categories are three types of statisticians styled by Healy (1973) as intelligence statistician (a pseudonym for official statistician), experimental statistician and statistical mathematician respectively.

It is the training of official and experimental statisticians that engages the attention of many African countries. It is relevant to mention at this stage the intensified efforts of the United Nations Development Programme (UNDP) through the Economic Commission for Africa (ECA). At the instance of the conference of African Government statisticians (now succeeded by the Joint Conference of African planners, statisticians and demographers) a programme called the Statistical Training Programme for Africa (STPA) was started at the ECA in 1978. The main objective of the programme is to ensure adequate supply of trained statistical personnel through the coordination of the activities

of the institutions being helped by the UNDP.

The Centres in Ibadan, Legon, Lesotho, Botswana and Makerere have already been mentioned because they are run within university settings. The other centres are in Yaounde, Abidjan, Algiers, Rabat, Kigali, Dar es Salam and Paris. Details about these centres can be found in the ECA document Directory of Centres participating in the STPA.

Junior Cadre

Many National Statistical offices now offer internal training for enumerators, statistical assistants/clerks, who usually have GCE O-level or equivalent in mathematics as their basic entry requirement. Right from this level, economics and mathematics are offered along with statistics. The aim here is to train the assistants in the practical (not theoretical) aspects of their job. It is to be mentioned that any of the assistants with GCE O-level in statistics need not undergo this training.

Intermediate Level

In view of the traditional approach of the few existing universities in Africa and realising the need to produce enough statistical manpower, special middle-level training programmes were instituted in Accra, Ibadan, Dar es Salam, Rabat and Yaounde in the early and mid sixties. The programmes in Accra, Dar es Salam, Rabat and Yaounde were handled through independent training centres whilst the Ibadan programme was run as a special programme within the Faculty of Social Science of the University of Ibadan. Though rudiments of such courses as experimental design featured in some of the curricula, the intention was clearly on the training of what can be called middle-level economic (and to a lesser extent, social) statisticians, the prospective official statisticians.

All the centres had the support of the UNDP through the ECA and the UNESCO. The centre at Ivory Coast now also offers training at this level as an independent centre. In the English speaking countries the centres in Lesotho, Botswana and Nigeria offer certificate and diploma programmes (at intermediate levels) within the context of the programme of the University to which each is attached.

Centres in the French speaking countries offer programmes leading to Agent Technique and Adjoints Technique Diplomae. A centre at Lubango, Angola offers training for the Portuguese speaking countries.

There are also other training centres, not under the UNDP, which offer training at this level. An example is Ecole Nationale d'Economie Applique, Dakar, Senegal.

Curriculum for Intermediate Training: As earlier indicated, the curriculum of the centres was designed to train middle level official statisticians. Though the bulk of the graduates of these centres will by and large serve as such, some flexibility in the curriculum is now being attempted, because it is realised that the products of the centres may not all work in government statistics offices and that the more able ones among them may need to pursue higher studies.

For the Francophone centres a typical Agent Technique syllabus has the following as its core subject groupings: General studies comprising mathematics, numerical computation, French, English, economics and geography; basic statistics, applied statistics, computer science and practical project.

The Adjoint Technique syllabus has additional topics like statistical organisation and preparation; elements of civil and commercial law, development and planning statistical analysis, survey theory and practice.

The Anglophone syllabus falls broadly under Economics, including economic geography; mathematics, statistics theory and methods, applied statistics, practical project.

While statistics, mathematics and computer science are seen as core subjects in these programmes, the exact balance among them is still open.

It is, however, understood that the ECA is working towards the production of a guide syllabus flexible enough for adaptation at most African institutions.

Professional Level

All the five Francophone centres in Africa (Abidjan, Algiers, Kigali, Rabat and Yaounde) offer various degree equivalents in statistics, Ingenieur des travaux statistique (ITS), economic statistics, Ingenieur statisticien economistes (ISE), applied statistics, Ingenieur d'application de la statistique (IAS), and in one case information science, Ingenieur analystes information (IAI). The Centre Europeen de formation des Statisticiens, economistes de pays en Developpement (CESD) Paris was for a long time solely responsible for training Francophone Africans at the ISE level. It continues to do so alongside the growing centres in Africa.

In the English speaking countries, no special independent training centres exist as such, other than those already mentioned in the section 'Teaching Statistics in Higher Institutions in Africa'.

Curriculum for Professional Training: A typical ITS syllabus would have courses similar to those contained in the B.Sc. (Statistics) syllabus given earlier, but with more pronounced courses in economics and planning e.g. economic

theory, world economic systems, national accounts projects selection and evaluation theory. A lot of weight is attached to the practical work that normally follows classroom lectures. Lectures and practical work in the major subject areas are spread over the period of training according to some relative weighting, e.g.

mathematics 15%, statistics and demography 31%,
economics 28%, computer science 10%, languages 8%,
miscellaneous and case studies 8%.

Special Training: To produce further training for serving statisticians, post-graduate courses, seminars and workshops of varying duration are now being run in some institutions and centres.

For example, in Ibadan a one-year postgraduate diploma is being run primarily for serving statisticians who are non-statistics graduates.

PROBLEMS AND PROSPECTS

We conclude this review by highlighting some problems inherent in statistical education in Africa. An insight into the future is then undertaken.

Because the educational systems of the two main foreign language groups (Anglophone and Francophone) are different no excessive generalisations can be attempted. Nonetheless we examine the problems under the headings: curriculum and standards, staffing, facilities, organisational set up and manpower needs.

Curriculum and Standards

In the context of the work of official statisticians, the syllabuses of the Francophone centres at both middle and high levels seem to be broadly based. While statistics, mathematics and economics remain common topics in all the programmes, the type of degree awarded appears dictated by the emphasis. Thus we have ITS, and ISE being awarded within the same institution. This practice obviates the problem of finding a suitable name for a degree in which some depth in statistics has to be sacrificed for other subjects. However, with such other subjects as business, law, accounting being an integral part of the syllabuses, a kind of general degree would be implied. Perhaps this is why statistical education in the Francophone countries is not normally part of a university setting.

In the Anglophone countries, there seems to be some confusion between the middle level and higher (degree) level training. From the foregoing account, one can notice an attempt to 'professionalise' the training at the middle level. Though many of the universities mentioned attempt to practicalise their degree programmes, yet this practicalisation is not to the same extent and level as

in the Francophone institutions.

If such professionalisation were attempted within the existing university set up, many problems of logistics would arise. For example without a well established course system, a department of statistics trying to offer courses in national accounts may run into difficulties. With a course system, however, such topics can be offered in appropriate departments.

The Makerere experiment of running a parallel B.Stat. programme is yet to be evaluated for its advantages over a well designed B.Sc. In summary the basic question that still remains is that of determining how much of applied statistics topics (economic statistics, educational statistics, agricultural statistics) and miscellaneous topics (national accounts) should be taught in universities and training centres. On the other hand, it is relevant also to ask how far African statistics departments and training centres should go in teaching topics like probability theory, theory of sample survey and design of experiment as distinct from their applications. We do not pretend to have the answers. It seems clear, however, that more emphasis is likely to be placed on additional applied options even at the expense of the theoretical base. Suffice it to say that with continued discussion and practice, the right mix will appear at the appropriate time.

Staffing

With a normal statistics unit (department) offering standard statistics courses, lack of staff is a near universal problem. From our experience of universities in and outside Nigeria, we doubt if there is a single university in Africa which has enough qualified African staff to man its statistics department adequately. The problem becomes worse if departments are to teach some of the various applied options already mentioned. We have noticed in one or two instances when, even among the available staff, nobody was willing to teach some particular applied courses for reason of lack of either text books or interest.

We realise that a practice for combating these problems is to hire short term specialist staff from outside the institution itself. If well organised this may well be a way out. But it can be no substitute for tenured long term experienced staff.

It is pertinent to reiterate the help that UNDP has been giving through UNESCO and the ECA to many of our institutions. Many other multilateral and bilateral agencies have been assisting too. It appears that such help will be needed for some time to come.

Facilities

Lack of adequate text books and teaching materials has already been mentioned. The problem is all the more acute in areas where texts are needed to suit local conditions, for example official statistics (or at least topics in its methodology) and national accounts.

Other facilities that are usually not at hand are mainly computing (calculating) machines. The effect of this on what is taught is clearly obvious.

Organisational Set Up

We have noticed three or four organisational hybrids for teaching statistics through

(a) conventional statistics departments in universities or polytechnics;
(b) an existing university department like mathematics or economics;
(c) institute of statistics within university set up and;
(d) independent training institutions awarding their own diplomas, degrees, etc.

In the discussion under curriculum and staffing, mention has been made of the associated problems under each set up. It may be added that in some universities both situations (a) and (b) exist for one reason or the other. Whatever operates in any given institution in a country will certainly depend on many factors that sometimes have nothing to do with statistics.

Manpower

Perhaps a most disturbing fact is the slow output rate of statisticians from all the institutions. Measured against expected demand by 1987, which according to an ECA survey, is put at about 5,000 at middle level and 7,000 at professional level, the present output rate will hardly ever catch up with these demand figures. Even with the low output, not many of the statistics graduates from universities take up employment in public services statistical agencies. It may well be that the professionalised training centres are to be preferred to the university, as the graduates of the latter seem to have greater opportunities and wider avenues to choose from.

References

[1] Healy, M.J.R. (1973) The varieties of statistician. Journal of the Royal Statistical Society, Series A, 136, 71-74.

[2] 9th Conference of Commonwealth Statisticians. Summary of Proceedings, (1980) Nairobi.

[3] Directory of Centres participating in the statistical training programme of Africa. (Preliminary edition) ECA. (1981)

[4] Survey of statistics teaching in developing countries. UNESCO (1971).

STATISTICAL EDUCATION AND TRAINING IN DEVELOPING COUNTRIES OF ASIA AND THE PACIFIC

M N Murthy, Statistical Institute for Asia and the Pacific, Tokyo

INTRODUCTION

There have been considerable changes in the developing countries of Asia and the Pacific in the last three or four decades in the context of economic and social development in the post-war era. Statistics is increasingly being used in studying and finding solutions to problems of change. In this connexion, statistics has become an indispensable new technology for use in development efforts and assessment of gains due to such efforts. This has led to the governments of developing countries establishing and expanding statistical offices to meet the growing demands on statistics. Recently the semi-government and private sectors have also expanded reasonably well in these countries providing further opportunities for use of statistics. All these developments have invariably resulted in a growing need for trained and experienced statistical manpower to meet the challenges of change and development.

For a balanced development of statistical personnel it is important to develop both statistical education and statistical training facilities. It would have been desirable to have developed the statistical offices in the developing countries after establishing statistical education facilities to provide the necessary statistical methodology. However, the force of changes in economic and social development had necessitated the establishment and expansion of statistical offices even before statistical education and training facilities could be created and stabilized. The last two or three decades have seen considerable development of statistical education and statistical training facilities in the developing countries of Asia and the Pacific including establishment of regional statistical training centres such as the International Statistical Education Centre (ISEC) at Calcutta and the Statistical Institute for Asia and the Pacific (SIAP) at Tokyo.

Since its inception, SIAP has been interested in the training requirements of the developing countries of the region and has carried out surveys in the past for this purpose. Since statistical education and training are both required to produce effective statisticians with sufficient background knowledge of theory and practice, SIAP has recently conducted an enquiry on statistical education and training facilities in the developing countries of Asia and the Pacific. The results of this survey are reflected in this paper and the details are covered in the section on 'Reports from Countries'.

The main aim of this paper is to assess the present position regarding the extent of the availability of statistical education and training facilities in the developing countries of the ESCAP region and to discuss some of the main problems being

faced by these countries in developing the statistical manpower needed in the context of the development efforts. Early developments in the fields of statistical education and statistical training are briefly covered in the section 'Early Developments' and university education in statistics, regional and national training facilities are taken up in the sections 'University Education in Statistics', 'Regional Training Centres' and 'National Training Centres'; teaching of statistics in schools is briefly discussed in the section 'Concluding Remarks'. No attempt is made in this paper to review or evaluate the design and contents of the courses offered by the universities, institutes and training centres in the region.

Statistical education, training and research have been discussed at length by several authors and by committees set up by national and international organizations. Montenegro (1973) reviewed the efforts at promoting statistical education and training and referred to reports of various committees and conferences in this connexion. Bjerve (1976) in his presidential address at the Warsaw session of the International Statistical Institute brought into focus the gap that had developed between theoretical and practical statisticians and the problems relating to the inceasing specialization in various subject-matter areas in statistics. Since then, the SIAP, the American Statistical Association (ASA) and the ISI had set up panel discussion, conference and committee to discuss these problems, and they have come up with many useful recommendations (Murthy and Khan, 1977; ASA, 1978; Duncan and Durbin, 1980). ASA (1978) includes many useful papers on statistical education and training and particular reference may be made to the papers by Fellegi, Malinvaud and Murthy. United Nations (1979) has discussed human skills and training in the context of statistical work. Finker and Washabaugh (1977) have described many programmes available for training domestic and foreign government statisticians in the U.S.A.

The ISI and the UNESCO have been engaged in the promotion of statistical education, whereas the United Nations and its specialized agencies such as FAO, WHO and ILO, have been concerning themselves with the promotion of statistical training and standardization in data production. This division of work has evolved over the past 30 years without appreciable co-ordination of the activities, leading to some problems in relating statistical education to work requirements, and in turn giving rise to the situation referred to above. This is further discussed later in this paper.

EARLY DEVELOPMENTS

Though statistics has a long history in its practical aspect in the sense that data collection and compilation had been of interest to administrators from the very early times, statistical methodology is relatively of recent origin and owes its development and promotion to persons like Karl Pearson, Ronald Fisher,

Jerzy Neyman and P.C. Mahalanobis. In the Asia and Pacific region, the promotion of statistics and the development of statistical education and training have taken place in most of the developing countries during the last three decades of the post-war period. India has had a longer association with statistics and its development, and ranks among the relatively more developed countries in the field of statistics.

In India, there is a record of a statistics committee being set up in 1862 to propose modern statistical forms for collection and compilation of statistics in the fields of trade, finance, education, agriculture etc. which later led to the publication of statistical abstracts in 1868 (CSO, 1970). With the establishment of the Indian Statistical Institute in 1931, statistics became accepted in India and the universities gradually started offering courses in statistics, initially as part of departments of mathematics and later as separate departments of statistics. This process of statistical development was strengthened by the founding of the Indian Agricultural Statistics Research Institute in 1947 and the Central Statistical Organization in 1951. The development of statistics in India owes much to the dynamic leadership and sustained efforts of P.C. Mahalanobis.

The Statistical Office of the United Nations and the Statistical Sections and Divisions of the Specialized Agencies such as FAO, ILO, WHO and UNESCO have done commendable work in getting statistics accepted in the developing countries and in promoting and effecting statistical development by organizing international technical assistance programmes. Standardization of concepts, definitions and tabulation plans in population, housing, agriculture and industrial censuses has been advanced by the United Nations with the co-operation of all the countries including the developing ones. The United Nations Statistical Commission and its Sub-Commission on Statistical Sampling established in 1947 have done much to promote and develop statistics in the developing countries of the world, (United Nations 1947-51). Even now the United Nations and its specialized agencies are active in this field and there is currently the National Household Survey Capability Programme (NHSCP) of the United Nations which aims at setting up continuing household surveys in the developing countries of the world.

In this region, the ESCAP has been promoting and developing statistics in the developing countries through the organization of working groups, seminars and technical meetings on various aspects of data production and utilization as well as through technical assistance to these countries. In some countries, expansion of statistical data collection and compilation grew faster than the development of statistical education facilities leading to the employment of personnel lacking in suitable background for actual statistical work. Further, the univer-

sities initially started teaching statistics as part of mathematics department courses and naturally emphasized the theoretical aspects of statistics with a mathematical orientation at the expense of statistical operations and applications. This led to the gaps between theory and practice and between theoreticians and practitioners mentioned in the Introduction. This imbalance in statistical education and work requirements, which developed in the early period and persists even now, has necessitated the creation of national and international training facilities to remedy the situation. Recently there has been considerable constructive discussion on the problems of statistical education and statistical training and the current situation is discussed further later in this paper.

The situation mentioned above had been realized by the United Nations (1949) and by the ISI (1951) in the early stages. For instance, the Statistical Commission in its report in 1949 mentioned that it recognized there was a shortage of adequately educated and trained statisticians in many countries and that the burden of statistical work in public and private institutions was large, would increase and would have to be met by the education and training of more and better statisticians. The ISI has taken many steps to strengthen statistical education, one such step being the setting up of the International Statistical Education Centre at Calcutta in 1950. Also, the ISI has been holding round-table conferences on teaching of statistics and meetings of Directors of International Statistical Education Centres in the recent years. However, the present International Conference on Teaching Statistics being held at Sheffield in August 1982 under the sponsorship of the ISI is the first of its kind, as there has been no such conference so far in the world.

UNIVERSITY EDUCATION IN STATISTICS

Even in the early years of development of statistical education in universities considerable attention has been given to this topic as is evidenced by the report of the Council of the Royal Statistical Society (1947) and the report of the Commission on Teaching of Statistics of the Institute of Mathematical Statistics (1948). More recently, the ISI (1969) organized a round-table discussion on university teaching of statistics in developing countries. UNESCO (1974) has surveyed the situation regarding teaching of statistics in developing countries. The problems of teaching statistics at university level also get reflected in many seminars and conferences including the ISI biennial sessions. The ISI Committee on Integration of Statistics had also considered the problems of statistical education in universities (Duncan and Durbin, 1980). A useful discussion of teaching statistics in universities has been given by Daniels (1975). The periodical American Statistician has a teachers' corner to discuss problems of teaching statistics and the periodical Teaching Statistics provides a very good forum for teachers to exchange ideas on course designs, course contents and methods

of teaching. Also a number of teaching aids in the form of cassettes, slides, movies and video-tapes are becoming available, and their proper use will help increase the effectiveness of teaching statistics.

In the Asia and Pacific region, except India where university education in statistics developed much earlier, the development in this connexion has been of a recent origin in many developing countries and in some countries the university facilities are still non-existent or available only at a minimum level. Historically, statistics came under the Department of Mathematics in many universities and this naturally led to an emphasis on theoretical aspects of statistics with a mathematical orientation as mentioned earlier. Though statistics is basically an applied subject, it was and is being taught in many universities as a branch of mathematics. The graduates coming out of universities become the teachers of statistics and research workers, and this inbreeding ensures continuation of the stress on the theoretical aspect in teaching of statistics and research work in this field.

University education in statistics concentrates on the so-called mathematical statistics and even when applied subjects such as sample surveys and design and analysis of experiments are taught, the emphasis is on their mathematical orientation. However, after graduation most of the students generally find employment in statistical offices of government and semi-government agencies and in business and industry, where they have to design and carry out the statistical operations of a statistical investigation including data collection, compilation, analysis and interpretation, still not generally covered at all or covered in the universities with considerable abstraction.

The above situation in this region has been noted at a very early stage of development of statistical education in universities by Madhava (1951) when he said we had failed to take adequate care of practical work in statistics. The United Nations (1981, pp.42-43) states that in many countries the knowledge acquired in universities, both in subject-matter fields and in statistics, was often too general and abstract to be immediately well adapted to the needs of statistical agencies. It may be mentioned that it is desirable to take the general objectives of statistical education as developing a statistical sense with a view to understanding natural and physical phenomena through statistics and facilitating working with statistics for decision-making and teaching the methodology needed to produce appropriate statistics economically and to analyze and use them meaningfully.

Under pressure from the employers of statistical graduates such as government and non-governmental agencies, the universities in this region are attempting to reorient their courses towards application of statistics with some emphasis on practical aspects. This effort, though commendable, has generally been

with uneven benefits mainly because of lack of teachers with background and experience in practical aspects of applied statistics. Further, skill in statistical operations such as evolving definitions, classification, designing of questionnaires, data collection, editing, coding, etc. can only be acquired through on-the-job training which is not possible in the universities due to lack of funds and facilities. Thus, if practical orientation of the university curriculum is to succeed, the facilities for exposure of the students to actual work should be provided through collaboration with statistical agencies.

Notwithstanding what has been said above, it must be noted, as mentioned by Rao (1979), that while the discussion of practical problems in teaching statistics was essential, it must not be done at the expense of the theoretical and mathematical contents of the courses. Though it is important that a statistician should be equipped with the practical skills to enable him to choose the appropriate methodology for analysis of given data, he should have at the same time the theoretical knowledge to formulate a new statistical problem for study if the existing methodology is inadequate in any given practical situation. Thus, what is needed is that we produce statisticians with the proper knowledge of theory and a flair for applications and give them opportunities to improve their skills during the course of their work.

Another aspect of statistical education in universities is the teaching of statistics to students of various disciplines, as statistics has an application in many fields. Here, the promotion of statistics can be ensured only if the subject is taught in such a way as to bring out forcefully the contributions of statistical knowledge to the development of the main subject of interest to the student which can be economics, demography, sociology, etc. Thus, practically oriented courses in statistics with a particular subject bias are needed in this situation instead of abstract theoretical courses.

The advent of computers and their accessibility to the university statistics departments are having a profound effect on the teaching of statistics in the universities. It is important that the universities should take account of the development in computer technology and capabilities and build this knowledge into the university curriculum. Many universities have already done this in this region and there is a keen desire to do so among the others also subject only to the necessary resources becoming available. Computer science is acting as a very good catalytic agent in galvanizing the development of statistical education in universities and this situation should be understood and utilized well in the interest of further development of statistical methodology and its applications.

REGIONAL TRAINING CENTRES

From what has been said in the previous section on university education in statistics, it is clear that the statistics graduates have generally adequate theoretical background in the subject, but the gap between statistical education and work requirements exists in spite of the effort to reorient the university curriculum towards applied aspects of statistics. This necessitated the creation and strengthening of facilities to provide the skills essential in statistical operations before the statistical graduates could be fruitfully employed in statistical work involving data production, analysis and interpretation, (ISI, 1951). According to the SIAP survey, though 10 out of 26 responding countries reported availability of national training facilities, 20 mentioned that the existing training facilities in the region were not adequate to remedy the situation, notwithstanding the fact that two international statistical training centres have been set up in this region to meet the training needs of the developing countries. They are the International Statistical Education Centre (ISEC) at Calcutta and the Statistical Institute for Asia and the Pacific (SIAP) at Tokyo.

The activities of ISEC are briefly discussed in ISEC (1977, 1982). The ISEC was set up in Calcutta in October 1950 for the training of statistical officers from the countries of the Middle, South and Far East. The Centre was sponsored by the International Statistical Institute and the Indian Statistical Institute under the auspices of and with financial support from the UNESCO. This arrangement was discontinued in 1952 but the work of the Centre was continued with the help of grants from the Government of India to the Indian Statistical Institute and also to the International Statistical Institute which made available teachers from abroad for the ISEC to supplement the faculty provided by the Indian Statistical Institute. The main purpose of the Centre is to provide courses in theoretical and applied statistics at various levels to selected participants.

The teaching at the Centre is mainly undertaken by the staff of the Indian Statistical Institute at Calcutta and this faculty strength is supplemented by visiting teachers provided by the International Statistical Institute, the United Nations and other international organizations. The Government of India collaborates with the Indian Statistical Institute in conducting the courses of ISEC by providing facilities for training in official statistics at the Central Statistical Organization and other ministries of the government.

Most of the trainees at the Centre are financially supported by the Government of India training grants under the technical co-operation scheme of the Colombo Plan, the Special Commonwealth African Assistance Plan (SCAAP) or other schemes. Some participants are under the fellowships provided by the United Nations and other international organizations. A number of trainees are also supported by their respective national governments or employer organizations. Originally

the ISEC course was a six-month regular course, but some students were given the choice of continuing in the next term on some specialized topics. From 1954-55, the duration of the regular course was made nine months, but the choice of continuing in a subsequent term was no longer given to the participants. From 1966-67, the regular course has been made to be one of ten months' duration, thus the course length has been extended from the initial six months through the intermediate nine months to the present ten-month term. Also the Centre organizes special courses in collaboration with the CSO at places other than Calcutta but within India on specific topics, and these courses may be of varying duration.

The course design of the ten-month regular course consists of four phases of varying duration. During the first four months (June - September) at Calcutta, preparatory courses covering mathematics, practical work in probability and statistical theory and statistical projects such as sample surveys and data processing are conducted. This is accompanied by courses on economic statistics and national income statistics. During the second phase of six weeks (October - mid-November), the participants undergo training in official statistics organized by the Central Statistical Organization of the Government of India at New Delhi. The third phase covering mid-November to January and the fourth phase covering February - May are conducted at Calcutta. During the third phase of training, statistical methods, design of experiments, economics and economic statistics, sample surveys and demography are covered. During the fourth phase, each participant specializes in some field of study such as sample surveys, mechanical data processing, demography and vital statistics, statistical quality control and operations research, (ISEC, 1982).

The participants are evaluated on the basis of written tests and computations involved for analyzing data needed for the solution of practical problems. For the specialized training towards the end of the course, students are sometimes evaluated on short dissertations or reports on project work or through work assignments such as writing a computer programme. Candidates passing the final examination are awarded certificates of merit, while those who satisfactorily attend and complete the course but do not pass the final examination are awarded certificates of attendance.

For a detailed account of activities of SIAP, reference may be made to SIAP (1980) and Nair and Murthy (1971, 1973). The SIAP (formerly Asian Statistical Institute) was established in 1970 by the Governments of the countries of the Asian and Pacific region with the support of the United Nations Development Programme (UNDP) and the Government of Japan as an autonomous organization operating under the aegis of the Economic and Social Commission for Asia and the Pacific (SIAP). The Institute has completed two five-year phases and is

presently entering its third five-year phase. The purpose of the Institute is to meet the training needs of the developing countries of this region in the field of statistics. At present the Institute conducts annually the following courses:

(i) General Course for six months for junior statisticians;
(ii) A three-month course on Automatic Data Processing for statisticians;
(iii) Two Advanced Courses or Seminars for four to six weeks on specialized topics for senior statisticians; and
(iv) Country or Sub-regional Courses on specific topics.

The fellowships for the General and ADP Courses are provided by the Government of Japan and those for the Advanced Courses and Seminars by the Government of The Netherlands. The faculty of the Institute is international in its character and is provided by the United Nations and the faculty strength is supplemented from time to time by experts made available to the Institute on non-reimbursable basis by host and other donor governments and international organizations. In the case of Country Courses the cost is generally shared by the Institute and the concerned countries.

The course design of the six-month General Course consists of a mandatory part and an elective part. The mandatory part generally covers sampling methods and elements of sampling, food and agricultural statistics, economic statistics, national accounting statistics, social and demographic statistics, statistical operations and automatic data processing. During the last three months of the course, in addition to some of the mandatory subjects, each participant specializes in one or two elective subjects. The subjects offered as electives vary from year to year, and in 1981-82, the two electives offered were sampling methods and demographic analysis. Greater choice of subjects is offered in some years depending on the availability of adequate faculty strength and interest among the participants. The course also includes field visits to observe various aspects of data production in Japan.

The emphasis in the General Course is on methodological, practical and applied aspects of statistics most relevant to the statistical work in government offices. Originally, it was a ten-month course during the first five-year phase and the duration of the course was reduced to six months since the beginning of the second five-year phase due to budgetary considerations, various demands on faculty time and other factors.

The fellows of the General Course, who generally have at least a basic degree, are first given diagnostic tests at the beginning of the course to evaluate their knowledge in basic mathematics and economic analysis and those who do not have adequate background in these two subjects are given supplementary courses on these two topics. In the case of the Advanced Course or Seminar, generally

the fellows are senior statisticians with good academic background and considerable practical experience and hence the group turns out to be more homogeneous in their background and interest. Similarly, in the case of the Country Course on a specific topic, it is generally found that the background and experience of the fellows are basically of the same order. It may be mentioned that in an international course there is bound to be some amount of heterogeneity among the participants in respect of their educational background and work experience, and this should be kept in view in determining the course design and contents.

It is general practice in the Institute to provide the participants with a lecture outline or lecture notes and other relevant training documents well in advance of each lecture. Thus, over the last few years, the faculty members have prepared lectures notes in their areas, and these have been put in bound volumes for use in different courses. The notes generally include a brief description of the concepts, definitions, methods, operational procedures, illustrative examples, references and workshop assignments. These notes are kept in constant review in the case of the General Course and periodically updated using new data that become available. The Institute has produced lecture notes in the following subject areas: Statistical Methods, Sample Surveys, Agricultural Statistics, Demographic Statistics, Social Statistics, Economic Statistics, Economic Accounting Statistics, Statistical Operations, and Automatic Data Processing. Further, the documentation of the Advanced Courses are compiled together and kept in the library for the use of the faculty and the fellows. These lecture notes are available to the government statistical offices for their use in their training courses.

SIAP has recently initiated a UNDP-supported programme for the training of trainers of field enumerators and office processing staff. With financial support of the UNDP, SIAP collaborates with participating countries in producing relevant training manuals and training local trainers in the use of the training manuals and in training methods and techniques. The programme is underway in Indonesia and the Republic of Korea, and will eventually be expanded to include other interested countries. Its rationale is the belief that the quality of statistics produced is highly dependent on the motivation, knowledge and skill of field and office operatives.

In the case of the General Course, the participants' work is evaluated on the basis of a continuing system of written tests and assessment of completed workshop assignments. At the end of the course, the results of the evaluations are summarized in the form of consolidated scores on the basis of which certificates are issued to the participants. Those who satisfy the course requirements are given a certificate of successful completion and those who do not are given

a certificate of attendance. Also, the fellows are given an opportunity to make their own assessment of the course through the use of a structured questionnaire. On the basis of the results of this enquiry, the course design and contents of the General Course are reviewed and wherever considered necessary changes, if any, are made for the next general course. This practice of evaluating the participants by the faculty and of the course by the participants is carried out in the case of the Country Courses also.

NATIONAL TRAINING CENTRES

As mentioned earlier, training centres have been necessitated by the gap that persists between statistical education and work requirements. National governments are trying to reduce this gap by creating and strengthening training facilities, which is generally a time-consuming and costly process. From the SIAP's survey, it is clear that only 3 out of 26 reporting countries felt that their training facilities were adequate for their purpose, while others expressed that their training facilities at national level can be realized when the rapid expansion of statistical services over the years is kept in view. Further, the two international statistical training centres in the region can at best meet only partially the training needs of the developing countries of the region.

The need for training at primary level is keenly felt whereas the international training centres generally aim at training statistical personnel at intermediate and professional levels. However, the training of trainers programme of SIAP tries to help this situation by training teachers to conduct training courses at primary level. Though many countries mentioned that their training facilities were inadequate, it is interesting to note that 10 out of the 26 countries in the SIAP survey reported creation of training facilities to train their staff. What is needed in these countries is to strengthen these facilities through proper training of teachers, supply of lecture notes, other training documents and teaching aids such as overhead projectors, slides, movies, etc. and rationalizing course design and content. In the case of countries where no facilities exist, the possibility of having basic courses to train primary and middle level personnel should be explored. The international centres can help the countries in establishing and strengthening their training facilities to serve their requirements at primary and intermediate levels. The problems of recruiting, training and retaining of statisticians in government service in the region were discussed at the Seminar on Statistical Organization held at Wellington, New Zealand, (Miller, 1979).

In the case of very small countries the possibility of developing a sub-regional training centre may also be explored for the purpose of conducting basic courses for sub-professional staff, and the requirements at professional level can be met through training in international regional centres.

TEACHING OF STATISTICS IN SCHOOLS

Of the 25 countries responding to the SIAP survey, 14 countries are already teaching statistics in schools and the other 11 countries are not doing so. In countries where statistics is taught in schools, it is generally introduced at the secondary school level, though a few countries have reported introducing graphical presentation at primary level. Further, the coverage is usually confined to descriptive statistics consisting of charts and graphs, frequency distributions and measures of central tendency and dispersion. However, a few countries are taking up some statistical analysis at the higher secondary school level covering correlation and regression analysis.

Though teaching of statistics in schools is of fairly recent origin, this has been strongly recommended much earlier by the Committee on Teaching of Statistics of the Institute of Mathematical Statistics (1948), Council of Royal Statistical Society (1947) and Mahalanobis (1951). ISI had discussed this topic at length at the Fourth Round Table Conference on Teaching of Statistics (Breny, 1976).

It may be stressed that if teaching of statistics in schools is to be taken seriously, standard textbooks would have to be developed through the organization of task forces and these would have to be tried out in some selected schools with the help of trained teachers. Needless to say that the textbooks should be simple, illustrated with pictures and numerical examples and should bring out clearly the role of statistics in data production and inference. A concerted effort should be made to train the teachers through modification of the syllabus of the degree courses in education in the colleges and the universities. These points have also been brought out in the responses of some of the countries to the SIAP survey.

REPORTS FROM COUNTRIES

In the past, as part of its activities, the Statistical Institute for Asia and the Pacific (SIAP) has carried out surveys of training needs of the developing countries of the ESCAP region and of the extent of utilization by the countries of the training provided by the Institute through its various courses. These surveys had concentrated on the training aspects involved in producing an effective statistician as the Institute is basically a training institute, striving to provide the skills necessary for statistical work in the developing countries. Also, the Institute periodically tries to find out the specific requirements of each country in the region by correspondence or through the visits to selected countries by the faculty members as part of the activities of the Institute.

Recognizing that statistical education and statistical training are equally important to produce efficient statisticians, the Institute has been interested in both these fields to get a proper perspective of the statistical education and training facilities

available in the developing countries of the region. For this purpose and to support the First International Conference on Teaching Statistics to be held at Sheffield, United Kingdom, during 8-13 August 1982, the Institute carried out a survey among the developing countries of this region. There is considerable interest in the outcome of this effort, as it would help in orienting its future programmes and plans to better meet the requirements of the countries of the region. Further, it is hoped that the result of the survey will provide a solid basis for further discussions of the problems of statistical education and training at national and international level in the region. For the purpose of the survey the following four forms were used:

(a) Statistical education at college level;
(b) Training in statistics;
(c) Statistical education at school level; and
(d) Overall comments and suggestions.

In the survey, statistical education at college and school levels and statistical training were separated to facilitate supply of required information, as the sources of information on them are usually different in the countries of the region. Before actually conducting the survey, the forms were first sent to India, Indonesia and the Phillipines for pretesting, which proved to be very useful. On the basis of this experience, the forms were finalized and mailed to all the developing countries of the region. Twenty-six countries have responded to the survey and only a few are yet to respond.

An overview of the results of the SIAP's survey is given in this section and detailed summaries of the reports from the countries so far received are given in Annex 2. In the table on statistical education and training facilities of Asia and the Pacific given in Annex 1, summary of the responses from the countries on the following are given: university facilities, computer facilities, demand for statisticians over supply, gap between education and work requirements, training facilities and teaching of statistics in schools.

If there is any basic degree course in statistics offered by any university in the country, the country is said to have a university facility. However, if the available facilities are considered insufficient to meet the requirements in the sense of lack of qualified teachers, number of students educated being less than the demand for statisticians, etc., the facilities are considered inadequate. If the educatioin facilities meet fully the requirements of the country in quality and quantity, the country is said to have adequate educational facilities. Similarly, in the case of computer and training facilities also the countries are classified as those where they are adequate and those where they are not adequate. In the case of a gap between education and work requirements, the countries are classified as those which are stating that the gap exists and those which

do not. And for countries where no educational or training facilities exist this information is reported as not applicable.

From the SIAP survey, it emerged that 16 out of 26 responding countries had university facilities for teaching statistics and of these only 7 mentioned that these facilities were adequate. Of the 16 countries reporting university facilities, 14 mentioned that there was a gap between statistical education and work requirements necessitating further training of university graduates before their assignment to specific statistical work. Further, 21 out of the 26 countriers reported that the demand for statisticians was more than the supply with the available educational system. Regarding computer facilities, only 4 countries in the region reported that their facilities were adequate, whereas the others had mentioned that they either had no such facilities or that their facilities were inadequate and that they had to be strengthened.

Regarding training facilities, 10 of the 26 reporting countries mentioned that they had national training facilities in the field of statistics and the remaining 16 reported that they did not have formal training facilities. However, 23 of the reporting countries including 7 of the countries having training facilities reported that such facilities were inadequate. In fact, only 3 countries in the region, namely India, Hong Kong and Singapore, reported that they had adequate training facilities. In the case of teaching of statistics in schools, the 26 reporting countries were divided 15-11 between those teaching statistics in schools and those not teaching statistics at school level.

The general impression one gets from the SIAP survey is that statistical education and training facilities are gradually emerging in the countries of the region taking into account also the necessary computer facilities. In spite of the problem of not having trained and experienced faculty, lack of adequate teaching aids, books, etc., the countries have progressed in introducing statistics in graduate and post-graduate courses in the universities. The importance of statistical computing is well realized and appreciated and there is keen interest in expanding computer facilities in the universities in the context of statistics courses. Further, there is an increasing attempt to make the university courses in statistics oriented towards applied aspects of the subject. However this attempt has not fully succeeded due to various reasons including lack of teachers trained in practical aspects of statistical work. Thus, many countries reported that the gap between statistical education and work requirements exists and hence the countries have gradually tried to remedy this situation by creating and strengthening training facilities in their countries as well as in the region. Though statistics is taught in schools in many countries of the region, a number of problems regarding training of teachers, textbooks, etc. are yet to be discussed and resolved.

FUTURE DIRECTIONS

The complexities of modern economic and social life create a demand for information; and the advancement in computer technology and the availability of inexpensive, powerful computers make the creating of an adequate information base technically possible. Future needs therefore relate to our role as initiators, creators and users of the data base; that is, to:

(a) statistical awareness and motivation at policy levels in government and universities leading to the allocation of high priority and resources for the production and wide utilization of an adequate statistical data base;

(b) development of the quantity and quality of skilled manpower needed for the production and utilization of the data base; and

(c) education of users at all levels, particularly in government and business but also of the common people who need to understand and use quantified information for managing the affairs of ordinary life and even for intelligent understanding of news and events.

While (b) and (c) clearly give rise to needs for education and training, they are, at least in the short term, dependent on (a). If governments or universities do not recognize the central importance of quantitative thinking and statistics, they are unlikely to give adequate priority to statistical education in schools and universities. However, even a modest increase in the priority accorded to statistical education and training, particularly in schools, will eventually produce a more articulate and sophisticated demand for statistics and consequently an increase in priority and resources. Future efforts might therefore be concentrated in three main directions:

(i) increasing the statistical awareness and priority at policy levels in government;

(ii) widening of statistical curricula in universities to include practical aspects of applied statistics and promotion of the study of statistics; and

(iii) general statistical education in high schools.

More specifically, the following suggestions are made in respect of future needs in the field of statistical education and training in developing countries of Asia and the Pacific:

(i) initiative by national government statisticians to demonstrate the importance of statistics by making full use of existing statistics, particularly their analysis and presentation in forms which meet or partly meet the needs of official and other users. This inevitably increases awareness of the importance of statistics, sharpens the users' perceptions of their statistical requirements and leads to a higher priority being accorded to statistics;

(ii) Conduct of regular seminars by government statisticians and universities to introduce government and non-government users to the availability and uses of statistics and statistical methods;

(iii) Active promotion of statistics as a course of study and a career interest by official and other statisticians through campus interviews and promotional talks and the distribution of publicity brochures in universities and high schools;

(iv) Review of university curricula to reflect the practical aspects of applied statistics without diluting the conceptual and theoretical foundations necessary at the university stage;

(v) Revising the university curricula to include statistical computing as an essential part of the courses;

(vi) Training of teachers in universities and making computer specialists and practitioners available to universities to teach effectively statistical computing and practical aspects of applied statistics;

(vii) Achieving a balance in research work by encouraging research scholars to work on practical as well as theoretical problems, possibly through exchange of personnel between academic institutions and statistical offices of government, semi-government and other agencies;

(viii) Developing of national capability to train at least the personnel at primary and middle levels in statistical offices of government and other agencies;

(ix) Regional institutions such as SIAP helping the national efforts mentioned in (viii) through provision of training materials and aids in the form of lecture notes and structured training manuals, and with consultancy services;

(x) Encourage statisticians and statistical offices to work on methodological problems having relevance to present or future work of such offices and to publish their results in established journals and to attend seminars, professional meetings, etc.;

(xi) Collaboration between the United Nations Statistical Office, the United Nations Specialized Agencies, the International Statistical Institute and other relevant bodies in bringing the theoreticians and practitioners together for the development of statistics, and a much greater emphasis by all these agencies on the needs of the developing countries;

(xii) Professional journals reviewing and modifying their editorial policy to encourage publication of practical results of methodological interest and formulation of practical problems for possible solution by theoreticians, if necessary, through enlarging their editorial boards and referee panels

to include practitioners and subject-matter specialists to ensure publication of papers with a practical orientation;

(xiii) Strengthening of computer capabilities of the universities and statistical offices with a view to utilizing such facilities in teaching, training and practical work in an effective manner;

(xiv) Promotion of teaching of statistics in schools as a mandatory subject, at least, from the secondary school level;

(xv) Review and standardization of textbooks and teaching methods for teaching statistics at school level;

(xvi) Organization of national and international panel discussions, seminars and conferences on various aspects of statistical education and training; and

(xvii) Special attention to be given to the countries in the Pacific region by strengthening the existing University of the Pacific at Suva, the Papua New Guinea University at Port Moresby, the University of Samoa at Apia and by creating sub-regional facilities for training of personnel at sub-professional level.

CONCLUDING REMARKS

Though it is made out in the paper based on the SIAP survey that there has been a gradual progress in the development of statistical education and training facilities in developing countries of Asia and the Pacific, it must be stressed that there is no room for any complacency, as the facilities created are far from adequate in many countries and such facilities are non-existent in some countries. It is expected that the national and international agencies concerned with statistical development in the region would be interested in the results of the SIAP survey and the views expressed in this paper. In the interest of statistical development, it is urged that these agencies should carefully consider the suggestions given in the section 'Future Directions' in this field with a view to taking up necessary steps to implement these suggestions as early as possible.

References

[1] American Statistical Association (1978) Conference on Transfer of Methodology between Academic and Government Statisticians; Washington, D.C.

[2] Bjerve, P.F. (1976) Presidential address presented to the fortieth session of the International Statistical Institute; Bulletin of the ISI, XLVI, (1), 41-48.

[3] Breny, H. (ed) (1976) The teaching of Statistics in Schools; Proceedings of the Fourth ISI Round Table Conference on Teaching of Statistics, The Hague.

[4] Central Statistical Organization (1970) Statistical System in India; Department of Statistics, Government of India, New Delhi.

[5] Council of the Royal Statistical Society (1947) Report on the Teaching of statistics in universities and university colleges; Journal of the Royal Statistical Society, 110, 51-57.

[6] Daniels, H.E. 1975) Statistics in universities - a personal view: The Address of the President (with proceedings); Journal of Royal Statistical Society, Series A, 138, 1-17.

[7] Duncan, J.W. and Durbin, J. (1980) Report of the International Statistical Institute Committee on the Integration of Statistics; International Statistical Review, 48, 139-168.

[8] Fellegi, I.P. (1978) Some thoughts on training needed for government statistical work; Conference on Transfer of Methodology, 43-52, American Statistical Association, Washington, D.C.

[9] Finker, A.L. and Washabaugh, B. (1977) Education and training of domestic and foreign government statisticians in the United States; Bulletin of the International Statistical Institute, XLVII, (1), 447-469.

[10] Institute of Mathematical Statistics (1948) The teaching of statistics - Report of the Teaching of Statistics; Annals of Mathematical Statistics, 19, 95-115.

[11] International Statistical Education Centre (1977) Report on the Activities of the ISEC; Meeting of Directors of International Statistical Training Centres, Calcutta.

[12] International Statistical Education Centre (1982) Prospectus for the year 1982-83; Indian Statistical Institute, Calcutta.

[13] International Statistical Institute (1951) Statistical education programme; Report of the Statistical Education Committee; Bulletin of International Statistical Institute, XXXIII, (1), 116-137.

[14] International Statistical Institute (1969) Round Table Discussion on the University Teaching of Statistics in Developing countries; Review of the ISI, 37, (1), 86-98.

[15] Madhava, K.B. (1951) Problems of teaching statistics in South-East Asia; Bulletin of ISI, XXXIII, (V), 313-322.

[16] Mahalanobis, P.C. (1951) The role of mathematical statistics in secondary education; Bulletin of ISI, XXXIII, (V), 323-334.

[17] Malinvaud, E. (1978) University teaching for future official statisticians; Conference on Transfer of Methodology, 53-67, American Statistical Association, Washington, D.C.

[18] Miller, J.G. (1979) Status, recruitment, training and retention of statistical personnel in government office; mimeographed, Seminar on Statistical Organization, ESCAP, STAT/SSO/7.

[19] Montenegro, T.H. (1973) Recruitment, preparation and status of statistical personnel; Inter-regional Seminar on Statistical Organization, Ottawa, Canada.

[20] Murthy, M.N. and Khan, T.M. (eds) (1977) Proceedings of the Panel Discussion on Communication and Cooperation between Theoreticians and Practitioners; Occasional Papers No.6, Asian Statistical Institute, Tokyo.

[21] Murthy, M.N. (1978) Statistical education, training and research and their utilization; Conference on Transfer of Methodology, 68-93, American Statistical Association, Washington, D.C.

[22] Nair, K.R. and Murthy, M.N. (1971) Asian Statistical Institute; Journal of the Society of Agricultural Statistics, XXIII, (2), 152-162.

[23] Nair, K.R. and Murthy, M.N. (1973) Evaluation of the courses and fellows' performance; Proceedings of the Meeting of the Directors of International Statistical Training Centres held at Szentendre, Hungary, 215-226.

[24] Rao, C.R. (1979) Perspectives in statistics; Presidential address at the 42nd session of the International Statistical Institute, Manila.

[25] Statistical Institute for Asia and the Pacific (1980) 10th Anniversary 1970-80; SIAP, Tokyo.

[26] United Nations (1947-51) Reports of the Sub-Commission on Statistical Sampling to the Statistical Commission; E/CN/, 3/37, 52, 83, 114, 140, United Nations, New York.

[27] United Nations (1949) An international programme of education and training in statistics; United Nations, E/CN/, 3/56, New York.

[28] United Nations (1981) Handbook of Statistical Organization; Volume I, Studies in Methods, F, 28, New York.

[29] UNESCO (1974) A survey of statistics teaching in developing countries; Paris.

ANNEX 1

STATISTICAL EDUCATION AND TRAINING FACILITIES
IN DEVELOPING COUNTRIES OF ASIA AND THE PACIFIC

Country [1]	University facilities		Adequacy of computer facilities	Demand of statisticians more than supply	Gap between education and work requirement	Training facilities		Teaching of statistics in schools
	Availability	Adequacy				Availability	Adequacy	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Afghanistan	Y	N	N	Y	Y	N	N	N
Bangladesh	Y	Y	N	N	Y	Y	N	N
Brunei	N		N		na	N	N	
China	Y	N	N	Y	Y	Y	N	Y
Cook Islands	N	N	N	Y	na	N	N	N
Hong Kong	Y	Y	Y	N	Y	Y	Y	N
India	Y	Y		N	Y	Y	Y	Y
Indonesia	Y	N	N	Y	Y	Y	N	Y
Kiribati	N	N	N	Y	na	N	N	N
Korea	Y	Y	Y	Y	N		N	
Malaysia	Y	N	N	Y		N	N	Y
Maldives	N	N	N	Y	na	N	N	Y
Nauru	N	N	N	Y	na	N	N	N
Nepal	Y	N	N	Y	Y	N	N	Y
Niue	N	N	N	Y	na	N	N	Y
Pakistan	Y	N	N	N	Y	N	N	N
Philippines	Y	N	N	Y	Y	Y	N	Y
Singapore	Y	N	Y	Y	na	Y	Y	Y
Solomon Islands	N	N	N	Y	Y	N	N	Y
Sri Lanka	Y	N	N	Y	Y	Y	N	N
Thailand	Y	Y	N	Y	na	Y	N	Y
Tonga	N	N	N	Y	na	N	N	Y
Tuvalu	N	N	N	Y	na	N	N	N
Vanuatu	N	N	N	Y	Y	N	N	Y
Viet Nam	Y	Y	N	N	N	Y	N	N
Western Samoa	Y	N	N	Y	Y	N	N	Y
SUMMARY	Y-16 N-10	Y-7 N-19	Y-4 N-22	Y-21 N-5	Y-14 N-2	Y-10 N-16	Y-3 N-23	Y-15 N-11

[1] responding to the SIAP survey.

Abbreviations: Y - Yes; N - No; na - not applicable.

ANNEX 2

SUMMARY OF REPORTS FROM COUNTRIES RESPONDING TO THE SIAP SURVEY

1. Afghanistan

The University of Kabul provides a B.Stat. course for twenty candidates per year, emphasizing equally theoretical and practical aspects of statistics. Statistics is also taught in B.A.(Econ.) course. The university courses are supplemented by the training opportunities offered by the international agencies in and outside the country. Most of the graduates are absorbed in the government agencies. There is a shortage of professionally trained statisticians in the country, as there are no regular training facilities. The facilities currently available in the universities are also limited and minimal.

There is an informal on-the-job training carried out by the employees in the data processing in the Central Statistical office (CSO) and once a year instruction in data collection is offered to the provincial personnel. The CSO considers that there is a gap between statistical education and work requirements of a statistician and to reduce this gap the CSO is in consultation with governmental and academic organizations and would support them qualitatively and quantitatively in ensuring that practical procedures are introduced in the curriculum. There is a plan to establish an Administrative Training Centre to meet the training requirements of government personnel.

It may be mentioned that the SIAP has conducted two courses in Afghanistan, one is a workshop on basic aspects of general applied statistics in 1974, and a country course in demographic and social statistics with special reference to the population census in 1976. The country expects to have more contributions from SIAP in reducing the gap that exists between supply of and demand for statisticians.

2. Bangladesh

The University of Dacca offers graduate and post-graduate courses in statistics in the country. As part of the University, the Institute of Statistical Research and Training (ISRT) is participating in statistical education, training and research and it brings out the Journal of Statistical Research. There is the Bangladesh Statistical Association headquartered at the ISRT in Dacca. There is a training section in the Statistics Division in the Government of Bangladesh to co-ordinate the training activities in the country and many training courses have been held to impart specific skills to statistical personnel at different levels. SIAP has conducted training courses on sample surveys in 1977, management of statistical

operations in 1979 and national accounting statistics in 1980. Also, the country sends its statistical personnel abroad for training in statistics. Bangladesh has hosted two FAO regional workshops on agricultural census and agricultural statistics in 1978 and 1981.

3. Brunei

There are no statistical education and training facilities in the country. Students wishing to take up degree or diploma courses have now to go abroad and most of them go to the United Kingdom for statistical education. Presently the demand for statisticians exceeds the supply. Basic elements of statistics are introduced in schools right from the lower primary classes. In the secondary forms, many subjects such as science and economics include statistics as part of the curriculum.

4. China

The People's University of China and the Universities of Xiamen, Liaoning and Hebei, the finance and economic colleges and the finance and trade colleges have been offering many courses in statistics in the country. At the bachelor's level, the most commonly offered courses are principles of statistics, mathematical statistics, industrial statistics, agricultural statistics and commercial statistics.

The facilities for statistical education and training available in China are relatively less developed and in most cases the faculty strength and capabilities are limited. computer facilities are available only in a few universities and colleges and the concerned personnel lack the necessary experience in handling them. There is shortage of trained statistical personnel in the country.

There is considerable interest in learning statistics, the modern technology, by increasing education and training facilities, sending appropriate personnel abroad, inviting experts to give lectures and conduct seminars and courses, and procuring books. The present aim is to train the teaching and professional staff concerned with statistical work by stages and in groups.

There is a plan to set up a planning and statistical institute and extend statistical institutions, which already exist at present at some places, to more places in the country. It would be useful if it is possible to train in about five years' time the present statistical personnel who do not have proper background in statistics. At present statistics is not being taught at school level in China.

5. Cook Islands

The country offers correspondence course in applied statistics with the co-operation of the University of the South Pacific in its diploma programme. Apart from this, the country has no statistical education or training facilities at present and depends to a large extent on external assistance on developing its statistical personnel.

6. Hong Kong

The University of Hong Kong and the Chinese University of Hong Kong provide graduate and post-graduate courses in statistics and the Hong Kong Polytechnic offers a diploma course in statistics. The bachelor's degree course is for three years at the University of Hong Kong and for four years at the Chinese University of Hong Kong. The master's degree course is for two to four years. In some of the courses, computer science is offered as a minor. Computer facilities available are adequate for purposes of statistical education and training. The Hong Kong Statistical Society was founded in 1977 and has a present membership of about 200.

The gap that exists between statistical education and statistical work requirements is being reduced by organizing training for the new recruits at the Census and Statistics Department, applying job rotation system, sending appropriate personnel abroad for specialized courses and attachment of personnel to statistical offices of developed countries. Statistics is taught in the secondary schools covering graphical presentation and measures of central tendency in first to third year and covering measures of dispersion of standard distribution and linear regression in fourth to seventh year.

7. India

India has about 50 universities offering statistics at graduate and post-graduate levels with the bachelor's degree being for three to four years and the master's degree being for two years. The Indian Statistical Institute, established in 1931, and the Indian Agricultural Statistics Research Institute founded in 1947 also offer graduate and post-graduate courses in statistics and carry out research work in both mathematical statistics and applied statistics.

The country has more than adequate facilities for teaching statistics and guiding research in this field. Though the curriculum in most of the courses concentrates on mathematical statistics, attempts are being made to make the courses include some emphasis on statistical operations and applications. Mechanical and electronic desk calculating machines are usually available for use in classrooms and computer facilities are increasingly becoming available. Many foreign students study statistics in Indian universities and institutes. The local demand for statisticians is adequately being met by the present educational system. In fact, there is some problem of unemployment among the graduate statisticians in the country.

In spite of the attempts to make the graduate and post-graduate courses oriented towards practical aspects of statistics, there is still a gap between statistical education and professional work requirements in this field. Training facilities have been created in the two institutes mentioned above and in the Central Statistical Organization (CSO) to bridge this gap. The new recruits in the

statistical service of the country are given intensive training for two years before they are assigned to regular jobs in the government. Appropriate personnel are sent abroad for further training in case of specialized fields. SIAP participated in the training seminar on use of sampling in census held at New Delhi in 1980.

Recognizing the need for practical orientation in teaching statistics, some universities send their students to the CSO to study official statistics as part of their curriculum. The International Statistical Education Centre, which runs a long-term statistical training course for participants from various countries, is being supported by the Indian Statistical Institute, the Government of India and the International Statistical Institute. The fellowships for this course are open to candidates from any country in the world. The activities of this Centre are covered in greater detail in the section 'Regional Training Centres' in this paper.

Statistics is being taught in higher secondary and high schools in India. Basically, graphic representation is covered in the first five years of the school's curriculum and in the remaining years collection and compilation of data and descriptive statistics are covered.

The Indian Statistical Institute Association established in 1931, the Indian Society of Agricultural Statistics established in 1947, and the Calcutta Statistical Association founded in 1945 are well known statistical professional societies in India and they bring out Sankhyā, the Indian Journal of Statistics, Journal of the Indian Society of Agricultural Statistics and the Bulletin of the Calcutta Statistical Association, as part of their professional activities in statistics.

8. Indonesia

Statistics courses are offered by the Academy of Statistics (Jakarta), Institute of Technology (Bandung) and some universities. There is a continuing attempt to improve the facilities by having new classrooms, more teaching staff at the M.A. and Ph.D. levels, library expansion and acquiring of more computer capability. Further, the existing staff capability is also proposed to be upgraded through further education abroad.

The gap that exists between statistical education and work requirements in statistics is being bridged to some extent by creating training facilities at the Statistical Education Training Centre which conducts many training courses including some correspondence courses. There is a plan to upgrade and expand the training facilities by establishing training centres in some provinces. The country is participating in the SIAP's training of trainers programme. SIAP has conducted training courses on national accounts in 1975 and on distributive and service trade statistics in 1978 and has participated in the workshop on sampling in 1980. Statistics is taught at high school level in the country.

The Indonesian Statistical Society is headquartered at the Central Bureau of Statistics in Jakarta.

9. Kiribati

The country utilizes the University of the Pacific branch in Tarawa, which conducts a course on elementary statistics for six months. The country has also participated in the course on statistical operations and procedures of the south Pacific Commission conducted for six weeks in 1979. It is felt that one graduate statistician is badly needed to organize and carry out the statistical functions in the country. Statistics is being taught at the secondary school level in Kiribati.

10. Korea

Graduate and post-graduate courses in statistics are offered by universities in Korea. The bachelor's degree courses are usually for three to four years whereas the M.A. degree courses are for two years. The universities generally have adequate computer facilities for use in their courses. The demand for statisticians in the country is about 20% more than the supply available from the present educational system. Most of the courses in the universities have been oriented to applications of statistics to a large extent and this has helped in reducing the gap that generally exists between statistical education and work requirements in this field. However, even with this reorientation of the courses to practical applications, there is some gap that is felt and to reduce this gap it is planned to set up a centre for statistical survey and its analysis at Donaook University. Further, there is also a plan to have larger computers in the universities and also to expand the available educational facilities by about 20%.

SIAP has conducted short-term training courses at Seoul on National Accounts in 1972 and on Statistical Methods with emphasis on practical aspects in 1979. The country is participating in the SIAP's programme of training of trainers.

Statistics is taught at school level in Korea, the coverage being statistical tables, histograms, pictographs, arithmetic mean and frequency distribution at the primary level; relative frequency, cumulative frequency, meaning of sample and population at the junior high school level; and mean, standard deviation, frequency distributions and normal distribution at the high school level.

The country has a Statistical Association of Korea headquartered at the Bureau of Statistics at Seoul.

11. Malaysia

The University of Malaya, the National University and the Agricultural University offer courses in statistics at graduate level. There is a great need for statisticians with an economics orientation and attempts are made to make the statistical courses application-oriented in the universities. The computer facilities available at present are not adequate. The demand for statisticians exceeds supply in the case of applied statisticians and the situation is the reverse in the case of theoretical statisticians. The country has no training facilities as such at present. It is felt that there is considerable need to upgrade the skill of teaching staff through exposure to applied aspects of statistics at SIAP and other institutions abroad.

Statistics is taught at the GCE O-level in schools. However, the present syllabus in high school certificate course in the field of statistics is too advanced and proper training for teachers is necessary if these courses are to succeed.

12. Maldives

The country has no university or college. The statistical education requirements of the country are presently being satisfied by sending appropriate personnel abroad for this purpose. The shortage of statisticians is being keenly felt in the context of the development of the country's economy and social life. The country has no regular training facilities and is sending suitable candidates for training at international institutes such as SIAP at Tokyo, ISEC at Calcutta and ISPC (International Statistical Programme Centre) in Washington D.C. Recently, the UNDP expert on statistical development conducted an elementary basic statistical course on collection of data for about three months at Malé. There is a need to meet the requirements for statistical training personnel through training abroad at the present stage.

13. Nauru

At present the country has no statistical education and training facilities. As the population of the country is rather small, the scope for statistics is also limited. However, rudiments of statistics are taught in secondary schools as part of mathematics curriculum. Nauru is also a member of the University of the South Pacific, Suva (Fiji), where short-term and degree courses in statistics are offered to the nationals of member countries. But so far Nauru could not avail of this facility.

14. Nepal

The Tribhuvan University offers bachelor's and master's degree courses in statistics. Except for calculating machines, there are no computer facilities readily available for use in statistical education at university level. Since the

university courses are theoretical in nature, and since applied courses in statistics are not available, there is a gap between statistical education and work requirements. The demand for statistics still exceeds the supply, though the position at present is better than before.

There are no training facilities in the country in the field of statistics. A few agencies in the government such as the Central Bureau of Statistics have benefited from getting some of their staff trained at institutions such as ISEC and SIAP. However, such opportunities are not easily available to the educational institutions to upgrade the quality of their staff. It is to be noted that the success of developmental planning in future will depend on the upgrading of the skills of the producers and the users of statistics in the country.

Elementary ideas of statistics are taught in schools from the eighth grade onwards.

16. Pakistan

The universities in Pakistan offer graduate and post-graduate courses in statistics and the facilities for teaching statistics are quite adequate and there is not much imbalance between supply and demand of statistics in the country. Pakistan Statistical Association is headquartered at the Institute of Statistics in the University of Punjab at Lahore.

There is a gap between statistical education and work requirements and this gap is being bridged through orientation and introductory courses conducted by the Statistics Division for the statistical personnel. However, at present, there is no institutional set-up for training as such in the country. The Statistics Division has training section and it is expected that the role of this section would be expanded to facilitate training of the statistical personnel to enable them to perform their duties effectively. There will still be the need of foreign facilities to upgrade the capabilities of the statistical personnel in specialized fields. In future it is expected that the direct recruitment at middle level in the statistical cadre would be through competitive examinations covering both theory and practice of statistics. The country presently utilizes the available fellowships under various plans such as the Colombo Plan, UNDP programmes, SIAP, ESCAP and East-West Centre, but it is found that these facilities are still not adequate to meet the growing requirements within the country. SIAP has conducted short-term training courses at Karachi on Sample Surveys in 1979 and on National Accounts in 1981.

Statistics is not being taught in schools in the country.

17. Philippines

The University of the Philippines and its Los Banos Center, the Polytechnic University and the University of the East offer graduate and post-graduate courses in the field of statistics. The graduate courses have been oriented towards practical aspects to the extent feasible. The aims of these courses are to produce professional statisticians and teachers and research workers. In spite of these facilities, there is a serious shortage of qualified statisticians. Appropriate statistical personnel are sent abroad for education and training, but these facilities are limited and are not sufficient to meet the growing demand for technical expertise in the country. The University of the Philippines at Los Banos is planning to include use of computers in its courses. It also has programmes to develop a Ph.D. course and statistical computing facilities. Similarly, the University of the Philippines is planning to revise its curriculum to reflect the specialists' needs from its present generalist orientation.

There is a gap between statistical education and work requirements in statistics and to bridge this gap to some extent training facilities are being created at the University of the Philippines at Los Banos, Statistical Centre of the University of the Philippines National Census and Statistics Office and National Economic Development Authority (NEDA). The NCSO provides on-the-job training for non-statisticians and also conducts a training course on EDP. Also, NEDA conducts non-degree, work-orientation courses for its staff.

A national convention of statisticians is held biennially to exchange views and discuss problems among professional statisticians in the country. There is an imbalance between supply and demand of statisticians in the country as about 48% of the staff have had no formal training in statistics. The NCSO is undertaking a continuing statistical training programme to develop various statistical training models for different levels of statisticians to assist other government agencies in the development of statistical methodology, to set up an effective method of administration of statistical training and staff development by providing and expanding scope of statistical training to its personnel in co-operation with the universities and other agencies in the country and abroad. It is also proposed to develop and maintain a system of evaluating the results of such efforts. The Statistical Centre of the University of the Philippines is proposing to have a department of training and administrative services to provide continuing education for statistical centre alumni, training for statistics teachers and consultancy services for government agencies.

The NEDA is considering the possibility of setting up a statistical institute for the Philippines which would conduct professional degree and non-degree courses in applied statistics and also organize seminars and workshops in the specialized fields in statistics for improving statistical knowledge and upgrading

skills of the appropriate personnel. The institute will also develop training materials, conduct studies on more effective methods for conventional and specialized training in statistical work and carry out research on recent developments in applied statistical methods, data processing and analytical techniques.

It is felt that SIAP should increase training programmes and fellowships for trainees on the advanced courses in statistics and in computer sciences. SIAP has conducted a short-term course on National Accounts in the country in 1977 and has participated in a data processing course of the National Census and Statistics Office in 1981. The United Nations and its agencies should help developing countries acquire hardware for computer usage.

Statistics is taught in schools. Elementary statistics is an elective subject in the fourth year of secondary school courses but only a few schools offer this and very few students take up this option.

The Philippines Statistical Association has a membership of about 350 and brings out the periodicals Philippines Statistical and PSA newsletter.

18. Singapore

The National University of Singapore offers bachelor's degree course emphasizing both theory and practice of statistics. The course is of four years' duration and trains sufficient number of candidates for supplying the local demand. The universities have adequate computer facilities but however in practice little training is possible on the computer itself. There is no difficulty in graduate employment in the country. It is expected that in the near future the statistics education facilities will be increased by about 20 per cent. Also statistics is taught as a subsidiary subject in many other courses with an orientation towards applications.

In spite of the practical orientation of the university courses, there is a gap between statistical education and work requirements and to reduce this gap short orientation courses are given to the new recruits and also appropriate personnel are sent abroad for further training or practical attachment with statistical offices of developed countries in specialized fields. With the latest emphasis on computerization in the civil service, statistical personnel will also be trained to handle computer equipment and operate micro- and mini-computers to upgrade the operations of calculation, processing and compilation of statistics. There is still a shortage of trained statisticians at graduate level as many statistical graduates find employment as economists or administrators.

Statistics is taught at the secondary school level and no difficulties are experienced in this effort.

19. Solomon Islands

The Solomon Islands Extension Centre of the University of the South Pacific offers a course in applied statistics as part of its diploma in accounting studies. The country has no other course in statistics and the present arrangement with the University of the South Pacific is considered adequate. The country has participated in the SPC's course on statistical operations and procedures (SOAP), which is a six-week course with practical orientation. There is a specific need in the country for sub-professional training and long-term overseas courses are not feasible as the number of personnel involved in the statistical office is very few. It is expected that some candidates will make use of the Institute of Statisticians' Examination offered by correspondence courses. It is desirable to supplement the SPC's SOAP course by correspondence or extension courses in applied statistics.

Statistics forms part of mathematics and social science syllabus in Solomon Islands' school certificate taken up to five years of secondary education and the emphasis is on basic numeracy and using statistics in real life situations.

20. Sri Lanka

The Universities of Colombo, Jaffna and Jayawardanapura offer graduate and post-graduate courses in statistics, emphasizing both theoretical aspects, the bachelor's degree course being for three years and the master's degree course being for one to one and a half years and the diploma post-graduate course being for one year. Statistics is also taught as a subsidiary subject in a number of other courses. The faculty generally faces shortage of teaching expertise and computer facilities in conducting the courses. The need to introduce automatic data processing as part of statistics curriculum is being increasingly felt. There does not seem to be any serious imbalance between supply and demand of statisticians in the country. There is also need for strengthening the statistics faculty in the universities and also to make available better computer facilities to carry out the statistics courses more effectively. The University of Colombo has link arrangement with the Department of Applied Statistics of the University of Reading and such arrangements are considered to be very beneficial for upgrading the teaching capability of the faculty in the universities.

The gap that exists between statistical education and work requirements is keenly felt and the Training Division in the Department of Statistics conducts occasional training courses at different levels to impart the necessary skills to its staff. The training facilities locally available are not adequate to meet the growing demands and foreign training facilities at SIAP, East-West Centre, ISEC and ISPC are being utilized to meet the demand for trained statisticians. SIAP has conducted short-term courses in Colombo on sample surveys and statistical

indicators in 1975, agricultural statistics in 1977, demographic and social statistics with special reference to population census in 1978 and national accounts in 1979.

Statistics is not being taught in the schools in the country.

Sri Lanka Statistical Society has its headquarters at the Department of Census and Statistics in Colombo.

21. Thailand

The universities in Thailand offer graduate and post-graduate courses in statistics with a mix of theory and practice, the bachelor's degree course being for four years and master's degree course being for two years. Desk calculating machines and many computers are generally available for use in the courses. There is not much of an imbalance between supply and demand of statisticians in the country. It is expected to revise the course contents in statistics with a view to orient it towards applications of statistics. Also it is intended to enlarge the computer facilities.

In spite of the practical orientation of the university courses in statistics, there is a gap between statistical education and work requirements, as students in colleges have little working knowledge of statistical operations as carried out in real-life situations. The National Statistical Office organizes training programmes for four to six months to reduce the gap between statistical education and work requirements by imparting statistical skills to its staff. Thus, the balance between supply and demand of statisticians is only in quantity but not in quality as in practice the graduates do require further training to make them useful in statistical offices. The country also utilizes training facilities available abroad. SIAP has conducted courses on evaluation analysis and use of statistical data in 1978 and on assessment and control of errors in 1981.

Statistics is taught at school level in the country covering interpretation of graphs and charts at the primary level, elementary basic statistics at the secondary levels and elements of statistical theory at the high school level.

Thailand's Statistical Association is headquartered at the National Statistical Office at Bangkok.

22. Tonga

The country has no statistical education and training facilities and relies mainly on assistance from abroad to meet its statistical personnel requirements. The country has had the SPC's SOAP course for six weeks and also specialized country course for middle-level personnel for four weeks. Calculating machines are generally available for use in courses and in work. The shortage of trained

personnel is particularly felt at middle-level and there is need for organizing primary and middle-level training programmes to upgrade the skills of the statistical personnel.

Statistics is taught in schools from form four level and it is expected to introduce more advanced statistics at form seven level.

23. Tuvalu

The country has no statistical education and training facilities and there are no plans for including statistics in the school curriculum. The present requirements of statistical personnel are met through external assistance in this field.

24. Vanuatu

The country has no statistical education and training facilities. With a population of 113,000 the country has limited scope for statistics. The country has utilized the SPC's courses on external and internal migration and the statistical operations and procedures. Calculating machines are available for use in the courses and in work. The country has three experts in the field of statistics who are expatriate staff, and it is expected that this staff will conduct training for the local staff at sub-professional level. For the purpose of professional training, it may be necessary to use the facilities available abroad.

Statistics is taught in schools with simple charts and statistical measures of central tendency being covered in three years of post-primary education and statistics and probability being covered at the GCE O-level.

25. Vietnam

The Economics Planning Universities in Hanoi and Ho-Chi-Minh City conduct economic statistics courses for four years with emphasis on practical aspects. Statistics is also taught as part of the book-keeping courses of the central statistical cadre and secondary statistical school. The emphasis in these courses is on statistical survey methods and reporting of forms. There is no gap between supply and demand as supply is planned on the basis of the requirements of the government and other agencies. Specialized courses on statistics balance and on sample surveys have been conducted in the country. For higher education and research, appropriate personnel are sent abroad. The country expects to benefit from the course design and course contents used at SIAP.

Statistics is not taught in schools in the country.

26. Western Samoa

The Extension Center of the University of the South Pacific at Apia and the University of Samoa offer five-year bachelor's degree courses in statistics and

mathematics covering only the basic aspects of the subjects. There is a serious imbalance between supply and demand of graduate statisticians in the country. The practising statisticians generally receive the necessary training in statistics overseas in institutions such as the Statistical Institute for Asia and the Pacific at Tokyo, the International Institute for Population Studies in Bombay and the East-West Center at Hawaii. The training facilities offered by ESCAP and SPC are also utilized.

The primary and secondary teachers colleges offer courses in mathematics covering elementary and basic aspects of statistics in their short-term courses, which are oriented towards primary and secondary school teachers. Statistics is taught in schools in the country covering basic elementary aspects of statistics in form 2 as part of mathematics syllabus and from form 3 statistics is taught as a part of applied mathematics.

It is suggested that the policy-makers should undergo an appreciation course in statistics, the teachers and lecturers should have intensive/university education and the practicing statisticians should have in-service training in some recognized statistical offices after completing statistical training abroad.

COMMENTS ON THE ABOVE PAPERS

K T de Graft Johnson, Economic Commission for Africa, Addis Ababa

The three papers give brief descriptions of the development of statistical education and training in Africa, Asia and the Pacific and Latin America and discuss problems created by the gap between theory and practice which in spite of attempts to bridge it still persists to this day. Before commenting on general issues, I will make some remarks on each paper separately.

Adichie and Afonja state that what is taught in statistical institutions is dictated by the priorities of the country and the university. I doubt whether that reflects accurately what is actually happening in most of the African universities and training centres. For example, energy is currently the most important element in the development of the countries of the region, whether these countries are oil-producing or not, but how many of these educational and training institutions include a course on energy statistics in their curriculum? The truth is that most of the courses are carbon copies of what is taught in British and American universities. It is only recently that conscious efforts are being made at least in the field of applications to shift the focus to be in line with national priorities.

The authors also refer to the admission of statistics as a discipline in its own right as a recent event. This is true of Africa but the underlying reasons are rather complex. While, for example, the University Colleges of Ibadan and the Gold Coast were first established with predominantly British academic staff, it was only natural that they should follow the Oxford and Cambridge rather than, say, the Iowa State tradition. The majority of students in the initial stages were general rather than special or honours degree potential. My own graduating class in 1953 had only two candidates for the special degree in mathematics; and of course there were only three questions out of ten in one of the eight papers required for the B.Sc. (Special in Mathematics) examination of the University of London which related to statistics. You could therefore take the whole examination without answering a single question in statistics. It may be recalled that during the period under reference, the general degree was of a two years' duration while the honours degree lasted three years, both after the intermediate examination. As seniority in the civil service was based on year of entry, honours or special degree holders had the disadvantage of working under their classmates who had opted for the general degree. My own recollection of the first person in the then Gold Coast (now Ghana) who could be described as a professional statistician having done a degree in mathe-

matics and also passed the Institute of Statisticians' examinations with honours, was that he was refused entry into the statistical office because he was too specialised to be useful to the Central Bureau of Statistics. Thus job openings and the academic inheritance of the African universities were the main problems encountered in statistical training in the initial stages in English-speaking Africa. The position of the French-speaking countries was only slightly better, since they were trained at the CESD, but it should be noted that this was modelled on the lines of the French ENSAE.

Adichie and Afonja also list two major constraints in statistical education, namely, inadequacy of qualified teachers and non-availability of teaching material in applied areas. However, as implied in my preceding remarks the position is more complex.

I next turn to the paper by Cansado. He refers to the problem of finding suitable staff, namely persons with a good theoretical and practical background for the statistical training centres. This is a serious problem for all training centres in the developing world. I will revert to this issue in my concluding remarks.

Cansado also states "A characteristic of university teaching in this part of the world..... is that in the formal teaching a heavy weight is given to statistical theory and a very light one to applications and actual data analysis". The question is "need this be so?" It is my view, for example, that in teaching sampling or design of experiments there is great scope for illustrating the theory by reference to applications and actual data analysis. If this is done, most of the criticisms directed against the so-called academic statisticians by the government statisticians would no longer be valid.

Another valid point made by Cansado is the very little research done in Latin American Universities. He quotes the recommendations of a 1980 Working Group on this subject which I hope will be discussed extensively here. There is one question which should be addressed to all authors of papers for this session in connection with the recommendations. How much, and what type of research, is any developing area government willing to pay for? My experience in a number of developing countries has been that governments are willing to pay very little for pure or applied research. Most research projects have thus to be camouflaged by emphasising the immediate policy relevance of the results rather than the wider and long-term benefits.

I will now deal with only two issues in the paper by Murthy; namely future directions and his suggestions for action. He emphasises three main areas in which actions should be taken:

"(i) increasing the statistical awareness and priority at policy levels in government;

(ii) widening of statistical curricula in universities to include practical aspects of applied statistics and promotion of the study of statistics; and

(iii) general statistical edcuation in high schools."

These recommendations are generally in line with those which have been made by various working groups and meetings of directors of African statistical training centres and approved by the African Conferences of Ministers of Planning. In fact the region has gone even further to recommend that national statistical development programmes should be included as part of national development plans. This view although accepted by the OAU meeting of heads of state has yet to be widely implemented. This is an example of a different type of gap - between resolutions and their implementation.

As stated earlier, I will now deal with general issues arising out of the three papers. These relate to the implied dichotomies between education and training, and academic vs government statisticians as well as the role of research in statistical development.

My dictionary defines training as "bringing a person to the desired state or standard of efficiency, etc., by instruction and practice", and education as "giving intellectual and moral training to a person". Thus as far as statistical development is concerned there should have been considerable overlap between education and training. Failure to expand this overlapping area is the cause of the present unsatisfactory situation. It is worth noting that in one African University, the Institute of Statistics and Applied Economics at Makerere University has attempted to expand this area of intersection between training and education by introducing a degree course B.STAT in addition to the traditional B.Sc.Stat. The former deals extensively with data handling techniques while the latter only briefly considers official statistics such as agricultural and industrial statistics. In line with my own views on the matter, I would have been much happier if the institute could have evolved only one degree course for statistical specialists which brought together the best elements in the B.STAT and B.Sc. courses.

Implicit in all three papers is the rather erroneous assumption that there are only two groups of statisticians - academic and government. The ISI Committee on Integration of Statistics identified 15 categories, namely:

Academic	(including university, college, school)
Business	(including actuarial, marketing, economic, econometrics)
Government	(including federal, state, local, municipal)
Industrial	(including operations researcher, quality controller, reliability worker, physical, engineering, transport)
Practitioner	(including applied statistics, data analyst)
Research	(including applied statistics, methodological, theoretical, mathematical, probabilist)
Student	
Biometric	(including ecological, agricultural, forestry)
Medical	(including biostatistician, health, epidemiologist, pharmaceutical)
Social	(including demographic, political, judicial, criminal, behavioral, psychological)
Administrator	(including manager, policy former, planner, forecaster, auditor)
Consulting	
Computational	(including simulator, computer programmer, package writer)
Survey	(including pollster)
Other	(including sports, numerologist)

In my view, it is more relevant and useful to deal with the content of statistical training in developing countries in the broader spectrum of 15 different categories than in the narrower context of a simple dichotomy of academic and government statisticians.

At this stage, it is worth discussing the main recommendations of a 1981 Working Group on Statistical Organization convened by the United Nations Economic Commission for Africa which considered inter alia, the integration of statistics and its implications for teaching. The meeting "considered that integration was desirable but that it could be achieved more effectively by improvements in the methods of training which produced a proper mix of theory, methods and applications. ..... Stress was laid on the period of apprenticeship either as part of a degree programme or outside it".

In conclusion, Mr. Chairman, I would like to refer to the need for research - both substantive and methodological - by all 15 categories of statisticians in developing countries. This emphasis is implied in Cansado's paper and I endorse it whole-heartedly. Research can also bridge the gap between the academic and government statisticians, if both engage in it. In addition those who teach official statistics and data handling techniques in universities should have had considerable practical experience in these fields. This regrettably is not the case in most developing countries.

REPORT ON SESSION ON STATISTICAL EDUCATION AND TRAINING IN DEVELOPING COUNTRIES

B P Adhikari, Indian Statistical Institute

The session consisted of a paper reading meeting and two workshops.

At the paper reading meeting 3 papers were presented:

1. Statistical Education and Training in Latin America - Enrique Cansado, CIENES, Santiago, Chile.
2. Statistical Education and Training in the Developing Countries of Africa - J.N. Adichie, Department of Statistics, University of Nigeria, Nsukka and Biyi Afonja, Department of Statistics, University of Ibadan, Ibadan.

and

3. Statistical Education and Training in Developing Countries of Asia and the Pacific - M.N. Murthy, Statistical Institute for Asia and the Pacific, Tokyo.

The first two papers were presented by the authors themselves. Since Dr. Murthy was not present, the salient points made in his paper were presented by the Chairman. Copies of all three papers had been sent in advance to Dr. K.T. de Graft-Johnson of the Statistics Division, Economic Commission for Africa, Addis Ababa and he presented his comments. There was an open discussion on the papers which, together, gave a broad picture over all the developing countries of the world. Since copies of the paper had not been made available in advance to participants and the abstracts too brief, it was difficult for them to focus on specific issues raised in the papers. Participants from several countries disagreed with some of the findings of the survey reported in Dr. Murthy's paper. Critical comments on his paper were made particularly by participants from India and the South Pacific region.

The two workshops partly dealt with the contents of the three papers and partly on other issues raised on the floor. Some of the points made were:

1. The level and spread of statistical education and training were unequal in the different developing countries, some being quite advanced and comparable to the developed countries and some at lower levels. Advantage of this heterogeneity could be taken by utilising the expertise available in those countries with higher levels of statistical education.
2. The gap between academic teaching of statistics in universities and colleges and the requirements of practitioners was universally noted. In some countries steps were being taken to achieve a closer collaboration between university departments and the official statistical system. Some participants were of the view that this was useful. Others suggested that universities might have two streams of curricula - academic and professional.

3. It was pointed out that the quality of data and statistical expertise depended on the degree of realisation by the governments of the usefulness of statistics for policy making. More and more countries were progressing towards this realisation.

4. Commendable work was being done in remote and inaccessible islands of the South Pacific.

5. School level education in statistics was either non-existent or in a rudimentary stage in most countries.

6. International bodies like the UNESCO should be approached for help in coordination, holding of meetings etc. and, in general, in helping in the spread of statistical education and training in developing countries.

TOPIC 11: CO-OPERATION BETWEEN ACADEMIC AND PRACTISING STATISTICIANS

ORGANISER: PROFESSOR W G HUNTER

CO-OPERATION BETWEEN UNIVERSITY AND INDUSTRY STATISTICIANS IN THE UNITED STATES

R D Snee, E I du Pont de Nemours and Company, Wilmington

INTRODUCTION

Rarely a week goes by without an article on cooperative interactions between academic and industrial communities appearing in Science, Nature, or other journals of science and engineering. This dialogue has been stimulated in large part by the decreasing growth of student populations and the reduced Federal funding for academic research. These developments have also been a concern to statisticians in the U.S. and the American Statistical Association (ASA) has formed committees to study the issues and to make recommendations (ASA 1980, ASA 1982). These reports have in turn produced comment by several university and industry statisticians (e.g. Bradley 1981, 1982; Hill 1982; Hunter 1981, 1982; Rustagi and Wolf 1982, Snee 1982).

Cooperation between industry and academia in the U.S. has a long history. There are many models to follow (e.g. Russell and Turner 1968; Wicks 1981) and new approaches continue to be proposed (Roy 1981). Tebo (1982) emphasizes that cooperative programs can be viewed as a mechanism of increasing university productivity. There has also been considerable cooperation in the field of statistics. It is apparent, however, that this activity must be increased and broadened if our profession is to successfully meet the challenges that face us. Certainly statistics has no God-given right to survive as a discipline and profession and we must continually look for ways to increase our impact and influence.

Various methods of interaction and cooperation are reviewed in this paper and suggestions for new approaches are presented for consideration. Other suggestions can be found in ASA (1980). It is apparent that the development of university-industry cooperative programs will result in a broader role for statisticians in our society. In Snee (1982) it was emphasized that focusing on opportunities such as training for industry, development of improved communication skills, statistical quality control training and research, and the establishment of consulting intern programs are all part of developing a broader view of statistics. Interactions with industry can help universities take full advantage of these opportunities.

KEYS TO SUCCESS

We must recognise at the very beginning that successful cooperative programs do not happen by themselves; they require careful planning and a recognition of some key ingredients that must be present (Table 1). Fruitful interactions begin with both parties getting to know each other and understanding each other's needs and objectives. It has been my experience that unsolicited proposals from unknown parties with no prior contact are almost always rejected out of hand. Industry must communicate who they are and what they do and suggest programs that are of interest to them. Universities must identify their goals and the constraints of their environment. I know of several instances in which both parties have held unrealistic views of the other's situation.

TABLE 1

COOPERATION BETWEEN UNIVERSITY AND INDUSTRY
CHARACTERISTICS OF SUCCESSFUL PROGRAMS

- Both parties must get to know each other.
- Either party can initiate the discussion.
- There must be mutual benefit to both parties.
- Interaction with nearby universities is easiest to implement.
- Programs can start small and include a plan for future growth.
- No single approach will work in all cases.

In order to get the interaction started someone must take the first step. Both parties have the responsibility to do this. If the proper groundwork has been done, it is not unlikely that the proposal will be met with the response, "we think that your proposal is worth pursuing; in fact, we have been thinking along the same lines ourselves. Let's get together to discuss it further". Experiences of other disciplines (e.g. Russell and Turner 1968; Tebo 1982) indicate that contacts and interactions between interested individuals at the university and company (frequently one person in each organization) are needed to define and carry out the cooperative programs. The success of the programs depends heavily on the dedication of the individuals involved.

It should also be recognized that both the university and the company must benefit from the interaction. Fusfeld (1980), Director, Center for Science and Technology Policy, New York University, in a Science editorial, points out that:

> "Initial approaches were made by universities with overtones of 'with your money and our brains' - not an endearing note, and surely not the best one on which to begin a relationship. ... Universities and industry are now building toward long-term relations that take into account each other's needs and contributions, the functions that each serves in our society. This is, in short, a sound base for a sensible working partnership".

E. G. Jefferson, Chairman of The Du Pont Company, in commenting on collaborative research efforts, emphasizes that these agreements

> "... are built on the sound foundation of mutual benefit, ..." and they emphasize "... not what industry can do for academia, but what we can do for each other".

There are many ways that industry can help beyond the awarding of unrestricted grants. Innovative thinking is needed to take full advantage of the possibilities.

The interaction is usually easiest to implement with nearby universities because there is greater opportunity for extended face-to-face contact, visits to the other party's site, adjunct faculty appointments, intern programs, etc. This greatly facilitates the "getting to know each other" process.

In getting a cooperative program off the ground it is usually easier to start small and include a plan for future growth. Starting small (e.g. one or two adjunct faculty appointments and a student intern program) provides flexibility and is easier for university administration and company management to approach. As these efforts prove successful and the parties get to know each other there will be stimulus and support for other programs such as joint research, visiting appointments, etc. A useful parallel is that of developing a consulting relationship with an organization. One usually begins with an important project with modest objectives. The successful completion of the effort often leads to other projects with broader scope and greater impact.

It is not possible to detail a prescription that will work in all situations. A variety of approaches is possible. Each program should be tailored to the needs of the organizations involved. The following sections describe various forms of cooperative programs that have been developed. The discussion is organized by those interactions that take place with a university and those outside the university including those through professional societies such as the ASA and the American Society for Quality Control (ASQC).

INTERACTIONS WITH A UNIVERSITY

Interactions between a university statistics department and a company can take a number of different forms (Table 2). One of the easiest programs to implement is visits to universities by industrial statisticians and vice versa.

Many universities and companies extend such invitations several times each year, usually for a day consisting of a seminar and one-on-one discussions. Visits of two days to one week are increasing in popularity and are particularly useful in getting a feel for what life is like in a given site. Visits for longer periods of time are also possible although they appear easier for university statisticians to make. Statisticians working in industry tend to have responsibilities that are not easily shifted from one person to another.

TABLE 2

INTERACTIONS WITH A UNIVERSITY

- Visiting lecturers from industry.
- Appointment of industry statisticians as adjunct faculty.
- Consulting intern programs in industry.
- Inclusion of industry statisticians on university advisory committees.
- Special programs and short courses.
- Problem solving for industry.

Industry statisticians often serve as adjunct faculty teaching courses in their specialty, helping out with teaching overloads, advising students, and directing thesis research. Students generally respond positively to the opportunity to take courses from statisticians who work on industrial problems on a regular basis and provide a different perspective on the role of the statistician.

Programs such as visiting lecturers from industry and appointment of industry statisticians as adjunct faculty members have been in existence at a few universities for many years. These programs will naturally expand as the interaction between universities and industries increases. Both parties should initiate discussions to speed up the process.

Employment of students in industry as consulting interns or as summer employees has been implemented by a few universities; however, it appears that much more can be done (e.g. Jones and Kanji 1980; Stoline 1982). It should be recognized that such programs require time to administer effectively and may be available to only a few students because an industrial organization can generally handle only one or two students per year. Consulting intern programs in industry may be difficult to implement if the university is far from an area with a large number of companies.

A university should also consider having a consulting intern program at the university in conjunction with its statistics laboratory (Snee 1982). Ideally, a university should have all three programs (consulting intern in industry

or at the university and summer employee in industry) with students participating in those programs that will do them the most good. The key objective is for the student to work on real problems under the guidance of a statistician experienced in statistical practice (ASA 1980, Section 3.4). This can be done either in industry or at the university. I worked, during my graduate study, as a summer employee for Jones and Laughlin Steel Corporation and as a consulting intern at Rutgers University (Snee 1982). I found both of these experiences very worthwhile in learning the realities of statistical practice and how statistics is used in solving the problems of science and engineering.

Consulting intern programs of all types are of interest to industry and they enthusiastically endorse such activities. Through the efforts of William J. Hill, the Allied Foundation has provided funding for a Workshop on Consulting Intern Programs to be held at the University of Wisconsin in November 1982. The workshop, which was organized by Brian L. Joiner, includes leaders in consultant training from industry, government, and academia.

Some universities (e.g. Brigham Young, Ohio State, Oakland) are including industry statisticians on their departmental advisory and evaluation committees. This can be beneficial to both parties and should be pursued more actively. It can give the program added credibility with industry and be a source of ideas, energy, and publicity.

The activities discussed above are well known, if not widely implemented, and have received considerable discussion (e.g. ASA 1980). The opportunities associated with programs such as short courses for industry, research on unsolved statistical problems, and the investigation of subject matter problems are beginning to be recognized by universities and merit consideration and discussion by both universities and industry.

Special Programs and Short Courses for Industry

Statisticians working in industry have a need to keep up with recent developments in the field and to learn new technology. Brigham Young University recognized this need and has run a one-week Summer Institute on Applied Statistics since 1976. A well-known statistician serves as the primary lecturer with supplementary presentations by BYU faculty and statisticians working in the area. This institute benefits not only the persons who attend (mainly practising statisticians), many of which are BYU alumni, but also BYU faculty and students. Other universities should find this kind of program attractive because it gives the university's alumni an opportunity to return to the campus.

The University of Massachusetts offers each year a three-day Seminar on Statistics in Biopharmaceutical Research and a two-week Summer Program in Biostatistics. These programs are sponsored by the Biopharmaceutical Research Unit which

is funded through grants from several pharmaceutical companies.

Scientists and engineers working in industry have a need to learn what statistics can do for them and their companies. The University of Wisconsin School of Engineering and Statistics Department held a two-day seminar on Statistics In Industry in February 1982 to fill this void and plans to present similar programs in the future. Speakers from Consolidated Papers, Du Pont, Kodak, 3H Industries, 3M Company, and the University of Wisconsin emphasized how statistics is used and the impact it can have in industry. Videotapes of the sessions are available for distribution. For further information, write Mr. Robert Perras, Kurt Wendt Engineering Library, University of Wisconsin, 215 N. Randall Avenue, Madison, WI 53706.

The University of Tennessee has focused on another critical industrial problem; that of improving product quality through the use of statistical quality control (ASA 1981). The University has formed The Institute for Productivity Through Quality which offers an intensive three-week training program spread over a three-month period (one week/month). Many more such programs are needed to satisfy the demand in the U.S. for training in statistical quality control. Universities can play a leadership role here.

Problem Solving for Industry

In developing cooperative programs with industry, it is logical to think of research projects on statistical methodology. Bradley (1981, 1982) suggests that industry should pose unsolved statistics problems for universities to investigate. This is certainly appropriate for, as Box (1980) points out, interest in practical problems has led many statisticians (e.g. Gauss, Gosset, Fisher, Wald) to make important theoretical advances. Companies are more likely to do so if they have been approached by a university expressing interest in such problems. Some persons argue that if a problem is important to a company, they will solve it themselves, perhaps with the aid of a consultant. A company is also unlikely to recognize a statistics research problem unless it has someone on its staff with a knowledge of statistics. It seems that it would be particularly worthwhile for a university to initiate contact with those companies whose personnel have little formal training in statistics.

The identification of important unsolved statistics problems can be done through our journals; however, a more effective method is for industry and academic statisticians to discuss problems of mutual interest in a conference/workshop setting. This could be done through a professional society or organized by a specific university or group of universities.

The Mathematics Department of Duke University is holding such a program that is patterned after a similar workshop held for more than ten years at

Oxford University in England (see SIAM News, June 1979). Problems are solicited from industry and government and reviewed by the faculty to determine interest in the problem and potential for solution. The proposers of the selected problems are invited to a week-long conference in which a small study group is assigned to tackle each problem. The conference meets as a group at the end of the week to discuss their progress. It is not unreasonable that the interaction and dialogue might continue after the conference.

The key ingredient in such an arrangement is university and industry statisticians who have a mutual interest. A small seminar might be a good way to continue the dialogue.

The solution of subject-matter problems is another opportunity for cooperative programs. Environmental studies is an area where there is tremendous need that can be filled by university statistics departments. Research on these problems can be funded through trade associations because the problems are of concern to several companies. These problems often require specialized statistical methodology and the development of new techniques. Such activity is characteristic of the university. Expertise and reputation in a particular area of statistics and an interest in working on practical problems are key considerations when industrial groups seek such help.

The University of Wisconsin did considerable work in the mid-1970's on the analysis of Los Angeles air pollution data with funding from the American Petroleum Institute (Tiao, Box, and Hamming 1975a, 1975b; Tiao, Phadke, and Box 1976). More recently, Wisconsin and Princeton University have made extensive analyses of trends in stratospheric ozone (Reinsel, et al 1981, Bloomfield, et al 1982) with support from the Chemical Manufacturers Association. Time series methodology played a key role in both of these projects.

INTERACTIONS THROUGH PROFESSIONAL SOCIETIES

Interaction between university and industry statisticians can also take place in conjunction with professional society activities (Table 3). These contacts enable both parties to get to know each other and provide a mechanism through which they can address concerns of mutual interest.

TABLE 3

INTERACTIONS THROUGH PROFESSIONAL SOCIETIES

- Industry support of professional society activities.
- Meetings with speakers from both industry and universities
- Conferences on unsolved problems and other key issues.
- Committees to address common problems.

Industry has a long history of support of professional and learned societies in the United States. Industry not only sends their statisticians to meetings but also supports their activities as meeting organizers, speakers, moderators, committee members, and society officers. Part of the ASA's program on filming of distinguished statisticians has been supported by the efforts and funding of the Pfizer, General Motors, and Du Pont Corporations. Industrial statisticians have also written numerous articles describing statistical consulting in industry (ASA 1980; Greenfield 1979; Hooke 1979; Marquardt 1979, 1981) and presenting results of their research and statistical studies (e.g. see Biometrics, Technometrics, J. Quality Technology).

Interaction and cooperation can be stimulated by insisting that conference committees and speaker lists include representatives from both academia and industry. This has long been a guiding principle of the Gordon Research Conference on Statistics in Chemistry and Chemical Engineering and the Fall Technical Conference of the Chemical Division of the American Society for Quality Control (ASQC) and the ASA Section on Physical and Engineering Sciences. A goal of the ASQC Chemical Division's short course program, which began in the 1950's, has been that the instructors in each course should be one person from industry and one person from a university. This ensured a variety of viewpoints and encouraged cooperation between university and industry. Applied Regression Analysis, the first text to address the practice of regression analysis, grew out of the notes for one of the ASQC Chemical Division's courses.

The key to success of any conference is a balanced program that has something for everyone. It is disappointing to see a meeting with an underrepresentation of one group or another. The balanced program is, of course, more difficult to develop and takes greater effort but is more valuable in the long run and is usually well attended.

Conferences also present unique opportunities for students to meet and get to know statisticians from industry. Universities can arrange for students to attend. It is usually possible to negotiate reduced registration fees or find ways for the students to help out in lieu of the fee. It was under such an arrangement that, as a graduate student, I first met R.W. Kennard of Du Pont who was to later hire me into Du Pont's Applied Statistics Group.

Two other examples of cooperative efforts through professional societies are the founding of Technometrics, a journal of statistics for the physical and engineering sciences, and the Biopharmaceutical Section of ASA. Technometrics was started with matching funds from the ASQC Chemical Division and ASA. The Biopharmaceutical Section of ASA resulted from the rapid growth of statistics in the pharmaceutical industry and the need for industry and university

statisticians working on pharmaceutical problems to meet and exchange ideas and experiences.

In my visits to universities I have found that many faculty and students desire more information on statistics in industry. I have suggested that professional societies can help by developing a list of companies that employ statisticians, preparing a speakers list of statisticians in industry who are interested in visiting universities to discuss opportunities in industry, and by arranging workshops and question-and-answer panel discussions at local and national meetings concerning industry-academic programs and related issues (Snee 1980).

Professional societies can play an important role in getting university and industry statisticians together in conference and workshops to discuss key issues. A conference on unsolved statistical problems, as discussed in 'Problem Solving for Industry', could be sponsored by a professional society. Improved public perception concerning the role and contribution of statisticians and methods of recruiting students into the profession are possible topics for such conferences. These meetings could be held at either the local or national levels and could be in the form of "statistics days" in which students can learn about careers and opportunities in the field of statistics.

The formulation of committees of professional societies to address issues of common interest is another good way to develop cooperation between university and industry statisticians. The ASA Committee on Training of Statisticians for Industry (ASA 1980) is a good example where fruitful interactions have taken place. ASA has also recently formed a continuing education advisory committee composed of representatives of both industry and academia.

OTHER PROFESSIONS ARE FACING SIMILAR PROBLEMS

It was pointed out at the beginning that interest in cooperative programs between industry and universities is due in large part to decreasing student populations and reduced Federal funding for academic research. Statisticians must recognize that computer science and some fields of engineering are experiencing many of the same pressures as statistics and that companies will view statistics in competition with these fields when developing and funding cooperative programs. The Du Pont Company, for example, made a $6 million grant to Harvard University for research in genetic engineering (Stein 1981) and recently established a program to encourage chemical engineering students to pursue the PhD degree (Tebo 1982). Both of these areas are of keen interest to Du Pont and play a significant role in the future of the Company.

We must also keep in mind that statistics, particularly in industry, is a young profession (industrial statistics groups were first established in the U.S. in the early 1950's) relative to other fields of science and engineering. It will

take many years (perhaps a generation or two) for statisticians to have the impact and gain the influence now enjoyed by other professions.

Statistical studies also result in a report rather than a piece of hardware (i.e., a machine or process). It has been my experience, and that of many others, that hardware always wins in the sense that it is easier to convince people of the utility of hardware and, in the case of computers, its associated software, than of concepts, theories, and procedures. Statisticians should take advantage of this by emphasizing the role of the computer both in implementing the procedures of statistical practice and in statistical research. Industry is a good source of examples of the use of the computer in statistical studies.

I make these comments not to discourage the profession but to point out the realities of the situation so that we are not disappointed in our progress because of unrealistic expectations. University-industry cooperation has benefited the profession in the past and increased activity in this area will serve to stimulate even greater advances in the future.

ACKNOWLEDGEMENTS

The author is indebted to D.W. Marquardt and P.V. Tebo for stimulating conversations regarding the interaction between universities and industry and for their helpful comments and suggestions on the presentation of this paper.

References

[1] ASA Committee on Training of Statisticians for Industry (1980) "Preparing Statisticians for Careers in Industry" (With Discussion). The American Statistician, 34, 65-80/

[2] ASA Committee on Training of Statisticians for Industry (1981) Letter to the Editor. The American Statistician, 35, 269. Need for training in statistical quality control is discussed.

[3] ASA Committee on Training of Statisticians for Federal Government (1982) "Preparing Statisticians for Careers in Federal Government". The American Statistician, 36, 69-89/

[4] Bloomfield, P., Thompson, M.L., Watson, G.S. and Zeger, S. (1982) "Frequency Domain Estimation of Trends in Stratospheric Ozone". J.Geophysical Research, In Press.

[5] Box, G.E.P. (1980) "Comment on Preparing of Statisticians for Careers in Industry" The American Statistician, 34, 76-77.

[6] Bradley, R.A. (1981) "What Industry Can Do For Statistics". Presented at the Fall Technical Conference on the ASQC Chemical Division and ASA Section on Physical and Engineering Sciences, Gatlinburg, TN, October 1981.

[7] Bradley, R.A. (1982) "The Future of Statistics as a Discipline" J.American Statistical Association, 77, 1-10.

[8] Draper, N.R. and Smith, H. Jr. (1966) Applied Regression Analysis, John Wiley and Sons, Inc., New York. Second Edition appeared in 1981.

[9] Fusfeld, H.I. (1980) "The Bridge Between University and Industry". Science, 209, No.4453, 221.

[10] Greenfield, A.A. (1979) "Statisticians In Industrial Research: The Role and Training of the Industrial Consultant". The Statistician, 28, 71-82.

[11] Hill, W.J. (1982) "Training of Statisticians for Employment in Industry". In Proceedings of the Ohio State University Conference on Teaching of Statistics and Statistical Consulting, eds. J.S. Rustagi and D.A. Wolfe, New York: Academic Press.

[12] Hooke, R. (1979) "Getting People to Use Statistics as God and Sir Ronald Fisher Intended". The American Statistician, 33, 39-42.

[13] Hunter, W.G. (1981) "The Practice of Statistics: The Real World is an Idea Whose Time Has Come". The American Statistician, 35, 71-76.

[14] Hunter, W.G. (1982) "Graduate Programs in Statistics". In Proceedings of the Ohio State University Conference on Teaching of Statistics and Statistical Consulting, eds. J.S. Rustagi and D.A. Wolfe, New York: Academic Press.

[15] Jefferson, E.G. (1982) "Chemical Research: An Industry View". CHEMTECH, 12, 260-262.

[16] Jones, J.D. and Kanji, G.K. (1980) "The Role of Professional Experience in Statistical Education". The Statistician, 29, 196-203.

[17] Marquardt, D.W. (1979) "Statistical Consulting In Industry". The American Statistician, 33, 102-107.

[18] Marquardt, D.W. (1981) "Criteria for Evaluating Statistical Consultants in Industry". The American Statistician, 35, 216-219.

[19] Reinsel, G., Tiao, G.C., Wang, M.N., Lewis, R. and Nychka, D. (1981) "Statistical Analysis of Stratospheric Ozone Data for the Detection of Trend". Atmospheric Environment, 15, 1569-1577.

[20] Roy, R. (1981) "Graduate Universities - A New Model". Science, 214, 1177.

[21] Russell, T.W.F. and Turner, H.E. (1968) "Internship in Chemical Engineering Design". Chemical Engineering Education, Spring 1977, 74-77.

[22] Rustagi, J.S. and Wolfe, D.A. (1982) Proceedings of the Ohio State University Conference on Teaching of Statistics and Statistical Consulting, New York: Academic Press

[23] Snee, R.D. (1980) "The Bridge Between University and Industry". Amstat News, No.70, December, 1.

[24] Snee, R.D. (1982) "Comment on Preparing Statisticians for Careers in Federal Government". The American Statistician, 36, 82-87. Comparison of training needed by industry and government is made.

[25] Stein, M.D. (1981) "Industry and Research: Du Pont Joins In". Nature, 292, 191.

[26] Stoline, M.R. (1982) Western Michigan University Biostatistics Internship Program. Personal Communication.

[27] Tebo, P.V. (1982) "Programs Contributing to Increased University Productivity: A Grass Roots Approach". Presented at The American Society for Engineering Education, Texas A&M University, College Station, TX, June 23, 1982. The Du Pont Company's university financial assistance and partnership programs are discussed.

[28] Tiao, G.C., Box, G.E.P. and Hamming, W.J. (1975a) "Analysis of the Los Angeles Photochemical Smog Data: A Statistical Overview". J. Air Pollution Control Association, 25, 260-268.

[29] Tiao, G.C., Box, G.E.P. and Hamming, W.J. (1975b) "A Statistical Analysis of the Los Angeles Ambient Carbon Monoxide Data 1955-1972". J. Air Pollution Control Association, 25, 1129-1136.

[30] Tiao, G.C., Phadke, M.S. and Box, G.E.P. (1976) "Some Empirical Models for the Los Angeles Photochemical Smog Data". J. Air Pollution Control Assocaition, 26, 485-490.

[31] Wicks, Z.W. Jr. (1981) "Industry and Academia Can Work Together - A Case History. CHEMTECH, 11, 740-745.

SOME OBSERVATIONS FROM EUROPE

Lars-Erik Öller, Ministry of Finance, Helsinki

STATISTICS AND SEMANTICS

To the best of my knowledge, in all Indo-European languages the word "statistics" is used to denote two different things, namely a collection of data *and* an academic discipline. In the only non-Indo-European language I know, i.e. in Finnish, there are separate words for the two concepts. (*)

I think this semantic flaw has resulted in a lot of confusion. There is, and should be, a difference between the aims and means of a statistical office and those of a research oriented statistical unit. The former has to meet the demand of society for certain data. The main task of the office is to collect those data and make them available to the public. This involves application of some statistical techniques, such as sampling theory, but to a much larger extent it requires subject matter insight and thereby other disciplines. National accounts imply knowledge of economics, social statistics is supported by sociology, meteorological statistics needs meteorology etc.

If statisticians think that they and only they are fit to assume full responsibility for routine processing of social data I think we are aiming at the wrong target.

Instead of getting involved in semantics I suggest we concentrate our efforts on three areas where the need for statistical knowhow, if not overt, at least potentially, seems to be most urgent. These are: *digestion of existing data*, *general assistance in empirical research* and *developing better methods to achieve these goals*.

DIGESTING EXISTING DATA

In an industrialized European country, like the one I represent, a vast amount of data is produced by the central statistical office. It is very hard to give a precise figure of how much data is supplied. A moderate estimate is that two million time series are published in print each year. On request 20-30 times that amount is available to the public. The office has been working for nearly one hundred years.

These figures apply more to the level of development of a country than to its size. The latter affects more the magnitudes of the figures than the amount of data.

(*) "Tilasto" meaning a collection of data and "tilastotiede" meaning the science for (a collection of) data.

Since further analysis of this data mainly takes place outside the production unit it is even harder to say how much the data is further analyzed. My guess is that an extremely small fraction of it is ever analyzed using statistical methods more sophisticated than comparing the magnitudes of a few figures. Averages are sometimes applied but measures of disperson are almost unknown. In some cases graphical displays are used. Ad hoc time series methods, such as seasonal adjustment, are common in processing economic time series. The amount of time series that would go into any more advanced statistical processing probably is closer to a three digit number than to a five digit one.

The core of statistical theory is, in my view, making inference from quantitative observations. In parametric statistics we are interested in distributional properties of data, such as location, uni- and multivariate dispersion, skewness etc. Available data provides limitless possibilities to apply what methods we know and to develop new ones.

To give a concrete example of the possibilities let us think of a manageable collection of time series. Their joint distribution can be classified into three main cases:

(i) Each series is white noise and they are mutually independent for all lags,
(ii) The series are mutually independent but they are serially correlated,
(iii) The time series are both serially and cross correlated.

I will study these cases from a time series analyst's point of view. The best a statistician can do in case (i) is to show that the variables can be regarded as white noise, show what the series look like and estimate the means and the variances.

In case (ii) a statistician can apply univariate ARIMA techniques to find a model for each time series. The models act as filters that turn mutually independent autocorrelated time series into independent white noise series, i.e. to case (i). The models can be used for forecasting. ARIMA and spectral techniques can be combined for decomposing the series (seasonal adjustment) etc.

When joint dependence prevails between the time series, as in case (iii) they can be filtered into mutually independent white noise using e.g. multivariate ARIMA techniques. Thereby we may get an idea about the true behavioural mechanisms that are reflected in the data and we would be able to simulate them.

It is rather difficult to find good examples of time series that are to be regarded as mutually independent white noise. However, just estimating a mean and maybe a variance means that from a statistician's point of view we are acting as if the data were white noise.

It took us a century to build the data base we have today. Let's now start building a research structure on this base. Processing equipment is becoming better and cheaper every day and we possess the necessary knowhow to start building. An easy starting point is to build univariate ARIMA models producing forecasts that would routinely be published together with observed data. A more distant goal is multivariate models, the errors (shocks) of which may be of enough interest to be published together with the observed series. Decision makers would focus their interest on impulses, not directly osbervable, reaching the system and on structural changes more than on the normal, model-specified working of society.

ASSISTANCE IN EMPIRICAL RESEARCH

It is hard to think of any empirical research that could not benefit from statistics. In fact scientific experiment is the cradle of modern statistical thought as introduced by R.A. Fisher. Still, I have the impression that in many, if not in most, empirical research teams the professional statistician is absent. We may ask: how do biologists, physicists, economists etc. manage to do empirical research without the assistance of statisticians?

As I see it there are three ways to avoid a statistician:

(i) You conduct your experiment as much as possible in the same way as others have done before you. Assuming that the method is applicable in your case the only thing that can go wrong is that the method was wrong in the first place. But then you would be in good company.

(ii) Empirical scholars are given basic training in statistics. A biologist can apply a t-test, a physicist knows something about stochastic processes and an economist has taken a basic course in econometrics. Instead of using this knowledge as a working language for communicating with a statistician you can stick to the rule of "do it yourself".

(iii) Collect your data as well as you can and bring it to the computer center. They will have somebody with some statistical training who knows all the ready-made program packages. At least one will fit your purposes.

Why would people avoid statisticians? I see two important reasons: *ignorance* and *convenience*. Many statisticians have a large share in the first sin and we are not completely innocent on the second account either.

Ignorance first. People simply do not know that we could be of assistance. Too few statisticians have endeavoured to market their knowhow. For that it is necessary to meet the empiricist halfway by learning at least as much about his/her subject as he/she knows about statistics.

Convenience. It is more convenient for the empiricist to try some of the avoiding tricks than to have to rely on a strange theoretician who may even turn all you have done upside down and with whom you would have to share the glory. For the statistician it may be much more convenient to indulge in some theoretical problem that few, if any, people can understand and criticize. This kind of research gets green light more because of its contribution to other statisticians' feeling of ignorance than due to its contribution to knowledge.

What can we do to improve the situation? I think that the answer lies in changing attitudes, ours and theirs. Every statistician should get training in soiling his hands with real - not only textbook - data. In practice this implies that every statistics department practices consulting where students under the supervision of senior statisticians, help to solve empirical research problems. The students would also have to acquire at least basic knowledge of an applied field. On the other hand, students of an applied field would get accustomed to the idea of conducting joint research with statisticians. This policy aims at changing attitudes where they can be changed: in young age groups.

METHODS

If some of the criticisms of theoretically minded statisticians for being ignorant and arrogant can be well founded, some of it may be beside the point. No theories are final and methods applied in practice are often just the tip of an iceberg of theory. During the development of that theory it was not always clear where it could be put into practice.

Practitioners must realize that there is no substitute available for mathematical language. It is an auxiliary device that helps us to understand. Those for whom mathematics is an end in itself would more properly be called mathematicians and those at the other end who substitute numerical observations for symbolic variables are applied people in the subject proper. Both are needed but let's keep the road open between the two ends.

These thoughts could be summarized in one of our favorite gadgets, the two by two table:

Theory \ World	simple	complicated
simple	A	B
complicated	C	D

Sometimes our problems are simple and require simple solutions (A) but maybe more often the world is complicated and problem solution requires complicated theory (D). Suboptimal behaviour occurs in the off diagonal entries. In C we are talking beside the point and in B we want to see the world as simpler than it really is. It is not always just theory that is hard to understand - it may be the world.

DISCUSSION

The theme of this meeting is a popular one. In Scandinavia alone it has been a topic of discussion at four international meetings since 1970 (cf. Statistical Review of the Swedish National Central Bureau of Statistics 2/1978). On a world scale the discussion has been going on for decades. Many of the ideas put forward at this meeting have probably been expressed before. I have no idea about the effect these discussions may have had. The time has come to talk about the implementation of all the good suggestions.

Here the International Statistical Institute (I.S.I.) comes to my mind. It has the necessary authority to issue international recommendations that could improve and coordinate the teaching of statistics.

The I.S.I. has traditionally published bibliographical studies. It could, as a next step, help to improve the status of statistics by publishing papers on the history of statistics. Not only should every statistician know the historical background of his knowledge but in a popular form this narrative should be distributed as widely as possible. It should be as natural to connect Bayes and Fisher (*) with statistics as Linneaeus with botany and Keynes with economics.

Finally, the discussions about the role of statistics have a social aspect. This is the occasion when we can laugh at ourselves. We go as far as quoting classical poems, such as those about slaying dragons and identifying elephants. In the following poem I have tried to contribute to that tradition:

(*) Cf. J. Fisher-Box: "R.A. Fisher, the Life of a Scientist", John Wiley and Sons Inc., New York, 1978.

STATISTICS

Statistics to the uninformed
Is columns, rows that on us stormed
When handling data became the feat
Of dilettantes you couldn't beat
In asking questions about you
When you had many things to do.

But the informed would simply frown
And tell a story, much of his own
Of what statistics really is
(Letting all his s's fizz).
And not without a bit of pride
He thus would unto you confide:

No! Statisticians never bother you.
They write no numbers, except maybe one or two
Expressing confidence in what they believe
(Some do not even these conceive).
They assume the world has been or is to be observed
By which their peace of mind can greatly be preserved.

They use their hieroglyphs to show where you should look
(Rules for the design can be found in many a big book).
And when the observations have been picked
They tell you how to make inference, how to predict,
And suddenly a chaos of complexity
Seems to have turned into a system of simplicity.

Alas! The world does not always fit into our abstraction
And this can hardly be improved by mere obstruction.
So let's not forget the guy messed up in data.
And if, for fun, we sometimes get absorbed by *epsilon* and *eta*
At least it should be something new and true and nice
For such extravagance to be worth its price.

BUILDING BRIDGES BETWEEN THE ACADEMIC AND REAL WORLDS - SOME OBSERVATIONS FROM SOUTH AMERICA

C A Peres and P A Morettin, University of São Paulo

INTRODUCTION

The purpose of this paper is to summarize the present state of the statistical education and profession in South America, with special emphasis on Brazil. We analyze the problems of statistical education, training and research and indicate who the main users of the statistical methodology are.

In recent years some attention has been devoted to the problem of closing the existing gap between the academic statisticians and the professional statisticians working in government agencies, industries, etc. We refer in particular to the Conference on Transfer of Methodology Between Academic and Government Statisticians (1978), sponsored by the American Statistical Association (which will be referred to briefly as CTM).

Concerned especially with Latin America we mention the SELENES (Latin-American Seminar on the Teaching of Statistics) (1978) and a Meeting of the Working Group to Study the Status of the Statistical Profession in Western Hemisphere (1980), supported by the Inter-American Statistical Institute (IASI) (referred as WG for brevity).

We conclude by discussing an experience at the University of São Paulo, showing how a consulting course, which is taught to undergraduate students, can help as one means of closing this gap betwen the academic and real worlds.

EDUCATION AND TRAINING OF STATISTICAL PERSONNEL

In this section we try to summarize some sparse information, available from different sources, on the stage of the development of the education and training of statisticians in South America.

By education we mean that an academic degree is granted, at the undergraduate or graduate level. By training it is meant that special courses are offered in order to supply limited knowledge for special tasks (see Gura, 1980). Usually, this training is an in-service training, that is, given after the employee is employed by the organization. See CTM, Chapter IV.

Murthy (1978) states that the objectives of education are "to develop a statistical sense with a view to understanding socio-economic situations and natural and physical phenomena through statistics and teaching the methodology appropriate to produce statistics economically and to analyze and use them meaningfully", while the objective of training is "to impart special skills required to plan and undertake statistical operations necessary in producing and analyzing

statistics".

Concerning undergraduate education in South America, thirteen countries have programs in Statistics; in Brazil, seventeen schools have such programs. At the graduate level, seven countries have programs leading to a master's degree and none to a doctoral degree. In the case of Brazil, seven centers offer graduate study at master's level. See Appendix A and B.

Two main conclusions were drawn in WG referring to education and training:

(a) "The principal limiting factor for the establishment of appropriate programs is the absence of teaching personnel adequately qualified for such responsibilities. As a result, the students participating in various programs, frequently receive low quality training or training that is not oriented toward the needs of government and industry, the most likely sources of employment;

(b) the absence of well qualified teaching personnel is principally due to the scarcity of resources allocated to graduate statistical training in most of the countries of the region".

In some countries (Brazil and Chile, for example), efforts are being made to implement reciprocal cooperation with more advanced centers in Europe and North America, with the objective of establishing doctoral programs.

Due to this lack of graduate programs (at least at the doctoral level), the solution has been to send people abroad to obtain advance training. This is of course costly and many countries do not have the resources to support adequately such programs. This is also a long run policy and as mentioned by Bradley (1982a) "the first priority for the use of available qualified statisticians should be the well-supported staffing of needed training programs in Statistics, programs developed centrally on a national basis or, when appropriate or necessary, on a regional basis".

Statistics is relatively recent as a discipline. Only in the last three decades university education has had an increasing expansion. Due to historical factors, the teaching of statistical disciplines were done in Mathematics departments, hence the principal emphasis has been on the theoretical side. The implication is that the students receiving this biased formation do not fill the needs of industry or government offices.

Most of the countries try to provide "ad hoc" training facilities, generally through National Statistical Offices, in order to prepare the personnel for their needs. The exception, perhaps unique in South America, is a course offered by CIENES, called "Program in Statistical Techniques", which is intended to train those persons who will work in statistical offices in government.

In some countries the needs vary from one region to another. In any event,

as we mentioned above, the government seems to be the main source of jobs for statisticians coming out of universities.

In this connection it is very important that the curriculum of statistical programs at the universities should be composed of disciplines relevant to the immediate applications of the region. For example, if the principal needs are statisticians for statistical offices, a curriculum centered in economics seems to be the adequate answer, with courses including principles of sampling and survey design, survey methods, construction of tables, graphical methods, data analysis, etc. (CTM, Chapter III). If industry is a major source of jobs, the emphasis should be on regression analysis, design of experiments, response surfaces, quality control, time series, etc. (Marquardt, 1979).

A way to balance the statistical education of future statisticians is suggested in Appendix C. The idea is to use a flexible structure of modules in order that the students acquire a substantial subject-matter content from other fields, as social sciences, economics, biology, etc. In one semester the students receive the basic information in the field (Economics, Epidemiology, Psychology etc.) and in the following semester a statistically oriented course (Econometrics, Biometrics, Psychometrics, etc., respectively).

Bradley (1982a) advocates that "on-campus statistical consulting in support of university research programs may be initiated". This recommendation is further discussed in "A Suggestion for Building the Bridge between the Academic and Real Worlds" further on.

A point of apprehension is that the enrolments in statistical courses is decreasing in the last years. This is discussed by Bradley (1982b) for the USA case, but it is true for South American countries also. A recent survey among freshmen in Mathematical Sciences at the University of Sao Paulo, showed that 80% of the students prefer to pursue studies in Computer Sciences, while only 6% indicated a preference for Statistics.

This fact is part of an overall situation, where enrolment of talented students in Mathematical Science courses (except Computer Science) has been decreasing steadily, due mainly to the low salaries paid to high school and university professionals, compared with those paid by industries, for example.

USERS OF STATISTICS

The major sources of jobs for statisticians in South America are the government and the universities. This is true even in heavy industrialized regions, as São Paulo, Brazil, where industry, commerce and market research could generate large amounts of jobs.

The main reason for this situation is the lack of knowledge of most other sectors that a statistician could be very useful in helping them increase productivity and maximize profit. They prefer to hire an economist or an engineer, to perform a job which is better suited to a statistician.

We think there is not enough promotion of the statistical profession in South American countries to show the importance of the statistician in any decision-making process in industry, government, management, etc.

A further drawback is that most of the industries that operate in the region are multinational companies, and they bring from their home countries all the necessary technology to function, including the statistical technology. This introduces a further limitation in the work of statisticians and surprisingly enough, in some instances, the supply of statisticians is greater than the demand.

Of course the topic on promotion of the statistical profession is intimately tied to other factors already mentioned, like education and training, the existence of professional and scientific meetings on the field, statistical publications, and others.

From what has been said before, it might be argued that an effective promotion of the statistical profession could be counter-productive, since on the whole we would not have, in the South American region, enough statisticians prepared to reach the demand for them. On the other hand, in some areas, the number of statisticians coming out from undergraduate courses each year is more than enough to meet the demand.

To further confuse the situation, in Brazil Statistics is a profession regulated by law, as are engineering and medicine, and only those statisticians who had their B.S. in Statistics have permission to work professionally, the only exception being the university personnel. It follows that most of the professionals working in industry, government offices and other places besides universities have only a B.S. in Statistics, while those who pursued some graduate program usually work in universities.

It is a fact that some universities in the USA, Europe and other countries are trying to reformulate their master's programs in Statistics, in order to prepare the graduate students to attack real world problems. That means that the master's program usually has a more "applied" character. Even if these students did not have an undergraduate degree in Statistics, they will have a greater maturity than those students who had an undergraduate program in Statistics and immediately after coming out of the university, find a job in industry or government.

In this connection, to try to remedy the situation caused by this regulation, an experiment was introduced in the Statistics undergraduate program at the University of São Paulo, which we discuss in the next section.

A SUGGESTION FOR BUILDING THE BRIDGE BETWEEN THE ACADEMIC AND REAL WORLDS

The undergraduate program leading to a B.S. degree in Statistics, at the University of São Paulo, is a four-year program, the first year being a basic one for all mathematical science students (see Appendix C).

A two-semester course in consulting was introduced in 1978, in the last year of the program. We describe now how this course was designed and make some comments on its results.

Two members of the staff of the Statistics Department are responsible for the course, each year, although every professor can supervise the statistical analysis of some project. The course is composed of three parts.

Part I - Design and analysis of research projects developed in several areas of the University;

Part II - Training period in some firm or government office;

Part III - Development of statistical methodologies used frequently in Part I and II and not included in the undergraduate curriculum (complementary subject).

The Statistics Department has a Statistical Laboratory, where Part I takes place.

Description of Part I

i) A researcher (from now on called simply a "user") comes to the Statistical Laboratory and fills in appropriate forms with information about himself, the objective of the research and the research itself;

ii) each project is supervised by one member of the staff of the Statistics Department and performed by two students. An appropriate statistical methodology is chosen, after an interview with the user;

iii) the supervisor and the students make an outline of the project to all students of the consulting course. The user and the professors responsible for the course attend the exposition. During the discussion the students usually ask many questions and the user plays an important role in the discussion;

iv) after the statistical analysis is completed, the students have to write a report (always in close contact with the supervisor), which is submitted

to the appreciation of the user before it is typed;

v) a copy of the original data and a copy of the report (in the form of an "Applied Statistics Report") are kept in the files of the Statistical Laboratory. Usually they can be used as "case studies" in classes of several disciplines.

The report does not have a fixed structure, in order to motivate the creativity of the students. But usually they contain the description of the experiment, the statistical analysis, a summary of the conclusions and a reference list. In Appendix D we present a table listing the projects concluded in 1981.

Description of Part II

Since Statistics is a professional career in Brazil, the undergraduate program must contain a minimum of disciplines regulated by federal law. This includes a training period during two semesters. Most of the students do this training period in government offices, private banks and a few in industries. At the end of each semester we send questionnaires to these places with the purpose of evaluating the performance of the students.

Description of Part III

Part of the work load of the consulting course (2 hours) is devoted to lectures covering statistical methodologies not covered during the normal academic program. For example, this year, the lectures include topics such as categorical data analysis and cluster analysis.

We remark that the time spent in the training period is extra to the class period. The final grade of the student in each course is an average of those in all three activities discussed above.

FINAL COMMENTS

It is a fact (WG, p.5) that statistical activities in South America are basically carried out at the academic and government level, and to close the gap between the objectives of each one, it is recommended that a greater quantity of subject matter useful in the governmental area should be included in the academic curricula. Moreover, in those areas where industry, banking, commerce, market research may be sources of jobs, a balanced curriculum should be taught.

It is fundamental that the statistical profession be promoted adequately, and one appropriate tool for this purpose is the establishment of national statistical societies, holding regular scientific meetings and providing a substantial coverage of all sectors of the profession. Publications at the national or regional level are also of great importance.

The curricula of the university level statistical training should offer an appropriate balance of theory and methodology, together with practical work, carried out either in agencies that generate official statistics or in other sectors of society that use statistical techniques.

A suggestion is made of a consulting course which has shown to be of great value in order to achieve the goals mentioned above.

References

[1] American Statistical Association (1978) Report on the ASA Conference on Transfer of Methodology Between Academic and Governmental Statisticians. Washington, D.C.

[2] Bradley, R.A. (1982a) Statistical Training in Developing Countries, Int. Statist. Educ. Newsletter, Vol.2, No.2, a publication of ISI.

[3] Bradley, R.A. (1982b) The Future of Statistics as a Discipline, Journal of the American Statistical Association, Vol.77, p.1.

[4] Gura, B. (1980) La Profession Estadistica en el Hemisfério Occidental: Situación, Educación y Capacitación (Con Referencia Especial a la America Latina), IASI, Washington, D.C.

[5] Interamerican Statistical Institute (1980) Report on the Meeting of the Working Group to Study the Status of the Statistical Profession in the Western Hemisphere, Santiago, Chile.

[6] Marquardt, D.W. (1979) Statistical Consulting in Industry, The American Statistician, Vol.33, p.102.

[7] Murthy, M.N. (1978) Statistical Education, Training and Research and Their Utilization, in CTM, Chapter III, p.68.

[8] Seminário Latinoamericano Sobre Enseñanza de Estadistica (1978) Informe Final, CIENES, Santiago, Chile.

APPENDIX A

Schools with a Statistical Program in Brazil, Undergraduate Level (by state)

1. São Paulo - University of São Paulo
   Federal University of São Carlos
   State University of Campinas
   Faculdade Paes de Barros

2. Rio de Janeiro - National School of Statistics
   Federal University of Rio de Janeiro
   State University of Rio de Janeiro
   Catholic University at Rio de Janeiro

3. Pernambuco - Federal University of Pernambuco
   Catholic University at Recife

4. Ceará - Federal University of Ceará

5. Rio Grande do Norte - Federal University of Rio Grande do Norte

6. Minas Gerais - Federal University of Minas Gerais

7. Paraná - Federal University of Paraná

8. Bahia - School of Statistics of Bahia

9. Rio Grande do Sul - Federal University of Rio Grande do Sul

10. Brasília - University of Brasília

11. Paraíba - Federal University of Paraíba

APPENDIX B

Schools With a Master's Program in Statistics (by state)

1. São Paulo - University of São Paulo
   University of São Paulo at São Carlos
   University of São Paulo at Ribeirão Preto (Biostatistics)
   University of São Paulo at Piracicaba (Agriculture)
   State University of Campinas

2. Rio de Janeiro - Federal University of Rio de Janeiro
   Institute of Pure and Applied Mathematics,
   IMPA (National Research Council)

3. Pernambuco - Federal University of Pernambuco

4. Brasília - University of Brasília

APPENDIX C

Suggestion of a Curriculum for an Undergraduate Course in Statistics
(Where the main demands are for statisticians working in industry and government)

- Basic for all students in Mathematical Sciences

1st SEMESTER - Calculus I
Vector and Geometry
Introduction to Computer Science I
Introduction to Statistics I
Physics I

2nd SEMESTER - Calculus II
Linear Algebra I
Introduction to Statistics II
Introduction to Computer Science II
Physics II

- B.S. in Statistics

3rd SEMESTER - Calculus III
Linear Algebra II
Computer Programming
Statistical Inference I

4th SEMESTER - Calculus IV
Complex Variables
Exploratory Data Analysis
Probability
Statistical Inference II

5th SEMESTER - Introduction to Stochastic Processes
Nonparametrics
Sampling Techniques
Design of Experiments I

6th SEMESTER - Design of Experiments II
Regression Analysis
Multivariate Analysis
Applied Stochastic Proesses

7th SEMESTER - Statistical Consulting I
Modulus I
Modulus II
Modulus III

8th SEMESTER - Time Series Analysis
Statistical Consulting II
Modulus I
Modulus II
Modulus III

The disciplines in the modules may vary from one year to another. An example follows:

Module I (subject-matter discipline allied to a discipline with a statistical character)

Microeconomics and Econometrics
Epidemiology and Biometry
Vital Statistics and Demography
Psychology and Psychometrics
Biology and Biometry

Module II (Quantitative specialization)

(Math.) Mathematical Analysis I and Mathematical Analysis II

(Bus. & Econ.) Operations Research I and Operations Research II
Accounting I and Accounting II

(Eng.) Industrial Organization I and Industrial Organization II

(Comp. Sci.) Mathematical Programming and Simulation
Mathematical Programming and Graph Theory

(Physics) Statistical Mechanics and Statistical Methods in Experimental Physics

Module III (Disciplines in the Statistics Area)

Game Theory, Information Theory, Decision Theory, Statistics in Market Research, Categorical Data Analysis, etc.

APPENDIX D

Research Projects, 1981 - Statistical Laboratory, University of São Paulo

PURPOSE / AREA	Research Leading to: Doctor Degree	Master Degree	Publication	TOTAL
Architecture	-	-	1	1
Marine Biology	-	-	1	1
Biosciences	2	3	1	6
Botany	1	-	-	1
Biomedical Sciences	2	-	1	3
Pharmaceutical Sc.	2	2	-	4
Education	-	-	1	1
Nursing	-	2	1	3
Engineering	-	-	1	1
Social Sciences	1	-	1	2
Medicine	4	3	4	11
Dentistry	-	2	1	3
Psychology	1	5	1	7
Public Health	1	-	-	1
TOTAL	14	17	14	45

BUILDING BRIDGES BETWEEN THE ACADEMIC AND REAL WORLDS

B Jones, B J T Morgan and G B Wetherill, University of Kent at Canterbury

INTRODUCTION

In the following we describe how statisticians at the University of Kent have, over the last ten years, managed to bridge the gap between the 'academic' and 'real' worlds. This has not been accomplished easily but by dint of very hard work. Our description is divided into five sections. In the first of these we list some of the collaborative Ph.D. projects that have been undertaken by students working in the Department and in the following two sections we discuss some of the advantages and disadvantages associated with this type of collaboration. Following this we mention other types of collaboration and end with some final remarks. A list of references to some of the published work done by the staff and students at Kent is given at the end.

TITLES OF SOME Ph.D. PROJECTS UNDERTAKEN WITH OUTSIDE BODIES

1. S.H. Bishop - Three-way and higher cross-over designs in clinical trials - Pfizer Central Research, Sandwich, Kent.
2. P. Bravo - Acceptance sampling plans by variables for percent defective, with emphasis on double sampling plans - in association with the Ministry of Defence.
3. J.B. Curram - Modelling case studies in economic quality control - Amersham International Ltd., Buckinghamshire.
4. C. Daffin - Statistical studies in long term economic forecasting - I.C.I. Mond Division, Runcorn, Cheshire.
5. K.V. Darby - Statistical examination of British Bird observatories data - B.T.O., Tring, Hertfordshire.
6. P. Duncombe - Some problems in writing user-friendly intelligent regression software - I.C.I. Mond Division, Runcorn, Cheshire.
7. T.M.M. Farley - Statistical problems of chronic toxicity - Pfizer Central Research, Sandwich, Kent.
8. J.P. Hinde - Statistical principles of quality control procedures for saturation analysis assay kits - Amersham International Ltd., Buckinghamshire.
9. B. Jones - computer assisted design of experiments - I.C.I., Mond Division, Runcorn, Cheshire.
10. P.M. North - Statistical methods in ornithology - B.T.O., Tring, Hertfordshire.
11. C.C. Robertson - Statistical applications in memory studies - In association with a number of psychologists at Cambridge and Oxford.

12. S.A. Watts - Stochastic aspects in the use of compartment models related to pesticide evaluation - Shell Research Ltd., Sittingbourne, Kent.

THE ADVANTAGES

Statistics is a practical subject and through collaborative work we have been trying to develop practical statistical methods. For instance the two research studentships conducted in conjunction with the British Trust for Ornithology involved large sets of data which had been gathered by the Trust but not looked at in any detail and certainly not by a statistician. Results of the collaboration here have been a variety of new approaches, new methods and research papers, all of which have helped to bring the data alive. In this case, a large numbr of voluntary workers are involved in the data collection and it is extremely important that they should receive some form of feedback on the interpretation of the data. Birds are important indicators of changes in the environment and statistical methods here are clearly playing a wider, more important, rôle. Out of this collaboration arose a recent One-Day Biometrics Conference in London on "Statistical Methods in Ornithology". The research project of Simon Watts, in collaboration with Shell Research at Sittingbourne, dealt with stochastic aspects in the use of compartment models for describing experiments to investigate pesticide evaluation. Here again Statistics was being used with the protection and monitoring of the environment in mind.

The results of a collaboration can also be the development of useful tools which non-academic statisticians can use as part of their day to day work. For example, algorithms to search for optimal experimental designs were developed as part of a project with I.C.I. These algorithms are being used in a number of centres worldwide. The Caribbean Agricultural Research and Development Institute in Trinidad have versions which run on quite small computers and are used regularly to find designs for non-standard situations. A further example is provided by another I.C.I. project in which a user-friendly, interactive computer program has been developed for analysing regression data. The main feature of the program is that it has been developed for use by non-statisticians. In the development of this program the student worked as part of a larger team but nevertheless played a major rôle.

The contact with I.C.I. for example has led to further research contracts to extend the work done by the students. This extended contract has led to a useful working relationship which has proved profitable to both sides. As is apparent from the list in section 2, a number of contacts have supported more than one project over the years.

THE DISADVANTAGES

It is difficult to set up collaborative awards. A large amount of groundwork has to be undertaken and goodwill has to be established and maintained with the collaborating body. This frequently involves the production of regular research projects, which, in the last resort, do help the student with the writing-up stage. Furthermore a large amount of travelling may be involved, if, for instance, the collaborating body is in Runcorn or Oxford or in Tring, say. Another disadvantage is that it may not always be possible to follow-up the research and the result of a CASE Award may just be the scratching of the surface, the opening up of areas which may then not be developed to their full potential. A final disadvantage is that these projects, when the aim is for a Ph.D., do have to try to satisfy a double aim, that is, the student wants to produce an academically respectable and acceptable dissertation but, at the same time, the collaborative body has to be kept happy. This suggests that the period of study should, in fact, be longer than the usual three years and a further pointer to the same conclusion lies in the need of the student to spend time and regular visits with the collaborating body, as well as spend quite possibly an appreciable amount of time becoming familiar with the practical background, quite apart from the necessary study of the theoretical background.

Supervision of these projects can also be additionally difficult and time-consuming since the supervisor also has to become acquainted with what could be an unfamiliar area.

There are also other disadvantages, not the least of these is finding a suitable student. As many of these projects sound specific and quite difficult, possibly being worded in jargon which is not current in statistics, it can be difficult to find a student. On occasions detailed investigations shows that the problem is not there! On other occasions the original industrial contact vanishes in between the signing-up of the student and the student starting. Similarly, on occasions, the industrial goes 'busy' and then later comes up with some vital information which was required at a much earlier stage.

OTHER FORMS OF COLLABORATION

Some other forms of collaboration with non-academic contacts that have been undertaken are those in which the collaborating body has financed one or two members of the lecturing staff to undertake specific research tasks. Two in particular are

(a) Cluster analysis of the elderly population of England.
(b) The Household Formation Project.

Three members of staff were involved in (a) and in fact the original project which was with the Department of the Environment (DOE) was extended and further financed by the Department of Health and Social Security (DHSS) and the DOE. One outcome of this work was a number of published papers not only in statistical journals but in those aimed directly at workers in the field.

In project (b) two members of staff and two research associates were contracted by the DHSS to consider how social and economic information might be used to forecast numbers of households.

Further contract work currently underway involves

(i) I.C.I. Mond Division - development of an interactive user-friendly computer program for the analysis of regression data.

(ii) The Overseas Development Administration - development of a user-friendly survey analysis package for the micro-computer.

(iii) The D.H.S.S. - cluster analysis of the English Personal Social Services Authorities.

The above projects have been organised by the Applied Statistics Research Unit at Kent which has been set up specifically to undertake contract work from outside bodies. The Unit, which is run by a Director and an Assistant Director, employs a number of staff to undertake the routine statistical work. The pharmaceutical industry in particular has made wide use of the Unit's facilities. Members of the lecturing staff are also involved in the Unit's work either in a consultancy role or directly contracted to undertake statistical work.

A two-year project on intercropping led by Professor S.C. Pearce and financed by the Overseas Development Administration was also recently completed at Kent.

Another form of collaboration is through the Collaborative Training Award scheme financed by the S.E.R.C. At the present time one student is working on the Papadakis method in collaboration with Shell Research at Sittingbourne, Kent, and in October a student begins work on normality testing for Smith, Kline and French Ltd. at Welwyn Garden City, Hertfordshire.

SOME FINAL REMARKS

Some concluding remarks are as follows:

The solution of practical problems so often comes up with the need for theoretical research. Examples of this can be given very easily from the examples given above and references given below.

Some industries are severely put off by their contact with universities. Attempts by universities to build bridges on their own terms will do more damage to this cause than anything else. Some people are interested in bridges with the real world provided, for example, the real world is interested in their particular stochastic process. Alternatively the original problem posed by the industry is diverted into something of more interest academically. One has to realise that for the industrial statistician he has a problem and he wants it solved - anything else is, for him, a pure by-product.

References

[1] Bassett, E.E., Clewer, A., Gilbert, P. and Morgan, B.J.T. (1980) Forecasting Numbers of Households: The Value of Social and Economic Information, Report - U.K.C.

[2] Bravo, P.C. (1980) Sampling inspection by variables with known variance. Bulletin in Applied Statistics, Vol.7, No.2, pp.203-222.

[3] Bremner, J.M. (1978) Algorithm AS 123. Mixtures of beta distributions. Applied Statistics, 27, pp.104-109.

[4] Jolliffe, I.T. (1980) Multivariate analysis and weather forecasting. Probabilistic and Statistical Methods in Weather Forecasting, pp.103-108, World Meteorological Organisation.

[5] —— (1979) How accurate are long-range weather forecasts? New Scientist, No.84, pp.192-194.

[6] Jolliffe, I.T., Jones, B. and Morgan, B.J.T. (1982) Utilizing clusters: a case study involving the elderly. Journal of the Royal Statistical Society, Series A, Vol.145, pp.224-236.

[7] Jones, B. (with J.A. Eccleston) (1980) Exchange and interchange procedures to search for optimal designs. Journal of the Royal Statistical Society, Series B, Vol.42, pp.238-243.

[8] —— (1980) Exchange and interchange procedures to search for optimal row-and-column designs. Journal of the royal Statistical Society, Series B, Vol.42, pp.372-376.

[9] Jones, B. (1980) Combining two component designs to form a row-and-column design. Applied Statistics, Vol.29, pp.334-337.

[10] Jones, B. (1980) The Computer as a Statistical Consultant. Bulletin in Applied Statistics, Vol.7, No.2, pp.168-195.

[11] Köllerström, J. (with T.C. Tozer) (1980) Penalties of hard decision in signal detection. Electronics Letters, No.16, pp.199-200.

[12] Köllerström, J. and Wetherill, G.B. (1981) Sampling by variables with a normal-gamma prior. In: H.-J. Lenz, G.B. Wetherill and P.-Th. Wilrich (eds) Frontiers in Statistical Quality Control, Physica-Verlag.

[13] Morgan, B.J.T. and Robertson, C.C. (1980) Short-term memory models for choice behaviour. Journal of Mathematical Psychology, Vol.21, pp.30-52.

[14] Morgan, B.J.T. and Watts, S.A. (1980) On modelling microbial infections. Biometrics, Vol.36, pp.317-321.

[15] Morgan, B.J.T. and North, P.M. (1980) On using cluster-analysis and multi-dimensional scaling for describing animal movements and bird territories. Biometrical Journal, Vol.22, pp.525-533.

[16] North, P.M. and Morgan, B.J.T. (1979) Modelling Heron survival using weather data. Biometrics, Vol.35, pp.667-681.

[17] North, P.M. (1977) A novel clustering method for estimating numbers of bird territories. Applied Statistics, Vol.26, pp.149-155.

[18] North, P.M. (1981) Models for a simple capture-recapture procedure to estimate population size. Biometrics, Vol.37, pp.661-672.

[19] North, P.M. (with R.M. Cormack) (1981) On Seber's method for estimating age-specific bird survival rates from ringing recoveries. Biometrics, Vol.37, pp.103-112.

[20] Paul, S.R. (1979) A clumped beta-binomial model for the analysis of clustered attribute data. Biometrics, Vol.35, No.4, pp.821-824.

[21] Paul, S.R. (1979) Models and estimation procedures for the calibration of examiners. British Journal of Mathematical and Statistical Psychology, Vol.32, pp.242-251.

[22] Pearce, S.C. (with B. Gulliver) (1979) Graphical assessment of intercropping methods. Journal of Agricultural Science, Vol.93, pp.51-58.

[23] Pearce, S.C. (1980) Randomized blocks and some alternatives: a study in tropical conditions. Tropical Agriculture (Trinidad), Vol.57, pp.1-10.

[24] Pearce, S.C. The agricultural field experiment: a statistical examination of theory and practice. John Wiley, Chichester.

[25] Robertson, C. (1982) Modelling recall order in free recall experiments. British Journal of Mathematical and Statistical Psychology (to appear).

[26] Wetherill, G.B. and Köllerström, J. (1979) SPRT's for the normal correlation coefficient. Journal of the American Statistical Association, Vol.74, No.368, pp.815-821.

[27] Wetherill, G.B. (with P.C. Bravo) (1980) The matching of sampling plans and the design of double sampling plans. Journal of the Royal Statistical Society, Series A, Vol.143, Part 1, pp.49-67.

[28] Wetherill, G.B. (with H.-J. Lenz and P.Th-Wilrich) (1981) Editors of Frontiers in Statistical Quality Control, Physica-Verlag.

[29] Wetherill, G.B. et al. Advanced Regression Analysis (in preparation).

CLOSING THE GAP BETWEEN STATISTICIANS AND STATISTICS USERS

Guido E Del Pino and Carlos Prado-Campos, Catholic University of Chile*

SUMMARY

In this paper we discuss :

a) The present state of the teaching of statistics in Chile, with particular emphasis on the Catholic University of Chile at Santiago (U.C.).

b) Our experience with attending the demands of several users.

c) Some possible solutions to close the gap between statisticians and statistics users.

We also offer some comments on the importance of closing the gap mentioned in c).

INTRODUCTION

Closing the gap between statisticians and experts in other fields is very important on at least two counts. On the one hand, major statistical developments have been made by the users themselves. This is the case, for instance, with psychologists and multivariate analysis, economists and Bayesian analysis, astronomers and variance components analysis, to mention only a few. On the other hand, if statisticians are not acquainted with some field of applications, statistics becomes something too theorical to be of any use. Since statistics is a relatively new science, there tends to be a natural gap between statistics and the needs of its users.

It is then not surprising that important statistical associations like the International Statistical Institute (ISI), the American Statistical Association (ASA) and the Inter-American Statistical Institute (IASI) have been seriously concerned with the study of this problem through several meetings, such as Conference on Transfer of Methodology between Academic and Government Statisticians, Reston, Virginia, USA (ASA), the Latin-American Seminar on the Teaching of Statistics (SELENES) Santiago, Chile, 1977 (IASI) and the Round Table Discussion on the University Teaching of Statistics in Developing Countries, The Hague, Holland, 1968.

THE TEACHING OF STATISTICS IN CHILE

A prominent role in the teaching of statistics in Chile has been played by the Inter-American Statistical Training Center (CIENES) and by its predecessor,

the Inter-American Center for Financial Statistics (CIEF) that was established in 1952, both organisations being dependent on the Organization of American States (OAS).

It is a fact that there is no university in Chile with a statistics program whose faculty does not have a graduate from CIENES. On top of this, it is also true that most Chilean statisticians with a Ph.D. degree have been previously on the CIENES programme.

Undergraduate statistical programmes offered by the universities were initially three years long and they led to the professional degree of Statistical Technician. This situation has now changed and at present some of the programmes offered are at least four years long. The universities with programmes in Statistics are shown in exhibit 1.

The curricula of the undergraduate programmes offered by the Catholic University of Chile and by the Catholic University of Valparaiso are shown in Appendix 1.

THE TEACHING OF STATISTICS AT THE CATHOLIC UNIVERSITY OF CHILE

The first step was taken in 1971 with the creation of a professional programme of statistical technician, with a length of three years. The aim of this programme was to prepare a professional to support users undertaking studies that needed statistical methodology. For this reason the curriculum included courses in sampling theory, analysis of variance, computer science and quality control.

This programme was in force for five years. During this period the need was detected for producing a professional of higher category, able to assume bigger responsibilities and consequently earning a higher income. For this reason the programme was replaced by a four year programme leading to the degree of statistician, which incorporated a broader range of courses of formative value.

The results otbained to date are satisfactory, since ninety per cent of the graduates have reasonable jobs in different firms, public and private institutions and research institutes.

It was felt, however, that it was necessary for students to have a stronger background in basic sciences and a new five year programme was created leading to the degree of Licentiate in Mathematics with major in Statistics and to a new professional degree of Statistician. The new programme includes a one semester consulting course in statistics and a one semester professional practice course.

UNIVERSITY	CITY	ACADEMIC OR PROFESSIONAL DEGREE	LENGTH
Iquique	Iquique	Statistician	4
Valparaiso	Valparaiso	Licentiate in Statistics	5
Catolica de Valparaiso	Valparaiso	Licentiate in Statistics	4
		Magister in Statistics	2
Catolica de Chile	Santiago	Licentiate in Mathematics with Major in Statistics	5
		Doctor in Mathematics with Major in Statistics	4
Chile	Santiago	Mathematical Engineer with Major in Statistics	6
Concepcion	Concepcion	Magister in Statistics	2*
Austral de Chile	Valdivia	Statistician	4

* Graduate programme

EXHIBIT 1

The curricula of the programmes mentioned are presented in Appendix 1. Some of the professional practices entered in different institutions are presented in Appendix 2.

EXPERIENCE WITH USERS

There are two basic classes of problems that typically arise in contacts with users of statistics.

a) Lack of a suitable common language to establish good communication.

b) Application of statistical methodology with disregard of the basic assumptions.

The first problem is, doubtless, mainly responsible for the break between statisticians and users. Although the second problem is solvable, if it becomes systematic it brings discredit to the statistical profession. One should point out that these problems are foreseeable. As T.A. Bancroft says "the present advance and complexity of sciences makes it difficult for the specialist to keep up to date in his own field. For this reason it is virtually impossible for him to be aware of the advances in statistical methodology".

SOME SOLUTIONS TO CLOSE THE GAP BETWEEN STATISTICIAN AND THE STATISTICS USER

To close the gap between statisticians and statistics users we make the following suggestions:

a) To offer graduate programmes directed at professionals from different disciplines.

b) To offer short courses on statistical topics of general interest, or on recent advances in statistical methodology.

c) To give a statistical consulting service to the rest of the university and outside users. This must be connected to a statistical consulting course for students in statistics.

d) To give students in statistics in-depth experience of the language and concepts of a field of application. After learning this first language, it is felt that the students would, easily learn the language of a second field of application.

e) To strengthen the role of National Statistical Associations.

f) To keep in close contact with graduates of statistical programmes. Their experience should be of great use in updating the curriculum and at the same time they would benefit by learning recent advances in statistics.

g) The statistics group in the university should offer service courses to all disciplines.

h) To illustrate methodological courses with real life examples brought from different fields of application.

i) To keep close contacts with information collecting organizations, especially with National Statistical Institutes. One way to do this is through the periodic realization of the professional practices of statistics students.

A P P E N D I X I

CURRICULUM OF THE LICENTIATE IN MATHEMATICS WITH MAJOR IN STATISTICS AT THE CATHOLIC UNIVERSITY OF CHILE

SEMESTER I	Number of Credits
Algebra and Introduction to Calculus	12
Geometry I	12
Introduction to Computer Sciences I	10
Introduction to Probability and Statistics	10
SEMESTER II	
Calculus I	12
Geometry II	12
Computer Sciences II	15
General Studies	5
SEMESTER III	
Calculus II	8
Linear Algebra I	10
Introduction to Foundations of Mathematics	10
Statistical Methods	14
General Studies	5
SEMESTER IV	
Calculus III	12
Differential Equations	8
Probability I	10
Linear Algebra	10
Introduction to Numerical Algorithms	5
SEMESTER V	
Deal Analysis	16
Probability II	16
Regression Analysis	16
SEMESTER VI	
Statistical Inference	16
Analysis of Variance	16
Sampling Theory	8
Numerical and Optimization Methods	10
SEMESTER VII	
Stochastic Processes	14
Introduction to Multivariate Analysis	12
General Studies	5
Optional Courses	20

SEMESTER VIII	Number of Credits
Consulting Laboratory	15
General Studies	5
Optional Courses	30
SEMESTER IX	
Professional Practice	20
General Studies	5
Optional Courses	20
TOTAL	424

LIST OF OPTIONAL COURSES IN STATISTICS

1.	General Methods	
	Time Series Analysis	12
	Discrete Variable Analysis	12
	Non-parametric Methods	12
	Design of Experiments I	12
	Design of Experiments II	12
	Data Analysis	10
	Topics in Statistical Methods	10
2.	Statistical Theory	
	Foundations of Statistical Inference	10
	Bayesian Inference and Decision Theory	10
	Statistics Seminar	10
3.	Specialized Statistical Methods	
	Statistical Methods in Econometrics	12
	Statistical Quality Control	10
	Introduction to Biostatistics	10
	Topics in Biostatistics	10
4.	Probability theory and Stochastic Processes	12
	Stochastic Processes II	
	Topics in Probability Theory	
5.	Applied Probability Models	
	Reliability theory	10
	Queuing Theory	10
	Inventory analysis	10
	Stochastic Processes Applied to Biology and Health	10
6.	Statistical Computing	
	Simulation	10
	Computational Methods in Statistics	10
	Workshop on Statistical Software	10

CURRICULUM OF THE DEGREE OF STATISTICIAN AT THE CATHOLIC UNIVERSITY OF VALPARAISO

SEMESTER I	Number of Credits
Algebra	6
Calculus I	6
Elements of Statistics I	6
Physical Education	1
Religious Culture I	2
SEMESTER II	
Linear Algebra	4
Calculus II A	4
Elements of Statistics II	8
Physical Education II	1
Religious Culture II	2
SEMESTER III	
Complements of Linear Algebra	4
Calculus II B	4
Probability Theory	6
Survey Design	3
Physical Education III	1
Religious Culture III	2
SEMESTER IV	
Elements of Numerical Analysis	4
Complements of Calculus	6
Statistical Inference	6
General Studies	2
SEMESTER V	
Operations Research	6
Sampling Theory	6
Optional Course	4
General Studies	2
SEMESTER VI	
Experimental Design	6
Optional Course	4
Optional Course	4
Optional Course	4
SEMESTER VII	
Statistical Workshop I	6
Regression Analysis	6
Optional Course	4
SEMESTER VIII	
Statistical Workshop II	6
Statistics Seminar	6
General Studies	3
TOTAL	145

APPENDIX II

REPORTS ON THE PROFESSIONAL PRACTICES MADE BY STUDENTS OF STATISTICS OF THE CATHOLIC UNIVERSITY OF CHILE

In the following list are registered the name of the professional practices entered by the students and the place where the work was undertaken.

	NAME	PLACE
1.	Fit of a nonlinear regression model to an agronomic problem.	School of Agriculture, U.C.
2.	An application of statistics to forestry research.	National Forestry Research
3.	Statistical analysis of private undergraduate educations programmes.	National Statistical Institute
4.	Statistical analysis of movie preferences in Santiago.	Compañía distribuidora de películas (CONATE)
5.	Statistical applications data processing.	National Electricity Company
6.	Marketing research on powder milk.	CRAVAL
7.	Study of the academic efficiency of Psychology students of the Universidad Catolica.	School of Psychology of the U.C.
8.	Statistical use of a data bank.	Las Viscachas Company.
9.	An application of statistics to agronomy research.	National Institute of Agronomy Research
10.	An application of statistics to quality control.	CTI
11.	Estimation of the mean number of failures in electronic instruments.	Subway of Santiago
12.	Market research in fabrics.	Yarur S.A.
13.	Efficiency of looms in the Yarur company.	Yarur S.A.
14.	Distribution of compensations and other guarantees in bakery workers.	Departamento de Indemnizacíon de obreros molineros y panificadores
15.	Evaluation of a quality control system.	Phillips Chilena
16.	Estimation of the long distance call income.	Chilean Telephone Company
17.	Level of fertility research in Chile.	Latin American Center of Demography
18.	Inventory control.	Chilean Telephone Company
19.	Application of PERT analysis to geological problems.	Geology Institute

20.	Applications of different statistical methods in agronomy research.	National Institute of Agronomy Research
21.	Research on the trend of mortality by sex and age.	Latin American Center of Demography
22.	Sampling and checking an Inventory in the Andina Mining Company.	Andina Mining Company
23.	Time series analysis of the demand for mushrooms.	Pro-Chile
24.	Forecasts of illiteracy (1975-2000).	National Statistical Institute
25.	Error analysis in determining the laws of copper and molibdenum.	Research Center for Mining and Metallurgy
26.	Forecast of automobile demand between 1980 and 1989.	Banco Hipotecario de fomento
27.	Quality control of raw materials in a pharmaceutical industry	Schering Laboratory
28.	Portfolio analysis	U.C. Computer Center
29.	Analysis of lattice designs.	National Seeds Company
30.	Market research in housing.	Colocadora National de Valores
31.	Forecasting textile imports.	Textile Institute of Chile
32.	Analysis of automobile accidents	BHC Insurance Co.
33.	analysis of malnutrition in Chile.	CONPIN
34.	Study of the quality control of services of 3M-chile.	3M-Chile
35.	Forecasting models for food imports.	SOFOFA
36.	Study on the consumption of fertilizers and pesticides in Chile.	computing Promotional Co.

CO-OPERATION BETWEEN ACADEMIC AND PRACTISING STATISTICIANS

A Prat Bartés, AVDA, Barcelona

The main purpose of this contribution is to illustrate some of the benefits that co-operation between our Department of Statistics at the Polytechnic University of Barcelona and the public administration and/or private industrial sector in Barcelona, Spain, has produced.

As George Box (1980) pointed out, some discoveries of the stature of least squares, ratio estimators, design of experiments and others, have been the result of the interaction with reality. Our Department is firmly convinced that co-operation with the industrial environment in real problem solving is very important in order to stimulate theoretical research.

Our involvement with real problem solving has been possible in part due to the engineering background of most of the personnel in our Department. Knowledge of the physics of many industrial systems and familiarity with computers have proved very useful tools when used in connection with statistical theory and methods.

Some examples of how reality has stimulated theoretical research in our Department are the following:

1. During our co-operation with Electricity Utilities of the Barcelona area in a research project about load curves, a problem of non-linearity in the relationship between temperature and daily electricity consumption arose. This has stimulated a Ph.D. research project on the topic of non-linearity in transfer function models.

2. When co-operating with a banking organization in Barcelona with some 1,000 branches we found that it would be almost impossible to forecast the deposits of each of the branches using the Box-Jenkins approach. The need for a reliable automatic forecasting system arose.

One of our students, Moisés Valls, has almost completed his Ph.D. work which gives a solution to the problem of automatic modelling by defining a statistical distance between theoretical autocorrelation functions and inverse autocorrelation function, and the corresponding sample functions. The method seems to work very well with real data.

There are many other benefits arising from this co-operation, especially in getting good material for teaching. With the permission of the Electricity Utilities we use a very important data-base obtained during the research project as real data for teaching purposes. At the beginning of the semester, each student

gets a portion of the data-base and all students are trained in using a statistical package (SPSS in our case). Most of the general statistical course is based on this real data set. A great number of statistical methods can be used here: checking the quality of the data, exploratory data analysis, linear models, mechanistic and empirical model building, time series, influential observations, etc. Using real data increases the motivation of our students to learn the theory necessary to understand the methods required to analyse those data.

Furthermore, as Ron Snee pointed out in his talk here at ICOTS, through this co-operation with local industry we have been able to use companies' big computers that were not available to us in our University at this time.

Another area of co-operation with local industry has been the training of industrial statisticians in the most recent developments of our subject matter, through short courses designed specifically for this purpose. Some of these short courses have been given "in house" to the industrial practitioners, and others have been lectured in our Department.

Those one-week courses have covered subjects such as linear models, time series analysis and survey sampling.

When dealing with small and medium size industries our efforts have not been very successful yet. Our Department offers short courses in Statistics through the CPI (Center for Continuous Education of the Catalan Engineering Association) that are attended by some of the engineers working in these types of industries. We are developing statistical software tailored to the needs of these industries and specifically designed for minicomputers. Our feeling is that much effort should be developed in this area.

As a way of summarizing this talk we would say that the interaction between reality and theory in our Department coming from our co-operation with local industries and public administration, has been extremely fruitful from the point of view of both academic and industrial institutions. It has provided real problems to be solved through our research programme and excellent material for teaching.

The industries, on the other hand, have been introduced to the most recent methods of statistics and in some cases have experienced great reductions in production costs through a sound use of data analysis.

Finally, we feel that more effort should be deployed in the co-operation with small and medium size industries.

Reference

[1] Box, G.E.P. (1980) Comment on the paper: "Preparing Statisticians for Careers in Industry: Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Industry". American Statistician, 34, 2, 76.

TOPIC 12: TEACHING SURVEY SAMPLING

ORGANISER: DR D SINGH

THE TEACHING OF SAMPLE DESIGN TO GOVERNMENT STATISTICIANS FROM DEVELOPING COUNTRIES

Joseph K Garrett, U S Bureau of the Census, Washington, D C

INTRODUCTION

The International Statistical Programs Center (ISPC) of the U.S. Bureau of the Census has conducted training programs for government statisticians from developing countries for more than 30 years. Through these programs, ISPC has contributed substantially to statistical development in many countries. The programs are conducted under the auspices of the Agency for International Development and serve urgent and changing needs of developing countries for trained personnel to collect, process, and analyze statistical data. Currently, 1-year training programs are offered in each of the following specializations:

(1) Sampling and Statistical Methods
(2) Computer Data Systems
(3) Survey Methods
(4) Agricultural Statistics
(5) Demographic Statistics
(6) Economic Statistics

These six programs are directed toward training statisticians in the specific skills needed to carry out many different phases of ongoing statistical activities involved primarily in conducting socioeconomic household surveys, agricultural surveys, and retail establishment surveys. Although participants specialize in one program, they are introduced to clearly related and auxiliary fields and skills to give them an understanding of how their specialty is linked with the total statistical system.

This paper discusses the education and training given to participants specializing in Sampling and Statistical Methods at the ISPC. This education and training in survey sampling is conducted through linkage of sampling concepts with the various other survey-related skills and fields. This means, among other things, that it is not sufficient to merely offer a course on sampling theory when the objective is to train participants to be functioning sampling practitioners and technicians. While this paper primarily discusses the training in sampling conducted at the ISPC for statisticians from developing countries, it is nevertheless felt that most of the remarks are relevant to the teaching of survey sampling in general.

THE CONCEPT OF LINKAGE

It would seem self-evident to assert that when teaching any subject, or when training in any subject-matter area, that subject needs to be linked to related disciplines and auxiliary fields. While this appears obvious, the concept of linkage is often absent in traditional courses on survey sampling. More often than not, the sampling course becomes a course in algebra. Algebraic results showing an estimator to be unbiased or of smaller variance than that of a competing estimator seem to be the dominating feature of many class lectures and homework problems (Murthy, 1980). While such results are indeed important and should be included in sampling courses, too much emphasis on algebra at the expense of linkage tends to create "conditional samplers". That is, students of such courses may become adept at solving problems such as: "Given the simple random sample $x_1, x_2, \ldots, x_n$, show that $E(\bar{x})=\bar{X}$ and that $V(\bar{x})=(1-f)\frac{S^2}{n}$ ". While developing a student's ability in this type of algebraic endeavor is a necessary exercise in order to build his or her understanding of the mathematical-statistical framework upon which population sampling is grounded, its overemphasis would tend to ignore the fundamental problem of how a simple random sample of n observations is actually obtained. To the extent that this conditional approach dominates the more complicated (and most often used) topics of stratified, clustered, and multi-stage sampling, the student is not suitably prepared to deal with the applied issues of data collection that are involved in survey sampling. To such students, the concept of a primary sampling unit may be nothing more than an algebraic entity.

In teaching survey sampling, one must not only develop the mathematics of survey sampling, but also provide for the integration of sampling topics with the various other important survey-related topics. Development of a survey sampling statistician requires linkage to be developed in the context of the overall objectives of the training program.

THE DEVELOPMENT OF LINKAGE

In order to develop the concept of linkage in the training of survey sampling to statisticians from developing countries, two broad tasks need to be accomplished. First, fields and skills clearly related and auxiliary to survey sampling need to be identified. These include, among others, design of tables and questionnaires, editing and coding principles, methods of imputation, data processing, survey budgeting and scheduling, survey management, control and evaluation of nonsampling error, geography and mapping, statistical quality control, enumerator recruitment and training, and presentation and publication of data. Second, an assessment of the general survey conditions pertinent to survey sampling in developing countries needs to be done. These two tasks

are not entirely mutually exclusive. Rather, in assessing the general survey conditions, some needed skills clearly related to survey sampling in developing countries are identified. Moreover, such assessment provides the basis for emphasis that certain topics receive during the training program. While it is not the purpose of this paper to fully characterize the general survey conditions of the developing countries of the world, it is possible to list a few that are inherent in many developing countries:

- Available maps are often inadequate or outdated.
- Transportation in many parts of the country is difficult; good roads are few and far between.
- Complete frames which can be used for sampling arenot available; frame construction is an essential.
- Independent sources of data are skimpy and often unreliable.
- Nonsampling error is often the dominant source of error in survey estimates.
- Survey budgets are quite limited and managed, at times, in a haphazard manner.
- Overall quality control is poor and often not regarded as important.
- The "sampler" does not havedesigning samples as his/her only job. He/she must not only have specific sampling knowledge, but also general survey knowledge.
- Subject-matter concepts are often poorly or inconsistently defined.
- Training of enumerators is generally not of sufficient quality.
- Communication among various survey personnel can be improved.
- Housing stock is generally poor and living conditions are quite crowded.

Using this fairly broad assessment of general survey conditions in developing countries in conjunction with identification of clearly related and auxiliary fields and skills, the concept of linkage can be developed. Not only is there the pedagogical need to link the concepts of sample design with other important survey-related activities in the actual teaching process, but participants should also be introducd to the various other survey-related activities so as to provide a meaningful setting for the linkage. For example, in order to link the sampling concept of "determination of sample size" with the survey concept of "output tables", participants need to have a conceptual grasp on what "output tables" are.

At ISPC, the development of linkage in the training of survey sampling to government statisticians from developing countries is accomplished on two levels. In the actual teaching of sampling concepts, frequent reference is made to the other important survey-related activities. In addition, to cement the concept

of linkage, education and training in these other important survey-related activities is also conducted concurrently. This is all done through classroom lecture and discussion, "learning by doing" laboratories, and case study workshops. In addition, a large-scale demonstration survey is planned, designed, managed, and ultimately conducted in the field by the participants.

While emphasis in this paper is on the application of the concept of linkage of the sampling concepts to the various other survey-related concepts, a description of a typical training plan for ISPC sampling participants is given in Appendix A.

APPLICATION OF LINKAGE IN THE TEACHING OF SAMPLE DESIGN

To illustrate linkage in the teaching of sample design to government statisticians from developing countries, several examples will be given. These examples are primarily concerned with socioeconomic household surveys and agricultural surveys conducted in developing countries and consequently emphasize general survey conditions present in many of these countries. However, several of the remarks are relevant for surveys conducted in developed countries as well. On a more general level, the application of linkage of sampling concepts to related survey topics is a principle applicable in varying degrees to any course on survey sampling. The examples that follow focus on important and basic topics such as survey error, determination of sample size, stratification, clustering, multi-stage sampling, area sampling, point estimation, and variance estimation. While these do not exhaust survey sampling topics, they suffice to illustrate the approach by which sampling topics can be integrated with and linked to other survey-related activities in sample design courses.

Survey Error: Sampling error is, quite obviously, an extremely important concept in survey sampling. Controlling sampling error to a tolerable level is a required design feature. In many developing countries, however, nonsampling error is often the dominant source of error in survey estimates. Thus, in teaching sampling to developing country statisticians, it is better to introduce from the beginning a broad framework for sample surveys which includes both sampling and nonsampling errors (Hansen, Hurwitz and Bershad, 1961; Murthy, 1980). This introduces the notion of "total error" of a survey estimate. For discussing nonsampling error, this author has used a framework which classifies sources of nonsampling error by the respective survey stage that the source initially appears (Garrett, 1979). Thus, sources of nonsampling error are found in the preparatory stages of surveys, the data collection stage, the data processing stage, and the publication stage. This framework serves two purposes. It provides a method for introducing a concept which by its very nature is diverse and pervasive. Moreover, it serves as an approach to discuss

controlling nonsampling error throughout the survey cycle. Preparatory stage errors can occur due to problems of specifying survey goals and objectives, improper planning and scheduling of survey operations, use or development of inadequate maps, use or development of faulty frames, and inadequate recruitment and training of enumerators. Errors in the data collection stage include errors of coverage, content, and nonresponse. Errors in the data processing stage include errors of editing, coding, imputation, key punching, computer programming, tabulation, and calculation. Publication stage errors include inconsistent tables, statistically inaccurate statements, and misleading narrative.

Insofar as the first step toward controlling nonsampling error is identification of its sources, the above outline indicates many of the more common sources of nonsampling error in surveys planned and conducted in developing countries. The concept of controlling sampling error needs to be linked to the concept of controlling nonsampling error so as to establish control of total error. Thus, the concept of nonsampling error should be developed concurrently with that of sampling error and discussions related to the control of nonsampling error should point out relevant survey procedures, methods, and activities that provide such control. These include necessary improvements in mapping, questionnaire and table design, planning of field operations, and overall survey quality control.

Determination of Sample Size: The sampling problem of determining an appropriate sample size should be linked with the survey concept of output tables. Reviewing intended survey output tables gives both the survey sponsor and the sampling technician a feeling for the magnitude of the proposed survey (Ford and Tortora, 1978). Throughout such review, attention can be directed toward the desired reliability requirements for the various major tables and cells within. Additionally, review of output tables helps in determining possible stratification criteria such as "urban" and "rural". As estimates of sample sizes for major characteristics of various table grand estimates, marginal estimates, and cellular estimates are made, it may be found that certain cells or tables need to be collapsed. Also, it may be found that the survey is overloaded with major characteristics which are too diverse. In many developing countries, survey budgets are quite limited and management of such budgets can be improved if proper attention is given to careful review of output tables in connection with determining sample size. Available data for estimating variances required for estimating sample sizes is skimpy or of poor quality in many developing countries and crude estimates often need to be resorted to. Rough estimates of variance can be calculated using information on the approximate range and shape of the distribution of the major survey characteristics (Deming, 1960) and these approximation methods need to be stressed in class.

Stratification: The topic of stratification offers many opportunities to introduce clearly related survey activities. While many texts emphasize size stratification, geographic stratification is often used for multipurpose surveys in developing countries. Topographic maps are often used in conjunction with subject-matter knowledge in developing stratification for, say, an agricultural survey. So-called "land use strata" are delineated on maps when defining strata to be used with area sampling. Satellite imagery, which is available for most countries, is often useful for stratification. For household surveys, geographic stratification makes sense if major household survey characteristics are similar within broad socioeconomic neighborhoods or regions. In the teaching of the principles of geographic stratification, laboratory exercises, case-study workshops, and demonstration surveys are used at ISPC to allow sampling participants the opportunity for "hands-on" experience in developing strata. These exercises emphasize that when delineating stratum boundaries, use of natural physical features which can be easily identified in the field is a primary consideration. In many developing countries, lack of quality data generally prohibits an extremely rigorous quantitative search for stratifying variables and, hence, geographic stratification is often a reasonable alternative. Thus, the topic of stratification can be linked with availability of quality maps, aerial photographs, satellite imagery, and the existence of adequate census data.

Clustering, Multi-Stage Sampling, Area Sampling: The sampling topics of clustering, multi-stage sampling, and area sampling are topics which, quite obviously, involve mathematics. In conjunction with stratification, these topics are also very applied notions which have definite relations to other survey activities. These relationships deserve proper recognition during the teaching of survey sampling. For socioeconomic household surveys or agricultural surveys in developing as well as developed countries, the general sample design is often a stratified, clustered design involving more than one stage of sample selection. Often, the first-stage sampling units are geographic areas. While the necessary algebra for this general design needs to be developed, it is also the situation that these topics need to be linked with a variety of other survey activities. The topic of clustering basically involves the choice of a suitable sampling unit. It should be stressed that existing administrative units should be reviewed to see if any would also be appropriate sampling units. For national socioeconomic household surveys this review may provide primary sampling units which are suitable from not only a statisical viewpoint, but also from a cost-benefit approach. For agricultural surveys, nonadministraive geographic areas delineated on topographic maps (if such are available) within land use strata may represent reasonable primary units. Thus, complementary to the mathematics, an applied, spatial notion of "primary sampling unit" needs to be developed. This can often be done in teaching by relating the concept of "primary sampling unit"

with existing administrative units and available maps.

Generally, subsampling within primary units is necessary due to budgetary and staffing constraints. This is especially true in developing countries. During teaching, the algebra of subsampling needs to be balanced with the details of this process in the field. The survey topics of map preparation for selected primary units, segmentation of primary units, field listing and screening, and enumerator recruitment and training should be linked with the sampling topic of multi-stage sampling. Additionally, emphasis should be given to the control of nonsampling error during these field operations. If second (and further) stage sampling units are geographic areas, the use of physical boundaries in the mapping process defining such units needs to be emphasized.

Area sampling in which completely enumerated second-stage sampling units are geographic areas within selected first-stage area units is a commonly used design for agricultural surveys. The methodology involved during data collection with this design has impact on questionnaire wording and is a function of the characteristics being measured. For example, crop information is generally gathered for the entirety of land area within sample segments operated by farmers living either in or out of the sample segment. However, information on, say, cattle is collected only from farm operators residing within the sample segments. For multi-purpose agricultural surveys of this design, the sample design is closely related to design of the questionnaire. Failure to observe this relationship may result in estimation problems once the data are collected. Thus, relationship between sample design and questionnaire design needs to be observed when conducting sampling courses.

The topics of clustering, multi-stage sampling, and area sampling should have the requisite mathematics developed while concurrently developing the linkage of these topics to geography and mapping activities, review of available administrative areas which might serve as statistical units, details of various field operations, enumerator recruitment and training, questionnaire design, and survey budgeting.

Point Estimation, Variance Estimation: The survey activities that occur after data are collected need to be emphasized. It is not a good portrayal of survey reality to discuss estimation of various population parameters abstracted from such important survey activities as editing, coding, keypunching, adjustments for non-response, and other data "cleaning" and data preparation activities. To discuss point estimation as simply a parochial and clinical algebraic process does not suffice. The sampling technician in developing countries is involved, directly or indirectly, in all of these activities. During the teaching process, frequent reference should be made to the importance of these activities. The

usefulness of the estimating formulae in sampling can be linked directly to the quality with which these data processing activities are completed.

After data preparation activities are completed, the sample data needs to be "weighted" so as to obtain population estimates. In survey practice, full communication between the sampling technician and the data processing specialist needs to be realized. Communication among survey personnel can be a problem in developed as well as developing countries. It is the sampling technician's responsibility to provide lucid specifications precisely describing how this weighting should be done - including all necessary adjustments to the basic weight derived from the selection probabilities. For large-scale national surveys, the sampling concept of "self-weighting design" can be readily linked to this possible communication problem. In the sampling course, development of weighting specifications needs proper consideration and involves much more than simply describing the algebraic form of the various estimators.

For large-scale, multi-purpose surveys involving an extraordinarily large publication stub, variance estimation is a complex undertaking. For national-level surveys involving numerous characteristics, adjustments of various kinds for nonresponse and undercoverage, estimation features such as ratio adjustments and large observation procedures, and design features such as stratification, multi-stage sampling, and unequal selection probabilities, the variance estimation approach needs to consider not only accuracy, but also timeliness, cost, simplicity, and other administrative considerations (Wolter, 1979). In developing countries constrained by computing facilities, survey budgets, and survey operations staff, these considerations become very important. In addition to customary (textbook) variance formulae, other variance estimation methods used in practice such as random group procedures, balanced half-sample techniques, and methods for systematic sampling need to be introduced. These ideas can also be related to report writing and the need for statistical standards in describing the findings of the survey.

REINFORCEMENT OF LINKAGE IN THE TRAINING OF SAMPLING

While constant reference to related survey topics, skills, and fields should be made during the teaching of the sampling course, it is also essential to allow students the opportunity to actually design a sample. At ISPC, the participants specializing in Sampling and Statistical Methods design either a socioeconomic household survey or an agricultural survey, depending on their interests. For both surveys, the target population is an entire county in the U.S. Working with participants specializing in other areas, these surveys are conceived, planned and ultimately conducted in the field. Beginning with definition of survey goals and objectives, proceeding through the survey cycle, and ending with

"publication" of the survey results, the integrated nature of survey activities becomes apparent. In particular, the sampling participants experience for themselves that survey sampling is more than an algebraic endeavor. While the surveys are viewed as training or demonstration surveys, many real survey features are present. The socioeconomic household survey generally has a sample size of approximately 500 households selected through a stratified (urban/rural), clustered (existing administrative or statistical units), multi-stage (area segments as second-stage, listed households as third-stage) design. The agricultural survey generally utilizes a sample of area segments selected within land-use strata (intensive agriculture, extensive agriculture, forests, mountains, and urbanized) through a two-stage process. During the data collection phase of the survey, participants become enumerators and face first-hand some of the real problems involved in implementing the details of field operations. The experience emphasizes the need for a quality enumerator training program for proper data collection activities. Throughout the six-month period of involvement in the survey, participants not only perceive the various connections of sampling to other survey features, they witness the effects of ignoring relevant linkage.

The purpose of these demonstration surveys is two-fold: (1) To allow participants the opportunity to develop a technically sound sample design, complete with documented details on stratification, clustering, sample selection, and estimation; and, (2) to show the relevant and necessary linkage that binds sample design to overall survey design. For many, it may be the first time they have viewed a sample survey as an integrated process. This experience translates into better, more well-coordinated surveys when they return home.

CONCLUDING REMARKS

Sample design is basically an applied field. While probability theory and algebra are necessary ingredients in the application of sampling principles, they do not represent the entirety of sampling. In the teaching of sample design, whether it be to government statisticians from developing countries or any other group, the basic teaching problem is one of balance - balancing necessary mathematics with necessary application. Real application of sampling principles to subject-matter problems involves development, application, and maintenance of linkage of sampling concepts to other important survey skills, activities, and fields. Consequently, incorporating linkage more fully in the teaching process is necessary. It seems, as stated earlier, that this idea is obvious and self-evident. Why, then, do many sampling courses become unduly balanced by excessive algebra? I can think of two possible reasons: (1) Algebra is the most important part of sample design and should be the dominating feature, (2) experience in algebra is more fully developed in the instructor than is experience in actual application of sampling principles. If a sample design were developed to reliably estimate

the proportion of instructors teaching sampling who possess the characteristic "experience in sample design application", the resulting estimate would probably be very small. Insofar as this may be true, the second reason given above may very well imply the first reason.

References

[1] Cochran, W.G. (1977) Sampling Techniques. John Wiley & Sons, New York.

[2] Deming, W.E. (1960) Sample Design in Business Research. John Wiley & Sons, New York.

[3] Des Raj (1968) Sampling Theory. McGraw Hill Book Co., New York.

[4] Deighton, R.E., Poland, Jr., et al. (1978) Statistical Policy Working Paper 4, Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistics. U.S. Department of Commerce, Washington, D.C.

[5] Ford, B.L. and Tortora, R.D. (1978) "A Consulting Aid to Sample Design". Biometrics, 34, 299-304.

[6] Garrett, J.K. (1979) "Assessing the Reliability of Data: Nonsampling Error". Paper presented at 1979 Symposium on the Analysis of Labor Statistics: Washington, D.C.

[7] Gonzalez, M., Ogus, J.L., Shapiro, G., and Tepping, B.J. (1975) "Standards for Discussion and Presentation of Errors in Survey and Census Data". Journal of the American Statistical Association, 70, No.351, Part II.

[8] Hansen, M.H., Hurwitz, W.N. and Bershad, M.A. (1961) "Measurement of Error in Census and Survey". Bulletin of International Statistical Institute, 38, (2), 359-374.

[9] Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953) Sample Survey Methods and Theory, Volumes I and II. John Wiley & Sons, New York.

[10] Kish, L. (1965) Survey Sampling. John Wiley & Sons, New York.

[11] Murthy, M.N. (1980) "Some Thoughts on Teaching of Sampling". Paper presented at 1980 Symposium on Teaching Statistics, Goteburg, Sweden.

[12] Wolter, K. (1979) Introduction to Variance Estimation (draft copy of text to be published).

[13] Zarkovich, S.S. (1966) Quality of Statistical Data, Food and Agricultural Organization, Rome.

APPENDIX

COURSES AND WORKSHOPS FOR PARTICIPANTS SPECIALIZING IN SAMPLING AND STATISTICAL METHODS

		Total Class Sessions
First Training Period		
104-1	Design of Sample Surveys	15
105-1	Survey Sampling Laboratory	5
600	Introduction to Design of Surveys and Censuses	15
Second Training Period		
104-2	Design of Sample Surveys - continued	30
105-2	Survey Sampling Laboratory - continued	10
601	Introduction to Data Processing for Subject-Matter Specialists	15
602	Design of Tables and Questionnaires	30
604	Budgeting, Scheduling and Controlling of Survey-Census Operations	15
	Optional:	
203-1	Agricultural Survey Operations and Methods	10
405	Applications of Economic Survey-Census Techniques	10
Third Training Period		
103	Applied Regression and Correlation Analysis	20
106	Variance Estimation in Sample Surveys	20
107	Laboratory in Variance Estimation Techniques	10
111	Introduction to the Analysis of Data	20
120-1	Field Demonstration Laboratory in Sampling and Statistical Methods	10
150-1	Seminar: Selected Topics in Sampling and Statistical Methods	10
605	Editing, Coding and Imputation Principles	30
609	Workshop: Training for Statistical Activities	1 week
695	Special Programs in Statistical Technology	10
	Optional:	
203-2	Agricultural Survey Operations and Methods - continued	10
204	Agricultural Sample Survey Design	20
205	Frame Construction for Agricultural Samples	20
430	Economic Enterprise Survey: Field Demonstration Laboratory	10
603-1	Geography and Mapping for Surveys and Censuses	20
Fourth Training Period		
108	Control and Evaluation of Nonsampling Error	20
109	Quality Control in Censuses and Surveys	10
110	Applications of Current Survey Design and Evaluation Techniques	20
120-2	Field Demonstration Laboratory in Sampling and Statistical Methods - continued	10
150-2	Seminar: Selected Topics in Sampling and Statistical Methods - continued	10
606	Management of Statistical Activities	20
608	Presentation, Publication and Dissemination of Data	10
	Optional:	
206	Objective Measurement of Area and Yield	30
308	Alternative Strategies for the Collection and Analysis of Demographic Data	10
603-2	Geography and Mapping for Surveys and Censuses - continued	20
Fifth Training Period		
190	Case Study in Sampling: Workshop and Field Demonstration	Full Time

TEACHING OF SURVEY SAMPLING

D Singh, UNDP, Amman

INTRODUCTION

Survey sampling is one of the most commonly applied disciplines of statistical science. The subject is of such a common use that even a layman without any formal training in the subject may sometime depend upon its application for solving his problems and arriving at the right decision. The decision of a consumer to purchase a bag of grain by examining only a handful of grain is one of the many examples of the application of the sampling technique by a layman. Because of the apparent simplicity of the technique, the subject of survey sampling did not receive the respectability by the mathematicians that it deserved. It was during and after the Second World War when the demand for various kinds of statistics in diverse fields of human activities for taking quick policy decisions increased, the attention of theoretical statisticians engaged in war efforts or later on in socio-economic developmental programmes was drawn to the problem of collection of data by following sampling techniques.

Immediately after the Second World War, the demand for socio-economic data became so urgent that the United Nations Sub-Committee on Statistical Sampling was set up in 1974 to mainly review the scope of sampling technique in its application to censuses and surveys. As a direct consequence of the report of this Sub-Commission, the United Nations and its specialised agencies took necessary steps for promoting the use of sampling methodology, and considerable success, particularly in developing countries has been achieved over the last three decades. In several industrialised countries, particularly in UK and USA, which already had some experience in applying sampling technique even prior to the Second World War, systematic research both in theoretical aspect as well as in its application was undertaken during the war period, and it was introduced as a part of statistics courses in a number of universities. In the Indian sub-continent, the research on the application of sampling methods for estimation of crop productions was initiated by the Indian Statistical Institute and Indian Council of Agricultural Research more than four decades ago, and later on, the method so developed was adopted by the Indian Government as a routine procedure for crop estimation.

Many developing countries in Asia, Africa and Latin America which became free from colonial rule after the Second World War started socio-economic development to fulfil the aspirations of their people. One of the major problems

that these countries faced was the lack of basic information needed for the purpose. With the aid of some international agencies or friendly countries, many of them planned censuses and surveys depending upon the specific needs, and the monetary resources and manpower that they could mobilise. It was at this stage that these countries felt acute shortage of trained manpower at different levels to continue the census and survey activities on the regular basis so urgently needed. To fill up the gap of manpower at least at the intermediate level, several regional training institutes, each meeting the requirements of a group of countries of the region were established. Although these institutes impart courses in general statistics as applied to official work of national governments, an elementary course on sampling methods is an important component of the whole programme. The quality of these courses, however, depends on the quality of the teachers' and trainees' background. Many of these institutes have already functioned for more than a decade. It may be very appropriate at this stage that an independent international agency like ISI may evaluate as to how far these institutes have fulfilled the objectives for which they were established. Several national governments, besides encouraging the universities to impart courses on sampling methods as part of a degree course in statistics, established separate specialised training and research institutes. For example, in India, besides the Indian Statistical Institute, the Indian Agricultural Statistics Research Institute established immediately after the Second World War imparts advanced courses in sampling as applied to agriculture. A new Institute dealing with medical statistics recently established in New Delhi under the Indian Council for Medical Research, imparts courses on sampling as applied to medical problems. In USA, the Institute of Social Research and Statistics, University of Michigan organises special courses on sampling as applied to social sciences. The US Bureau of the Census, Washington D.C., known for its research in application of sampling techniques to census problems, organises special training programmes in sampling for the benefit of developing countries. Many such organizations came up all over the world.

With increased use of survey sampling technique, a number of books on sampling - Yates (1947, 1960), Cochran (1953, 1960), Sukhatme (1953, 1970), Hansen (1953), Murty (1967), Kish (1967), Des Raj (1968) etc. - were brought out as guides to practising statisticians and as reference literature to professional teachers. Fortunately, most of these books were written by practising statisticians and therefore, these reflect individual author's experiences. UN Statistical Office, New York, started separately monitoring information on current sample surveys conducted by the member countries, and making them available for wider use. Many national governments also bring out annual publications

on the current sample surveys. All these literatures are of considerable use particularly to practising sampling statisticians.

SCOPE OF THE COURSES

All the institutions engaged in teaching of survey sampling at present may be broadly classified into three categories (i) Universities (ii) Training and Research Institutes and (iii) National Statistical Offices.

In the universities, the subject of sampling is taught as a part of the course on mathematical statistics and emphasis is mostly laid on mathematics of sampling, and derivation of various formulas relating to estimates of population parameters and their variances. Even if the subject is taught as a part of applied statistics courses, the treatment is mostly mathematically oriented because of lack of application of the teachers in the practical application of the subject. The students may successfully complete the course without developing any confidence in its application to solve the practical problems of designing a sample survey. This is one of the reasons that in many countries a large number of statistical posts in government and public undertakings remain unfilled, in spite of the fact that many university graduates in statistics are unemployed.

The training and research institutes are supposed to give a balanced course, which may consist of underlying theory of sampling and its application. This balance can be obviously maintained only if the teachers themselves are engaged in research in the development of sampling techniques and their applications to solve the practical national problems. If the need of theoretical research arises, the teachers should have the capacity and interest in undertaking such research programmes. These are the institutions where teaching, research and extension programmes are supposed to be integrated.

The training courses organised at the national statistical offices are of purely applied type. In such courses, sampling methods are mechanically taught, describing the procedure of their applications. Even if some aspects of theory are discussed, use of algebra is avoided. Participants in such training programmes come mostly from government and public undertakings which adopt sampling technique as a tool for collection of data. Sometimes, these trainees learn the full details of all stages, from formulating the questionnaires and field instructions to preparing the survey report of a specific sampling project chosen from the area of the individual's interest.

CONTENTS OF THE COURSE

It is noticed that whatever may be the level of the students, the contents of the sampling course are almost the same. Most of the teachers try to cover the topics described in a standard textbook on sampling prescribed for the course. If the course is being delivered to the postgraduate students of mathematical statistics, it is assumed that they have already covered the relevant topics on statistical methods and probability, and know adequate algebra and calculus for following the course on sampling. Descriptions on the details of the individual topics depend upon the interest of the teacher and time allotted for it. In fact the subject is taught like any other topic of a mathematical statistics course, forgetting that the sampling is basically an applied subject and relevance of the course lies in inculcating interest among the students in learning it for solving the practical problems. The result is that unless such students are given further training on application aspects, they completely lose interest in the subject and in a short period forget it. If the objective of such a course is to produce teachers and research workers, the quality of teaching and relevance of research will again depend upon the way they look to the subject of sampling.

Because of the wide use of the technique, the courses on sampling are also these days imparted to students of other disciplines, such as economics, commerce, agriculture, medicine, engineering, public administration etc. Here the objective of the course is not to train the professional statisticians but to make the students familiar with sampling methods as a scientific tool for collection of data, which may be needed in their future careers. The number of students of this category is unusually large. The topics listed in the course are mainly the same as prescribed for the students of statistics, and teachers also employed for teaching this course have mostly training in mathematical statistics without any orientation in the application of the technique to specific fields. Consequently, their teaching style and way of communication are generally mathematically oriented. They introduce the topics through algebraic notations and develop the underlying concepts of sampling through hypothetical populations. Since these categories of students generally have a poor background in mathematics, they start resisting the subject and lose interest completely in learning it further. This is the greatest harm done to the subject. These are the persons who in future become decision makers and because of lack of appreciation of what the sampling technique can do in helping them to arrive at right and quick decisions, many useful sampling projects are abandoned on the pretext of lack of financial resources. The contents of the course for such categories of students should, therefore, lay emphasis on basic philosophy of the subject, and introduction of basic

concepts such as sampling and non-sampling errors, probability sampling, confidence intervals, unbiased estimate, accuracy etc. through live examples and numerical exercises.

To understand some of the concepts of sampling, training in basic statistical methods is necessary. For example, the concept of confidence interval cannot be clearly understood without proper knowledge of the properties of the normal distribution, and sampling distributions of relevant statistics. While discussing the sampling distribution of mean, the underlying idea of central limit theorem should be clearly understood, particularly for the reason that most of the survey populations are finite while the normal distribution concept is applicable to infinite and continuous population. Similarly, concepts of regression and correlation coefficients and their tests of significance should be considered as a prerequisite to learning the ratio and regression methods of estimation in sample surveys. These topics as a prerequisite to learning of sampling methods are normally taught as a part of the courses on statistical methods. However, to make the contents of the courses on sampling methods self-sufficient, the relevant topics relating to statistical methods may be considered as an integral part of the course on sampling, particularly for the students having either no training or little training in statistical methods. Most of the textbooks on sampling methods presuppose that the participants in the sampling course already have the necessary bckground in statistical methods and straightway start introducing the concept of sampling. It becomes the responsibility of the teacher to assess the background of the students and arrange the course and its level accordingly.

Sampling methods may be treated as technology for production of information on diverse fields of human activities. A professional sampling technologist must receive adequate training in all the relevant disciplines which, when used together judiciously, turn into the shape of sampling technology. As in the case of engineering, the basic disciplines being mathematics, physics and chemistry, likewise basic disciplines for sampling technology may be mathematics, theory of probability and statistical methods. In addition, a student specialising in sampling technique must acquire adequate knowledge of the area of application either through formal course work or through self-study. For example, a sampling statistician in agriculture must have adequate knowledge of the economy of agriculture and crop husbandry practices. The contents of the sampling course must, therefore, include an integrated course on the field of application. Preparation of a critical report on a sample survey project already completed may also be a part of the course contents. Such an exercise may help a student in knowing the various stages of a sample survey project.

SEQUENCE OF THE COURSE

As mentioned earlier, background in elementary algebra and calculus, and statistical methods may be considered a prerequisite to learning the subject of survey sampling. If a student has not got adequate training in these subjects, he must be advised to undergo a course containing these topics before the teacher starts the introduction of courses on sampling.

While introducing the subject of survey methodology, the first task of the teacher should be to generate interest among the students in learning it as a powerful tool for gathering information in the quickest possible time with least cost. The concept of a survey population, units of observations, character under study, parameters to be estimated etc. should be introduced with the help of live examples. The algebraic notations may be introduced only when the students have fully grasped the underlying concepts. This is the stage when merits and demerits of the census and sample survey technique as methods for collecting data should be thoroughly discussed. The communication may become more effective if the lecture is illustrated by an actual example of a census recently conducted. If the census has been conducted according to the procedure laid down and if all the arithmetical processes are correct, the census results should be free from any error. However, the same cannot be said about the results obtained from a sample in spite of the fact that all the prescribed rules have been strictly followed. This is because of the fact that the sample results are based on a fraction of the individuals constituting the population. Concept of sampling error becomes important here. The importance of sampling error as a measure of reliability of an estimate may be discussed now. Although the sampling distribution of error, unbiased estimate and confidence intervals are the topics to be generally considered as a part of the course on statistical methods, it may be useful, even at the cost of some time for repetition, to explain the underlying principles with the help of live numerical data, and electronic calculators for arithmetical calculations. The author's experience of teaching survey sampling for the last three decades to both categories of students, discipline of statistics and others, is that they understand the statistical concepts better when illustrated by actual data.

The concepts of standard error and sampling variance are of great importance in the entire theory of statistics, but these are of greater importance in developing the sample survey theory. In fact, the entire sample survey theory centers round the search for the smallest sampling variance under a given set of conditions. The students should get full insight into the utilities of standard error in presenting any sample estimate of a parameter, say mean or total. Once the students have got a good grasp of the concepts of standard

error and sampling variance, the concept of non-sampling errors and statistical biases, which are not commonly understood by the non-survey statisticians, may be introduced, emphasizing their effects on the interpretation of the census and survey results. The existence of large non-sampling errors may spoil even the results of a census, which are free from sampling errors. It is one of the strongest arguments in favour of sample surveys that it is possible to control or at least minimise the effect of non-sampling errors in a well designed and well conducted sample survey, which may not be possible in a census unless enormous resources are spent for the purpose. In general, therefore, the sample estimates are subjected to both sampling and non-sampling errors. Within the framework, the efforts needed to control and balance the two errors in sample surveys are to be stressed.

The concept of sampling error is valid only under the probability sampling scheme or popularly known as random sampling. Unless the sample is randomly drawn, the calculation of sampling error has got little meaning and statistical theory is hardly of any use to interpret the results. What is random sampling, how to draw a random sample, how to generate a population of samples from a given parental population, how the result of each random sample has got the same interpretation, etc. should be thoroughly explained with the help of numerical examples. Although the concept of random sampling that each unit of the population has got a preassigned probability of selection in the sample may be introduced in a general way, the theory of the simple random sampling, may be discussed first in greater detail as it is easier to explain the concept of unbiased estimate, the dependence of sampling variance on variations among the units and the units and the size of the sample, and the effect of the finite population. Besides deriving the algebraic expression for sample variance of the simple arithmetic mean, say $\overline{y}_n$, n being the sample size selected by simple random method without replacement, which may appeal more to the mathematically trained students, it may be interestingly shown with the help of numerical exercises as to how the sampling variance

$$V(\overline{y}_n) = \frac{N-n}{N} \frac{S^2}{n},$$

where N and S^2 have the usual meaning.

This is a very important formula in the theory of sampling. The development of various sampling designs can be linked with this formula.

The sampling variance of the mean depends upon S^2 and n. By increasing n, the sample size, the sampling variance can be reduced to any level. But an increase of the sample by an additional unit will result in an additional cost. Here the cost concept in sample survey may be introduced. The major advantage of using sampling method is to obtain maximum information at a

given cost or alternatively to get the required information at a minimum cost. At this stage, the various sampling strategies which may help in reducing the sampling variance, particularly the factor S^2 may be introduced. In many situations, a considerable amount of supplementary information about the units of the population may be available, or can be easily collected from the records maintained by public and private agencies. Depending upon the way the information is available, the improved sampling designs may be introduced in sequence of simple random sampling. In most of the textbooks on sampling, immediately after the simple random sampling, stratified random sampling has been discussed. In the opinion of the author, immediately after simple random sampling with and without replacement, sampling schemes, selection of units with probability proportional to size (pps), and ratio and regression methods of estimation, which use supplementary information for improving the accuracy of the estimate may be taught. Courses relating to double sampling and multi-phase sampling, and even that of successive sampling may be taken as an extension of ratio and regression methods of estimation. Stratified sampling may be introduced as another way of using the supplementary information for improving the estimate. Sometimes geographical stratification is used for administrative reasons. All the topics listed above may again be discussed under stratification of the population and the theoretical implication, particularly with regard to distribution of sample size among different strata, may be highlighted.

After covering the courses relating to the sampling techniques which use supplementary information on sampling units, as well as their combinations, the cluster sampling and multi-stage sampling may be taught in sequence. Since multi-stage design is of most common use in actual sample surveys, sufficient time should be devoted to discussing various aspects, such as methods of selection of sampling units at different stages, method of estimation including ratio and regression methods, variance components, determination of the sample size and its distribution at various stages, cost function, relative efficiency, self weighting design etc.

Systematic sampling and its variations, such as linear, circular, balanced and two dimensional may preferably be discussed as a separate topic after the students have completed the course on all aspects relating to random sampling. It may then be easy to explain its similarity to cluster sampling and stratified sampling. How important it is to avoid the systematic method of selection of the primary sampling unit in multi-stage sampling design can be better explained when the students have learnt all about the multi-stage sampling.

In a large scale complex survey, the calculation of sampling errors of various estimators becomes a time-consuming process and this is the reason that in a majority of sample survey reports, the estimates are presented without corresponding standard errors. It does not need any emphasis to say as to how difficult it is to give proper interpretations to such results. As the users are unaware of the limitation of the statistics calculated from sample surveys, they tend to utilise these figures without realising that they are subject to varying degrees of sampling and probability non-sampling errors. It may, therefore, be useful to give a course on short-cut methods of calculating the sampling errors. Many such short-cut methods have been discussed by Zarkovich (1965). In this connection, it may be useful to introduce the concept of independent interpenetrating or replicated sub-samples, as this technique is very useful for estimating sampling variances in case of complicated sampling designs. Further, this technique is also useful in measuring some types of biases and non-sampling errors.

Most of the sampling designs listed in the textbooks relate to uni-purpose or uni-character surveys. Owing to considerations of economy, operational convenience, and availability of limited qualified personnel, multi-purpose surveys have, however, become very common, particularly in many developing countries. The proposed United Nations Household Survey Capability Programme is one of the examples of multi-purpose surveys. One of the major problems concerning the multi-purpose surveys is to formulate a sampling design to estimate multiple characteristics with possibly different levels of accuracy, necessitating a search for an optimum or near optimum balance between diverse requirements, usually under cost and other operational restrictions. Although considerable work has been done on the subject and adequate experience has been gained in planning and conducting multi-purpose surveys, particularly in developing countries (Murty, 1974), it is hardly reflected in the textbooks on sampling. The teachers may have to collect the materials from statistical journals and supply survey reports and communicate the information to the students, highlighting the importance of the subject.

The last course may relate to biases and non-sampling errors in sample surveys. Although a reference to biases and non-sampling errors should always be made while initially introducing the concept of errors in sample surveys, detailed discussions on the causes of these errors and methods for minimising their effects may necessarily be conducted at the final stage of lecture courses. Fortunately, most of the textbooks on sampling have listed the causes of non-sampling errors and possible methods to minimise their effects, and these can be the basis for preparing teaching materials on this topic. However, unlike the sampling error, no systematic theory for treating all kinds of non-

sampling errors has so far been developed. Each kind of non-sampling errors and biases has to be dealt with separately and this can be better supported with numerous examples of actually conducted sample surveys and censuses.

LABORATORY WORK, DEMONSTRATIONS AND SEMINARS

Survey sampling may be taught as technology for data production. Although the theory of sampling, which mainly depends upon probability and mathematics for its development, is very general, its application is broadly governed by the area of application. Therefore, the application aspects of sampling may be taught better in the statistical laboratory through practical exercises based on live data, demonstrations involving actual field work, and seminars on special types of surveys.

Lectures on a sampling design should follow a number of practical exercises. If some recent census data relating to agriculture or human population are available, that may be treated as statistical population and parameters of the population may be determined. Under various sampling designs, sample estimates may be obtained. If no census data with necessary details are available, data relating to some large scale sample survey may be used, treating it as population. From such data, depending upon the sampling design, estimate, estimate of sampling variance, confidence interval, relative efficiency of the design under realistic cost function models, the sample size needed for given precision, etc. may be worked out by each student. If in certain exercises the computation involved becomes large, the work may be suitably divided among students and results of computation may be then pooled to give the final picture.

Since the use of electronic computers and electronic calculators has now become common, the students may be encouraged to utilise these facilities in their computational work. In fact this may, incidentally, help the students in learning computer programming, now an important qualification for a statsitician.

Practical exercises may also include some empirical studies. It is now well known that there does not exist a uniformly best design in sample surveys (Godambe, 1955). Only through the empirical studies, superiority of one sampling design over the other under some conditions can at best be demonstrated.

Unfortunately, not much attention has been paid to this aspect of teaching of survey sampling. Teachers, particularly in universities, find great difficulty in gathering live census and survey data for preparation of practical exercises. To make the practical exercises effective, these should be based on national data. Realising this need, a 'Hand Book on Sampling' containing practical exercises on sampling techniques based on live data collected through sample

surveys conducted in India was brought out (Singh, 1978). Similar publications are needed in other countries, where survey sampling is taught as a separate discipline.

Practical demonstration of application of sampling technique, from the initial stage of planning to the final stage of preparation of survey report should be considered as an integral part of the course. If the teaching centre itself is engaged in the conduct of sample surveys, students may be associated with the projects. They should participate in the preparation of sampling design and sampling frame, drafting of questionnaires and field instructions including concepts and definitions involved, training of field and supervisory staff, fieldwork for collection of data, scruting and editing of questionnaires received from the field, development of appropriate estimation procedure consistent with the sampling design, formulation of tabulation programme, and finally drafting of the project report. However, if the teaching centre has no such sample survey project of its own, it may formulate a live project in consultation with other teaching and research departments of the university or an agency which can finance such a project. Students under the supervision of the teachers may be made responsible for implementing it.

The demonstrations of conduct of an actual sample survey project help in establishing the firm link between sampling theory and its practical applications and bring out numerous problems faced by the survey statisticians. A student of a mathematically oriented sample survey course may not be able to imagine and appreciate the types of problems faced by a statistician responsible for planning and implementing sample survey projects unless he is brought into contact with actual work at all its stages.

The seminars should be made a regular feature of the sample survey course. The topics of common applications, but of controversial nature such as purposive sampling versus probability sampling, census versus sampling, repeat and rotation sampling, evaluation studies involving estimation of frequencies and percentages, etc. may be the subjects of discussions at the seminars. Besides students and teachers, other experienced survey statisticians employed in the government organizations should participate in such seminars.

TEACHER AND INSTITUTION

It may not be out of the way, at this stage, to say a few words about the qualifications of the teachers and teaching institutions. No matter how fine the materials and contents of the course may be, ultimately it is the quality of human resources, which decides success of a programme. We are well aware of the necessity of high standards of training in both theory and

practice. Since sampling is basically an applied subject, a balance between the theory and its practice has to be maintained while organising the course. The effectiveness of the teaching programme will ultimately depend upon the qualities of teachers and the taught.

The teacher of such a course should have training not only in mathematical aspects of the theory of sampling, but should also have adequate experience in its application for collection of data and their interpretation. Unfortuantely, the teachers in most of the universities which are the main centres for teaching of survey sampling, have hardly any orientation in the application aspect. As a result, the training remains lop-sided and when a student after successfully completing such a course takes an appointment in a private or public undertaking where his main job is to design and conduct sample surveys he finds himself completely unfit and loses confidence in his ability to do the job. That is why a paradoxical situation has arisen in many countries that unemployed statisical personnel and vacant positions exist side by side.

The universities and other teaching centres have, therefore, to orient their teaching programmes to meet the pressing needs of the country. Ideally, the best institutions for teaching an applied subject like survey sampling are those which are simultaneously responsible for an integrated programme of education, research and extension. Conceptually, extension may mean here the use of training and research to solve the practical problems faced by the country. This system will automatically produce teachers who are fully acquainted with the latest development of the subject through research, and have realisation of the relevance of the course through practical applications. The author discussed this particular aspect in a more general way in his presidential address at the Indian Science Congress (Singh, 1976).

The question is now whether the universities are capable of developing such a system or alternatively, whether new institutions with specific objectives in view can be created. Happily, some such institutions, to name a few, Survey Research Centre, Institute of Social Research, Univeristy of Michigan; International Statistical Programme Center, US Bureau of the Census, Washington D.C.; Indian Statistical Institute, Calcutta; Indian Agricultural Statistics Research Institute, New Delhi etc., with an integrated programme of teaching, research and extension of survey techniques already exist. The requirement of survey statisticians is indeed great. Many more such centres need to be established. Some of the universities with strong departments of mathematical statistics and other necessary infrastructures may orient their teaching and research programme to meet the requirement of trained survey statisticians. The International Statistical Institute with the active cooperation of the Inter-

national Association of Survey Statisticians, and with the financial aid from international agencies, may consider the establishment of some centres for training and research in survey sampling to meet the needs of developing countries. The Institute has already shown the way by organizing the World Fertility Survey on a world-wide basis.

CONCLUDING REMARKS

Sampling technique may be considered as technology for information production needed for planning socio-economic development. In teaching the sampling course, a balance between sampling theory and its application is required to be maintained. This can be achieved only if a course on design follows practical exercises based on live data. Demonstrations of conduct of actual surveys by involving students from the initial stage of planning to the final stage of reporting the results may be considered as an integral part of the course. A proper sequence in planning the course may be helpful from the communication point of view. Adequate emphasis on the role of sampling and non-sampling errors in interpreting the survey results may be laid. To maximise the utilities of training and research in sampling techniques, there is a need of developing an integrated programme of teaching, research and extension. To meet the increasing demand of trained survey statisticians, particularly in the developing countries, the teaching programme in universities may need orientation. There may be justification for establishing regional centres for training and research in sampling in developing countries with the aid of international financing agencies. The International Statistical Institute may take a lead in the matter.

References

[1] Cochran, W.G. (1953,1963) Sampling Techniques. John Wiley & Sons, New York.

[2] Deming, W.E. (1960) Sample Design in Business Research. John Wiley & Sons, New York.

[3] Des Raj (1968) Sampling Theory. McGraw Hill Book Co., New York.

[4] Godambe, V.P. (1955) A unified theory of sampling from finite population. Journal of the Royal Statistical Society, (B), 17, 269-278.

[5] Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953) Sample Survey Methods and Theory, Volumes I and II. John Wiley & Sons, New York.

[6] Kish, L. (1967) Survey Sampling. John Wiley & Sons, New York.

[7] Murty, M.N. (1967) Sampling Theory and Methods. Statistical Publishing Society, Calcutta.

[8] Murty, M.N. (1974) Evaluation of Multi-subject Sample Survey Systems. International Statistical Review, 45, 175-192.

[9] Singh, D. (1976) Economic Planning of Rural India, Statistics, and Statisticians; Presidential Address, Statistics Section, Indian Science Congress, 63rd Session, Indian Science Congress Association.

[10] Singh, D., Singh, P. and Pranesh Kumar (1978) Handbook on Sampling. Indian Agricultural Statistics Research Institute, New Delhi.

[11] Sukhatme, P.V. (1953) Sampling Theory of Surveys with Applications. Iowa State College Press, Ames, Iowa, second edition with Sukhatme, B.V. (1970).

[12] United Nations (1947-1951) Reports of the Sub-commission on Statistical Sampling; E/CN 3/37, 52, 83, 140, New York.

[13] Yates, F. (1949,1960) Sampling Methods for Census and Surveys. Charles Griffin & Co., London.

[14] Zarkovich, S.S. (1965) Sampling Methods and Censuses, FAO, Rome.

TEACHING OF SURVEY SAMPLING

P K Bose, Calcutta University

Statistical thinking is a form of logical thinking and is a part of our daily lives. When we say something or somebody is typical we are thinking in terms of statistical averages and departure from type is statistical variation. When we generalise from a few cases to a very large number, we employ sampling processes. A conclusion that two things always go together involves a pattern of thinking found in statistical correlation. Statistical thinking is not foreign to everyday thinking but is a scientific form of it. Statistics is a fundamental activity of mankind.

At present we are noticing two important trends in statistics which need some rethinking in teaching statistics in the Universities/Colleges and other institutions. One is its massive impact on society and the other is the extensive growth of statistical theory.

The merging of these two trends into a single stream makes new demands on our statistical personnel. Neither the isolated theorist nor the practitioner is able to keep abreast of the current. The specialist in mathematical statistics must acquaint himself with statistical problems and equally working statisticians must familiarize themselves with the logic of statistical theory.

It is necessary to bring students into awareness of and harmony with the statistical content of society. This content is extensive. Such large scale activities should be integrated with academic concepts such as probability, sampling, estimation and testing of hypotheses.

It is felt that the tools of statistics should be brought out and sharpened only after the need for them has been realised.

Reasoning from sample to universe is called statistical induction. Inductive reasoning involves observing individual cases and arguing from them to a general principle. In statistical induction we make inferences concerning the characteristics of a large group, sometimes of an infinitely large group, from the characteristics of a specifically chosen, usually small, part of the group that we observe. We draw conclusions from the smaller group and apply these conclusions to a larger mass.

In statistical induction we are able to add to our generalisations a statement as to their precision.

It involves the following stages.

1. Define the trait to be measured.
2. Measure this trait in a sample.
3. Apply the measure to the universe from which the sample is drawn.
4. Estimate the error involved in this application.

Basic sampling technique deals with the four stages enumerated above.

For teaching survey sampling the first course should be statistical reasoning, it should be followed by major statistical tools and the third course should be sampling methods.

Sampling plays a vital role in research design involving human populations, public health, business administration, public administration, commerce and industry and various disciplines of social, biological and physical sciences.

Excepting routine teaching on sampling in various universities and colleges, a very few institutions organise training programmes on survey sampling. In the Indian Statistical Institute we carried out experiments in statistical survey sampling for over a decade which has laid the foundation of survey sampling on a large scale in the world.

Sample enquiries fall into two broad categories. One is sample survey on an extensive scale covering a wide area in which the object is to obtain reliable estimates of the acreage, rate of yield per acre and total production of food and fibre crop and of economic or demographic factors relating to indebtedness, unemployment, distribution etc. The second type of work is an enquiry which is more localised and related to cost and level of living, housing, consumption of food etc.

The difference between the two is great so far as quantity is concerned and this brings in practice a change in quality. As Professor P.C. Mahalanobis suggested, large scale sample surveys may be appropriately called statistical engineering. Teaching of survey sampling should be concerned more with statistical engineering rather than with the pure theory of sampling.

In preparing the design of each sample survey the three most important considerations are time, cost and the human agency - all of course in relation to the permissible margin of error. I have already indicated that the teaching of survey sampling should consist of three sections, namely fundamentals of statistical reasoning, statistical tools and sampling methods. I am not discussing the first two; they are readily available in most text books on statistics.

Regarding the third I have already discussed the essential difference in concepts between a small sample enquiry and large scale sample surveys. Sampling theory as applied to small or localised enquiries is also elaborately dealt with in many text books. Teaching is carried out mostly on the basis of such books.

Unfortunately, very little attention is given to experiments to be carried out in the laboratories; unless the students carry out some experiments themselves, the fundamentals of sampling will never be clear.

I may give a few illustrations.

Experiment one

Conduct opinion poll sampling of student body; construct questionnaire, design sample, discuss interview technique, tabulate results, make inferences about population.

If the above experiment is carried out in the laboratory the students will enjoy this experience, and will ask many penetrating questions. This may open opportunities for sound instruction.

Experiment two

Extend sampling experience by use of coins or dice or cards. Record result of each toss or draw.

Verify probability laws empirically.

Experiment three

Use bowl containing equal number of beads of two colours or use table of random numbers with equal numbers of even and odd digits. Draw many samples of ten and record the number of successes.

Emphasise sampling variation and observe improved reliability when samples are combined into large samples.

Method of teaching statistics, and more especially sampling, should be based on experiments which combine observations, calculations and inference.

In teaching sampling theory emphasis should be laid on this aspect.

The second important aspect which is overlooked while teaching is the quality problem of statistical data. No method of collecting and processing data is safe from errors. Some ideas about errors that statisticians are expected to study and evaluate should be given. Errors in the broad sense of the word may be classified in three groups.

A. Errors resulting from inadequate preparations.
B. Errors committed in the stage of data collecting.
C. Processing errors.

Necessary precautions should be taken for reducing these components of errors. Large scale sample surveys which form the major part of survey sampling have a number of distinguishing features, namely the concepts of optimum design of surveys, pilot surveys and inter-penetrating network of sub-samples and the

control of non-sampling errors. In survey sampling the magnitude of non-sampling error which is due to various causes is more than the sampling errors. A detailed discussion of these topics should form part of a teaching programme.

The pioneering work of P.C. Mahalanobis in the field of large scale sample surveys is widely recognised. He demonstrated the possibility of using large scale surveys to collect information quickly and economically with sufficient accuracy.

COST FUNCTION

Large scale surveys are necessarily expensive and considerations of cost are therefore of great importance. Total cost of operations depends on the size of the grids and their total number. Working within a fixed budget this means once the size of grids is settled, their total number must also be determined by the budget limit.

Professor Mahalanobis obtained the total cost function in his study on sample census of Jute as

$$\phi = c_o + c_1 y_i - c_2 y_i^2 + c_3 x_i y_i$$

where x = size of sample unit and y = density of sample unit.

VARIANCE FUNCTION

The precision of the final result is important. The variance of grids or sample units decreases as the size of the grid is increased.

Professor Mahalanobis obtained the variance function as

$$v = \frac{\Sigma A_i \alpha p_i q_i}{y_i (x_i)^g}$$

where A_i = Area, p_i = proportion of land under crop.

OPTIMUM DESIGN

It is now possible in terms of the cost and variance functions to state the conditions for an optimum or most economical design for sampling. The aim may be stated in either of the following two alternative forms:

a) to determine the size and distribution of sample units in such a way that the variance of the final estimate is a minimum when the total cost of the survey is fixed;

or

b) to determine the size and distribution of sample units so as to reduce the variance of the final estimate to any desired level at a minimum cost.

Professor Mahalanobis obtained the most economic design for sampling with the help of an empirical knowledge of variance and cost functions.

SEQUENTIAL DEVELOPMENT

In large scale work, the survey has often to be continued from year to year or at suitable intervals. It is possible to adopt an exploratory method in which a sample survey is carried out on a very small scale with the primary object of collecting basic information required for preparing an efficient design for later surveys.

In the exploratory stage the object is to gather as quickly as possible some rough idea of the variance function, cost of operations from one region to another type and nature of human agency likely to be suitable for field work. As basic information begins to be gathered the design of the survey steadily improves and the cost per unit rapidly decreases in subsequent surveys.

INTER-PENETRATING SUB-SAMPLES

One of the chief objects of the sample survey is to test the accuracy of results attained in practice by a field trial on a large scale. The information is collected in the form of two independent half-samples, each of which consists of nearly half the total number of sample units. In each half-sample the sample units are located at random over the whole area and the field operations for each half-sample are carried out by different sets of investigators. The difference between the results obtained from the two half-samples will furnish a good idea of the margin of error inherent in the present method.

The teaching of survey sampling units should include mainly the following.

A. Introductory courses on statistical reasoning and statistical tools.
B. Fundamentals of survey sampling.
C. Special problems etc. and techniques of large scale sample surveys.

Topics related to A and B are well known. They have been discussed in different platforms and good text books are also available on them. Regarding C it is still uncharted. A few items covering the same are given below.

(a) Terminology and classification of sample surveys.
(b) Cost function; variance function; optimum solutions.
(c) Sequential development technique.
(d) Planning of surveys.

(e) Classification of errors.
(f) Controls at the stage of statistical processing.

References

[1] Mahalanobis, P.C. Experiments in Statistical Sampling. Asis Publishing House, Calcutta.

[2] Mahalanobis, P.C. Sample Census of Area under Jute in Bengal 1940. Statistical Publishing Society, Calcutta.

[3] Mahalanobis, P.C. (1944) Large scale sample survey. Philosophical Transactions of the Royal Society, London Series, 231.

[4] Sukhatme, P.V. (1947) Problem of plot size in sample surveys, JASA.

[5] Snedecor, G.W. (1948) A proposed basic course in statistics, JASA.

[6] Seng, Y.P. (1951) Historical surveys of the development of sampling theories and practice, JRSS (A).

[7] Zarcovich (1956) History of sampling in Russia, JRSS (A).

TOPIC 13: TEACHING DESIGN AND ANALYSIS OF EXPERIMENTS

ORGANISER: DR D A PREECE

THE DESIGN AND ANALYSIS OF EXPERIMENTS - WHAT IS IT ABOUT?

D A Preece, Rothamsted Experimental Station

ABSTRACT

Basic definitions are given. Then the subject is divided into six "components", which are discussed in turn:-

I: Planning, design and layout.
II: Management of the experiment or experiments.
III: Data recording.
IV: Scrutiny and editing of the data.
V: Computational analysis.
VI: Interpreting and reporting the results.

Over-theoretical, over-mathematical teaching of the subject is criticised. The importance of practical considerations is stressed.

Key Words:
analysis of variance; Balances; Blocking; Checking copying and calculations; confounding; Covariance analysis; Data recording; Data scrutiny; Experimental error; Factorial designs; Null hypotheses; Randomisation; Repeated measures; Replication; Report-writing; Significance; Split-plot designs; Transformations.

INTRODUCTION

Within the subject of this paper, an "experiment" is "comparative" (Anscombe, 1948, Cox, 1958, Section 1.1; Finney, 1960, Section 1.2): it compares different "treatments", which might be different varieties of wheat, different diets for pigs, different drugs for insomniacs, or different acid concentrations in an industrial process. Many comparative experiments are "factorial" (Yates, 1937; Cox, 1958, Chapter 6), permitting simultaneous study of different "factors", each at different "levels". Thus in a fruit-tree experiment, three pruning methods (factor A, with 3 levels) might be tried in conjunction with four sprays (factor B, with 4 levels); in this "3x4 factorial experiment", the twelve combinations of pruning method and spray are the 12 "treatment combinations". A factor may be "qualitative" (e.g. different colours in a psychological experiment) or "quantitative" (e.g. different doses of Vitamin C).

Each experiment comprises a set of "experimental units" (often called "plots", and sometimes "runs), with different treatments or treatment combinations applied to different experimental units. Plots may be, for example, pens of animals, individual animals or plants, individual leaves on growing plants,

or (in an inoculation experiment on plants) half-leaves. Commonly, treatments are "replicated", i.e. there is more than one plot per treatment. However, factorial experiments include "single replicate experiments" (one plot per treatment combination) and "fractional replicates" (not all treatment combinations present).

The comparison between treatments, between factor levels, and between treatment combinations are based on data recorded after the treatments have had a chance to affect the experimental material. These data are one or more "variates" (often called "variables" - although this term is sometimes used for "factors"), each variate consisting of one value for each plot (except possibly for "missing values"). Variates may be "continuous" (e.g. weights, heights, temperatures, pressures) or "discrete" (counts, scores); they may be "primary" (obtained by direct reading of a scale, by counting, or by scoring) or "derived" (obtained from one or more other variates by arithmetic calculation). In statistical terminology, "analysis of an experiment" is elliptic for the analysis of a variate or variates, possibly in conjunction with other data pertaining to the experiment. For a factorial experiment with quantitative factors, the analysis of a variate may consist of investigating a "response surface"; for present purposes, experiments designed to study response surfaces are regarded as comparative.

"Design" is more difficult to define than "analysis", not least as it means both "designing" (part of "planning") and "a specific design". Finney (1955, Section 1.2; 1960, Section 1.5) defined the "design" of an experiment to mean

(i) The set of treatments selected for comparison;
(ii) The specification of the plots;
(iii) The rules for allocating the treatments to the plots;
(iv) The specification of the measurements or other records to be made on each plot.

A fifth component should perhaps be listed as being a legitimate concern of "designers", whether or not it is strictly a part of design:

(v) Specification of details of the management of the experiment.

But the "design and analysis of experiments" deals not only with individual experiments, but also with series of experiments, indeed with whole experimental programmes. Within a series, there may be experiments in different locations, or on different occasions (perhaps in different years), or both (John and Quenouille, 1977, Chapter 12).

The need for a statistical subject of "design and analysis of experiments" arises from variability inherent in the experimental material, environment, and management. This variability leads to uncontrolled - indeed uncontrollable - variability in the observed variate values, i.e. to "experimental error", where "error"

does not mean "mistake". A variate affected differentially by the treatments thus includes variability from one treatment to another, plus the "error" variability that would have been present even if a single treatment (or none) had been applied throughout. The experimental error can be regarded as "background noise" such as the "static" heard when radio reception is poor. Differences in effect between treatments can then be likened to a broadcast message; if the message is spoken quietly (treatment differences small), it may be drowned out by the static.

COMPONENTS OF THE SUBJECT

The main principles, tooks and skills of our subject can perhaps be summarised in the following "Top Twelve" headings, which I have tried to arrange in a logical order:

(1) The basic principles of statistical inference.
(2) The ability to select treatments, experimental material, experimental units, and variates that will permit "useful" results to be obtained.
(3) The ability to recognise relevant sources of variability, and to make preliminary assessments of their importance and relative magnitudes.
(4) The basic principles of "good" design.
(5) Mathematical methods for constructing designs and obtaining analyses, particularly analyses of variance and covariance.
(6) Trained personnel for setting up and managing experiments.
(7) Proper instruments (balances, etc.) for obtaining accurate, precise data values, and trained personnel to use the equipment.
(8) Training in data transcription (with checking).
(9) Ability to recognise and interpret patterns and anomalies in data.
(10) Understanding of the assumptions underlying standard statistical analyses, and the ability to recognise when these assumptions are seriously violated in practice.
(11) Facility in spotting mistakes in analyses.
(12) Skill in reporting conclusions in clear, simple prose, with appropriate tabulations and graphs.

In the teaching of the design and analysis of experiments - whether in textbooks or lecture-rooms, whether in syllabuses or examinations - some of the above twelve have received much less time and attention than they deserve. The emphasis has often been on the theoretical and the mathematical, to the detriment of the applied and the practical. Derivable formulae, provable theorems and repeatable mathematical procedures have taken precedence over studies of the environments in which experiments are done, over getting a

proper "feel" for experimental data, over thinking about the validity of assumptions and of results, and over report writing. The emphasis given to different parts of the subject must, of course, depend on the audiences, but insufficient attention to the Real World of real experiments transforms the subject from statistics to mathematics, from science to ritual (Preece, 1982b).

So the words "design and analysis" of themselves provide a very deficient picture of the content of our subject. To avoid both the over-simple dichotomy between "design" and "analysis" and the over-elaborate set of twelve headings above, I now break up the subject into six components, each demonstrably of statistical concern, that may be labelled roughly and simply as follows:-

I Planning, design and layout;
II Management;
III Data recording;
IV Scrutiny and editing of the data;
V Computational analysis;
VI Interpreting and reporting the rsults.

Elsewhere (Preece, 1982a) I did not include the present component II; however, as bad management of an experiment can introduce major variability and biases, I now believe that management merits specific mention. Other lists can of course be produced, with different amounts of detail and different emphases; the nineteen "Main Features of a Clinical Trial" given by Pocock (1982, p.3) are a good example.

COMPONENT I: Planning, Design and Layout

In the vocabulary of Jeffers (1978), this first of my six components includes

(a) Stating the objectives [of an experiment];
(b) Defining the population about which the inferences are to be made;
(c) Selection of experimental treatments;
(d) [Choosing the] Plot shape and size.

The choices (c) and (d) cannot be made separately from the choices of treatments to be included and of number of plots; all these choices, and the allocation of treatments to plots, are to be made in accord with the three basic principles (or "three R's") of Fisherian experimental design: Replication, Randomisation and Blocking. The three were discussed in the first edition of "Statistical Methods for Research Workers" by R.A. Fisher (1925); their explicit formulation can be taken to mark the start of the statisical subject of the "Design and Analysis of Experiments". [Writing for agricultural research workers, Little and Hills (1978, p.9) listed a different "three R's": Replicate, Randomize, and Request help! Das and Giri (1979, p.7) gave the Fisherian three as

Randomization, Replication and Error Control, with "Error" again meaning "experimental error".]

Replication of a treatment involves the treatment being applied to more than one plot; the number of plots carrying the treatment is the number of "replications" or "replicates" of the treatment. Confusingly, the word "replication" is also used for a set of plots within which each treatment occurs exactly once. Replication enables estimates of experiments error to be made.

Randomization (Cox, 1958, Chapter 5; Finney, 1960, Section 1.6) is a procedure for allocating the treatments to the plots. The allocation adopted is chosen at random from a set of possible allocations, the set being such that the experiment shall provide unbiased estimates of treatment differences and of experimental error. In many experiments, randomization gives all the plots an equal chance of receiving any particular treatment or treatment combination; there will not however be equality in a properly randomized experiment with twice as many plots for a control treatment as for any other treatment. Randomization avoids the biases that can arise from subjective allocation of treatments to plots.

Blocking is a grouping of the experimental units into "blocks" (often all containing the same number of plots) with the aim of reducing the experimental error or "background noise" in variates to be studied. The grouping is done with the hope of capturing any large components of the noise as "between blocks" variability which can then be filtered out in the statistical analysis: plots likely to give the most similar results in the absence of treatment effects are, as far as possible, grouped together within blocks. Perhaps the commonest experimental design is the "randomized complete block design" (often just "randomized block design"), in which each block contains exactly one plot for each treatment, the allocation of treatments being randomized independently within each block; this design is used for many agricultural field experiments, often with each block a compact group of adjacent plots. Some experiments, notably on fruit-trees in orchards, take the idea of blocking further, in having a "row-and-column design" with two overlaid systems of blocks, namely "rows" and "columns", at right angles to each other.

Adherence to the Fisherian "three R's" helps to satisfy the "Requirements for a Good Experiment", which Cox (1958, Section 1.2) gave as

(A) Absence of systematic error;
(B) Precision (suitably small random errors of estimation);
(C) A wide range of validity for the conclusions that are to be drawn;
(D) Simplicity;
(E) A proper statistical analysis to be available which, without artificial assumptions, permits calculation of the uncertainty in the estimates of treatment differences.

Because of the need to choose suitable experimental material and an appropriate design for that material, it is often desirable for statisticians involved to visit the location of an experiment, to join the experimenters in examining the material, and to have discussions not only with the organisers but also with supervisory and experimental-management staff.

Planning discussions about (a), (b) and (c) must often cover the possible inclusion of a control treatment. Sometimes a strictly "untreated" control may seem appropriate, sometimes - in medical vocabulary - a "placebo".

Basic ideas about factorial experiments were set out by Yates (1937) and rehearsed by Cox (1958, Chapter 6). Fisher (1926) pointed out that an experiment studying several factors simultaneously can have various advantages over a succession of experiments each on one factor only; as paraphrased by Yates (1964, p.312), the advantages are:-

(1.) the plots are used several times over to determine the average effects of the different factors;
(2.) only by factorial design can any information be obtained on how responses to one factor are affected by another (i.e. on interactions);
(3.) factorial experiments provide a wider inductive basis for conclusions on the effects of the factors.

Some factorial experiments have "split-plot" designs, the simplest of which has two factors, A and B, with the levels of A allocated to "main plots" or "whole plots", each subdivided into as many "sub-plots" as B has levels; the levels of B are assigned at random to the sub-plots within each whole plot, a separate randomisation being used for each whole plot. Because of the two sizes of plot, and the different randomisations for A and B, the analysis of variance for the design has <u>two</u> error terms - "whole-plot error" and "sub-plot error" - not just a single "(experimental) error". There are three possible reasons for using a split-plot design (Finney, 1960, Section 5.13, p.85):-

(α) Practical considerations (e.g. size of machinery) may make it undesirable or impossible to apply the levels of some factor(s) to small plots, whereas these levels may be feasible on large subdivided plots.
(β) an experimenter may wish to add some further factor(s) into an existing experiment, and this may be possible by putting the new factor(s) "on sub-plots" within the existing plots.
(γ) Often, error variability between whole plots is greater than that between sub-plots; then, the effects of sub-plot factors, and the interactions between sub-plot and whole-plot factors, are estimated more precisely than if the sub-plots were not grouped into whole plots, whereas there is a loss of precision on main-plot factors.

The last of these reasons seems to be of little importance in practice. Despite reports of mis-use of split-plot designs (e.g. Mead, 1982), plot-splitting is very important in the design and analysis of experiments. The omission of split-plot designs from an admittedly mathematical "systematic treatise" on factorial designs (Raktoe, Hedayat and Federer, 1981) is unsatisfactory and bizarre.

Some factorial experiments have fewer plots per block than there are treatment combinations. Important effects of the factors are, however, still estimable if suitable subsets of the treatment combinations are allocated to the blocks, the subsets being obtainable from simple mathematical relationships that define a "confounding scheme". Then the design may be said to be "confounded"; more strictly, certain effects (usually interactions between as many factors as possible) are "confounded with blocks", totally or partially (Yates, 1937; Cox, 1958, Section 12.3). Interactions that have been confounded are usually ones in which the experimenter has little interest, or which the experimenter believes to be negligible.

Some <u>non</u>-factorial designs have fewer plots per block than there are treatments. Such "incomplete block designs" include the "balanced incomplete block designs" (Cox, 1958, Section 11.2; Finney, 1960, Sections 6.2 to 6.6) that have been of so much interest to combinatorialists.

COMPONENT II: Management

The administration, supervision and management of an experiment need systematic care if reliable results are to be obtained. If the plots are pieces of land, these must be measured up and planted properly. If the treatments are different rates or doses, these must be determined accurately. Steps must be taken to try to ensure that each plot receives the intended treatment or treatment combination; for field and orchard experiments, this requires a plan of the physical layout, whereas various options are available for clinical trials (Pocock, 1982, p.4).

Management of a field experiment continues to be statistically important after the treatments have been applied. Extra variability can be introduced into the results by differential management of the plots, e.g. by having some plots weeded more carefully or frequently than others in a field experiment where treatment differences are not intended to include weeding differences. Such circumstances may require different blocks to be weeded by different people, so that unwanted "between labourers" differences are assimilated in "between blocks" differences. Likewise, if an experiment has to be irrigated throughout, but there is insufficient equipment to irrigate it all at once, the irrigation should be "by blocks". Also, such operations as scoring for disease - and

usually harvesting itself - should be "by blocks" (Dyke, 1974, p.43).

COMPONENT III: Data Recording

The need for suitable balances, etc., for obtaining accurate, precise data values has already been mentioned. When weights, lengths, and other continuous variates are to be recorded, a "degree of precision of recording" should be specified and adhered to, e.g. "to the nearest 5gm", "to the nearest 25cm", etc. A handy rule-of-thumb specifies a variate as having an acceptable degree of precision of recording if the "grouping interval" (5gm, 25cm, etc.) is less than one-quarter of the expected standard error per plot (Cochran and Cox, 1957, Section 3.33, p.60). Sometimes, up to one-half of the expected standard error per plot is considered acceptable (e.g. Dyke, 1974, Appendix A, pp.162-164).

Data should be recorded on prepared data sheets having proper headings. Copying of data introduces errors and should be avoided if possible; unavoidable Copying should be checked by "calling over" (Dyke, 1974, p.49), and good practice requires copies to be signed by the two people involved in the calling over.

In field experimentation and industrial experimentation, some variates are obtained from sampling within the plots (see, e.g. Gomez and Gomez, 1976, Chapter 13). A pre-determined sampling procedure must then be adhered to for each plot, so that the recorder's selection is not influenced by what is found in a plot.

Variates consisting of scores need special care, as scoring can be very subjective. Scoring schemes need to be well worked out in advance, and scorers need training and practice. Scoring should usually be "by blocks" (Dyke, 1974, p.43).

Data recorders themselves sometimes calculate derived variates. Such calculations should be checked, and everyone involved should have been taught that the worst way to check calculations is merely to repeat them.

COMPONENT IV: Scrutiny and Editing of Data

Science is not well served when experimental data are merely "processed" in a standard computer analysis, without anyone first "sniffing over" the data to note any salient patterns or anomalies. All too many experimenters do not adopt a critical attitude towards their data and, in this respect, cease to be scientists. Anomalous values (e.g. values much larger or smaller than might be expected, values not recorded to the chosen degree of precision, etc.) inevitably occur, and some can readily be found to result from erroneous

copying or flawed experimentation; also, an unexpected value could well be the most important result from an experiment.

Skill in "sniffing over" data is an important part of "Statisticianship", which is to statistics what musicianship is to music; a person without musicianship (or statisticianship?) hardly deserves to be called a "musician" (or "statistician"?). Data-sniffing skill is gained through experience rather than learnt (as are many of a musician's skills), but opportunities for gaining some of the experience can easily be provided when the design and analysis of experiments is being taught. Some statisticians claim to handle far too much data for them to have time to look through most of the data before analysis; however, a statistician with a well-developed data-reading skill can pick out important anomalous values very quickly indeed, with the ease with which an orchestral conductor, with his score-reading skill, can pinpoint misprints in a full orchestral score. Similarly, skill in detecting heterogeneity of error can be highly developed.

Editing of data clearly involves correction of known mistakes, but what is to be done about unexplained outliers and other doubtful observations? Hard and fast rules cannot be provided for this, and judgement must often be used. Discussion between experimenter and statistician may be needed or desirable.

COMPONENT V: Analysis

Writing of the experimental designs considered in his text-book, Lewis (1968, p.11) said that:

> All assume measurement on an ordinal or interval scale and so permit an analysis of variance.

Even apart from the unsatisfactory use of "assume", this statement is dangerous, in that it might lead the unwary and unthinking - and the educational researcher for whom the book was intended - to suppose or believe that any continuous variate, or any variate consisting of scores or counts, can reasonably be analysed in an analysis of variance. Sadly, many researchers do indeed seem to have such a belief, and to perceive analysis of variance as a "black box" for converting almost any variate into publishable tabulations and into statements about "significance". Such uncritical use of analysis of variance can lead to publication of tables such as the following (ICRISAT, 1981, p.61, part of Table 8), where I have substituted code letters for treatment names:-

Treatment	Downy mildew incidence (%) at		
	26 days	47 days	73 days
A	0	0	0.6
B	0	0	0.9
C	0	0.7	1.5
D	0	1.9	3.5
Untreated	63.9	64.3	65.0
SE± [a]	(0.72)	(1.20)	(1.63)

[a] Standard error of mean, based on arcsin transformation of the percentage values.

Here, each standard error is presumably intended to apply to each of the five means above it, and seems to have been calculated from all plots of the experiment. But this is absurd as, at 26 days, there was no mildew on any of the treated plots, and so no variability of incidence on these plots. (The standard errors are also glaringly wrong in having seemingly been obtained by illegitimate back-transformation from standard errors calculated for means of angular-transformed percentages. And the use of the "angular transformation" - mentioned later in this section - is unlikely to have been appropriate even where the means are not zero).

R.A. Fisher's well-known statement (Fisher, 1935, Section 8) that

> Every experiment may be said to exist in order to give the facts a chance of disproving the null hypothesis.

is not above criticism. A factorial experiment can be said to have several null hypotheses (for "main effects" of individual factors, and for "interactions" between factors), not just one. But more fundamentally, as Yates (1964, p.320) pointed out:-

> In many experiments ... it is known that the null hypothesis customarily tested, i.e. that the treatments produce no effects, is certainly untrue; such experiments are in fact undertaken with the different purpose of assessing the magnitude of the effects. Fisher himself was of course well aware of this ...

Estimated means and standard errors are thus the main if not sole aims of many analyses of experimental results. Even when significance tests may seem to be called for, we must remember a point made by Cox (1977, p.61):-

> statistical significance is quite different from scientific significance ... therefore estimation, at least roughly, of the magnitude of effects is in general essential regardless of whether statistically significant departure from the null hypothesis is achieved.

A continuing need to stress the true role of significance tests in experimental work has recently been recognised by authors who have contributed the following to the medical literature: "The inadequacy of statistical significance" (Lutz and Nimmo, 1977); "Significant or important?" (Wade and Waterhouse, 1977); "Significance, importance and equality - three basic concepts in the analysis of a difference" (Taube, 1980).

Overemphasis on significance testing or hypothesis testing has been carried to extremes by workers who seem to believe that "multiple comparison" tests (see, e.g., O'Neill and Wetherill, 1971) must always be applied to treatment means from an experiment. "Duncan's multiple range test" is particularly conspicuous in many journals of agricultural research; indeed Martynov and Krupnov (1980) reported that, of one hundred publications in the Agronomy Journal for 1973, sixty-eight used the Duncan test. Particularly disheartening are two common "glaring abuses" that were roundly condemned by Chew (1976): the use of a multiple range test in disregard of a natural ordering to the treatments and in disregard of a factorial structure to the treatments.

Some theoretical background to the analysis of variance must of course be taught to students of the design and analysis of experiments. But much of the supporting mathematical and conceptual framework in the literature is unnecessary and indeed harmful. For example, as Pearce (1982) wrote in a book-review:-

> one experimenter may be studying four inhibitors at two temperatures, while another has four diets, which he tries out on rats on two occasions. Why regard them together as examples of "two-way classifications" and then go on for pages about model I and model II, fixed and random effects? The first investigator is interested in both inhibitors and temperatures and in the way they interact, the other only in diets, the occasions merely providing confirmation. Their logical problems are unrelated.

A student, faced with data from either of these experiments, ought to be able to use a desk- or pocket-calculator to produce a correct analysis, without conscious reference to an algebraic formulation.

Likewise, a student faced with data from a simple split-plot experiment ought to be able to produce a standard analysis of variance without recourse to cumbersome mathematics. Except for the calculation of some of the standard errors, the analysis of variance for a split-plot design is simple (Preece, 1982a, Section 6.8), and not the perilous matter that it has sometimes been presented as.

Some experiments involve measuring the same attribute on several successive occasions on each plot, the occasions being the same for all plots; each plot receives its treatment just once, before the first measurement, or receives

the same treatment repeatedly or continually. Such an experiment is often said to be a "repeated measurements experiment" or to have a "repeated-measures design", although "successive" would be much better than "repeated". Sometimes the successive occasions for each plot have been likened to sub-plots; indeed Little and Hills (1978, p.137) horrifyingly said that repeated-measurements data "are most properly analyzed as the split-plot design", and Aitkin (1981) that repeated-measurements designs "are called 'split-plot' designs". These statements notwithstanding, the two situations are formally very different: the split-plot design has randomisation within each whole plot, whereas the successive occasions cannot be randomised and have an inherent ordering unparalleled in the split-plot design. In some circumstances, a "split-plot" analysis may perhaps turn out to be a satisfactory approximation for a repeated-measurements analysis. However, the important paper by Rowell and Walters (1976) considered five diverse repeated-measurements experiments for all of which a "split-plot" analysis was manifestly invalid, and suggested that a simple alternative analysis should always be used for "repeated observations"; the alternative consists of separate analyses of different "contrasts over time", each contrast having been evaluated for each plot.

Preliminary scrutiny of a variate may suggest that it does not satisfy the assumptions lying behind standard analyses, but that the assumptions might be satisfied after the variate has been "transformed". Several "transformations" are commonly used, including the "logarithmic", "square root" and "angular". A log transformation, for example, simply involves replacing each variate value x by log x. The angular transformation, useful for <u>some</u> variates consisting of percentages or proportions, involves replacing each proportion p by arcsin $\sqrt{p}$, i.e. by $\frac{1}{2}$ arc-cos (1-2p); this was presumably the transformation used in the ICRISAT analysis mentioned above. Rules-of-thumb for the choice of transformations have their place in our subject (see, e.g., Goulden, 1952, p.98), but should be used critically.

"Covariance analysis" is another important topic within our subject. At its simplest, covariance analysis involves two variates (each with one value per plot), which may be denoted as y, the variate of primary interest, and x, the "covariate" consisting of "concomitant observations" unaffected by the treatments (Cox, 1958, Chapter 4). Often, the x-values are obtained from the plots before the treatments are applied. the supplementary information from x is used with the hope of reducing the experimental error in y, and so as to "adjust" the y treatment means for initial plot differences in x. The analysis is a combination of linear regression analysis and of a standard analysis of variance of y. Covariance analysis can also be used, with care, with an x-variate that <u>is</u> affected by the treatments (Cochran and Cox, 1957, Section 3.88).

Combining results from different experiments within a series is yet another important analysis topic, but is too complex to discuss here. Sadly, standard textbooks give it little attention, except notably those of Cochran and Cox, 1957, Chapter 14) and John and Quenouille (1977, Chapter 13).

COMPONENT VI: Interpreting and Reporting the Results

Finney (1960, pp.13-14) wrote that

> Too often a statistician imagines that his work is complete when he has designed an experiment and has later computed the analysis of variance (or other appropriate procedure) for the results. It cannot be too strongly emphasized that his responsibility extends further, and that he must play his part, of course in close association with the experimenter, in the interpretation of the statistical analysis. A bare statement of an analysis of variance, with tests of significance and standard errors of means, is no more than the scaffolding around which a report is to be constructed... .

Likewise Yates and Healy (1964, p.205) complained of

> reports on experiments such as "the Z is not significant" without any numerical results.

Despite such clear pronouncements from such eminent authorities, the interpretation of statistical analyses has been badly neglected in teaching of the design and analysis of experiments. As I wrote elsewhere (Preece, 1982a, Section 7):-

> Many academics consider it to be beneath their dignity to provide students with guidance on report-writing; a report written by a student is awarded good marks if it is considered to seem to suggest that the student understands his subject, even if the expression is incoherent and almost devoid of syntax. The subject of the Design of analysis of Experiments has suffered accordingly... .

Skill in writing reports on experiments requires a good feel for clear, concise prose, and ability in both tabular and graphical presentation of means and their standard errors. The writer should know how many significant figures to quote in means, regression coefficients and standard errors, and should be no more shy of using "Plain Numbers" than Civil Servants should be of "Plain Words". As I complained before (Preece, 1982a, Section 7):

> Time and again, when the reader wants to know that, say, Magnaquat made the treated plants grow x ± y cm more than untreated plants, he is fobbed off with "The application of Magnaquat caused a marked increase in plant height" (or "pronounced increase", or "considerable increase", or "notable increase").

It is pleasant to be able to note, in conclusion, that good report writing is encourged by part of "Principle 3" of the Council for National Academic Awards (CNAA, 1979, p.10, paragraph 3.5):-

The programme of studies must encourage the skills of clear communication and logical argument... .

Acknowledgement

This paper was written whilst the author was Biometrics Liaison Officer (Field Corps) of the Overseas Development Administration (ODA).

References

[1] Aitkin, M. (1981) Response to Query on "Regression models for repeated measurements". Biometrics, 37, 831-832.

[2] Anscombe, F.J. (1948) The validity of comparative experiments (with Discussion). J.Roy.Statist.Soc., 111, 181-211.

[3] Chew, V. (1976) Comparing treatment means: a compendium. Hort Science, 11, 348-357.

[4] CNAA (1979) Principles and Regulations for the Award of the Council's First Degrees and Diploma of Higher Education. London: Council for National Academic Awards.

[5] Cochran, W.G. and Cox, Gertrude M. (1957) Experimental Designs, 2nd edition. New York: Wiley. [Available with soft covers].

[6] Cox, D.R. (1958) Planning of Experiments. New York: Wiley. [Available with soft covers].

[7] Cox, D.R. (1977) The role of significance tests. Scand.J.Statist., 4, 49-70.

[8] Das, M.N. and Giri, N.C. (1979) Design and Analysis of Experiments. New Delhi: Wiley Eastern

[9] Dyke, G.V. (1974) Comparative Experiments with Field Crops. London: Butterworths.

[10] Finney, D.J. (1955) Experimental Design and its Statistical Basis. Univ. Chicago Press.

[11] Finney, D.J. (1960) An Introduction to the theory of Experimental Design. Univ. Chicago Press. [Available with soft covers].

[12] Fisher, R.A. (1925) Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd.

[13] Fisher, R.A. (1926) The arrangement of field experiments. J.Min.Agric., 33, 503-513.

[14] Fisher, R.A. (1935) The Design of Experiments. Edinburgh: Oliver and Boyd.

[15] Gomez, Kwanchai A. and Gomez, A.A. (1976) Statistical Procedures for Agricultural Research, with Emphasis on Rice. Los Banos, Philipines: Internat. Rice Res. Inst.

[16] Goulden, C.H. (1952) Methods of Statistical Analysis, 2nd edition. New York: Wiley.

[17] ICRISAT (1981) Annual Report 1979/80. Patancheru, India: Internat. Crops Res. Inst. for the Semi-Arid Tropics.

[18] Jeffers, J.N.R. (1978) Statistical Checklist 1: Design of Experiments. Cambridge: Inst. Terrestrial Ecology, Natural Environment Res.Council.

[19] John, J.A. and Quenouille, M.H. (1977) Experiments: Design and Analysis, 2nd edition. London: Griffin.

[20] Lewis, D.G. (1968) Experimental Design in Education. Univ. London Press.

[21] Little, T.M. and Hills, F.J. (1978) Agricultural Experimentation: Design and Analysis. new York: Wiley.

[22] Lutz, W. nd Nimmo, I.A. (1977) The inadequacy of statistical significance. Europ. J. Clinical Investigation, 7, 77-78.

[23] Martynov, S.P. and Krupnov, V.A. (1980) Estimation of the significance of differences between varieties. (In Russian) Tsitologiya i Genetika, 14, 83-87. [English translation: Cytology and Genetics, 14, 73-77].

[24] Mead, R. (1982) Contribution to Discussion of "Regression diagnostics, transformations and constructed variables" by A.C. Atkinson. J.Roy.Statist.Soc. B, 44, 30.

[25] O'Neill, R. and Wetherill, G.B. (1971) The present state of multiple comparison methods (with Discussion). J.Roy.Statist.Soc. B, 33, 218-250.

[26] Pearce, S.C. (1982) Review of "Biometry: The Principles and Practice of Statistics in Biological Research, 2nd edition" by R.R. Sokal and J.F. Rohlf, and of "Statistical Tables, 2nd edition" by J.F. Rohlf and R.R. Sokal. Times Higher Ed. supp., No.487 (March 5), 17.

[27] Pocock, S.J. (1982) Statistical aspects of clinical trial design. Statistician, 31, 1-18.

[28] Preece, D.A. (1982a) The design and analysis of experiments: What has gone wrong? Utilitas Mathematica (to appear)

[29] Preece, D.A. (1982b) La biometrie; pas rites mais science. Proc. 11th Internat. Biometric Conf. (to appear)

[30] Raktoe, B.L., Hedayat, A. and Federer, W.T. (1981) Factorial Designs. New York: Wiley.

[31] Rowell, J.G. and Walters, D.E. (1976) Analysing data with repeated observations on each experimental unit. J.Agric.Sci., Cambridge, 87, 423-432.

[32] Taube, A. (1980) Significance, importance and equality - three basic concepts in the analysis of a difference. Upsala J.Med.Sci., 85, 97-102.

[33] Wade, O.L. and Waterhouse, J.A.H. (1977) Significant or important? Brit.J.Clin.Pharmacol., 4, 411-412.

[34] Yates, F. (1937) The Design and Analysis of Factorial Experiments. Tech. Commun. No.35 of the Imperial Bureau of Soil Science (subsequently Commonwealth Bureau of Soils). Harpenden: Imperial Bureau of Soil Science (now Farnham Royal: Commonwealth Agricultural Bureaux).

[35] Yates, F. (1964) Sir Ronald Fisher and the design of experiments. Biometrics, 20, 307-321.

[36] Yates, F. and Healy, M.J.R. (1964) How should we reform the teaching of statistics? J.Roy.Statist.Soc. A, 127, 199-210.

TEACHING EXPERIMENTAL DESIGN IN THE COMPUTER AGE

R Mead, University of Reading

INTRODUCTION

The basic principles of experimental design, as it is usually taught today, both to students specialising in statistics and to those in other disciplines, were established nearly fifty years ago. For many statisticians and users of statistics, the major textbook on Design is still Cochran and Cox (1957), written before the advent of the computer. During the last twenty-five years, our computational habits have changed radically, and it is important to ask whether the fundamental ideas of experimental design remain unchanged, or whether these ideas were influenced by the computational environment in which they were developed.

Do we know the answers to the computer-orientated experimenter who says "There is no need to plan the allocation of treatments to experimental units since we can use the programs for general linear models to sort out the effects of treatments from those of the various environmental factors which we can recognise most effectively after the experiment has begun?" Or the more subtle version, which says "Just try out any treatment allocation on the computer and if the standard errors look OK go ahead and use that allocation."

I shall not be talking about design by computer, but about the teaching of design liberated by the computer from restrictions imposed by the need to analyse the data. I shall attempt to consider not only the teaching of design now, but into the future.

THE GENERAL PRINCIPLES OF DESIGN AND ANALYSIS

Much of what passes for the statistical theory of experimental design is really mathematics and not statistics. The matrix theory of the general linear model in terms of generalised inverses involves no statistical ideas, no implications for the practical collection and interpretation of quantitative data. The extensive theory of partially balanced designs is really only a branch of mathematics. Nevertheless, there are statistical principles which have become submerged in both the mathematics and the details of standard designs, and these principles are what should be taught.

The fundamental principle for the analysis of experimental data is that of fitting sets of parameters, representing the effects of applied treatments, added sequentially to a basic model representing the environmental variation between experimental

units. The interpretation of results starts from the analysis of variance (or deviance) which shows the division of the variation between observations into component sums of squares. The most important initial information for interpretation of results is the extra sum of squares due to fitting the latest terms added to the model, combined with the variance-covariance matrix of parameter estimates.

The model and its implications are exemplified most clearly by computer programs, most notably GLIM, which simply allow the addition of sets of terms to the model, calculating the extra sum of squares and providing the variance-covariance matrix. In addition, GLIM recognises the number of distinguishable contrasts it is being required to estimate. The use of GLIM, or some similar program, as a component of an advanced course on Experimental Design is, I believe, essential in the teaching of experimental design. Sometimes GLIM can provide worrying results because of its coding of parameter information, but I have found that the sorting out of apparent paradoxes invariably leads to greater insight into experimental design ideas. One example of the potential of GLIM to worry the innocent user is the analysis of the Latin Square with a missing row shown in Figure 1. If the model

$$Y_{ij} = t_{k(ij)} + r_i + c_j + \varepsilon_{ij}$$

is fitted, the terms in the model being written in the order of fitting then the standard errors of the five treatment effect estimates given by GLIM, $\hat{t}_k$ are not all the same. The reason for fitting the treatment effects first was to obtain distinct estimates for all five treatment means, by avoiding constraints on the t_k and gives a clue as to why an obviously balanced design appears unbalanced. A further clue is that the treatment effect for which the standard error is different from the other four is treatment B.

A	D	B	C	E
C	E	A	D	B
D	B	E	A	C
E	A	C	B	D

Figure 1: Latin Square with a missing row

The ideas of statistical design which I believe to be of primary importance are:

(i) Replication
(ii) Blocking
(iii) Randomisation
(iv) Orthogonality
(v) Factorial Treatment Structure
(vi) Confounding
(vii) Covariance
(viii) Quantitative Treatment Factors

Of these, I shall not consider in this paper the ideas of replication, randomisation and factorial structure. This does not imply any lack of devotion to the ideas, each of which figure prominently in any course on experimental design which I teach, but that the presentation of their ideas is not, in my view, essentially altered by the advent of computers.

BLOCKING

The development of the idea of blocking in experimental design depended heavily on the requirement of being able to analyse the results from a blocked experiment. This meant that the Least Squares equations must have a considerable degree of pattern. The sequence of introduction of new design ideas displays a gradual move away from the very simplest form of least squares equations (orthogonality) through balance and on to various patterns of partial balance. The designs thus developed were:

(a) Randomised Block Design
(b) Balanced Incomplete Block Design
(c) Latin Square Designs
(d) Graeco-Latin Square Designs
(e) Partially balanced incomplete block designs
  (i) Lattice Designs
  (ii) Cyclic Designs

It is important to realise that the requirement, per se, of a simple solution for least squares equations is no longer valid. Therefore, the clever tricks for solving sets of equations, represented by the various forms of partial balance, are no longer needed. The relatively simple PBIBD's must therefore be justified as superior to other incomplete block designs in terms of precision of results, not of mathematical nicety.

The importance of orthogonality can now be reassessed. It is that the different sets of effects in the model can be assessed independently without the implicit ambiguity of the non-orthogonal design. The separation of treatment and block effects in the randomised block design is an ideal. However, it is an ideal which in some forms of experiment, for example, clinical trials (Pocock

and Simon, 1975) is extremely difficult to achieve. And it is an ideal that is easily disturbed by missing observations. Therefore, while accepting it as an ideal, we must consider how to design experiments when the ideal is unachievable.

In teaching experimental design at all levels, I believe in starting with a problem in which the ideal, orthogonal solution is impossible. The exercise, TOMATO, set to students, and which will be used and discussed in the workshop associated with this invited paper session is described by Mead, Pollock and Ross-Parker (1979). Briefly, a set of twelve glasshouse compartments in two natural groups of six compartments each is available for the experiment. An experimental treatment is necessarily applied to a complete compartment so that there are twelve available experimental units in two blocks of six. The problem presented to the students is to find the best treatment combination from:

2 levels of Heating,
2 levels of Lighting,
2 varieties of tomato.

The problem then is to compare eight treatment combinations in two blocks of six experimental units each.

For students already indoctrinated with the ideas of Randomised Blocks or Confounded Designs, this can be a very worrying problem. Many such students attempt to decide which two treatment combinations to omit so that they can use a Randomised Block Design. However, the logic of choosing a design is very straightforward. I shall leave detailed discussion to the workshop session, but with eight treatments in two blocks of six, it is inevitable that four treatments will be replicated and four not. The replication provides the basis for comparing all eight treatments, and the analysis follows simply in terms of the comparisons within each block.

Having, hopefully, disturbed the presumptions of the students, I then get them to 'solve' in class a sequence of more or less complex design problems. The purpose of this process is to set up the need for some principles for the choice of particular design allocation patterns. The problems can be solved as easily by non-specialist students as by specialist students though only the specialists can follow the mathematics of the theory, the study of which is a natural consequence of the attempts to solve the problems. The ideas are presented in a very intuitive way. The intuition which I seek to develop is an understanding of the statistical theory of experimental design. The mathematics provides further justification if it is needed.

The sequence of problems, all for the comparison of six treatments, all comparisons being assumed equally important, is as follows:

(i) Four blocks of six units each.

Obviously a simple problem to start with, once the students have realised there are no tricks.

(ii) Six blocks of five units each.

Symmetry leads very rapidly to the solution of omitting one treatment in each block, each treatment being omitted from just one block.

(iii) Six blocks of four units each.

It becomes apparent fairly rapidly that there is no satisfying symmetric solution. Two treatments must be omitted from each block. It can readily be seen that there are more than six possible pairs. The students rapidly arrive at a solution of the form shown in Figure 2, although they may get diverted by attempting to write the treatments into the blocks in the order ABCD/EFAB/CDEF/ ABCD/EFAB/CDEF/

(iv) Ten blocks of three units each.

Once they have appreciated that the idea of pairs of treatments occurring together in a block is crucial, and this comes from TOMATO and the solution to problems (ii) and (iii), students come to appreciate that a design with all pairs of treatments occurring once together in a block is possible. And that the construction is almost inevitable.

(v) Six blocks of seven units each.

Again, symmetry rules.

(vi) Three blocks of five units each.

Four blocks of four units each.

Three blocks of three units each.

This, although not solvable in terms of providing equality of precision for all comparisons, is seen to be a genuine problem and to motivate discussion of principles.

For students with the necessary mathematical background, it is now appropriate to consider the least squares equations for a general block-treatment design, and to derive the result due to Pearce (1964) for balance of treatment comparisons. If n_{ij} = the number of observations for treatment j in block i, then a sufficient condition for the equal precision of estimates of treatment differences, for all pairs of treatments, is that the sum of weighted concurrences,

$$\sum_i (n_{ij} n_{ij'} / (\sum_j n_{ij})),$$

should be invariant over all treatment pairs (j, j′).

Now although this condition for the balance of all treatment comparisons is clear, it is not clear what criterion should be used in cases when it is not

Block 1	A	B	C	D
Block 2	B	C	D	E
Block 3	C	D	E	F
Block 4	D	E	F	A
Block 5	E	F	A	B
Block 6	F	A	B	C

Figure 2: Six Treatments compared in 6 blocks of 4 units each

possible for the sum of weighted concurrences to be the same for all treatment pairs. Studies of the precision of treatment comparisons in unbalanced designs appear to suggest that the variances of treatment differences are as closely related to the unweighted sum of concurrences, $\sum_i n_{ij} n_{ij'}$, as they are to the weighted sum of concurrences. Further research is needed here.

The inequality of precision in some of the examples should lead to a discussion of why equality of precision should be regarded as an ideal. When the treatment comparisons are markedly unequally precise, then the simplest reliable criterion appears to be the unweighted sum of concurrences. By discussing a range of examples, the effectiveness of this criterion can be demonstrated.

To summarise the ideas I believe should be taught about blocking:

(i) The relevance of orthogonality is changed though still important for avoiding ambiguity,
(ii) Balance and near balance,
(iii) Simple properties of the classical BIB designs,
(iv) The practical desirability of unequal precision.

Particular designs which should be discussed include the Randomised Block Design, Balanced Incomplete Block Design, Latin and possibly Graeco-Latin Squares. Designs which should not be taught include Partially Incomplete Block Designs (including Lattice and Cyclic Designs) and Completely Orthogonal Squares.

To complete this section, I shall use a problem I met recently in an advisory consultation to illustrate the kind of thinking that I believe we should be aiming at in our teaching of blocking. The experimenter wanted to compare 5 diets for cattle, and he had available 9 animals for four successive time periods. The experimental units can be represented as in Figure 3(a). When he arrived for the consultation, the experimenter had convinced himself that he would have to use only 4 treatments in two 4 x 4 Latin Squares, ignoring one animal.

		Animals								
		1	2	3	4	5	6	7	8	9
Periods	I	X	X	X	X	X	X	X	X	X
	II	X	X	X	X	X	X	X	X	X
	III	X	X	X	X	X	X	X	X	X
	IV	X	X	X	X	X	X	X	X	X

(a) The structure of units

A	D	B	C	E	C	E	B	A
C	E	A	D	B	E	D	C	B
D	B	E	A	C	A	C	E	D
E	A	C	B	D	B	A	D	C

(b) The proposed design

Figure 3: Comparison of five diets using 4 period for 9 animals

The first amendment to this was to split the 9 animals into a group of 5 and a group of 4. A 4 x 4 Latin Square could be used for the four more important treatments using the group of 4; and a 5 x 5 Latin Square with a missing row, for all five treatments using the group of 5. This is clearly more efficient than the original design and there is no way of getting more information out of the group of 5. But the group of 4 animals could be used to compare all 5 treatments by using a Latin Square with one row and one column missing. One treatment will now be more precisely estimated than the other four, and there may be problems in choosing which treatment that is to be. However, in comparing five treatments, the symmetry or near symmetry of the design ensures its optimality. The resulting design is shown in Figure 3(b). The supremacy of the intermediate design is clearly arguable, there being advantages for each.

COVARIANCE

The primary importance of covariance in the teaching of design has been through its relationship to blocking. The need in many practical situations to choose between several different blocking criteria leads naturally to consideration of how to allow for variation of factors ignored in the blocking structure. The advantages of using quantitative factors as covariates with qualitative

factors for blocking is just one consideration that should be discussed. The problem of the use of covariates is probably particularly important in clinical trials, but does also have importance in other disciplines, such as animal production and psychology. A thorough understanding of the principles and practice of covariance analysis is therefore important in teaching experimental design.

A second, and probably even more important, reason for considering covariance analysis in field crop experiments where the plots are spatially arranged is to introduce the idea of adjustment by covariance for local fertility effects. The method of Papadakis (Bartlett, 1978) has been much discussed recently and must be investigated further before it begins to have much impact on experimental design. However, this is another change in analysis which is encouraged by the easy availability of computers and its general implications for design should not be ignored.

One of the implications of the possible use of Papadakis-style adjustments is that, instead of thinking solely in terms of groups of reasonably homogeneous plots, we should think rather about pairs of adjacent plots. An extreme form would involve considering as 'blocks' all pairs of adjacent plots so that each plot will occur, usually, in four blocks. There are many opportunities here for research, but we should not lose sight of the need to introduce ideas based on Papadakis adjustments into our courses on experimental design in the not-too-distant future.

QUANTITATIVE FACTOR LEVELS

Traditionally, the investigation of the effects of varying the level of a quantitative factor has involved the use of three equally spaced levels. The choice of equal spacing implies the use of a quadratic response function for which the equal spacing gives optimal estimation of the position of any optimum level. The use of just three levels is also optimal unless a response function with more than three parameters is to be fitted, when it is obviously inappropriate.

The computer allows us to use, rather more easily, response functions other than the simple polynomials. We can fit many different response functions and we can also consider designs to discriminate between alternative response functions.

Not only does the computer allow us to consider a wide range of response functions, but it enables us, through approximate results or simulations, to consider optimality of designs for various criteria for any particular response function. As yet, the amount of research on the optimal designs for practically relevant criteria is rather limited, although there is a considerable amount of general theory which indicates clearly the advantages for estimation of using

the minimum number of levels. There is a need for much more research on optimal or near-optimal designs for alternative criteria, so that some general patterns relevant to the practical design of experiments can be discerned and taught. Until more information is available, the scope for teaching about the choice of quantitative factors levels in experimental design is limited, but I suggest that this should not provide an excuse for ignoring the subject in our teaching. We can show how computers can already provide the information about optimal design for any particular practical problem and, by teaching the rather sparse information currently available, we can perhaps motivate academic statisticians to find the answers to real problems.

CONFOUNDING

Experimenters believe that confounding is a form of black magic. Textbooks, and sometimes even statisticians, reinforce this view. Most computer programs make the analysis of data from confounded experiments easy, since there is no real difficulty in such analyses; but also it is possible to check, using a computer program such as GLIM, the extent to which effects are confounded in quite complex confounded experiments in which the number of levels per factor varies. I think it is therefore time for a completely new look at how we teach confounding. The motivation for this reconsideration is only partially provided by computers, but the almost total mistrust of confounding by those who should be using confounded experiments provides ample justification.

I believe that confounding is usually taught from quite the wrong viewpoint. Instead of considering primarily the negative question of 'which effects should we confound?', we should ask the more positive question, 'which effects will be unconfounded and therefore available?'.

I suggest that the teaching of confounded experiments should start with the consideration of a complex but practical problem. To illustrate this approach, consider a problem in intercropping experimentation on how to grow a mixture of maize and cowpea crops. The initial experiment consists of three randomised blocks of twelve experimental treatments. The treatments are all combinations of three densities of cowpea (D), two cowpea varieties (C) and two maize varieties (M). The experimenter would ideally also like to include three nitrogen levels (N), in his experiment. Quite apart from the difficulty of having no explicit replication, he would not consider this addition seriously since it would involve the abandonment of blocks, or the division of treatment combinations into blocks, which he knows means confounding. Also, most textbooks do not include a confounded 3 x 3 x 2 x 2 design.

However, consider the problem of adding the nitrogen levels to the initial randomised blocks. To avoid the nitrogen effects being confounded with blocks, we must try to have each of the three nitrogen levels occur four times in each block. We must also consider the allocation of nitrogen levels to the twelve combinations in each block. If we ensure that each nitrogen level occurs with each combination of cowpea and maize varieties, then the interactions NC, NM and NCM will also be unconfounded. The nine combinations of the density and nitrogen factors will necessarily be at least partly confounded with differences between blocks.

We have so far talked only about what we would ideally like. In practice, however, there is no difficulty in allocating the nitrogen levels so as to allow unconfounded estimation of all the effects we hoped for. The resulting design is shown in Figure 4. The construction can be traced by considering the restrictions on the placement of N levels in the order of the observations of block 1 and subsequently of block 2, which in fact requires only a cyclic rotation. The restrictions only start to apply from the fifth combination, except that in the absence of other restrictions, we should try as far as possible to have even replication of the D x N combinations.

The design of Figure 4 is, I believe, a sensible practical experimental design, of a type which should be used far more frequently. This form of design should also be taught at all levels, both specialist and non-specialist. If we consider again the impact of the computer on the analysis and design of experiments, then we can observe two benefits. First, a program such as GLIM is able, if the experimenter is in doubt about the properties of his design, to demonstrate which effects are unconfounded. This should, if confounded designs are properly understood, provide more than a checking facility. The second advantage is related to the ideas discussed in the section on blocking. If we consider the D x N interaction, then we have nine treatment combinations displayed in Figure 5(a), which are allocated to blocks following the design shown in Figure 4 as shown in Figure 5(b). Now, if we had been required to allocate nine treatments to three blocks of twelve plots each, then we should have inexorably come to an allocation with each block containing all nine treatments, plus three duplicated treatments, the three duplicated treatments being different in each block - exactly as in Figure 5.

The computer enables the D x N interactions effects to be estimated. These effects are not orthogonal to blocks, though they are partially balanced (a fact which is quite unnecessary to the appreciation of the advantages of the design). The description of the status of the D x N interaction is not clear from the literature. It is clearly not completely confounded, it is certainly

Block I	Block II	Block III
$d_1\ c_1\ m_1\ n_0$	$d_1\ c_1\ m_1\ n_1$	$d_1\ c_1\ m_1\ n_2$
$d_1\ c_1\ m_2\ n_1$	$d_1\ c_1\ m_2\ n_2$	$d_1\ c_1\ m_2\ n_0$
$d_1\ c_2\ m_1\ n_2$	$d_1\ c_2\ m_1\ n_0$	$d_1\ c_2\ m_1\ n_1$
$d_1\ c_2\ m_2\ n_0$	$d_1\ c_2\ m_2\ n_1$	$d_1\ c_2\ m_2\ n_2$
$d_2\ c_1\ m_1\ n_1$	$d_2\ c_1\ m_1\ n_2$	$d_2\ c_1\ m_1\ n_0$
$d_2\ c_1\ m_2\ n_2$	$d_2\ c_1\ m_2\ n_0$	$d_2\ c_1\ m_2\ n_1$
$d_2\ c_2\ m_1\ n_0$	$d_2\ c_2\ m_1\ n_1$	$d_2\ c_2\ m_1\ n_2$
$d_2\ c_2\ m_2\ n_1$	$d_2\ c_2\ m_2\ n_2$	$d_2\ c_2\ m_2\ n_0$
$d_3\ c_1\ m_1\ n_2$	$d_3\ c_1\ m_1\ n_0$	$d_3\ c_1\ m_1\ n_1$
$d_3\ c_1\ m_2\ n_0$	$d_3\ c_1\ m_2\ n_1$	$d_3\ c_1\ m_2\ n_2$
$d_3\ c_2\ m_1\ n_1$	$d_3\ c_2\ m_1\ n_2$	$d_3\ c_2\ m_2\ n_0$
$d_3\ c_2\ m_2\ n_2$	$d_3\ c_2\ m_2\ n_0$	$d_3\ c_2\ m_2\ n_1$

Figure 4: One replicate of a 3 x 3 x 2 x 2 experiment in 3 randomised blocks

	d_1	d_2	d_3
n_0	1	2	3
n_1	4	5	6
n_2	7	8	9

(a) Coding of the nine treatment combinations from N x D

Block I	Block II	Block III
1	4	7
4	7	1
7	1	4
1	4	7
5	8	2
8	2	5
2	5	8
5	8	2
9	3	6
3	6	9
6	9	3
9	3	6

(b) Allocation of N x D combinations to blocks (from Figure 4)

Figure 5: the non-orthogonal arrangement of the ND Interaction in blocks

not unconfounded, and the term 'partially confounded' is otherwise defined. The natural description would seem to be 'partly confounded'.

CONCLUSION

This has been a very personal view of the teaching of experimental design. Doubtless some of the ideas expressed herein will be felt to be too extreme; possibly some will be regarded as too traditional. My primary concern, however, is to dispel the view of experimental design as a rather staid aspect of statistics, where a range of standard designs exist, all of which can be looked up in the appropriate index in Cochran and Cox, or successors to that worthy volume. This view is only justifiable through a wilful ignorance of the impact of computers. The subject of experimental design is alive and exciting, and there is no reason why it should not be so taught at all levels of teaching.

References

[1] Cochran, W.G. and Cox, G.M. (1957) Experimental Designs, 2nd edition, New York, Wiley.

[2] Mead, R., Pollock, K.H. and Ross-Parker, H.M. (1979) A sequence of games useful in teaching experimental design to agricultural students. American Statistician, 33, 70-76.

[3] Pearce, S.C. (1964) Experimenting with blocks of natural size. Biometrics, 20, 699-706.

[4] Pocock, S.J. and Simon, R. (1975) Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics, 31, 103-115.

TEACHING CONCEPTS OF DESIGN AND ANALYSIS USING REAL AND SIMULATED EXPERIMENTS

J R Elliott, Wilfrid Laurier University and J C Robinson, University of Waterloo

INTRODUCTION

Courses in experimental design usually provide students with a set of possible designs and corresponding methods of analysis but fail to offer them opportunities for using these new "tools" to solve actual problems. This situation is unfortunate since some of these students may eventually be required to give advice to researchers planning experiments without ever having experienced the joys, frustrations, or compromises involved in conducting an experiment. In fact, with the advent of co-operative, internship, and work-study programs, the student is being requested to give advice even sooner than in the past. How can these students, who have never been confronted by time or cost constraints, choice of appropriate factors, or design of follow-up experiments fully appreciate the problems their clients face?

With the recent development of more powerful and more accessible computer hardware and software, this situation is gradually changing. Several computer-oriented texts are now available containing large sets of "real" data (Afifi and Azen, 1972; Draper and Smith, 1981; Gunst and Mason, 1980). These data are typically used not only for examples and exercises but also to provide a unifying theme for the analytical procedures discussed in the text. However, this approach has not as yet permeated the field of experimental design where the use of small "cooked" data sets is still the norm, even in applied courses. Moreover, this "large data-set" approach is most useful for illustrating the analytical techniques available and tends to ignore the strategies involved in data collection.

The latter, although central to any experimental design or sampling course, are often given a superficial treatment. They are discussed and illustrated, but as noted above few students receive practical experience in their use. W.G. Hunter has suggested, in The American Statistician, two possible remedies; namely, providing the students with computer routines for generating data from simulated experiments or having them obtain "real" data by formulating and performing an actual experiment (Hunter, 1977). Emphasizing the latter approach, he has described a number of simple factorial experiments conceived and conducted by students at the University of Wisconsin. While favouring this method, Hunter noted that, although the total number of possible experiments was large, the range of practical experiments of this type was rather limited.

A more recent article by Pollock et al concentrated on data from simulated experiments and described a sequence of "games" developed at the University of Reading which could be conducted either on the computer or in the classroom (Pollock et al, 1979). These simulations - TOMATO, CHICK and SELECT - were used in teaching design to agriculture students and provided an opportunity for these students to experience firsthand the effects of concepts such as randomization, replication, interaction and blocking on the data-gathering process.

In fact, both approaches have considerable merit as teaching aids in design courses and we too have used both in attempting to make statistical concepts more understandable. Although the mini-research project approach more closely approximates reality, it does have the all-too realistic disadvantage of failing to provide useful results on occasion. Primarily for this reason, we have favoured the use of simulated experiments and have attempted to counter the absence of problem formulation by encompassing a wider variety of design types than could be achieved with real experiments. This paper describes in detail the computer package which we have developed for this purpose and outlines its use as a teaching aid. Following more than ten years of combined experience with this package, we believe that it is a unique, extremely versatile, and powerful tool not only for use in experimental design courses, but in regression, sampling, multivariate, and introductory courses as well.

EXPERIMENT SIMULATION PACKAGE

The Experiment Simulation Package or ESP consists of a series of interactive programs or functions, written in the conversational APL language, and currently available for Honeywell Model 6000, DP58 or level 66 computers with a GCOS3 or GCOS8 operating system, or for Xerox Model Sigma 7 computers with a CP-V version COO (or greater) operating system* (Appendix II). The functions model the carbon monoxide, tar, nicotine and hydrogen cyanide deliveries of cigarettes manufactured by the hypothetical "Ruthless Tobacco Company". The yields of these four combustion products are presumably influenced by six factors; namely, amount of chemical additive, cigarette diameter, filter efficiency, filter overwrap, paper porosity, and cigarette weight (Table 1). (For example, see Baskevitch, 1981; Resnik et al, 1978; Robinson and Forbes, 1975). A different response surface is generated for each of the four combustion products, with carbon monoxide, tar, and nicotine deliveries being defined directly in terms of the six factors, while hydrogen cyanide delivery is a simple quadratic function of both carbon monoxide and tar deliveries (Table 2). The shapes of the curves are different, particularly with respect to the numbers

* These programs could be easily transferred to SUPERPET or IBM computers as well, and conversion from APL to FORTRAN should be possible.

TABLE 1: FACTORS OF INTEREST AND ASSOCIATED RANGES

A: CHEMICAL ADDITIVE - a substance added to the tobacco in small amounts which influences deliveries by altering the burning rate.

RANGE: 0 to 0.8mg per cigarette

D: DIAMETER - interacts with weight through density and influences deliveries by diluting the smoke and/or altering combustion.

RANGE: 0.73 to 0.78cm

E: FILTER EFFICIENCY - the material comprising the filter influences deliveries by differential removal of varying amounts of solid particles from the smoke and is available in only four types

TYPES: 15%, 30%, 45%, 60% efficient

O: FILTER OVERWRAP - the cork covering the filter is available in only three porosities and influences deliveries by diluting the smoke

TYPES: Standard, Porous, Very Porous
(Coded: 1, 2, 3 respectively).

P: PAPER TYPE AND POROSITY - the paper covering the tobacco can be treated either with citrate or phosphorus salts and is available in three porosities (the time in seconds required to draw 35ml of air through one square centimeter of paper). It influences deliveries by diluting the smoke since the burning rate increases as the paper becomes more porous.

TYPES: Citrate paper - 10, 20, 30 sec porosity
Phosphorus paper - 10, 20, 30 sec porosity
(Coded 1, 2, 3, 4, 5, 6 respectively)

W: WEIGHT - interacts with diameter through density and influences deliveries by diluting the smoke and/or altering combustion.

RANGE: 0.97 to 1.02gm per cigarette

and types of interactions present. (For example, additive and porosity interact for both tar and carbon monoxide, but not for nicotine). Some of the factors have little or no effect on one response, but have a major influence on others. (For example, increasing filter efficiency has no effect on carbon monoxide delivery, but produces a moderate decrease in nicotine delivery and a massive drop in tar delivery). Consequently, the set of levels for the six factors which minimizes a particular response differs slightly from that which optimizes other responses. This feature allows the instructor to focus on one response in class or on the assignments without providing too much information on the optimal solution for the remaining responses. There is, however, enough similarity in the four responses so that predictions can be made, to some extent, for one response based on results from another; that is, the responses are positively correlated through a similar reaction to changes in the levels of various factors such as diameter and weight (Dalhamn, 1975).

In addition to these deterministic components, all four responses are subject to several possible sources of variation or "noise" due to environmental effects. These consist of a pure error component, a random machine run effect, a fixed port effect, and a fixed cigarette lot effect (Table 2). The random error and run components consist of normally distributed variables generated using Marsaglia's polar modification of the Box-Muller method (Box and Muller, 1958; Marsaglia and Bray, 1964), whereas the fixed port and cigarette lot components have U-shaped distributions with the maximum effect being at the extremes. The instructor has control over the magnitude of each component and may easily suppress or implement any combination of these sources of variation. In fact, all variation may be suppressed so that the effect of any combination of levels of the six treatment factors can be determined. This is particularly useful for examining the true effect of supposed "optima" obtained in student projects.

Using the program PARAMETERS, the instructor determines the factor or factors which must be specified by the student, imposes restrictions on the size of the experiment, introduces the error components, determines the responses that will be simulated, and establishes a due data (Figure 1(a)). This program, together with the functions RESOURCES, SIMULATE and DESGEN, and the built-in APL function SETLINK, is permanently stored in a protected workspace currently called ESP..ICOTS. All functions are locked for further protection. Through an unprotected workspace called EXPERIMENT, the student is given access to the locked programs RESOURCES, EXPERIMENT and END. The RESOURCES program informs the student of all restrictions imposed on the experiment by the instructor except for the composition of the error component (Figure 1(b)). The EXPERIMENT program copies the main routine, SIMULATE, from the protected

workspace, while the END program erases SIMULATE and saves the responses for the student when the simulation is completed. The heart of ESP is the SIMULATE program which performs the actual simulation of the responses according to the conditions set by the PARAMETERS program. SIMULATE requires as input a design matrix which is usually generated by the student prior to running the simulation (Figure 2(a)). However, if such a matrix is unavailable at the time the SIMULATE program is being executed, then the subroutine DESGEN is retrieved from the protected workspace to aid the student in preparing the proper design matrix. In either case, the SIMULATE program conducts extensive checks of the design matrix for unacceptable, missing, or out-of-range factor levels before proceeding with the simulation. It also denies access to any student attempting a simulation after the due date. A typical terminal output from the SIMULATE program is given in Figure 2(b) corresponding to the design matrix in Figure 2(a). Note that a total cost is printed for the experiment, together with production and testing costs associated with each type of cigarette or treatment combination so that an additional restriction can be imposed on the size of the experiment through budget constraints. Detailed cost information is provided in Appendix I.

USE OF ESP

ESP has been used as a teaching aid in introductory statistics courses, as well as in the more advanced design and regression courses. In the introductory course, or initially in the more advanced course, the levels of all but one factor are fixed, a random error component is introduced and one of the responses, usually carbon monoxide, is selected for study. Students are required to simulate carbon monoxide deliveries for one or more levels of the factor and simple test statistics are calculated by the students based on their individual responses. Familiarity with the use of ESP, and more generally with the entire APL language, is quickly gained in this fashion. (Some lecture time is devoted to teaching APL which is not a prerequisite for the courses). As the design course progresses, first port and then run effects are introduced so that the student learns to use randomization and blocking to minimize or eliminate such sources of bias. By now, the students have been given assignments, both on the design and on the analysis of completely randomized, randomized block, and Latin square experiments. APL programs* for the analysis of data from such experiments are also provided in the EXPERIMENT workspace).

Factorial designs are introduced and the carbon monoxide deliveries are examined when first two and, finally, all six factors are varied simultaneously. The 2^n and 3^n series designs, including the concepts of confounding and fractional

* Program listings are available from the authors.

TABLE 2: RESPONSE SURFACE MODELS

RESPONSE	MATHEMATICAL MODEL
CARBON MONOXIDE (grams)	$CO = 10 + 4(A - C[P])^2 + (3 - O)^2 + C[P](3 - (O - 2)^2) + 120(1.77 - 0.2C[P] - W/D^2)^2 + \text{Error}^*$ where $C[P] = 0.1, 0.35, 0.6, 0, 0.25, 0.4$
TAR (milligrams)	$Tar = (1 - E)\{12.5 + 5(0.64 - (A - T[P])^2) + 2T[P] + O.5(O-3)^2(1 + D^2/W) + 20(1.77 - 0.2T[P] - W/D^2)^2\} + \text{Error}^*$ where $T[P] = 0.15, 0.2, 0.5, 0, 0.1, 0.6$
NICOTINE (milligrams)	$Nic = \sqrt{1 - E}\{1.028 + 0.25N[P] + 0.1(O - 3)^2(1 + D^2/W) + 6(1.77 - 02N[P] - W/D^2)^2\} + \text{Error}^*$ where $N[P] = 0, 0.15, 0.4, 0, 0.15, 0.4$
HYDROGEN CYANIDE (micrograms)	$HCN = \left(\frac{CO + 2Tar}{4}\right)^2$

* The error component is determined by multiplying the average value of the response for the particular experiment by the sum of the pure error, machine run, machine port and cigarette lot effects, if any, appropriate to the cigarette being tested. The pure error and machine run effects are obtained by multiplying randomly generated standard normal values by the corresponding coefficients of variation specified through the PARAMETERS program. The port effect is obtained by multiplying the maximum port error specified through the PARAMETERS program by $\left(\frac{2X_p - N_p - 1}{N_p}\right)^4$ where X_p is the number of the port at which the cigarette is being tested and N_p is the total number of ports available (i.e. the "size" of the smoking machine). This simulates a machine with a single centrally-located suction tube and imposes the maximum variability on ports at each end of the machine (i.e., for $X_p = 1, 2, 3$ or $N_p - 2, N_p - 1, N_p$). The cigarette lot effect is determined in a similar manner except that the 10th power is used so that the range of unaffected or marginally affected lots is substantially larger relative to that for ports.

TABLE 2: RESPONSE SURFACE MODELS - continued...

MINIMUM	FACTOR LEVELS AT MINIMUM A	D	E	O	P	W
10g	0	§	†	VP	P10	§
5mg	0.8	§	0.6	VP	P10	§
0.650mg	†	§	0.6	VP	C10 or P10	§
31.8µg	†	§	0.6	VP	P10	§

NOTES:

§ All responses are related to diameter and weight through density, W/D^2, and are minimized for a density of 1.77. This produces an infinite number of optimal solutions and requires the final solution to be determined on the basis of manufacturing costs.

† Carbon monoxide is unaffected by filter efficiency. Cost considerations would dictate 0.15 or 15% as the appropriate choice. Similarly, nicotine is unaffected by additive and attains its minimum for both 10 second porosity citrate and phosphorus papers. Again, cost considerations would suggest 10 second citrate paper and no additive as the most appropriate choices. Hydrogen cyanide is also independent of additive so that no additive would produce the least expensive optimal cigarette.

FIGURE 1: SAMPLE OUTPUTS FROM PARAMETERS AND RESOURCES PROGRAMS

A: OUTPUT FROM PARAMETERS PROGRAM

```
            PARAMETER RESTRICTIONS AND FACTOR SPECIFICATIONS
            ************************************************

EFFECT OF ADDITIVE,DIAMETER,OVERWRAP,POROSITY AND WEIGHT ON CO DELIVERY

RESTRICTIONS ON SIZE OF EXPERIMENT:      {VARIABLE "SIZERES"}
************************************

                        OBS/RUN      RUNS
MAXIMUM NUMBER:           20          10

SPECIFICATIONS FOR DESIGN MATRIX:      {VARIABLES "FACLEV","FACSTUD"}
*********************************

FACTORS WHOSE LEVELS MUST BE SPECIFIED BY STUDENTS:

    ADDITIVE    DIAMETER    OVERWRAP    POROSITY    WEIGHT
     PLUS PORT NUMBERS BETWEEN 1 AND 20 INCLUSIVE

FACTORS WHOSE LEVELS HAVE BEEN PRE"SPECIFIED:

  EFFICIENCY:   0.6

SPECIFICATIONS FOR ERROR AND BIAS TERMS:    {VARIABLE "ERRSPEC"}
***************************************

EXPERIMENTAL ERROR:   COEFFICIENT OF VARIATION =  10%

MACHINE RUNS:   COEFFICIENT OF VARIATION =  5%

MACHINE PORTS:   MAXIMUM EFFECT =  1%

CIGARETTE LOTS:   NO EFFECT

RESPONSES TO BE SIMULATED AND PRINTED:      {VARIABLE "RESP"}
*************************************

     CO

THE ASSIGNMENT DUE DATE IS *****  AUG. 31, 1982 *****

ESP          SAVED  20:38 MAY 26,'82
```


B: OUTPUT FROM RESOURCES PROGRAM

```
      RESOURCES

EXPERIMENTAL CONDITIONS:
************************

     THE RESTRICTIONS LISTED BELOW HAVE BEEN IMPOSED ON YOUR EXPERIMENT.
PLEASE REMEMBER THEM WHEN FORMULATING YOUR DESIGN. NOTE ESPECIALLY THAT
THE TOTAL NUMBER OF RUNS MAY NOT EXCEED ***   10   *** AND YOU CAN
HAVE AT MOST ***   20   *** OBSERVATIONS IN EACH RUN SO THE
MAXIMUM NUMBER OF OBSERVATIONS IN YOUR EXPERIMENT IS ***   200   ***.
     YOU MUST SPECIFY IN YOUR DESIGN MATRIX THE NUMBER OF THE PORT AT
WHICH EACH CIGARETTE IS TO BE SMOKED. IN THE PRESENT EXPERIMENT, THE RANGE OF
PORT NUMBERS IS ***   1 TO 20   *** INCLUSIVE AND PORT NUMBERS SHOULD
NOT BE REPEATED WITHIN A RUN UNLESS YOU HAVE BEEN INFORMED OTHERWISE.


     CIGARETTE LOT NUMBERS HAVE BEEN RANDOMLY ASSIGNED BY THE LAB, SO
YOU DO NOT NEED TO SPECIFY THEM.


RESTRICTIONS ON FACTOR LEVELS:
******************************

     YOU MUST SPECIFY LEVELS FOR EACH OF THE FOLLOWING FACTORS
IN YOUR DESIGN MATRIX.

FACTOR:   ADDITIVE    DIAMETER    OVERWRAP    POROSITY    WEIGHT
RANGE:    {0,0.8}     {.73,.78}    {1,3}       {1,6}     {.97,1.02}

     THE FOLLOWING FACTORS HAVE BEEN PRE"SET AT THE LEVELS INDICATED AND
SHOULD NOT BE INCLUDED IN YOUR DESIGN MATRIX.

 EFFICIENCY :   0.6

     RESPONSES WILL CONSIST OF ***      CO      ***
DELIVERIES FOR EACH CIGARETTE.

     YOUR DESIGN AND/OR ANALYSIS MUST BE SUBMITTED BY 5:00 P.M. ON
                ***   AUG. 31, 1982   ***.


YOU MAY NOW CREATE YOUR DESIGN MATRIX, BEARING IN MIND THE OBJECTIVES AND
RESTRICTIONS OUTLINED ABOVE, OR YOU MAY BEGIN YOUR SIMULATION IMMEDIATELY
BY TYPING

                    EXPERIMENT
                    SIMULATE

AND RESPONDING TO THE PROMPTS FROM THE PROGRAM. THE SIMULATION ROUTINE
HAS A BUILT"IN FACILITY FOR HELPING YOU TO CONSTRUCT THE DESIGN MATRIX.
```

FIGURE 2

A: DESIGN MATRIX FOR SIMULATION

(RUN 1: ROWS 1-16 RUN 2: ROWS 17-32)

ADDITIVE	DIAMETER	OVERWRAP	POROSITY	WEIGHT	PORT NO
0.00	0.730	1	2	0.970	15
0.80	0.780	1	2	0.970	8
0.80	0.730	3	2	0.970	17
0.80	0.730	1	5	0.970	16
0.80	0.730	1	2	1.020	6
0.00	0.780	3	2	0.970	11
0.00	0.780	1	5	0.970	13
0.00	0.780	1	2	1.020	5
0.00	0.730	3	5	0.970	14
0.00	0.730	3	2	1.020	3
0.00	0.730	1	5	1.020	12
0.80	0.780	3	5	0.970	7
0.80	0.780	3	2	1.020	9
0.80	0.780	1	5	1.020	18
0.80	0.730	3	5	1.020	10
0.00	0.780	3	5	1.020	4
0.80	0.730	1	2	0.970	17
0.00	0.780	1	2	0.970	5
0.00	0.730	3	2	0.970	14
0.00	0.730	1	5	0.970	8
0.00	0.730	1	2	1.020	15
0.80	0.780	3	2	0.970	4
0.80	0.780	1	5	0.970	9
0.80	0.780	1	2	1.020	12
0.80	0.730	3	5	0.970	7
0.80	0.730	3	2	1.020	13
0.80	0.730	1	5	1.020	11
0.00	0.780	3	5	0.970	18
0.00	0.780	3	2	1.020	10
0.00	0.780	1	5	1.020	6
0.00	0.730	3	5	1.020	3
0.80	0.780	3	5	1.020	16

B: SAMPLE OUTPUT FROM SIMULATE PROGRAM

RUTHLESS TOBACCO COMPANY
****

INVOICE NO.: 3361717
DATE: MAY 27, 1982

TO: RESEARCH BRANCH 251839
ATTENTION: J.R.ELLIOTT

ADDITIVE	DIAMETER	EFFICIENCY	OVERWRAP	POROSITY	WEIGHT	PORT NO.	LOT NO.	COSTS OF TEST SAMPLES	CO(MG)
0.00	0.730	0.60	S	C20	0.970	15		$245.93	20.05
0.80	0.780	0.60	S	C20	0.970	8		$270.47	18.50
0.80	0.730	0.60	VP	C20	0.970	17		$266.33	14.95
0.80	0.730	0.60	S	P20	0.970	16		$275.72	16.90
0.80	0.730	0.60	S	C20	1.020	6		$258.33	23.94
0.00	0.780	0.60	VP	C20	0.970	11		$270.07	12.41
0.00	0.780	0.60	S	P20	0.970	13		$280.68	16.12
0.00	0.780	0.60	S	C20	1.020	5		$262.07	16.81
0.00	0.730	0.60	VP	P20	0.970	14		$275.32	10.92
0.00	0.730	0.60	VP	C20	1.020	3		$257.93	17.53
0.00	0.730	0.60	S	P20	1.020	12		$267.32	19.00
0.80	0.780	0.60	VP	P20	0.970	7		$301.08	12.65
0.80	0.780	0.60	VP	C20	1.020	9		$282.47	13.56
0.80	0.780	0.60	S	P20	1.020	18		$293.08	16.76
0.80	0.730	0.60	VP	P20	1.020	10		$287.72	14.35
0.00	0.780	0.60	VP	P20	1.020	4		$292.68	10.79
0.80	0.730	0.60	S	C20	0.970	17		$256.33	20.48
0.00	0.780	0.60	S	C20	0.970	5		$260.07	18.89
0.00	0.730	0.60	VP	C20	0.970	14		$255.93	10.96
0.00	0.730	0.60	S	P20	0.970	8		$265.32	16.74
0.00	0.730	0.60	S	C20	1.020	15		$247.93	18.87
0.80	0.780	0.60	VP	C20	0.970	4		$280.47	14.31
0.80	0.780	0.60	S	P20	0.970	9		$291.08	16.99
0.80	0.780	0.60	S	C20	1.020	12		$272.47	16.20
0.80	0.730	0.60	VP	P20	0.970	7		$285.72	11.34
0.80	0.730	0.60	VP	C20	1.020	13		$268.33	16.81
0.80	0.730	0.60	S	P20	1.020	11		$277.72	21.08
0.00	0.780	0.60	VP	P20	0.970	18		$290.68	10.69
0.00	0.780	0.60	VP	C20	1.020	10		$272.07	10.36
0.00	0.780	0.60	S	P20	1.020	6		$282.68	14.19
0.00	0.730	0.60	VP	P20	1.020	3		$277.32	14.58
0.80	0.780	0.60	VP	P20	1.020	16		$303.08	11.65

TOTAL COST OF TEST SAMPLES: $ 8774.40
LABORATORY CHARGE FOR TESTING: $ 1440.00

TOTAL COST OF EXPERIMENT: $ 10214.40
(32 DETERMINATIONS)

C: SAMPLE OUTPUT FROM YATES' ANALYSIS PROGRAM
CALCULATIONS FROM YATES ALGORITHM
****

EFFECT OF ADDITIVE,DIAMETER,OVERWRAP,POROSITY AND WEIGHT ON CO DELIVERY,

FACTORS
****

A - ADDITIVE
D - DIAMETER
O - OVERWRAP
P - POROSITY
W - WEIGHT

EFFECT	MEAN SQUARE	SOURCE
1.34652389	14.50501264	A
-2.35146344	44.23504256	D
-0.05287523	0.02236632	AD
-5.22897135	218.73713124	O
0.07553895	0.04564906	AO
0.47182882	1.78097949	DO
0.60979595	2.97480879	ADO
-1.86477804	27.81917706	P
-0.26173253	0.54803133	AP
0.46694587	1.74430755	DP
0.53071160	2.25323840	ADP
0.12902434	0.13317825	OP
-0.41018827	1.34603531	AOP
0.05773496	0.02666661	DOP
-0.43157086	1.49002726	ADOP
0.84766564	5.74829625	W
0.17952511	0.25783411	AW
-2.12865819	36.24948566	DW
0.03128599	0.00783050	ADW
0.57676670	2.66127864	OW
-0.82418173	5.43420425	AOW
-0.22028360	0.38819892	DOW
0.66440226	3.53144294	ADOW
0.40679330	1.32384630	PW
0.05336085	0.02277904	APW
0.10865603	0.09444906	DPW
-0.11315684	0.10243576	ADPW
-0.39036221	1.21906125	OPW
0.15264198	0.18639660	AOPW
0.34829098	0.97045283	DOPW
-0.69515916	3.86597001	ADOPW

TABLE OF FACTOR MEAN RESPONSES:
****

FACTOR	LOW LEVEL	HIGH LEVEL	DIFFERENCE
ADDITIVE	14.93242535	16.27894924	1.34652389
DIAMETER	16.78141901	14.42995557	-2.35146344
OVERWRAP	18.22017297	12.99120162	-5.22897135
POROSITY	16.53807631	14.67329827	-1.86477804
WEIGHT	15.18185448	16.02952011	0.84766564

replication, are discussed in detail and illustrated with the simulator. The design matrix, simulated results, and partial analysis for a typical assignment are given in Figure 2. (A program* for performing an analysis of variance using Yates' method is also provided in the EXPERIMENT workspace).

Following consideration of the 2^n series, the students are assigned a "real" data project of the type described by W.G. Hunter (Hunter, 1977). The general factorial design is discussed and illustrated, as are miscellaneous topics such as composite designs, split plot experiments, incomplete block designs, and simple response surface methodology. At this point, the students have enough background to attack a problem based upon the full complexities of the simulator and a major project is assigned requiring each student or team of students to determine the levels of the six factors which will minimize one of the remaining responses; that is, tar, nicotine, or hydrogen cyanide. In larger classes, different responses can be assigned to different teams. However, it is our impression that providing everyone with the same response stimulates competition and improves the quality of the final products. At the end of the projects, a written report of all findings is submitted, sometimes accompanied by a verbal "in-class" presentation. Each separate experiment is described as to choice of design (including rationale), results from analysis of responses, and conclusions regarding the direction of future experiments. Recommendations regarding the optimal cigarette are presented, along with the total cost of experimentation. Project evaluation is based upon use and understanding of design, coherence of the sequence of experiments, proximity of recommended cigarette to the optimum, and economy of experimentation (with respect to total cost and to number of experiments). Typically, the two projects (the real and simulated experiments) account for a substantial proportion of the student's final grade in the course.

DISCUSSION AND CONCLUSIONS

Like most instructors who teach experimental design as a service course, we have attempted to provide our students with a variety of methods for scientific (that is, statistically valid) data collection and analysis. They are presented with a variety of distinct data collection strategies (that is, designs), each of which applies to a slightly different experimental situation and each of which is accompanied by a related but distinct method of analysis. As with D.R. Cox's approach in "Planning of Experiments" (Cox, 1958), our goal is to instill in our students an appreciation of which design is most appropriate for a given experimental setting, stressing at all times that there is really no "uniquely correct" answer but rather a series of tradeoffs between cost and precision (as defined in the broad sense).

We have found it difficult to accomplish this objective using lectures based solely on traditional experimental design texts (Cochran and Cox, 1957; Mendenhall, 1968; Snedecor and Cochran, 1967). Simply rehashing other people's data does not expose the students to the very real problems of factor and design selection that will be encountered when conducting their own experiments. Thus, like Hunter, we have supplemented this text material with a combination of real and simulated experiments. The former allow the students to experience first-hand the joys and trials of data collection and analysis; they must formulate a problem, select relevant responses and suspected influential factors, conduct an experiment, analyze the results, and prepare a coherent written report of their findings. This approach is extremely exciting to the students who become deeply involved in their projects. Yet, by their very nature, such projects are limited in scope, and must not be too large or too complicated. Moreover, they are subject to all the frustrations of the real research projects which they model; in particular, they have an annoying habit of revealing absolutely no statistically significant differences between the levels of the factors selected. Still as noted by Hunter, discussion of such outcomes can be a learning experience as well, provided neither the instructor nor the students place too much emphasis on the actual results of the experiments.

In contrast, with simulated experiments, the student sacrifices the opportunity to formulate a problem and select relevant responses. This is more than balanced, in our opinion, by the versatility through exposure to a wide variety of situations that such simulators in general, and ESP in particular, possess. In addition, the occurrence of missing observations is avoided and the time between selection of a design and production of results is greatly shortened. For these reasons, we have made ESP an integral part of our approach to experimental design. Besides using it for assignments and examples, we have written a text to accompany the course which makes extensive reference to the situation being modelled in illustrating all of the designs considered. This has a unifying influence on the course similar to that achieved with large data sets in multiple regression texts; namely, it provides a common context in which different designs can be compared on a variety of dimensions (for example, efficiency, degree of complexity, or cost). Moreover, the extreme versatility of ESP permits rapid production of data from numerous design situations - simple or complex, single or multi-factor, complete or incomplete block, whole or split plot - by slight adjustment to the model using the PARAMETERS program. This virtually limitless capability for producing material for assignments and examples frees the instructor from an otherwise time-consuming task.

To date, student reaction to the package has been extremely positive. On assignments, many students initially discount the possibility of port or run

biases and use completely randomized designs. However, when their classmates achieve a smaller experimental error in fewer observations with randomized block or Latin square designs, most begin to appreciate the sacrifice in precision that accompanies underdesigned experiments. In contrast, the more cautious students complicate their experiments unnecessarily by over-blocking or by using an excessive number of runs and/or factor levels. Frequently they explore one factor at a time and thus miss important interactions; failing to obtain significant effects or finding that the results of an experiment contradict previous conclusions, they seek advice from their instructor. Discussion of these "failed" experiments individually or in class is beneficial for illustrating, in a concrete manner, the advantage of applying factorial designs with minimal replication early in the experimental program. Moreover, concepts such as confounding and fractional replication, which at first seem somewhat elusive to the student, become valuable tools during the project. Slowly, sometimes painfully, most of the students come to realize that experimental design is not simply a theory but rather a very practical and useful concept.

Students find the simulation package simple to use and appear to derive much personal satisfaction from conducting the project. This positive attitude towards the simulation program seems to stem in part from the ease with which students can relate to the situation being modelled, a feature shared with those described by Pollock et al. Almost everyone, regardless of smoking status or academic discipline, has some knowledge of cigarette construction and the products of tobacco combustion. Terms such as "filter cigarette", "additive", "diameter", "weight", "tobacco", "tar", "nicotine" and "carbon monoxide" are familiar, so that both the factors and the responses have an intuitive appeal. Thus, students can readily appreciate the nature of the factors and reasons why they might influence tobacco combustion. A knowledge of cigarette manufacturing technology is not necessary for students to predict, for example, that increasing filter efficiency should decrease tar and nicotine delivery. Similarly, knowledge of human physiology is not essential for understanding that reducing carbon monoxide, tar, nicotine, or hydrogen cyanide output should produce a less hazardous cigarette. In these many ways, the simulation package serves as a driving force to make the theory of design a reality and obscure concepts more concrete.

References

[1] Afifi, A.A. and Azen, S.P. (1972) Statistical Analysis: A Computer Oriented Approach. New York: Academic Press.

[2] Baskevitch, N. (1981) Continuing Upward Trend in Low Tar and Nicotine Cigarettes. Tabak Journal International, 4, 304-306, 308, 310.

[3] Box, G.E.P. and Muller, M.E. (1958) A note on the generation of random normal deviates. Annals of Mathematical Statistics, 29, 610-611.

[4] Cochran, W.G. and Cox, G.M. (1957) Experimental Designs, 2nd edition. New York: Wiley.

[5] Cox, D.R. (1958) Planning of Experiments. New York: Wiley.

[6] Dalhamn, T. (1975) Chronic Pulmonary Disease - The Etiological Factors in Smoke. Proceedings of the Third World Conference on Smoking and Health, 415-426.

[7] Draper, N.R. and Smith, H. (1981) Applied Regression Analysis, 2nd edition. New York: Wiley.

[8] Gunst, R.F. and Mason, R.L. (1980) Regression Analysis and its Application: A Data-Oriented Approach. New York: Marcel Dekker.

[9] Hunter, W.G. (1977) Some Ideas About Teaching Design of Experiments, with 2^5 Examples of Experiments Conducted by Students. The American Statistician, 31, 12-17.

[10] Marsaglia, G. and Bray, T.A. (1964) A Convenient Method for Generating Normal Variables. Society for Industrial and Applied Mathematics: Review, 6, 260-264.

[11] Mendenhall, W. (1968) Introduction to Linear Models and the Design and Analysis of Experiments. Belmont: Duxbury Press.

[12] Pollock, K.H., Ross-Parker, H.M. and Mead, R. (1979) A Sequence of Games Useful in Teaching Experimental Design to Agriculture Students. The American Statistician, 33, 70-76.

[13] Resnik, F.E., Houck, W.G. et al (1978) Factors Affecting Static Burning Rate. Tobacco Science, 21, 103-107.

[14] Robinson, J.C. and Forbes, W.F. (1975) The Role of Carbon Monoxide in Cigarette Smoking: I. Carbon Monoxide Yield from Cigarettes. Archives of Environmental Health, 31, 425-434.

[15] Snedecor, G.W. and Cochran, W.G. (1967) Statistical Methods, 6th edition. Ames: Iowa State University Press.

APPENDIX I: PRODUCTION AND TESTING COSTS

PRODUCTION COSTS FOR TEST SAMPLES

The standard cigarette contains no additive, is 0.755cm in diameter with a 0.215cm overlap, weighs approximately 0.97gm and is covered with 30 second porosity citrate paper. The filter has a 15% efficiency and a standard overwrap.

Cigarettes of each type are manufactured in lots of 10,000 and the standard cigarette costs $233.00 per lot to produce (including government tax).

Additive costs $0.50 per gram and must be applied by hand at a flat rate of $10 per lot.

Changing the diameter from the standard value of 0.755cm increases or decreases cost proportional to the amount and type of paper required.

Filter material of 15% efficiency is included in the basic cost. Increasing the efficiency to 30% adds $5 per lot, to 45% adds $10 per lot, and to 60% adds $20 per lot.

Using porous or very porous overwraps adds $5 or $10, respectively, to the base cost per lot.

Citrate paper, regardless of porosity, adds nothing to the base cost, whereas phosphorus paper increases the cost by $30 per lot for 10 sec, $20 per lot for 20 sec, and $10 per lot for 30 sec porosity.

Weights above 0.97gm per cigarette increase the cost by $4 per kg of additional tobacco required.

LABORATORY COST FOR TESTING CIGARETTES

The first run costs $500 and yields tar deliveries whether they are requested or not. Carbon monoxide and nicotine determinations each add $300 to the base cost, while hydrogen cyanide determination adds a further $500. Thus, for example, a single run for carbon monoxide delivery only would cost $800 and a single run in which all four responses were monitored would cost $1600.

Each additional run adds 80% of the cost of the first run to the total.

APPENDIX II: PROGRAM LISTINGS FOR EXPERIMENT SIMULATION PACKAGE .

A: PROGRAM PARAMETERS

```
    ∇ PARAMETERS;LABEL;LOC;NF;I
{1}   '
RESTRICTIONS ON SIZE OF EXPERIMENT:
**********************************

MAXIMUM NUMBER OF:
     OBSERVATIONS PER RUN?'
{2}   SIZERES←□
{3}   '     RUNS?'
{4}   SIZERES←SIZERES,□
{5}   '
SPECIFICATION OF FACTOR LEVELS:
*******************************

(TYPE 9 IF FACTOR TO BE SPECIFIED BY STUDENT; OTHERWISE, SPECIFY LEVEL)

ADDITIVE?     {0.0 TO 0.8 MG PER CIGARETTE}'
{6}   FACLEV←□
{7}   'DIAMETER?     {0.73 TO 0.78 CM}'
{8}   FACLEV←FACLEV,□
{9}   'EFFICIENCY?     {.15,.30,.45,.60}'
{10}  FACLEV←FACLEV,□
{11}  'OVERWRAP?     {STANDARD,POROUS,VERY POROUS¯¯¯CODED: 1,2,3 }'
{12}  FACLEV←FACLEV,□
{13}  'POROSITY?     {C10,C20,C30,P10,P20,P30¯¯¯CODED: 1,2,3,4,5,6 }'
{14}  FACLEV←FACLEV,□
{15}  'WEIGHT?     {0.97 TO 1.02 GM PER CIGARETTE}'
{16}  FACSTUD←9=FACLEV←FACLEV,□
{17}  '
SPECIFICATION OF EXPERIMENTAL ERROR AND BIAS TERMS:
***************************************************

(TYPE 0 TO SUPPRESS EXPERIMENTAL ERROR OR SPECIFIC BIAS TERMS)

EXPERIMENTAL ERROR:     {NORMAL RANDOM VARIABLE}
     COEFFICIENT OF VARIATION (%)?   {USUALLY 5¯10 %}'
{18}  ERRSPEC←□
{19}  'MACHINE RUNS:     {NORMAL RANDOM VARIABLE}
     COEFFICIENT OF VARIATION (%)?   {USUALLY 5¯10 %}'
{20}  ERRSPEC←ERRSPEC,□
{21}  'MACHINE PORTS:     {U¯SHAPED FIXED EFFECT}
     NUMBER OF PORTS PER RUN?     {USUALLY 5¯20}'
{22}  →(0=¯1↑FACSTUD←FACSTUD,(0≠¯1↑FACLEV←FACLEV,□))/L0
{23}  '     MAXIMUM EFFECT (%)?   {PORT EFFECT IS PROPORTIONAL
                                   TO 4TH POWER OF DISTANCE FROM CENTER}'
{24}  ERRSPEC←ERRSPEC,□
{25}  →L1
{26}  L0:ERRSPEC←ERRSPEC,0
{27}  L1:'CIGARETTE LOTS:     {U¯SHAPED FIXED EFFECT}
     NUMBER OF CIGARETTE LOTS?   {USUALLY 100¯1000}'
{28}  →(0=¯1↑FACSTUD←FACSTUD,(0≠¯1↑FACLEV←FACLEV,□))/L2
{29}  '     MAXIMUM EFFECT (%)?   {LOT EFFECT IS PROPORTIONAL TO 10TH
                                        POWER OF DISTANCE FROM CENTER}'
{30}  ERRSPEC←ERRSPEC,□
{31}  →L4
{32}  L2:ERRSPEC←ERRSPEC,0
{33}  L4:'
RESPONSES TO BE SIMULATED:
**************************

(TYPE S IF RESPONSE IS TO BE SIMULATED; OTHERWISE, TYPE N)

CARBON MONOXIDE?  {10¯20 GM}'
{34}  RESP←1-'N'∈□
{35}  'TAR?  {5¯20 MG}     '
{36}  RESP←RESP,1-'N'∈□
{37}  'NICOTINE?'
{38}  RESP←RESP,1-'N'∈□
{39}  'HCN?'
{40}  RESP←RESP,1-'N'∈□
{41}  '
ASSIGNMENT DUE DATE:     MONTH/DAY/YEAR     {E.G. 10 25 82}
*******************'
{42}  DUEDATE←□
{43}  '
ENTER A TITLE FOR THE PARAMETER SETTINGS, THEN ADVANCE PAPER TO A FRESH
PAGE AND PRESS THE RETURN BUTTON'
{44}  LABEL←□
{45}  '

           PARAMETER RESTRICTIONS AND FACTOR SPECIFICATIONS
           ***********************************************

';LABEL
{46}  TEMP←14↑TO
```


A: PARAMETERS (CONT'D)

```
{47}  '

RESTRICTIONS ON SIZE OF EXPERIMENT:     {VARIABLE ¯SIZERES¯}
**********************************

                    OBS/RUN     RUNS
MAXIMUM NUMBER:';,'I10,I11'TEMP 1 2⍴SIZERES
{48}  ε')ERASE TEMP'
{49}  '

SPECIFICATIONS FOR DESIGN MATRIX:     {VARIABLES ¯FACLEV¯,¯FACSTUD¯}
**********************************

FACTORS WHOSE LEVELS MUST BE SPECIFIED BY STUDENTS:
     '
{50}  ,(LOC←FACSTUD{⍳6})/{⍵}LABEL←6 12⍴'    ADDITIVE    DIAMETER  EFFICIENCY    OVERWRAP    POROSITY    WEIGHT  '
{51}  →(~FACSTUD{7})/L5
{52}  '     PLUS PORT NUMBERS BETWEEN 1 AND ';FACLEV{7};' INCLUSIVE'
{53}  →(~FACSTUD{8})/L5
{54}  '     AND CIGARETTE LOT NUMBERS BETWEEN 1 AND ';FACLEV{8};' INCLUSIVE'
{55}  L5:→(6=+/LOC)/L6
{56}  '
FACTORS WHOSE LEVELS HAVE BEEN PRE¯SPECIFIED:
     '
{57}  I←1
{58}  NF←⍴LOC←(1-LOC)/⍳6
{59}  L7:→(LR,LR,LR,LO,LP,LR){LOC[I]}
{60}  LO:'     OVERWRAP:    ';(3 11⍴'STANDARD   POROUS     VERY POROUS'){FACLEV{4};}
{61}  →I←LR
{62}  LP:'     POROSITY:    ';(6 3⍴'C10C20C30P10P20P30'){FACLEV{5};}
{63}  →LR←I+1
{64}  LR:LABEL{LOC[I];};':   ';FACLEV{LOC[I]}
{65}  →(NF≥I←I+1)/L7
{66}  L6:'

SPECIFICATIONS FOR ERROR AND BIAS TERMS:     {VARIABLE ¯ERRSPEC¯}
*****************************************

EXPERIMENTAL ERROR:   COEFFICIENT OF VARIATION =  ';ERRSPEC{1};' %'
{67}  →(0=ERRSPEC{2})/L61
{68}  '
MACHINE RUNS:   COEFFICIENT OF VARIATION =  ';ERRSPEC{2};' %'
{69}  →L611
{70}  L61:'
MACHINE RUNS:   NO EFFECT'
{71}  L611:→(0=ERRSPEC{3})/L62
{72}  '
MACHINE PORTS:   MAXIMUM EFFECT =  ';ERRSPEC{3};' %'
{73}  →L621
{74}  L62:'
MACHINE PORTS:   NO EFFECT'
{75}  L621:→(0=ERRSPEC{4})/L63
{76}  '
CIGARETTE LOTS:   MAXIMUM EFFECT =  ';ERRSPEC{4};' %'
{77}  →L8
{78}  L63:'
CIGARETTE LOTS:   NO EFFECT'
{79}  L8:'

RESPONSES TO BE SIMULATED AND PRINTED:     {VARIABLE ¯RESP¯}
*************************************

     ';,RESP/{⍳4} 8⍴'CO      TAR     NICOTINE     HCN'
{80}  LABEL←12 5⍴' JAN. FEB. MAR. APR.  MAY JUNE JULY AUG.SEPT. OCT. NOV. DEC.'
{81}  '


THE ASSIGNMENT DUE DATE IS ***** ';LABEL{DUEDATE{1};};' ';DUEDATE{2};', 19';DUEDATE{3};' *****'
{82}  ε')GROUP GLOBVARS'
{83}  ε')GROUP GLOBVARS SIZERES FACLEV FACSTUD ERRSPEC RESP DUEDATE'
{84}  '

     '
{85}  ε')SAVE'

      ∇
```

B: PROGRAM RESOURCES

```
    ∇ RESOURCES;LABEL;NF;LOC;I
{1}   ε')QCOPY ESP..ICOTS GLOBVARS'
{2}   '

EXPERIMENTAL CONDITIONS:
************************

     THE RESTRICTIONS LISTED BELOW HAVE BEEN IMPOSED ON YOUR EXPERIMENT.
PLEASE REMEMBER THEM WHEN FORMULATING YOUR DESIGN. NOTE ESPECIALLY THAT
THE TOTAL NUMBER OF RUNS MAY NOT EXCEED ***   ';SIZERES{2};'   *** AND YOU CAN
HAVE AT MOST ***   ';SIZERES{1};'   *** OBSERVATIONS IN EACH RUN SO THE
MAXIMUM NUMBER OF OBSERVATIONS IN YOUR EXPERIMENT IS ***   ';(×/SIZERES);'   ***.'
{3}   LABEL←2 6 12⍴'   ADDITIVE    DIAMETER  EFFICIENCY    OVERWRAP    POROSITY    WEIGHT      {0,0.8}     {.73,.78}
    {.15,.60}     {1,3}       {1,6}    {.97,1.02}'
{4}   →(~FACSTUD{7})/L1
{5}   '     YOU MUST SPECIFY IN YOUR DESIGN MATRIX THE NUMBER OF THE PORT AT
WHICH EACH CIGARETTE IS TO BE SMOKED. IN THE PRESENT EXPERIMENT, THE RANGE OF
PORT NUMBERS IS ***   1 TO ';FACLEV{7};'   *** INCLUSIVE AND PORT NUMBERS SHOULD
NOT BE REPEATED WITHIN A RUN UNLESS YOU HAVE BEEN INFORMED OTHERWISE.
    '
{6}   →(~FACSTUD{8})/L2
{7}   '
     YOU MUST SPECIFY IN YOUR DESIGN MATRIX THE NUMBER OF THE LOT FROM
WHICH THE CIGARETTE WAS CHOSEN. IN THE PRESENT EXPERIMENT, THE RANGE OF
LOT NUMBERS IS ***   1 TO ';FACLEV{8};'   *** INCLUSIVE. LOT NUMBERS MAY BE
REPEATED IF NECESSARY, EVEN WITHIN THE SAME TYPE OF CIGARETTE, BUT THIS
IS NOT ADVISEABLE.'
{8}   →L3
{9}   L2:'
     CIGARETTE LOT NUMBERS HAVE BEEN RANDOMLY ASSIGNED BY THE LAB, SO
YOU DO NOT NEED TO SPECIFY THEM.
    '
{10}  →L3
{11}  L1:'
     BOTH PORT AND CIGARETTE LOT NUMBERS HAVE BEEN RANDOMLY ASSIGNED
BY THE LAB, SO YOU DO NOT NEED TO SPECIFY THEM.
    '
{12}  L3:'

RESTRICTIONS ON FACTOR LEVELS:
****************************** '
{13}  '
     YOU MUST SPECIFY LEVELS FOR EACH OF THE FOLLOWING FACTORS
IN YOUR DESIGN MATRIX.
    '
{14}  'FACTOR:';,(LOC←FACSTUD{⍳6})/{⍵}LABEL{1;;}
{15}  'RANGE: ';,LOC/{⍵}LABEL{2;;}
{16}  →(6=+/LOC)/L4
{17}  NF←⍴LOC←(1-LOC)/⍳6
{18}  '
     THE FOLLOWING FACTORS HAVE BEEN PRE¯SET AT THE LEVELS INDICATED AND
SHOULD NOT BE INCLUDED IN YOUR DESIGN MATRIX.
    '
{19}  I←1
{20}  L3:→(LR,LR,LR,LO,LP,LR){LOC{I}}
{21}  LO:'     OVERWRAP :    ';(3 11⍴'STANDARD   POROUS     VERY POROUS'){FACLEV{4};}
{22}  →1+LR
{23}  LP:'     POROSITY :    ';(6 3⍴'C10C20C30P10P20P30'){FACLEV{5};}
{24}  →1+LR
{25}  LR:LABEL{1;LOC{I};};':   ';FACLEV{LOC{I}}
{26}  →(NF≥I←I+1)/L5
{27}  L4:'

     RESPONSES WILL CONSIST OF ***   ';,(RESP/{⍵}4 8⍴'   CO      TAR    NICOTINE     HCN ');'   ***
DELIVERIES FOR EACH CIGARETTE.'
{28}  LABEL←12 5⍴' JAN. FEB. MAR. APR.  MAY JUNE JULY AUG.SEPT. OCT. NOV. DEC.'
{29}  '

     YOUR DESIGN AND/OR ANALYSIS MUST BE SUBMITTED BY 5:00 P.M. ON
                ***   ';LABEL{DUEDATE{1};};' ';DUEDATE{2};', 19';DUEDATE{3};'   ***.
    '
{30}  '


YOU MAY NOW CREATE YOUR DESIGN MATRIX, BEARING IN MIND THE OBJECTIVES AND
RESTRICTIONS OUTLINED ABOVE, OR YOU MAY BEGIN YOUR SIMULATION IMMEDIATELY
BY TYPING

                    EXPERIMENT
                    SIMULATE

AND RESPONDING TO THE PROMPTS FROM THE PROGRAM. THE SIMULATION ROUTINE
HAS A BUILT¯IN FACILITY FOR HELPING YOU TO CONSTRUCT THE DESIGN MATRIX.
    '
{31}  ε')ERASE GLOBVARS'
    ∇
```


C: PROGRAM EXPERIMENT

```
    ∇ EXPERIMENT
{1}   ε')QCOPY ESP..ICOTS SIMULATE'
    ∇
```

D: PROGRAM SIMULATE

```
    ∇ SIMULATE;I;J;DMAT;DESMAT;NOPR;FIXERR;RANDERR;NRUNS;NOBS;LAB;LOC1;LOC2;LOC3;LOC4
{1}   ε')QCOPY ESP..ICOTS  GLOBVARS'
{2}   LAB←12 5⍴' JAN. FEB. MAR. APR.  MAY JUNE JULY AUG.SEPT. OCT. NOV. DEC.'
{3}   ⍝
{4}   ⍝   CURRENT DATE COMPARED WITH DUE DATE AND PROGRAM
{5}   ⍝   TERMINATED IF TOO LATE.
{6}   ⍝
{7}   →(DUEDATE{3}>¯1↓LOC1←LOC1,(100×LOC2-1↓LOC1←LOC1,(⌊LOC2←100×LOC2-LOC1←⌊LOC2←(I25)÷10000)))/P0
{8}   →(DUEDATE{1}>LOC1{1})/P0
{9}   →(DUEDATE{1}<LOC1{1})/TURF
{10}  →(DUEDATE{2}≥LOC1{2})/P0
{11}  TURF:'MANAGEMENT MADE THE DECISION ON ';LAB{DUEDATE{1};};' ';DUEDATE{2};', 19';DUEDATE{3};'.
YOUR ADVICE IS A LITTLE LATE.   CLEAN OUT YOUR DESK¯¯¯YOU ARE FIRED ***'
{12}  →LERASE
{13}  ⍝
{14}  ⍝   OBTAINING THE DESIGN MATRIX
{15}  ⍝
{16}  P0:'IS A DESIGN MATRIX AVAILABLE FOR INPUT?   {YES OR NO}'
{17}  →('N'∈⍞)/PDES
{18}  'INPUT THE DESIGN MATRIX.'
{19}  DESMAT←⍞
{20}  →DMCHK
{21}  ⍝
{22}  ⍝   DESIGN MATRIX ¯DESMAT¯ IS NOW CREATED
{23}  ⍝   BY THE SUBPROGRAM 'DESGEN' IF NOT AVAILABLE
{24}  ⍝   EXTERNALLY TO 'SIMULATE'
{25}  ⍝
{26}  PDES:ε')QCOPY ESP..ICOTS   DESGEN'
{27}  DESGEN
{28}  ε')ERASE DESGEN'
{29}  ⍝
{30}  ⍝   THE DESIGN MATRIX ¯DESMAT¯ IS NOW TESTED TO INSURE
{31}  ⍝   THAT IT CONFORMS TO ALL RESTRICTIONS IMPOSED
{32}  ⍝   BY THE PROGRAM  'PARAMETERS'.
{33}  ⍝
{34}  DMCHK:→((NOBS←1↓⍴DESMAT)≤×/SIZERES)/P1
{35}  'THE MAXIMUM NUMBER OF OBSERVATIONS ALLOWED IS ***   ';(×/SIZERES);'   ***.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{36}  →LERASE
{37}  P1:'NUMBER OF RUNS?'
{38}  →(SIZERES{2}≥NRUNS←⍞)/P3
{39}  'YOU ARE ONLY ALLOWED ***   ';SIZERES{2};'   *** RUNS.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{40}  →LERASE
{41}  P3:→(1=NOPR÷⌊NOPR←NOBS÷NRUNS)/P2
{42}  'THE TOTAL NUMBER OF OBSERVATIONS IS NOT AN
INTEGRAL MULTIPLE OF THE NUMBER OF RUNS.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{43}  →LERASE
{44}  P2:→(NOPR≤SIZERES{1})/P4
{45}  'YOU ARE ONLY ALLOWED ***   ';SIZERES{1};'   *** OBSERVATIONS PER RUN.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{46}  →LERASE
{47}  ⍝
{48}  ⍝   THE DESIGN MATRIX  IS NOW EMBEDDED IN AN
{49}  ⍝   NOBS BY 8 MATRIX CALLED ¯DMAT¯ AND IS TESTED
{50}  ⍝   TO VERIFY THAT ALL VALUES ARE IN THE PROPER RANGES.
{51}  ⍝
{52}  P4:DMAT←(NOBS,8)⍴FACLEV
{53}  DMAT{;LOC1←FACSTUD{⍳8}+DESMAT
{54}  LOC2←8 10⍴' ADDITIVE  DIAMETER EFFICIENCY OVERWRAP  POROSITY   WEIGHT   PORT NO.   LOT NO.  '
{55}  LOC3←8 2⍴0 0.8 0.73 0.78 0.15 0.6 1 3 1 6 0.97 1.02 1,FACLEV{7},1,FACLEV{8}
{56}  I←1
{57}  LOC4←⍴LOC1
{58}  L1:→(0=+/(LOC3{LOC1{I};1}>DMAT{;LOC1{I}}),LOC3{LOC1{I};2}<DMAT{;LOC1{I}})/L2
{59}  '
AT LEAST ONE OF THE SPECIFIED ***   ';LOC2{LOC1{I};};'   *** LEVELS IS
OUTSIDE THE ACCEPTABLE RANGE ***   { ';LOC3{LOC1{I};1};', ';LOC3{LOC1{I};2};' }   ***.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{60}  →LERASE
{61}  L2:→(LOC4≥I←I+1)/L1
{62}  →(~FACSTUD{3})/NEFF
{63}  →(0=+/LOC4-⌊LOC4←DMAT{;3}÷0.15)/NEFF
{64}  '
AT LEAST ONE OF THE SPECIFIED ***   EFFICIENCY   *** LEVELS IS
NOT 0.15,0.30,0.45 OR 0.60 , THE ONLY VALUES PERMITTED.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{65}  →LERASE
```

```
{66}  NEFF:→(0=+/FACSTUD{4 5 7 8})/NOQUAL
{67}  I←1;LOC4←⍴LOC1←FACSTUD{4 5 7 8}/4 5 7 8
{68}  LQ2:→(0=+/DMAT{;LOC1{I}}-LDMAT{;LOC1{I}})/LQ1
{69}  '
AT LEAST ONE OF THE SPECIFIED ***   ';LOC2{LOC1{I};};'   *** LEVELS IS
NOT AN INTEGER BETWEEN ***   ';LOC3{LOC1{I};1};' AND ';LOC3{LOC1{I};2};'   ***.
***** RE¯CHECK THE DESIGN MATRIX AND TRY AGAIN *****'
{70}  →LERASE
{71}  LQ1:→(LOC4≥I←I+1)/LQ2
{72}  ⍝
{73}  ⍝   THE REQUIRED RESPONSES ARE NOW SIMULATED.
{74}  ⍝
{75}  ⍝
{76}  ⍝   RANDOM N(0,CVE*2) ERRORS ARE GENERATED USING
{77}  ⍝   THE POLAR METHOD OF MARSAGLIA AND ARE STORED
{78}  ⍝   IN THE VECTOR ¯RANDERR¯.
{79}  ⍝
{80}  NOQUAL:→(0≠+/ERRSPEC)/B1
{81}  RANDERR←FIXERR←NOBS⍴0
{82}  LOC1←I20
{83}  →L3
{84}  B1:→(0≠+/ERRSPEC{1 2})/B2
{85}  RANDERR←NOBS⍴0
{86}  →B3
{87}  B2:ε')QCOPY ESP..ICOTS SETLINK'
{88}  LOC2←SETLINK LOC1←I20
{89}  →(0≠ERRSPEC{1})/B4
{90}  RANDERR←NOBS⍴0
{91}  →B5
{92}  B4:RANDERR←l0
{93}  I←J←1
{94}  LRE:→(0<J←-J)/LRE2
{95}  LRE1:LAB←+/LOC3×LOC3←¯1+2×LOC3←(2?100000)÷100001
{96}  →(LAB>1)/LRE1
{97}  RANDERR←RANDERR,(LOC4←LOC3×LAB←((¯2×⍟LAB)÷LAB)*0.5){1}
{98}  →LREND
{99}  LRE2:RANDERR←RANDERR,LOC4{2}
{100} LREND:→(NOBS≥I←I+1)/LRE
{101} RANDERR←RANDERR×0.01×ERRSPEC{1}
{102} B5:→(0=ERRSPEC{2})/LL1
{103} ⍝
{104} ⍝   RANDOM N(0,CVR*2) RUN BIAS INTRODUCED {IF REQUIRED}
{105} ⍝   AND ADDED TO ¯RANDERR¯.
{106} ⍝
{107} DESMAT←l0
{108} I←J←1
{109} LRR:→(0<J←-J)/LRR2
{110} LRR1:LAB←+/LOC3×LOC3←¯1+2×LOC3←(2?100000)÷100001
{111} →(LAB>1)/LRR1
{112} DESMAT←DESMAT,(LOC4←LOC3×LAB←((¯2×⍟LAB)÷LAB)*0.5){1}
{113} →LRRND
{114} LRR2:DESMAT←DESMAT,LOC4{2}
{115} LRRND:→(NRUNS≥I←I+1)/LRR
{116} RANDERR←RANDERR←(,⍉(NOPR,NRUNS)⍴DESMAT×0.01×ERRSPEC{2})
{117} LL1:ε')ERASE SETLINK'
{118} B3:FIXERR←NOBS⍴0
{119} →(0=ERRSPEC{3})/L31
{120} ⍝
{121} ⍝   FIXED EFFECT PORT BIAS INTRODUCED {IF REQUIRED}.
{122} ⍝   EFFECT RELATED TO 4TH POWER OF DISTANCE FROM CENTER.
{123} ⍝   MAXIMUM EFFECT (÷) IS CONTROLLED.
{124} ⍝
{125} FIXERR←FIXERR+0.01×ERRSPEC{3}×(((2×DMAT{;7})-1+FACLEV{7})÷FACLEV{7})*4
{126} L31:→(0=ERRSPEC{4})/L3
{127} ⍝
{128} ⍝   FIXED EFFECT CIGARETTE LOT BIAS INTRODUCED {IF REQUIRED}.
{129} ⍝   EFFECT RELATED TO 10TH POWER OF DISTANCE FROM CENTER.
{130} ⍝   MAXIMUM EFFECT (÷) IS CONTROLLED.
{131} ⍝
{132} FIXERR←FIXERR+0.01×ERRSPEC{4}×(((2×DMAT{;8})-1+FACLEV{8})÷FACLEV{8})*10
{133} L3:DMAT←DMAT,{2}NOBS⍴0
{134} ⍝
{135} ⍝   REQUIRED RESPONSES ARE NOW GENERATED
{136} ⍝   AND STORED IN THE VECTORS ¯CO¯, ¯TAR¯,¯NICOTINE¯, AND/OR ¯HCN¯.
{137} ⍝
{138} →(~1∈RESP{1 4})/P5
{139} LOC2←0.1,0.35,0.6,0,0.25,0.4
{140} CO←10+(4×(DMAT{;1}-LOC2{DMAT{;5}})*2)÷((3-DMAT{;4})*2)+(LOC2{DMAT{;5}}×(3-(DMAT{;4}-2)*2))+120×(1.77-(0.2×LOC2{
      DMAT{;5}})+DMAT{;6}÷DMAT{;2}*2)*2
{141} DMAT←DMAT,{2}CO←CO+(FIXERR+RANDERR)×+/CO÷NOBS
{142} P5:→(~1∈RESP{2 4})/P6
{143} LOC2←0.15,0.2,0.5,0,0.1,0.6
{144} TAR←(1-DMAT{;3})×12.5+(5×0.64-(DMAT{;1}-LOC2{DMAT{;5}})*2)+(2×LOC2{DMAT{;5}})+(0.5×(1+(DMAT{;2}*2)÷DMAT{;6})×(
      DMAT{;4}-3)*2)+20×(1.77-(0.2×LOC2{DMAT{;5}}))+DMAT{;6}÷DMAT{;2}*2)*2
{145} DMAT←DMAT,{2}TAR←TAR+(FIXERR+RANDERR)×+/TAR÷NOBS
{146} P6:→(~RESP{3})/P7
{147} LOC2←0 0.15 0.4 0 0.15 0.4
{148} NICOTINE←((1-DMAT{;3})*0.5)×1.028+(0.25×LOC2{DMAT{;5}})+(0.1×(1+(DMAT{;2}*2)÷DMAT{;6})×(DMAT{;4}-3)*2)+6×(1.77-(
      0.2×LOC2{DMAT{;5}})+DMAT{;6}÷DMAT{;2}*2)*2
{149} DMAT←DMAT,{2}NICOTINE←NICOTINE+(FIXERR+RANDERR)×+/NICOTINE÷NOBS
```



```
{150} P7:→(~RESP{4})/P8
{151} DMAT←DMAT,{2}HCN←(0.25×CO+2×TAR)*2
{152} DMAT←DMAT{;l9},((0 9)↓DMAT){;RESP/(1,2,3,3+RESP{3})}
{153} P8:→(4=+/RESP)/P9
{154} ε')ERASE  ',,(~RESP)/{1}4 9⍴'       CO      TAR NICOTINE    HCN  '
{155} ⍝
{156} ⍝   LABORATORY TEST COSTS ARE COMPUTED
{157} ⍝   FOR EACH TREATMENT COMBINATION TYPE OF CIGARETTE
{158} ⍝   AND STORED IN COLUMN 9 OF THE MATRIX ¯DMAT¯.
{159} ⍝
{160} P9:DMAT{;9}←0.01×L0.5+100×((10×DMAT{;1}≠0)+0.5×DMAT{;1})+(0,5,10,20){LDMAT{;3}÷0.15}+(40×DMAT{;6}-0.97)+(5×¯1+
      DMAT{;4})+((0 0 0 30 20 10){DMAT{;5}}+233)×1+1.2144×DMAT{;2}-0.755
{161} 'TYPE IN YOUR NAME:'
{162} LOC2←⍞
{163} 'TYPE IN YOUR I.D. NUMBER THEN ADVANCE PAPER TO A FRESH
PAGE AND PRESS THE RETURN BUTTON.

NOTE:   RESPONSES ARE STORED IN:  ';,RESP/{1}4 9⍴'       CO      TAR NICOTINE      HCN'
{164} ⍝
{165} ⍝   TREATMENT COMBINATIONS,COSTS AND RESPONSES ARE PRINTED.
{166} ⍝
{167} LOC4←⍞
{168} LAB←12 5⍴' JAN. FEB. MAR. APR.  MAY JUNE JULY AUG.SEPT. OCT. NOV. DEC.'
{169} '


                        RUTHLESS TOBACCO COMPANY
                        ************************

                                                    INVOICE NO.: ';LOC1
{170} '                                                      DATE: ';LAB{(LLOC1÷10000);};' ';(L(10000|LOC1)÷100);', 19';l0
      LOC1←I25
{171} 'TO: RESEARCH BRANCH ';LOC4;'
ATTENTION: ';LOC2
{172} '
                                                          COSTS
ADDI¯   DIA¯   EFFI¯  OVER¯  POR¯          PORT  LOT   OF TEST'
{173} 'TIVE   METER  CIENCY  WRAP  OSITY  WEIGHT   NO.  NO.   SAMPLES ',,RESP/{1}4 9⍴'   CO(MG)  TAR(MG)  NIC(MG)  HCN
      (UG)'
{174} LAB←'F8.3,BI6,BI6,X3,⍞$⍞,F6.2,',(,RESP/{1}4 5⍴'F9.2,F9.2,F9.3,F9.1,'),'X1'
{175} J←0
{176} LOC1←3 5⍴'    S    P   VP'
{177} LOC2←6 7⍴'    C10    C20    C30    P10    P20    P30'
{178} LOC3←4++/RESP
{179} LOC4←+/DMAT{;9}
{180} DESMAT←14⍴0
{181} LOOP:' '
{182} I←1
{183} LOOP1:;,'F4.2,F8.3,F7.2,X1 'DESMAT(1 3⍴DMAT{I;1 2 3});LOC1{DMAT{I;4};};LOC2{DMAT{I;5};};,LAB DESMAT(1,LOC3)⍴5↑
      DMAT{I;}
{184} →((NOPR+1)>I←I+1)/LOOP1
{185} →(NRUNS>J←J+1)/DMATDRP
{186} →LCOST
{187} DMATDRP:DMAT←(NOPR,0)↓DMAT
{188} →LOOP
{189} LCOST:LOC1←0.01×L0.5+100×(0.2+0.8×NRUNS)×500++/RESP/300 0 300 500
{190} '                                                        ¯¯¯¯¯¯¯'
{191} '                     TOTAL COST OF TEST SAMPLES:    $';,'F9.2'DESMAT LOC4
{192} '                     LABORATORY CHARGE FOR TESTING: $';,'F9.2'DESMAT LOC1
{193} '                                                    ¯¯¯¯¯¯¯¯¯
                  TOTAL COST OF EXPERIMENT:      $';,'F9.2'DESMAT(LOC1+LOC4);,'
                       ( ';NOBS;' DETERMINATIONS)'
{194} LERASE:ε')ERASE  DESMAT GLOBVARS'



E:  PROGRAM END



    ∇ END
{1}   ε')ERASE SIMULATE'
{2}   ε')SAVE'
    ∇
```

```
F:  PROGRAM DESGEN


    ∇ DESGEN;LOC1;LOC2;LOC3;J;LOC4
{1}   'YOU MUST SPECIFY LEVELS FOR EACH OF THE FOLLOWING:'
{2}   LOC1←2 6 12⍴'   ADDITIVE    DIAMETER  EFFICIENCY    OVERWRAP    POROSITY    WEIGHT      {0,0.8}     {.73,.78}
   {.15,.60}     {1,3}       {1,6}     {.97,1.02}'
{3}   'FACTOR:';,((LOC2←FACSTUD{16})/{1}LOC1{1;;})
{4}   'RANGE: ';,(LOC2/{1}LOC1{2;;})
{5}   '
     ALL OTHER FACTOR LEVELS FOR EACH TREATMENT COMBINATION HAVE BEEN
PRE¯SET AS INDICATED PREVIOUSLY.'
{6}   →(0=+/FACSTUD{7 8})/L1
{7}   '
     IN ADDITION, YOU MUST PROVIDE THE FOLLOWING:'
{8}   →(~FACSTUD{7})/L11
{9}   '
       A NUMBER BETWEEN *** 1 AND ';FACLEV{7};'   *** WHICH SPECIFIES
       THE PORT AT WHICH THAT PARTICULAR CIGARETTE WILL BE SMOKED.'
{10}  L11:→(~FACSTUD{8})/L1
{11}  '                              ';(2 3⍴'   AND'){+/FACSTUD{7 8};};'
       A NUMBER BETWEEN *** 1 AND ';FACLEV{8};'   *** WHICH SPECIFIES
       THE LOT FROM WHICH THAT CIGARETTE WAS SELECTED.'
{12}  L1:'
     ENTER THE TOTAL NUMBER OF OBSERVATIONS REQUIRED. THIS NUMBER
MUST NOT EXCEED ***   ';(×/SIZERES);'   ***.'
{13}  LOC1←⍬
{14}  DESIGN←(LOC1,LOC3←+/FACSTUD)⍴0
{15}  LOC2←+/LOC2
{16}  '
     ENTER THE TOTAL NUMBER OF RUNS REQUIRED. THIS MUST NOT
EXCEED ***   ';SIZERES{2};'   *** RUNS.
     '
{17}  LOC4←LOC1÷⍬
{18}  L2:'
     YOU MAY NOW INPUT YOUR DESIGN, ONE TREATMENT COMBINATION AT A TIME.
FOR EACH TREATMENT COMBINATION, SPECIFY THE LEVELS OF THE ***   ';LOC2;' ***
FACTORS REQUIRED IN THE ORDER LISTED ABOVE {I.E., ALPHABETICALLY}.'
{19}  ,FACSTUD{7 8}/{1}2 29⍴'THEN, GIVE THE PORT NUMBER.  FINALLY, GIVE THE LOT NUMBER.'
{20}  '
     START WITH THE FIRST TREATMENT COMBINATION IN THE FIRST RUN
AND ENTER THE REQUIRED VALUES AS A VECTOR OF LENGTH ';LOC3;',LEAVING
A BLANK BETWEEN EACH NUMBER.
     '
{21}  LOC1←LOC1-1
{22}  J←0
{23}  R1:'
SPECIFICATIONS FOR:   RUN *** ';(⌈(1+J)÷LOC4);' *** ; TREATMENT COMBINATION *** ';(1+LOC4|J);' ***'
{24}  →(LOC3=⍴LOC2←⍬)/R2
{25}  '
*** INPUT ERROR IN NUMBER OF LEVELS¯¯¯TRY AGAIN ***
     '
{26}  →R1
{27}  R2:DESIGN{J+1;}←LOC2
{28}  →(LOC1≥J←J+1)/R1
{29}  '

     YOUR DESIGN SPECIFICATIONS ARE NOW STORED IN THE ';(1+LOC1);' BY ';(+/FACSTUD);'
MATRIX ¯DESIGN¯.  DO YOU WISH TO CHECK THE DESIGN MATRIX BEFORE CONTINUING
WITH YOUR SIMULATION?   {YES OR NO}'
{30}  →('N'∈⍞)/DCONT
{31}  '


'
{32}  ⎕PP←4'
{33}  DESIGN
{34}  ⎕PP←10'
{35}  '


DO YOU WISH TO MAKE CHANGES IN THE MATRIX ¯DESIGN¯?   {YES OR NO}'
{36}  →('N'∈⍞)/DCONT
{37}  LCH:'SPECIFY CHANGE AS FOLLOWS: ROW,COLUMN,NEW VALUE'
{38}  DESIGN{LOC1{1};LOC1{2}}←(LOC1←⍞){3}
{39}  FCHG:'DO YOU WISH TO MAKE ADDITIONAL CHANGES?  {YES OR NO}'
{40}  →('N'∈⍞)/DCONT1
{41}  →LCH
{42}  DCONT1:'DO YOU WISH TO RE¯CHECK ¯DESIGN¯?   {YES OR NO}'
{43}  →('N'∈⍞)/DCONT
{44}  '


'
{45}  ⎕PP←4'
{46}  DESIGN
{47}  ⎕PP←10'
{48}  '


'
{49}  →FCHG
{50}  DCONT:DESMAT←DESIGN
    ∇
```


TEACHING EXPERIMENTAL DESIGN: PRESCRIBED BY A MEDICAL STATISTICIAN

Sheila M Gore, MRC Biostatistics Unit, Cambridge

INTRODUCTION

The saying : those who can do, those who can't teach, and those who can't teach write about it inclines me to "had my wheesit" to save face. Only the radical papers by Dr. Mead and Dr. Preece inspire me foolhardily to write a partial prescription for the teaching of experimental design which would have a high compliance amongst medical statisticians. The course I prescribe has been tested on statistics honours students at Aberdeen University and aspects of it, such as practical methods of randomization and teaching by real examples, have found favour with doctors. Adequate trial size to guarantee a fair test of new medicines is a central issue in clinical studies and is discussed in the section 'Adequate Trial Size'. The collaborative nature of much applied statistical work means that the statistician must find a common language with colleagues from other disciplines - clinical trials for example, concern patients, not experimental units. The section 'Collaboration and Consultancy' considers how to give students practice in statistical consultancy. Unreliable data do not warrant the investment of statistical effort; undergraduate statisticians should learn data scrutiny techniques. These are discussed in the section 'Reliable Data'.

PRACTICAL RANDOMIZATION

Undergraduate statisticians should encounter methods of randomization which are practical in the context of (1) sequential patient entry (2) multi-centre trials and (3) ingenious investigators who try to discern the randomization "pattern" by fair means and foul (for example, keeping track of block length or resealing envelopes). Balanced block randomization with fixed small block length is insufficiently secure when treatments are identifiable (Gore and Altman, 1981). Simple randomization has no place except in the largest clinical trials because it appears careless when an investigator ends up with grossly unequal treatment numbers if he intended a randomization ratio of 1:1. Stratification by more than one or two major covariates risks there being many strata which are represented by few patients.

Sequential patient entry cannot be manipulated to satisfy elegant, balanced designs of the type that traditional teaching delights in. Instead medical statisticians use regression models in clinical trial analysis, allowing for most covariates only retrospectively, a practice which is at odds with the rigour of classic experimental design.

As a compromize solution there is a practical method of randomizing patients which minimizes marginal treatment imbalance over several covariates. The method is called minimization (see Pocock, 1979, White and Freedman 1978) and can be programmed easily or kept track of using card indices. Minimization has the advantage of defeating the ingenious investigator. The following is an example of what happens:- a doctor telephones the central randomization office to register his trial patient. Eligibility criteria are checked and relevant covariate information is given - for example : tumour size ≧ 5cms, premenopausal patient, teratment centre Cambridge. These data are keyed in. The minimization routine looks up the marginal treatment totals so far for (1) all Cambridge patients (2) premenopausal women and (3) patients with tumour size ≧ 5cms and works out the treatment allocation which leads to least imbalance, summed over the marginal treatment distributions for the three covariates. In fact the routine will display a ranking of allocations from the treatment which induces least imbalance to the assignment which most exaggerates the imbalance. A probability distribution is associated with this treatment ranking and favours the treatment assignment leading to least imbalance. For example, in a trial which compares early intervention after myocardial infarction by intravenous beta-blockade with no intervention, the probability distribution might be chosen so that the chance is 2/3 that the patient is allocated to the treatment which reduces the marginal imbalance between treatment numbers.

REAL EXAMPLES

Even trivial designs (as in clinical trials) are challenging in application. Consider a within-patient study of dinitrochlorobenzene (DNCB) for the treatment of alopecia areata, a condition which entails partial or complete hair loss. DNCB dissolved in acetone was painted onto one half of the patient's scalp, the other half being untreated. Even discounting the punk appearance (in 1977!) of those patients for whom DNCB induced effective hair growth, the design has two other limitations. Firstly, the action of DNCB may be systemic as well as topical and so the untreated side may benefit from a systemic action; secondly there is contamination of the untreated area when the patient rubs or scratches his head. Choosing an appropriate study design is one of the most challenging aspects of problems in applied statistics and so experimental design must surely be taught through real examples which are not "shorn" of context.

Curiosity is a help to the applied statistician in discovering the context of a problem:

"I have six honest serving men,
They taught me all I knew,
Their names are what and why and when,
And how and where and who".

Rudyard Kipling

ADEQUATE TRIAL SIZE

A fundamental problem in the design of clinical experiments is one to which neither Dr. Preece nor Dr. Mead accorded much speculation. It is patient accrual or adequate trial size. In a survey of 38 clinical trials in The Lancet (July-December 1977) I found that no study mentioned how trial size was determined. More recently Dersimonian et al. (1982) reported on 67 trials published in four leading general medical journals. Power considerations were omitted from 85% of the reports. Pocock (1978), on the other hand, surveyed trial organisers and found that in a majority of randomized cancer studies statistical methods were used to determine trial size. Not until these statistical aspects are routinely written in to published reports will the ideas become familiar to doctors. Differences which are not statistically significant in trials with inadequate accrual are often wrongly interpreted (Freiman et al. 1979) as meaning that there is no difference between treatments. Were the corresponding confidence interval reported, it would be seen to be hopelessly wide and a ridiculous basis for deciding between treatments. I regard studies which have less than 50% power to establish a worthwhile treatment difference as bad science. Dr. Preece points out that there is virtue in combining the results from such trials. Indeed there is, provided that one avoids selection bias. Because of the tendency for unpromising studies to fall by the wayside and go unreported, there is an inherent bias in combining the results from published trials only. The pharmaceutical industry is the best place to keep and make available a register of all clinical trials for which it releases supplies of new medicines.

Therapeutic advances - such as the use of anticoagulants (Chalmers et al. 1977) and intervention by beta-blockade after myocardial infarction - have been delayed by inadequate trial size and failure to report confidence intervals. It is perhaps more than mere coincidence that the first trial of beta-blockade which showed a significant reduction in mortality after myocardial infarction (Norwegian Multicenter Study Group 1981) was also the first which was large enough to detect a worthwhile treatment difference. "Not statistically significant" does not mean "there is no difference". Too often it means still: "there has not been a fair test of this treatment"!

That the teaching of experimental design does not emphasize enough how to determine patient numbers was evident at a recent SERC conference in Reading for postgraduate students in statistics, few of whom knew how to calculate trial size. Is it any wonder that doctors get it wrong sometimes? An incidental advantgae of the promotion of stopping rules and group sequential methods (Pocock 1982) for clinical studies is that they impose an explicit discipline on trial size.

COLLABORATION AND CONSULTANCY

When I gave an honours course on clinical trials in the Department of Statistics at Aberdeen I invited clinical colleagues to attend. The presence of a doctor in the lecture room was valuable support for me when explaining the medical background to problems and was a reminder to students of the collaborative nature of much applied statistical work. Of course, I gave notice of the more technical lectures which doctors would not find it profitable to attend but always included on their return a resume of the essence of the statistical method we had been studying. Thus I encouraged students to recognise that what is important for a medical audience is being able to understand the idea behind the formal mathematics, and that communicating such ideas is part of their responsibility as statisticians. Likewise, my recommendation about interpreting and reporting results, especially in relation to complex data, was for the use of graphical methods and descriptive statistics to record initial impressions. These are the stepping-stones by which the statistician reaches a final analysis and are no less important for the reader if he is to have some feel for the data and understand better the results of statistical modelling.

Even before report-writing, students must develop a skill for statistical consultancy. I have learned that divulging to students the steps which I take to become conversant with a medical problem before the first consultation and the strategems I have for verifying what a doctor tells me is no substitute for their practising this for themselves. When lecturing on clinical trials I therefore posed to each honours student (numbers varied from three to seven students) a separate clinical trial problem, as it had been brought to me. The students arranged a meeting with me, who played the part of the clinician, and they got only such additional information about the problem as they thought to ask for. This is precisely what happens in practice: if one fails to ask a relevant question, the information is not forthcoming as the doctor may not appreciate its statistical value. Several students said how relieved they were that their first experience of consultation was "in house" because their fund of questions was too soon exhausted. Only then did they take seriously the points I had

ventured about being prepared for a consultation! Moreover, many found it necessary to request further meetings when they got down to writing the protocol for the clinical trial because they became aware of salient gaps in their knowledge of the problem.

I suspect that too often, when lecturers talk about practical classes and practical exercises, what they mean is practice at routine design and analysis in relation to some rounded problem. But it is the ragged edges that statistical consultancy is all about. Discovering these and weighing them in the balance is why applied problems are seldom routine, and are never dull.

RELIABLE DATA

At the end of a lecture on the checking and editing of medical data I was surprised to be asked "Who has the time to do all that?". My reply was "Who can afford not to do it!". Students should learn about and have practice in data scrutiny. No amount of analysis can make bad data good.

Transcription errors may be checked by reading across or by double-punching; they can be avoided by keying in data directly to a microcomputer which has a well designed and tested screen format. Range checks and checks on the accuracy of recording are elementary and can be programmed, as can logical checks: for example it is inconsistent that a patient who is 60 years old at diagnosis should be menstruating regularly, impossible that date of treatment is later than date of death and so on ... Arithmetic errors can be avoided by programming the relevant calculations - in two separate studies I have detected 15% and 40% respectively error rates for age last birthday. Sorting data by trial number and/or date of randomization, date of birth may detect duplicate patients or missing serial numbers. Even after all such checks have been made there may be residual errors which have transgressed neither logical nor range checks. To estimate the residual error rate, one selects a random sample of patients and compares their computer records with hospital case notes. I did such an exercise for the Western General breast cancer series and discovered an important coding anomaly which had been overlooked hitherto. Of the 361 patients whose tumours were ulcerated, fixation to the overlying skin was recorded as nil or complete for all but 9 patients, the extremes being noted with similar frequency. The explanation was that no coding convention had been determined, and the consultant had argued variously:

(1) the tumour is ulcerated and so has broken through the skin, having been previously completely fixed to it. Code: complete fixation.

(2) ulceration supersedes fixation, which is thus irrelevant. Code: nil fixation.

A salutary tale on which to end.

References

[1] Dersimonia, R., Charette, L.J., McPeek, B. and Mosteller, F. (1982) Reporting on methods in clinical trials. N.Engl.J.Med, 306, 1332-7.

[2] Freiman, J.A., Chalmers, T.C., Smith, H. and Kuebler, R.R. (1978) The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. N.Engl.J.Med, 299, 690-4.

[3] Gore, S.M. and Altman, D.G. (1982) Statistics in Practice, British Medical Association, London, 44-47.

[4] Norwegian Multicenter Study Group (1981) Timolol-induced reduction in mortality and reinfarction in patients surviving acute myocardial infarction. N.Engl.J.Med., 304, 801-7.

[5] White, S.J. and Freedman, L.S. (1978) Allocation of patients to treatment groups in a controlled clinical trial. Br.J.Cancer, 37, 849-857.

[6] Pocock, S.J. (1978) Size of cancer clinical trials and stopping rules. Br.J.Cancer. 757-66.

[7] Pocock, S.J. (1979) Allocation of patients to treatment in clinical trials. Biometrics, 35, 183-97.

[8] Pocock, S.J. (1982) Interim analyses for randomized clinical trials. Biometrics, 38, 153-62.

CONTRIBUTION TO DISCUSSION OF THE ABOVE PAPERS

B Jones, University of Kent

Dr. Preece has indeed told us "what it is about", and drawn our attention to examples of some of the misconceptions and careless thinking which has crept into some published accounts of designed experiments. However, an important question I would like to put to him is "How to teach such a large subject in a typical university lecture course?". In Table 1 I have given a division of Dr. Preece's "top 12" components. The upper half includes what I consider tends to get taught, certainly to students of mathematics, to those studying the design of experiments. The lower half contains those topics which I feel are better placed in a separate and later course which aims to cover the practical application of the theory. I regard both halves of the table to be of equal importance but such a subdivision makes it more convenient from the teaching point of view. A typical lecture course consists of 30-40 hours of lectures and so each half could be accommodated in a standard lecture program.

TABLE 1: A division of DAP's "Top Twelve" headings

Theory

1. The basic principles of statistical inference.
4. The basic principles of "good" design.
5. Mathematical methods for constructing designs and obtaining analyses, particularly analyses of variance and covariance.
9. Ability to recognise and interpret patterns and anomalies in data.
10. Understanding of the assumptions underlying standard statistical analyses, and the ability to recognise when these assumptions are seriously violated in practice.
11. Facility in spotting mistakes in analyses.
12. Skill in reporting conclusions in clear, simple prose, with appropriate tabulations and graphs.

Practice

2. The ability to select treatments, experimental units, and variates that will permit "useful" results to be obtained.
3. The ability to recognise relevant sources of variability, and to make preliminary assessments of their importance and relative magnitudes.
6. Trained personnel for setting up and managing experiments.
7. Proper instruments (balances, etc.) for obtaining accurate, precise data values, and trained personnel to use the equipment.
8. Training in data transcription (with checking).

Arranging for students to gain real experience of the practical side of the subject is difficult. Apart from a few visits to an industry or research station to see real experiments in progress I think most practical skills can be learned by carefully studying case histories. The particular university course I teach is quite short and so I am able only to give a flavour of the practical side. This I do by making use of a large collection of projector slides of photographs of real experiments. These are used to illustrate what experiments look like, the size and shape of plots and the details recorded on the working plan of the design e.g. sites of former roads or streams in an agricultural trial. Spending a lot of time visiting one real experiment is usually not justified because a student only gets exposed to a limited point of view. The student certainly gets a feel for the problems involved but only gets experience of one experimental situation. For example, studying the practical difficulties of designing a raspberry trial may not be much use to a student who is eventually employed to design trials on cowpea in a developing country.

Another question I would like to put concerns the approach we ought to adopt when teaching the theory. Should we present the subject by way of numerous special cases of interest or should we start with a general theory and then begin to specialize? Having tried both I am clear that they each have certain advantages and disadvantages.

With reference to the "three R's" of Fisherian experimental design it seems to me that the one corresponding to randomization stands out as being far more difficult to teach than the other two. All three are quite intuitively reasonable but the large theoretical framework associated with randomization is often daunting to most students.

Mr. Mead's ideas on using computers as an aid to teaching are very useful. The use of simulation games provides a worthwhile alternative to lengthy visits to experimental sites. I have since tried his Tomato game with an M.Sc. in Statistics class and they found it enjoyable and instructive.

His overall approach is attractive: we can now let computers take the burden of analysis so let's concentrate on getting the most appropriate design. Often in the past experiments have been designed a particular way because the analysis was available in a textbook, without any regard for whether or not the most suitable experimental design was used. The use of some partially balanced incomplete block designs certainly could be criticised from this point of view. The use of a computer program such as GLIM is definitely required to make the most of Mr. Mead's ideas. However, teaching students to learn GLIM just for a design course, would probably not be practical. Therefore, if such a program is to be used it should have already been taught as part of the overall degree course in Statistics.

His lack of emphasis on the usefulness of factorial experiments is regrettable as these are widely used, particularly in industry. However, his suggestion that we should concentrate on getting desirable patterns of treatment concurrences is most important but has received very little obvious attention in standard text books. In a factorial experiment there are usually some effects, e.g. high order interactions which are of little interest, while others e.g. main effects must be well estimated. Finding patterns of treatment concurrences which lead to designs where the most important effects are well estimated is often not easy but provides a useful basis for introducing students to the idea that standard text book designs may be of little use on some occasions. Also it helps introduce another feature of the computer - that of being programmed to search for 'optimal' designs. To illustrate this point of view I have used an algorithm (Jones and Eccleston, 1980) to search for designs in 6 blocks of 4 plots for a 2 x 3 factorial experiment. The two factors I have labelled as L and N. Factor L has two levels A and B and N has 3 levels 1, 2 and 3. A set of five possible contrasts of interest is given in Table 2, where each contrast is scaled so that the sum of squares for its coefficients is unity. Table 3 gives the variances of the five contrasts obtained from separate runs of the algorithm. In each run a degree of importance or weight was assigned to each contrast, where a weight of 1 indicates full importance and 0 indicates no importance. The algorithm

TABLE 2: Treatment Contrasts

	A1	A2	A3	B1	B2	B3	Scaling value	
1	1	1	1	-1	-1	-1	$1/\sqrt{6}$	
2	1	0	-1	1	0	-1	$1/\sqrt{4}$	MAIN EFFECTS
3	1	-2	1	1	-2	1	$1/\sqrt{12}$	
4	1	0	-1	-1	0	1	$1/\sqrt{4}$	INTER-ACTIONS
5	1	-2	1	-1	2	-1	$1/\sqrt{12}$	

requires a starting design and the one used in each run was that given by Mr. Mead in his Figure 2. It will be noted that Mr. Mead's design cannot be improved upon if all contrasts are of equal importance, i.e. the weights are (1, 1, 1, 1, 1). It is clear that there is a wide choice of design depending on the experimenter's interests. The algorithm used searches for optimal concurrences of treatments and so is fully in the spirit of Mr. Mead's suggestions.

TABLE 3: Jones and Eccleston (1980) Designs obtained from Roger Mead's Starting Design

Contrast variances

WEIGHTS	1	2	3	4	5
1 1 1 1 1	0.30	0.27	0.27	0.31	0.26
1 1 1 $\frac{3}{4}$ $\frac{3}{4}$	0.25	0.27	0.27	0.31	0.31
1 1 1 $\frac{1}{2}$ $\frac{1}{2}$	0.25	0.27	0.27	0.31	0.31
1 1 1 $\frac{1}{4}$ $\frac{1}{4}$	0.25	0.27	0.27	0.31	0.31
1 1 1 0 0	0.25	0.27	0.27	0.8	0.31
$\frac{3}{4}$ $\frac{3}{4}$ $\frac{3}{4}$ 1 1	0.30	0.31	0.27	0.27	0.26
$\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ 1 1	0.27	0.32	0.30	0.26	0.25
$\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$ 1 1	0.25	0.33	0.33	0.25	0.25
0 0 0 1 1	0.25	0.33	0.33	0.25	0.25

A point I find puzzling however is his suggestion that only unweighted concurrences should be considered. I feel it is only the weighted concurrences which matter and give in Tables 4 and 5 some examples which support this view. In each example the experimenter is likely to be misled by looking at the unweighted concurrences.

In conclusion I congratulate both speakers on making plain the fact that the subject of experimental design is full of interest and is able to take advantage of the advent of computers.

Reference

[1] Jones, B. and Eccleston, J.A. (1980) Exchange and interchange procedures to search for optimal designs, J.R. Statist. Soc., B, 42, 238-243.

TABLE 4: Example of Unequal concurrences and Unequal Block Sizes
6 Treatment in 18 Blocks - Block Sizes 6 x 4, 12 x 2.

BLOCK						
I	1	2	3	4		
II	1	2	5	6		
III	3	4	5	6		
IV	1	2	3	4		
V	1	2	5	6		
VI	3	4	5	6		
VII	1	3				
VIII	1	4				
IX	1	5				
X	1	6				
XI	2	3				
XII	2	4				
XIII	2	5				
XIV	2	6				
XV	3	5				
XVI	3	6				
XVII	4	5				
XVIII	4	6				

Note: In this design there are 4 concurrences of (1,2), (3,4) and (5,6) and 3 concurrences of all other pairs.

If only unweighted concurrences are considered we should expect treatment differences involving (1,2), (3,4) and (5,6) to have smaller variances than the rest.

However this design is totally balanced - the sum of the weighted concurrences of every pair is equal.

In particular every standardized treatment difference has variance $\sigma^2/6$.

TABLE 5: Example of equal concurrences and unequal block sizes

BLOCK						
I	1	2	3	4		
II	1	2	5	6		
III	3	4	5	6		
IV	1	3				
V	1	4				
VI	1	5				
VII	1	6				
VIII	2	3				
IX	2	4				
X	2	5				
XI	2	6				
XII	3	5				
XIII	3	6				
XIV	4	5				
XV	4	6				

Notes: Here the unweighted concurrences are equal but not all treatment differences are estimated with the same precision.
E.g. Variance of 1 vs 2 is $0.25\sigma^2$
Variance of 1 vs 3 is $0.23611\sigma^2$.

Indeed the design is group divisible - so again the weighted concurrences are the only things of interest.

TOPIC 14: TRAINING INDUSTRIAL PRACTITIONERS

ORGANISER: DR R D SNEE

THE POLYMATH CONSULTANT

A A Greenfield, The Queen's University of Belfast

INTRODUCTION

I speak from experience as a consultant. This experience was garnered mainly in two very different industries: the steel industry and the medical faculty of a university. It is supported by a brief consulting experience in the pharmaceutical industry, by employment in quite different roles in the mining and publishing industries, and by brief contacts with the gas, oil, chemical and quarrying industries. It was through all these diverse experiences that I eventually reached two conclusions: one is that there is a profitable role in life for a polymath; the second is that a man or a woman can be taught to be a polymath.

After I had reached these conclusions, I made a very pleasing discovery. This was that my friends at Sheffield City Polytechnic, Kanji and Gilchrist (1981), had reached the same conclusions by an entirely different route. But instead of the word 'polymath' they use the term 'applied statistician'. Furthermore, they are doing something about it: they are running the best course, in terms of practical value, that I have been able to find anywhere.

A polymath is a person who knows many arts and sciences. In an earlier paper (Greenfield, 1979) I described the role of the polymath in industry, the great benefits he would bring to it, and the pleasure he would derive from such employment. In this paper, I shall briefly restate those points and then proceed to propose the outline curriculum of a three year undergraduate course leading to a Bachelor of Science. I shall leave it to you to imagine the fine detail of the course and hope that I shall stimulate you to add some constructive suggestions.

Since writing that paper, I was greatly encouraged to read a report by the American Statistical Association's committee on Training of Statisticians for Industry, entitled "Preparing Statisticians for Careers in Industry". In this, they acknowledged a close agreement with my ideas and they discussed the major approach to their task. However, even they did not go far enough in describing the best possible education for an industrial consultant. Even though the broad education they propose, and Sheffield City Polytechnic and others provide, will produce excellent applied statisticians, and even though I agree that statistics is essential to my proposal, there is a need in industry at large for a brighter jewel; the applied polymath. And I believe that he

could be produced, post A-level, in three hard years of study and active practical learning.

WHO NEEDS POLYMATHS?

British industry as a whole needs polymaths because Britain lacks many of the natural resources that other countries possess and we depend for our prosperity on manufactured products. There has been a widespread failure in recent years to put more resources into product development. This has not usually been for lack of finance, nor for lack of government investment incentives. Nor is it for the lack of advice on how to do it, because academics and management consultants are for ever telling the industrial managers how to do it. I believe it is because management structures have become trees of little boxes, each filled with a specialist who knows his own place. Creativity has been cramped. The way to overcome this is to employ super-statisticians. Not super in the sense of theoretical and academic brilliance, but super in the sense of knowing something, preferably a lot, about most things that make industry tick: the sciences and technologies, people and markets, accountancy, finance, economics and business systems, and, above all, communications because it is communication that cuts across all those barriers.

But why a statistician? Why not a super-something-else? Professor Michael Healy (1979) gave the answer in his inaugural lecture:
"The statistician can contribute technological suggestions about the conduct of the research which may greatly increase the impact or reduce the cost; but he can contribute a great deal more than this if he has an adequate broad knowledge of the subject matter. He is the one who is most likely to spot the analogies between the problem in hand and other problems elsewhere in medicine, in agriculture or in industry. He may be the only one to suspect that such an analogy exists. He will also be the least likely to carry such analogies across uncritically into the new area and the most capable of adapting them to the new requirements".

Broadness of view and ability to see the big picture are needed in any dynamic and enterprising business, because of the diversity of problems and situations that will be met. It troubles me that so many statistical studies are confined: statisticians restrict themselves to the experimental design and analysis in a formal quasi-academic cage without allowing themselves to reach out into the wider issues which they could so easily influence if they tried. Hunter (1981) argues similarly from his own experience.

SOME EXAMPLES

In my earlier paper (1979) I briefly described the structure of the steel industry and explained one process, steelmaking, in more detail to illustrate the need for a useful statistician to be able to understand, communicate with, and help such diverse specialists as: physicists, chemists, metallurgists, psychologists, accountants and engineers of all sorts: electronics, chemical, mechanical and production.

Figure 1 shows that in the pharmaceutical industry there is a similar diversity of links. This is contrary to popular opinion.

When I asked the academics in an English university: "What should be the academic training of statisticians in the pharmaceutical industry?" they all gave me the same answer. I asked in the statistics department. "Clinical trials", they told me. I asked in the pharmacology department. "Clinical trials", they said. And in the departments of physiology, biochemistry and community medicine, they all told me: "Clinical trials".

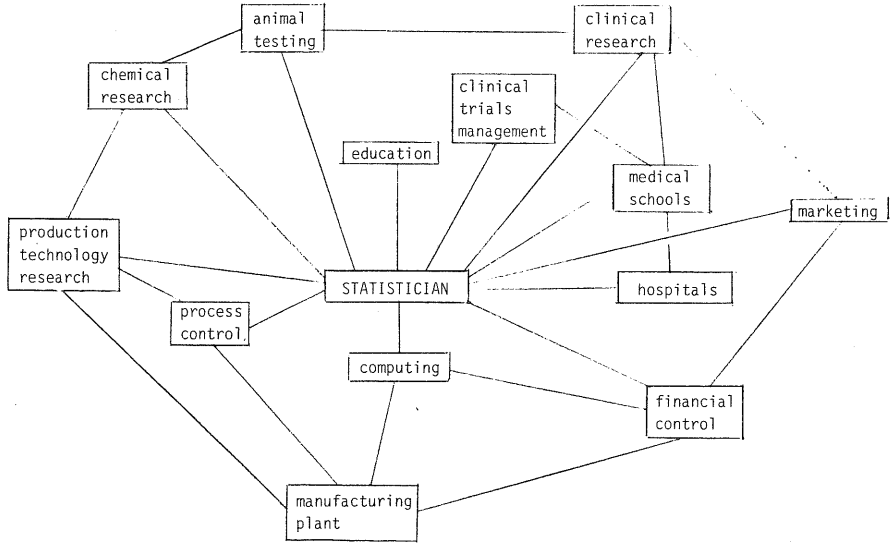


FIGURE 1: The Statistician in the Pharmaceutical Industry

In examining the problem, I have tried to imagine the life of a statistical consultant in the industry with a view to drawing the links between his life and his needs for academic training. But in thinking about his life, I found myself dwelling more on his desirable characteristics, his background and his interests rather than what is currently regarded as formal curriculum studies. I hope to identify those qualities by describing the possible role of the statistician in the manufacturing, commercial and research environment. My example is based on a brief attachment to a pharmaceutical company although the conclusions relate to all other industries based on the physical, chemical, biological, or engineering sciences.

Opportunities for, and worth of statisticians in all industries are far from fully appreciated. The well-trained statistician is more suited than any other specialist to take a leadership role in the solution of company problems, whether in research, development, manufacturing, commercial or personnel areas. But what is well-trained?

The well-trained statistician is the one with personality and panache, with an understanding of company structure and operation, with a wide based knowledge of science and engineering, and with an ability to communicate. Each of these is as important as a thorough grasp of statistical theory and methodology.

Just look at the structure of the pharmaceutical industry. Many companies are now international. Some have more than 40 peripheral associate or subsidiary companies distributed across the face of the globe constituting the nodes of a network of manufacturing, research and commercial activities. The statistician must be prepared to work in such an organisation and contribute generally to its success.

Think of a single unit within such a great complex: one research laboratory. This laboratory is manned largely by scientists and engineers who have been educated deeply and narrowly in special fields, often to the doctorate level, plus associated technicians. There are medics, psychologists, biochemists, botanists, zoologists, and maybe plant engineers and control engineers. The roles of these specialists are to invent, develop and test new products and processes. These aspects of research and development are, in my view, the most exciting, the most intellectually appealing, and the most rewarding. The rewards are in the pleasure of personal achievement, in contributions to scientific progress, and in enhancing industrial profitability. But these specialists should also be in touch with their counterparts in industrial production (on the works) and in commerce (in the marketing companies). Here their role, associated with plant in use and with current products, is to identify problems affecting the efficiency, and therefore the costs, of working and marketing practice, and to solve these problems. These may be associated

with energy, manpower, health and safety, consumable materials, product quality, test procedures, plant maintenance and life, packaging, distribution, storage, and legal and ethical conditions related to the products.

Within that organisation, let us identify what a statistician actually does, or what his colleagues expect him to do, rather than what his job description says he does. Think too, about what a statistician could do, to the benefit of himself and of his employer, IF he had the somewhat broader training that I shall advocate. Furthermore, consider what he would do if that training included practical instruction on how to assert himself.

All of you with experience in similar organisations, know that the statistician has daily to practice the skills of consulting, or being consulted; he is involved in operations research, in simulation, in computing, in running courses, in chairing committees, in speaking in public, and in writing reports. His involvement, for example, in clinical trials is not restricted to the design of the trial and the analysis of the data. He must advise the medical specialists on what is practical as well as scientific and ethical and he must press them to keep to the timetable and to the protocol. He must set up and monitor management procedures for processing records and he must ensure that there is the proper flow of material in and out of the computing department. He must communicate with, and educate, the staff in peripheral companies and participating hospitals, and he must ensure that adequate supplies of materials and forms are reaching the area manager and being forwarded by him. He must keep the clinical research director informed about the daily progress and report to him all the things that do go wrong with big trials with suggestions as to how to put them right. Even without describing any projects in detail, you will appreciate that the statistician will be expected to understand and help medics, scientists, and business men, even though all these specialists will arrogantly remain ignorant of the statistician's arts.

Thus the mathematical or statistical consultant must have a wide scientific and technical education if he is to succeed in industrial research. That education must include written and spoken communication with scientists and engineers of different disciplines. Such polymaths already exist, but usually their wide abilities have developed over a decade or more after a more specialised formal edcuation. That formal education should and could be modified to produce more polymaths so that they may be more useful to research and development and the rest of industry earlier in their careers. There is a helplessness and potential bankruptcy in specialisation, but there is security and potential wealth in versatility. Education should aim for that versatility with particular qualities: a thorough grasp of scientific method acquired through formulating and solving practical problems; a broad knowledge of scientific and technical subjects relating

to the major industries; an understanding of industrial organisation and management; and, emphatically, a well-developed ability to communicate at all levels and in many different circumstances.

This list of qualities, you may object, are a long way from statistics. I think not. The statistician is central to any situation where data are collected, analysed, interpreted and used. There is no occasion in industry or industrial research where this does not apply. A statistical education is one of two keystones to this degree in polymathy that I envisage. The need for good communication is the second keystone and it is universally agreed. Yet little is done about it after O-levels. It is a sad fact that those students who choose the sciences rarely have any instruction, practice, or correction in communication after the age of 16, when they take English language at O-level. Then, two or three decades later, after they have been employed in industry for several years, either they or their managers realise how thoroughly bad they are at communication. They are sent on brief, superficial, and expensive management courses. Those who employ new graduates are often heard to grumble about the poor articulation, the bad grammar, and the wrong spelling of their recruits.

THE POLYMATH COURSE

An approximation to polymathy in some degree institutions is the modular method by which students may choose modules from the full arts and sciences spectrum. They win points towards their final degrees as they succeed in each module. This does not seem to me to be such a good idea because it does not yield a coherence of approach, knowledge and experience. Chemistry, for example, is distinct from English which is distinct from statistics. On the other hand, my proposal, although it could be seen as a list of many disparate subjects, is planned to be a merging of those subjects into a unified whole. It is a reflection of creativity, which can be defined as an ability to associate those things which are not associated so as to produce something new. Perhaps the best way to explain how this can be done is by describing my draft three year timetable shown in Figure 2.

This timetable is only a first draft. I hope that discussion will lead to improvements. But it will serve for the present to illustrate my intentions.

There are two main features of the course. One of these is what you would normally expect in any undergraduate course, but do not always find: it takes the final school year material, in the sciences, and uses it to illustrate and practice the philosophy of the course before introducing new material and building towards a greater structure of knowledge, experience, understanding and ability. The second feature is that each subject is taught in such a way

as to link closely with all other subjects. None of them is a subject on its own. They are all mutually intermeshed. For example, chemistry, writing, and statistics are mutually dependent on each other. Thus, to illustrate both main features, sixth-form scientific laboratory work will be repeated in the first term, but with emphasis on data recording, analysis and presentation, written reports, conference discussion, and touching on the needs for mathematical modelling and experimental design.

COMMUNICATION

We all know well enough that the statistical industrial consultant must be able to interview a client, to persuade him to talk about and help him to formulate his problems. He must be able to write memoranda and letters explaining his work and seeking information and action. He should be able to write reports and articles describing how he has tackled and solved problems. He should be able to speak in public, spontaneously as well as in the formal presentation of papers. He should know how to behave on committees and working groups, as chairman, secretary, or contributor. He should have a working knowledge of making diagrams and charts, slides and transparencies, block and plate making, typography, printing, proof-reading and layout of publications. He should be trained in self-communication: note-taking, transcription, and filing.

Although there are volumes about the theory of communications, the real worth of any course is in the eventual ability of the students. We should therefore aim at the development of practical skills in terms of speaking ability, writing ability, and the ability to use the media of communication. One of the most useful of these abilities is typing and I propose that this should be required in the first term of the course. A further advantage of this skill is its value in using computer terminals and word processors. This should be followed by other physical techniques of communication: printing, photography, graphic design, and visual aids for lectures, displays and exhibitions, including the making of slides, overhead projector foils, projection techniques, models, films, video equipment, and computer developed graphics.

The skills of writing would be developed by formal instruction about the writing and editing of project reports, scientific papers, proposals for research and investment, technical articles, instruction manuals, memoranda, committee minutes, letters for business and newspaper publication, the use of word processing. There would be sessions on the analysis of published papers and reports.

All of these written and mechanical communication abilities would be brought together in the third year when three hours a week would be devoted to the production of two course publications: a weekly newspaper and a monthly

magazine. These would contain the products of other parts of the course, which I shall describe later, as well as any other material produced by staff and students.

Public speaking abilities would be developed in the first year together with interviewing skills. Public speaking would include the arts of impromptu speaking, debating, speechmaking, teaching, and public lecturing. Interviewing would be helped by third year students who would act as interviewees, about simulated or real problems, for the first year students. In all these activities, mutual criticism and competition would be encouraged to develop personality and confidence and to overcome faults.

Speaking exercises would continue through into the final year when there would be a weekly debate on current affairs. This would be organised with formal rules of debate and course points would be awarded for participation. Committee work would be simulated and there would be brainstorming sessions to stimulate creativity.

All practical work would be reported by the students, using visual aids, as if they were teaching or delivering papers, and they would be encouraged to criticise each other's presentation as well as content. These presentation and discussion sessions appear prominently on my draft timetable. The weekly climax of these activities would, however, be the course conference, occupying the last one and a half hours of the Friday afternoon. The students of all three years would gather together for this. Although presentations would mainly be by senior students, others would be encouraged to join in.

All of this practical learning would be supported by more formal instruction: in the first year about library and information systems, abstracting, indexing, referencing, researching, and about the theory and art of communication: in the second year about the structure of business and public institutions and the flow and control of information; and in the third year about committee structures, their organisation and operation, about public relations, and about the organisation of courses and meetings.

During the second and third year, each student would study in each term:

1. A paper from one of the more mathematical scientific or statistical journals. He would prepare from this:
   a) an oral presentation, with visual aids, aimed at a general scientific audience.
   b) a written paper, version of (a), for which he would do all the editorial preparation including graphics for production in the monthly journal.
   c) a newspaper account of the work, with explanations of possible social, industrial, or scientific implications.

   d) an interview role, to give the first year students experience of interviewing.

2. A paper from a trade or technical publication, which:
   a) he would criticise constructively, including analysis of data published with the article, or supporting data retrieved from other sources, in a written paper for publication in the monthly journal.
   b) a newspaper account of the work.

PRACTICAL PROJECTS

My proposal for dealing with practical projects illustrates the integrated nature of the course. In most courses, for example, if a student were asked to prepare some ethyl acetate he would be exercised only in his chemical laboratory ability. In this course, however, he would also be exercised in model formulation, the statistical aspects of experimental design and analysis, in data collection and recording, in writing and graphic presentation, and finally in oral presentation to the class the following day.

Such projects would be based first of all on subjects with which the incoming student is already familiar, because they would be drawn from A-level science material in chemistry, physics, biology and mechanics. They would be used further to develop his broad scientific knowledge, so that in the second term of the first year his projects would be to do with psychology, electronics and physiology, while in the third term some industrial projects would be simulated in the laboratory. One afternoon of every week in the second year would be devoted to an environmental project, extending his interests into geology, meteorology, pollution, industry, transport, architecture, housing, education and other public services. All such projects would be designed to give experience in hypothesis statement, modelling, data collection, library research, interviewing, data analysis, reporting, and interpretation, as well as written and spoken presentation. Other project work in the second year would be laboratory based, with computer simulation of industrial and other institutional situations and processes. There would also be data analysis projects.

All these would lead to larger and individual research projects for the students in the third year when there would also be half day attachments each week to other institutions, such as manufacturing industry, commerce, research laboratories, and local government and services. Kanji (1979) has discussed the role of projects in statistical education.

MATHEMATICS, STATISTICS AND COMPUTING

Mathematics is not only a basis for studying and understanding the sciences and other subjects the student will meet, it is an international language, a medium of communication. It should therefore be spread throughout the course, as are the other communications studies, rather than being confined to the early terms. There are some important aspects of mathematics which must be included to support other subjects, particularly probability, statistics, operational research, computing, electronics and instrumentation. While these may continue throughout, the later terms should be devoted more to the analysis and discussion of published mathematical applications papers.

The most important statistical techniques, in my experience, are: graphical display and data summary; design of experiments; regression analysis; response surface methodology including evolutionary operation; discriminant analysis; simulation, both process and event; time series analysis, including the analysis of continuous multivariate processes with feedback; numerical analysis; surveys, reliability; quality control and acceptance sampling. Others have views about an applied statistics curriculum and it is being covered well in other papers so I shall not dwell on it here. However, a vital feature is the statistics laboratory, described by Kanji (1974) which he has shown will add greatly to the student's practical understanding of the subject.

The study of computers must be planned so as to produce not only a good user of computers, but also a good adviser about them. It should cover the full range from hand held machines, through micros, minis and up to the large mainframes. Hardware studies should include electronics and instrumentation because these are so often at the root of data collection in industry. Sadly, I have seen too many statisticians assume that the data presented to them must be correct and they spend time analysing errors which they may be able to eliminate if they understood the plant and transducers which produce peripheral equipment: typewriter terminals, VDUs, printers graphics terminals and plotters. They must learn something about computer packages, interactive statistical analysis, textual storage and retrieval, word processing, computer based learning, and simulation. They should be able to program in several high level languages (Basic, Fortran 77, Pascal and Cobol) as well as in an assembler. They should be familiar with documentation, standards, file handling, and structure of languages and programs. Their practical work should yield fully documented programs, illustrating their mathematical, logical, artistic, creative, technical and communications skills.

GOVERNMENT, INDUSTRY AND BUSINESS

It is surprising that so few courses that prepare students for lives in industry actually equip them with any general knowledge about industry. We should start with a study of government and public institutions, banks and insurance companies, progress through the structures of industry and types of business and the links between government, industry and business. Links with other subjects on the curriculum would be studied through business management, economics, financial planning, accounting, manpower planning, industrial relations, operational research and public relations. This may seem to be rather much, but I do not suggest that these studies, except for operational research, should be in any depth. Rather an understanding of the meanings and relationships of these subjects is intended.

CONCLUSION

Perhaps my proposal for a degree course in polymathy is still sketchy, but if I have succeeded in conveying the intended philosophy we may be able to use this paper as a basis for generating further ideas and a detailed course proposal. The advantages are clear. We should produce graduates who are capable of ferreting out problems at any level in industry, solving them, and communicating the solutions so that they are put to profitable effect. Although our graduates would be good applied statisticians, and better than most because of their communicating abilities and broad understanding of the sciences, they would also be potential leaders of dynamic and enterprising businesses within any industry. Furthermore, our graduates would be educated to enjoy interesting lives and to be paid well for living them.

References

[1] Greenfield, A.A. (1979) Statisticians in industrial research: The role and training of the industrial consultant. The Statistician, 28, 71-82.

[2] Healy, M.J.R. (1979) Does Medical Statistics Exist, BIAS, 6, 2, 137-182.

[3] Hunter, W.G. (1981) Six statistical tales. The Statistician, 30, 107-117.

[4] Kanji, G.K. (1974) The role of the statistical laboratory in the teaching of statistics. Int.J.Math.Educ.Sci.Tech., 5, 53-57.

[5] Kanji, G.K. (1979) The role of projects in statistical education. The Statistician, 28, 19-27.

[6] Kanji, G.K., Gilchrist, W.G. (1981) A degree in applied statistics. The Statistician, 30, 193-198.

[7] Preparing statisticians for careers in industry: Report of the ASA section on Statistical Education Committee on Training Statisticians for Industry (1980). American Statistician, 34, 65-75.

FIGURE 2: Draft Timetable for B.Sc. course in polymath consultancy

FIRST TERM

FIRST YEAR

	Monday	Tuesday	Wednesday	Thursday	Friday
9.00	MATHS	MATHS	MATHS	MATHS	MATHS
10.00	COMPUTING	Library and Information Presentation chemistry	Library and Information Presentation physics	COMPUTING	Psychology Presentation biology
11.00					
12.00	Institutions	Elementary Statistics	Elementary Statistics	Public speaking Interviewing	Statistics Laboratory
2.00	Chemistry Laboratory	Physics Laboratory	GAMES AFTERNOON	Biology Laboratory	STATS LAB.
4.00	TYPING	TYPING		TYPING	COURSE CONFERENCE

SECOND YEAR

9.00	Experimental Design	Experimental Design	COMPUTING	Experimental Design	Experimental Design
10.00	MATHS	MATHS		MATHS	MATHS
11.00	Business and O.R.	COMPUTING	Philosophy of science	Business and O.R.	COMPUTING
12.00	Presentation environment		Presentation of data analysis proj.	Statistics Theory	
2.00	SIMULATION THEORY AND PRACTICE	DATA ANALYSIS PROJECTS	GAMES AFTERNOON	ENVIRONMENT PROJECTS	Data analysis project
4.00	Photography	Graphic Design			COURSE CONFERENCE

THIRD YEAR

9.00	Control theory	Business and O.R.	INSTITUTION ATTACHMENTS	Control Theory	Attachment discussions
10.00	Statistics	Statistics		Statistics	Statistics
11.00	Economics	Current Affairs DEBATE		Project Discussions	Business and O.R.
12.00	MATHS	MATHS		Committees Brainstorming Interviewing	MATHS
2.00	PROJECTS	PROJECTS	GAMES AFTERNOON	PROJECTS	PROJECTS
4.00	EDITING	EDITING		EDITING	COURSE CONFERENCE

FIGURE 2: continued...

SECOND TERM

FIRST YEAR

	Monday	Tuesday	Wednesday	Thursday	Friday
9.00	MATHS	MATHS	MATHS	MATHS	MATHS
10.00	COMPUTING	Electronics theory and design	Computing	COMPUTING	Instrument'n.
11.00		Presentation Psychology	Presentation Electronics		Presentation Psychology
12.00	Institutions	Statistics	Statistics	Public speaking Interviewing	Statistics
2.00	Psychology Laboratory	Electronics Laboratory	GAMES AFTERNOON	Physiology Laboratory	STATISTICS LABORATORY
4.00	WRITING	WRITING		WRITING	COURSE CONFERENCE

SECOND YEAR

9.00	Surveys	Surveys	COMPUTING	Surveys	Surveys
10.00	MATHS	MATHS		MATHS	MATHS
11.00	Business and O.R.	COMPUTING	Business and O.R.	Business and O.R.	COMPUTING
12.00	Statistics		Presentation of data analysis project	Statistics	
2.00	SIMULATION THEORY and PRACTICE	DATA ANALYSIS PROJECTS	GAMES AFTERNOON	ENVIRONMENT PROJECTS	Data analysis project
4.00	Visual aids	Visual aids			COURSE CONFERENCE

THIRD YEAR

9.00	Control theory	Business and O.R.	INSTITUTION ATTACHMENTS	Control theory	Attachment discussions
10.00	Statistics	Statistics		Public Relations	Statistics
11.00	Economics	Current Affairs Debate		Project Discussions	Business and O.R.
12.00	MATHS	MATHS		Committees Brainstorming Interviewing	MATHS
2.00	PROJECTS	PROJECTS	GAMES AFTERNOON	PROJECTS	PROJECTS
4.00	EDITING	EDITING		EDITING	COURSE CONFERENCE

FIGURE 2: continued...

THIRD TERM

		Monday	Tuesday	Wednesday	Thursday	Friday
FIRST YEAR	9.00	MATHS	MATHS	MATHS	MATHS	MATHS
	10.00	COMPUTING	Theory & Art COMMUNICTN	Theory & Art COMMUNICTN.	COMPUTING	Instrumentat'n.
	11.00		Presentation industrial lab	Presentation industrial lab		Presentation Electronics
	12.00	Institutions	Statistics	Statistics	Public speaking Interviewing	Statistics
	2.00	Industrial Process Laboratory	Industrial Process Laboratory	GAMES AFTERNOON	Electronics & Instrumentation Laboratory	STATS LAB
	4.00	PRINTING	PRINTING		PRINTING	COURSE CONFERENCE
SECOND YEAR	9.00	Stats Comp Packages	Stats comp Packages	COMPUTING	Stats Comp Packages	Stats Comp Packages
	10.00	MATHS	MATHS		MATHS	MATHS
	11.00	Business and O.R.	COMPUTING	Business and O.R.	Business and O.R.	COMPUTING
	12.00	Statistics		Presentation of data analysis project	Statistics	
	2.00	SIMULATION THEORY and PRACTICE	DATA ANALYSIS PROJECTS	GAMES AFTERNOON	ENVIRONMENT PROJECTS	Data analysis project
	4.00	Word Processing	Exhibitions Design			COURSE CONFERENCE
THIRD YEAR	9.00	Accounting	Consulting	INSTITUTION ATTACHMENTS	Accounting	Attachment discussions
	10.00	Urban and Social studies	Writing papers, reports		Educational Systems	Government & Industrial Finance
	11.00	Economics	Business and O.R.		Business and O.R.	Communications Survey
	12.00	MATHS	MATHS		MATHS	MATHS
	2.00	PROJECTS	PROJECTS	GAMES AFTERNOON	PROJECTS	PROJECTS
	4.00	EDITING	EDITING		EDITING	COURSE CONFERENCE

TEACHING APPLIED STATISTICS FOR AND/OR IN INDUSTRY

H C Hamaker, Huygenslaan 11, Eindhoven

INTRODUCTION

The quality specifications a statistician who wishes to follow a career in industry has to satisfy have been detailed in a report prepared by a Committee especially established for that purpose by the American Statistical Association [1] and in a paper by A.A. Greenfield [2]. These qualifications are concerned partly with purely personal characteristics, which can not be taught, and partly with the statistical and technical knowledge needed or desirable. With these we will mainly be concerned.

To my mind the requirements listed raise the important question whether any person trained primarily as a physicist, chemist, or engineer with some secondary training in statistics is not better prepared for a career as statistician in industry than a person trained primarily as a statistician.

To assist in planning experiments in industry the statistician has (1) first to diagnose the problem which is often incompletely and/or incorrectly stated, (2) to take into account the technical possibilities and the circumstances under which the experiment will have to be carried out, (3) to issue clear and unambiguous instructions for performing the experiment, and (4) to present the results in a practical form by graphs or tables that can be read by non-statisticians, and are directly related to the problem the experiment meant to solve. In these activities technical insight and practical experience in experimentation are invaluable.

Further a physicist, chemist, or engineer with some basic training in mathematics can easily acquire the essential statistical knowledge from text books and the literature, if need be assisted by the extra-mural courses organised today by universities and statistical societies.

Last but not least, and perhaps most important, a person with practical experience in technology or the natural sciences will be better equipped to see through the idolatry that dominates the statistical science of the day. This will be my main topic.

At birth and in its early youth statistics was nothing but exploratory data analysis. Pearson's set of frequency curves were invented to describe observed populations which did not fit the normal distribution; student's t-test was derived to deal with small series of data in a brewery; and the design of experiments with the F-test had its roots in the complicated problems of agricultural experimentation. But since then the science of statistics has degenerated

into a branch of mathematics, so much so that Exploratory Data Analysis, or EDA, is introduced as something new. Most of the literature nowadays consist in what Sir Claus Moser [3] in his Presidential Address to the Royal Statistical Society has termed "theories looking for data rather than real problems needing theoretical treatment". And the data used to illustrate such theories are then all too often wrested from their context and distorted in order to fit the theory, as some examples to be discussed further on will illustrate.

DESIGNING EXPERIMENTS I

One of the main functions of a statistician in industry is to assist in designing experiments; but sometimes no experiment is needed.

Example 1:
A factory developing a hot iron has trouble with the thermostat and asked for an experimental design. Five influencing factors were listed, but when I asked what was the influence of each of these, varied singly, my clients simply did not know. I therefore advised them to perform some one-factor experiments first; without some information it is not possible to design an experiment on a reasonable basis. My clients never came back because they had soon found out what was the source of their trouble.

Example 2:
An experiment was requested to find the sources of the variability in the "blowing time" of small electrical fuses, which was said to be far too large. The factors to be considered were: (1) the material, (2) the diameter, and (3) the length of fuse wire, (4) the current strength and (5) the type of current (direct or alternating), (6) the diameter and (7) the wall thickness of the glass tube, and (8) the nature of the gas filling (air or an inert gas). But when I asked "How large is the variability in the blowing time?" the answer was "Very large but we do not know more than that". So I wanted to see this first and instructed my clients to blow sets of fuses of one type at different currents and show me the results. The variability turned out to be quite reasonable except at one specific current, where there appeared a two-humped distribution. The physical explanation was evident; at this current the fuse wire was heated to near its melting point; for some that point was reached and these blew almost at once, while others stayed below the melting point and had to blow by evaporation and/or oxidation which is a much slower process. This simply was not known and threw quite a new light on the problem. No further experiments were needed.

Example 3:
A tool used to press small products out of a metal ribbon broke down too often and an experiment was requested to find the reason why? Factors were (1) the material of the tool, (2) the tool maker, (3) the hardening process, (4) the machine setter, and (5) the material on which the tool operated; but the difficulty was that these always blamed the defects on one of the others. I then advised to establish a group with representatives from all the departments concerned and jointly inspect a number of defective tools. It was then discoverd that 80% of the defects were due to one weak spot in the construction of the tools, which had not been realised before.

In such situations little or no statistics are required, only some common sense is needed. This I believe has to be taught by practical experience in the factories.

In teaching the design of experiments it should, in my opinion, be emphasized that the basic ideas originated in agriculture; where one has the spring and summer to carry out the experiment, and the autumn and winter to harvest and analyse the data, and plan another experiment for the following year. Working at that speed it is essential to draw the maximum of information from each experiment, and complicated designs with confounding, split plots, etc. are in order. But in industry the execution of a designed experiment will often not take more than a day, and seldom more than a week. And in that situation such complicated experiments including many factors are not the most effective solution. A series of experiments with only 3 or 4 factors may be much better. This has several advantages; one does not have to play all cards at once; one can gain experience in the subject matter as the series develops; and the results of a 3- or 4-factor experiment are much easier to explain than those of an experiment involving a larger number of factors.

Example 4
The following factors were to be included in a design to determine the visibility of light signals: (1) day or night, (2) the signal is continuous or flashing, (3) blinding of the observer by direct or indirect light, (4) colour: red, green, yellow, white, (5) distance: 100, 300 and 1,000 meters, (6) 11 observers. The responses were decided to be the intensity of the light source when the signal was judged: just visible, just right, and very clearly visible.

A full scale experiment with all factors crossed, yielding in all 3,168 data, had been carried out before I was consulted. Technically this may well have been the only possibility, because a special set-up had to be arranged, and the persons taking part could not be called off their work or off home too often.

An important question in this case would also be what the real purpose of the experiment is? If it is only to find conditions under which the light signals are invariably either easily or very clearly visible, then the data can be thrown into a computer and sifted out in some simple way. But if the purpose should be to express the responses as functions of the factor levels, and if the statistician had complete freedom to arrange the experiment as he likes best, then it might be much better to proceed carefully; for instance by first investigating the reproducibility of the responses taking into account weather conditions and fatigue as well; and then to study the relation between distance and visibility on a modest scale to find a suitable model to be used in further research; etc.

Example 5

A new type of radio tube developed in factory A gave a poor electron emission of the cathode when produced in factory B. In this case it is known at the outset that there exists a very large effect, the only purpose is to find its source, which is likely to be due to one small difference in production. A careful analysis showed that there were 18 steps in the production where the two factories slightly differed. experimentation is now very simple; the 18 conceivable factors can be split up in two blocks of 9. A set of tubes is produced with block 1 carried out in factory A and block 2 in factory B. This will at once reveal in which of these two blocks the source of trouble must be sought. That block can then be split up in the next experiment, etc. It turned out that for some reason the glass bulbs were in factory A rinsed out with water before going in production, while in factory B this was not considered necessary. The chemistry behind this phenomenon need not consider us here.

The conclusion to be drawn from these examples is that the actual planning of an experiment has many more aspects than are taught in text books on the subject. An experiment may have many different purposes, and may have to take account of technical restrictions which nearly always exist. Also the statistician has always to take a critical attitude with regard to the situation as it is represented by his clients, because these look at their problems from their preconceived points of view. For instance, if this is not mentioned by his clients, the statistician should, in the case of the light signals, bring up the problem of weather conditions, fatigue, color blindness, the use of spectacles, the selection of the observers, etc. In other cases he may have to reckon with the possibility of hidden or "lurking" factors, which have not been mentioned but appear to exist by a careful analysis of the data. B.L. Joiner [4] has collected a number of examples.

So much for my own experiences, let us now look at some examples taken from the literature.

Example 6 - taken from N.L. Johnson and F.C. Leone [5, p.135]

The response is the measured viscosity of a certain polymer, and the experiment includes four factors which are all nested within each other : U, V nested within U, W nested within UV, and Y nested within UVW; respectively with 2, 10, 2, and 2 levels. The analysis of the data is presented by the following analysis of variance :

TABLE 1

Analysis of variance of Example 6

	As given by Johnson and Leone			Corrected	
Source	SS	DF	MS	DF	MS
U	0.07	1	0.07	1	0.07
V within U	503.48	9	55.94^{+++}	18	27.97^{+++}
W within UV	454.72	20	22.74^{+++}	20	22.74^{+++}
Y within UVW	109.13	49	2.23	40	2.73

A check revealed that the SS's are correct but the DF's as assigned by the authors are wrong; the MS for V was overestimated by a factor 2 and turns out to be not significant.

But another feature is, that if W is nested within U the MS for U is suspiciously low; $F(20.1) = 22.74/0.07 = 324$ exceeds the 5% level of significance; this might indicate that W is not nested within U. This finally led us to carry out a complete analysis of variance with U, V, and W crossed, and leaving out Y as of minor importance. Result:

Analysis of variance considering U, V, and W crossed

Source	SS	DF	MS
U	0.07	1	0.07
V	488.00	9	54.22
W	12.09	1	12.09
UV	15.47	9	1.72
UW	1.08	1	1.08
VW	432.05	9	48.00
UVW	9.50	9	1.06

TABLE 2

The MS's for V and VW are of the same order which tallies with the view that W is nested within V, and that V itself does not contribute to the variance. But the MS's for U and all interactions within U are an order of magnitude smaller; and this can only be understood if U was crossed with both V and W.

The description of the experiment does not help: "... the original sample was subdivided, for ease of handling, into 2 random subsamples which can be represented by random variables U_1 and U_2. Each of these was then divided into 10 aliquots ($V_1 \ldots V_{10}$). The first step in the analysis of the polymer was then carried out. Each aliquot was again divided into 2 (W_1 and W_2) and the second step of the analysis was performed." But what the steps in the analysis were and how they were carried out is not mentioned.

Anyhow something is definitely wrong; the experiment has been misunderstood, or the data have been misrepresented.

Example 7 - taken from C.R. Hicks [6, p.94]
The object of the experiment was to study the thrust force in drilling at different speeds, feeds, and in different materials. The data for the second material are reproduced below.

TABLE 3

Feed/Speed		100	220	475	715	870	The signs of $R_1 - R_2$				
0.004	R_1	192	136	122	108	136	+	+	+	+	+
	R_2	170	130	85	75	75					
0.008	R_1	386	333	318	472	499	+	+	-	+	+
	R_2	365	300	300	350	390					
0.014	R_1	810	779	810	893	1820	+	+	+	+	+
	R_2	725	670	750	890	890					

Now in presenting these data it is stated that "The order of the experiment was completely randomized", and this is at once belied by the signs of the differences between the replicates R_1, R_2. And what is meant by "completely randomized" taking into account that a drill gets blunted in the course of the operation?

Further it is stated that "the levels of all factors are fixed", which is evidently meant for the pleasure of the statistician, but for a technologist is perfectly meaningless. If further experiments of the same kind are contemplated the engineers will have complete freedom to choose different speeds, feeds, and materials.

Then one of the replicates at the end of the bottom row is, to say the least, suspect as an outlier, which occurs at the highest speed and feed; this might well be due to some instability, for instance when a drill gets blunted.

The data are given as an exercise, the 3 tasks imposed being:
1) Do a complete analysis and state your conclusions,
2) Plot any results which are significant,
3) Set up tests on means where suitable and draw conclusions.

From a technical point of view these questions do not make much sense; a complete analysis requires that we decide what to do with the suspected outlier first; a plot of the data is needed and is practically all that is needed, but plotting only significant results is impossible without plotting insignificant results as well. And the significant effects are visible enough and do not require a test.

Nevertheless such a set of data may be very useful in the classroom provided they are used to discuss what are the conceivable lurking factors that have not been mentioned, but should be taken into account or else left out wittingly. I have, for this specific case, listed a number of questions that might be asked elsewhere and I will not repeat them here: Hamaker [7].

Example 8 - again borrowed from N.L. Johnson and F.C. Leone [5, p.233]
The response is the reflectance of a paint, the factors being different operations in the production of that paint, namely: (1) two liquids in which and (2) two different mills by which the pigment is dispersed, (3) three replicates nested within L and M, (4) two times of aging before the paint is finally mixed, (5) two ways of applying the paint, and (6) two wavelengths at which the response is determined; respectively L,M,R,T,S and W. T,S and W are crossed with L,M and R. The data are reproduced below.

Percentage reflectance of a paint:
L two Liquids, M two ways of Milling, T two aging Times,
S two ways of applying the paint, W two Wavelengths, R three Replicates

			L_1						L_2					
			M_1			M_2			M_1			M_2		
			R_1	R_2	R_3	R_1	R_2	R_3	R_1	R_2	R_3	R_1	R_2	R_3
T_1	S_1	W_1	44.5	44.5	45.0	47.0	50.5	48.0	44.0	44.5	44.0	47.5	48.5	46.5
		W_2	7.5	8.0	8.5	8.5	8.5	8.5	7.5	7.5	7.5	9.0	9.0	8.5
	S_2	W_1	43.5	42.5	44.0	45.5	44.5	45.0	43.0	44.0	43.0	44.5	43.0	43.0
		W_2	10.0	9.5	10.0	12.0	12.5	12.5	9.0	10.5	9.0	10.5	9.5	10.0
T_2	S_1	W_1	45.0	44.0	45.5	47.0	46.5	45.5	44.0	44.5	43.5	48.5	46.5	46.5
		W_2	7.5	7.5	8.0	8.0	8.0	8.0	7.5	7.5	7.0	9.5	8.5	8.0
	S_2	W_1	44.0	43.0	44.5	45.0	44.0	45.5	43.5	44.0	43.5	43.5	43.5	43.0
		W_2	10.5	10.0	11.0	12.0	11.0	11.0	9.0	9.0	9.0	9.5	10.5	10.0

TABLE 4

I first give my own analysis and then discuss that of the authors.

A) At a glance it is noticed that the level of the response is quite different for W_1 and W_2, and that the differences between replicates are considerably smaller for W_2 than for W_1. Hence it is best to treat the reflectances for the two wavelengths as two separate responses. Therefore the averages over the three replicates should be represented as in Table 5.

TABLE 5

Average percentage reflectances for W_1 and W_2

		W_1				W_2			
		L_1		L_2		L_1		L_2	
		M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
T_1	S_1	44.7	48.5	44.2	47.5	8.0	8.5	7.5	8.8
	S_2	43.3	45.0	43.3	43.5	9.8	12.3	9.5	10.0
T_2	S_1	44.8	46.3	44.0	47.2	7.7	8.0	7.3	8.7
	S_2	43.8	44.8	43.7	43.3	10.5	11.3	9.0	10.0

Approximate 95% confidence intervals for these averages are:

± 1.0% ± 0.5%

B) The confidence intervals have been based on the pooled standard deviations derived from the sets of replicates, which are s = 0.86 and 0.44% for W_1 and W_2 respectively; this may not be quite correct, but serves to indicate the order of magnitude.

C) Some significant effects are obvious: the highest reflectances are obtained by the combination M_2S_1 for W_1, and by M_2S_2 for W_2. There may of course be other effects that are statistically significant, but it seems questionable whether these are worthwhile from a technical point of view, and the statistician in charge should first consult with the production manager or research department who requested this experiment; perhaps Table 5 is sufficient for their purposes.

D) Let us now look at the analysis by Johnson and Leone. They see this as an example of a split-plot experiment, the underlying model being stated as:

TABLE 6. The model for the data of Table 4

$$x_{ml\alpha tij} = A + M_m + L_l + (ML)_{ml} + u_{\alpha(ml)} \quad \left.\right\}\text{Main Plots}$$

$$+ T_t + (MT)_{mt} + (LT)_{lt} + (MLT)_{mlt} + v_{\alpha(mlt)} \quad \left.\right\}\text{Sub Plots}$$

$$+ S_i + (MS)_{mi} + (LS)_{li} + (MLS)_{mli} + (TS)_{ti} + (MTS)_{mti} + (LTS)_{lti} + (MLTS)_{mlti} + (u'_{\alpha(mli)} + v_{(mlti)}) \quad \left.\right\}\text{Sub}^2\text{-Plots}$$

$$+ W_j + (MW)_{mj} + (LW)_{lj} + (MLW)_{mlj} + (TW)_{tj} + (MTW)_{mtj} + (LTW)_{ltj} + (MLTW)_{mltj} + (SW)_{ij} + (MSW)_{mij} + (LSW)_{lij} + (MLSW)_{mlij} + (TSW)_{tij} + (MTSW)_{mtij} + (LTSW)_{ltij} + (MLTSW)_{mltij} + (u''_{\alpha(mlj)} + v''_{\alpha(mltj)} + u'''_{\alpha(mlij)} + v'''_{\alpha(mltij)}) \quad \left.\right\}\text{Sub}^3\text{-Plots}$$

with a corresponding analysis of variance as shown in Table 7. The conclusion is drawn that there are strong interactions between L, M and S; that SW, MSW and MW are also significant; and that MT, MLT, MS, LS, MLS, LSW and MLTSW though significant are not very large. But no attempt whatsoever is made to explain what anyone of these interactions means from a technical point of view.

E) The interpretation as a split plot design is based on the assumption that the factors T, S and W not only correspond with a fixed effect but also introduce each a random error. The errors due to S are then nested within T, and those due to W nested within T and S; we should in that case expect the sub-plot error MS's to decrease from sub-plot (ii) to sub-plot (iv); as these are nearly equal the conclusion seems justified that T and S do not significantly contribute to the random error, and that the errors in the determination of the reflectance are the ones that matter.

TABLE 7

Analysis of variance of the data of Table 4 according to the model of Table 6, according to Johnson and Leone

ANOVA—Percentage Reflectance of Phthalocyanine Blue

SOURCE	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARE
Mill (*M*)	1	48.17	48.17***
Liquid (*L*)	1	9.38	9.38*
ML	1	0.26	0.26
Whole plot error	8	8.48	1.06
Residual (*i*)			
Sub-Plots			
Time Aging (*T*)	1	1.50	1.50
MT	1	2.34	2.34*
LT	1	0.26	0.26
MLT	1	2.04	2.04*
Sub-plot error	8	2.35	0.29
Residual (*ii*)			
Sub²-Plots			
Spray or Flow Coat (*S*)	1	0.26	0.26
MS	1	6.00	6.00**
LS	1	5.04	5.04**
MLS	1	5.51	5.51**
TS	1	1.04	1.04
MTS	1	0.01	0.01
LTS	1	0.51	0.51
MLTS	1	0.00	0.00
Sub²-plot error	16	8.50	0.53
Residual (*iii*)			
Sub³-Plots			
Wave length (*W*)	1	30566.34	30566.34***
MW	1	3.38	3.38***
LW	1	0.04	0.04
TW	1	0.00	0.00
MLW	1	0.26	0.26
MTW	1	0.51	0.51
LTW	1	0.09	0.09
MLTW	1	0.00	0.00
SW	1	110.51	110.51***
MSW	1	10.67	10.67***
LSW	1	1.50	1.50*
MLSW	1	0.01	0.01
TSW	1	0.67	0.67
MTSW	1	0.51	0.51
LTSW	1	0.09	0.09
MLTSW	1	1.50	1.50*
Sub³-plot error	32	7.68	0.24
Residual (*iv*)			
Total	95	30805.41	

F) But the split-plot analysis disregards the difference in the error variances of W_1 and W_2, which are mixed up. Moreover, if T and S contribute random errors it is conceivable that the error variances will differ between S_1 and S_2, and between T_1 and T_2. If these error variances are of practical importance they should be determined by experiments specially designed for that purpose, which can be of a very simple type.

G) Another problem is that the reflectance of 50.5% in the top row of Table 4 seems rather high in comparison with the values on both sides, so that it might be suspected as an outlier.

The mean range for the 16 sets of replicates for W_1 is $\bar{R}$ = 13.4%, and the range for the top group under L_1M_2, which is R = 3.5, falls just outside the 3σ limit, which confirms our suspicions.

Further, for the combination $L_1M_2W_1$ we have the following sets of replicates:

				Analysis of variance			After replacing 50.5 by 47.0	
					MS	DF	MS	DF
T_1S_1	47.0	50.5	48.0					
T_1S_2	45.5	44.5	45.0	Treatments	8.61	3	4.18	3
T_2S_1	47.0	46.5	45.5	Replicates	0.15	2	0.44	2
T_2S_2	45.0	46.5	44.0	Residuals	1.51	6	0.44	5

If we reject the value 50.5% as suspicious its replacement values is found to be 47.0%; and, as shown, the replacement has a considerable effect on the MS's in the analysis of variance. Following this trail one step further it is found that by this replacement the residual error variances in the analysis of variance of Table 6

which are	1.06	0.29	0.53	and	0.24
change into	1.11	0.13	0.40	and	0.15

respectively. This will not materially change the conclusions with regard to the significant effects in Table 6, but it illustrates the influence a single suspected datum can have on the error estimates.

H) I also believe that an experiment as that of Table 4 should not in an industrial environment be called a split-plot experiment. Split-plots belong to agriculture where they came from; but for industrial purposes the design illustrated by Table 4 is more correctly described as a full 2^5 design with factors LMTSW and with 3 replicates nested within L and M but crossed with T, S and W; and it is only a split-plot design when T, S and W not only represent fixed effects, but are also considered as sources of random error components.

A STATISTICAL IDIOSYNCRASY

To me the statistical analysis presented in Table 6 illustrates an idiosyncrasy prevalent among statisticians, the chief symptom of which is the canonization of tests of significance. No attempt whatsoever is made to present the results of the experiment in a practical form; the only purpose of the experiment seems to be to test significances, and thereby the problem is considered solved.

One reason for believing that a training as physicist, chemist or engineer might be preferable as a basis for a statistical career in indstry is that this training may render people less susceptible to that idiosyncrasy. And one of the purposes of preparing statisticians for careers in industry should be to inoculate them; or to disenchant them if they are already infected. Here is a list of views they should be made familiar with.

a) Standard deviations σ are never known, except perhaps with simulated data. all we can dispose of is some estimate $\hat{\sigma}$ which we are prepared to use as if σ were known. But even if that estimate is based on 50 degrees of freedom, which is quite a lot, it may still be 10%, and occasionally 20%, in error; and under repeated use of that estimate the level of significance of a test may lie anywhere between 90% and 98%, when we pretend it to be 95%.

b) And this defect cannot be remedied by having recourse to the t-distribution; for that distribution presupposes that every t is based on a fresh estimate of σ ; if we use an estimate s repeatedly in a series of t-tests the level of significance may be as much in error.

Hence with a table for the t-test it should be explained that the underlying assumption is, that the estimate s in t is used only once; and that, if the same estimate is used repeatedly the level of significance is only approximate. Otherwise we are cheating.

Of course the experience shows that this does not matter in the least in statistical practice; the real value of a significance test does not lie in the exact value of the associated probability; it resides in the fact that such tests provide us with simple and universally accepted rules for reaching decisions.

c) The normal distribution should be taught in the same vein, as an accepted convention for arriving at conclusions we are willing to abide by as long as this convention is unchallenged.

d) In the analysis of an experiment every mean square corresponds to an effect, and these effects, and not the mean squares, should be earmarked according to their degree of significance. For example in the case of Table 7 the effects are

Grand Mean ; 27.03, W : -17.84, M ; +0.71, MW : -0.19

and the average reflectances for the combination of M and W are

$$M_1W_1: 27.03 + 17.84 - 0.71 - 0.19 = 43.97\% ,$$
$$M_2W_1: 27.03 + 17.84 + 0.71 + 0.19 = 45.77\% ,$$
$$M_1W_2: 27.03 - 17.84 - 0.71 + 0.19 = 8.66\% ,$$
$$M_2W_2: 27.03 - 17.84 + 0.71 - 0.19 = 9.71\% ,$$

as it can be verified from Table 4.

The mean squares for W and for MW, respectively 3056.63 and 3.38, are both marked as highly significant, but it is very questionable whether the effect of MW = 0.19% has any significance at all from an industrial standpoint; and an analysis of variance giving only the mean squares without stating the magnitude of the effects leaves out the information that matters in the first place.

Also the effect of MW is only about 1% of that of W, and when such drastic differences occur it may always be simpler and more effective to split up the analysis in parts corresponding to the levels of the factor with the largest effect; as we have done in Table 5.

e) The actual planning of an experiment can best be learned in the factory. The situations in a chemical industry, a steel industry, a textile industry, an electronic industry, etc. are all different. If there is a statistical department in the industry new members of the staff can be taken along and learn from their experienced seniors. In Dutch universities it is also a common practice to place students for a training period of about half a year in some industry before their final exam, and this is very helpful. I always made it a point to visit the factory departments that requested a designed experiment and take one of my assistants along with me; you have to see what is going on, for example in order to become aware of possible complications and lurking factors. Effective teaching of the planning of experiments in the classroom is to my mind much more difficult; perhaps a classroom discussion of such a situation as that of Example 4 (visibility of light signals) and all the many possible ways to organise a restricted experiment for a start might be useful.

f) In industrial experiments outliers do occur and have to be reckoned with!! In a recent document on procedures to validate the performance of methods of analysis B. Horwitz [8] states that collaborative test programs (in which a number of laboratories take part), "seem to generate an inherent 5 to 15% outliers; above 20% without an explanation (e.g. failure to follow the method) is ordinarily excessive. The sooner an outlier is investigated, the more likelihood of finding an explanation". In keeping with this statement chemists often carry out their analysis in duplicate, and the same practice may be recommended

for industrial experimentation, at least if the tests are not too expensive. In the case of Example 8 this could have eliminated the uncertainty concerning the one suspect reflectance of 50.5%. With duplicate tests we should then start by investigating the differences between duplicates before embarking on a full scale analysis of variance; or the ranges or standard deviations of series of more than 2 tests. Also it would be advisable as a preliminary study to determine the repeatability of the test method, in order that during the execution of the final experiment a criterion is available by which the occurrences of outliers can be judged while the tests are made.

Significance tests for outliers are of little use. Barnett and Lewis [9] list no less than 43 different tests for a single series: σ and μ both known, one of them unknown, both unknown, one upper outlier, two upper outliers, one or two extreme outliers, etc. This is all due to the exalted importance attached to the 5% and 1% level of significance. Such a variety of tests is, however, only confusing.

From another chapter in Barnett and Lewis it is seen that detecting and testing outliers in the data from designed experiments is considerably more difficult. The case is further complicated because in an experiment of some size there may easily be more than one outlier and we do not know how many.

Nevertheless, outliers must be reckoned with and it is surprising to find that text books on the design of experiments pay no attention to this problem; in many of them the term "outlier" does not even occur in the index.

We can distinguish three types of outliers: (1) those that can be technically explained, (2) those that are so outrageous that there can be no doubt, and (3) the doubtful cases for which there is no technical explanation; these last are the real source of trouble.

The detection of suspect data will in the first place depend on a critical examination of the data or of residuals. In doubtful cases there is in my mind only one reasonable way to deal with outliers. After concluding his analysis the statistician has to draw conclusions and make recommendations. In the face of possible outliers the analysis should then be carried out with the suspected data in place and after rejecting and replacing them. If the conclusions and recommendation are unaffected there is no problem, but if they are materially changed this should be stated so that it can be taken into account.

Anyhow "detecting and dealing with outliers" should form part of the training of statisticians who aspire after a career in industry.

g) On the other hand they should be warned against Bayesian theories. Every successful industrial statistician has to be a Bayesian in the sense that he has (1) to make effective use of past experiences, (2) to consider the costs

of further research, and (3) to take into account the economic consequences of his recommendations. Conceptually there may exist some economic optimum. But as soon as we set ourselves the task of reaching that optimum, we first have to define that optimum, and this invariably leads to unrealistic assumptions and illusive solutions.

In sampling inspection for instance, Bayesian theories have to force past experience into a stationary a-priori distribution, and as costs they only consider the cost of inspecting one item and the losses due to accepted defectives. This is all wrong, because the purpose of sampling inspection is not to prevent the acceptance of defectives but to prevent the production of defectives. A producer who knows that his lots will be inspected will pay more attention to the quality of his products; it is the interaction between sampling inspection and product quality that is important. Bayesian theories may easily lead to solution under which 50% or so of the lots received have to be rejected; a condition utterly impossible in industry, because it will upset the regular production process and stock policy.

Conceptually the problems a statistician has to solve, or assist in solving, may have an economic optimum solution, but this solution must be accepted as lying out of reach. By his work a statistician can help to achieve improvements and with that he has to be satisfied.

h) Last but not least, statisticians should stop fooling themselves and cheating others; and should learn to see applied statistics in its true light.

By producing a standard for sampling inspection by variables based on the assumption that σ is known we are cheating, if we do not specify what we mean by that.

By defining the AQL as "satisfactory as a process average" we are cheating; the probability of acceptance at AQL is only 90% for some of the sampling plans in the Mil.Std.105D, and in that case a producer has to establish a process average considerably better than the AQL to stay out of trouble.

To wait until 2 out of 5 successive lots have been rejected before switching to Tightened Inspection is absurd, at least at the present time. Every rejected lot is a source of trouble and will stimulate some action; and the quality requirements of electronic components are so high, that even a single defective may bring in its train an inquiry as to the source of the defect and the steps taken to prevent it from recurring.

In two standards, ISO 5725 and ISO 4259, the repeatability r and the reproducibility R of a standard method of test are stated to be associated with a probability level of 95%; but as they are only estimates based on a limited number

of degrees of freedom and on a set of assumptions which we know very well can never be quite true, it would be much more honest to state that the associated probability is "of the order of 95%".

In ISO 4259 a transformation is applied to the data which is quite superfluous, and a relation between a set of standard deviations s_i and corresponding means $\bar{x}_i$, which by a least squares fit should be $s = a(\bar{x})^{0.574}$ is arbitrarily changed into $s = a(\bar{x})^{2/3}$, whereby a systematic bias of the order of 60% is introduced in the resulting repeatabilities r. As a statistician I protest, but as a practitioner I am inclined to recognize that, despite these statistical imperfections, the standard satisfies practical requirements. The field of application of the repeatability and reproducibility is too insensitive to bother about such statistical refinements.

Perhaps my views are best expressed by saying that statisticians are too finicky, fastidious, or pedantic about their levels of significance; and this may not only hold for industry but equally well elsewhere.

A MOTTO TO CONCLUDE THIS SERMON

After reviewing the statistical theories of his day Walter Shewhart [10] introduced his famous 3σ limits for control charts in the following words:

"Experience Indicates that t = 3 Seems To Be An Acceptable Economic Value"

There is no mention of a normal distribution !
There is no level of significance attached !
It is recognized that we are at bottom facing an economic problem !
For which we have to be satisfied with an acceptable solution !
Based primarily on experience !
It is in this spirit statistical practitioners for industry should be trained !

References

[1] Preparing Statisticians for Careers in Industry: Report of the ASA Section on Statistical Education Committee on Training Statisticians for Industry (1980) The American Statistician, Vol.34, No.2.

[2] Greenfield, A.A. (1979) Statisticians in Industrial Research: The Role and Training of an Industrial Consultant. The Statistician, Vol.28, No.2.

[3] Moser, Sir Claus (1980) Statistics and Public Policy. Journal of the Royal Statistical Society A, 143, 1-33.

[4] Joiner, B.L. (1981) Lurking Factors. The American Statistician, Vol.35, No.4.

[5] Johnson, N.L. and Leone, F.C. (1964) Statistics and Experimental Design in Engineering and the Physical Sciences, Vol.II, Wiley, New York.

[6] Hicks, C.R. (1964) Fundamental Concepts in the Design of Experiments. Holt, Rinehart and Winston, New York.

[7] Hamaker, H.C. (1971) New Techniques of Statistical Teaching, Review of the International Statistical Institute, Vol.39, No.3.

[8] Horwitz, W. (1982) Outline of Collaborative Study Procedure to Validate Performance of a Method of Analysis. Prepared for the Committee on Collaborative Studies of the Association of Official Analytical chemists.

[9] Barnett, V. and Lewis, T. (1977) Outliers in Statistical Data, Wiley, New York.

[10] Shewhart, W. (1931) the Economic Control of Quality of Manufactured Product. Von Nostrand, Princeton.

PRACTICALS, CASE STUDIES, PROJECTS AND PROFESSIONAL EXPERIENCE: SOME METHODS OF DEVELOPING SKILLS AMONG BUDDING STATISTICIANS

G K Kanji, Sheffield City Polytechnic

INTRODUCTION

Over the last 50 years, in response to the increasing demand for statisticians in research, industry and commerce, the teaching of statistics in the British Academic institutions has expanded greatly. Most Universities, Polytechnics and Colleges have either a department of statistics or a statistics section to provide various courses in statistics.

The demand for statistical education at the academic institution comes broadly from three classes of students:-

a) those aiming at becoming academic statisticians i.e. future teachers of statistics,

b) those aiming to be applied (practising) statisticians (i.e. industrial statisticians, social statisticians, etc.)

c) those merely requiring numerate training.

In addition to the above requirements there is an increasing amount of statistical education in schools.

The theme of this session is "Training of Industrial Practitioners" (i.e. training budding statisticians to be industrial practitioners) and I will therefore try to explore some of the methods available to academic statisticians to help them train as well as teach the future applied statisticians. From the outset this teaching and training of future applied statisticians requires two things. Firstly the provision of an appropriate core of theoretical knowledge and secondly assistance with the development of appropriate skills in order to apply this knowledge fruitfully. In short we have to teach "Applied Statistics". Many people will be asking whether it is possible to train applied statisticians at an academic institution and if so, how? My reaction to these questions will be that it is not possible in the real sense but we may go some way towards this goal. Let us therefore explore how far down the road of the development of skills we may go and illustrate the methods available to academic institutions.

By tradition, academic statisticians over a period of 50 years have developed academic works with great expertise and excellence. It is therefore no problem whatsoever to provide budding statisticians with theoretical knowledge (i.e. to teach) but we encounter difficulties when we try to help them to develop their practical skills (i.e. to train them) to solve the types of problems which exist in various (applied) environments. Part of the difficulty lies in the fact that academic statisticians throughout the world do their jobs so well that

they often find it almost impossible to change. It is ironic that if what they do well is no longer wholly appropriate to the world of which they are part then their very proficiency makes it more difficult for them to change. However, in spite of my unbounded admiration for the academics, I feel that some of us (academics) are not completely fulfilling our roles as educators in an applied subject.

Yates and Healy (1964) discussing about the contents of courses for professional statisticians suggested that the course should be judged on the following criteria:

(i) how far the techniques to which it gives rise are likely to be required by practising statisticians
(ii) how far its mastery contributes to general understanding of statistical reasoning and improvement of statistical skill
(iii) the ease with which it can be mastered.

Speaking about the training of statisticians for economics and business Dean (1964) said that "the students should in their course, be taught not only how to formulate a problem or construct an econometric model, not only how to draw sound conclusions for themselves, but also to assess the findings in everyday language".

Presenting his views about the training of government statisticians Bishop (1964) said that "the relationship between theory and practice, and a feeling for the logic of this bridging process and the usefulness of models, can only be acquired by using crude, real life figures - preferably a lot of them - and hammering out some conclusions, under guidance of course. Ideally syndicate or project work would provide opportunities for team work, discussion of principles, writing up results and perhaps for computer work".

Jeffers (1976) talking about teaching said that "the formal teaching will never be enough to provide the necessary knowledge and experience for an applied statistician. Our teaching must be allied to a form of practical training which builds on, and widens the experience gained through formal education".

Spicer (1964) expressing his doubts about the training of medical statisticians wondered "whether the teaching of statistics is not more heavily weighted by considerations of rigour and mathematical technique than, say, engineering".

At this point it is worth reiterating the suggestions on the content of courses made by the I.S.I. report on the integration of statistics (1980), namely

a) better balance between theory and practice,
b) the illustration and exercise from real life situations in the classroom setting,
c) assistance from outside in the development of case materials which can be introduced into the curriculum,

d) joint programmes to establish the relationship between theoretical training and applications in order to yield case studies for course work, and
e) the development of the special purpose teaching materials to bridge the gap between academic training and applied statistical work.

In another report, prepared by the A.S.A. section on Statistical Education Committee on training of statisticians for industry (1980), recommended

i) inclusion of a wide variety of statistical techniques in the course work,
ii) the provision of experience of real problems through a co-operative programme, and
iii) the development of communciation skills as part of the graduate curriculum.

In describing the philosophy of an applied statistics course Kanji and Gilchrist (1981) pointed out that "the aim of the course is to provide an education for the future applied statistician as distinct from an education in statistics. Thus the aim is not just to impart a knowledge of statistics to a student but to impart the necessary skills that are required to make use of that knowledge as an effective applied statistician".

It is therefore very clear from these views, reports and Kanji and Gilchrist's idea of a degree in applied statistics that skills in application are crucial in the role of the statistician and that he is useless to any employer without some problem-solving ability. Hence we will have failed as educators if we ask the employer to provide all the opportunities for problem-solving skills. In the following sections I would like to illustrate and discuss some of the methods that are possible to practice within the academic framework and which I believe help trainee statisticians to develop the skills they require in order to do their job efficiently, namely practical work; case studies; projects and professional training.

PRACTICAL WORK

We all know that statistics is an applied subject which lives on the information collected from other disciplines such as Biology, Agriculture, Medicine, Industry, Public Health etc. For the analysis and interpretation of such information (data) it is necessary to call for assistance from the theory of statistics which is in turn based on the science of mathematics together with the art of good reasoning. This gives the subject a hazy nature which the student, in the early stage of his statistical studies, finds rather confusing. However, the student new to statistics may be helped to iron out his confusion by conducting a series of simple experiments. The collection of information for and the analysis of data obtained from these experiments offers an experience of some of the difficulties encountered in more sophisticated statistical analyses. By meeting

and attempting to overcome these difficulties, the student gradually builds up his own confidence and ability and it is for this reason that the need for statistical practical work is of prime importance.

In describing the role of a statistical laboratory in the teaching of statistics Kanji (1974) has mentioned that "the greatest difficulty for students in the learning of statistics is the transfer of thoughts i.e. from their basic knowledge of mathematics to basic ideas of statistics". Statistical practicals help this process and are necessary for developing a better understanding of the theoretical aspects of the subject. Such practicals should not of course be seen as isolated events but rather as part of a sequence of practicals gradually increasing in sophistication. Therefore the provision of facilities for conducting experiments, the purpose of which is the collection of data and their analysis, is an important role of statistical education.

There are at least two ways one can present practical problems in the classroom. Firstly by asking students to conduct simple experiments and collect the data and then analyse it. A typical such example, taken from our Statistical Laboratory Manual, is given below.

STATISTICS LABORATORY — Experiment No.32

Statistical Concept - Linear Regression

Approximate time: 30 minutes — No. of students: 2

AIM — The experiment is designed to illustrate the method of fitting a regression line to data generated by the student.

EQUIPMENT — One toy 'hotwheel' type car, a sloping length of track and a tape measure.

EXPERIMENT

(a) Hold the car steady on the line drawn 5cm from the bottom of the track - then let it go. Measure carefully the distance it travels from the end of the track to where it comes to rest.

(b) Repeat this with the car starting on the 10cm, 15cm, 20cm, 30cm, 40cm, 50cm and 70cm lines in turn, making eight readings altogether.

ANALYSIS

1. Summarise your results in the form:
   Distance up the Track (x) 5 10 15 20 30 40 50 70
   Distance Travelled (y)

2. Draw a scatter diagram of the data.

3. If 'n' is the number of readings, we have
   $n = 8$, $\Sigma x = 240$, $\Sigma x^2 = 10650$

4. If x and y are connected by a relationship of the form $y = a + bx$ (the regression line) then:
   $\Sigma y = na + b\Sigma x$ and $\Sigma xy = a\Sigma x + b\Sigma x^2$
   Substitute the values obtained in (3) into these equations and hence solve them for 'a' and 'b'. State the equation of the regression line and superimpose the line on your scatter diagram in (2).

5. Using your regression line, estimate the distance the car would travel from (a) the 60cm line and (b) the 90cm line and then try it out in practice, and compare the agreement between the predicted results and the experimental ones.

State clearly your conclusions as a result of the experiment, commenting particularly on whether or not you feel an assumption of a linear relation was correct.

Secondly, a variety of practical problems which have been collected from industrial organisations, government establishments and research centres could be considered. Some will be relatively short and will be worked by students on an individual basis, others may take several hours or even last for more than a week and will be tackled by groups of students. In some cases the problem may involve only one technique but most will bring together ideas from different parts of a theory course. A typical practical problem of this kind is as follows.

PRACTICAL

Title: Some Problems Concerning Mould Temperatures

An investigation was carried out on a five section double gobbed IS machine to study the relationship between the operating temperatures of the mould and dimensional quality of the bottles produced.

It was anticipated that the rate at which glass is cooled by different parts of a mould during forming would have some effect on the shape and glass distribution in the bottles produced.

The mould surface temperature is known to be one process variable which affects the rate of heat transfer from the glass and has been shown by previous measurements to vary considerably between the moulds on an IS machine.

A land fibre optic pyrometer was used to measure the mould surface temperatures.

Fig.1 illustrates the positions on the blank moulds at which the temperatures were measured. Readings were taken at each of these positions for both the inner and the outer mould. The mould surface temperature measurements are given in Table 1. The wall thickness of the bottles, measured using callipers after annealing, are given in Table 2. This was made on the flat sides of the bottles (Top, Middle and Bottom), as illustrated in Fig.2.

Analyse the data in detail and write a report discussing whether mould temperature may have a useful relationship with the occurrence of rejectable ware.

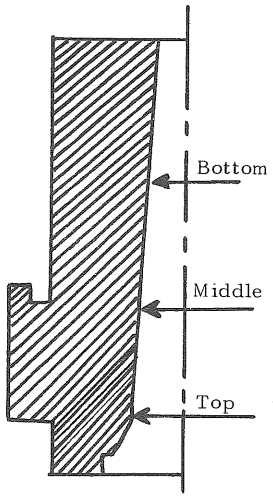


FIG.1 Positions on the Blank Mould Halves at which the temperature measurements were taken.

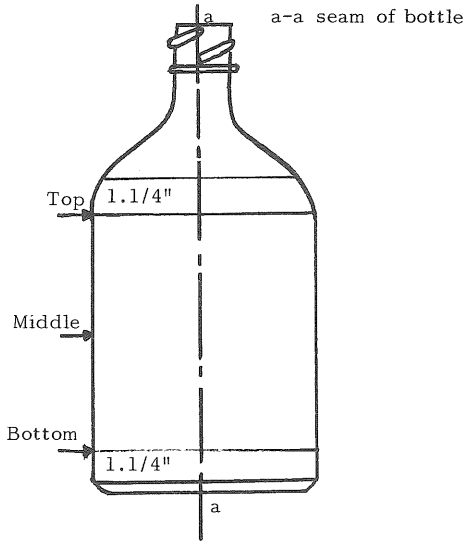


FIG.2 Positions on the bottle at which wall thickness measurements were taken.

TABLE 1

Blank Mould Surface Temperatures in °c.

		Section 1		Section 2		Section 3		Section 4		Section 5	
		LHS	RHS	LHS	RHS	LHS	RHS	LHS	RHS	LHS	RHS
Inner	T	473	484	442	491	426	392	440	442	409	418
	M	521	506	455	497	430	385	491	478	466	430
	B	485	479	449	461	410	382	466	460	452	424
Outer	T	518	517	474	462	417	430	404	409	402	414
	M	547	547	491	489	446	434	449	464	443	449
	B	527	528	468	464	434	410	454	463	436	426

TABLE 2

Wall Thickness Measurements of the Bottles (in inches)

		Section 1		Section 2		Section 3		Section 4		Section 5	
		LHS	RHS	LHS	RHS	LHS	RHS	LHS	RHS	LHS	RHS
Inner	T	.131	.080	.125	.120	.105	.120	.128	.120	.095	.110
	M	.106	.070	.112	.102	.088	.104	.105	.095	.095	.096
	B	.130	.080	.120	.108	.088	.130	.130	.106	.111	.112
Outer	T	.110	.100	.100	.110	.090	.125	.095	.122	.095	.120
	M	.098	.100	.080	.120	.082	.130	.086	.124	.095	.090
	B	.085	.098	.095	.140	.090	.125	.092	.110	.092	.120

CASE STUDIES

The importance of case studies in the study of applied statistics is that they help the student to understand the uses and abuses of statistical techniques and their applications. The most obvious benefit is that they can be used to demonstrate the application of statistical reasoning and techniques in an applied area without having to wait for a long period of time encountered by a real life problem. A further point is that whereas periods of professional experience (which I will discuss later) are restricted to one or two firms and hence in general to a limited number of problems, case studies may be drawn from a wide variety of areas. Thus the future statisticians can be made aware of problems in a variety of different areas including social, business, government, marketing, biometrics and so on.

In describing the use of case studies in teaching statistics Bissell (1975) has pointed out that case studies can be presented in different ways. "At one extreme, the problem and its solution may simply be described by the expert responsible for dealing with it. Whilst this may provide a fascinating (if not bewildering) glimpse of the artist at work, it often fails to involve the audience in the situation. At the other extreme, the problem may be formulated and then tackled by syndicates representing board meetings or management committees. The main difficulty here is that the leader or presenter of the study has to steer the discussion towards a solution that he knows to exist; rather than evaluate possible alternative actions that might have been either disastrous or even more successful than that actually adopted".

In spite of the difficulties I believe that the presentation of case studies similar to the latter method mentioned by Bissell is much more useful to students because it gives them an opportunity to participate and work as a group to solve the problem. It may be worth remarking that this method is the harder of the two for the presenter i.e. for the academic statistician. A typical case

study which could be used by the students either singularly or in a group is as follows.

CASE STUDY

Title : Replacement of a Wearing Part

a) A part on a machine is subject to wear and has to be replaced regularly. As it ages the part tends to break more and more of the items produced by the machine. Thus the part incurs two types of cost - cost of replacement of the part, and cost of breakages. It is required to minimise total cost. Explain how such minimisation can be achieved for each of the following situations.

i) Where the number of breakages depends directly upon the age (in days) of the part.

ii) As (i), but with the condition that the supplier of the part, by way of guarantee, agrees to cover the cost of all items broken in the first 'k' days of usage of the part.

iii) Where the number of breakages appears to be a quadratic function of age (in days) of the part.

In any explanation, denote the cost of replacement of the part by C_R(£), the cost per breakage by C_B(£) and the number of breakages during the t^{th} day from time of installation of the part by n_t.

b) The part to be replaced at regular intervals is available from two suppliers - A, who charges £55 per part for his brand, and B who charges £50 per part for his particular brand. The cost per breakage is 50p. As the result of controlled trials the firm producing the items has available the following data on breakages.

Day from time of installation of part	Breakages during that day Brand A part	Brand B part
5	2	2
10	4	5
15	4	7
20	7	8
25	8	10
30	11	14

Recommend the brand the firm should buy, and how often the part should be replaced in this case.

Discuss the possible effects of randomness in the data on the recommendations.

Suppose the supplier whose brand has not been recommended is willing to offer a guarantee to cover the cost of breakages during the first 5 days of usage of the part. Examine the effect of this guarantee, if any, upon the recommendations previously made.

Note : The total number of items produced each day can be taken as fixed (constant).

c) Suppose the part is available from a third supplier - C, who charges only £40 per part for his brand. Whilst brand C is known to perform well during the early days of its usage (causing few breakages), it is also known to produce more breakages as it ages than either brand A or brand B. The following data concerning breakages is available for analysis.

Day from time of installation of part	Breakages during that day for brand C part
5	1
10	5
15	8
20	14
25	26
30	34

Consider the effects of this information upon any recommendations previously made.

Just as the practical programme allows the student to progress from relatively simple to relatively sophisticated materials so will the case study programme.

PROJECTS

The inclusion of statistical projects in any training programme for statisticians is vital. Many academic institutions are not very keen to introduce project work at the undergraduate level as part of their curriculum, reasons for this unwillingness being the difficulties in generating enough appropriate and distinct projects and the necessary supervision.

In describing the aims of the project Kanji (1979) said that "the project is the most personalised aspect of the course in statistics and aims to provide students with the opportunity to apply and integrate the knowledge gained during the course by an extended study and to develop an organised and methodical approach. It also provides them with the opportunity

(a) to improve their ability to write critically and exercise judgement in a chosen area of statistics,

(b) to exercise their skills in reviewing and condensing published information,
(c) to develop their powers of written expression,
(d) to assist them in the use of the library and the various sources of literature,
(e) to help maintain a continuous thread in the main branches of statistics".

A major aspect of project work is that it makes it necessary for the student to go beyond the scope of the work covered during the more conventional aspects of their course work.

Clearly these aims are interlinked and it is therefore expected that an undergraduate student might take the following steps

(i) understand and define the problem clearly
(ii) organise the work in an effective manner
(iii) extend his/her knowledge through guided reading if appropriate to the development of a solution
(iv) interpret results and draw valid conclusions, and
(v) present his/her work in the form of a logical and coherent report.

It is difficult to put forward a strict rule about the suitability of a project because projects of a fairly broad nature could meet the above aims just as adequately as highly specialised ones. However, emphasis should be put on the maintenance of continuity in one of the branches of statistics and on a thorough understanding of the statistical analysis involved in the project.

A typical example of a project proposal for an undergraduate student might be the following.

PROJECT PROPOSAL

Title of Project : Environmental Impact Analysis

Aims of Project :

A survey was carried out by the Greater Manchester Passenger Transport Executive into the attitudes of households in the Hillcock area concerning the environmental impacts caused by bus services. The residents were asked to indicate the possible environmental impacts which bothered them. The results obtained indicate that though traffic in general is a greater source of bother than bus services in particular, the following aspects of bus services are a source of bother to at least 20% of the residents in the sample in Hillcock.

a) Risk of accident
b) Noise
c) Vibration
d) Dust and dirt

The purpose of this work is to set up a model to investigate the factors (i.e. characteristics of householder) affecting the level of concern within each of the various aspects of bus services.

Sequence of project

1. Survey of relevant literature
2. Familiarisation with the detailed theory of the Log-Linear models and possible modelling procedures
3. Familiarisation with GLIM
4. Fitting a Log-Linear model using GLIM
5. Analysis
6. Conclusions and Report

Resources available - Data and GLIM Manual
Supervisor - G K Kanji

PROFESSIONAL EXPERIENCE

An applied statistics course should be designed to develop the students' understanding of statistics and of associated methods and their ability to apply this knowledge to problems related to the application of statistics in industry in commerce and in government etc. One way to achieve this is to develop a course which includes a period of professional experience.

In describing aims and objectives of professional experience Jones and Kanji (1980) have pointed out that the main aim is to use work experience so that students can set their theoretical knowledge into perspective. The objective is to allow the student to work for an organisation within which he or she is performing a useful and necessary function in a problem solving capacity, thus developing his working skills both technical and personal.

Oakes (1981) has echoed Jones and Kanji's sentiment by describing the aims and objectives of professional training as follows.

1) To integrate theory and practice and to reinforce learning.
2) To enable students to relate their specialism to the wider social context and to bring their critical faculties in a non academic environment.
3) To increase knowledge of subjects outside of the academic course and of the world outside the University.
4) To provide a supplement to or a substitute for, some University practical work and to enable the student to learn and to appraise practical skills.
5) To gain insight into personal and group relationships in the employment situation.
6) To increase the maturity of the student.

7) To clarify and enhance career prospects.
8) To improve the students' ability to communicate orally and in writing.
9) To provide opportunities for the exercise of judgement and improve powers of decision making. Also to be able to meet deadlines.

The professional experience may be regarded broadly as an extension of the continuous assessment of the student into the industrial period. The student is therefore required to write a report at the end of the period and this is read and commented upon by the tutor who made the visit. The purpose of the report is two-fold, one to gain experience of report writing and two to illustrate to the student that all aspects of working life have to be accounted for by some means.

In order to give you some idea about a typical students' activity while he/she is having a professional training I would like to provide you with a very brief summary of a full student's report.

SUMMARY OF INDUSTRIAL PLACEMENT AT CIBA-GEIGY, HORSHAM, JANUARY-SEPTEMBER 1981

This particular student placement was with Ciba-Geigy. It is one of the largest multi-national companies in the world which originates, manufacturers and markets organic chemical products. The Group is made up of affiliates in some fifty countries, the parent company being based in Basle, Switzerland.

The student was involved with the pharmaceutical area of the company at Horsham working with a group of ten statisticians.

The work done in the statistics department is mainly concerned with clinical trials on drugs, in phase three and four, that is licence for studies to be carried out on human beings. The trials are designed by the department in consultation with the medical department and the record forms are also designed by statistics. This ensures that the trial can answer the initial questions set out by the medical department with the appropriate statistics. The design of the trials are usually double-blind with a within patient crossover including a placebo for control. This gives rise to some interesting analysis including analysis of variance, transformations, t-tests, Hotelling T^2, basic statistics such as frequency counts, means, modes, median standard error; curve fitting etc. with which he was involved on most projects.

An interesting part of the students placement was a visit by Professor V. Barnett from the University of Sheffield. He gave a seminar on Outliers in Statistical Data. This showed how the department was not only concerned with pumping out figures, but interested in continuing updating their knowledge.

The computer facilities were very good as the department has linked to the company IBM 4341 computer which gave them the ability to use many different languages or statistical packages such as SAS (statistical analysis system), STATPK, GLIM, BMD through a Librarian system. The statistics department also has two microprocessors with plotter and printer linked to them and also a word processor for the design of record forms.

The two placements with this firm have given the student an opportunity to widen his theoretical knowledge. It has also given him the experience of data handling and analysis using computers. Finally he has gained some administrative experience in preparing data and co-ordinating data preparation.

It is clear that the role of the industrial supervisor, who is after all an experienced industrial practitioner, is crucial.

CONCLUSION

The practical work, case studies, projects and professional experience therefore should show the students how statistical ideas, mathematical techniques, computer usage and communication together with common sense, are all necessary to find a reasonable solution to practical problems. The object is to show that there is often no 'right' answer but that the applied statistician needs to be able to find the best possible method with the given resources bearing in mind the existing constraints. The message that there is often no 'right' answer should become apparent to the students during the continuing case study programme and should be well established by the time they undertake their project.

The author believes that under the present educational system inclusion of practical work, case studies, projects and professional experience in statistics courses at the undergraduate level can bridge the gap between theoretical statistics and real world statistical problems. This in turn will provide the student with the necessary skills that are required to make use of statistical knowledge in order to become an industrial practitioner.

Acknowledgement

I wish to thank my colleagues especially Mr. D.F. Phipps, for their help, comments and suggestions in the preparation of the paper.

References

[1] Bishop, H.E. (1964) The Training of Government Statisticians. Journal of the Royal Statistical Society, A, 127, 211-215.

[2] Bissell, A.F. (1975) The Use of Case Studies in Teaching Statistics. BIAS, 2, 29-32.

[3] Deane, M. (1964) The Training of Statisticians for Economics and Business. Journal of the Royal Statistical Society, A, 127, 216-218.

[4] Jeffers, J.N.R. (1976) Can Applied Statistics Be Taught? BIAS, 3, 17-29.

[5] Jones, J.D. and Kanji, G.K. (1980) The role of professional experience in statistical education. Statistician, 29, 196-203.

[6] Kanji, G.K. (1974) The role of the Statistical Laboratory in the Teaching of Statistics. Int.J.Math.Educ.Sci.Technol., 5, 53-57.

[7] Kanji, G.K. (1979) The role of projects in Statistical Education. Statistician, 28, 19-27.

[8] Kanji, G.K. (1982) Degree with a Difference. In: Proceedings of Concordia Conference on Applied Statistics, Marcel Dekker Inc. (To appear).

[9] Kanji, G.K. and Gilchrist, W.G. (1981) A Degree in Applied Statistics. Statistician, 30, 193-198.

[10] Oakes, M.B. (1981) Evaluating the Learning Outcomes. Proceedings of the Second World Conference on Co-operative Education, Northeastern University, Boston, USA, pp.142-146.

[11] Preparing Statisticians for Careers in Industry: Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Industry (1980), American Statistician, 34, 65-75.

[12] Report of the International Statistical Institute Committee on the Integration of Statistics (1980) International Statistical Review, 48, 139-168.

[13] Spicer, C.C. (1964) The Training of Medical Statisticians. Journal of Royal Statistical Society, A, 127, 219-221.

[14] Yates, F. and Healy, M.J.R. (1964) How should we reform the teaching of statistics? Journal of the Royal Statistical Society, A, 127, 199-210.

TRAINING IN STATISTICAL QUALITY CONTROL

A F Bissell, Abergavenny

INTRODUCTION

The observations on SQC training in this introduction to the Workshop are based on my involvement, both statistically and technologically, in quality achievement and control over a period of over 30 years. My early technical experiences led me to recognise the important role statistics can play in quality operations, and eventually to specialise in statistics rather than in technology. Since that switch, I have been concerned with statistics (including SQC) in manufacturing industry; with the preparation of standards and other "official" documents incorporating statistical methods; with computer or calculator-based systems design for quality or quantity control; and with the training of both statisticians and industrial technologists in industrial applications of statistics.

RANGE OF APPLICATIONS OF SQC

Despite my declared interest in industry and technology, I wish to emphasise that SQC is not a topic of purely industrial application. If I may again quote my own experience, I have recently encountered important uses of what are often regarded as SQC techniques in such diverse fields (other than manufacturing industry) as:-

Control of quality of information from tests in haematology laboratories;
Monitoring of rates of occurrence of defects in ocean-going ships (for insurance purposes);
Appraisal of results of social and other improvements in reducing congenital abnormalities;
Retrospective analysis of accident and sickness rates;
Control of quality of output of clerical and data preparation departments;
Monitoring fuel consumption in buildings, vehicles, etc.

To these, many others could be added, including the monitoring of residuals from experiments or predictive models.

My object in making these remarks is that, in determining the target population for training in SQC, attention should not be confined to those concerned with industry in general, let alone with technology in particular. Some acquaintance with the fundamental philosophy and methods of SQC is useful to anyone who may encounter data collected in some kind of ordered sequence, or with using or designing procedures to discriminate between desirable and undesirable

situations or operations.

TARGET POPULATION FOR SQC TRAINING

In the previous section, I emphasised the breadth of applications of SQC. I now turn to another dimension, that of the position of the recipient of the training on the hierarchal ladder of his organisation.

SQC training has tended to concentrate on those at the lower end of this hierarchy - with "SQC practitioners". My own experience, and that of many authors on the subject, is that the training of SQC practitioners without the active involvement of higher management is virtually useless. The success of certain industrial nations in penetrating the markets of Western Europe and the USA is attributable not only to their productivity (a field where the techniques of SQC can be usefully employed, incidentally) but to their top management interest in quality at all stages from design, through manufacture, to follow-up service.

It follows, therefore, that SQC should form a part of courses in business management, as well as in technology. At present, such business courses emphasise financial and marketing activities. The importance of quality, methods for investigating failure to achieve it, for assessing improvements and for monitoring performance, should be essential topics for all managers - the higher the current or potential status, the greater becomes the need for their interest, direction and inspiration in the area of quality.

Obviously the content of SQC training will vary according to the place of the individual in the organisation. The target population, nevertheless, will include production staff, designers, salesmen, market analysts, accountants, as well as the technical specialist. Many of these are likely, much to their surprise, to find that SQC offers insight and applications in their own areas of activity, in addition to those in connection with the products to which their specialism is related.

I need hardly add, as a rider, that a knowledge of the philosophy and methods of SQC is essential to the statistical consultant, and a course in SQC should have as high a status as one in regression (or experimental design, or multivariate analysis) in graduate statistics courses.

CONTENT OF SQC TRAINING

In view of my remarks on the target population for SQC training, and the consequent variety of needs, both in breadth and in depth, that this will generate, I can only offer here a list of topics which bear upon SQC. In order to promote discussion, I divide them into some broad categories. Thus,

in considering how a trainee statistician might be catered for, we might reasonably assume his familiarity with the aspects listed under Basic Statistics and Statistical Techniques. On the other hand, the trainee technologist may need detailed training in these fundamentals. Again, the trainee manager will not require such a detailed training in the techniques listed under SQC and Reliability, but rather with principles and some examples of applications.

Basic Statistics	Data types and levels of measurement. Data presentation - summary statistics and graphical/tabular methods. Probability and distribution theory - with emphasis on models, not laws. Sampling methods.
Statistical Techniques	Inference, with estimation occupying at least as important a role as significance tests. ANOVA - including its use in estimation of variance components. Correlation - and its proper use in measuring association and dimensionality. Regression and linear models. Experimental design, with emphasis on the non-exclusivity of various designs, and including response surface representation.
SQC techniques	Sampling inspection (attributes and variables). Basis and procedures for control charts. Cusum methods - monitoring and retrospective analysis. Process capability measurement. Sample design for cost-effective control.
Reliability	Probability plotting (in my view, this might be better under basic statistics). Hazard functions and time to failure. Distribution models (Exponential, Gamma, Weibull, Lognormal, etc.) their applications and limitations. System reliability.
Other topics	National and International standards in quality, reliability and statistics. Use of statistical packages and systems. Simulation as a tool for Q. and R. study. Management practice (financial, personnel, technical, production, marketing) vis-a-vis QC attainment.

TEACHING AND TRAINING PRINCIPLES

The target population must be kept in mind when deciding in what manner and in what depth to teach those topics pertinent to the students or course members' needs. For training QC practitioners already working in industry, a much less theoretical approach is appropriate than when training statisticians for applied work in quality control.

Some general points that need to be noted are:-

We must strive to impart understanding of methods, not merely provide recipes.

Statistical methods are based on assumptions and principles which need justifying in the situations and data on which they are used. They also depend on probability and distribution models, which are not laws.

Methods of checking (calculations and assumptions) should be taught, and the importance of operating SQC procedures on regularly updated parameters needs stressing.

Impress on the QC practitioner that:-

Your data costs Time
Money
Effort

Collect it wisely
Analyse it effectively
Use it efficiently

In presenting topics to actual or potential QC practitioners especially if they lack a strong background in general statistical methods, a highly practical approach, with emphasis on visual aids, is appropriate. Many topics which cause difficulty when written as formulae or algorithms (e.g. probability rules, central limit theorem, components of variance, confidence interval estimation) lend themselves to dynamic visual aid presentation. A dash of humour is not amiss to liven up the subject.

TEXTS AND REFERENCES

In an introduction of this kind, it would be impossible to cover adequately the selection of texts and papers for use in SQC training. The choice will depend on the situation - general or "in-house" training? which industry? in what country? at what level? In any case, you will wish to express your preferences during the workshop discussions, and I will not therefore pre-

empt or prejudice these discussions.

A few titles are obvious - general courses are likely to include (at least parts of) Juran's Quality Control Handbook, and perhaps a specific SQC text such as:-

Statistical Methods for Quality Control
Jardine, Mcfarlane & Greensted (Heineman)

Statistical Quality Control Methods
Burr (Marcel Dekker)

Sampling Inspection in SQC
Guenther (Griffin)

Sampling-Inspection & Quality Control
Wetherill (Chapman & Hall)

Appropriate national and international standards will be useful in many cases, and more general statistical texts in some, e.g.

Statistical Methods in Research and Production
ed. Davies and Goldsmith (Oliver & Boyd)

Statistics for Technology
Chatfield (Chapman & Hall)

Statistical Theory with Engineering Applications
Hald (Wiley)

At the appreciation level, "Facts from Figures" by M.J. Moroney (Pelican Books) provides a readable introduction covering a surprising amount of ground.

These are personal choices (at various levels of theoretical treatment) and others will no doubt appear during the workshop.

CONCLUSION

May I be permitted a personal fancy? Some 12 years ago, I was forced by a particular turn of events on a major project to take an interest in Markov Chains - a topic I had, till then, dismissed as of little relevance in my activities. Although the project was only vaguely connected with quality, I discovered I could then better appreciate certain papers in the literature which invoked Markov Chains in SQC techniques. The insight which this appreciation provided into the underlying unity of many SQC procedures, and the power of the Markov Chain representation as a tool for analysis of properties of SQC procedures, has convinced me that they should appear in courses for anyone who will be involved in SQC or Reliability theory, and any statistician who may be

drawn into SQC or Reliability applications. With this possibly controversial suggestion, I conclude this introduction and offer the foregoing thoughts as a starting point for workshop discussion.

DISCUSSION OF PAPERS BY GREENFIELD, HAMAKER AND KANJI

A F Bissell, Abergavenny

The papers presented in this session deal with very different aspects of the training of industrial practitioners, but I detect three threads common to all three contributions. These are (not necessarily in order of importance) firstly that the industrial practitioner must be familiar with a broad spectrum of ideas, methods, techniques - some of them not generally included in conventional statistics courses, some quite outside the usual syllabus but nevertheless relevant. Secondly, statistics is an essentially practical subject, and this practical nature must be emphasised in training its practitioners. Thirdly, and closely linked to the previous point, the statistician must be able to communicate, at all stages of any project, with his clients, and the potential practician needs training in the various forms of communication currently available.

It is in the methods of imparting or cultivating these ideas that the speakers differ (though they do not fundamentally disagree). Professor Greenfield would teach a very wide variety of subject matter to young people about to embark on a career. Dr. Hamaker would, broadly speaking, take experienced practitioners in scientific disciplines and make statisticians of them (the route by which I came to statistics, incidentally). Dr. Kanji advocates learning by the experience of others, and by simulated experience, as part of a broad statistical training. The various approaches are not mutually exclusive. Each, whilst having the merits that the speakers have indicated, raises its own problems, and I now turn to consider each paper in order of presentation.

The list of subjects, topics, talents, techniques that Professor Greenfield presents is impressive, and one must agree that they are all relevant to the "compleate angler" trying to wrest all he can from the stream of data in which he wades. One might even add to the list - accounting methods, company law, human relations, etc. The danger seems to me that in order to squeeze these useful items into a degree course, we must sacrifice something else, either by omission, compression or superficiality. I wonder also whether the selection of (presumably) 18-20 year-olds for such a course can be effective - is this not too early to recognise potential "super-stats", let alone real polymaths? There would be problems at the end of their training too - we have already witnessed the failure (to reach the echelons of top management where the real decisions are made) of management trainees from business schools, and perhaps of O.R. scientists too. In an economic climate where only cash flow and return on capital are regarded as meaningful data, will the young polymath fare any better? Or will they, like their predecessors in numerate sciences, be dubbed with whatever

is the topical derogatory term synonymous with "whizz-kid" or "boy-wonder", accused of having knowledge but not wisdom?

I agree that Britain, and other nations, needs polymaths, but is it perhaps the mature individuals, whom Dr. Hamaker recognises as potential applied statisticians, that we should be training in the way Professor Greenfield suggests? It would be essential to catch them at the right age - I suggest that they may not be sufficiently mature until 30-35, yet by 50 they will, with present industrial attitudes which seem likely to persist indefinitely, become candidates for so-called early retirement.

Dr. Hamaker sets less ambitious targets for his scientists or engineers. He aims at producing, or converting other practitioners into, good practical statisticians. I would question whether they can acquire an adequate statistical knowledge from text-books or extra-mural courses. Indeed, Dr. Hamaker's criticisms of some reputable text books could be construed as an implication that this is a route beset with dangers. A properly integrated course is essential to avoid the very dangers that Dr. Hamaker has pointed out. Statistics learned (or picked up) in an ad hoc fashion is likely to be compartmental, and will lack the depth of understanding necessary to avoid pitfalls like those in the examples quoted. Suitable structures already exist - the Institute of Statisticians examinations, for instance or the Sheffield Polytechnic course might be condensed to provide an M.Sc. or Diploma for mature students from indsutry. In the U.K., at least, this would raise problems of financing suitable candidates, and might need modifications to regulations in order to offer entry to those with the right experience but lacking formal first degree qualifications.

Dr. Hamaker, like Professor Greenfield, makes the point that statisticians should challenge the data they are presented with. I support and encourage such numerate scepticism, and would extend it to the questioning of foregone conclusions from previous investigations from which new projects often develop. For example, I recall being consulted about a specification which required that the ratio of two variables, say A/B, must be controlled to 1.25 ± 0.04. This was based on some earlier work which purported to demonstrate how critical A/B was. Elsewhere in the same report was a record of subsidiary investigations from which it was evident that A and B could each be measured only to the nearest 2% - I leave you to consider the implication!

The first two papers dealt with what we may call the strategy of training industrial practitioners. Dr. Kanji has covered this elsewhere, not only in his papers in 1981 (with Gilchrist) and 1982, but in the actual development of the Sheffield Polytechnic course. In the present paper, he turns to tactics, and in particular the provision of vicarious or simulated experience. I agree whole-heartedly with the use of case studies, syndicate work and data generation equipment

in statistical teaching at any level and for any field of application. The case study evidently needs to be oriented to the particular field of application, but equipment like the hot-wheel car can be used on any course. My version of the car-and-slope experiment is even simpler - a marble, a chute and a strip of carpet! One of the beauties of such demonstrations or data-gathering exercises is that the students can see the relevance of mathematical models in statistical work. In this experiment, if the car (or marble) is placed at the foot of the slope, its rolling distance is zero. Should we therefore constrain the constant term in the model to zero? And if so, is conventional least-squares the proper way to estimate b? I offer another example I use in teaching experimental design and analysis. Three participants tackle three jig-saw puzzles of differing sizes, timing their efforts, to assess whether the puzzles vary in difficulty. A Latin Square experiment takes only 15 minutes to execute, and provides data which illustrate, inter alia:-

(i) the use of a model to estimate effects and residuals;

(ii) the equivalence of a rule-of-thumb ANOVA to the analysis of a linear model;

(iii) the value of replication - three more participants, i.e. two replicates of the square, will quadruple the residual degrees of freedom and vastly improve the power of the experiment;

(iv) the dangers of using rule-of-thumb methods - in fact, a multiplicative model is more appropriate, leading to conventional analysis via log transformation;

(v) the necessity to re-interpret an analysis based on a transformation in terms of the original dimensions. Here, anti-logs of effects become "effect ratios" around a geometric mean.

Those interested in further examples of simple demonstration models will find a selection in "Statistical Teaching Aids" published by the Institute of Statisticians (Bury-St.-Edmunds, England).

I should like to take up a few points of detail in the three papers. Professor Greenfield's example of the restrictive view of the needs of statisticians in pharmaceuticals carries generally into the industrial field (and no doubt elsewhere). Indeed, the draft programme for this conference session lists SQC, Acceptance sampling and reliability as topics in industrial statistics:- my own experience is that there is (for the industrial statistics consultant) some of the first and third, little of the second, but much of the experimentation (not just experimental design), data sleuthing and effective data presentation.

I sympathise with Dr. Hamaker in his views that over-emphasis on "significance probability", rather than formal acceptance or rejection of hypotheses, is a move in the right direction - a procedure which some of us have been using furtively for years! I am less sure about the over-abundance of test methods. Certainly every statistician, industrial or otherwise, needs to know about robust and non-parametric methods.

Perhaps Dr. Hamaker is too emphatic about Bayesian methods. There is sometimes meaningful prior information which can strengthen inference about (for example) a regression coefficient or failure rate or mean life and I agree we must teach caution in the use of priors, especially when real data is sparse and the prior information may swamp it.

Do I detect, in Dr. Kanji's paper, a degree of disagreement with Professor Greenfield when he says that "it is not possible to train applied statisticians in the real sense at academic institutions"?. He makes this point although the Sheffield Polytechnic Course has an industrial experience layer in a 4-year sandwich (Professor Greenfield's is a 3-year non-sandwich course proposal).

All three papers give us much to consider, and (in the workshop sessions) to discuss and even argue about. I look forward to taking up in more detail one aspect of the training of industrial practitioners in Thursday's workshop on quality control.

DISCUSSION OF PAPERS BY GREENFIELD, HAMAKER AND KANJI

D G Watts, Queen's University at Kingston, Ontario

Mr. Chairman, speakers, and ladies and gentlemen. I would like to preface my discussion by thanking the Chairman for inviting me to participate in this session, and for organizing a very interesting and stimulating period. It is a pleasure to be here and to be involved in this important first International Conference on the Teaching of Statistics.

We have had the privilege of hearing three well-written and well presented papers today. Professor Hamaker's paper was concise, with an excellent summary, demonstrating the good communication skills recommended by all three speakers. He raised many good points, most of which I agree with, a few I didn't understand fully, and one with which I quite strongly disagree. Professor Greenfield's paper I found interesting and with some neat innovative ideas, but rather imprecise to serve as a blueprint for curriculum design. I will make some specific suggestions on that aspect later. Professor Kanji's paper was the most limited of the three, and offered some persuasive arguments for inclusion of projects in statistical training.

There were some obvious agreements of opinion in the three papers, the two most important to me I describe as

1) communication abilities, and
2) the statistician as scientist.

I strongly identify with both of these.

Some of you may remember a Conference on the Future of Statistics held in Madison, Wisconsin in 1967, the proceedings of which were published by Academic Press. One of the major themes of that conference turned out to be communication, not because of design, but because of its INTRINSIC IMPORTANCE TO THE DISCIPLINE. I commend those proceedings to you for some excellent reading on this important topic, which is now receiving a more appropriate share of attention. One thing you may be struck with is that not much progress has been made. Perhaps with enthusiasts like Professors Greenfield, Kanji, Hamaker and Snee and others, and hopefully influential reports like the ASA Section on Education reports on Training of Industrial and Government Statisticians, and the ISI report on Integration of Statistics, progress will be accelerated.

I suggest statisticians are going to continue to have trouble getting communication skills into the curriculum for a couple of reasons, however: many, if not most academic statisticians have no real experience in writing such reports, and therefore will resist - because it will be hard for them to come to grips with

(e.g. marking is very time consuming, albeit rewarding) and because they don't appreciate the importance of these skills and therefore will not accord them a high enough priority in the curriculum: a manifestation of the "it can't be important if there is no equation on the board" attitude. To effect this change, then will require assistance from outside - from the consumers of our product - by demanding these skills.

One of the things that is being recognized (or discovered) by statisticians is that communication abilities are skills which can be learned, but that they must be learned by DOING. Report writing, presentations, and interpretations must be practised, ...and practised, ...and practised. Professor Greenfield suggests some nifty ideas to make this more enjoyable, through a weekly newspaper and a monthly magazine, and through a course conference. Another great way to encourage report writing and analysis interpretation skills is through projects and real experiences, as recognized by Professors Kanji and Hamaker, although I wish one could find more stimulating and relevant projects than the example of hotwheel regression or one we use at Queen's which measures rebound height of a dropped golfball. The golfball example at least has the virtue that a transformation of the response should be made to stabilize variance and subsequently the same transformation should be made to the drop-height to reinstate a straight line relationship.

One of the best ways to improve communication skills, in a nicely motivated way, is to have the students do their own experiments or to have them assist someone on the campus - projects, as recommended by all three speakers. I have had very good success in this, by having 3 or 4 students form a group and then find a researcher or graduate student on campus who needs to analyze some data. Each group must prepare PROPOSALS on at least two projects which I then use for evaluating the project as regards suitability for the course and viability. The students then use the project throughout the course to motivate them and as a learning agent. They must prepare a final report for submission to me for grading (35 - 50% of course mark) and to the researchers for their use.

The second issue which was strongly evident in all the papers was that of what I term the statistician as scientist, or what Professor Greenfield refers to as the polymath. One of the great advantages I believe my engineering training gave me over my statistical colleagues was that it took the STATIC out of statistics. That is, most statisticians are steeped in theory which assumes no dynamics - the independent part of I I D.

Three courses from my undergraduate program which I have found, and continue to find, particularly valuable, were very much concerned with problem solving. The first was a first year Civil Engineering tutorial course given in conjunction with the calculus course, in which we had to set up the differential equations

for given physical situations, then solve the equations and finally write up the results in a workmanlike fashion. The second was a third year mathematical physics course, again where we developed mathematical models for various physical systems and solved the equations, very often using approximate methods. The third was a fourth year Electrical Engineering course dealing with the setting up of differential equations for dynamic systems, and again solving them. The emphasis was on SETTING UP, not solving. I have often felt that these three courses have been the most useful of my undergraduate courses, and this seems to concur with Professors Greenfield's and Hamaker's contentions.

Another advantage I believe my engineering undergraduate training gave me is something which is implicit in what Professors Greenfield and Hamaker say, and that is a different PERSPECTIVE from liberal arts students. Crudely, one can shoot the breeze with medical people or engineers more easily after having been faced with many of the same difficulties, which engineers refer to as "The perversity of inanimate objects". I also had to do many experiments and write them up.

I disagree with Professor Hamaker that Bayesian methods are to be avoided because "they make unrealistic assumptions". As an engineering statistician I find the Bayesian approach much more realistic and sensible than the sampling theory approach, and much more intellectually and scientifically honest. Bayes doesn't cause you to make unrealistic assumptions, it makes you ADMIT YOUR ASSUMPTIONS.

Which brings me to another perspective on perspective: one thing about being a scientist statistician (one identifying attribute being that you continue to consult and collaborate wherever you are) is that you are forced to adopt a decent perspective about assumptions and about analyses. For example, one realizes that many of the problems academic statisticians worry about are of about tertiary importance compared with the gross assumptions and simplifications one must make to formulate most engineering or medical problems.

I was a bit surprised about the comments on computing. Professors Hamaker and Kanji gave them no or little import, while Professor Greenfield recommended considerable programming skills in several languages. I disagree with all three speakers in this instance, because it seems to me that computer packages are becoming more useful and more powerful, and hence statisticians should not have to be adept in programming in several languages. We certainly need to be able to USE computers EFFECTIVELY, of course, so computing must be in the curriculum, but as "solution facilitators" I think we should be primarily computing MANAGERS, not computing PROGRAMMERS.

In conclusion, I would like to mention curriculum design, which Professors Greenfield and Kanji discuss. I believe that both these gentlemen and their respective Departments and students, - and we in other departments - would benefit if they would apply modern educational techniques of curriculum design in the style propounded by Robert Mager, especially in his book with Kenneth Beach Jr. "Developing Vocational Instruction", Fearon Publishers, California. The technique is very systematic and involves first writing a job description, which is "a general statement about what a person on the job does, and tells something about the conditions under which he does them". Next is the task analysis which requires "listing all the tasks that might be included in the job ..... and then the steps involved in each of the tasks on the list in terms of what the person does when performing the step". This is followed by development of prerequisites and course objectives, for example as applied to regression analysis in my paper "A Task-analysis Approach to Designing a Regression Analysis Course", The American Statistician, May 1981, pp.77-84. This approach is a marvelous one for clarifying exactly what should be IN a course as well as what should be OUT; it gives a wonderful operational perspective.

TOPIC 15: PRINCIPLES OF LEARNING PROBABILITY AND STATISTICS

ORGANISER: PROFESSOR E FISCHBEIN

LEARNING FROM EXPERIENCE - BAYES THEOREM: A MODEL FOR STOCHASTIC LEARNING SITUATIONS?

H Steinbring and G von Harten, University of Bielefeld

INTRODUCTION: THE BAYESIAN APPROACH - AN INTEGRATED BASIS FOR MULTIPLE DEMANDS?

One of the most difficult, and presumably most far-reaching, questions for stochastics education is that of the relationship between probability theory and statistics (Cf. Dörfler/Fischer 1981). From this perspective, it must be welcomed that at least one of the "Invited Paper Sessions and Workshops" within the frame of the First International Conference on the Teaching of Statistics will be devoted to the problem of "Principles of learning Probability and Statistics".

In his well-known book "Mathematics as an Educational Task", Freudenthal has warningly described the negative effects the separation and mutual dissociation between probability theory and statistics may have on mathematics education. Without the reference to the mathematical concepts and interpretations from probability theory, statistics can be reduced to a collection of cooking recipes. Cut off from its real applications, and practised purely within the framework of set theory, combinatorics, or measure theory, probability theory becomes a formal mathematical system. Summarising this comparison, Freudenthal draws the following conclusion: "We have given here a rough sketch of two kinds of teaching degeneration: probability as an abstract system cut off from reality, and on the other hand a system of patterns of computations to be filled out by numerical data. They look like opposites of each other but they are complementary in a most natural way!" (Freudenthal, 1973, p.591).

To oppose probability theory to statistics along the dividing line between mathematical theory and its applications also has a negative effect on the educational conceptions of teaching stochastics. Above all, this split can be rediscovered in the fundamental concept of stochastics, i.e. in the probability concept itself. The question is whether probability in mathematics instruction is to be introduced starting from applications, i.e. from statistics by means of the concept of relative frequency, or whether it is more appropriate to take the path via ideal gambles and thus via the classical conception which stresses the theoretical-combinatorial aspect of probability - a "wrong" alternative in our eyes.

This indissoluble tension between probability theory and statistics - "Probability theory is so close to its applications that it can even boast a theory how it is applied, I mean mathematical statistics". (Freudenthal, p.610) - has

led, in the historical development of probability theory as well, again and again to debate and to attempts at founding this theory anew which from the very beginning strove to end this separation.

Thus, for example, the renowned probability theorist Richard von Mises attempted to reach this goal by means of axiomatising principles of experience, the so-called primordial phenomena (Urphänomene). Accordingly, he thought that the distinction was not to be made between probability theory and a theory of its application.

The Neo-Bayesian approach also pretends to formulate a unified point of view integrating both theoretical and empirical aspects which ends the separation between probability theory and statistics, or at least strongly relativates the latter. In the following, we shall discuss and analyse this point of view in more detail.

If we try to imagine such a point of view with regard to stochastics, we shall at once get into a situation which is typical, as a whole, for stochastics education. We are talking of a fundamental problem which is mainly relevant for mathematics instruction, and for the mathematics teacher: stochastics works with a variety of models (cf. Freudenthal's statement: "No scientist is as model-minded as the statistician."). Didactics, too, makes use of this promising model aspect. Just think of Engel's abacus, of the urn model, also of dice, coin-tossing and the wheel of fortune (which are objective representations of models). Other forms of representation, too, reflect aspects of stochastic models: attention may be called to tree diagrams, even set calculus, to other forms of graphs, and also the formulas used in connection with stochastics. Stochastics involves a whole range of most different types of modelling: starting from concrete models in objective form, across representations of these models getting ever more abstract, up to representing their mathematical form by means of fundamental equations and formulas. Within the context of such models and their various modes of representation, the attempt is to provide a most simple mathematical calculus, a basic and effective technique. As opposed to this technical simplicity of modelling, there are, however, the many problems of interpreting these models and of applying them to random situations. It is mainly this interpretation and application of these ideal models which constitute the fundamental problem: in the first place bridging the gap between the models, their forms of representation, graphs and formulas on the one hand, which are themselves at first given as fixed in their mathematical structure, and which basically function in a deterministic way, and real chance events on the other, concerns the aspect proper of stochastics, and may provoke stochastic thinking.

The mathematical calculus underlying the Bayesian approach is also of simple structure. Here again, the problem is not situated in mathematics, but in this model's interpretation, in the way it is applied. It is seen that such an approach which claims for itself tohave overcome the separation between probability theory and statistics is just as well affected by these problems in the same way, on a higher level. These issues concerning the relation between the mathematical form of representation and its interpretation, between simple theoretical models and their effective application, are, in the first place, crucial for learning stochastics, and for developing stochastic thinking.

The Bayesian approach shall serve us in this respect: which are the advantages and drawbacks of such a model, mainly with regard to a better understanding for stochastics education? This question refers to two different aspects: to both the exemplary use of such an approach in instruction, and to the benefits and more comprehensive knowledge resulting in principle from such a didactical way of phrasing the question.

Accordingly, and from an educational perspective, this is not simply a matter of deciding whether the Bayesian approach is right or wrong - i.e. of contributing to the controversy about the subjective and the objective concept of probability - but rather this debate is intended to help us obtain new insights into the relationship between probability theory and statistics which are not confined to stating either a strict separation, or an identity, between probability theory and statistics.

BAYES THEOREM AND "LEARNING THROUGH EXPERIENCE"

It is remarkable that the theorem from which the opposed groups of "Bayesians" and "Antibayesians" derive their names is rather simple from a mathematical point of view. The crux of this controversy is not about the question whether this theorem is mathematically correct, but about the problem of its application and interpretation.

Reverend Thomas Bayes had wanted to tackle the following problem: "Given the number of times in which an unknown event has happened and failed: Required the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named." (Bayes, 1763/1958, p.298).

His friend Mr. Price, who published Bayes manuscript after his death, thought that Bayes had mathematically solved the following problem: "Common sense is indeed sufficient to shew us that, from the observation of what has in former instances been the consequence of a certain cause or action, one may make a judgement what is likely to be the consequence of it another

time ..." (Bayes, 1763/1958, p.296).

Loosely speaking, Bayes has solved the problem of estimating from an observation the probability of its causes. For a finite number of causes, Bayes' theorem has the following form:

If only the causes E_1, E_2, ..., E_n are possible and if ω_i is the "a priori" (i.e. before observation) probability of the cause E_i, and further if p_i is the "a priori" probability of an event A in case the cause E_i is given, then the probability to be estimated $p_i = P(E_i|A)$ is proportional to $\omega_i \cdot p_i$, that is

$$(1) \quad P(E_i|A) = C \cdot P(E_i) \cdot P(A|E_i)$$

Summing over i yields :

$$C = \sum_{i=1}^{n} P(E_i) \cdot P(A\,E_i)^{-1}$$

and hence this is Bayes' theorem :

$$(2) \quad P(E_i|A) = \frac{P(E_i) \cdot P(A|E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A|E_i)}$$

Let us look at the following application:

(3) Given three urns with white and black balls. The probabilities of drawing a white ball out of one urn are $p_1 = 1$, $p_2 = \frac{2}{3}$ and $p_3 = \frac{1}{3}$, respectively. The urns are equally probable.
We have chosen an urn and drawn a ball: it is white. By (2) the probability that the ball belongs to the first urn is:

$$P_1 = \frac{1}{1 + \frac{2}{3} + \frac{1}{3}} = \frac{1}{2}$$

This application of Bayes' theorem is undisputed. The result can even be verified experimentally.

Let us look at another example:

(4) We choose at random n+1 points P_o, P_1, ..., P_n in the interval [0,1]. The point P_o chosen first has, like any other point, a uniform distribution. Knowing its position p, the event that any particular one of the other points P_1, P_2, ..., P_n falls left of P_o will have the probability p, independent of the others. If p is not known, the probability that k out of the other n points fall to the left of P_o is $\frac{1}{n+1}$ for every k, because P is equally likely to be any one of the n+1 ordered points. If we assume that we know there to be k out of the n points to the left of P_o, then the probability of a success with the next point (i.e. that P_{n+1} falls to the left of P_o) chosen at random

is given by $\frac{k+1}{n+1}$, because the n+1 points divide the interval into n+2 pieces, k+1 of which are to the left of P_o. The probability distribution of P_o, which is initially uniform, is no longer such if we know that out of n random points k have fallen to the left of P_o. It is, instead, the beta distribution $f(p) = K.p^k(1-p)^{(n-k)}$, because P_o is then the (k+1)th point from the left of n+1 random points.

This example can be seen too as an application of Bayes' theorem, and the problem is very similar to another solved by Bayes in his article. What we require is a continuous version of Bayes' theorem:

Take an "a priori" distribution (for the sake of simplicity we take a distribution with a density) of the parameter p i.e. f (p). A is an observation. Then we have:

$$f_{\pi|A}(p) = \frac{f_{A|\pi}(p) \cdot f_{\pi}(p)}{\int_o^1 f_{A|\pi}(p) \cdot f_{\pi}(p)\, dp}$$

If the values of p are "a priori" equally probable (i.e. f (p) = 1) we get:

$$f_{\pi|A}(p) = \frac{p^k \cdot (1-p)^{n-k}}{\int_o^1 p^k (1-p)^{n-k}\, dp}$$

In our example, we have to estimate the expectation i.e.:

$$E = \frac{\int_o^1 p^{k+1} \cdot (1-p)^{n-k}\, dp}{\int_o^1 p^k \cdot (1-p)^{n-k}\, dp} = \frac{k+1}{n+2}$$

In classical statistical analysis, the probability p, for instance, that a die will show a six, is viewed as a constant, unknown but fixed. Bayes' theorem now opens the possibility of a radically different approach, i.e. to consider the unknown value of the parameter p as a value of a random variable.

For the above problem, everybody will find this approach justified. In this instance, there is no controversy about the "a priori" distribution of the parameter p; it is given naturally in the problem.

But consider the following problem:

(5) I have a coin. I hold it in my hand and I am relatively convinced that this coin is fair. But I also toss the coin one hundred times. My "a priori" distribution (the density) of the probability of a head would, for example, have the following form:

$$f_{\pi}(p) = \frac{(2m+1)!}{(m!)^2} \cdot p^m \quad (1-p)^m$$

This distribution shows a maximum at $p = \frac{1}{2}$.

I toss the coin one hundred times and it turns up heads for k = 48 times. Now the probability for heads in the next toss (for m = 10) would be:

$$p = \frac{59}{121} = 0.4876$$

But if I had chosen the following "a priori" distribution

$$f(p) = \frac{101!}{99!} p^1(1-p)^{99}$$

I get according to Bayes' theorem:

$$p = \frac{50}{202} = 0.2475$$

My probability distribution would get better if I continued to toss, but for 100 tosses, these two estimates are rather different.

Another important dispute arose about the class of problems where probability can be applied.

(6) Suppose you meet a stranger in a bar who offers to toss a coin to decide who shall pay for the drinks. He will pay if the coin falls tails, you will pay if it falls heads. The thought crosses your mind that the coin may be two-headed, but you are too delicate to ask to examine it. What is the probability of the coin being two-headed?

Some people may say that the question is meaningless, whereas a Bayesian needs only ask himself what is his personal "a priori" probability of the coin being two-headed.

This may be π; thus the probability of "coin is fair" is $1-\pi$.

Suppose the stranger tosses the coin and it falls heads. Now by Bayes' law, you have to adjust your probability in the following way:

$$P(\text{coin is two-headed} \mid \text{heads}) = \frac{\pi}{\pi \cdot 1 + (1-\pi) \cdot \frac{1}{2}} = \frac{2\pi}{1+\pi}$$

The following table shows how the probability of "coin is two-headed" after the toss (of heads) depends on your "a priori" probability:

prior	0	0.01	0.1	0.5	0.9	0.99	1
posterior	0	0.0198	0.0182	0.667	0.947	0.995	1

Note that if you are certain before the experiment that the coin is two-headed (π=1) or that the coin is fair (π=0), the experiment will not change your probability. This is generally true for any statement.

J. Berkson has pointed out the problem of the range of application of Bayes' theorem by means of the following example:

(7) A lady in the supermarket approaches the meat counter. "Here", says the meat man, "comes the lady who has bought 10 pounds of flour every Monday for the last five years; I'll bet she does it again". The lady has mistakenly put four instead of five 2-pound packages in the paper bag. The man places it on the scales which read 8 pounds. Trained as a Bayesian, he computes the posterior mean and charges the lady for 9.99 pounds.

Let us briefly summarise the two alternative views of probability i.e. "personal" or "subjective" and "frequentist" or "objective", not with the objective of deciding which one is "better", but in order to underline the difference between them.

One difference is that the subjectivist is willing to apply probability ideas to a wider class of problems, as we have seen in example (6).

Another problem in the subjectivist view is that once you have established your "a priori" distribution there is a kind of automatism, i.e. Bayes' theorem, that tells you how you have to change it. There is no place for an unexpected hypothesis, Subject and object are bound together into a knot, which, once established, never can be untied.

For the frequentist, on the other hand, probability is linked to the object. As Dinges put it "... the probability is given by the experimental design, it is an objective value ... For an objectivist, the statement 'the probability of A is p', where A is a well-described event and p is a number, is either true or false... There is no place for the notion of hypothesis." (Dinges, 1978, p.116).

So both views, we think, can be characterised by the relation they assign to subject and object.

INDEPENDENCE AND EXCHANGEABILITY

Independence is given in case the possibility of observed results influencing subsequent evaluations is specifically excluded. This says, for the subjectivist, that any possibility of "learning through experience" is excluded if the original opinion is based on independence, because the latter, by definition, requires that the original opinion will not be modified on the basis of any osbervation of results. If, on the other hand, you consider independence as a hypothesis, there are statistical tests that can support or reject this hypothesis.

From the subjectivist point of view, however, this problem is the motive for considering the notion of exchangeability. This is a condition for the common distribution of a finite or infinite number of random variables, just as independence is. Mathematically these notions have the following form:

A family $\{X_i\}$ of random variables is called independent, if for every finite collection:

$$F_{X_1, \ldots, X_n} = F_{X_1} \cdot \ldots \cdot F_{X_n}$$

and is called exchangeable, if for every finite collection and every permutation σ:

$$F_{X_1, \ldots, X_n} = F_{X_{\sigma(1)}, \ldots, X_{\sigma(n)}}$$

where F denotes the distribution function.

If independent random variables are equally distributed, they are exchangeable. But there are interesting examples of exchangeable variables which are not independent.

For example, suppose that an urn contains m balls, r of which have a "1" written on them and m-r of which have a "0". If these balls are drawn from the urn, one at a time at random without replacement, and X_k denotes the digit on the kth ball drawn, then the sequence $X_1, \ldots, X_n$ is an exchangeable sequence; however, the random variables are clearly not independent. With this notion, one can easily solve the following problem:

(8) In an urn, there are 50 balls: 49 black and 1 white. Two persons alternately draw without replacement. The one who draws first the white ball has won. Does the first to begin or the second have the better chance to win?
The variables are exchagneable. Therefore the chances to win are equal.

We have thus seen that exchangeability, from a mathematical point of view, is an interesting notion. It is, for instance, possible to prove the strong law of large numbers for exchangeable sequences of random variables. Unfortunately, this is rarely done in text-books.

Let us briefly summarise some advantages of this notion:

a) As we have seen in example (8), it is a very natural notion (i.e. two events A,B are exchangeable if $P(A|B)=P(B|A)$), which has a wide range of application.

b) It can be shown that an infinite sequence of exchangeable variables is a probability mixture of independent, identical distributed variables. This is de Finetti's well-known representation theorem. Thus this notion has a technical meaning for an objectivist, too.

We have thus two intuitive and mathematical concepts: independence and exchangeability.

One may wonder why a subjectivist denies the notion of independence. First of all, independence, as we have mentioned, offers no possibility of

learning through experience. On the other hand, de Finetti, for instance, argues that independence is a very vague and unprecise notion.

We, however, consider both notions offer the possibility of building a model of a real phenomenon. In this respect, they are links between theory and application, and are faced with the problem of theoretical terms versus real phenomena.

THE BAYESIAN APPROACH - ONE AMONG MANY IN MATHEMATICS INSTRUCTION

It is a fact certainly undisputed that the demands addressed to stochastics, as they have been formulated by the subjectivist school, have initiated positive changes in stochastics instruction as well. As compared to a widespread conception of probability as a rigid, objectivist concept which is quasi attached to the object of cognition, a more comprehensive view of the complex relationship between the subject and the object of cognition was increasingly developed. Stochastic activities cannot be simply confined to the aspect of the objectively observable and measurable; the "subjective" side of arguing, estimating, and deciding must also be taken into account.

Such a conception of stochastics which, for the purposes of instruction, is expressed in a qualitatively new relationship between subject and object, between the learner and the subject matter to be learned, and which is usually not so explicitly thematised in mathematics instruction, has been formulated by Dinges in a general and fundamental way as a proposition: "I should like to propose to begin stochastics instruction always with decisions under uncertainty naively conceived by the learners. Subsequently, they are to formulate the procedure of decision. finally, there is a check how far mathematical methods can be developed to criticise the procedures of decision. Hence, the starting point will always be a situation in which the facts to be explained and the concepts to be introduced are concretely given, and a decision is required. After that, there is an exercise at more precise formulation, and finally mathematics with an apparatus of signs." (Dinges, 1980, our translation).

A broad scope associated with the probability concept, which intends to link the objective and the subjective aspects of stochastic learning, is promising both for instruction itself and for research into stochastics education; this tension between subjective and objective aspects, between the learner and the object of learning makes possible constructively interdisciplinary work of learning psychology and stochastics education.

If we turn to the various concrete approachces to stochastics for the purposes of mathematics education, we have to examine a double question: What can

be learned by means of stochastic concepts and methods, and what is learned, at the same time, about stochastics? We shall use this double question to analyse concretely the Bayesian approach to mathematics instruction.

Quite in keeping with Dinges' proposition, Riemer attempts, in his doctoral thesis (1981), to present an approach to stochastics which places the focus on the subjective concept of probability, and on Bayes' formula. The important fundamental idea is that certain (previous) assumptions, which are given subjectively and "a priori" by the students, be checked by means of an experiment in order to be finally revised. This basic idea of an "a priori" distribution and of modifying this distribution on the basis of experimental data is, for one thing, the prime concern of the Bayesian approach; in part, a promising educational aspect emerges here as well.

The following stages can be sketched for the stochastics course proposed; the introductory stage requires the students to make subjective estimates which are then revised by means of experiments and relative frequencies. The second stage then presents experiments which are to be repeated just once (resp. few times). In this connection, the Bayesian formula is developed via the inverted tree diagram. In a third stage, the Bayesian formula is applied several times in the sense of an iteration; computer simulation enables one to make several hundred experiments and evaluations by means of the Bayesian formula.

On the whole, it must be noted that this approach via the Bayesian formula does not just refer, as textbooks usually do, to the technical and calculus-related aspects of this formula, but attempts, by repeatedly applying this formula, to develop the Bayesian learning's basic concern of learning through experience for mathematics instruction.

These positive opportunities offered by an approach to stochastics via the Bayesian formula thus conceived is, however, subsequently relativated by the fact that an implicit claim to uniqueness is raised; this claim is justified, among other things, by the assumption that only the Bayesian approach enables one to establish a close link between probability and statistics.

In our opinion, this leads to drastic limitations of stochastics, and deprives the Bayesian approach of its advantages as compared to the rigid objectivist concept of probability. This limitation becomes acute mainly in the following two points.

- In a way similar to the fundamental problem discussed in part 2 ("in which way can the Bayesian approach cope with the unexpected hypothesis?") there is no enquiry, in this didactical connection either, after the limitations and conditions of the Bayesian approach. How can the

Bayesian cognitive situation itself be questioned?

- The one-sided attachment of probability to the subject of cognition entirely loses from view the reference to the object. The prerequisites and conditions of application and usefulness of the Bayesian formula can only be discussed if there is a representation of stochastic's object of application and its specific particularities. The specific particularities of stochastic object express themselves in different aspects; they can be physical, biological, psychological, and social characteristics, but also simple symmetrical conditions. These object-related characteristics are introduced in the shape of previous assumptions, but are also aspects which are explored by means of stochastics.

One of the most basic and abstract expressions of stochastics' object-relatedness is stochastic independence. For the Bayesian approach, the lack of self-reflection of the conditions of application shows, among other things, in the fact that the concept of independence plays an unimportant role. "Subjectivists... know independences only as something very much derived, as extremes in a symmetrically dependent model. This inability to develop productively the independence concept is, in our eyes, a grave defect of the subjective approach." (Dinges, April 1981).

We shall use a basic concept of stochastics, the inverted tree diagram, which is also use in the proposed course described, as an example to explain our proposition.

The inverted tree diagram is used to conclude inductively from an effect (or several) to possible causes. The possible causes are assigned an "a priori" distribution, and additional information is used to draw conclusions as to the effect, resp. the "a priori" distribution is modified according to the additional information. Riemer assesses the possible didactical use of the inverted tree diagram as follows: "The fact that students sometimes read off the Bayesian rule from the inverted probability tree even before the 'central question' is posed, shows the generative power this diagram has." (Riemer, 1981, p.24).

In its structure, the inverted probability tree can be compared to the so-called fault tree which plays a crucial role in many applications of stochastics. First of all with regard to the probability forecasts in the field of safety of nuclear power plants, the way this seemingly simple and elementary concept of the fault tree is used and interpreted raises the most fundamental problems. The following quotation shows in an exemplary fashion the actual difficulties which beset the concept of the fault tree and its use: for one thing, this

concerns the question whether really all possible causes have been included to find out the effect, and second, it concerns the question as to the reference to the object in the context of the independence concept. "This methodology faces a number of serious problems. The most fundamental is that unknown or unsuspected failure mechanisms cannot be included in the analysis... Sometimes, the probabilities of breaching each safety barrier are not necessarily independent since "common-mode" failures can increase the likelihood of failure of one barrier once another has been penetrated. The probability of such common mode failure is uncertain unless the physical mechanism coupling the supposedly independent barriers is understood" (Nuclear Power, Issues and Choices, 1977, p.227).

Both objective and subjective approaches to stochastics should avoid one-sidedness and beware of setting this absolute. Rather, there should be an effort, for stochastics instruction, to develop multiple approaches in a parallel way, fruitfully referring to one another; there can be no unique and solely valid conception of stochastics. "Stochastics in school must organically link notions taken from various philosophical traditions, in particular

- stochastics as the mathematics of mass phenomena
- stochastics as the logic of uncertainty
- statistics as the technology of transforming data into insights
- stochastics as a theory of decisions." (Dinges, 1981, p.51).

References

[1] Bayes, Th. (1958) An essay towards solving a Problem in the Doctrine of Chances, Biometrika, 45, 296-315.

[2] Berkson, J. (1977) My Encounter with neo-Bayesianism, Internat.Stat. Review, 45, 1-8.

[3] Dinges, H. (1978) Schwierigkeiten mit der Bayesschen Regel, Math. Phys.Sem., XXV, 113-156.

[4] Dinges, H. (1980) Thesenpapier, Manuscript, Frankfurt.

[5] Dinges, H. (1981) Bemerkungen zur Unabhängigkeit, Manuscript, Frankfurt.

[6] Dinges, H. Gedanken zum Aufsatz des Rev. Th. Bayes (1702-1761), unpublished Manuscript.

[7] Dorfler, W. and Fischer, R. (1981) Stochastik im Schulunterricht, Wien.

[8] de Finetti, B. (1972) Probability, Induction and Statistics, Wiley & Sons.

[9] de Finetti, B. (1974) Theory of Probability, Vols. 1 and 2, Wiley & Sons.

[10] de Finetti, B. (1979) Probability and Exchangeability from a Subjective Point of View, Internat.Stat.Review, 47, 129-135.

[11] Freudenthal, H. (1973) Mathematics as an educational task, Riedel, Dordrecht.

[12] Hacking, I. (1965) Logic of Statistical Inference, Cambridge.

[13] Hacking, I. (1975) The Emergence of Probability, Cambridge.

[14] Hamaker, H.C. (1977) Subjective Probabilities and Exchangeability from an Objective Point of View, Internat.Stat.Review, 45, 223-231.

[15] Heath, D. and Sudderth, W. (1976) De Finetti's Theorem on Exchangeable Variables, American Stat., Vol.30, No.4 188-189.

[16] Hewitt, E. and Savage, L.J. (1956) Symmetric measures on Cartesian Products, Trans.AMS, Vol.80, 470-501.

[17] Hodges, J.L. and Lehmann, E.L. (1970) Elements of Finite Probability. San Francisco.

[18] Humphreys, P.W. (1977) Randomness, Independence and Hypothesis. Synthese, 36, 415-426.

[19] Jordan, K. (1972) Chapters on the Classical Calculus of Probability. Budapest.

[20] Lindley, D.V. and Phillips, L.D. (1976) Inference for a Bernoulli Process (a Bayesian view), American Stat., Vol.30, No.3, 112-119.

[21] Nuclear Power, Issues and Choices. (1977) Ballinger, Cambridge, Mass.

[22] Riemer, W. (1981) Heuristische Begründungen und Modelle zu zentralen Begriffen und Sätzen der Stochastik, Dissertation, Wuppertal.

CHILDREN'S CHOICE BEHAVIOUR IN PROBABILISTIC SITUATIONS

Ruma Falk, The Hebrew University, Jerusalem

Most decisions in uncertain situations involve comparisons of the probabilities of success under the alternative choices, and selection of that alternative where the chances of success are higher. Hence the ability to discriminate between probabilities of varying discrepancies is crucial in such tasks. The development of that capacity was examined in two experiments.

In a previous study (Falk, Falk and Levin, 1980), 61 children, from 4 to 11 years old were repeatedly presented with two sets, each containing blue and yellow elements. Each time, one colour was pointed out as the payoff colour (POC). The child had to choose the set from which he or she would draw at random a POC element in order to be rewarded. The sets were of varying size with different proportions of the two colours. Three kinds of materials were used: Pairs of urns with blue and yellow beads, pairs of roulettes divided into blue and yellow sectors, and pairs of spinning tops, likewise divided into two colours. It was found that roughly around the age of six, children started to select the higher of the two probabilities systematically. The dominant error, that characterized the choices of the youngest children and decreased with age, was selecting the set with the greater absolute number of POC elements. No other systematic erroneous strategies were unequivocally discovered.

However, the conclusions of that study were derived mainly from analysis of group results. The pooling of responses of children in an age group might obscure individual choice strategies. Furthermore, some indications of incorrect choice strategies, different from "choose the set with more elements of the POC", could not be followed up because of shortcomings of the experimental design, so that a definite determination of such strategies was not possible. For example, several children seemed to select the urn where there were absolutely less elements of the non payoff colour (NPOC), but the design included too few trials where choice according to that rule differed from the correct choice. Moreover, there was no differentiation between the pattern of correct choices and the one dictated by the strategy; "Choose the set where the difference between number of payoff elements and number of non-payoff elements is greater". Thus, the latter principle could not be diagnosed.

THE EXPERIMENTS

The purpose of the present study is to identify the principles of choice in individual response patterns.

Design. The design was similar to that of Falk, Falk and Levin (1980). In each trial, two urns were presented, each containing blue and yellow beads in different numbers and in different proportions. One colour was designated as the payoff colour (POC). The children had to choose the urn from which they would try to blindly draw a bead of the required colour. Success in drawing a bead of the POC granted the child a prize. Hence, in order to be maximally rewarded, the subject's problem was to identify the set with the higher probability of the desired colour. The set of trials was designed to allow differentiation between relative and absolute quantitative strategies. By "relative strategy" I mean the correct choice of the set with a higher proportion of the POC elements. An "absolute strategy" could mean either selection of the set with more POC elements, or selection of the set with less NPOC elements, or possibly even selection according to a larger (or smaller) total number of elements. Suppose, for example, that the urn on the left contains one blue bead and three yellow ones, and the urn on the right contains three blue beads and four yellow ones, and yellow is the POC. The correct choice, by a relative rule, would be that of the left-hand urn, since $3/4 > 4/7$, however, the absolute strategy: "select the urn with a greater number of POC beads" will result in choice of the right-hand urn, since $4 > 3$.

The trials of our previous study had been classified into three levels of difficulty (Falk et al., 1980): "Easy" problems, in which one of the two proportions is larger than $\frac{1}{2}$ and the other smaller than $\frac{1}{2}$; "Medium" problems, involving comparisons of $\frac{1}{2}$ with a different proportion; and "Difficult" problems, in which both proportions are either larger or smaller than $\frac{1}{2}$. The unique status of the fraction $\frac{1}{2}$ is justified by the following: A child who is confronted with either an easy or a medium problem may compare absolute quantities of the two colours within each set and select that set in which the POC is a majority. That consideration would lead to the correct choice. Only with difficult problems is there no way to avoid comparing ratios between sets. Indeed, some of our subjects had verbalised such an absolute principle of choice: Sarah (9:9) (9 years and 9 months old) responded to an "easy" trial of the type (4 yellow, 6 blue) on the left versus (2 yellow, 1 blue) on the right, with blue = POC, by selecting the left set, and explained: "Here (on the left) there are more blues than yellows, whereas here (on the right) there are less blues than yellows". We found that level of difficulty did affect the rate of correct choices, the subset of "difficult" trials evoking the lowest rate.

Since less than half the trials in that study were "difficult", the inclusion of a majority of "easy" and "medium" trials might have partly accounted for the early age of emergence of the ability to discriminate between probabilities. In the present study, most of the trials are of the difficult category, i.e. presenting two probabilities on the same side of $\frac{1}{2}$.

In all trials, in both experiments, the urn containing more beads of the POC, also contains more beads of the NPOC (relative to the other urn). Note that is is impossible to construct a comparison where the correct choice, i.e., the urn with the higher proportion of POC beads, will include less POC beads than the other urn and more NPOC beads. It is possible, of course, for the correct choice to include more beads of the POC than the incorrect choice and less beads of the NPOC. However, in such a trial both strategies of choice "more of the POC" and "less of the NPOC" converge on the same urn, without any conflict. Therefore, in order to be able to diagnose the particular strategy by the pattern of choices, the urn with more beads of the POC (and more beads of the NPOC) was always different from that with less beads of the NPOC.

In addition to that, the design enables identification of consistent choice of the set where the (absolute) difference between the number of elements of the POC and the number of elements of the NPOC is greater.

Experiment 1 comprised 32 trials controlling for all the above mentioned potential strategies.

Consider, for example, the following trial: (3b; 9y) vs. (5b; 10y) with b = POC (i.e., the left urn contains 3 blue beads and 9 yellow ones, and the right one contains 5 blue beads and 10 yellow ones, and blue is designated as POC). Obviously, the right-hand urn is the correct choice since 5/15 > 3/12, however, the strategy of selecting the set with "less of the NPOC" will result in selection of the left urn, since 9 < 10.

Let us now consider (1b; 2y) vs. (6b; 9y) with y = POC. The correct choice is the left urn, since 2/3 > 9/15; however, the strategy of selecting the urn where the difference between the numbers of elements of the two colours in favour of the POC is greater, will lead to choosing the right-hand urn, since 9-6 > 2-1.

All 32 trials were of the "difficult" category.

Care was taken to present all possible cross-combinations of trials that would turn up to be correct or incorrect by employing all potential strategies. Table 1 presents the distribution of all the trials (along with those of Experiment 2) by the quantitative characteristics of the correct urn, relative to the other urn. The first line in each cell gives the number of trials of that category and the second line gives an example of a trial of that kind. The examples always give the numbers of the two colours written in the same order in both sets, with the number of POC elements written first (irrespective of whether it was blue or yellow). The arrow points at the size of the subset of the POC in the correct choice.

TABLE 1

Classification of trials by relative magnitude of numbers of beads of both colours (POC and NPOC) on each side, and the difference in favour of the POC on each side.

Experiment 1

Difference between number of beads of the POC and number of beads of the NPOC	Number of beads of both colours: Greater on correct side	Number of beads of both colours: Smaller on correct side	Total
Greater on correct side	4 (8 : 12) > (3 : 9) ↑	8 (4 : 1) > (5 : 3) ↑	12
Smaller on correct side	4 (5 : 10) > (1 : 3) ↑	8 (5 : 1) > (12 : 6) ↑	12
Total	8	16	24
Conservation trials	8 (1 : 3) vs. (3 : 9) ;	(5 : 5) vs. (2 : 2)	8
		Total	32

Experiment 2

Difference between number of beads of the POC and number of beads of the NPOC	Number of beads of both colours: Greater on correct side	Number of beads of both colours: Smaller on correct side	Total
Greater on correct side	5 (6 : 4) > (1 : 2) ↑	5 (2 : 1) > (4 : 6) ↑	10
Smaller on correct side	3 (3 : 6) > (1 : 3) ↑	3 (3 : 1) > (6 ; 3) ↑	6
Equal on both sides	4 (2 : 3) > (1 : 2) ↑	4 (2 : 1) > (3 : 2) ↑	8
Total	12	12	24

Note: Cells give the number of trials (upper line), and an example of a trial (below). The arrow indicates the number of POC beads in the correct choice. The second number is that of NPOC beads. The two numbers in the other parenthesis give, respectively, the same magnitudes in the incorrect choice.

The trial described in the upper right cell of Table 1, for example, presents a comparison between an urn with 4 POC beads and 1 NPOC bead, and an urn with 5 POC and 3 NPOC beads. The first urn (that in reality was assigned a right or left position as randomly determined) is the correct choice, since 4/5 > 5/8. Numbers of elements of both colours are smaller on the correct side (4 < 5 and 1 < 3), and the difference between number of elements of the POC and number of elements of the NPOC is greater on the correct choice (4-1 > 5-3). In contrast to that, in the trial described in the left cell of the second row of the table, that same difference is smaller on the correct choice, since 5-10 < 1-3.

Eight of the 32 trials were "conservation problems" presenting the same proportion of POC elements constructed of different absolute numbers of beads (an expanded fraction versus a reduced one). Two examples are given in Table 1. Obviously, there is no "correct choice" in conservation trials. These trials were introduced in order to test the prediction that a subject selecting the urn with "more of the POC" on non-conservation trials, will select the expanded of two equal fractions, and a child whose strategy is to select by "less of the NPOC" will choose the reduced fraction of the two.

Half of the non-conservation trials presented the correct choice on the subject's right side and half on the left. Likewise, the POC was evenly split between blue and yellow. All the features of the tasks were combined in a counter balanced and randomised design. The 32 trials of the experiment were divided into "Half A" and "Half B". The two halves were balanced with respect to all the variables of the design.

Experiment 2 consisted of 24 trials. The design was essentially the same as that of Experiment 1, precluding conservation items. See Table 1 for the distribution and for examples of trials.

A special feature of that experiment was that all trials came in pairs. The same two urns, with exactly the same two distributions of blue/yellow beads, were presented in the two trials of each pair (although their right-left location was randomly determined). Only the POC was blue in one trial and yellow in the other one (see Table 2 for examples of trials with their "mirror images"). Recall that the urn with the greater number of POC beads, always included also more of the NPOC beads. That way, the design enables to distinguish subjects who employ a relative strategy (the correct one) from those employing an absolute strategy. Consider, for example, the trial (6b; 4y) vs. (1b; 2y) with b = POC (see Table 1, first row of Experiment 2). One should select the urn (6b; 4y), since $6/10 > 1/3$. If the POC is changed to yellow, the other urn becomes the correct choice, since $2/3 > 4/10$. Such a switch of choices would signify the application of a relative strategy, relying on comparison of proportions. However, imagine a subject who systematically selects the urn with "more of the POC". He or she will select the (6b; 4y) urn in both trials of that pair, irrespective of the stated POC. Likewise, a child who selects the set with "less of the NPOC", will choose (1b; 2y) on both occasions. Such a pattern of results clearly violates the reversal of the direction of inequalities with change of sign (i.e., the rule stating that if $P_1 > P_2$, then $1-P_1 < 1-P_2$). A consistent pattern of correct choices will result in reversal of choices on all 12 pairs of trials, whereas a systematic reliance on an absolute strategy will yield 12 pairs of identical choices. Fluctuating strategies, or mere guesses, are expected to yield switches in about half the number of

pairs. Thus, the number of pairs (out of 12) in which a subject chooses the same urn, will be indicative of the tendency of that subject to depend on an absolute strategy, although the specific nature of that strategy can be determined only by examining the choices. The two trials of a pair did not necessarily appear successively; their position was randomly determined among the 24 trials.

Experiment 2 was planned for younger subjects, on the average, than those of Experiment 1. Therefore, the length (and duration) of the experiment was shortened by omitting the eight conservation trials. Another change relative to Experiment 1, was that four of the 24 trials presented paired comparisons of probabilities on opposite sides of the probability $\frac{1}{2}$ ("easy problems"). An example of a pair of such trials is given in the first row of Experiment 2 in Table 1. This was introduced in order to test the adherence of subjects to absolute strategies when carried to the extreme. The puzzling question was whether a child who consistently selects the urn with "more of the POC" (or "less of the NPOC") out of the two urns, on difficult items, will stick to that principle also when confronted with an easy trial. In the latter case one can resort to selection of the urn within which the POC is a majority, as in $(2 : 1) > (4 : 6)$.

Subjects. Thirty five children took part in the research. Twenty in Experiment 1 and 15 in Experiment 2. These were mostly children of highly educated parents (most of the children were from Jerusalem and only five were Americans from Oregon). Five children of each of the ages 6, 7, 8 and 9 were tested in Experiment 1. Experiment 2 included five children in each of the ages 5, 6 and 7. The children were invited to participate in a "lottery game" and came voluntarily (and gladly).

Procedure. Each child was tested individually. On both experiments "Half A" was run first and "Half B" second, for half the children, and vice versa for the other half.

A preliminary lottery was performed three times, tossing a die and using candies as prizes for otbaining an outcome that had been stated beforehand by the child. The purpose of this procedure was to acquaint the child with the idea that a win in a lottery is conditioned on the outcome of an uncertain event, and to prepare the child for possible disappointments.

Two experimenters were present. One gave instructions and presented the materials, and the other one recorded the responses.

One preliminary demonstration-trial with urns and beads preceded the experiment proper. Then came the two halves of each experiment with a short intermission for rest in between.

The standard procedure of each trial was as follows: The child was offered a selection of 3 to 4 prizes such as coloured candies, plastic toys or marbles. He or she selected the most attractive prize to try to win. The prize, together with a bead of the POC of that trial, were placed in front of the child in a small plate. The pair of urns were now presented and placed on both sides of the prize in equal distances. The beads of the two colours were put into the urns, according to the required composition, while the child was watching, then the experimenter explained: "You will soon be given a chance to draw one bead, with your eyes closed, from one of these urns. If you draw a blue (yellow) bead (indicating the bead next to the prize), you will win the prize. Which of the urns would you like to try to draw from?". The child then pointed at the selected urn, and closed his/her eyes. The selected urn was shaken by the experimenter and presented to the child's groping hand. At the end of each trial the child either won or did not win the prize, according to the rules of the game. No verbal reinforcements were given for correct choices.

Verbal responses were not required of the child throughout the experiment. However, each spontaneous comment or explanation of the child was recorded. A short post-experimental interview was conducted with each subject. A couple of trials that had appeared in the experiment were presented again and the child was asked to explain his/her choice.

The experiment lasted about an hour (somewhat less for Experiment 2).

RESULTS AND CONCLUSIONS

All the responses (to non-conservation trials) were scored by either + or -, depending on whether the higher or the lower of the two probabilities had been selected, respectively.

The overall rate of correct responses, for the different ages, is presented in Table 2. The results describe both experiments, and are pooled across subjects and across all categories of trials.

Clearly, there is improvement in performance with age. The rates of correct responses are lower than in our previous study for all ages (see Falk et al., 1980, Table III, page 192). This was expected, considering that a much higher proportion of trials in the present study are "difficult", and that all trials present more beads of the two colours in the same urn. The conjunction of these two factors probably caused the performance in this research to be lower than in our previous study, even for the subset of trials with smaller number of POC elements on the correct side (Falk et al., 1980, Table IV), and for the subset of "difficult" trials (Table V). However, the overall rate

of correct responses presents group results, eventually including individual subjects with contrasting response tendencies.

TABLE 2

Analysis of Results of Experiments 1 and 2 by Age of Subjects

Age Group	Number of subjects	Rate (a) of correct responses to all trials	Proportion (b) of subjects deemed to:- master the task	(1) select by "more of the POC"	(2) select by "less of the POC"	(3) select by "greater difference in favour of POC"
5 - 6	5	0.56	-	-	0.40	-
6 - 7	10	0.57	0.10	0.40	0.10	-
7 - 8	10	0.73	0.50	0.10	-	0.20
8 - 9	5	0.78	0.60	-	0.20	0.20
9 - 10	5	0.84	0.80	-	-	-
Total	35	0.68	0.37	0.14	0.11	0.09

(a) Rates were computed out of 120 responses when 5 subjects were involved, and 240 responses in the case of 10 subjects.
(b) Proportions were calculated out of the number of subjects in the age group.

Each individual subject's pattern of choices was analysed in order to determine whether that subject had mastered the task of selecting the higher of two probabilities, and in order to diagnose possible strategies of choice.

The criterion for classifying a subject as employing the "correct principle", was that he/she achieved at least 18 correct choices out of 24 (non-conservation) comparisons. This means that no more than six errors were tolerated. Such a high level of performance would have been obtained by chance (were the child merely guessing) with probability 0.011. The same (arbitrary) level of proximity was also applied for determining that a response pattern agrees with the pattern predicted by any of the strategies studied. (There was only one exception: As can be seen in Table 1, there are only 16 definite predictions of responses resulting from the strategy: "select the urn with a greater difference between number of POC beads and number of NPOC beads", in Experiment 2. In that experiment the difference between the numbers of beads of the two colours is equal in both urns, in 8 trials. Therefore, the criterion for determining application of that strategy, was at least 13 responses, out of 16, matching the strategy. The probability of satisfying the criterion, by chance, is 0.011).

Table 2 presents the proportion of subjects, in each age, that either performed correctly (close enough to perfection), or applied one of the three strategies consistently. The three strategies that were "predicted" by the design, and did appear in the experiments, were: select by (1) "more of the POC", (2) "less of the NPOC", and (3) "a greater difference in favour of the POC". Note, that when the unit of analysis changes from trial to subject, numbers become small. The proportions in Table 2 were based on 10 or 5 subjects. Thirteen of the 35 subjects applied the correct principle, whereas 12 were sticking to different incorrect strategies: 5 subjects to strategy (1), 4 to (2), and 3 to (3). The mean number of errors for the 13 subjects deemed to master the task was 2.8. The mean number of deviations from a perfect "strategy-pattern", for subjects classified as applying specific erroneous strategies was 2.0.

A count of the number of identical choices in "symmetric" pairs of trials (i.e. pairs of trials presenting the same two urns with blue as the POC in one and yellow in the other) in Experiment 2, turned out as follows: There were 5 subjects who applied an absolute strategy (either "more of the POC" or "less of the NPOC"). These subjects made, on the average, 10.2 identical choices out of 12 pairs[(1)]. All other subjects of that experiment, who did not apply an absolute strategy, scored a mean of 5.6 identical choices. Thus, the number of choices of the same urn, in spite of change of the POC, supports the impression that some of the subjects were committed to their erroneous absolute strategies. Furthermore, these 5 subjects stuck to their erroneous absolute strategies even on "easy" trials, that served as test-trials in Experiment 2.

Some individual results. A few of the children who matched the pattern of responses predicted by an absolute strategy, seemed very dedicated to their principle. Thus, Tali (6:0), in Experiment 1, who applied the strategy "more of the POC" in 23 out of the 24 non-conservation trials, verbalised her rule of choice explicitly in the interview: "Here (on the left) there are more blues (the POC of that trial), so I thought it would be easier for me to find a blue one". When we remarked that on the left there are also more yellow beads, and asked whether she had considered that, she replied: "That does not matter at all!". On one occasion, when we started filling the urns for a certain trial, and had already put the required numbers of POC beads in both urns, she immediately selected the urn where we put more beads. When told: "But we did not put the beads of the other colour, yet" she said: "This does not make any difference!". Not surprisingly, on all 8 conservation trials, Tali selected the urn with "more of the POC", i.e. the expanded fraction.

(1) The probability of making by chance at least 10 identical choices, out of 12, is 0.019.

Beni (6:0) also selected by "more of the POC" and consistently selected the "expanded urn" on all conservation trials. Yotam (8:1) applied the strategy of selecting "less of the NPOC". He too carried that pattern of choices over to conservation trials, where he chose the urn with less beads of both colours on 7 trials out of the 8. Ron's (8:6) principle was to select the urn with "greater difference in favour of the POC", and he verbalised that principle clearly. In 4 conservation trials, where the probability was smaller than $\frac{1}{2}$, Ron stuck to his principle. Thus, in (2b; 6y) vs. (1b; 3y) with b = POC, for example, he selected the right-hand urn, since 1-3 > 2-6. (In the other 4 conservation trials the probability was $\frac{1}{2}$ and choice by difference was impossible).

Among other children, who were not committed to any absolute strategy, all the range of the distribution of responses to conservation trials could be found. Curiously enough, a few of the children who responded correctly (beyond criterion), chose the reduced urn consistently on all 8 conservation trials. A refreshing reaction was that of Oren (9:7): When he was confronted with conservation problems, after coping with paired comparisons of different probabilities, he burst out in laughter, and contemptuously selected via chanting "En Den Dino" (a well known Israeli nursery rhyme for chance selection of one from a group). The procedure proved efficient... his responses were equally divided between the reduced and the expanded fractions.

Another absolute strategy, that appeared in Experiment 2, deserves mentioning: Tobi (5:3), who was classified (in Table 2) as choosing by "less of the NPOC", was actually selecting the urn with a smaller total number of elements. Since it was the same urn, on all trials, that contained less of the POC and less of the NPOC, that urn also had a smaller total; consequently there was no differentiation between the patterns of choices resulting from the two strategies. However, Tobi counted the beads aloud, in the course of the experiment, and he was obviously comparing totals, and he said so clearly in the post-experimental interview. His adherence to the strategy was perfectly consistent on all 24 trials, even in the "easy" trials of that experiment. Tobi's number of identical choices in the symmetric pairs of Experiment 2 was 12, i.e., in all pairs of trials he chose the same urn, in spite of reversal of POC.

A "dramatic" insight, in the course of the Experiment, was displayed by Shahar (6:3). He started by selecting "more of the POC", and did so consistently on all the trials of "Half A" and on the first 4 trials of "Half B" (he was therefore classified among the subjects applying that strategy in Table 2). At that point he announced spontaneously: "Now I shall choose where there are fewer" (probably meaning: "fewer of the NPOC"). He did switch to that strategy for a few trials, but later abandoned it and kept fluctuating. The

point of interest for us was watching him suddenly discover "the other colour", and later on exhibiting some conflict and lack of knowledge of how to utilize the information on both colours simultaneously.

Naomi (7:4) spread her 7 errors over 7 different symmetric pairs of trials. She was focusing on the same urn in these seven pairs (and consequently made identical choices on 7 of the 12 pairs). She expressed her special interest in the urn when at one point of the experiment, after displaying some conflict, she said: "I am much attracted to this urn".

The issue of stages. Not all the subjects could be classified as selecting according to a given strategy, correct or incorrect. There were 10 children (of 35) whose "stage" could not be determined even with a less than perfect criterion of agreement. Some children seem not to stick to a single strategy, They either have some rule in mind and make errors in using it, or they fluctuate between strategies, or even make chance decisions (at least from the point of view of our analysis).

The conclusions that can be drawn from the results, concerning the developmental course of the ability to discriminate between probabilities, are limited. I shall try to delineate a tentative developmental course based on logical analysis, and partly on suggestions derived from the results.

It seems reasonable that absolute quantities would initially be dominant in that kind of judgement. Especially, the absolute size of the winning category (POC) would be salient in capturing the judge's attention. Hence, "select by more of the POC" would appear to be a candidate for the first quantitative strategy. Later, would come the understanding that the size of the set of elements of the other colour is also related to the probability of success. The awareness of the relevance of the number of NPOC elements may be dominant for a while and the child may focus on that number exclusively. That way, "less of the NPOC" would be the "second stage". The reliance on number of elements of only one colour may fluctuate between the two colours for a while, until the conflict is resolved by a realisation of the need to consider both quantities simultaneously. The integration of these two numbers via computation of proportion (or ratio) is not expected to occur instantaneously. Different attempts to integrate these two numbers may appear, combined with temporary "regressions" into judgement by only one of the quantities. One of the easiest and tempting ways to try to combine the two numbers, is via computing differences. Additive operations come first in education and hence are easily available to the child. This suggests that "select the urn with a greater difference in favour of the POC" could be the third stage, one before finally combining the numbers via proportion and mastering the conservation of that concept.

However, there is not enough evidence in the results to establish that sequence of strategies as an empirical finding. More data should be collected in order to learn about the prevalence of the various strategies in different age groups. Also a longitudinal study, following the same children for a number of years, may determine whether those stages appear sequentially, and whether they come necessarily in the same order.

Concluding remarks. One of the educational implications of this study is that the teaching of probability to children should be carefully planned. The lesson to be learned from the results of the present research, in conjunction with earlier studies, is that the concept of proportion (and probability) is very elusive. Children can easily avoid directly coping with proportional computation. They may find alternative strategies and appear as if they have mastered the problem. Some of the incorrect choice strategies that were diagnosed in the elaborated design of the two experiments, were not and could not have been discovered in some of the previous designs. For example, Yost, Siegel and Andrews (1962) presented paired comparisons only between complementary probabilities. Two unequal complementary probabilities are always on different sides of $\frac{1}{2}$. As we saw above, such "easy" problems can be correctly solved by absolute considerations within an urn, whereas the experimenter, and teacher, may entertain the illusion that the children have mastered the task. Yost et al. (1962), Davies (1965), and Goldberg (1966) concluded that preschool children (even 4 year olds) were capable of systematic probability judgements. However, their comparisons always included an equal number of elements of one of the colours, the colour preferred by the child, in both sets. Hence, when the NPOC appeared in equal numbers the child could choose by "more of the POC", and when the POC came in equal numbers the child could choose by "less of the NPOC". That child would be counted as "capable of systematic probability judgement". In a similar vein, the few children who were judged in the present study as choosing by "a greater difference in favour of the POC" would have been counted as selecting correctly in Falk et al. (1980), because of the limitations of their design. We should take care that this would not happen in the teaching of probability. The teacher's responsibility should be to present the concept in its most general form and to make sure that all the components of the concept (namely, the numbers of elements of both categories, the difference between them, and the total) will appear in all their possible mutual relations and will vary in a free interplay.

Comparison of probabilities is but one of the judgements involved in choices in probabilistic situations. Decision making under uncertainty requires also assessing the costs and payoffs associated with the alternative choices.

Next comes the need to integrate the probability and the payoff (or utility) of each alternative into one magnitude characterising its expected value. A natural extension of the present research would confront the child, on each trial, with a task of choosing between two gambles. Each gamble would consist of its own probability of success and the payoff (gain) for winning. A conflict could be created between the result of comparing the two probabilities to that of comparing the payoffs, so as to diagnose possible choice strategies of children. Such a research is now under way.

Acknowledgements

This study was supported by the Human Development Center, The Hebrew University of Jerusalem.

Thanks are due to my husband Raphael Falk for his helpful comments and to my son Oren for assistance in conducting the experiments.

References

[1] Davies, C.M. (1965) Development of the probability concept in children. Child Development, 36, 779-788.

[2] Falk, R., Falk, R. and Levin, I. (1980) A potential for learning probability in young children. Educational Studies in Mathematics, 11, 181-204.

[3] Goldberg, S. (1966) Probability judgements by preschool children: Task conditions and performance. Child Development, 37, 157-167.

[4] Yost, P.A., Siegel, A.E. and Andrews, J.M. (1962) Nonverbal probability judgements by young children. Child Development, 33, 769-780.

AN IDEA OF STRUCTURING PROBABILITY TEACHING BASED ON DIENES'S SIX STAGES

Anna Olecka, Guardamangia, Malta

INTRODUCTION

Zoltan P. Dienes has proposed a way of developing mathematical concepts going step by step from concrete to formal operations. He has subdivided this process into six consecutive stages. These [1] are briefly:

(1) Initial interaction.
(2) Discovery of regularities in situations.
(3) The comparison of several games possessing the same structure; this is the stage of searching for isomorphisms.
(4) The representation of isomorphic situations in one all-embracing form.
(5) The study of the representation by the description of its properties. In this way all the isomorphic situations are being studied. Usually symbols will be needed at this stage.
(6) The formalisation of the system.

The purpose of this paper is to suggest how Dienes' idea can be applied as a framework for organising probability teaching. If we try to organise probability teaching globally in this manner, it turns out too complicated and inadequate; too many concepts are involved at the same time. So I believe that for the sake of didactical analysis it may be useful to separate particular concepts and to model the teaching according to Dienes' stages - each concept separately - and finally combine these models.

This approach may be a starting point for designing probability curricula for beginners: children, adolescents and even college students and teachers (if they are learning probability for the first time). Of course, young children will need more time for experimentation. Also the time allotted for the choice of apparatus, use of graphical schemes etc. would depend on the kind of student for whom it is intended. But the general pattern will follow the same line. And it is just this general pattern which is presented here. Neither methods nor teaching aids presented in this paper are new. They have been widely discussed in the literature for the last few years. I have taken advantage of them to fill up the idea of Dienes' stages with some practical contents. The literature on which this paper is based is listed at the end.

The concepts considered in this paper are: probability multistage experiments, expectation and conditional probability. However, only stages I - V will be discussed here. Stage VI could be common to all of them as a formal presentation of an elementary probability course in a systematic way, after

completing stage V for the concepts involved.

THE CONCEPT OF PROBABILITY

Stage I. Initial interaction. Random experiments with various devices (coins, spinners etc.). Experiments may be connected with games, but the games should have very simple rules.

Stage II. Discovering regularities. The aim of this stage is to help the children to observe the stabilisation of the relative frequencies when the number of trials is large enough and to compare the frequencies with the chances assigned "a priori", intuitively or with some reasoning.

Two kinds of students' activities are expected at this stage:
* Experimentation, recording and organising data, computing the relative frequency, presenting through graphs both the data and the relative frequency.
* Some simple reasoning involving symmetry or proportions, leading to the theoretical values of probabilities in very simple cases (one- or two-stage experiments). Some graphical aids (tables, tree, diagrams, etc.) are useful here. In order to present the random experiment from the theoretical point of view, theoretical graphs (the different outcomes on one of the axes, the probabilities on the other) can be made.

Comparing theoretical chances with those otbained from estimation of the relative frequences is a significant, indispensable step towards understanding probability. It can be done by comparison between theoretical and statistical graphs.
For some experiments it is impossible or at least very hard to assign theoretical probabilities of outcomes. Yet it is important to show some situations of this kind and to try to estimate probabilities by relative frequencies. A cumulative relative frequency graph will be a very useful tool at this point.

Stage III Searching for isomorphisms. During stage II, students carry out research on each experiment separately. In the next stage, it is time to make comparisons between the different experiments.

The aim of stage III is to show that some random experiments can be replaced by others.

At the beginning we shall consider cases where the random devices are isomorphic in some obvious sense. For example it is significant that students can grasp the idea that a spinner (Fig.1) can be used instead of a coin. This will not be so obvious if we divide the spinners into more than two parts (see Fig.2).

FIG.1

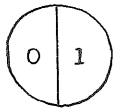


FIG.2

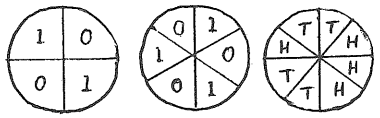


FIG.3



If we divide the spinner into unequal parts (as in Fig.3) it would be even more difficult. A comparison between theoretical graphs of the experiments yields a good criterion for isomorphism of the random experiments involved. When it is intuitively understood by students we can go on to the difficult area of comparing experiments with quite different structures. Example:

A. Spinning the spinner as in Fig.4 twice and recording the sum of the two numbers obtained.

B. Spinning the spinner as in Fig.5 once and recording the number obtained.

FIG.4

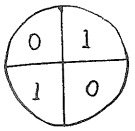


FIG.5

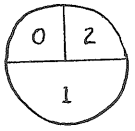
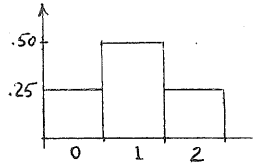


FIG.6



Although the experiments A and B are different their sets of outcomes and the corresponding probabilities are the same. Students can discover this while constructing theoretical diagrams for both of the experiments (Fig.6). Also circle diagrams for both of the experiments are the same: this is a spinner as in Fig.5.

This stage seems to be the crucial point in the learning process. So we have to choose adequate activities which would enable the students to simulate some probabilistic situations with others.

Stage IV. Representation. In order to describe a random experiment one has to describe the set of all possible outcomes and then to assign the probabilities of the outcomes. Thus the following table is a very useful representation:

outcome	w_1	w_2	w_3	...	w_n
probability	p_1	p_2	p_3	...	p_n

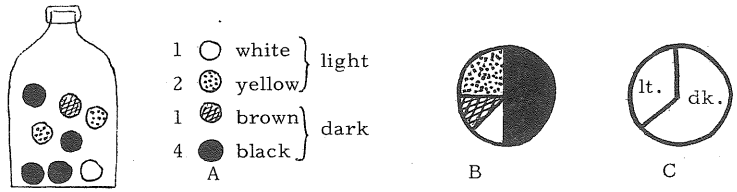
$p_1 \leq p_2 \leq \ldots \leq p_n$

A circle diagram (Fig.7) or a one-step tree diagram can be another suitable representation. The circle diagram helps to see that $p_1+p_2+\ldots+p_n = 1$, as we just divide the whole circle into parts. The tree diagram will become very useful later on, when multistage experiments are considered.

Stage V. The properties of the representation. The main property to be discovered now is a rule for computing the probabilities of events. Sometimes one can do it just by changing the way the set of outcomes is described. For example, both of the circle-diagrams (as in Fig.7) represent an experiment of drawing a marble. The set of given marbles is the same in both cases

(Fig.7A). But in the first case (Fig.7B) the colour is recorded whereas in the other case (Fig.7C) only two possibilities are considered: "dark" or "light".

FIG.7



One can notice that we can get the same result by simply adding the probabilities of corresponding outcomes in the first case. So we get a rule for assigning probability to an event E:

* to describe (in any convenient way) all the possible outcomes and to assign them probabilities;

* to add the probabilities of all the outcomes favourable to E.

The classical scheme (when all the outcomes are equally likely) will appear at this stage as a particular, but significant, case.

Now some properties of probability may be studied (such as probability of the complementary event etc.).

MULTISTAGE EXPERIMENTS

Stage I. Initial interaction. At this stage the students will be doing and watching various multistage experiments (perhaps connected with simple games).

Stage II. Discovering regularities. Each experiment will now be separately investigated. The basic skill developed here is the ability to distinguish successive stages of a multistage experiment. In order to describe such an experiment the student will need some graphical representation. Tree diagrams are the most convenient, but for the two-stage experiments square diagrams or just simple tables are also useful.

Examples:

A. Drawing balls from the box |O ⊠ ⊠ ⊠| without replacement until a white ball is drawn. A tree digram describing the experiment is shown in Fig.8.

B. Drawing two balls from the box |O ⊠ ⊠ ⊠| without replacement. Constructing a square diagram describing the experiment is shown in Fig.9.

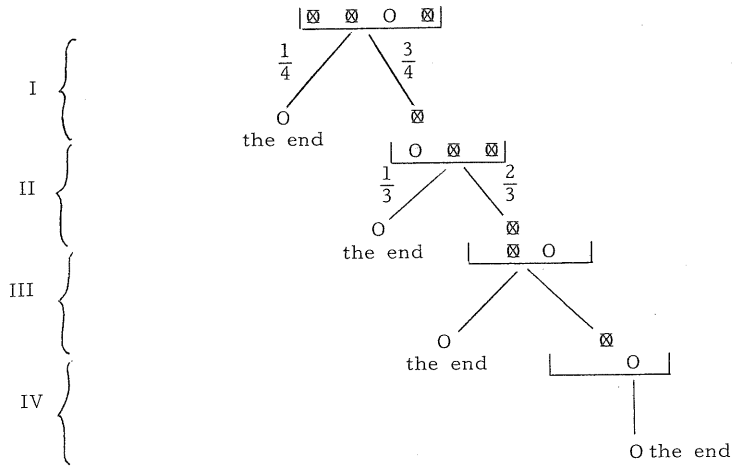


FIG.8

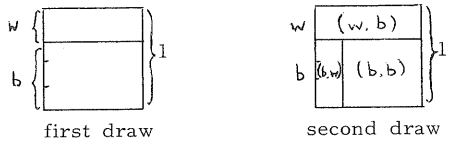


FIG.9

We assume that students are already able to assign correct probabilities at each stage separately, but computing probabilities for the entire experiment would be too difficult at the moment. So they would only be able to estimate them by repeating the experiment many times. A tree diagram can be helpful: we put some chips at the top of the tree and then each chip will follow a pattern down the branches according to the results of the experimentation. Finally we just count how many chips get to the end of each branch, and calculate the relative frequencies (number of chips at the end of a branch divided by the total number of chips used).

Stage III. Searching for isomorphisms. Here we shall compare various experiments. We can do it either by comparing all the possibilities they yield or by comparing their tree diagrams.

Stage IV. Representation. Two experiments will be considered isomorphic if they can be described by the same tree diagram (up to permutation of branches). Therefore a tree diagram could be a good representation of a multistage experiment.

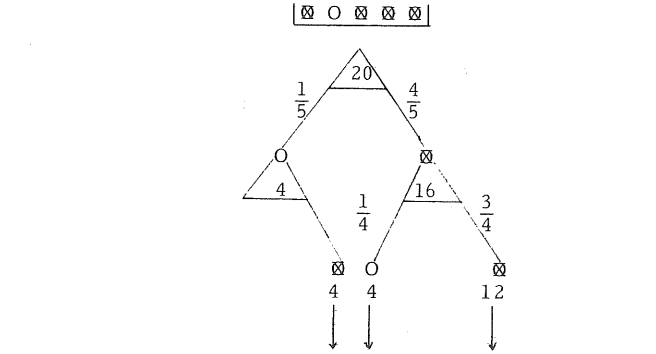
Now the question arises how to compute the probabilities of outcomes. One can use the special case of the probabilistic abacus of Arthur Engel [4], namely with the tree diagram. The probabilistic abacus is a kind of idealised simulation. The chips are moved along the branches and they are divided into groups, at each stage according to the probabilities of branches. Eventually each of them reaches the end of the branches (we supply more chips at the start if necessary). We can now define the probability of the whole branch as the fraction: number of chips at the end of the branch divided by the total number of chips used.

Using the probabilistic abacus on the tree diagrams may seem trivial, but it can be a significant step between estimating probabilities and computing them.

Finally, a natural representation of the multistage experiment is the tree diagram supplied with the probabilities of the outcomes. Fig.10 shows the representation of the experiment of drawing two balls from the box |⊠ O ⊠ ⊠ ⊠| without replacement while two possibilities are considered: "the two balls drawn have the same colour", "each of them has a different colour".

Stage V. The properties of the representation. The above way of computing probabilities will lead to discovering the two fundamental rules: the rule of multilying probabilities along a branch and the rule of adding the probabilities of the favourable branches (see [2]). Later on we can examine carefully the tree diagrams in order to find properties of probabilities in multistage experiments.

FIG.10



outcome	different colours	the same colour
probability	$\frac{4+4}{20} = 0.4$	$\frac{12}{20} = 0.6$

CONDITIONAL PROBABILITY

Stage I and stage II to some extent can be done together with multistage experiments. Indeed in describing the experiments with tree diagrams one has to assign conditional probabilities to the branches. Therefore we may regard these activities as the intuitive base for conditional probability.

Stage II. Discovering regularities (continuation). Our starting point was that the concept of probability should develop as a synthesis of both of the approaches: "a priori" and"a posteriori". The same idea governs our approach to conditional probability. Example:

Take the spinner shown in Fig.11 and record in the table below the numbers obtained, taking into account at the same time the colours of the fields obtained:

FIG.11



white	2 2 1 1 2 1 1 2 2
not white	1 1 1 1

After having recorded the data for, say, 100 trials, we now disregard the lower part of the table and we calculate the relative frequencies of numbers 1 and 2 for those times when the spinner landed on the white field.

Another example:

Toss a coin 3 times and consider the event A: "we get an odd number of tails". Let us record the results of experimentation in the following manner:

A	(T,H,H), (H,H,T), (T,H,H), (H,T,H), (T,T,T), (H,H,T), ...
A'	(T,T,H), (T,H,T), (H,H,H), (T,T,H), ...

Now we consider another event B: "we get one tail only". We shall find the relative frequencies of B (i) from the combined results from both parts of the table (ii) from the upper and lower parts separately (the last one equals zero, of course).

Stage III. Searching for isomorphisms. Examining the activities of the type discussed above we next move on from relative frequencies to probabilities. Therefore we consider all the possible outcomes split into two parts, A and A':

A	(THH)	(HTH)	(HHT)	(TTT)
A'	(HTT)	(THT)	(TTH)	(HHH)

Since all the outcomes have the same probability we infer that $p(B)=\frac{3}{8}$, $p_A(B)=\frac{3}{4}$.

Now we can go back to the previous activities and use the same method to find conditional probabilities.

Stage IV. Representation. In stage III students calculate conditional probabilities simply by following the method of the first example (with three coins). Although the way of calculating conditional probabilities was always the same there was no general rule formulated. Now is the time to formulate the rule.

Let us assume that a set of all possible outcomes consists of N equiprobable outcomes. Let us consider now two events A and B and let us decompose the N outcomes according to the criterion of favourability to each of the events A and B. So we get our set of N outcomes sub-divided into four subsets. The cardinalities of the subsets are shown in the following table:

(*)

	B	B'
A	k	t
A'	m	n

k+t+m+n=N

The conditional probability is then equal to

(**) $$p_A(B) = \frac{k}{k+t}$$

We may then ask the question - What is an adequate graphical representation of the conditional probability?

We may find it didactically attractive to replace the table (*) by a square diagram (Fig.12). It is a more general graphical representation because the outcomes need not be equiprobable.

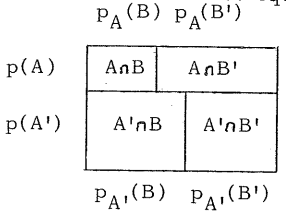


FIG.12

If the events A and B correspond to two successive stages of an experiment, a tree diagram is also a useful representation (Fig.13)

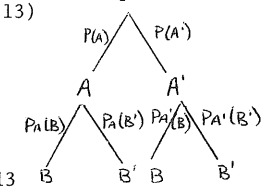


FIG.13

Stage V. The properties of the representation. In this stage we shall examine the representation in order to find properties of conditional probabilities ($P(A \cap B) = p(A) \cdot p_A(B)$ etc.). Studying the square diagrams one notices that in the case when $p_A(B)$ is independent of A, the diagram turns out as in Fig.14. Therefore square diagrams may be a good representation also of independent events. Didactically, there are two choices: either to teach conditional probability and independence together from the beginning (stages II-V) or to treat them separately in stages I-IV and to link them together in stage V.

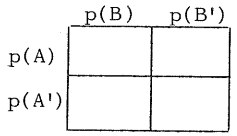


FIG.14

EXPECTATION

Stage I. Initial interaction. Games and experiments considered now will always have the outcomes being rational numbers. Students will have an opportunity to gain some experience observing the duration of the game, the value of the win etc.

Stage II. Discovering regularities. We analyse each game separately and calculate experimental mean values in various situations. This can be motivated by asking questions about the fairness of the game etc.

Stage III. Searching for isomorphisms. If we compare the procedures used to calculate the mean values for the above experiments we will notice that the procedure follows the same pattern:

(*) $f_1w_1 + f_2w_2 + \ldots + f_nw_n$ where f_i is the relative frequency of the outcome w_i.

Stage IV. Representation. Take the following problem: We want to play a game a large number of times. Can we predict or estimate the mean value before playing?

If the number of trials increases, the relative frequencies will come closer and closer to the probabilities. Thus we may think of the "ideal mean" which will replace the formula (*):

(**) $w_1p_1 + w_2p_2 + \ldots + w_np_n$ where p_i is the probability of the outcome w_i.

Now we are ready to compute expectations in the above games and experiments and compare them with our experimental results.

Stage V. Properties of the representation. We will now examine the formula (**) and look for properties of the expectation. In other words we will now investigate all the above situations at the same time by simply examining their common representation.

The approach to the expectation proposed above, based on abstracting from concrete situations, has the advantage of avoiding the concept of a random variable which is difficult for beginners. Consequently, it makes it possible to talk about expectation much earlier, presumably as soon as students are able to assign the probabilities.

Later on, in more advanced courses, the above approach will be just a particular

case where the outcomes are numbers and the random variable is an identity.

The special case considered here includes most of the typical examples of expectations in an elementary probability course. If we need the expectation of two random variables on the same space we simply think of two different experiments with different numbers as outcomes. This is sufficient at such an elementary level.

* * *

In a similar way, developing other important concepts (Bernoulli scheme, geometrical probabilities etc.) can be also put in the same framework of stages I-V.

As has been said in the introduction, a didactical consideration of isolated concepts is not enough. The problem now is to link them. This can, of course, be done in various ways. The optimum way seems to depend on the reasonable compromise between the following two postulates:

1. To start with the important concepts as soon as possible in order to provide sufficient time for deeper understanding.
2. For each concept stages I-IV should be in close succession to each other so that the student will not have forgotten the stage immediately preceding.

So, it seems to me that the compromise would be to have stage II immediately and stages III-V staggered, provided in the meantime the students can apply the knowledge they gain. The following table represents a possible order of concepts and stages. However, this has not yet been a subject of global experimental verification.

Concepts	Order of stages																				
Probability	I	II	III			IV		V													
Expectation				I	II											III	IV				V
Multistage exp.						I		II		III		IV	V								
Condit. prob.						I		II			III			IV		V					
Independency								I			II			III	IV	V					
Bernouli's sch.																		I-III	IV-V		
Random walks																		I-III		IV-V	

References

[1] Dienes, Z.P. (1971) An Example of the Passage from Concrete to the Manipulation of Formal Systems. Educ.Stud.in Math., 3.

[2] Engel, A. (1973) An Introduction to Probability. St.Louis, CERMEL.

[3] Engel, A. and Rade, L. (1973) Topics in Prob. & Stat., St.Louis, CERMEL.

[4] Engel, A. (1975) The Probability Abacus. Educ.Stud. in Math., 6.

[5] Engel, A., Varga, T. and Walser, W. (1973) Zufall oder Strategie.

[6] Fischbein, E. (1975) The Intuitive Sources of Probabilistic Thinking. Dordrecht, Holland, Reidel Publ. Company.

[7] Freudenthal, H. (1974) The Crux of Course Design in Probability. Educ.Stud. in Math., 5.

[8] Look on Luck. (1974) IOWO, Utrecht.

[9] Symposium on Combinatorics and Probability in Primary Schools (1975) (proceedings), Warsaw.

[10] Teaching of Probability and Statistics. (1970) L. Rade (Editor), Stockholm.

[11] Teaching Statistics and Probability. (1981) A.P. Shulte and J.R. Smart (Editors), National Council of Teachers of Mathematics, Yearbook.

DOES THE TEACHING OF PROBABILITY IMPROVE PROBABILISTIC INTUITIONS?

E Fischbein and A Gazit, Tel Aviv University

In the introduction to his well known treatise on probability, W. Feller has emphasized the necessity to develop adequate probabilistic intuitions as a basic component of the teaching of probabilities (Feller, 1968). But so far, little is known about probabilistic intuitions and their development under the influence of systematic instruction. In their main work on the chance concept in children, Piaget and Inhelder have concluded that, during the intuitive period (before the age of 6-7), the child is not able to distinguish clearly between chance and necessary phenomena. That distinction appears during the concrete-operational period together with elementary forms of probabilistic estimations. The concept of probability, as a formal construct, develops only during the formal operational stage and it represents a synthesis between the necessary and the possible (Piaget and Inhelder, 1951). This interpretation has been contested by various authors. Some of them consider that elementary probabilistic estimations may be found even in preschool children (Yost et al, 1962; Goldberg, 1966; etc.). Fischbein has also claimed that correct probabilistic intuitions may be identified in preschool children, but, because of social influences and present school curricula, only the deterministic aspect is actually developed (Fischbein, 1975). This situation may explain, partially, the fact that, frequently, the spontaneous probabilistic judgements of people are biased. Such incorrect evaluations of uncertain situations are often expressed as preconceived opinions and superstition. A good account of various classes of probabilistic misconceptions may be found in the work of Twersky and Kahneman (Kahneman and Twersky, 1972; Twersky and Kahneman, 1973). A more recent work dealing with children's probabilistic subjective evaluations is that of Falk et al. (1980).

The present investigation is concerned especially with the influence that the teaching of probabilities may have indirectly on intuitive probabilistic judgements. As mentioned above, there is very little information available about this problem. In an earlier work, Ojemann et al. have reported a positive, indirect effect of probability lessons on predictions made by their subjects (8-10 years old) in probability learning tasks (Ojemann et al., 1965).

The term "intuition" used in the present paper means, basically, a global synthetic, non-explicitly justified evaluation or prediction. Such a global cognition is felt by the subject as being self-evident, self-consistent, hardly questionable. For instance, the statement: "Through a point outside a line one and only one parallel may be drawn to that line". That statement is

accepted intuitively, according to the above definition. An intuition is somehow a cognitive belief; sometimes, such cognitive beliefs may coincide with scientific accepted statements but sometimes they may not or may contradict them.

In the teaching process such intuitive attitudes must not be ignored. If they are correct they help the learner to acquire and integrate the corresponding scientific concepts. If they are not objectively acceptable they must be eliminated and adequate intuitive representations must be developed instead. If the teaching programme does not care about possible intuitive biases they will continue to mislead the learner despite the conceptual structures he has been taught.

Our point of view is that new intuitive attitudes can be developed only through personal involvement of the learner in a practical activity. Intuitions (cognitive beliefs) cannot be modified by verbal explanations. Therefore, a teaching programme which intends to develop an improved and efficient intuitive background for probability concepts and strategies along with the corresponding formal knowledge, must provide the learner with frequent opportunities to live actively, even emotionally stochastic situations. In such situations, he will confront his plausible expectations with empirically obtained outcomes.

The present research was chiefly devoted to evaluating the indirect effect of a course on probabilities on typical probabilistic, intuitively based, misconceptions.

The teaching programme itself, as a whole, was intended to determine the capacity of children 10-13 years old, to assimilate, correctly and productively, basic concepts and solving procedures in probability. According to our point of view the children have been presented with various situations which offered them the opportunity to be active in calculating probabilities, predicting outcomes, using operations with dice, coins and marbles for watching, recording and summing up different sets of outcomes. It is that reciprocal dynamics of theoretically computed probabilities and observed relative frequencies that may best contribute to the development of efficient probabilistic intuitions.

METHOD

The teaching programme (12 lessons) included: the concepts of certain, possible and impossible events, outcomes and events in a chance experiment: the concepts of chance and of measuring chances, the concepts of probability and relative frequency and the relation between them; counting outcomes; simple and compound events and their probabilities.

The subjects: a) the programme has been attended by 285 pupils; 70 grade 5 (2 classes); 160 grade 7 (5 classes) and 55 grade 7 (2 classes). Parallel

control classes have also been investigated: 50 pupils (2 classes) grade 5; 200 pupils grade 6 (5 classes) and 55 pupils (2 classes) grade 7; (305 subjects in total in the control classes).

The assessment of the teaching effects has been obtained by means of two questionnaires: Questionnaire A was devised to assess to what extent the pupils, in the experimental classes, have assimilated and were able to use the concepts and procedures they have been taught. It has been administered only to the experimental classes. Questionnaire B was devised to assess the indirect effect of instruction on the children's intuitively based misconceptions. That questionnaire did not require special knowledge in probabilities and was administered to both experimental and control classes.

THE QUESTIONNAIRES

Questionnaire A (which will be described only briefly in the present paper) contained problems referring to all the notions mentioned above (the experimental lessons) such as: giving examples of certain, possible and impossible events; the probability of extracting a marble of a certain colour, when knowing the number of marbles of different colours contained in a box; the probability of various, simple and compound, events when rolling one or two dice; to give examples of simple and compound events in stochastic experiments; to sovle a problem based on proportional reasoning.

Questionnaire B, directly related to the topic of our paper, is presented below:

1. In a box there are 4 red, 3 green and 2 white marbles. How many marbles has one to draw in order to be sure that one will get at least one marble of each colour?

2. Joseph endeavours to enter the classroom each day by putting his right foot first. He claims that this increases his chance of getting good marks. What is your opinion?

3. Shula filled up once a lottery form with the following numbers: 1; 7; 13; 21; 22; 36, and she won. Therefore, she claimed that she must always play the same group of numbers, because it turned out to be a lucky one. What is your opinion on this?

4. Ruth prefers, when she participates in a lottery, to choose consecutive numbers like 1, 2, 3, 4, 5, 6. She claims that in this way she increases her chance of winning. On the other hand Jenny claims that the chance of getting six consecutive numbers like 1, 2, 3, 4, 5, 6 is smaller than that

of getting a random sequence of numbers. She says that a lottery is something chancy and therefore there is no chance of getting a sequence of consecutive numbers. What is your opinion with regard to the two attitudes, that of Ruth and that of Jenny?

5. Gill has participated in a weekly lottery during the last two months. So far he has never won but he has decided to go on for the following reason: "Lottery is a game based on chance, sometimes you win sometimes you lose. I have already played many times and I have never won. Therefore, I am sure more than before that I shall win in one of the next games". What is your opinion with regard to Gill's explanation?

6. Inbar has 100 white marbles and 50 black ones in a box. Itay has in his box 200 white marbles and 100 black ones. Without looking, each of them takes a marble from his own box. Compare their chances of extracting a black marble. Mark with an x the correct answer:

A. Inbar has a bigger chance of extracting a black marble.
B. Itay has a bigger chance of extracting a black marble.
C. Their chances of extracting a black marble are equal.
D. It is impossible to get an answer from the above data.

7. Gilla is 10 years old. In her box, there are 40 white marbles and 20 black ones. Ronit is 8 years old. In her box there are 30 white marbles and 15 black ones. Each of them draws one marble from her own box, without looking. Ronit claims that Gilla has a greater chance of extracting a white marble because she is the older one, and therefore she is the cleverest of both of them. What is your opinion on this?

8. Uri has, in his box, 10 white marbles and 20 black ones. Guy has in his box 30 white marbles and 60 black ones. They play a game of chance. The winner is the child who pulls out a white marble first. If both simultaneously take out a white marble no one is the winner and the game has to go on. Uri claims that the game is not fair because in Guy's box there are more white marbles than in his box. What is your opinion on this?

Question 1 checks the capacity of the subjects to identify an outcome which is certain in a stochastic experience. Questions 2 and 3 refer to the idea of luck, in fact to the acceptance of a prediction made in disregard to a fair evaluation of chances. Question 4 refers to a typical probabilistic misconception called by Twersky and Kahneman (1972, p.134) "representativeness". A sequence of consecutive numbers appears, intuitively, to be less frequent in chance extractions than a random one because the first seems to be less "representative". Question 5 deals with what is called a negative recency effect. The intuitive conception on which that effect is based is that of a necessary equilibration, in the long run, between the frequencies of various

possible outcomes. It expresses a confusion between the probability of getting a single result (which, in fact, does not depend on the previous outcomes) and the probability of getting an expected set of outcomes which depends on the structure of the set (see Fischbein, 1975, pp.58-62). Question 7 checks a misconception described by Falk et al. 1980: children may assign a role, in chance events, to the personal qualities of the player, though, objectively such an effect does not exist. Questions 6 and 8 deal with proportioal reasoning. Proportional reasoning is considered to be a basic tool of probabilistic reasoning. We wanted to check this relation.

RESULTS

We intend to describe in detail the data obtained with Questionnaire A in a different paper. Nevertheless, a brief presentation of the basic results obtained with A seems to be necessary (the conceptual gains as a direct result of instruction).

Questionnaire A. There is a steady increase with age of proportions of correct answers, with differences reaching as big as 30-40% between two consecutive grades. The majority of the children of grade 5 did not perform well at most of the items (except those requiring examples of certain and possible events). Items referring to computations of probabilities of simple and compound events were solved by 60-70% of the subjects in grade 6 and by 80-100% in grade 7. For instance with regard to the question: "What is the probability of obtaining "1" when rolling a die?" we obtained the following percentages of correct answers: 53.5 in grade 5; 87.5 in grade 6; 100 in grade 7. The probability of obtaining a number smaller than 6: 32.4 in grade 5; 63.7 in grade 6; 92.3 in grade 7. Computations referring to compound events expressed in sums were solved by 10-20% of the subjects in grade 6 and by 30-40% in grade 7 (Ex.: One rolls a pair of dice and one sums up the figures obtained. What is the probability of obtaining an even number? Percentages of correct answers: 2.8 in grade 5; 16.1 in grade 6; 44.2 in grade 7). To the question referring to proportions, the following percentages of correct answers have been obtained: 1.4 in grade 5; 23.2 in grade 6; 44.2 in grade 7.

Questionnaire B. Question 1 refers to the notion of certain events. The data appear in Table 1. This question does not refer explicitly to probabilistic estimations. But, as a matter of fact, it checks the capacity of the child to produce, mentally, a situation which corresponds to a certain event and this is one of the basic concepts of the theory of probabilities.

As one can see by inspecting Table 1, there is a steady improvement with age for both control and experimental classes. The experimental classes got, at each level, higher percentages of correct answers than the control ones.

TABLE 1

Percentages of correct and incorrect answers for question 1*

Answers	Grades 5 E	5 C	6 E	6 C	7 E	7 C
Correct: Eight marbles	14.1	2.0	31.5	21.8	53.8	45.5
Incorrect number of marbles:						
3	23.9	26.5	23.2	23.8	11.5	18.2
6	12.7	22.4	11.3	13.9	19.2	25.5
2 or 1	32.1	34.7	9.6	7.9	5.8	5.4
7	1.4	0.0	3.6	1.5	5.8	1.8
Total incorrect answers	70.4	83.6	47.7	47.1	43.3	50.9
No answer	15.5	14.4	20.8	31.1	2.9	3.6

*E stands for experimental and C for control classes.

Comparing the figures, it turns out that factor age is more effective than factor instruction in the given conditions (control classes grade 7 perform better than experimental classes grade 6).

Question 2: The subjects have been asked whether entering the classroom with the right foot will increase their chances of getting high marks in the respective day (see Table 2).

TABLE 2

Percentages of correct and incorrect answers, Question 2

Answers	Grades 5 E	5 C	6 E	6 C	7 E	7 C
"No" correct	60.0	81.4	89.9	72.3	97.3	83.0
"Yes" (incorrect)	9.0	11.8	3.6	22.8	1.9	13.8
No answer	31.0	6.8	6.5	4.9	0.8	3.2

Generally, this misconception has been, spontaneously, rejected by most of the children at each age level. In the experimental classes there is an increase with age of the proportions of correct answers, an important leap taking place from grade 5 to grade 6. It is clear that in grade 5 instruction did not produce any improvement in those children who first hesitated or accepted the misconception. On the contrary, much more non-instructed children rejected the superstitious claim than the instructed ones. According to this data it turns out that the teaching process - probably inadequate to that age - has had a negative effect on the children's chance estimations!

On the other hand, in grades 6 and 7 the experimental classes performed better than the control ones - although the differences were not very high. Nevertheless it is a very important result: it proves that the process of teaching probabilities may have not only a formal-instructive effect (transmission of knowledge) but also an indirect positive effect on the children's intuitive estimations in uncertain situations.

Question 3 asked whether in a lottery game a certain group of numbers which once won may be considered as being an especially fortunate one (see Table 3).

TABLE 3

Percentages of correct and incorrect answers - Question 3

'The fortunate number'

	Grades 5		6		7	
Answers	E	C	E	C	E	C
"No" (correct)	21.1	12.2	52.4	40.6	53.9	52.7
"Yes" (incorrect)	0.0	4.1	2.4	6.9	0.0	7.3
The same number cannot win any more (incorrect)	47.9	63.3	29.2	46.0	21.2	29.1
No answer	31.0	20.4	16.0	6.5	24.9	10.9

A highly frequent mistake consists of considering that a number which once won will not win any more. In both control and experimetal classes there is a clear drop with age of this type of basic misconception. Age seems then to be an essential factor. Very few subjects admit the possibility of a lucky number. At each level, experimental classes perform better (more correct answers) than control ones.

Question 4 asked about the chances that a group of consecutive numbers will win in a lottery game. See Table 4 for results.

TABLE 4

Percentages of answers referring to the probability of a set of consecutive numbers winning in a chance game

	Grades 5		6		7	
Answers	E	C	E	C	E	C
There is no relation (correct)	21.2	38.8	27.4	24.2	32.7	12.8
Random numbers have a higher probability of winning (incorrect)	28.2	34.7	44.0	52.5	42.3	41.8
Lottery is a chancy game (ambiguous)	5.6	10.2	16.1	10.4	25.0	40.0
No answer	45.0	16.3	12.5	12.9	0.0	5.4

Many children do not refer directly to the probability that a group of consecutive numbers will win but give an ambiguous answer ('lottery is a chancy game') which may be interpreted in two opposite ways: (a) A lottery is a chancy game and, therefore, the nature of the numbers does not influence the probability of winning or not; (b) A lottery being a chancy game, a group of consecutive numbers has a smaller chance of being extracted.

The proportion of clearly correct answers (the possibility of a set of numbers being extracted does not depend on its content) is higher in the experimental classes 6 and 7 than in the corresponding control classes. The difference is impressive in grade 7. One can also see that in the experimental classes the proportions of correct answers grow steadily with age (grade) while in the control classes there is a decrease with age in the proportions of subjects who answer correctly. The presence of the ambiguous type of answer ('A lottery game is chancy') makes a further interpretation of the results rather difficult. One may say that in general, with age, children become more aware of the non-predictability of chance phenomena (as opposed to certain ones). But, at the same time, the internal conflictive nature of the intuition of chance becomes more acute with age. Intuitively, a lottery game being a chancy

one, every set of numbers must have the same probability of winning. At the same time, one can hardly accept, intuitively, that the set 1, 2, 3, 4, 5, 6 has the same chance of winning as an "irregular" group of numbers. The picture in grade 7 is highly instructive. In the control classes about 40% of the subjects prefer non-consecutive numbers, about 40% affirm that a lottery is a chancy game (ambiguous) and only 12.8% admit, explicitly, the independence of chance in relation to the nature of the numbers. In the experimental classes, again, about 40% prefer a random set of numbers. But this time only 25% answer ambiguously while 32.7% affirm explicitly the independence of probability. This is a drastic change of a very complex intuitive attitude, as an indirect effect of an instructive process.

With regard to the categories of answers which may be characterised as ambiguous ones, one can see that there is a drop at grade 7 especially in experimental classes.

Question 5 was intended to check the childrens' attitude with regard to the same problem as question 4: the independence of probabilities. This time, the subjects had to refer to the independence of probabilities in a sequence of trials which, objectively, leave the probabilities unchanged. (Gill has not won so far, though he has tried many times. He is convinced that he must win in one of the following games). See Table 5 for results.

TABLE 5

Percentages of various categories of answers to question 5: previous successive losses may increase the probability of a further positive outcome?

	Grades					
	5		6		7	
Answers	E	C	E	C	E	C
"No" (correct)	18.3	18.4	44.6	33.7	73.1	49.0
"Yes" (error)	12.7	22.4	10.7	21.3	5.8	14.5
"Lottery is chancy"	18.3	28.5	23.8	23.2	9.6	14.5
Undecided ("possible")	9.9	12.2	14.9	16.8	7.7	18.2
No answer	40.8	18.5	6.0	5.0	3.8	3.8

There is a clear and steady increase with age (grade) in proportions of correct answers ("The odds are the same at every trial") in both experimental and control classes. On the other hand, in grades 6 and 7 the proportions of correct answers in the experimental classes are neatly higher than in the control classes. In grade 5 they are equal. It is, therefore, clear that

both factors, age and instruction, are acting in the same direction, leading to a more adequate intuitive approach with regard to chance estimations in sequences of trials. We must again emphasise that instruction acted here only indirectly: no direct mention has been made during the experimental lesson of that matter.

Question 6 deals with the comparison of probabilities based on a comparison of ratios: 100 white and 50 black marbles versus 200 white and 100 black marbles. The subject has to compare the odds to extract a black marble. (See Table 6).

TABLE 6

Percentages of correct answers when comparing the odds to extract a black marble from (100 w; 50 b) versus (200 w; 100 b)

	Grades					
	5		6		7	
Answers	E	C	E	C	E	C
The odds are equal (correct)	21.1	22.4	42.9	51.5	80.8	85.5
In 200/100 the odds are bigger (error)	32.4	20.4	36.9	20.3	15.4	9.1
In 100/50 the odds are bigger (error)	8.5	6.1	6.5	4.0	1.9	0.0
One cannot decide (error)	11.3	12.9	6.5	21.8	0.0	3.6
No answer	26.3	8.2	7.2	2.4	1.9	1.8

There is an increase with age (grade) in proportions of correct answers in both experimental and control classes almost doubling themselves from grade to grade. The age factor acts very strongly in this respect. But, at the same time, at all three levels considered, the control classes performed better than the experimental ones: Instruction had rather a negative effect. Though the differences are not very high, this tendency seems to be a systematic one.

Specifically, there are higher proportions of children who choose the box with 200 W/100 B as favouring the extraction of a black marble. This means that for many chidlren (20-30%) the absolute quantities are still decisive with regard to the probability of chance events.

Question 7. Do intelligence and age influence the outcome of a chance event? See Table 7 for results.

TABLE 7

Percentages of various categories of answers -
"Do intelligence and age influence the outcome of a chance event?"

Answers	Grades 5 E	5 C	6 E	6 C	7 E	7 C
"No" (correct)	49.4	63.2	67.8	84.2	82.6	87.3
"Yes" (error)	14.1	10.2	22.6	9.9	9.6	3.6
No answer	6.5	26.6	9.6	5.9	7.8	9.1

The majority of subjects, at each level, reject the claim that intelligence and age may affect the outcome of a purely chancy game (except the experimental classes of grade 5). There is an increase in the proportion of correct answers with age in both experimental and control classes. But, very surprisingly, there were higher percentages of correct answers in the control classes. Again, a certain factor (objectively a non-relevant one) is considered by a number of children as being effective in deciding the outcome in a chance situation. Instruction seems to have played rather a negative role with regard to that problem.

The last question (the 8th) was devoted again to a comparison of odds based on a comparison of ratios (10/20 versus 30/60) but using a different wording from that used in question 6 (see Table 8). Does the fact that in (30W/60B) there are more white marbles than in (10W/20B) determine a higher probability of extracting a white marble from the first box?

TABLE 8

Categories of answers in percentages. Comparison between the chances of extracting a white marble from (10W/20B) versus (30W/60B)

Answers	Grades 5 E	5 C	6 E	6 C	7 E	7 C
The chances are equal (correct)	21.1	28.6	53.0	50.4	69.2	72.7
The equality of ratios made explicit*	(11.3)	(10.2)	(38.7)	(26.7)	(65.4)	(61.8)
Higher chances for 30W/60B (incorrect)	35.2	40.8	35.7	35.1	17.3	12.7
No answer	43.7	30.6	11.3	14.5	13.5	14.6

* The figures in paranthesis are extracted from those referring to the previous category (the chances are equal) in which they are also included.

It seems that the teaching of probability according to our programme did not improve the capacity of the pupils to make use of proportional reasoning for reaching a correct estimation of chances. Again, the picture presented by grade 7 is highly instructive. There are a few more subjects, at that level, in the experimental classes, who are able to justify their correct answer by referring explicitly to the equivalence of ratios. But, despite this, there are also more subjects in the control classes who answer correctly. The differences are very small, statistically non-significant. It is clear that the teaching programme did not improve the global estimation of odds when proportionality was involved. But some of those pupils who were ready to estimate correctly the equality of chances have learned, as a result of our programme, to use, explicitly, the formal justification (equality of ratios). Considering only those subjects who justified explicitly their correct answers one can see that, at each age level, there were more subjects in the experimental classes than in the control classes who were able to offer such an explanation.

The factor age seems to play a decisive role: the increase with age in the proportions of correct answers is evident, the great leap taking place from grade 5 to grade 6.

SUMMARY AND CONCLUSIONS

1. There is a clear increase with age (grade) in proportions of correct answers in both experimental and control classes at all the items. While 5th grade children (10-11 years old) seem not to be intellectually ready to assimilate the basic concepts and procedures of probability reasoning, 6 graders seem to be mentally ready for that task. About 60-70% in grade 6 and about 80-100% in grade 7, performed well at almost all the items referring to elementary questions with simple and compound events.

2. With regard to the indirect effect of teaching probabilities on the intuitive estimations of children in uncertain situations the picture is rather complex:
a) There were two typical biased intuitions on which the teaching process has had a clear improving effect (in grades 6 and 7): the 'representativeness' tendency, described by Tversky and Kahneman and the well known negative recency effect. As a matter of fact, there were still about 40% of subjects in grades 6 and 7 who believed that a set of consecutive numbers had a smaller chance of winning than an 'irregular' set of numbers. But, at the same time, there were more subjects who answered correctly in the experimental classes than in the control ones.

A third question with regard to which the teaching process had an improving effect was that referring to a superstition: would the fact of entering the class first with the right foot ensure higher marks? That claim is rejected by most of the pupils, but in grades 6 and 7 more subjects from the experimental classes than from the control ones rejected it.

There was one item with regard to which the control classes performed clearly better than the experimental ones. Do age and intelligence increase the chances of winning in a (fair) chancy game? The majority of subjects rejected the possibility of such an influence. Nevertheless it was surprising to see that while 84.2% of subjects answered correctly in the control classes, in the experimental classes only 67.8 answered correctly (grade 6).

A possible explanation may be the following: There is an intuitive correct estimation in children (aged 12-13) of the fairness of a chancy game (in this case based on proportions). Instruction may have drawn the attention of the subjects to the fact that even in chancy games, some kind of prediction is possible based on reasoning and knowledge. Perhaps some of our subjects have been confused by this intervention. This phenomenon deserves further research.

3. Two items were devoted to proportional reasoning. In relation to both items, the experimental classes performed, generally, worse than the control ones. Questionnaire A referred also to a proportion problem and even in grade 7, 60% of the subjects were not able to answer it correctly. These findings are in accordance with those frequently mentioned in recent years in relation to the same topic. (See Karplus, 1976; Hart, 1981, pp.66-101). In our opinion, an important theoretical conclusion may be drawn from the above data: Probabilistic thinking and proportional reasoning are based on two distinct mental schemata. A progress obtained, as an effect of instruction, in one direction does not imply an improvement on the other. Certainly, probability computations may require ratio comparisons and calculations, but probability, as a specific mental attitude, does not, necessarily imply a formal understanding of proportion concepts. This finding is in accordance with other results obtained before. Yost et al. (1962), Goldberg (1966) have shown that one can identify, even in preschool children, correct probability estimations while proportional reasoning is still deficient. But it is also possible that at a very basic intuitive level, the two schemata share the same roots, namely an intuition called by us the intuition of relative frequency, which is revealed especially in probability learning tasks (Fischbein, 1975, pp.58-65). It refers to the subject's capacity to adapt intuitively his reactions to the relative frequencies of a group of stimuli. Such a common base may explain

Ojemann's finding quoted above. But, in fact, the two forms of reasoning, the probabilistic and the proportional ones, include, besides this hypothetical common root, clusters of intuitions and technical procedures which make them distinct. By emphasizing (via systematic instruction) specific probability viewpoints and procedures one may disturb the subject's proportional reasoning, still fragile in many adolescents. When we implemented our program we were not yet aware about that danger. However, we suppose now that, by using adequate means, this negative side effect can be avoided.

References

[1] Falk, R., Falk, R. and Levin, J. (1980) A potential for learning probability in young children. Educational Studies in Mathematics, 11, 181-204.

[2] Feller, W. (1968) An introduction to probability theory and its applications. John Wiley, New York.

[3] Fischbein, E. (1975) The intuitive sources of probability thinking in children. D. Reidel, Dordrecht.

[4] Goldberg, E. (1966) Probability judgement by pre-school children; task, condition and performance. Child Development, 37, 157-167.

[5] Hart, K.M. (Editor) (1981) Children's understanding of mathematics. John Murray, London.

[6] Kahnemann, D. and Tversky, A. (1972) Subjective probability; a judgement of representativeness. Cognitive Psychology, 3, No.3, 430-454.

[7] Karplus, R. and Karplus, E. (1971) Proportional reasoning and control of variables. Cambridge, Mass.: Division for Study and Research in Education.

[8] Ojemann, R.H., Maxey, E.J. and Snider, B.C. (1965) Effects of guided learning experiences in developing probability concepts at the fifth grade level. Perceptual and Motor Skills, 21, 415-427.

[9] Piaget, J. and Inhelder, B. (1951) La genèse de l'idée de hasard chez l'enfant. P.U.F. Paris.

[10] Tversky, A. and Kahneman, D. (1973) Availability: a heuristic for judging frequency and probability. Cognitive Psychology, No.5, 207-232.

[11] Yost, P., Siegel, A.E. and Andrews, J.N. (1962) Non-verbal probability judgement by young children. Child Development, 33, 769-780.

A P P E N D I X

Chi square tests for comparisons of frequencies of correct answers between experimental and control classes separately for each grade level (df=1)! and each item of Questionnaire B.

Grades

Questions	5		6		7	
	Chi square	P	Chi square	P	Chi square	P
1	3.91	P<0.05	3.92	P<0.05	N.S.	
2	5.63	P<0.02	24.6	P<0.001	5.39	P<0.05
3	N.S.		4.19	P<0.05	N.S.	
4	N.S.		N.S.		5.20	P<0.05
5	N.S.		4.03	P<0.05	5.53	P<0.02
6	N.S.		N.S.		N.S.	
7	N.S.		19.80	P<0.001	N.S.	
8	N.S.		N.S.		N.S.	

STOCHASTICS TEACHING BASED ON COMMON SENSE

H-J Bentz, University of Osnabrück

The following list could give a rough description of contents and goals of usually practised stochastics courses (for simplicity, this list will later also serve as a guide-line for this presentation):

- Discussion of the properties of P (Probability), comparison with f (relative frequency
- Algebra of events (analysis and composition of events)
- Laplace-P, Combinatorics, Bayes'-Formula, Independence
- Random variables (RV), Expectation (E), Variance (Var)
- Distributions, e.g. Binomial, Hypergeometric, Poisson, Gauss or Normal Distribution
- Chebyshev's Inequality, laws of Large Numbers
- Descriptive Statistics
- Tests (Parameters, Independence, Chi-Square, etc.)

Normally the theoretical background is described by Kolmogoroff's Axioms, the Algebra of sets and Measure theory.

We, the teachers, know about difficulties (in learning, understanding and handling) which may arise at nearly every stage during the course. The reasons are various. Partly they come from misconceptions (or no conceptions), misfeelings, lacks in techniques, difficulties in understanding and transforming literal texts,even from psychological obstacles - to name only some of those deficiencies which I observe again and again.

Maybe some of the very important contents of our courses could be taught with more success (concerning the students) and less effort (concerning the teacher) by using a different concept. The following outline (which is based on a joint paper [1] with G. Palm) will describe a stochastics course which essentially differs from the usual ones, namely e.g.:

in theory: it does not use the sample space Ω, the algebra of events, set- and measure-theory, nor Kolmogoroff-Axioms, but is based on the concept of a random-variable and its expected value.

in concepts: Probability is not described by axioms or by comparison with frequencies, but is defined as (is the name of) a certain expected value. So, expectation is not defined as $\Sigma\ p_i\omega_i$, but is one of the fundamental notions itself.

in proceeding: the aim is to "analyse" random variables (RV), i.e. to simplify RV by decomposition in simpler parts. So, composition and decomposition play an important role.

in notation: a certain notation which allows us to describe decompositions is introduced. In addition it is very helpful in transforming literal texts into "mathematical" symbols.

in models: coins, dice, cards, urns, ..., are not used as standard models, but a special "wheel of fortune" serves as first and global model. this "spinner" has a front on which one can write with chalk and wipe clean again. With it (practically) all RV can be represented, even continuous ones.

The reader should not feel deterred by this list of (as I mean partly severe) differences; the following explanations will take care of the contents mentioned at the very beginning. But the key-words in the more detailed description below will be: Random Variable (RV), Expectation (E), Decomposition, Simple and, last but not least, Spinner (SP).

The text is divided in places into global remarks on the theoretical background and the procedure in the classroom. So it is rather more a description than an instruction for the students, although many examples and hints mentioned stem from classroom work.

I. Basic notions: RV, W and E. The space *RV*

A RV is to be looked at as an accidental outcome (= numbers) of an (actual) experiment. The E of complicated RVs has to be determined by use of simpler RVs. So the aim will be to analyse RVs and to decompose them. What do "decomposition" and "simple" mean? The space *RV* of all RVs, what structure should it have?

Each RV X has outcomes. These numbers are collected in a set W(X). Besides this, X has an expected value. So there have to be two mappings, a positive linear functional E : *RV* → IR,
and an unspecified W : *RV* → Power set (IR).

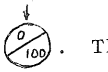
Simple then means W(X) ≤ {0,1}, *decomposition* means linear combination. Therefore *RV* should be an ordered vector space at first (and a Jordan-Banach-Algebra later). More details are given by axioms, see [1].

It should already be stressed here that the course continually utilises the property E(X+Y) = E(X) + E(Y), i.e. the linearity of E.

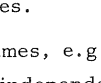
As a model for RVs a spinner is used. With it the main properties of RVs and important stochastic principles are discussed. RVs upon spinners are also called *games*.

II. Learning and playing of games

First examples:

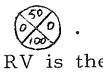
1. Introduce games like 

. The spinner is divided into two equal parts, the possible outcomes are 0, 100, so W(•) = {0,100}. The questions are, what will we expect to win on the average? What amount would we ask for, to offer this game to a player. What amount would we bet on this game? ... Continue with similar examples.

2. Introduce the *sum* of games, e.g. 

+

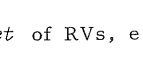
. One will win the sum of the (not necessarily independent) outcomes. W = {0, 40, 60, 100} if both run independently of each other. W may vary according to the type of dependence.

3. Introduce the *multiple* of a RV, e.g. 3. 

. One will win the outcome of the spinner times the factor. Note, this RV is the same as 

.
W = {0, 150, 300}.

4. Introduce the *product* of RVs, e.g. 

• 

. W = {0, 20, 40, -60} if both run independently.

5. (Eventually) introduce games of continuous type, e.g. 
.W = [0,100], here one wins that fraction of the radius which is "cut" by the curve (total length ≙ 100). In this respect this model is in fact different from urns, dice, coins, ...

With these examples and types the students get a first impression about RVs. The main emphasis in this stage lies on the "learning" of RVs. The meaning and properties of E will be examined afterwards.

III. Expectation

Start with games like 

,

, ..., i.e. games with two partitions and W = {0,a}. Here, E is easily found (≙ average outcome). Students either use symmetry or statistical arguments. "Half of the time we get a, half 0, so we expect a/2". After a while, E is immediately found, if the game is of type a • 

.

Continue with the problem X = 

, E(X) = ?

1st solution: X = 40 +  , the latter is well known, so E = 40 + 5.

2nd solution: "statistical argument" : 40·3/4 + 60·1/4 .

3rd solution: X =  +  , both run strictly synchronous. $E=E_1+E_2$.

4th solution: "à la *Huygens*". This method is to be explained by a simpler game, namely Y = 

Two partners are assumed to play against each other, both stake x. Let it be agreed that the one who wins shall give a to the one who loses. This agreement is fair. Everyone has the same chance to get a (loser) or to get $2x-a$ (winner). To have the game Y, set $b = 2x-a$, then one has the same chance for a and b.

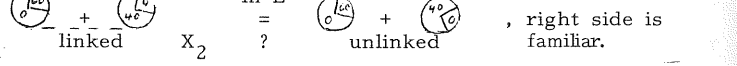
As consequence $x = (a+b)/2$ for the value of the game Y.
In like manner one obtains $x = (a+b+c)/3$ as the value of the game 

(Three partners have to stake x with the agreement to pay b to the second and c to the third, if the first is winner, etc.).

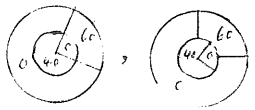
With these preliminaries our X from above equals 
and E = (40 + 40 + 40 + 60)/4.

(For a detailed description and various remarks, see [3], [6]).

Remark: It should be stressed that E remains invariant (as long as the total sum of the outcomes remains unchanged and the partitions are all equal), following Huygens' arguments, even if the single amounts or parts of them are distributed arbitrarily in the spinner-field.

5th solution: 
(X_1 linked X_2 ; in E = ? ; unlinked ; right side is familiar.)

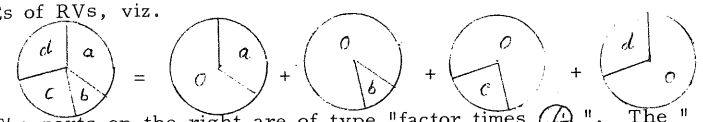
The equality sign does not hold for the games themselves, as W("left") = {40,60} ≠ W ("right") = {0,40,60,100} but it holds in E. This very fundamental fact should be discussed extendedly. Hints for procedure: Look at the left and there at X_1. Its E_1 is already known, $E_1 = 15$. This E_1 is not influenced by the other game, because the outcomes are determined only by chance. The analogue holds for the second game on the left, its E_2 equals 30. So, if both outcomes were to flow into one single "money-box", $E_1 + E_2$ would be expected to be in there. This argument also shows that the total E does not depend on the initial position of the two separate games. Altogether one gets the following understanding: If two games are played independently (with two spinners) the total E is the same as if they are played with one spinner, e.g.



so the rule $\boxed{E(X+Y) = E(X) + E(Y)}$ holds in general. Check your intuition by the question: "How would I like to link the two games X_1, X_2 ?".
(Note that there is a certain correspondence between this "equality in E" of the examples in the brackets, just above, and those stated as remark to Huygens' solution).
Some didactical comments to this 5th solution are given in [2].

IV. Decomposition

All the above solutions are worth mentioning in the classroom. We prefer to continue the course, referring to the last one. The advantage lies in the generality of its application. We use it in a straightforward manner to determine Es of RVs, viz.


, and so on.

The parts on the right are of type "factor times  ". The "" is already familiar to us and is now called *simple* RV.
Its W is a subset of {0,1}, so e.g.  , , are simple games.

V. Notation

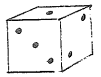
Simple RVs are written in the form $1_{[A]}$. The subscript "A" is an event or a condition (given in everyday language or in mathematical symbols). If A turns out to be true, then $1_{[A]}$ assumes the value 1, and it will be zero in the other case ("characteristic function").

Notes: (a) $1^2_{[A]} = 1_{[A]}$, so $1_{[A]}$ is idempotent. (b) $1 := $.

Examples. $1_{[\text{dropping a coin, heads will be uppermost}]} =$

$1_{[\text{the sum of two dice will be 1}]} = 1_{[2 > 7]} =$

$1_{[\text{rolling a die, we will see a number less than 10}]} =$.



$= D = 1_{[1 \text{ comes up}]} + \ldots + 6 \cdot 1_{[6 \text{ comes up}]}$

$= 1_{[D=1]} + 2 \cdot 1_{[D=2]} + \ldots + 6 \cdot 1_{[D=6]}$ (*"standard dec."*, see below)

$= 1_{[D\geq 1]} + 1_{[D\geq 2]} + \ldots + 1_{[D\geq 6]}$

As the die example shows the decomposition is not unique in general. We draw attention to one specific form. Let X be a RV with $W = \{a_1, a_2, \ldots, a_n\}$. Then the decomposition $X = a_1 \cdot 1_{[x=a_1]} + a_2 \cdot 1_{[x=a_2]} + \ldots + a_n \cdot 1_{[x=a_n]}$

is called the *standard decomposition* of X.

It is a central (and sometimes "funny") point in this concept to find the best (adequate, economic, ...) decomposition, according to the information given in the problem. Therefore, decompositions (and compositions) should be taught in the classroom. Some examples will serve as demonstration.

Example 1. A die is rolled until "3" comes up. Average waiting time ? T is the number of trials which have to be conducted until 3 comes up for the first time. This is a RV with the standard decomposition

$T = 1_{[T=1]} + 2 \cdot 1_{[T=2]} + \ldots + n \cdot 1_{[T=n]} + \ldots$. We find

$E(T) = \frac{1}{6} + 2 \cdot \frac{5}{6} \cdot \frac{1}{6} + \ldots + n \cdot (\frac{5}{6})^{(n-1)} \cdot \frac{1}{6} + \ldots$, which is not immediately calculable by our students.

Compare this with the decomposition

$T = 1_{[T \geq 1]} + 1_{[T \geq 2]} + \ldots + 1_{[T \geq n]} + \ldots$, and

$E(T) = 1 + \frac{5}{6} + \ldots + (\frac{5}{6})^{n-1} + \ldots = 6$ immediately.

Example 2. (see [5]). A bag contains the four numbers 1, 2, 3, 4 which are drawn in random order, every number in the correct place winning £1. Compute the expectation of this promise.

The RV is called F. Its standard dec. is given by

$F = 4 \cdot 1_{[1234]} + 2 \cdot 1_{[1243 \text{ or } 1324 \text{ or} \ldots \text{or } 4231]} + 1_{[1342 \text{ or} \ldots \text{or } 2314]}$

$E(F) = 4 \cdot \frac{1}{24} + 2 \cdot \frac{6}{24} + \frac{8}{24} = 1.$

Compare this with

$F = 1_{[1 \text{ at the first draw}]} + 1_{[2 \text{ at second}]} + \ldots + 1_{[4 \text{ at last draw}]}$

$E(F) = 1/4 + 1/4 + \ldots + 1/4 = 1.$

Of course, the standard decomposition is not always a "bad" one, but it is suggested in the "classical" courses because there we usually start with the sample space. In our course, decomposition is systematically trained, and there is no "favourite" generally given decomposition. (Look back again at Example 2, but now with 50 numbers, say. Working with the standard dec. soon turns out to be hopeless, so one is forced to look for a more economical one).

Remarks: The notation $1_{[\ ]}$ combined with the principle of (linear) decomposition seems to be very helpful in transforming literal texts into a form where the E-rule finds application. Remember that subscript-events are allowed to be formulated as well in ordinary language as in mathematical symbols. So this takes greatly into account the different abilities of the students. Very often it is possible to transform the problem directly in terms of

simple RVs (see Example 2, second solution). This was one reason for the label "Common Sense" Probability.

VI. Logical rules

It may happen that the event or condition "A" itself is a composition of other, simpler ones. To handle such cases the following rules should be kept in mind:

(1) $\boxed{1_{[A]} = 1 - 1_{[\text{not } A]}}$ (2) $\boxed{1_{[A \text{ and } B]} = 1_{[A]} \cdot 1_{[B]}}$

(3) $\boxed{1_{[A \text{ or } B]} = 1_{[A]} + 1_{[B]} - 1_{[A \text{ and } B]}}$ (proofs by trivial computation).

These identities reflect certain rules, which are well known to students but in terms of set algebra. However, (1), (2), (3) differ from those in several respects, namely:

1.) They show that characteristic functions of events "behave" like numbers; that simple games carry arithmetical properties.

2.) At this stage, these rules are easily motivated.

3.) Their inherent decomposition should be utilised, and in fact, E applies immediately in rules (1) and (3).

4.) The question if (and under which conditions) E also applies to (2) - in general: when does $E(X \cdot Y) = E(X) \cdot E(Y)$? - directly leads to the concepts of independence and dependence.

So, the algebra of events - often attacked for didactical reasons if posed as an introductory chapter to probability courses - may turn out to be favourable if used as "logical rules" in a context like this.

VII. Probability

In stochastic problems the word 'probability' is frequently used. We should not miss it. In this course *probability of an event* is <u>defined</u> in the following manner: $P(A) := E(1_{[A]})$.

So, the simple RVs provide a connection with 'probability'. P of an event A means the expected value of the attached simple game $1_{[A]}$.

Example: The probability of 5, rolling a die, is found as follows:



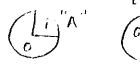
The expected value of the "inner" game is 1/6, so P(5) = 1/6.

(Note: Using the spinner, the distinction between "probability of an event" and "value of the outcome" is obvious).

VIII. Stochastic Independence

Students tend to apply the E-rule to $1_{[A \text{ and } B]} = 1_{[A]} \cdot 1_{[B]}$ as well, getting $E(1_{[A]} \cdot 1_{[B]}) = E(1_{[A]}) \cdot E(1_{[B]})$.

At this point a discussion about "independence" should start. Here we use two spinners. The game $1_{[A \text{ and } B]}$ played on two unlinked spinners, e.g.



has an E which differs (in general) from the E of $1_{[A \text{ and } B]}$, played on a single spinner, e.g.



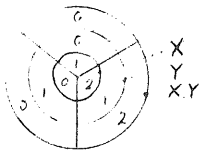
There is no doubt about this, because the information "A has come out to be true" does not influence our betting on "B" in the first case, whereas the same information changes our betting in the second case.

Note: The "overlap" of "A" and "B" in the second case may be chosen in such a manner that both E's are equal, and then the information on "A" does not tell anything about "B", just as in the unlinked situation, viz.



This leads to the definition: Two simple RVs X, Y are *independent* if $E(X \cdot Y) = E(X) \cdot E(Y)$ holds. Otherwise they are *dependent* (on each other). In general - if X, Y are not simple - the definition uses their standard-dec. (in the discrete case): $X = \sum_i a_i \cdot 1_{[X=a_i]}$, $Y = \sum_j b_j \cdot 1_{[Y=b_j]}$. X and Y are called *independent* if $E(1_{[X=a_i]} \cdot 1_{[Y=b_j]}) = E(1_{[X=a_i]})$. $E(1_{[Y=b_j]})$ holds for all i,j.

This extension is necessary, because there exist counterexamples, such as



$E(X)=1$, $E(Y)=2/3$, $E(X \cdot Y) = 2/3$ but the information about X changes the betting on Y. Reason for this: $E(1_{[X=2]} \cdot 1_{[Y=1]}) = 1/3$ and $E(1_{[X=2]}) \cdot E(1_{[Y=1]}) = 1/3 \cdot 1/3 = 1/9$.

IX. Variance, Standard deviation

This is introduced as usual. We define $Var(X) := E((X-E(X))^2) = E(X^2) - E^2(X)$, (the second "=" holds by linearity of E), and $\sigma(X) := \sqrt{Var(X)}$.

As an example (which is referred to below) take Bernoulli trials. There are n independent copies of with E = p. Let H_n:= Number of 'successes'.

$H_n = 1_{[\text{success at the first trial}]} + \ldots + 1_{[\text{success at the } n^{th} \text{ trial}]}$

$E(H_n) = p + \ldots + p = np$

Now to the variance. $Var(1_{[\text{success in the } i^{th} \text{ trial}]}) = E(1^2_{[i]}) - E^2(1_{[i]}) = p^2 - p = p(1-p) =: pq$.

Then $Var(H_n) = \sum^n Var(1_{[i]}) = npq$.

X. Conditional Probability, Bayes' rule

Define (1) $P_B(A)$ (or $P(A|B)$) $:= \frac{P(A \text{ and } B)}{P(B)}$, $P(B) \neq 0$.

(2) $E_B(X) := \sum_i a_i \, P_B(X=a_i)$, if $W(X)=\{a_1, a_2, \ldots, a_n\}$.

(1) gives $P(A \wedge B) = P_B(A) \cdot P(B)$ and by reasons of symmetry

(2) gives $P(A \wedge B) = P_A(B) \cdot P(A)$. From these two, one has the important equality (3) $P_A(B) \cdot P(A) = P_B(A) \cdot P(B)$, which actually contains Bayes' rule.

Now, if $1_{[A_1]} + \ldots + 1_{[A_n]}$ is a mutually exclusive decomp. of 1, and B is an arbitrary event, then

a) $1_{[B]} = 1_{[A_1 \wedge B]} + \ldots + 1_{[A_n \wedge B]}$, trivial.

b) $P(B) = P(A_1 \wedge B) + \ldots + P(A_n \wedge B)$, apply E to a).

By definition (1):

c) $P(B) = P(A_1)P(B|A_1) + \ldots + P(A_n)P(B|A_n)$

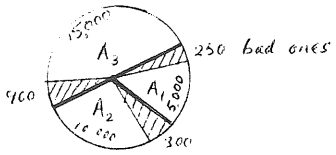
d) $P(A_i|B) = P(A_i)\, P(B|A_i) \,/ P(B)$, according to (3).

which gives *Bayes' Rule* " $= \frac{P(A_i)\, P(B|A_i)}{P(A_1)P(B|A_1)+\ldots+P(A_n)P(B|A_n)}$

How do (should) we work with Bayes' rule? We use the spinner, viz.

Example: Three machines A_1, A_2, A_3 produce nails, 5,000, 10,000, 15,000 per day, respectively. The dross rates are 5%, 3%, and 6% respectively. In the evening one nail is taken by chance. What is the probability that it is "bad"? If it is bad, what is the probability that it comes from A_1?

The situation is modelled with the spinner. The first question corresponds to the expectation that the spinner stops in one of the hatched sectors. The second means it has stopped there and now we are interested in the chance that it is the one belonging to A_1. The solutions are found immediately:



(proportions not correct!)

$P(\text{bad}) = \frac{300+900+250}{5,000+10,000+15,000} = \frac{1,450}{30,000} \approx 4.83\%$.

$P(A_1|\text{bad}) = \frac{250}{250+300+900} = \frac{250}{1,450} \approx 17.24\%$

XI. Conditional Expectation, General product rule

As $E(X \cdot Y) = E(X) \cdot E(Y)$ is not always valid, one is led to look for a similar formula, which holds in general.

(1) It is easy to show $E(X \cdot 1_{[A]}) = E(E_A(X) \cdot 1_{[A]})$. This is a special case of a general rule. To get it, define the conditional expectation. Let X, Y be RVs with standard decompositions $\Sigma\, a_i 1_{[X=a_i]}$, $\Sigma b_j 1_{[Y=b_j]}$. Then $E_Y(X) := \Sigma_j E(X|Y=b_j) \cdot 1_{[Y=b_j]}$ is called the *expectation of X under the condition Y.*

Up to now the condition was of type "A", fixed. Now, A may vary according to the outcomes of Y. Therefore $E_Y(X)$ is not a number, but a RV itself. With this the general product rule reads

(2) $$E(E_Y(X) \cdot Y) = E(X \cdot Y).$$

This equation holds in both cases X, Y dependent, independent (the latter is trivial). (Note: the values of $E_Y(X)$ are uniquely defined except for values b_j of Y with $P(Y=b_j) = 0$.)

This rule allows to calcualte E(XY) without discussing XY.

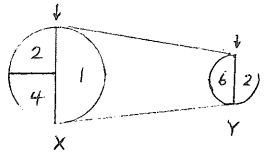
Example: X and Y are linked (e.g.) in a way that the Sp which represents Y runs two times as fast as the spinner-game X:

The dependency is expressed by the table:

	2	6
1	1/4	1/4
2	0	1/4
4	1/4	0

Y

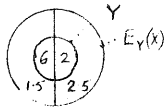
X

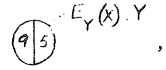


Construction of $E_Y(X)$ by calculating E(X|Y 2) and E(X|Y=6).

$E(X|Y=2) = 1 \cdot \frac{1/4}{1/2} + 2.0 + 4 \cdot \frac{1/4}{1/2} = 2.5$, and similar $E(X|Y=6)=1 \cdot 5$

These values have to be written "upon" Y:



The product $E_Y(X) \cdot Y$ looks like 

, E=7.

So we also know E(XY) = 7, by (2).

XII. Chebyshev's inequality, Laws of large numbers.

Usually, Chebyshev's inequality is easily derived. Nevertheless, a different proof is discussed here, which could be of interest for methodological reasons. Take X with μ and σ. Then $|X-\mu|$ and $1_{[|X-\mu| \geq \varepsilon]}$ are RVs and with real $\varepsilon > 0$ $\varepsilon \cdot 1_{[|X-\mu| \geq \varepsilon]} \leq |X-\mu|$ holds (check the truth-values of both sides.

As $1^m_{[A]} = 1_{[A]}$ we obtain $\varepsilon^m \cdot 1_{[|X-\mu| \geq \varepsilon]} \leq |X-\mu|^m$, $m \in \mathbb{N}$.

Dividing by ε^m and taking expectations (1) $P(|X-\mu| \geq \varepsilon) \leq \frac{E(|X-\mu|^m)}{\varepsilon^m}$

which contains *Chebyshev's inequality* (m=2).

Note: The proof holds as well if X is a continuous RV.

Now, *Bernoulli's theorem* is soon derived by (1). Remember the example in IX and define $h_n := H_n/n$. Then $\sigma^2_{h_n} = pq/n$, $E(h_n) = p$. Inequality (1) and the identifications $X=h_n$, $\mu=p$, m=2, give $P(|h_n-p| \geq \varepsilon) \leq \frac{pq}{n\varepsilon^2}$, which is Bernoulli's law.

With similar arguments the weak law of large numbers for a sequence (X_n) of independent RVs with existing μ_n and σ^2_n may be derived if

$\frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 \xrightarrow[n \to \infty]{} 0$ is assumed to hold.

As to the strong law of large numbers take again H_n and h_n. Formula (1), m=4, a rather lengthy (but elementary) calculation shows that $P(\lim_{n\to\infty} h_n = p) = 1$ holds too. (More general statements need different techniques and more effort).

XIII. Probability distribution functions.

Probability distributions may be introduced as usual. But one particular hint should be added here. We have stressed several times that E applies both for independent and dependent sums of RVs. And so, the calculations of the Es of the binomial and hypergeometric distributions run quite analogously, as follows. The binomial RV gives the number of successes in n Bernoulli trials, for example X is the number of successes in a sample of n elements drawn with replacement from a population of N elements of two kinds: N_1 elements labelled "success" and $N-N_1$ labelled "failure".

The hypergeometric RV is the number of successes in a similar experiment, but now drawn without replacement. The decomposition of X and Y may be chosen as follows:

$X = 1_{[\text{success at first draw}]} + \ldots + 1_{[\text{succ. at } n^{th}]}$, $Y = 1_{[\text{succ. at 1st}]} + \ldots + 1_{[\text{succ. at } n^{th}]}$

In both cases $E(1_{[\text{success at the ith draw}]})$ equals $\frac{N_1}{N}$ (=: p).

So E(X) = np and E(Y) = np. For a full discussion of this, see [4].

XIV. Statistics.

To show one important difference between probability (theory) and statistics, we present (in one of the first lessons in this course) a spinner-game which is placed in a way that the students in the class cannot see its face. Now a partition is drawn and two outcomes, 0 and 100 say, are given. The students

know these values, but not the size of the sectors. How can we find the expected value? Let us play! The teacher only tells the outcomes. The longer the list, the better the insight in this game. Later this action may be "justified" in the context of the laws of large numbers.

I feel that this situation (and its generalisations) reflects typical statistical aspects.

Some specific methods which describe 'decision under uncertainty' are worth being taught. Therefore we pick up some statistical tests. I will not omit to refer to two recent papers, which are dedicated to this problem and contain hints for classroom work, see [4], [7].

One example should be mentioned here. Let me enter the "runs", where we consider a finite 0-1-sequence. The number of runs (=maximal subsequences of symbols of like kind) may serve as an indicator of the randomness of the sequence. Too many runs (or too few) are incompatible with the idea of random action. Hence, it is essential to know the expected number of runs in a random sequence. We will derive this by our standard technique, i.e. adequate decomposition and application of E.

Let N, M be the number of zeros and ones, respectively, in the sequence; the total number of elements is L = N+M. Define a RV X which counts the number of runs and a RV Y which is the number of alternations in the sequence. Clearly $1 \le Y \le L-1$ and X = Y+1, so E(X) = E(Y) + 1. Determine E(Y).

$Y = 1_{[\text{alternation from the first to the second symbol}]} + \ldots + 1_{[\text{alt. from L-1 to L}]}$.

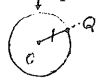
An elementary combinatorial calculation gives

$E(1_{[\text{alt. from i to i+1}]}) = N/L \cdot M/(L-1) + M/L \cdot N/(L-1) = 2\,\frac{NM}{L(L-1)}$, for all i.

Now $E(Y) = (L-1) \cdot 2\,\frac{NM}{L(L-1)} = 2NM/L$, and finally $E(X) = 1 + 2NM/L$.

Note: A full description of this was communicated by Ruma Falk, Jerusalem. Also she has extended the concept of a run to higher-dimensional "sequences". The calculations are similar there, because the above method works well again. See [4].

XV. Final remarks.

1) A last example of the usefulness of the spinning wheel as a model for RVs will be given with this simple game 

i.e. one will win 1 if Q comes to the pointer, and zero otherwise. What is E? The students themselves found the answer. ("If E were > 0, then we could find a sector which includes Q, but has a smaller expectation than "Q". So, E must be zero."). This simple example shows that the conclusion

p "Q") = 0 ⟶ "Q" impossible, does not hold in general. Furthermore, this problem can serve as motivation for 'infinity-problems'.

2) Most of the rules in which E is involved are valid for both discrete and continuous RVs (e.g. Chebyshev's inequality). This is based on the properties of E : *positive linear functional*. Students may already know integrals and that these carry those properties as well. From this point of view, the use of integrals in the context of continuous RVs seems quite natural. If students do not know integrals, our concept provides an introduction of this notion. (It is easier done - with the preparation by this course - as the catchword Stieltjes-, Lebesgue-, Daniell- integral may foreshadow).

3) Combinatorics - which is an important 'pillar' in the stochastics course - is only little influenced by our concept. The notations are easily adapted, the procedure runs as usual. Maybe some exercises will lose their effects, compare Example 2, V., or the 'run-test'-derivations in $\overline{X}$., where the distribution function of X or Y and the successive combinatorial E-calculations were bypassed.

References

[1] H.-J. Bentz and G. Palm (1980) Wahrscheinlichkeitstheorie ohne Mengenlehre, Mathematica Didactica, 3, 167-183.

[2] H.-J. Bentz (1980) Das Prinzip der 'getrennten Kassen' in der Wahrscheinlichkeitsrechnung, in: Beiträge zum Mathematik-Unterricht, Schroedel, Braunschweig.

[3] H.-J. Bentz (1982) Zum Wahrscheinlichkeitsbegriff von Chr. Huygens, to appear.

[4] R. Falk (1981) Some distributions and expectations - simplified, Preprint, Jerusalem. A German translation, entitled Vereinfachte Darstellung einiger Verteilungen und ihrer Erwartungswerte is to appear in: Probleme im Umgang mit dem Zufall, Der Mathematikunterricht, Klett, Stuttgart.

[5] H. Freudenthal (1973) Mathematics as an educational task, Reidel, Dordrecht.

[6] H. Freudenthal (1980) Huygens' foundation of probability, Historia Mathematica, 7, (2), 113-117.

[7] R. Ineichen (1982) Zufällig oder nicht zufällig? in: Probleme im Umgang mit dem Zufall, Der Mathematikunterricht, Klett, Stuttgart (to appear).

A SURVEY OF PROBABILITY CONCEPTS IN 3000 PUPILS AGED 11-16 YEARS

D R Green, University of Loughborough, England

AIMS OF THE PROJECT

The aims of this project, sponsored by the Social Science Research Council from November 1978 to October 1981, were:

a) to survey the intuitions of chance and the concepts of probability possessed by English school pupils aged 11 to 16 years.

b) to establish patterns of development of probability concepts and to relate these to other mathematics concepts, supplementing the work of the SSRC funded project 'Concepts in Secondary Mathematics and Science' based at Chelsea College, 1974-1979 (refs. 1,2).

c) to investigate pupil responses to GCE 'O' Level and CSE probability items, with particular references to sex differences.

This paper concentrates on (a). Full details of the whole project are to be found in a published monograph (Ref. 3).

TEST DEVELOPMENT

A special Probability Concepts Test was devised over a two-yearperiod (1978-80) entailing six pilot versions before achieving the definite form. The Test comprised 26 questions, subdivided into 58 items. The full test and all item facilities are published elsewhere (Ref. 3).

DATA SOURCES

a) Altogether about four thousand pupils were tested which ultimately led to a stratified sample of 2930 pupils. Children in English state comprehensive mixed schools aged 11 to 16 years (i.e. in normal secondary school Years 1 to 5) were tested on a class basis with:

i) Probability Concepts Test (about 1 hour) (Ref. 3).

ii) AH2 Test of General Reasoning Ability (about 1 hour) (Refs. 4,5).

b) Some pupils additionally were given CSMS Tests on Fractions and/or Ratio and Proportion (Ref. 1), and 44 pupils were given individual follow-up interviews after the two main tests.

SAMPLE DETAILS

From the original four thousand subjects, those present for one test only were eliminated and from the remainder a stratified sample was drawn, to reflect the proportions of the various intellectual levels in the school population as a whole. This was based on the AH2 grade distribution which is purported to be as follows (Refs. 4,5).

A	B	C	D	E
Top 10%	Next 20%	Middle 40%	Next 20%	Bottom 10%

The final 'Sample' (N = 2930) was as indicated in Table 1.

TABLE 1: Sample numbers by School Year and by AH2 Grade

School Year	A	B	C	D	E	Overall	Mean age on taking AH2 test
1	64	128	256	128	64	640	11.7
2	67	134	268	134	67	670	12.7
3	67	134	268	134	67	670	13.7
4	54	108	216	108	54	540	14.7
5	41	82	164	82	41	410	15.7
TOTAL	293	586	1172	586	293	2930	

All the participating schools were selected from volunteers, obtained by writing to all relevant schools within 50 miles of Loughborough, England.

TESTING PROCEDURE

In most cases pupils took the Probability Concepts Test first and one or two weeks later took the AH2 Test of General Reasoning Ability. (The mean time lags between the tests for years 1 to 5 were respectively 11, 17, 12, 16, 15 days). Most testing was carried out from September to December 1980, although some took place in January to March 1981.

In order to minimise possible reading difficulties, and to ensure that pupils attended properly to each of the 26 questions of the Concepts Test, and to prove comparable conditions to all classes tested, it was decided that:

i) The test would be administered by members of the project team, and preceded by a prepared introduction explaining its nature and purpose.

ii) The test would be read out to all 1st, 2nd and 3rd year classes (ages 11-14), total time being 48 minutes.

iii) The reading out would be so controlled as to give the same amount of time to a question for all pupils.

iv) The Test would be read out to any 4th and 5th year classes (ages 14-16) for which this was deemed advisable by the class teacher. Alternatively, adequate time would be provided for silent working (55 minutes).

PROBABILITY CONCEPTS TEST - INDIVIDUAL QUESTIONS

In this paper just five of the twenty-six questions comprising the Concepts Test will be briefly discussed.

Question 1

A small round counter is red on one side and green on the other side. It is held with the red face up and tossed high in the air. It spins and then lands. Which side is more likely to be face up, or is there no difference? Tick the correct answer:

(A) The red side is more likely ☐
(B) The green side is more likely ☐
(C) There is no difference ☐
(D) Don't know ☐

Year	A	B	C*	D	Omits	Errors
1	18	28	45	9	0	0
2	15	24	55	6	0	0
3	11	20	66	2	0	0
4	12	13	71	4	0	0
5	11	12	74	4	0	0
All	14	20	61	5	0	0

Note: An asterisk (*) indicates the most appropriate response.

The facility levels for this apparently trivial question indicate that many children aged 11-16 have a very wayward probability intuition. There is some evidence of the phenomenon of 'Negative Recency' operating. It was hypothesised that by emphasising that the counter was red face up a larger number would opt for green being more likely next time than would opt for red. This has clearly been shown for Years 1, 2, 3. The phenomenon virtually disappears in the transition from Year 3 to Year 4, a surprising result by itself.

Question 6

This question comprised five items concerning 'comparison of odds' in the familiar context of coloured counters in bags. The items and the responses are summarised in Table 2. In each case Black was the required colour.

	First Bag		Second Bag			
Item	Black	White	Black	White	Diagram given	Overall facility
6a	3	1	2	1	Yes	88
6b	5	2	5	3	Yes	67
6c	2	2	4	4	Yes	60
6d	12	4	20	10	No	57
6e	3	1	6	2	No	33

TABLE 2: Details of items concerning comparison of odds

To show the nature of the items, 6(d) is presented in full:

Question 6(d)

Two other bags have black and white counters:

Bag G: 12 Black and 4 White
Bag H: 20 Black and 20 White

Which bag gives a better chance of picking a black counter?

(A) Same chance
(B) Bag G
(C) Bag H
(D) Don't know

Why? .....

Year	A	B*	C	D	Omits	Errors
1	8	38	51	2	0	0
2	6	50	41	4	0	0
3	6	62	31	1	0	0
4	5	68	26	0	0	0
5	4	74	20	2	0	0
All	6	57	35	2	0	0

Analysis of both verbal explanations and multiple choice responses was undertaken to throw light on the strategies employed by pupils in solving these items. Although many theoretical strategies can readily be imagined the incidence of any one strategy being consistently applied to the five items in Question 6 is small. The commoner identifiable strategies are indicated in Table 3.

Strategy Number	Strategy Description	Frequency of consistent application	Percentage of Sample
1	Choose the bag with more total counters	41	1.4
2	Choose the bag with more Black counters	221	7.5
3	Choose the bag with greater Black-White difference	92	3.1
4	Choose the bag with greater Black:White ratio	583	19.9
		937	32.0

TABLE 3: Frequency of consistent responses for various theoretical strategies

It is clear that relatively few pupils aged 11-16 years (32%) consistently apply any one pure strategy to all five problems. By far the commonest single pattern observed is for the correct proportional reasoning strategy (20%). This analysis immediately indicates how unlikely it is that a straightforward 'model' of pupil behaviour could be devised. Various strategies are invoked depending on the circumstances. This point was brought out very clearly in individual interviews - for Question 6(a) the almost unanimous reason given was simply "Because A has more Blacks". Only with great perseverance could the investigator occasionally evoke a reason employing the ratio concept, even from pupils who readily supplied just such a reason for 6(d) or 6(e) where such a concept is essential.

A more detailed Year by Year breakdown is provided in Table 4.

Strategy Number	Strategy Description	Perfect pattern frequency as % of Year				
		1	2	3	4	5
1	More counters (Total)	2	2	1	0	0
2	Comparing BLACK (Counting)	13	10	6	4	3
3	B-W (Difference)	6	2	3	3	1
4	B:W (Ratio)	7	11	21	28	42

TABLE 4: Incidence of strategy response patterns by Year

The growth with age of the usage of the Ratio concept (4) and the decline of counting (2) and Differencing (3) is pronounced. Differencing would appear to run parallel to Counting rather than supersede it as the pupils grow older. The growth of the Ratio strategy is on a scale much greater than can be attributed to its replacing Counting and Differencing.

Table 5 shows several other popular response patterns.

Response Pattern Number	Response Pattern 6(a)	6(b)	6(c)	6(d)	6(e)	Frequency	%
(i)	A	A	C	B	B	82	2.8
(ii)	A	A	C	B	C	316	10.8
(iii)	A	A	C	C	A	72	2.5
(iv)	A	A	C	C	C	214	7.3
(v)	A	A	B	B	C	93	3.2

(Underlined responses are correct)

TABLE 5: Popular response patterns not corresponding to a theoretical strategy

Response patterns (i) and (ii) are correct apart from item 6(e) and their relative frequencies show the clear preference for choosing the bag with more black counters (C instead of B). Patterns (iii) and (iv) both are in error by wrongly selecting for item 6(d) the bag with more black counters, and that is the flaw manifested in pattern (v) also. Together, these five common patterns represent about 27% of the Sample. The inference is that many pupils use a mixture of strategies, competently handling problems where only one colour varies in number, but being less sure in cases where both colours' numbers vary, and generally reverting to selecting the bag with more black or greater black/white difference. This would account for perhaps 50% of the population aged 11-16 years. More detailed analysis surprisingly showed that, with the exception of the development of the correct Ratio strategy, percentages were very nearly static with increasing age.

Also, an analysis by AH2 grade was undertaken. It can be seen from Tables 6 and 7 that use of the Ratio concept is very much a function of intelligence, more so than of age - compare with Table 4.

Strategy	AH2 Grade A	B	C	D	E
1 (Total, more)	0	1	1	2	3
2 (Comparing Black)	1	5	8	12	10
3 (B-W difference)	0	2	3	5	5
4 (B:W ratio)	54	33	16	5	2
1 to 4 combined	55	41	28	24	20

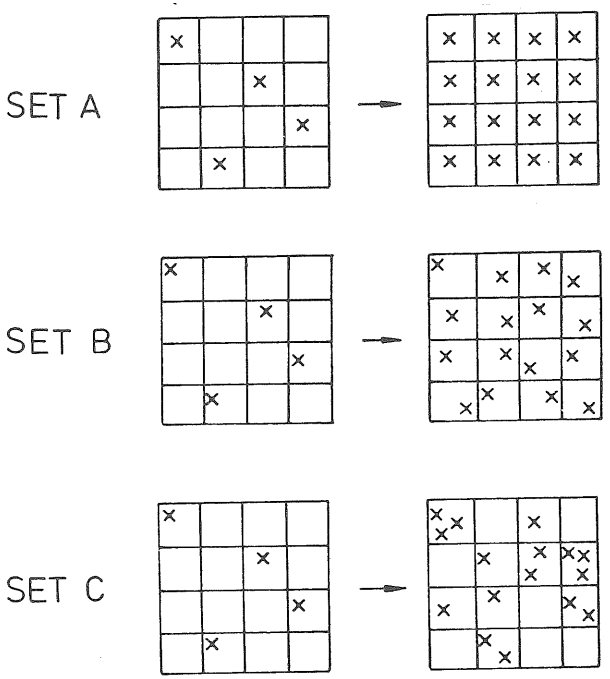
TABLE 6: Percentage of pupils in each Intelligence level whose response patterns match a theoretical strategy

Year	AH2 Grade A	B	C	D	E
1	31	13	4	1	0
2	36	23	6	1	0
3	64	36	15	5	0
4	70	45	25	7	6
5	83	62	43	17	7

TABLE 7: Percentage of pupils in each Intelligence level for each Year whose response patterns correspond to the Ratio strategy

Question 11

The flat roof of a garden shed has 16 square sections. It begins to snow. At first just a few snowflakes fall, then after a while more have landed. Below are three sets of two pictures. Each set shows the pattern of snowflakes building up - first 4 flakes, then 16 flakes.



QUESTION: Which of these sets best shows the kind of pattern you would expect to see as the snowflakes land?

- Set A ☐
- Set B ☐
- Set C ☐
- Sets B and C ☐
- Each kind of pattern is as likely ☐

Year	A	B	C*	B and C	'equal'	Omits	Errors
1	11	12	26	23	27	0	1
2	13	12	26	20	29	0	1
3	13	14	24	25	24	0	0
4	13	16	16	21	33	0	0
5	12	15	18	21	34	0	0
All	12	14	23	22	29	0	0

The astonishing thing about this item is that performance declines with age. These rsults confirm what had been found in Pilot Testing as reported earlier by Green (Ref. 6). It can be hypothesised that we see at work here two opposing tendencies - maturation/experience on the one hand and dominance of mathematical/scientific deductivism on the other which stifles the appreciation of randomness by seeking to codify and explain everything.

Question 20

A teacher asked Clare and Susan each to toss a coin a large number of times and to record every time whether the coin landed Heads or Tails. For each 'Heads' a 1 is recorded and for each 'Tails' an 0 is recorded.

Here are the two sets of results:

CLARE:
0101100110010101101101000111000110110101011001000 1
0101001110011010110011010011001001011101100110111
0101001011001010110001001101011001110111010110011

SUSAN:
1001110111101001110010011100100011101111110101010 1
1110000001000101001000001000110001010000000011001
0000000111110000110101001001001111111010011000110 00

Now one girl did it properly, by tossing the coin. The other girl cheated and just made it up.

(a) Which girl cheated? ANSWER .....

(b) How can you tell?
ANSWER .....
.....

This question was modelled on one suggested by M. Halmos (Ref. 7).

20(a)

Year	Clare*	Susan	Cannot tell	Omits
1	38	53	1	7
2	34	56	2	7
3	30	59	4	7
4	40	46	3	9
5	35	49	5	9
All	35	53	3	8

There is no improvement with age, and for all Years more pupils incorrectly conclude that Susan cheated rather than Clare.

Analysis of the reasons given is also interesting:

20(b)

	Facility	(1) (2) Pattern Regularity		(3) (4) Inequality of O, 1.		(5) (6) Equality of O, 1.		(7) (8) Run Lengths		Other	Omits
Year	(1)+(7)	C	S	C	S	C	S	C	S		
1	21	20	2	8	21	2	1	1	15	17	13
2	21	17	3	7	21	1	0	4	18	13	14
3	23	21	4	4	17	3	0	2	21	14	14
4	28	26	3	3	14	4	0	2	20	16	11
5	28	24	6	1	16	4	0	4	19	12	12
All	24	21	3	5	18	3	1	3	18	14	13

The reasons were generally of the following type, corresponding to the above abbreviated headings:

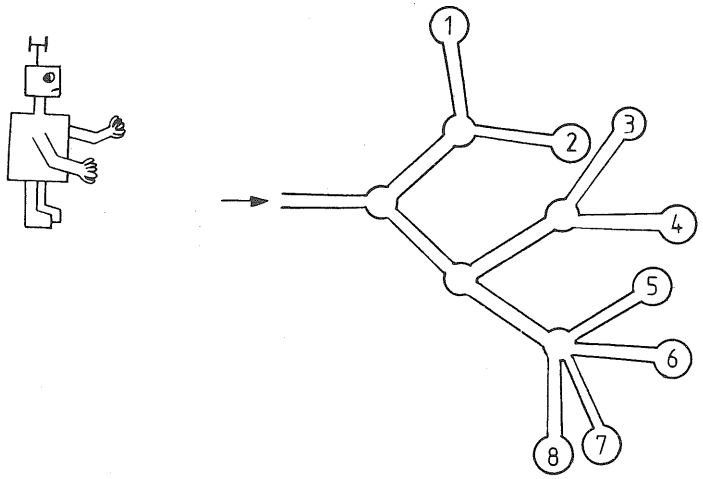
1) Clare too regular pattern (21%)
2) Susan too irregular pattern (3%)
3) Clare too many Is (or Os)/too far from 50-50 proportion (5%)
4) Susan too many Os (or Is)/too far from 50-50 proportion (18%)
5) Clare too close to 50-50 proportion (3%)
6) Susan too close to 50-50 proportion (1%)
7) Clare too short runs (3%)
8) Susan too long runs (18%)

Others + omits (28%)

Only (1) and (7) are valid reasons. Thus the true overall facility of this item is about 24%, and shows only slight improvement with age. Very few pupils thought Susan's pattern too irregular (3%) but their concern is shown in stating that her run lengths are too long (18%).

Question 22

A robot is put into a maze, which it begins to explore. At each junction the robot is as likely to go down any one path as any other (except it does not go back the way it came). There are eight traps at the ends of the eight paths (see picture). In which trap or traps is the robot most likely to finish up, or are all traps equally alike?



ANSWER: .....

Year	'1 and 2'*	1/2	5/6/7/8	'Same'	'Other'	Omits
1	5	7	22	57	6	3
2	4	5	20	62	6	2
3	7	3	19	62	7	3
4	9	7	16	59	7	3
5	13	4	15	59	5	3
All	7	5	19	60	6	3

This item, it was anticipated, would prove difficult for Years 1 to 3 (or 4) but would be considerably less difficult for Year 5 pupils who had met tree diagrams and multiplication of probabilities. However, in the event the item was extremely difficult for all years, the percentage considering all traps equally likely remaining virtually fixed at around 60% at each age level.

Even when treating correct a choice of either Trap 1 or Trap 2 singly the facility does not exceed 17%. Interviews with pupils showed that knowledge

of tree diagrams did not seem to help. Pupils who opted for "all are equally likely" often said that "all have a chance" and "no one can say which" were phrases which implied equality.

The abysmal failure of all Years on this item is in marked contrast to Fischbein et al's results (Ref. 8, appendix 11) in an equivalent but practical setting where facilities in the range 58% to 69% were reported.

PROBABILITY CONCEPTS TEST OVERALL RESULTS

Pupils' responses to the Concepts Test have been analysed by Year, Sex, Intelligence (AH2 grade) and Mathematical Ability ("MA"). As a further step each of the 26 questions was allocated to one of three sub-test categories: Verbal (VB), Combinatoric (CB), Probabilistic (PB).

Each item attracted a mark of 1 if correct and 0 if incorrect, yielding three sub-scores (maxima VB: 15, CB: 6, PB: 29) and a total Test Score (maximum 50). The Test Score breakdowns are shown in Table 8.

	TEST SCORE		VB SCORE		CB SCORE		PB SCORE	
Year	Boy	Girl	Boy	Girl	Boy	Girl	Boy	Girl
1	20.3	19.1	7.2	7.0	2.2	2.2	10.8	9.9
2	22.4	21.2	8.1	7.9	2.4	2.5	11.9	10.7
3	25.5	24.5	8.8	8.9	3.1	3.1	13.6	12.6
4	27.3	25.5	9.7	9.2	3.2	3.3	14.4	13.1
5	28.6	27.4	10.0	9.9	3.6	3.5	16.1	14.1

TABLE 8: Mean Test Scores broken down by School Year and Sex

Analysis of Variance shows that age, intelligence and sex are significant factors, the variation attributable to each being:

AH2 Grade: 44%, School Year: 14%, Sex: ½%.

The dominance of intelligence (AH2 Grade) is clear, its influence being three times that of age (School Year). With MA substituted for AH2 Grade a similar result was obtained (40% variance), indicating that MA is almost as good a guide to Test performance as is AH2 Grade. This is not surprising in view of the high correlation (+.72) between the two factors.

PROBABILITY CONCEPTS

In order to attempt to allocate each pupil unequivocally to a concept 'level', Guttman Scalogram Analysis was undertaken. The aim was to group together selected Concepts Test questions into subtests representing different levels, and to allocate a passmark for each level. To be at a particular level a pupil must pass each lower level and also pass the given level. The main outcome was a scale comprising three levels (plus zero level indicating total 'failure'). Unfortunately, but not surprisingly, many of the more unusual and interesting items in the Concepts Test could not be fitted into these scales! The distribution of Concept Levels is shown in Table 9.

		Probability Concept Level				
Year	Sex	0	1	2	3	Mean Level
1	Boy	15	51	31	3	1.21
	Girl	18	60	21	2	1.07
2	Boy	12	44	38	6	1.37
	Girl	11	55	27	6	1.28
3	Boy	5	31	48	16	1.75
	Girl	6	43	40	12	1.57
4	Boy	6	25	44	25	1.87
	Girl	6	34	40	20	1.74
5	Boy	2	16	50	32	2.11
	Girl	3	27	44	26	1.93

TABLE 9: Percentages of Main Sample Pupils in each Concept Level by School Year and Sex

A steady progression of Concept Level with increasing age is apparent. Boys show superiority to girls consistently, but the gap between them, as judged by the mean Concept Level, does not widen between the ages of 11-16 years. Further analysis by intelligence level throws light on the nature of the differences found (Table 10). The general trend is that whereas the more able pupils gain about 1 level during the five-year period, the least able, starting 1 level behind, fall even further behind during the same period, only gaining about ½ a level.

	AH2 Grade				
Year	A	B	C	D	E
1	1.84	1.52	1.08	0.80	0.62
2	2.15	1.77	1.30	0.86	0.67
3	2.54	2.12	1.65	1.29	0.67
4	2.67	2.27	1.89	1.25	0.77
5	2.83	2.44	2.10	1.50	1.10
All	2.36	1.98	1.55	1.10	0.74

TABLE 10: Mean Concept Level by Year and by General Reasoning Ability

CROSS-REFERENCE TO PIAGETIAN STAGES

By making a detailed study of pupils' responses to the 'comparison of odds' items (Question 6), and relating this to Piaget and Inhelder's own work (Ref. 9) it was possible to indicate the ages at which Piagetian Stages are normally achieved (Table 11). It must be concluded that most English school children do not achieve the level of formal operations, which is Piaget's Stage III or the Probability Concept Level 3, a result in line with findings on Science concepts reported by Shayer, Kuchemann and Wylam (Ref. 10).

Piagetian Stage	Corresponding Question 6 Item	AH2 ability band		
		Top 30%	Middle 40%	Bottom 30%
IB	(a)	<11	<11	11
IIA	(b)	11	13	16
IIB	(c)	12	14	-
III	(d) (e)	14	-	-

TABLE 11: Ages at which Piagetian Stages are normally achieved for comparison of odds problems, for three intelligence levels

SUMMARY OF FINDINGS RELATING TO THE PROBABILITY CONCEPTS TEST (N=2930)

1) Nearly all test items show an increase in facility with age and intellectual ability.

2) Often years 1 and 2 display quite similar response patterns, with substantial increases in performance for Years 3, 4 and 5.

3) Items which merely require the counting concept are well done by all Years, but those requiring the ratio concept are very poorly done, particularly below Year 4.

4) Various strategies are employed by pupils in answering comparison of odds problems. Counting and differencing are two popular strategies, but are rarely applied consistently. Computing ratios is markedly a function of intelligence and also of age.

5) Items involving an appreciation of randomness, the stability of frequencies and inference were particularly poorly done, with little improvement with age evident.

6) Although most pupils coped well with a combinatoric item at a concrete level (permutation of 3 objects: 75%) few could proceed to formal level items (permutations of 4 objects: 12%, 5 objects: 4%).

7) Pupils' verbal ability is often inadequate for accurately describing probabilistic situations.

8) Very often certainty and high probability are equated, as are impossibility and low probability. Pupils freely associated a 50% probability with (i) outcomes which might or might not happen, (ii) equi-probable outcomes (when more than 2 outcomes exist).

9) Generally, boys outscore girls on total Concepts Test Score, but not on combinatoric questions.

10) Generally, boys show superior performance to girls on many Concepts Test questions, with girls significantly superior on one question only (Question 20).

11) General reasoning ability (AH2 Grade) is the dominant factor in Test Score (44%), being three times as important as age (School Year) (14%). Sex is of minor significance ($\frac{1}{2}$%).

12) Concepts Level means rise consistently with increasing age (School Year) and intellectual level (AH2 Grade).

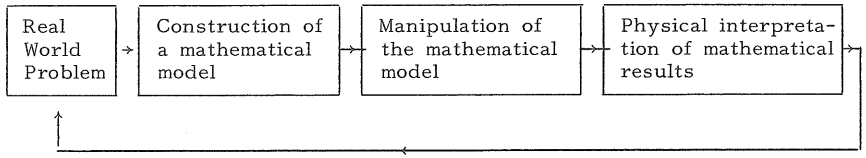
13) Generally, boys have a higher Concepts Level than girls.

14) Most pupils do not attain the level of 'formal operations' (Piagetian Stage III) by Year 5, and presumably leave school at 'concrete operations' level (Stage II).

15) Teaching style, along the 'teacher centred ... individualised learning' continuum does not appear to be a significant factor affecting Test Score.

SOME RECOMMENDATIONS

a) The subject of probability is a very appropriate topic whereby to emphasise the nature of mathematical modelling:



Question 1 is asking the pupil what he thinks is an appropriate model, which dictates his choice of response. (Is colour a relevant factor? Is past history?).

School mathematics very much emphasises mathematical manipulation and gives very little weight to either model building or interpretation. Probability teaching must attempt to redress the balance. (Does Question 1 only have one defensible answer?).

b) Individual interviews have confirmed that some pupils are quick to answer questions intuitively, with no obvious intervention of mathematical processes. However, it is unwise for the teacher to assume that all pupils can see what to the teacher appears obvious, even if some in the class effortlessly provide the 'correct' answers. Only an extensive programme of class-based activities is likely to provide the weight of evidence and experience to eliminate the fallacious thinking exhibited by pupils and adults alike.

c) The provision of diagrams seems to help pupils' understanding, and should be provided wherever possible, or pupils encouraged to interpret verbal problems pictorially before proceeding to solve them.

d) Various strategies are employed by pupils in solving comparison of odds questions but little consistency can be found. Much benefit can accrue from explicit classroom discussion of appropriate strategies for particular problems with a view to developing adequate discernment of types of problem or of mastery of a universally applicable strategy (the ratio concept).

e) The teacher needs to give attention to educating pupils in the normally accepted usage of various terms which indicate varying degrees of likelihood. Commonality of meaning is obviously most desirable, and should not tacitly be assumed. Discussion of events incorporating a chance element needs to be encouraged.

f) In our scientifically and technologically dominated society, with its deterministically oriented education system, it is perhaps inevitable that the concept of randomness should be neglected. Very little time or attention is given to this concept despite the fact that we are surrounded by random occurrences every day of our lives. It is unreasonable to treat so lightly such a pervasive property of our world.

g) The multiplicative principle is poorly understood, and pupils have little in the way of basic intuition on which it can be based. Response to Question 22 was very poor. Clearly little conceptual understanding has been achieved, and using a tree diagram was just a mechanical device.

h) Drawing inference from the outcomes of sampling experiments and appreciation of the stability of frequencies are two most important conceptual areas. Very few pupils were able to deal with either. Our results, and follow-up interviews, suggest that much practical activity and discussion is needed to develop an adequate appreciation of these concepts in the majority of pupils.

References

[1] Hart, K.M. (1980) Secondary School Children's Understanding of Mathematics (CSMS). Chelsea College, University of London.

[2] Hart, K.M. (1981) Children's Understanding of Mathematics: 11-16, John Murray.

[3] Green, D.R. (1982) Probability Concepts in 11-16 Year Old Pupils. Centre for Advancement of Mathematical Education in Technology, Loughborough University of Technology. 40 pages, £1.00.

[4] Heim, A.W., Watts, K.P. and Simmonds, V. (1974) AH2 and AH3: Parallel Tests of Reasoning, Brit.J.Psychol., 65, 1-11.

[5] Heim, A.W., Watts, K.P. and Simmonds, V. (1978) AH2/AH3 Manual, (2nd Edition) NFER.

[6] Green, D.R. (1979) The Chance and Probability Concepts Project, Teaching Statistics, 1, 3, 66-71.

[7] Halmos, M. (1975) First steps in teaching probability: in Statistics at the School Level, edited by L. Rade, Almqvist & Wiksell.

[8] Fischbein, E. (1975) The Intuitive Sources of Probabilistic Thinking in Children, D. Reidel.

[9] Piaget, J. and Inhelder, B. (1975) The Origin of the Idea of Chance in Children, Routledge and Kegan Paul.

[10] Shayer, M., Kuchemann, D.E. and Wylam, H. (1976) The Distribution of Piagetian Stages of Thinking in British Middle and Secondary School Children, Brit.J.Educ.Psychol., 46, 164-173.

MISCONCEPTIONS OF PROBABILITY, SYSTEMATIC AND OTHERWISE; TEACHING PROBABILITY AND STATISTICS SO AS TO OVERCOME SOME MISCONCEPTIONS

J M Shaughnessy, Oregon State University

Suppose you are teaching an introductory course on probability and statistics to first year college students or to fourth year secondary students. After several weeks, when you have covered some introductory material, you present these problems to your students. What answers will your students give for the questions?

1. The probability of having a baby boy is about 1/2. Which of these sequences is more likely to occur for having six children?

   i) a) BGGBGB b) BBBBGB c) about the same chance for each.

   ii) a) BGGBGB b) BBBGGG c) about the same chance for each.

2. Consider the grids below.

Grid A	Grid B
X X X X X X X	X X
X X X X X X X	X X
X X X X X X X	X X
	X X
	X X
	X X
	X X
	X X
	X X

   Are there

   a) more possible paths* in grid A?
   b) more possible paths in grid B?
   c) about the same number of paths in each grid?

(* A path is a polygonal chain of line segments, starting at top row and meeting one and only one symbol in each row. Examples have been shown to your students).

3. The chance that a baby is born a boy is about 1/2. Over an entire year, would there be more days when at least 60% of the babies born were boys in

   a) a large hospital
   b) a small hospital
   c) makes no difference

4. You must select committees from a group of 10 people. Would there be:

   a) more distinct possible committees of 8
   b) more distinct possible committees of 2
   c) about the same number of committees of 8 as committees of 2?

5. There are two cab companies, Green Cab and Blue Cab, in a city. Of all the cabs, 85% are Green and 15% are Blue.

   A cab is involved in an accident, and a witness testifies that the cab was Blue. Past data on the accuracy of witness identification indicates that witnesses make correct colour identifications in 80% of the cases, and incorrect in 20% of the cases.

   What is the probability that the cab in the accident was Blue rather than Green?

Of course, none of these questions is the type students in the first few weeks of an introductory course would be expected to do by applying theoretical principles. The point is: how do statistically naive students respond to such questions? Naive students would likely possess certain misconceptions of probability due to their inexperience with the mathematical laws of probability. Thus, it may be possible to clear up some misconceptions by familiarising students with probability concepts and theory. However, in the research that has been reported on these and similar problems, (Kahneman and Tversky 1972, 1973, 1974; Shaughnessy 1977, 1981; Bar-Hillel 1980; Tversky and Kahneman 1971, 1980; Well, Pollatsek and Konold 1982) the evidence suggests that some misconceptions of probability are rather deep seated and may not be the result of simple mathematical naivete. Some misconceptions appear to be of a psychological kind, and more exposure to the laws of probability and statistics may not be sufficient to overcome them, even in the best of courses. Indeed, some of these "non-mathematical" misconceptions appear to be systematic, that is, they can perhaps be characterised by certain psychological principles. On the other hand, some misconceptions may be very unsystematic and possibly the by-product of certain "belief systems" possessed by the students.

In this paper, we will discuss the types of misconceptions that arise from students' responses to the five questions above. Then we will consider the implications of these misconceptions for the teaching of probability and statistics, and suggest some approaches to probability that may be useful for confronting the misconceptions that our students possess.

TYPES OF MISCONCEPTIONS

Errors Based on Intuitive Heuristics

Consider once again Problem 1.

(1) Which sequence is more likely to occur, BGGBGB or BBBBGB? Here are the responses from a random sample of 80 college undergraduates prior to a course in probability. The numbers are frequency of responses.

BGGBGB	BBBBGB	Same
50	2	18

A large majority of students say that the sequence BGGBGB is more likely to occur (Kahneman and Tversky, 1972; Shaughnessy, 1977). When asked to give a reason for their choice they often say "because BGGBGB fits the expected 50:50 ratio of boys to girls". The sequences are actually equally likely to occur, each with probability, 1/64. These students have used the probability of a boy one try, 1/2, as an indication of the likelihood of a whole sequence of trials. Kahneman and Tversky refer to this strategy, or intuitive heuristic, as representativeness. According to respresentativeness, subjects tend to estimate likelihoods based on how similar an event is to the distribution from which it is drawn or upon how similar the event is to the process by which the outcomes are generated. Thus, BGGBGB is chosen by subjects to be more likely than either BBBBGB or BBBGGG. The former is not as "representative" of a 50-50 distribution as BGGBGB, and the latter does not reflect (represent) the random process of having children as well as BGGBGB.

The use of representativeness by subjects is also apparent in Problem 3, in a somewhat different manner.

(3) Are there more days during the year when at least 60% of the babies born are boys in a large hospital or in a small hospital?

Small	Large	No Difference
17	15	48

Many students felt that the distribution of boys in the daily sample had nothing to do with hospital size. However, there is a good chance that at least 60% of the babies are boys on a given day if the hospital has only 3 babies that day! The students used the 50:50 distribution of boys to girls in the population as a representative estimate, regardless of the sample size. Kahneman and Tversky (1972) suggest that the use of representativeness as a heuristic for estimating likelihoods can lead to errors that arise from

1. Insensitivity to prior probabilities and disregard for population proportions.
2. Insensitivity to the effects of sample size on predictive accuracy.

3. Unwarranted confidence in a prediction that is based upon invalid input data.
4. Misconceptions of chance, such as the gambler's fallacy (coin with memory).
5. Misconceptions about the tendency for data to regress to the mean.

Another systematic heuristic that is used to estimate likelihoods is availability. When people make predictions based on how accessible instances of an event are to memory, or on how easy it is to construct particular instances of the event, they are relying upon availability. For example, suppose a person is asked to estimate the percentage of smokers in her community. If she has frequent contact with people who smoke, her estimate will undoubtedly be much higher than an estimate by a person who seldom socialises with smokers. The ease with which positive instances of an event can be called to mind greatly influences estimates for the likelihood of an event.

Consider our introductory problems 2 and 4.

(2) Which grid has more paths?

A.
X X X X X X X X
X X X X X X X X
X X X X X X X X

B.
X X
X X
X X
X X
X X
X X
X X
X X
X X

Here are the results from the sample of 80 college students.

A	B	Same
53	8	15

Although it appears easier to construct instances of paths in grid A - there seems to be more choices available at each line - there are actually 512 distinct paths in each grid since $2^9 = 8^3 = 512$. When asked to give a reason why they chose grid A over grid B, students often said "It's easier to draw a path in grid A". Results of this question strongly suggest that the heuristic of availability has influenced many of the responses.

(4) Are there more distinct committees of 2 or of 8 that can be chosen from a group of 10 people?

Here are the responses.

Committees of 8	Committees of 2	Same
7	47	19

It is easier for people to generate a lot of different committees of 2 quickly than to create committees of 8. Thus, there would seem to be more committees of 2 because they are more "available". Most combinatorially naive people do not see the dual nature of the problem, that a committee of 8 is the same as a non-committee of 2.

It is perhaps not too surprising that combinatorially naive people, such as those in the sample above with little training in probability or statistics, would rely upon some intuitive heuristics for events. However, Kahneman and Tversky (1971) have reported that subjects with a substantial background in probability and statistics were subject to the same intuitive misconceptions about probability as the naive college students. This suggests that mere exposure to probability and statistics in course work may not be enough to overcome such systematic biases as those introduced by heuristics such as availability and representativeness. In fact, the use of such heuristics may often be perfectly justified and even beneficial. If you have seen several automobile collisions at a certain dangerous intersection, the use of availability to estimate the safety of the corner may make you cautious and save an accident. A doctor may be able to make a quick diagnosis because "there is a lot of that going around", again using availability as a guessing heuristic. Of course, there's always the unfortunate instance when this backfires, when the symptoms are the same but the diseases aren't! The challenge to those of us who teach probability and statistics is to sensitise our students to the existence of such intuitive heuristics in such a way that they may determine when they are useful and when they may be misleading.

Errors Arising From Neglect of Base Rate Data

There has been a considerable amount of research on the taxicab problem (number 5) given in the introductory section above (Tversky and Kahneman, 1980; Well et al., 1982). This problem, and similar ones, have been used to determine how subjects who are unfamiliar with probability make estimates for likelihoods when several pieces of information have to be integrated.

Consider the problem again.

(5) 85% of the cabs are Green and 15% are Blue. A cab has an accident. A witness who is accurate 80% of the time identifies the cab as Blue. What is the probability it was Blue rather than Green?

Tversky and Kahneman (1980) report that the median response among several hundred subjects was .8. They suggest that subjects tend to ignore the given base data of 85% Green cabs when making their decision. Furthermore, they argue that people will be more likely to integrate base rate data into their estimate if a "causal" influence is introduced into the base rate, for example by changing the problem to read, "...85% of accidents involving cabs involve Green cabs and 15% involve Blue cabs". The actual probability that the cab was Blue (B) given that it was identified as Blue (b) is

$$P(B/b) = \frac{P(B \wedge b)}{P(B \wedge b)+P(G \wedge b)} = \frac{.12}{.12+.17} = .41$$

Interest in these types of problems arose when Edwards et. al. (1963) suggested that people may make estimates for the likelihood of events according to Bayes' Theorem, only in a more conservative way than the theorem would predict. However, the results of Kahneman and Tversky imply that people do not act as intuitive Bayesians, at least not on the taxicab problem.

Recent research by Well et. al. (1982) suggests that the types of misconceptions that occur among naive subjects when confronted with Bayesian problems, such as the taxicab problem, are much more complicated and varied than simple neglect of base rate data or use of some causal schemas. Much of the research on misconceptions of probability has reported only average responses - means or medians - and has not asked the studetns why they responded in a particular way. Well carried out follow-up clinical interviews with students who had been given Bayesian problems, like the taxicab problem, in order to disclose their mental processes. He found large individual differences in the types of strategies employed and in the kind of misconceptions that were apparent among the students. In particular, he found no evidence to support the causal schemas contention of Tversky and Kahneman. Rather, some students used non-probabilistic strategies to estimate the likelihood of events, strategies which may be based on a deterministic model or even on personal belief systems. Well says, "It is difficult to infer a strategy from a numerical answer. We have observed subjects with markedly different strategies (in the interviews) arrive at the same numerical answer, and subjects with the same strategy arrive at different answers".

Errors Due to a Deterministic Model or Belief System

An estimate for the likelihood on an event may not be based upon a probabilistic model at all. For example, suppose someone decides that the witness is right in the cab problem. They will then conclude that the Blue cab was involved in the accident. There is no chance involved here, and no use of the percentages before the decision. In the interviews Well and his colleagues obtained estimates as high as 90-95% that it was a Blue cab. When asked why, these subjects said, "I think the witness was right", or "I think it was a Blue cab". In such cases, the decision is based on a belief or deterministic decision. The actual percentage given there is more like a confidence rating than a probability estimate. Base-rate information may be neglected if it violates preconceptions or beliefs of the subject.

This type of misconception may pose the most difficult challenge of all to those of us who teach probability and statistics. What do you do if your students just plain don't believe in probability or chance? Well (1982) reports a protocol excerpt in which a student is asked to pick which side he thinks will turn up if an irregularly shaped die (actually a bone) is rolled. During the interchange the student eventually settles on side C as the most likely to come up - for no apparent probabilistic reason. (In fact, there is no evidence of an understanding of chance by the student during the entire episode). The student is shown data from 1,000 rolls of the die which clearly suggest side D is most likely to occur, and then asked which would be the most likely. "You found D most likely. ... I still think C". This student has evidently formed a belief that C is most likely, and experimental data are not going to alter the belief.

A PROBLEM FOR TEACHERS OF PROBABILITY

Most of the literature referenced above comes from research on judgement and decision making by psychologists. However, this research has very serious implications for the teaching of probability and statistics. The student in our introductory courses, either in secondary school or in college, come already armed with a set of preconceptions and misconceptions about probability. The misconceptions may be of a systematic probabilistic nature such as representativeness and availability, or they may be non-probabilistic in nature, based on causal schemas, belief systems, or other deterministic attitudes. The intuitive probabilistic heuristics may be so ingrained that they are difficult to overcome In a course on probability and statistics, even one which is especially designed to confront them (Shaughnessy, 1977). The deterministic misconceptions of the students may pose a more serious problem for the teaching of probability. The teacher may have to break through a belief barrier to even reach the

students in the first place. Fischbein (1976) claims that science education traditionally has emphasized deterministic models and has tended to neglect probabilistic models and the study of uncertainty. Thus our intuition for probabilistic thinking may be distorted by an overemphasis on deterministic models in our schools, such as natural selection, Newton's laws of motion, or a deductive axiomatic approach to geometry. Can we then be too surprised at the deterministic reasoning of some of our students when we finally do get them in a course on probability?

What kinds of things should we do in our probability and statistics courses in order to confront and perhaps overcome some of these misconceptions? I do not claim to have a definitive answer to this question, but I do believe there are some things that will help. These suggestiosn certainly are not brand new, but I feel they are all the more important when one considers the misconceptions of probability indicated by psychological research.

TEACHING PROBABILITY TO OVERCOME MISCONCEPTIONS

Let us consider some recommendations which may help avert or overcome our students' misconceptions or probability and statistics.

1. Introductory probability and statistics should be activity-based and experimental. We can proceed from guesses to experiments to simulations of experiments, and only after considerable intuition has been developed, to a formal mathematical presentation. Confront students with problems in which misconceptions can arise, such as the introductory problems in this paper.
2. Emphasize the confounding role that belief systems can play in judgement and decision making. At the same time, teach the importance and usefulness of statistics for helping to make objective judgements and decisions.
3. Sensitize students to misuses of statistics.
4. Give students challenge problems which may require them to gather or simulate their own data, and then make a decision (i.e., synthesize 1-3 above).
5. Sensitize our students to some apparent statistical paradoxes.
6. Carry out clinical interviews with our own students to uncover their reasoning processes on probability problems.

Let us consider each of these briefly.

1. From guesses, to experiments, to simulation, to theory.

An experiment with a small-group problem solving, activity-based approach to introductory probability and statistics at the introductory college level was conducted by the author (Shaughnessy, 1977).

The results indicate that students have a better chance of overcoming misconceptions of probability that are due to intuitive heuristics such as representativeness or availability if they are introduced to probability in an experimental activity-based mode rather than in a formal lecture mode. Students can work in small groups, make guesses for the likelihood of events, perform an experiment, organise and analyse the data, and make estimates for the likelihood of events based on their data. More advanced students can perhaps attempt to build a theoretical probability model for the experiment and then compare theoretical predictions to experimental outcomes. The reasons for any discrepancies between experimental results and theoretical predictions, such as sources of bias in the data-gathering process, can be discussed.

An experimental approach to probability may, in fact, be the only way to break the "belief barrier" for those students who do not understand chance at all. Problems which confront misconceptions of statistics can be presented and solved by experiment. Below are several activities that have been used by the author during the first several meetings of an introductory college course on probability and statistics. More activities can be found in previous articles, activities that deal with the effects of sample size on mean, standard deviation, and probabilities of events.

Activity 1: Keep Your Head Up

Before doing this activity, write down your best guess for each of the following: If you flip 6 coins, what is the probability that you will get -

a) 6 heads b) 5 heads c) 4 heads d) 3 heads

Perform this activity in groups of six students:

1. Flip 6 coins. Record the number of heads. Repeat about 50 times. Use your data to answer each of these questions.

   a) What is the probability of getting 6 heads? 5 heads? 4 heads? 3 heads? 2 heads? 1 head? 0 heads?

   b) What is the probability of getting at least 1 head? At least 2 heads?

2. Make a list of all possible outcomes for flipping 6 coins.

   a) Develop a mathematical model to find the theoretical probability for the outcomes of flipping 6 coins.

   b) What is the theoretical probability for getting at least 1 head? At least 2 heads?

   c) What are the assumptions of your mathematical model?

3. Compare the experimental probabilities from part 1 with the theoretical probabilities in part 2. How well do they agree? Make a graph to compare the experimental (observed) probabilities of 50 flips for 0 heads, 1, ..., 6 heads with the theoretical probabilities. Plot the two graphs on the same set of coordinate axes. Where is there close agreement between the two graphs? Where is there not close agreement? Why do you suppose this happens?

4. What assumptions have you made in your experiment when flipping the coins? What suggestions do you have to improve the experiment?

5. List any other comments, questions, observations, or reactions that you might have concerning this activity.

Activity 2: A Pointed Affair

Three different colours (red, gold, and silver in this example) of ordinary household thumbtacks are distributed, one tack to each person.

1. Do this part of the activity by yourself.

   a) First, write down your best guess for the probability that your tack will land upright when dropped.

   b) Devise some uniform way of dropping your tack, and drop it 75 times. Arrange your data in an array with U to indicate upright and D to indicate down.

   c) On the basis of the data you have collected, calculate the probability that the tack will land upright; that it will land down.

2. Do this part of the activity in small groups. Get a person in your group with one each of the three colours of tacks. Use the probabilities calculated in part 1 to list the probabilities that (1) the red tack will land up, (2) the silver tack will land up, and (3) the gold tack will land up. Drop the three tacks together and record the results. Before performing this experiment, write down your best guess for each of the following probabilities:

   a) All the tacks land up

   b) No tack lands up (what's another way to say this?)

   c) At least one tack lands up (what's another way to say this?)

   d) The red tack lands up

   e) Two tacks land up and one lands down

   Drop the three tacks in some uniform way about 60 times. Record your data in an array by listing the reasons as triples; for example, UDD could stand for red tack up, gold tack down, and silver tack down.

   Use the data you have gathered to calculate the experimental probabilities for the events in a-e in part 2. Compare these calculations to your guesses. Any surprises? What assumptions have you made in doing the experiment? What suggestions do you have to improve this experiment?

3. Do this part of the activity in small groups.

a) Develop a mathematical model to assign theoretical probabilities to the outcomes of this experiment. First list all possible outcomes for the experiment; then devise a way of assigning a probability to each outcome.

b) Use your data from part 2 to determine experimental probabilities for each of the outcomes listed in 3a. Compare these experimental probabilities with the theoretical probabilities for the outcomes given by your model. Make a graph to compare your experimental probabilities with the theoretical probabilities. How well do the graphs agree?

c) What assumptions have you made in your mathematical model? Does the model for the tack experiment differ in any way from the model for the coin experiment? If so, how? Is there any similarity between the two experiments?

d) List any other comments, questions, osbervations, or reactions that you might have to this activity.

Activities such as these are aimed at correcting misconceptions. The tossing of six coins is equivalent to our representativeness problem (1) with the sequence of six children. The tack activity provides experience with unequally likely outcomes so that students begin to realise that if there are N possible outcomes, this does not mean that each outcome has a 1/N chance of occurring. Furthermore, the activities are in two stages. The first stage deals with experimental probability. The second stage encourages students to build their own theoretical models for assigning probabilities.

2. Judgements and Decisions: The confounding role of belief systems vs. the usefulness of statistics in helping to make objective decisions.

Suppose that you are taking a statistics course which has weekly quizzes with 10 questions each. Your average all year has been 7 correct out of 10. For the next quiz, you study with a friend and get 9 out of 10. Did studying with a friend help?

One possible response to this question is: "Of course it helped, isn't 9 better than 7?" Another possible response is: "Probably not, 9 could have just happened that time since your average of 7 is close to 9". If you ask this question of students in an introductory course in statistics, you will get both the answers above, in addition to a spectrum of "it depends" answers in between. These responses indicate that a belief system is being employed rather than that a calculation has been made. In order to resolve the argument in an objective way, the class needs some data. This is an excellent opportunity to carry out a simulation, say with random numbers, and see just how likely it is to obtain the event 9 out of 10 correct if you have probability of success P(s) = .7 on any one item. (It is not necessary - or advisable in this author's opinion - to do a calculation using the binomial distribution the first time a student sees this type of problem).

Suppose that a consumer's group feels that the gas tanks on a certain model of automobile are likely to explode on impact from the rear. This group has been formed because of the unfortunate death of a friend in just such an accident. The consumer's group has formed a belief about a car model. They may even have found other instances of such accidents. Should the car be recalled?

We must attempt to make an objective decision, if possible. Clearly there is some need for more data, and a class in introductory statistics could perhaps be led to suggest some reasonable experiment that can be carried out to determine the safety of the gas tanks. In a problem such as this, the importance of sampling in making decisions from statistics can also be brought out. (Would we do our experiments on cars all from one plant? Would we smash all the cars?).

If we focus on the usefulness and importance of statistics for making objective decisions, we may be able to temper the effects of belief systems or deterministic approaches to probability in our students.

3. Sensitise students to misuses of statistics.

If we demonstrate the importance of statistics in making objective decisions we are also obliged to show our students how statistics can be used to mislead or to "influence" our belief systems. Books such as How to Lie with Statistics (Huff, 1954) should be required reading in all introductory probability and statistics courses. Such books can sensitise our students to misuses of statistics, deliberate or otherwise, and encourage them to analyze probabilistic statements critically.

Students can be asked to keep a log of misuses of statistics that they discover. Articles can be taken from newspapers or periodicals and analyzed for correct or incorrect use of statistics. Students in my classes have ferreted out misleading graphs, inflated percentages, biased samples, insufficient sample size, and verbose statistical quantitative descriptions that had little or no basis in fact.

Many examples of statistics that were based on a biased sample have been analyzed by my students. One found an article on page 1 of the largest-selling newspaper in a major city that mentioned that 95 per cent of the people who were interviewed were against cross-district busing to achieve racial integration in the schools. A continuation of the article (on page 14) mentioned that 70 per cent of the people interviewed did not have children in the schools. Another student criticised the charts published by bookstores and record stores to indicate the top ten books or records for the week. These figures are often based on sales to a particualr class of people with special interests. Yet the list purports to reflect the population as a whole.

4. Give students challenge problems.

In order to synthesize the suggestiosn in 1-3 above (i.e., experiments, simulations, confront misconceptions, show usefulness of statistics, sentisize to misuses of statistics) I present challenge problems after several weeks in an introductory course. The students have to turn in a certain number of solutions before the end of the course. They may solve the problems any way they like, by experiment, by simulation, perhaps by a theoretical model if we are far enough in the course or if they do some outside reading. The assignment is called Stimulation with Simulation. Below are several examples of problems from past assignments. The problems have been gathered over the years from several sources or have been created by me. The sources include Mosteller (1962), Mosteller et. al. (1973), Shulte and Choate (1977), and Råde (1977).

STIMULATION BY SIMULATION

The questions in Stimulation by Simulation are somewhat open-ended in nature. In many cases it is necessary to first define clearly what the problem is, and then to adopt a mode of simulation which will enable you to "model" an experiment which may be too cumbersome to actually carry out.

Before doing any of these problems, always write down your best guess first, and only then carry out a simulation of the problem.

A simulation of a probability experiment invovles:

a) modelling an experiment by using some apparatus with known probabilities (i.e. coins, dice, spinners, random numbers).

b) "performing" the experiment many times with your apparatus, thus pooling results of several small groups may help to get a "large enough" sample.

c) gathering, organising, and analysing the data (some statistical skills).

d) calculating experimental probabilities, or other experimental outcomes (i.e. frequencies) from the data.

e) making inferences or drawing conclusions from the experimental outcomes, that is, looking back.

Problems

1. A cereal company has put plastic models of 5 Star Wars characters in boxes of SPACY-O'S. The company puts one character in each box. How many boxes would you expect to have to buy in order to get all five characters? (Write down a guess first).

2. Five friends are sitting around a table at a party discussing the merits of astrology. What is the probability that at least two of them have the same astrological sign? (Assume the signs are equal in duration).

3. Some data were obtained from the records of the Bureau of Meteorology in Perth, Australia. A sunny day was defined as a day on which the officially recorded number of hours of sunshine was 70% or more the maximum possible hours of sunshine for that day of the year. The months April to September in 1973 seemed to be typical months in terms of average number of hours of sunshine. A classification of each day in this period as "sunny" or "dull" (S or D) and recording of the events led to the following:

D followed by D:	79 occasions
D followed by S:	33 occasions
S followed by D:	33 occasions
S followed by S:	37 occasions

Thus the proportion of dull days followed by dull days was 79/112 = .70, and the proportion of sunny days followed by sunny days was 37/70 = .53.

a) What is the "average" length of runs of consecutive dull days?

b) What is the "average" length of runs of consecutive sunny days?

4. Two points (numbers) between 0 and 1 are chosen at random. The points divide the interval from 0 to 1 into three line segmetns. What is the probability that the three segments will form a triangle?

5. About one-half of the babies born are boys. Would there be more days during the calendar year when at least 60% of the babies born were boys:

i) In a small hospital

ii) In a large hospital

iii) Makes no difference

6. Suppose you are playing a game in which you have a 50% chance of winning. Bet a dollar. Each time you lose double our bet. Each time you win, go back to betting a dollar.

Is this a good betting system?

7. Fast Al suggests to Slow Sam that they play the following game. They will each throw either one or two fingers. If Sam throws one and Al throws two, Sam wins $30. If Sam throws two and Al throws one, Sam wins $10. If they throw the same number of fingers, Al wins $20.

a) If each player plays a 50-50 strategy between their two choices, who wins in the long run?

b) What is the best strategy for Slow Sam, i.e. how should he divide his choices?

8. From where he stands, one step toward a cliff would send the drunken man over the cliff. He takes random steps, either towards the cliff or away from it. At any step, the probability that he takes a step toward the cliff is 1/3, and the probability that he steps away from it is 2/3. What is his chance of escaping a fall?

9. 16 teams are entered in a double elimination softball tournament. What is the expected number of games a team will play:

   i) if the team has a 50% chance of winning each game?

   ii) if the team has a 70% chance of winning each game?

10. Several years ago a certain beer company claimed that they "tried their famous High Life taste test on many a beer expert, and no one had yet correctly identified which beer came from a can, which from a bottle, and which from a tap".

    How likely is it that the beer company was telling the truth? Or, perhaps another way to put it, how many beer experts do you suppose the company really tested?

11. Suppose a man faces north and tosses a coin to decide whether to take one step north or one step south. The man does this for twenty-five tries.

    i) On the average, how far is he from the starting point?

    ii) On how many steps is he on the north side of his starting point? On the south side?

    iii) How often does he return to starting point?

12. Three towers are built of 8, 6, and 4 blocks. A tower is chosen at random and a block is removed.

    a) What is the expected number of blocks that will be removed before one of the towers has been "levelled". (i.e., no blocks left?).

    b) What is the probability the 4-block tower will be destroyed?

    c) Can you think of a "real world" situation for which this tower game is a model?

13. (The big-bad-triangle, or three's a crowd). Three duellers, The Good, The Bad, and The Ugly, are to fight a three cornered pistol duel. All know that the Ugly's chance of hitting his target is .30, the Bad's is .50, and the Good (of course!) never misses. They are to fire at their choice of target in succession in the order Ugly, Good, Bad, cyclically until only one man is left unhit. A hit man loses further turns, and is no longer shot at. Ugly has the first shot. What should he do to have the best chance of survival? (Running away doesn't count!).

5. Expose Students to Statistical Paradoxes.

We have already mentioned the importance of sensitizing students to the misuses of statistics and emphasizing the role of statistics in making judgements and decisions. There are times when the way we choose to analyze data, or the underlying assumptions we make in our statistical model, will influence judgements or decisions. It is very important that we expose our students to instances of so-called "paradoxes" so that they remain highly critical of interpretation of data when reading results. If there is more than one possible interpretation

for the data, we should encourage our students to examine both interpretations. Several recent articles present instances of so-called paradoxes that are accessible to secondary school students in an introductory probability and statistics class (Reinhardt 1981; Wagner 1981). How, for example, is it possible for baseball player A to have a better batting average than player B both against right-handed pitching and against left-handed pitching, and yet player B gets to start games because she has a better overall average than player A? This example, taken from Reinhardt's article, points out the importance of a weighted average, and the resulting impossible dilemma. Of course when it comes to a decision as to which player should actually start against, say, a left hander, the decision depends on which model the manager chooses to use for the decision: overall average or average against lefties. This is the sort of paradox that can challenge the class to come up with a fairer decision making model.

Another paradox presented in both the Reinhardt and Wagner articles involves the conviction of a couple in a court case based upon six characteristics the witnesses testified were present in the couple at the scene of the crime. The prosecution computed the odds as 1 in 12 million that these characteristics would simultaneously occur, and obtained a conviction based primarily on this calculation. The verdict was subsequently overturned when a mathematician explained that the six characteristics mentioned by the witnesses were far from being clear independent events, and that in fact one could probably find quite a few couples near the scene of the crime with the same six characteristics (Time, 1968). The role of assumptions in our interpretative models is crucial, and we should impress this upon our students.

6. Collect Interview Data from Our Students

As Well et. al. have suggested in their analysis of interviews of college students answering the taxicab problem, there appears to be a great diversity of conceptions and misconceptions of probability and statistics among our naive students. Indeed, the way we tend to think about the subject is not at all the way they (our students) tend to think about the subject. It may be beneficial to both students and teachers if we explored our students' thinking processes in more detail, and reflected on their responses. A numerical estimate for the likelihood of an event, or even a number resulting from some sort of calculation, by itself gives very little indication of how or why that number was chosen. In a recent address, Labinowitz (1982) has suggested that teachers be trained in the art of administering clinical interviews. He has developed a course for preservice and inservice teachers which focuses on the use of audio taped interviews by teachers. The interviews are used to help teachers become aware of student's thinking processes, an exciting approach in an environment which is often all too product oriented. Doing some taped interviews may

help narrow the gap between the teacher's projection level and the students' reception level in statistics courses.

As we have seen, some recent research by psychologists and mathematics educators indicates that misconceptions of probability are clearly present in our students. Some misconceptions are intuitive, some are deterministic, and some may even be based upon belief systems. We have offered some possible approaches to the teaching of probability and statistics which may help confront, and hopefully overcome, some of these misconceptions.

References

[1] Bar-Hillel, M. (1980) The Base-Rate Fallacy in Probability Judgements. Acta Psychologica., 44, 211-233.

[2] Edwards, W., Lindman, H. and Savage, L.J. (1963) Bayesian Statistical Analysis for Psychological Research. Psychological Review. 70, 193-242.

[3] Fischbein, E. (1976) Probabilistic Thinking in Children and Adolescents. In: R. Bechauf (ed.) Forschung Zum Prozess des Mathematiklerners. Institut für Didaktik der Mathematik der Universität Bielefeld, 23-42.

[4] Huff, D. (1954) How to Lie with Statistics. London: W.W. Norton.

[5] Kahneman, D. and Tversky, A. (1972) Subjective Probability: A Judgement of Representativeness. Cognitive Psychology, 3, 430-454.

[6] Kahneman, D. and Tversky, A. (1973) Availability: A Heuristic for Judging Frequency and Probability. Cognitive Psychology, 5, 207-232.

[7] Kahneman, D. and Tversky, A. (1974) Judgement Under Uncertainty: Heuristic and Biases. Science, 185, 1124-1131.

[8] Labinowitz, E. Communicating the Results of Clinical Interviews to Teachers Paper presented at the research presessions of the 1982 National Council of Teachers of Mathematics meeting in Toronto. Based upon a writing project at California State University at Northridge.

[9] Mosteller, F. (1962) Fifty Challenging Problems in Probability with Solutions. Reading: Addison-Wesley.

[10] Mosteller, F., Kruskal, W.H., Link, R.F., Peiters, R.S. and Rising, G.R. (1973) Statistics by Example. Reading: Addison-Wesley.

[11] Råde, L. (1977) Take a Chance with Your Calculator: Probability Problems for Programmable Calculators. Forest Grove, Oregon: Dilithium Press.

[12] Reinhardt, H.E. (1981) Some Statistical Paradoxes. In: Teaching Statistics and Probability, the 1981 National Council of Teachers of Mathematics Yearbook, ed. by A. Shulte, 100-109.

[13] Shaughnessy, J.M. (1977) Misconceptions of Probability: an Experiment with a Small-Group, Activity-Based Model Building Approach to Introductory Probability. Educational Studies in Mathematics, 8, 295-316.

[14] Shaughnessy, J.M. (1981) Misconceptions of Probability: From Systematic Errors to Systematic Experiments and Decisions. In: Teaching Statistics and Probability, the 1981 National Council of Teachers of Mathematics Yearbook, Ed. by A. Shulte, 90-100.

[15] Shulte, A.P. and Choate, S. (1977) What are My Chances? Books A and B. Palo Alto, Calif: Creative Publications.

[16] "Trial by Mathematics". (1968) Time, 26, April, p.41.

[17] Tversky, A. and Kahneman, D. (1971) Belief in the Law of Small Numbers. Psychological Bulletin, 76, 105-110.

[18] Tversky, A. and Kahneman, D. (1980) Causal Schemata in Judgement under Uncertainty. In M. Fischbein (Ed.) Progress in Social Psychology. Hillsdale, N.J.: Erlbaum.

[19] Wagner, C.H. (1981) Paradoxes in Sampling. In: Teaching Statistics and Probability, the 1981 National Council of Teachers of Mathematics Yearbook, Ed. by A. Shulte, 162-167.

[20] Well, A.D., Pollatsek, A. and Konold, C. (1982) Probability Estimation and the Use and Neglect of Base-Rate Information. Draft copy of a paper presented at the research presession of the 1982 National Council of Teachers of Mathematics meeting in Toronto. Under revision at the University of Massachusetts, Amherst.

THE ROLE OF EXPLANATION IN TEACHING ELEMENTARY PROBABILITY

J L Myers, University of Massachusetts

Parents, educators, and mathematicians have been involved in recurrent debates about the relative merits of understanding mathematical concepts and rote memorizing computational formulas. In principle, such conflicts should be resolvable through research involving manipulation of the degree to which students understand mathematical material and observation of their performance on tests of mathematical knowledge. In actuality, we encounter difficulties because we lack both a precise definition of mathematical understanding and a specification of the processes such understanding might affect.

We have begun a research program to investigate the role of understanding in learning elementary probability by university undergraduates. Although we still lack a precise definition of understanding, we have identified three dimensions which have provided the basis for constructing short texts. The first such dimension, external connectedness, corresponds to the degree to which a text explicitly relates the concepts and formulas to be learned to the real-world knowledge the student presumably brings to the situation. The second dimension might be termed "internal connectedness", the degree to which relations among formulas and concepts are explicitly developed. For example, an equation to be learned may, or may not, be presented as a special case of a formula presented earlier in the text. The third dimension is explanation, the degree to which a rationale is provided for a particular formula. This refers not to derivations, but to intuitive justifications for procedures taught. For example one of our texts uses branch diagrams to explain why the probability of the joint occurrence of two independent events is a product of the marginal probabilities.

We will describe the texts constructed within this framework more fully in a later section. First, however, we wish to consider certain processes which should be involved in solving probability problems, and the ways in which these may be affected by the degree of understanding imparted by the text. We assume three stages: categorization of the problem, retrieval of the formula appropriate to that category, and translation of the problem - substitution of values from the statement of the problem into the retrieved formula. In general, understanding should aid both the categorization and translation of problems. Understanding may aid or hinder retrieval of formulas from memory. On the one hand, the linkage of ideas and formulas make the formulas easier to learn and recall; on the other hand, the additional explanatory and integrative material could detract from time and effort that should be spent learning the equations.

From our perspective, it makes little sense to examine total test scores as a function of the degree of understanding incorporated into the text used for instruction. Which text is better may well depend upon which processes are most important in the solution to a particular problem. For example, assume that texts which emphasize rote learning of formulas at the cost of de-emphasizing explanation do provide an advantage in memorizing formulas. Then, problems in which categorization and translation are not difficult may be more easily solved by students who have less understanding of the material. On the other hand, assume that a problem provides no key words which immediately signal the problem's category, or contains considerable irrelevant information which make categorization or translation difficult; a text which has emphasized understanding may provide a definite advantage here.

In view of these considerations, we have attempted a more molecular analysis of test performance than is usually the case. For each of several text conditions, we have examined variations in performance as a function of the type of problem. At a rather gross level, we have formula problems, which merely presented certain values (e.g. "P(A) = .6, P(B) = .4") and necessary conditions ("A and B are independent events"), and required the subject to set up the correct results (e.g. "What is P(A and B)?" requires .6 x .4 as the answer); and story problems which were considerably less transparent. Within the set of story problems, performance has been related to such variables as the presence or absence of key words (e.g. "and" implies the multiplication of probabilities), the presence of irrelevant values, and translation requirements (e.g. whether it was necessary to subtract a stated probability from one in order to solve).

Our perspective has also led us to require subjects to report their answers as operations on numbers; for example, a correct answer would be .4 x .2, not .08. This approach also provided more detailed knowledge of the types of errors made.

To summarize, we view understanding of mathematical material as a function of (1) connections of text concepts and formulas to real-world referrents; (2) integration of concepts and formulas within the text; and (3) explanation of formulas. This view has provided a basis for constructing three written treatments of elementary probability, presumably varying in the degree to which they convey understanding. Our view of the processes involved in solving problems has led us to use both formula and story problems, and to emphasize analyses of error protocols.

METHOD

Subjects

Forty-eight volunteers from an Introductory Psychology course at the University of Massachusetts were randomly assigned to three text conditions. None had previous exposure to probability or statistics.

Texts

There were three texts, each judged to be representative of actual chapters on probability in various introductory statistics textbooks. All contained the six formulas shown in Table 1, with one numerical example illustrating the application of each formula. All three texts began with a brief introduction indicating that the text contained a series of examples of probability calculations,

TABLE 1

Formulas Presented in the Three Texts

(1) P(A or B or C) = P(A) + P(B) + P(C)
(A, B and C are mutually exclusive)

(2) P(A or B) = P(A) + P(B) - P(A and B)
(general case)

(3) P(A and B and C) = P(A) x P(B) x P(C)
(A, B and C are independent)

(4) P(A, B, B, B, A) = P(A) x P(B) x P(B) x P(B) x P(A)

(5) P(A given B) = P(A and B)/P(B)

(6) P(A and B) = P(A given B) x P(B)

and that each example would be followed by a formula which could be used to solve problems of that type. All three texts then contained a section labelled "What are the chances?" which introduced the concept of probability and its limiting values. This section noted that an action such as tossing a coin or observing the day's weather could have several possible outcomes, that a probability value could be assigned to these outcomes, and that these values ranged from zero to one.

Despite these similarities, the texts differed in several ways. A high explanatory text defined probability as a relative frequency of events, and examples emphasized the relation between answers and counts of the numbers of ways that various events could occur. A standard text treated probability as a measure with certain well-defined properties (e.g. limits of one and zero)

and used familiar real-world examples (e.g. the probability of rain tomorrow), but did not explain probability in terms of counting processes. A low-explanatory text was more abstract than either of the first two; aside from an initial reference to dice-throwing, examples were stated in terms of such quantities as P(A) and P(B).

The texts also varied with respect to the degree of linkage among the equations presented. Only the high-explanatory text showed that one equation was a special case of another, or that one equation could be derived from another. For example, this text pointed out that the probability of the union of two mutually exclusive events (Equation 1), was a special case of the probability of the union of any two events (Equation 2).

Finally, the texts varied in the degree of explanation provided for each equation. In the high-explanatory text, Venn and branch diagrams were used to represent the information in the examples, and the equations applied in these examples were explained in terms of these diagrams. Neither of the other texts provided such explanations. For example, although all three texts stated that the probability of the joint occurrence of two independent events was the product of their probabilities and provided an example of this, only the high-explanatory text attempted to demonstrate why probabilities were multiplied in such cases.

In summary, all texts contained an example followed by the equation, as well as sufficient definitions of terms such as mutual exclusivity and independence. In the high-explanatory text, the relation of probability to counting was emphasized, pictorial aids were used, and the equation was presented as a natural development ensuing from the example.

Performance Tests

Two tests were constructed, each containign two sets of six story problems and two sets of six formula problems. Within each set, there was one problem testing each of the six equations in the text. The order of sets within a test was counterbalanced over subjects so that there were four sequences, with the constraint that story and formula sets would alternate. Half of the subjects received one test first and the remaining subjects received the other test first.

Formula and story problems differed in the degree to which the given information was explicit in the question. For example, a formula problem testing knowledge of Equation 1 would state "P(A) = .4 and P(B) = .3. If A and B are incompatible, what is the value of P(A or B)?". In contrast, a story problem testing the same knowledge would read as follows: "The probability is .20 that Jimmy will win a race and .30 that he will finish second. What is the probability that he will win or finish second?".

Procedure

Subjects were tested individually. They were instructed to read the text booklets silently and at their own rate, trying to understand the explanations and examples. The standard and low-explanatory groups, whose texts were four pages in length, were told that they would have 15 minutes to finish; the high-explanatory group, whose text was 14 pages in length, was given 25 minutes. These limits, based on pilot subjects, were imposed to discourage any individual from either quickly scanning the texts, or taking an inordinately long time. All subjects finished reading their texts in the given time period for their group. Subjects in the low-explanatory and standard groups tended to rate the study time more favourably than subjects in the high-explanatory text condition, but this difference was not significant.

At the completion of the study period, subjects were given two minutes to study a list of 15 words unrelated to probability and were then required to write down as many words as they could recall on a blank sheet of paper.

Upon completion of the interpolated task, subjects were given the test. There was one problem on a page with space to do any work which the subject felt might help in the solution. Subjects were given as much time as they required to complete the test; most required about 30 minutes. Upon completing the problems, subjects were asked to return 48 hours later. On the second test day, subjects were given the second form of the test. The order of presentation of the test forms was counterbalanced.

RESULTS AND DISCUSSION

Performance test data were scored according to both a strict and a lenient criterion. The strict criterion required that the subject set up the answer correctly in all respects; the lenient criterion gave credit if the subject used the correct equation but, for example, inserted the wrong values. This distinction applies only to performances on story problems; strict and lenient scores were the same for formula problems.

A Summary of Test Data

Figure 1 summarizes the data. The points on the lines are for the strict criterion; the disconnected open symbols are for the lenient criterion. There are two important observations to be made about the means. First, considering the strict-criterion means, the relative performances of the three text conditions strongly depend upon the type of problem ($p < .001$). The two non-explanatory groups appear to know the formulas better than the high-explanatory group but they have more difficulty in applying that knowledge to story problems.

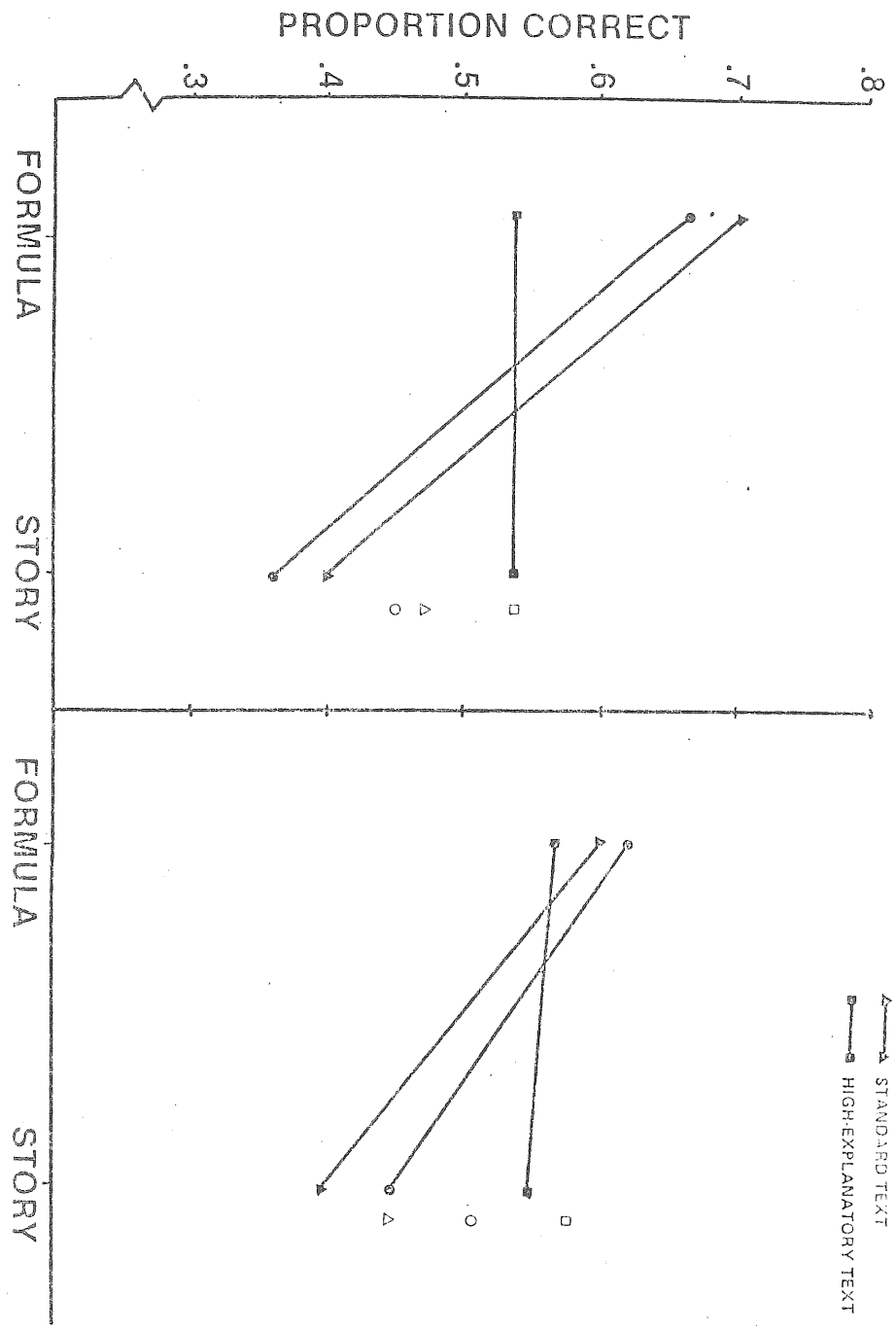


FIGURE 1

The second point to note in Figure 1 is that strict-criterion scores are significantly lower than lenient-criterion scores for the two non-explanatory groups ($p < .001$). Apparently, subjects in these two groups may correctly categorize the problem and recall the appropriate formula but are not able to translate the values in the problem statement. In contrast, when high-explanatory subjects know what operations are to be performed to solve a problem, they usually perform them on the correct set of values; there is little difference btweeen their strict and lenient-criterion scores.

A More Detailed Analysis of Individual Problems

It is impractical to examine the numbers and types of errors made to each of the 24 story problems for each of the three text conditions. Therefore, we will focus on those problems testing knowledge of Equations 1, 3, and 4: the problems within these sets reveal the widest variation in proportions of correct respones. In considering these problems, we ask: why is there marked variation in proportion correct to problems requiring the same equation for solution? And why is this pattern of variation affected by text condition?

The low-explanatory and standard groups performed in similar fashion on all problems. Therefore, in the interests of simplifying the presetnation, we have treated them as a single condition which we have labelled "non-explanatory." We will refer to the high-explanatory group as "explanatory".

Equation 1. Performances on these four problems indicate that subjects in the non-explanatory condition are sensitive to aspects of the problem statement which have little effect upon the performance of subjects in the explanatory condition. Proportions of correct responses in the non-explanatory group varied greatly across the four problems testing knowledge of Equation 1: (a) .375 (b) .4375 (c) .625 and (d) .7188. In contrast, the performances of those in the explanatory group are not only much better but also quite stable: (a) .9375 (b) .75 (c) .8125 and (d) .8125.

Examining the answers to (a) and (b) more closely, we find clear evidence of failures to categorize or translate these problems correctly, but only in the non-explanatory condition. Consider problem (b):

> "In the National Women's Ping Pong League, Los Angeles has a .3 chance of winning, San Francisco has a .2 chance, San Diego has a .1 chance, New York has a .1 chance, and Philadelphia has a .3 chance. What is the probability that a west coast team will win?"

Note that this problem contains irrelevant information; the stated probabilities for New York and Philadelphia are unnecessary. Subjects in the non-explanatory

condition attempted to use these probabilities; five subjects subtracted .1 + .3 from the corect sum (or from the product of the west coast probabilities), four others multiplied all five probabilities, and two others divided by .1 + .3. Only one subject in the explanatory condition attempted to incorporate the irrelevant values into his answer. Apparently, subjects who do not understand the rationale for using the equations attempt to use any information presented to them. In this example, the subjects cited misclassified the problem, using Equations 2, 4 or 5 in an attempt to incorporate all the values stated.

Problem (b) also indicates that subjects in the non-explanatory condition have difficulty when there is no key word to provide a basis for classification. Eight of the 32 subjects in this condition used Equation 3 rather than Equation 1, multiplying probabilities instead of adding them. It is revealing that the only other of the four problems in which this type of error was frequent (nine of 12 subjects who made errors) was (c); (b) and (c) are the only two problems in this set in which the problem does not specifically request the probability of something or something. It is also important to note that this misclassification error occurred only once for problem (b) and once for problem (c) in the explanatory condition.

Even when subjects in the non-explanatory condition used the correct equation, they may have had difficulty translating the story; that is, correctly inserting values from the story into the equation. This was a major source of difficulty for the non-explanatory group in problem (a). The problem presents frequencies of different coloured marbles in a jar, and then asks for the probability of drawing a red or white marble. Five subjects misclassified the problem, attempting to incorporate frequencies of orange and blue marbles into their answer. Most of the errors, however, reveal a lack of understanding of the relation between frequency and probability. Subjects often failed to divide the sum of the numbers of red and white marbles by the total number of marbles, or divided by 100. Note that this problem was the most difficult for the non-explanatory group but the easiest for the explanatory group who had the relation between frequency and probability explained in their text.

The performance of the non-explanatory group on problem (d) provides further evidence of the validity of our analysis of classification and translation errors. This problem evoked the best performance from the non-explanatory group. Not surprisingly, in view of our analysis, it states exactly two values, includes the key word "or", and requires no translation of the stated values.

Equation 3. The explanatory group did not perform as well on these problems as on those testing knowledge of Equation 1. We suspect that the rationale for multiplication of probabilities is more difficult to understand than that

for adding probabilities.

In one respect, performances on this problem set were similar to those on the first problem set: proportion correct was fairly stable over problems for the explanatory group and quite variable for the non-explanatory group. For the explanatory group, these proportions were (a) .5 (b) .5625 (c) .625 and (d) .5625; for the non-explanatory group, the proportions were (a) .2813 (b) .5 (c) .625 and (d) .8438. As with the first problem set, it appears that subjects in the non-explanatory condition are more sensitive to the statement of the problem than are subjects in the explanatory condition.

Much of the variability exhibited by the non-explanatory group can again be traced to variations in the problems with respect to irrelevant information and translation requirements. Problem (b) presented values that were redundant (in essence, presenting P(A), P(not-A), P(B), P(not-B)). Four subjects in the non-explanatory group used the wrong equation in an attempt to incorporate all four values in their answer (e.g. ".2 x .4 x .6 x .8") and one made what we would characterize as a translation error, multiplying .2 by .4. In contrast, no subject in the explanatory group used numbers other than .2 and .6 in their answers.

The answers to problem (a) indicate a serious translation difficulty. The problem states P(A) and P(not-B), and asks for P(A and B). Of 23 subjects in the non-explanatory condition who erred in response to this problem, 11 multiplied P(A) by P(B); only one subject in the explanatory condition did this.

In contrast to performances on problem (a), the non-explanatory group did very well in answering problem (d). This states:

> ".7 of the members of the state legislature favor a particular bill. .6 of the members are Democrats. If party and vote are independent, what is the probability that the next legislator you meet is a Democrat and votes for the bill?"

We suggest that this problem was very easy for the non-explanatory group because they have memorized the rule: "Whenever the probabilities of something and something else are required, multiply their individual probabilities". In problem (d), the word "and" is present and there are only two possible values to multiply. Fortunately, they are the right two values. The same circumstances hold for problem (a) except that the values given are not right; one value must be subtracted from 1.0 before multiplying.

Equation 4. This equation is the general case of Equation 3 and, therefore, it would seem that subjects should score about the same on this problem set as on the earlier one. That was the case for subjects in the explanatory group;

in the non-explanatory group, however, the average proportion correct was .22 lower for the problems testing knowledge of Equation 4 than for those testing knowledge of Equation 3. For the explanatory condition, the proportions correct were, as in the other problem sets, relatively stable: (a) .5 (b) .4375 (c) .625 and (d) .5625. Also, as before, the non-explanatory group exhibits greater variability of performance over the four problems: the relevant proportions were (a) .1563 (b) .1875 (c) .3438 and (d) .625.

That the non-explanatory had greater difficulty with these four problems than with the third problem set may reflect two factors. First, these problems typically omit explicit use of the key word "and"; this may cause classification problems. Second, the probability of the complementary event is often not explicitly stated; thus, there may be translation difficulties.

Both conditions exhibit about 16% classification errors as indicated by use of operations other than multiplication. More striking, at least in the non-explanatory condition, were errors of translation indicated by using the stated probability when its complement was appropriate. Averaging over problems (a), (b) and (c), 11.67 subjects in the non-explanatory group made such errors. Only 2.33 subjects did so in the explanatory group.

The non-explanatory group performed reasonably well on problem (d) which states:

> "A hockey team wins with probability .5, loses with probability of .3, and ties with probability of .2. What is the probability that in the next four games, the team wins, then loses, then wins, then loses?"

The statement of this problem more closely parallels the statement of Equation 4 than does that of the other problems in this set, thus reducing the likelihood of misclassification. Furthermore, all the required probabilities are stated; translation difficulties are presumably minimal.

CONCLUSIONS

Very different patterns of knowledge appear to be present in subjects in the non-explanatory (standard and low-explanatory texts) and explanatory conditions. Subjects in the two non-explanatory groups performed considerably less well on story than on formula problems, and often used the correct formula for a problem (that is, met the lenient criterion) but failed to solve it. A closer look at answers to story problems revealed that subjects in the non-explanatory conditions often required the explicit presence of key words which unambiguously pointed to certain operations, tended to misclassify problems in the presence of irrelevant or redundant information, and made many errors when the values

in the story required modification before insertion into the formula. In contrast, subjects in the highly-explanatory condition performed equally well on story and formula problems, tended to solve whenever they showed evidence of knowing the appropriate formula, and were considerably less hindered by absence of key words, the presence of irrelevant information, and the need to translate values in the story. The situation may be summarized by noting that, although all our subjects were novices, those in the high-explanatory condition seem to share more of the characteristics of experts than do those in the other two conditions.

The results of more recent studies lead us to believe that the solution processes in the explanatory and non-explanatory conditions are qualitatively different. In one study, subjects in a high-explanatory condition had only 60% correct recall of equations (as opposed to 90% for a standard group), but were able to perform correctly on 53% of story problems. This result suggests to us (as does the equivalence of formula and story scores in the study just detailed) that subjects in the high-explanatory condition often may not retrieve the formula at all. Another study in which subjects thought aloud while solving supports this conjecture; after studying a high-explanatory text, students tended to attempt to construct solutions, after using the diagrammatic aids provided by the text. Rather than recalling formulas, they attempted to recall examples from the text which they felt were similar to the problem at hand, and to use these as a model for solution.

Much remains to be done. High on the agenda must be the use of other measures of comprehension - both problems demanding more complex processing and transfer to new materials such as the Binomial and Bayes' theorems. We suspect that such research will provide further evidence of the role understanding can play in mathematical problem-solving.

Acknowledgements

This research was supported by a grant from the National Institute of Education, NIE G80-0126, and by a grant from the National Science Foundation, BNS 80-01181. Randall Hansen, Joan McCann, and Rikki Robson made major contributions to all phases of the research reported in this paper.

PROBABILITIES: LEARNING BY DOING OR LEARNING BY IMITATING

A Abele, Pädagogische Hochschule, Heidelberg

INTRODUCTION

Learning mathematical terms like 'frequency', 'random event', 'probability' and the like is closely connected with the means of illustration that you - the teacher or the pupil - choose. In solving descriptive mathematical problems, school children can be grouped into two different types: those who prefer graphs and graphical procedures like situational outlines or diagrams and the like; and those who prefer to choose a verbal form of expression and who like to work with symbolic means.

The learning situations for the first type of learners have to satisfy two conditions:

(1) The teacher has to organize his instruction as the external condition for the learning approach of the pupil in such a way that the pupil is able to develop his creative abilities in looking for 'roundabout solutions'.

(2) This is particularly successful when graphs can be used as an iconographical means that can be actively employed by the pupil. Even incorrect solutions may support the recognition process of the pupil, because a comparison of the graphical trial solution with the situation that is to be depicted that is undertaken by the pupil himself gives him much more insight than a correcting direct hint of the teacher.

INVENTION OF GRAPHICAL DESCRIPTIONS

The application of graphical means of solving an existing problem is, in a school situation, almost always enhanced by the teacher. One specific graph that had proved to be particularly apt for solving the pertaining problem is taught to the pupils as the standard procedure. As a rule, the pupil does not take part in the development process of this procedure, instead, it has already been 'invented' and is merely accepted by the pupil. If this attitude is to be corrected, one has to know how graphical solutions are arrived at via preliminary or even incorrect intermediary steps or phases.

The first questions are:

- What do children of different ages do in specific situations when the teacher wants them to invent a graphical description of a random process?
- Is it possible to bridge the gap between the pupil's proposal and the standard solution which is in the mind of the teacher?

We look at two typical examples:

1st example: Repetitive throwing of a coin (grade 7/8)
2nd example: Passing four roads by one/two/three cars by chance (grade 8).

First Example: Repetitive Throwing of a Coin

We throw a coin twice, three times, four times etc. Which combinations of head and tail could there be in such a series of trials? The solutions of the pupils are very informative. In confronting the pupils' proposals with the standard solution which is in the teacher's mind we recognize the gap between these two ways of solving the problem.

As a rule we solve the problem

(1) with graphical means using the tree diagram

The coin was thrown

for the first time

for the second time

for the third time

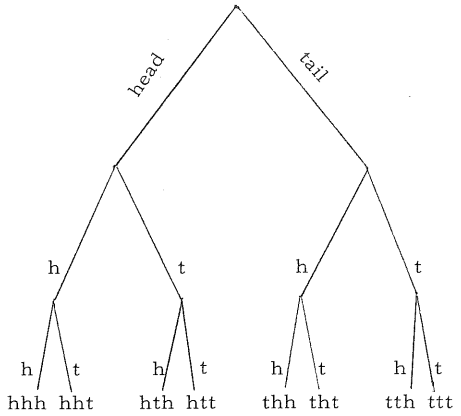


FIGURE 1

(2) or by calculating all combinations (variations with repetition), which are possible through two letters (or colours) on three places.

It has to be mentioned that the children had not been made familiar with either of the two methods. Some of the pupils declared they had not learned how to solve problems like this and waited for some help from their teacher. Others, however, set about solving the problem. In doing so one of the pupils proposed the following solution. He selected the iconographical level and came up with the following outline.

The coin was thrown

for the first time

for the second time

for the third time

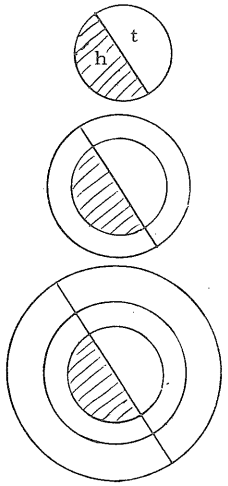


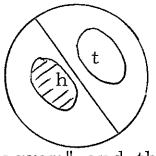
FIGURE 2

This illustration could only be applied to the situation when the coin was thrown for the first time. It is obvious that it could not be generalized.

Another pupil presented the following very remarkable solution.

The coin was thrown

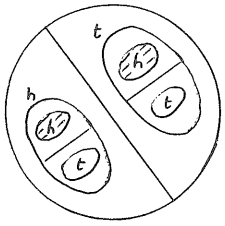
for the first time



He is using something similar to a "Venn diagram" and therefore in his thinking he has to put the elements of the set - the two events, which are represented by the two coloured zones - inside the boundary line of the Venn diagram.

What may the pupil do in applying his first proposal?

for the second time



for the third time

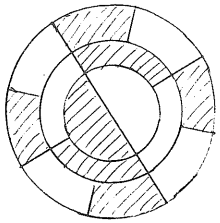


FIGURE 3

This is a correct solution and an illustration that could be generalized - found by the pupil. For that reason it must not be ignored even when we think it would be necessary for the children to learn the concept of decision-tree and its application. The approach taken by the pupil leads to the following consequence: From a methodological point of view it may often be wrong to have the pupils learn a standard procedure introduced by the teacher with the assumption that they will afterwards learn other procedures or inconographical representations. On the contrary we see the possibility of organizing the lesson in such a way that the pupil is able to develop his creative abilities in looking for self-made illustrations.

Second Example: Passing Four Roads by One, or Two, or Three... Cars by Chance

Four children of 8 grade are working with a map. Two villages A and B are connected by four roads.

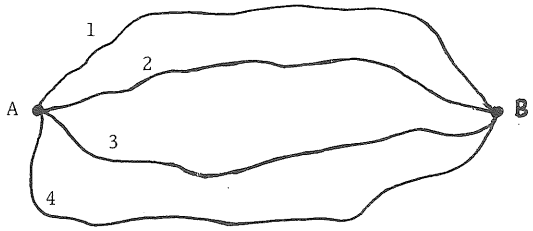


FIGURE 4

First problem: Two cars (red respectively yellow) run from A to B. How many possibilities have they to choose a road, if they run independently and by chance? The children try to solve the problem with the understanding that the yellow car takes another road than the red car. They presented their solution using two different illustrations.

first proposal

$\begin{pmatrix} \text{red} \\ \text{yellow} \end{pmatrix}$ $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$

FIGURE 5

second proposal

yellow ---------->

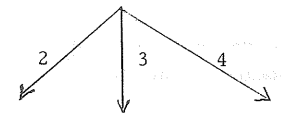


FIGURE 6

In every case the solution is not complete. The children do not notice the other possibilities. In these 9 cases the red car does not pass road one but road two or three, or four:

- red road 2 and yellow road 1 or road 3 or road 4
- red road 3 and yellow road 1 or road 2 or road 4
- red road 4 and yellow road 1 or road 2 or road 3

Second problem: How many possibilities have the two cars to choose one of the four roads by chance if the yellow car does not take necessarily a different road from the red car? The pupils suggest a new illustration.

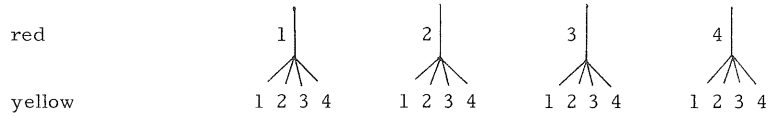


FIGURE 7

I want to accentuate two points:

- The branches relating to the red car do not start from one point
- The children do not use the illustration of ordered pairs $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, as they did above.

Third problem: One pupil makes the following suppositions:

- The red car and the yellow car take different roads.
- Each of the two cars gets two of the four roads to choose one of them, for instance: red gets the roads 1 and 2, yellow the roads 3 and 4.

For this part of the problem the solution is illustrated by the pupil twice:

"decision tree" as above

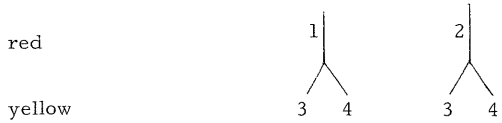


FIGURE 8

ordered pairs in the sequence (1,3) (2,3) (1,4) (2,4)

But nobody sees the other 20 possible pairs as further solutions. This is a typical phenomenon which is caused by the specific difficulty of probabilistic problems: For the pupils nothing is a real situation, which could have been mathematicized, all is hypothetical. If they pick out the four cases, which

can be realized at the same time, they have created their reality in their mind and the fictive totality of the 24 cases has disappeared.

Fourth problem: Three cars run from A to B. How many possibilities have they to choose a road, if they run independently and by chance?

First the children constructed the following illustration as a continuation of the figure they had used before.

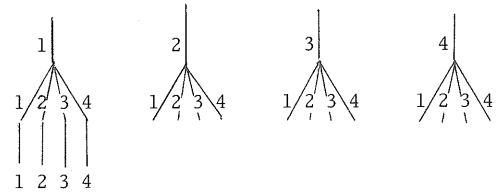


FIGURE 9

But this illustration is not suitable for a propitious description of this situation. After this first step one of the pupils proposes this figure



which is in use everywhere. FIGURE 10

Fifth problem: What about 4,5,6 etc. cars and the four roads - how many possibilities? This is now very easy. One of the pupils mentioned: "Always this is quadrupling". They no longer used an iconographic representation because now they had the general structure in their imagination.

PROPOSALS OF LEARNING ACTIVITIES IN THE "HAUPTSCHULE" (SECONDARY SCHOOL FOR SLOW LEARNERS)

Taking the observations above mentioned and the like as a basis we now want to offer some proposals for starting activities in the learning of probabilities in the "Hauptschule", which is a special secondary school for children on a very poor and low level. The proposals are both problem oriented, the first referring to some combinational problems which are very challenging for the pupils, the second starting with a well-known game whose analysis leads to the concept of probability.

Combinational problems

- Initial problem: Lotto, 6 numbers are drawn out of 49 numbers.
- Second problem (reduced): Lotto, 2 numbers out of 5.

- Third problem: Lotto, 2 numbers out of 6.
- Fourth problem: Lotto, 3 numbers out of 6.
- Fifth problem: Permutations
- Generalization
- Combinational basis problems

The concept of probability

- Initial game: "Mensch ärgere dich nicht" (game at dice).
- Analysis of the game by isolating single situations
- Probability, first definition.
- Decision tree and path diagram: Throwing a die two times, three times etc.
- Game at dice using four dice.

HOW DO PUPILS OF A SECOND GRADE HANDLE COMBINATIONAL PROBLEMS? DESIGNING COLOUR PATTERNS AND COMBINATORIAL COUNTING

As we all know, group instructional methods promote the learning of mathematics in situations of exercising. If - within a well-defined area - certain techniques could be trained and retained, group specific phases of work have a stimulating effect on the performance and the social behaviour of the pupil. In these cases group situations promote mathematics-related discussion and goal-oriented actions of the children. Therefore it is supposed that group instruction in mathematics namely in probabilities also has positive effects for the identification of new insight. So we expect - according to some preliminary studies - the following from the pupils' work:

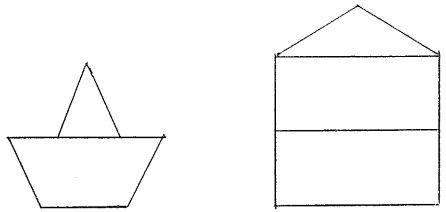
- The pupils can - without major assistance - dissect the problem into parts of problems and can thus work on those step by step before they reach a solution.
- The working material and the task assignment will enhance in the pupils a problem-specific willingness to discuss which is the basis for successful group work.
- The group situation will create possibilities to formulate solution hypotheses and to discuss and review conceptual suggestions on the basis of arrangements of the materials which can also be quickly discarded. Drawing, which takes time, and written notes, which bear the inherent and well-known difficulties of subsequent correction, are not necessary.

Didactical Analysis

Four children of one group of the second grade receive paper shapes in different colours with which they are supposed to lay out houses and ships. After a first, relatively free working phase, the teacher suggests to the pupils to

lay out ships from two paper shapes and houses from three paper shapes according to Figure 11 and to try to find as many colour combinations as possible.

FIGURE 11



Ship House

Through these instructions the mathematical problem is put into a concrete setting and is thus accessible for the pupils on the inherent active expressive level. The aim of the project is the gradual systemization of all colour variations (variations with repetition), which are possible through three colours in two places and through two colours in three places. The survey solution - in this form not really the direct aim of the pupils' work - is easily seen with the aid of a decision tree.

There are a total of nine possible colour combinations according to this rule.

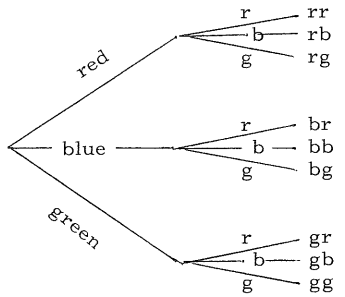


FIGURE 12 - Decision Tree A: Ships

There are a total of eight possible colour combinations according to this rule.

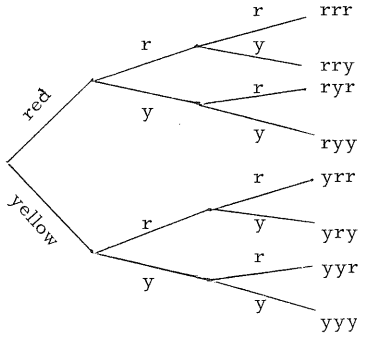


FIGURE 13 - Decision Tree B: Houses

Sequence of the group work

During the childrens' work one can differentiate clearly among five work phases:

First phase: Free construction

Second phase: Construction according to a rule - preliminary solution

Third phase: Review of the preliminary solution

Fourth phase: Finding individual aspects - beginning systemization

Fifth phase: Reaching a total solution through combination of the individual aspects - complete conscious systemization

Thus, without formulating it so specifically, the total problem is divided by the children into the following assignment parts:

1st part assignment: We build a great number of houses or ships (any kind)

2nd part assignment: We build a great number of houses or ships (specific)

3rd part assignment: We look for duplicates and try to fit the rest of the paper shapes in also

4th part assignment: We sort out all houses or ships

5th part assignment: We draw or note the results in a distinct form

The second phase was introduced by the teacher with the request that the pupils limit themselves to specific types of ships and houses; the fourth phase through the request to see "how many houses and ships are there". On the other hand, the transition from the second to the third and from the fourth to the fifth phase are fluent; they came about without the help of the teacher through discussion among the pupils themselves.

A careful analysis shows some remarkable details in the attitudes of the pupils.

First phase: Free construction

The pupils strive to construct a great variety of forms. They design two types of ship and four types of house (see Figures 14 and 15) which they offer in the various colour combinations.



FIGURE 14: Types of Ship

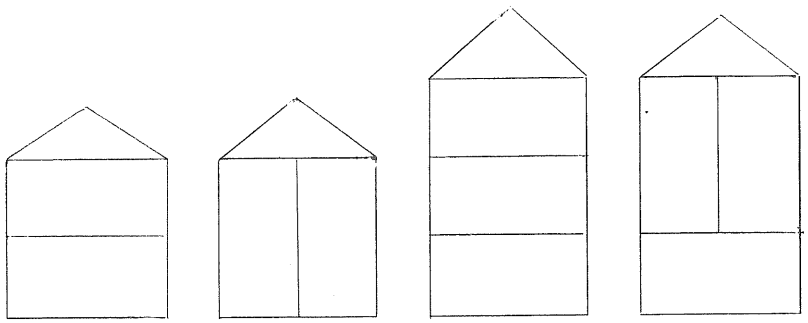


FIGURE 15: Types of House

For the variations aesthetic colour arrangements were above all decisive ("red and blue go well together"), the quantitative aspect, to find a great number of different colour combinations hardly played a role in this phase. The teacher intervenes when nine ships and nine houses are built and introduces the second phase with the additional rule.

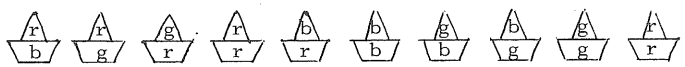
Second phase: Construction according to a rule

The pupils work in two directions, some rebuild the existing houses and ships to fit the requested forms, others construct new ships from the existing paper shapes. This work division without mutual consultation leads to the fact that some ships and houses occur twice. When this is noticed ("These are double, that we already have up there, now we have to put in a different colour"), the pupils begin to control the results that they have so far produced.

Third phase: Review of the preliminary solution

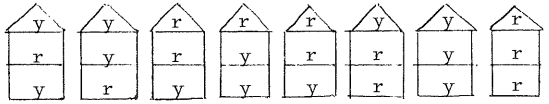
In order to find out whether "everything is all right", the children adopt two strategies: they sort out grossly separating ships from houses which, at the beginning, were mixed together and search for doubles. Moreover, the children try to build more figures with the rest of the shapes ("Maybe now we get it all done"). Toward the end of this control and supplementary phase ten ships and eight houses are laid out on the table:

Ships of two colours consisting of the colours red (r), blue (b) and green (g):



The combination  , which occurred for the second time was obviously not seen by the pupils.

Houses of two colours consisting of red (r) and yellow (y):



Here again, it was not seen that the combination  occurs twice.

The combination  is, however, missing.

In this control activity the pupils were invariably led by aesthetic viewpoints or the expectation of a certain regularity ("Look here, this is red-red, yellow-yellow", or: "yellow-red, red-yellow, those two there..." What was meant were the colour combinations  and ).

These beginnings of a systematization can be recognized more in the next phase, which is introduced through the teacher's question asking for the number of the figures.

Fourth phase: Finding individual aspects

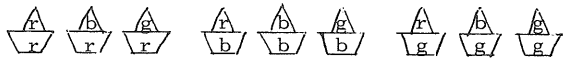
The question of how many houses or ships is initially quickly answered. When a more differentiated questioning occurs, the pupils follow different classification viewpoints.

Classification of ships

a) Grouping according to sails and boat hulls of the same colour. This method follows the taxonomy of the decision tree A (Figure 12) and offers the following three groups



or



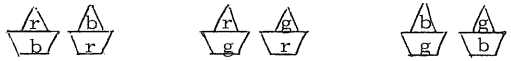
b) Groupings according to one-colour or multi-colour. You get three one-colour ships



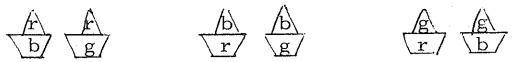
and 6 multi-colour ships.

c) First continuation of b)

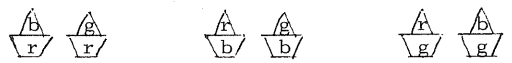
Subdivision of the multi-colour ships alternating in pairs



d) Second continuation of b)
Subdivision of the multi-colour ships in pairs according to sails

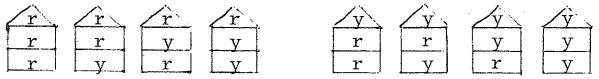


e) Third continuation of b)
Subdivision of the multi-colour ships in pairs according to hulls

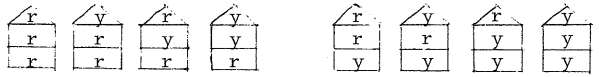


Classification of houses

f) Grouping according to roofs and floors with the same colours. This method again follows the taxonomy of decision tree B (Figure 13) and provides the two groups



or

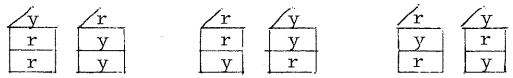


g) Grouping according to one-colour or multi-colour. One obtains two one-colour houses

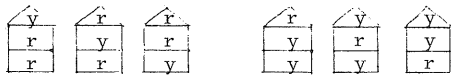


and 6 multi-colour houses.

h) First continuation of g)
Subdivision of the multi-colour houses alternating in pairs



i) Second continuation of g)
Subdivision of the multi-colour houses according to the distribution of the colours, two times red - one time yellow or one time red - two times yellow

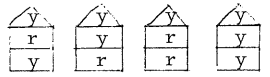


Since those continuing subdivisions c), d), e) and h), i) are not realizable at the same time, this variety of solution possibilities requires discussion within the group. Depending upon how the ships or houses look, the pupils will particularly notice different aspects; the different suggestions for a solution have to be coordinated and they have to consult about the future path. This is a typical situation in which the working group is under a strong pressure to talk, deriving from the mathematical state of affairs in its methodological development.

The pupils, who are not yet aware of the taxonomy above mentioned, but only start discovering it through their own activity, showed a behaviour that is typical in such situations. They first recite conspicuous single aspects, which at first stand next to each other in a very uncoordinated way; suspected regularities are formulated as statements which often determine the direction for the further proceedings and then prove to be right or wrong. In the fifth phase, the final coordination of the singular aspects towards a total solution has to be performed.

The individual analysis shows which solution strategies are favoured by the pupils:

Pupil: "Four yellow roofs"
He means



(Taxonomy f)

Pupil: "Wait a minute, see, and now four red roofs".
This is hypothetically formulated in analogy to the 4 houses with the yellow roofs since thus far there are only three houses with a red roof. The combination



is missing. Therefore, other pupils now react

Pupil: "No, six red roofs."
Pupil: "Five red roofs."
Pupil: "There, three red roofs."
With this, this attempt to systematize remains unfinished.

Pupil: "No, three red sails first."
He means taxonomy a)



This pupil had not taken part in the discussion on the houses and now pulls the other pupils into a discussion about the boats.

Pupil: "Three red boats". 

Pupil: "I'm sure there are a lot of mistakes there. Three red sails and three green sails."

Pupil: "How many blue ones? Three blue sails."
With this all the singular aspects of taxonomy a) have been formulated by the children, but still, they cannot see the total solution. The discussion on the grouping of the ships goes on. In doing so, the singular aspects are always related more closely to each other, so this can be considered to be about the transition to the fifth phase.

Fifth phase: Reaching a total solution through the combination of the singular aspects

Pupil: "Three blue ships! No, that one is red and below it's blue."

Pupil: "That's still right. Three blue ships. Should we now write three green bottom parts?"

Pupil: "Why don't you write: one green ship?"
(Change over from taxonomy a)

to taxonomy b)).

Pupil: "But this is just not possible."

Pupil: "Now I'm beginning to get it. I think that each ship is supposed to have the form red-red, blue-blue, green-green."

Pupil: "Write down, one green ship."

Pupil: "One green, one red, and one blue ship."

Pupil: "One red ship - there are more colours."

Pupil: "Six multi-colour ships." (Taxonomy b)).

Pupil: "Multi-colour - those are more colours."
Despite this objection, this attempt to solve the problem rests here after taxonomy b). A continuation according to taxonomies c), d), or e) is not tried by the pupils. They now draw their attention again to the houses and here go over to taxonomy g).

Pupil: "What now? Now we have to do the houses."

Pupil: "One yellow house, one red house."

Pupil: "Six multi-coloured houses."

Here, too, the pupils are satisfied with this partial solution. They do not have the need further to refine the general overview which they now possess. In order to continue their work, it would now be necessary for the teacher to step in perhaps in the form of a new problem-posing or other guidance. One could also think of structuring the pupils' work from the very beginning towards a more differentiated work result by handing out suitable empty forms on which ships or houses have to be grouped.

The result of the work was retained by the pupils in written form

	commentary
4 yellow roofs 4 red roofs	taxonomy f)
3 red sails 3 green sails 3 blue sails	taxonomy a)
1 green ship 1 blue ship 1 red ship 6 multi-colour ships	taxonomy b)
1 yellow house 1 red house 6 multi-colour houses	taxonomy g)

The continuing subdivisions were not explicitly formulated by the children. They had, however, on the active level, silently always been using them when, for example, they placed pairs of figures - taxonomy h).



Final Remarks

The analysis of this example lets us see that group instruction also has positive effects for the identification of new insights. Obviously, the group situation favours the willingness of the children to discuss. Their different suggestions for a solution and conjectures represent a constant challenge for the spontaneous exchange of argument and counterargument. The attempts of the children at reasoning, which were often just indicated on the active level, bring light

into the factual relationships, sometimes better than any verbal explanations on the part of the teacher, however well they may have been meant.

The group situation has proved successful when it comes to constructively building up a situation for solution especially in the third work phase, which dealt with the analysis of the situation which had been constructed by the pupils themselves. The joint commitment is often expressed by the pupils through statements like

- in the area of emotional concern - "Look, here, what I found."
- in the introduction of new problem-solving suggestions - "Or we can make everything yellow."
- in asking or looking for reasoning - "Why did you do that?"

On the basis of these observations, the idealism of many teachers who in their instruction classes rely on the curiosity and creative abilities of children, seems to be well-founded. Well-known theses are confirmed, and at the same time, taken out of the field of theory and are made more concrete in the school practice:

- Strong commitment and dedication on the part of the pupils with regard to the instruction can almost always be expected when the pupils are given the right to take a real part in making the decision for the directly following instruction. The participation of the pupils in the instruction gets more intensive because they have made the events in the class their own cause.
- The direct measures taken by the teacher are willingly accepted - they therefore have a positive effect on the learning progress - if they not only seem to concern the pupils' statements but also provide the pupils with a clear purpose.
- Appropriate combinational problems in favourable methodological development forms are particularly apt to be the natural basis of group discussion. The pupils can formulate solution hypotheses by themselves, attempt to verify them and give reasons for what they are doing and can independently carry out final control and control calculations.

TOPIC 19: TEACHING OF AGRICULTURAL STATISTICS

ORGANISER: DR D SINGH

TRAINING OF PERSONNEL FOR AGRICULTURAL STATISTICS IN JORDAN

Z R Takrouri, Department of Statistics, Amman

INTRODUCTION

Agricultural statistics is considered the basis of all agricultural decisions related to development plans in the agricultural sector. In most developing countries, planners try to benefit from the kinds of agricultural statistics readily available. In many cases they find that the statistics are either incorrect or incomplete. This makes them look for other sources of information available in the publications of international agencies, such as FAO, or try to compromise a figure taken from several sources.

The importance of getting fairly accurate agricultural statistics makes the programme of training the personnel of agricultural data producing agencies on better ways of improvement an urgent need. In Jordan, the government established in 1963 the Statistical Training Centre connected with the Department of Statistics. The purpose of setting up this centre was to train the governmental and non-governmental statistical personnel on how to raise the level of accuracy of the different statistical figures through learning statistical methods and their applications. A complete course (3 credit hours) in agricultural statistics was among the courses to be taught to the participants.

This paper will throw light on this kind of training of personnel, the problems faced during it, and suggestions toward improvement.

AGRICULTURAL STATISTICS

A. Definition: There is no fixed definition of agricultural statistics, but we can say that it is the statistics which deals with agriculture and covers all the agricultural aspects. But since agriculture is connected with other sectors of economy, one may be confused with some statistics whether one can consider them as agricultural statistics or not. Many examples can be given on these inter-related statistics. Agricultural constructions, food exports and imports, and loans given to farmers by credit corporations are some examples of statistics which deal with two sectors and can be considered in both sectors for specific purposes. In spite of this, one can consider any statistics which throw some information on any item related to agriculture as agricultural statistics. On the other hand, agricultural statistics covers many specific terms that ought to be defined. Usually in each of the censuses and surveys conducted in any country, some space is devoted to definitions and concepts

while reporting the results. This will make the reader or data user's job much easier. A page or two of definitions in the enumerator's guide will enable him to understand the questionnaires and fill them in accurately.

B. Need for training: Agricultural statistics deals mainly with quantitative descriptions of land and livestock. The work involved is voluminous and therefore, in many countries, particularly in developing ones, sample survey technique is used for collection of data. This requires training for the persons who work on such projects. Map reading and sketch preparing need special training. For quality check purposes, accurate area measurement of selected plots may be needed. Area measurement technique has its own methods and equipment, and it is not easy for any person to learn without training. Even the conduct of crop-cutting experiments to estimate the average yield of production per area unit needs a special type of training.

From what is stated above, one can conclude that a new graduate cannot enrol himself in the agricultural statistics professional team and work with them without getting special training in the work needed in connection with agricultural statistics.

C. Surveys conducted in Jordan: So far three agricultural censuses have been conducted in Jordan since the adoption of the present constitution. A fourth one is planned to be implemented in 1983. Besides the censuses, three annual surveys dealing with land are conducted. Two of them are on a sampling basis while the third is based on a complete enumeration of lands irrigated by underground water. From these surveys, the area and production statistics of crops are obtained in the country. We are now in the process of adopting a new scheme of livestock and poultry statistics. An agricultural household survey is planned to be conducted this year to study the socio-economic aspects of rural households. This survey will provide data users with many socio-economic indicators approved by the last conference of FAO for measuring the development of rural areas. Conducting of extra specialized surveys is handicapped by the unavailability of funds and qualified personnel.

Objective methods are followed to collect the information whenever resources are made available. Area surveying for selected plots is made to correct the farmers' responses. Annual crop-cutting surveys to estimate the average yield per area unit for wheat, barley and olives have been conducted in the country since 1969.

It is obvious that the conduct of such surveys needs qualified personnel to work in the agricultural statistics section in the Jordanian Department of Statistics. This requirement has been met mainly by organising training within the country and sometimes by deputing abroad.

QUALIFICATIONS

A. Trainer: It is not an easy job to train people in how to conduct censuses and surveys especially in agriculture. At the same time, it is difficult to find qualified trainers who can be entrusted with this duty. It is difficult in the developing countries to find a qualified statistician, but it is more difficult to find a qualified agricultural statistician. Maybe this is due to the difficulty associated with the fieldwork of agricultural surveys.

A good trainer should firstly have a good academic background, preferably a first degree in agriculture. The work relating to agricultural statistics needs this background since it deals with different aspects of agriculture and this is not easy if he is not well trained in agriculture. Agricultural faculties in universities in many developing countries neglect statistics courses. This means producing good agronomists with poor statistics knowledge. This problem can be solved by studying for one or two years in a statistical institute. An experience of at least five years in agricultural statistics work should be considered an essential qualification of a good trainer. This experience will teach him how to apply the academic study in the field work. He will learn the problems that he will face and the best way to solve them. He will learn how to deal with people, how to treat them, respect them, and convince them to respond correctly. During that period of experience he will be familiar with the geographical map of the country and the location of villages, roads and administrative divisions.

B. Trainee: All the trainees should be from organizations already engaged in statistical work, or those who will be employed in this work after completing the training. Their academic study should not be less than the general secondary certificate. This will enable them to understand mathematics, sampling and English terms without much difficulty. Trainees coming from rural areas and knowing agriculture will find the training in agricultural statistics very easy.

AGRICULTURAL STATISTICS COURSE IN THE JORDANIAN STATISTICAL TRAINING CENTRE

A. Contents: It is well known that statistics are obtained mainly through field work. The volume of field work differs from one kind of statistics to another. Agricultural statistics involve a big share of field work. Thus, a satisfactory course in agricultural statistics should not ignore the role of field work in the agricultural statistics process.

The syllabus of the agricultural statistics course being taught in the Jordanian Statistical Training Centre includes the following topics:

a. Agricultural census operations.
b. Methods of collecting data.
c. Kind of sampling frames.
d. Area measurement techniques.
e. Crop-cutting experiments.
f. Farm prices.
g. Agricultural production index numbers.
h. Cost of agricultural production statistics.
i. Agricultural income accounts.
j. Preparing and reading maps and sketches.
k. Data processing methods.
l. Publication of survey reports.
m. Questionnaires.
n. Field work instructions.
o. Plan of work and budget preparing.

B. Duration: The agricultural statistics course is among the courses taught to second year participants during the second semester. The duration of this course is about 15 weeks. Three lectures of one hour's duration are given in each week. Sometimes additional time is needed in order to be able to cover the whole topics especially when some holidays take place on the day of the lectures.

Field work operations are illustrated in the class room for the first time - conducting crop-cutting as an example - then one or two lecture periods are spent at a nearby location to demonstrate them on the spot. This actual demonstration makes the operation more clear and easy to stick in mind.

The Jordanian Statistical Training Centre organized two separate agricultural courses. The first was in June 1981 for three weeks' length about agricultural censuses in cooperation with the United Nations Food and Agriculture Organization. Twenty participants from local Jordanian organizations and Arab statistical agencies attended the course. The second course was held in March 1982 for five weeks' length covering different topics in agricultural statistics in cooperation with the Arab Institute for Statistical Research and Training. Thirty participants from local and Arab statistical agencies attended the course. The evaluation forms filled by the participants in the two courses indicated that they were very useful and successful and suggested that the duration of the course should be longer.

C. Field work: Field work training is considered as complementary to training in the class room. So, whenever there is a topic in the course in agricultural

statistics which needs an illustration in the field, a special time is devoted for this purpose. The field work training included the following operation:

a. Area measurement of the fields to obtain correction factors of the area reported by the respondents.

b. Crop-cutting experiments to estimate the average yield of production per area unit of the three principal crops in Jordan: wheat, barley and olives.

c. Pre-test operation to test the questionnaire and to conduct a pilot study on selected locations to examine the validity of the instructions for the whole operation, and how to obtain public cooperation.

In order to ensure the attendance of participants at the field work training, a weighting of about 40% of the course mark is devoted to it and the rest of the mark to the class examinations. Each participant is given the chance to practise by himself the field work operation by forming groups to do the work. Transportation facilities to and from the spot are provided by the Department of Statistics. Equipment required for this type of training is also provided by the Department.

PROBLEMS AND SUGGESTIONS FOR IMPROVEMENT

Training in the Jordanian Statistical Training Centre lasts for two years. The course in agricultural statistics is among the courses taught during the second year. Since that is the case, it is not easy to talk about the problems facing this course only. The following problems are faced in the course of organising the whole training programme including the agricultural statistics training:

a. The majority of the participants, especially those coming from the Gulf Countries, are poor in their academic background, while some possess very good background. This difference in their background makes the trainer's job very difficult. Such a problem can be solved by imposing strict conditions for admission of participants coming from countries having poor educational standards, or by introducing an examination for admission.

b. Many countries do not give any incentives to the successful participants, which makes them disinterested in the course.

c. In some cases the participant is not employed for statistical work after finishing his training programme. He may work with completely different work. This makes him feel disappointed toward his long training programme.

d. In Jordan there are two universities. Both do not give any recognition to study in this statistical centre. This makes the participants not eligible to complete their study in the university. Thus, many are discouraged

to enrol in the training programme at the centre. A continuous exchange of views to settle the matter is going on between the administrations of the two, the universities and the Department of Statistics.

e. An obvious weakness in the English language level among participants is a serious problem. To overcome this problem, a special course in English language is given during the first year of training to all participants.

f. Due to lack of funds, the payment per one hour lecture is low. This makes the instructors look for work in other institutes or with the universities, who usually pay them better. This can be solved by increasing the governmental financial support to the Department of Statistics budget.

g. Other miscellaneous problems can be mentioned which can be solved when additional funds are available, such as the unsuitability of the lecture rooms, both in size and in location, inadequacy of library facilities, etc.

Finally, we may say that in future additional time should be allowed for more courses in agricultural statistics. The course taught at present is very crowded and can be divided into two or three specialized courses.

CONCLUSION

Agricultural statistics is very important to development plans in the agricultural sector. A successful development plan should depend on accurate figures. Getting accurate statistics requires a good training of the personnel employed for agricultural statistics in the best methods to collect, tabulate, analyze, and publish data.

The Jordanian Department of Statistics realised the importance of teaching a course on agricultural statistics to the participants enrolled in the study and training programme at the Statistical Training Centre. This course covers many topics of agricultural statistics. Also, separate training programs were organized in co-operation with some international agencies to train participants from local and Arab statistical organizations on agricultural censuses or surveys.

There are some qualifications that should be possessed by the trainer in agricultural statistics. Likewise, the trainees should possess good academic background. The availability of these qualifications will lead to the best results.

Since the field work is a very important element in agricultural statistics, it is essential that the contents of the course should contain both theoretical lectures and practical field work. The minimum duration of such a course should not be less than 15 weeks.

A few problems that are being faced by the Statistical Training Centre in organising the course can be solved by suitable action and by making available additional fund resources.