THE CHALLENGES OF STATISTICAL TRAINING

R.R. Harris
Lancashire Polytechnic
Preston, Great Britain

G.K. Kanji
Sheffield City Polytechnic
Sheffield, Great Britain

Introduction

Although the drawing of inferences from statistical observations is (presumably) as old as time itself it is only relatively recently that this has been undertaken in a logical and ordered pay. It was in the 1920s that formal statistical education began in such places as University College, London and the University of Cambridge. As statistical theory developed through the 1930s, 40s and 50s so did statistical education in the Universities, being followed by a massive explosion in the 1960s and early 1970s.

In this period 1930 - 1970 most education of statisticians was in the University sector and Departments of Mathematics and primarily, but not exclusively, at the graduate level. There was a gradual recognition of the importance of statistical methods in the education of the other disciplines, for example biologists, sociologists, geographers, largely a consequence of an increased use of statistical analyses in the practical applications of such disciplines. It must also be remarked that the examining role of the Institute of Statisticians (IoS) since 1948 on a worldwide basis led to a number of correspondence courses undertaken by those in full time employment.

The 1970s saw an increase in the number of undergraduate statistics courses with a wider range of higher education institutions becoming involved; for example the polytechnics of the United Kingdom, technical colleges in other countries, and indeed specialist training centres, with sub-degree work also being developed. A particular feature of statistics teaching in the non-University sector was the applied nature of the work, not only for statisticians but also in the quantitative components of other disciplines' courses.

In the last five years the role of the statistician has shifted significantly due directly and indirectly to the development of computer technology. Directly in that vastly more data is being collected, stored and is accessible - witness the equipment on the executive's desk. Indirectly because this computer power allows for more sophisticated statistical methodology, for example in the numerical solution of equations with no explicit solution and the availability of simulation techniques; indeed the whole development of statistical modelling as we now know it would have been inconceivable even a decade ago. Thus the role of the "professional" statistician is developing from simply being one of a data analyst to a far more general one of information scientist, providing additionally in-depth professional support for statistical software and systems. The growth of ex-

pert and guidance systems may lead to more emphasis on a "trouble-shooting" role in non-standard situations as well as, necessarily, author and validator of such systems. A yet wider range of applications is appearing and the applicability of statistical techniques to these new areas must be carefully assessed.

The Professional Statistician

As implied above, more than ever before the professional statistician needs to be a polymath for, in addition to a sound knowledge of his own subject and the skill to apply the appropriate techniques, he needs a range of peripheral skills. He needs to be able to communicate effectively in a range of environments and this is not a one-way process. Surely he needs to be able to communicate his results but he also needs to be able to discuss what is required during the initial stages of an investigation. Often this means discussing with the client exactly how the current work fits into a wider context and consequently teasing more and more information, sometimes felt to be irrelevant, from the client. Thus he needs to be able to understand aspects of the discipline in which the problem lies; he needs to ask appropriate questions, sometimes apparently naive, with tack.

How then can this range of skills be incorporated into the programme of education of statisticians? It is clear that in the past statistical educators have concentrated too much on the theory of statistics and left the practice of statistics and the other, essentially non-statistical, skills referred to above to the employer. Statistics is not unique in this, how many times does one hear of the graduate physicist unable to change a fuse, and so on? Until the mid-1970s although the primary responsibility for the provision of skills of the graduate statistician lay with the educators, employers had to share some of the responsibility in that not many complained about having to "re-train" graduate entrants. Nowadays they do express an opinion, increasingly through an expression of what we, as statistical educators, should be teaching but more positively in which institutions they prefer their graduates to come from.

Courses are becoming more relevant and there are a number of ways this is being achieved, for example through practicals, case studies and industrial placements. These have all been discussed elsewhere and a major challenge to statistical educators is to make then as realistic as possible, that is to reflect the working environment as closely as possible. All aspects need to be included from automatic data collection devices through to expert systems and from individual statistical work to group projects. "Ah, but this all takes time!", comes the response and something must be lost from the existing curriculum. This is perhaps the major challenge we face and my response is that we do away with much of the theory. We put students onto packages straightaway and undertake case studies immediately building up the theory when it is needed. This means careful planning of the course of practicals and case studies so that the necessary theory forms a logical development in itself. We must also leave a lot more of the conventional statistical analysis training to the student themselves - we need only to give them a good grounding. We do not for example need to go through design after design once they have a grasp of the fundamentals

of experimental design - they can find out this by themselves later on (if they need to) using up-to-date information at that time.

This raises the important issue - and another challenge - of the continual updating of the qualified professional statistician, for they cannot afford to relax and think they "know it all" for the rest of their (statistical) working lives. We need to provide what the Institute of Statisticians call Continuing Professional Education (CPE), that is updating "events" in the practice (as opposed to the theory) of statistics. These may be on a range of topics from the latest hardware through to the legal aspects and such events must also be available for updating statistical educators and the "lecturers" at these events (which must be as varied in their presentation as initial courses of training of statisticians) must be primarily working statisticians. An important issue that arises is whether they should, in some sense, be compulsory for we must realise the dangers of statisticians becoming out of date!

Contemplation of the advantages of any programme of CPE and particularly the interaction between education and the workplace leads one to ask the question "Are full-time courses the best way to train professional statisticians?". For nearly forty years the IoS has examined students studying part-time and it may be argued that this is the best way - no need, for example, for too many practicals or case studies for the practising statistician! Education and industry need also to co-operate in discussing the levels at which statisticians are required and in what proportions, for example of sub-degree level and graduate level.

Other Types of Statisticians

We mentioned above the co-operation of industry and education in the training of professional statisticians, but there are two further types of person who require statistical training. These are the specialists in other disciplines who need an ability to undertake their own statistical analyses (the "user" statistician) and those who are consumers of statistics and the results of statistical analyses (the "layman" statistician).

What statistical skills and knowledge are required of the graduate geographer, say, compared to the graduate statistician? The knowledge base needs to be less in total but more specialised in nature and it changes with the technology currently available and being utilised by professional geographers. Basic statistical skills are needs for three major reasons. (i) to be able to undertake routine, usually specialised, statistical tasks, (ii) to be able to recognise when the help of a professional statistician is required and to communicate with him and (iii) to recognize statistical rubbish in the work of others. Clearly a practically based service course drawing scenarios from the major discipline is needed in which the mystic, not to say fear, of statistics is removed.

In the case of users of statistics and statistical analyses, for example the managing director, the advertising manager and so on, it is imperative that they gain (iii) above. They must have statistical appreciation rather than knowledge. Low level courses (statistically speaking) are required,

preferably based in their own workplace. Also, let us not forget the "man in the street" for statistics is impinging more and more on all our daily lives - we must mount courses and events to educate the general public.

Delivery Systems

In addition to providing the statistician with sophisticated tools with which to practice and develop their profession, the recent rapid development in computer technology has also provided educators and trainers with sophisticated tools. These can, indeed must, be utilised in the provision of statistical education and consequently we have available a range of delivery systems.

The greatest challenges in this area arise in the provision of non-full-time training, in particular in CPE, for in this situation we are, providing appropriate material to busy people. Much as we would like to think that CPE is important enough to take time sway from the workplace, this is often not possible. Distance learning has been with us for some tine now and the most appropriate mechanism for CPE is to use such a technique but not primarily via the written word. Use must be made of video, in particular, the interactive video disc, and local work stations so that "living' case studies may be accessed. He can easily develop our ideas of computer assisted learning to CPE and have different versions of the same basic material for different environments.

What about the layman? We need to involve television, the most powerful influence on thinking in most of the developed world. However we do not need initially programmes entitled "Introduction to Statistics" but we need programme makers to be more aware of statistics (and its power in communicating data) and allow it to pervade more programmes and then refer to learning more about statistics via videos, books, etc. Non-peakhour programming is worse than none at all. Also we need to start with the young and educate them about statistics . . . should not every sixteen year old understand the concept of a significance test? We have graduates who do not understand what a rate is!

What about the user-statistician - these would be classified as similar to CPE or to the full-time training given below, but the emphasis must be on appropriate material. What is more likely to prejudice an engineer against statistics than teaching him using medical examples? We now can utilize the technology to prepare several versions of the same material.

Finally, to return to the full-time training of professional statisticians. Others have written eloquently on the use of technology within their course of training but we would wish to emphasize the need for utilising CAL and distance learning-type material (perhaps the term "directed study" is preferable) even within a full-time course of study. In a working environnent we need to seek out information and this is a skill developed by appropriate CAL and directed study material.