Expanding Statistical Education: A New Zealand Retrospect

Geoff Jowett - Outram, New Zealand

1. Introduction

Expanding any kind of education has challenging difficulties, and expanding statistical education a very generous share. Some frustrating attitudes (slightly exaggerated!) have been:

> (general) "I didn't know you could teach statistics - I thought they just were!", "Statistics can prove anything.", "Dreary arithmetic."
> (pure mathematicians) "... a bit of the icing on the top of the mathematics cake, best added (if at all!) at the end."
> (applied mathematicians) "a minor topic worth about as much time as statics, but with an unjustifiably more pretentious name."
> (research workers - notably medical) "My paper has been faulted for lack of statistical analysis. Could you please append some and forward it to the editor.", "This conclusion is obvious - could you please demonstrate it statistically."

Then there has been the nature of the subject itself. The concepts and thought processes are rather subtle, not easy to grasp in the first place and rapidly lost if not well reinforced; once mastered they become obvious, so that the teacher tends to underestimate the students' difficulties. The terminology doesn't help - normal distributions that aren't, error and bias that aren't, goodness-of-fit tests that are tests of badness of fit, insignificant significance, and so on. And the main achievement seems rather negative: stopping people drawing conclusions (which is what they most like to do!).

2. Personal retrospect : industries, universities, agriculture

In the later years of the war, when almost all university teachers in Britain were on war work, I was one of a dozen or so raw mathematics graduates directed into a small section SR17 of the Ministry of Supply. After learning the relevant statistics from books and the handful of more experienced people already there we were sent round ordinance factories to install quality control charts and sampling inspection schemes, and so on. In convincing others that we had a good product to sell we convinced ourselves, and after the war the section exploded like a London bomb into missionary statistical occupations all over the country. I was a piece of shrapnel that landed in Sheffield to help A W Swan to start a statistical section in the United Steel Companies research division. We taught helpers in the factories (and also a visitor who was to try similar methods in India), taught management as much as we could, and gave expository talks to the RSS local group. Swan used to maintain that one could get a long way with little more than an understanding use of the standard error of the mean, and I agree with this. However, non-independence of successive observations can play havoc with the usual formula and its regression counterpart, and this can be far more serious than non-normality, though it gets little if any space in textbooks. It was meeting this in the steel industry that found me a field for research later on.

In 1947, when it was clear that the section was going to survive, I took an opportunity to move to the Mathematics Department at the University of Sheffield, my brief being to initiate appropriate statistics courses, give statistical help to other departments, and do research. It was essentially a one-man job, except that the statistical help (in increasing demand!) was used to justify the appointment of a mathematics graduate as a research assistant, to help with the computing and practical classes. Initially I gave a course in mathematical statistics in the final year of the mathematics honours course, a course in economic statistics for economics honours students, a course in statistical methods for staff and research students (which had some high-level takers), and an extra-mural course attended by all and sundry, including teachers and industrial workers. Later, when the last two of these had met the current need, I was able to reach engineering students by taking on their second year mathematics service course, into which I was allowed to put some statistics. To get the teaching of statistics going I have often had to teach it along with other topics: numerical analysis for the mathematicians, official statistics for the economics students, differential equations for engineering students. This expanded my own education but could make the load rather heavy and when I sought relief it was apt to come in unexpected ways involving some loss of ground. The engineering mathematics was taken over by an applied mathematician with concerns of his own, with a resulting contraction in statistical education. So the moral is that if you want to expand statistical education, follow the example of the politicians and try to stay in power.

In 1956 a small department of statistics was created, staffed by myself as senior-lecturer-in-charge, two research assistants, a computing assistant and a part-time technician, with some teaching help from the departments of Economics and Pharmacology. This meant that statistics was recognised as a subject in its own right, not just a topic in mathematics, and could be taken as a complete stage 1 subject by science students. It was put to me that it should be seen by students as substantial a commitment as the others, and that this could be achieved by matching lecture and laboratory hours with those of physics and chemistry. For a long time I had been promoting direct involvement with the generation of experimental data as an important component in expanding the quality of statistical education, and was able to exploit the matching of laboratory hours as an opportunity to do the involvement job properly. A typical year's laboratory programme was made to include half a dozen or so substantial projects involving interesting apparatus (mostly home-made and very simple but sometimes borrowed from other departments) which illustrated concepts and purposes, went hand in hand with lectures, and were designed to provide statistical experiences in fields of application and motivate the subject. The course went well; the practical work was something of a novelty at the time and is described in Jowett and Davies (1960). Creation of the department received a lot of support from the Faculty of Medicine, who were major users of its consulting services, and the department got a toe into the medical curriculum via three lectures in the Pharmacology course.

By the end of 1959 things were going well; a consulting statistician from another department had moved over to us as a lecturer, and one of the research assistants was ready for promotion. Professorial status for the subject seemed a long way away, however, and I was persuaded by the need for experience of advanced teaching, among other things, to move to the Department of Statistics at the University of Melbourne, where there was a three-year honours course in mathematical statistics and a lot of emphasis on consulting work for local hospitals and industry, as well as plans for new courses in operation research, much of which involves statistics. I was given facilities to introduce practical work of the Sheffield type into the courses I taught, which covered most of them since I stood in for Professor M H Belz during his absence on leave.

In 1964 I moved back into a mathematics department, this time to the new chair of statistics at the University of Otago. It was a bit like starting all over again except that I had more experience and influence, and that there was already a well-established one-year course in statistical mathematics, currently taught by an experienced statistician on the staff of the Department of Preventive Medicine who was an honorary lecturer in statistics. Once again it was a matter of making what progress I could just with a continuation of his help and that of a research assistant who had just graduated in mathematics and wanted to go on in statistics.

The first year went well, with legislation set in place for statistics options in mathematics that would provide an honours degree substantially equivalent to that in Melbourne, a service course Mathematical and Statistical Methods, and a foray into school mathematics to introduce some statistics (to be described later). The following year went less well, since the mathematics side of the department almost collapsed and I found myself holding the fort. In trying to cope, I found that in New Zealand it was not only statisticians that were in short supply but other sorts of mathematicians too. Graduating students usually took their scholarships overseas and often found posts there, and advertisements even for chairs brought few if any replies. Eventually the position in mathematics was righted, and a senior biometrician from the NZ Department of Agriculture came as a lecturer for a few years, being persuaded to return only when the group he had left was threatened with collapse. Staffing in statistics was difficult throughout my time, but even so we managed to educate several mathematical statisticians who are now lecturers or in senior posts of other kinds, and the service course which in my time grew from a dozen or so to 450 students has continued its phenomenal growth, with a substantial follow-on course established by Brian Manly who now occupies the statistics chair.

After a preliminary trial of statistics in the medical curriculum by three lectures contributed to the physiology course, it was expanded into a one-term course on its own. A postgraduate diploma in statistics was constructed by combining the second and third advanced-year papers in mathematical statistics with other appropriate papers, such as computer science and accounting. To make small staff resources go a long way it was often necessary to make one course serve two or more purposes, and this sometimes required very careful structuring of material, sometimes to make it acceptable to quite different types of student (as when the Geography Department arranged to send their students to the course for medicals), sometimes to make it possible to attend courses simultaneously that were normally attended sequentially.

One expansion technique that worked very well was collaboration with other departments in developing statistical methods courses. A big problem with service courses in statistics is that there is sometimes little reinforcement in the user departments of the material given, or if there is it is done in a different way and causes confusion. Partly to help and partly to learn, experienced members of staff from Commerce and Geography came as tutors in practical classes involving their students, and after my departure took over the statistics teaching of their own students.

In 1972 I moved to the public service Department of Agriculture as a biometrician (on the staff of my former lecturer) stationed at Invermay, a research station just over the hill from Dunedin. While nominally the job was to operate as a local consultant and to help scientists to get their results computed, it still had a strong content of expanding statistical education. An important aim was to get scientists interested and competent in the statistical analysis of their own data and improve technicians' understanding of the statistical purpose underlying the structure of the trials which were their everyday work. (Another was to stop them swamping us with requests for analyses that they didn't really want.) Much of the education was person-to-person, but sometimes by seminars or one-week courses. Occasionally these were at other centres (e.g. one for dairy scientists in the North Island, another for a group of top executives in Wellington).

It was during this period that personal computers arrived on the scene, with their tremendous potential for getting statistical ideas across, particularly those connected with sampling distributions and inference. We acquired an Apple computer which had a secondary use with big borrowed monitors as a lecturing aid, allowing on-the-spot statistical demonstrations and analyses in front of the class. Sadly it was thirty-five years too late for most of what I would have liked to have done with it, but at least I did have the experience.

3. Statistics in schools : teaching the teachers

During my time in university work opportunities arose to contribute to the expansion of statistical education in all three locations by becoming involved with public examining. Unfortunately this very influential activity tends to attract opprobrium rather than respect in academic circles, being regarded as a way of making a bit of extra cash that academics (who ought to be devoting their spare energies to research) are allowed to use.

In Sheffield the examinations were run by the Northern Universities' Joint Matriculation Board, and I was recruited just as the old examination structure was being changed to a more flexible General Certificate of Education with a new statistics option in a terminal seventh-form subject Mathematics. There were specimen papers and the first papers to set; the syllabus was unfortunately in place already. There was no stereotype for these new papers and few teachers knew anything about the subject or how it was used. I remembered from my own experience as a pupil how we worked endlessly from old examination papers, and judged that the most helpful way of expanding the interest of both teachers and pupils was to construct questions with an interesting practical background where the bit of statistics being tested was playing a clearly useful and important part. This was very challenging, because it had to be done without causing irrelevant comprehension problems for the candidates, or making the questions look alarming through length or unfamiliar terms. The syllabus was such as to make it difficult to set questions to test ability in statistics as a mathematical discipline (e.g. "Use (without proof) of the formula $\sigma\sqrt{n}$" without resources to use it for differences or other linear combinations of means).

A little later I was given the opportunity to design a syllabus myself, this time for a terminal mathematics paper at fifth-form level. Included in it along with the usual frequency tables and probability were topics such as weighted averages, trends and cycles and regression as a locus of means, all treated in a very elementary but mathematical way. For this subject too, questions were set with an interesting practical background.

In Melbourne statistics was already established as a quarter of the terminal sixth-form subject General Mathematics, and was being set and marked by mathematics staff who felt ill at ease with it, as did the teachers. My offer to take over the examining of the statistics section was accepted with alacrity, and from then on questions about the points of inflexion of the normal curve and permutations and combinations were less in evidence. The number of candidates taking statistics was much larger than in Sheffield, where statistics had been just an option, and the anxiety among teachers correspondingly greater. However, the Melbourne Matriculation Examination was a much more local examination, and I was able to give more help to teachers by visiting schools to give interest-arousing talks to classes, courses with practical projects and so on. The syllabus had bugs in it, such as the correlation coefficient in a void by itself - the only way it could be examined was to ask for it to be calculated from a frequency table - but it was not difficult to get it changed. I began to feel encouraged when teachers told me that the main-stream mathematics pupils were plaintively asking why they weren't doing this interesting stuff.

In Otago school mathematics was in the melting pot when I arrived because of the so-called "new maths", and the NZ Statistical Association was lobbying for the inclusion of some statistics but without any specific proposals. Even more than in Sheffield teachers lacked knowledge of the subject and were very apprehensive about finding themselves having to teach it. Luckily we found a good strategy for introducing it without too much opposition and in such a way as to give them time to cope. There was a subject at sixth-form level called mechanics, originally meant for technical school students who did not do physics, supposed to contain some experimental work, and which was being exploited by some mathematics teachers to get a few extra periods for mathematics teaching, since the syllabus was largely covered in physics which their pupils were taking anyway. We proposed that this subject should be re-named Applied

In Sheffield the examinations were run by the Northern Universities' Joint Matriculation Board, and I was recruited just as the old examination structure was being changed to a more flexible General Certificate of Education with a new statistics option in a terminal seventh-form subject Mathematics. There were specimen papers and the first papers to set; the syllabus was unfortunately in place already. There was no stereotype for these new papers and few teachers knew anything about the subject or how it was used. I remembered from my own experience as a pupil how we worked endlessly from old examination papers, and judged that the most helpful way of expanding the interest of both teachers and pupils was to construct questions with an interesting practical background where the bit of statistics being tested was playing a clearly useful and important part. This was very challenging, because it had to be done without causing irrelevant comprehension problems for the candidates, or making the questions look alarming through length or unfamiliar terms. The syllabus was such as to make it difficult to set questions to test ability in statistics as a mathematical discipline (e.g. "Use (without proof) of the formula $\sigma\sqrt{n}$" without resources to use it for differences or other linear combinations of means).

A little later I was given the opportunity to design a syllabus myself, this time for a terminal mathematics paper at fifth-form level. Included in it along with the usual frequency tables and probability were topics such as weighted averages, trends and cycles and regression as a locus of means, all treated in a very elementary but mathematical way. For this subject too, questions were set with an interesting practical background.

In Melbourne statistics was already established as a quarter of the terminal sixth-form subject General Mathematics, and was being set and marked by mathematics staff who felt ill at ease with it, as did the teachers. My offer to take over the examining of the statistics section was accepted with alacrity, and from then on questions about the points of inflexion of the normal curve and permutations and combinations were less in evidence. The number of candidates taking statistics was much larger than in Sheffield, where statistics had been just an option, and the anxiety among teachers correspondingly greater. However, the Melbourne Matriculation Examination was a much more local examination, and I was able to give more help to teachers by visiting schools to give interest-arousing talks to classes, courses with practical projects and so on. The syllabus had bugs in it, such as the correlation coefficient in a void by itself - the only way it could be examined was to ask for it to be calculated from a frequency table - but it was not difficult to get it changed. I began to feel encouraged when teachers told me that the main-stream mathematics pupils were plaintively asking why they weren't doing this interesting stuff.

In Otago school mathematics was in the melting pot when I arrived because of the so-called "new maths", and the NZ Statistical Association was lobbying for the inclusion of some statistics but without any specific proposals. Even more than in Sheffield teachers lacked knowledge of the subject and were very apprehensive about finding themselves having to teach it. Luckily we found a good strategy for introducing it without too much opposition and in such a way as to give them time to cope. There was a subject at sixth-form level called mechanics, originally meant for technical school students who did not do physics, supposed to contain some experimental work, and which was being exploited by some mathematics teachers to get a few extra periods for mathematics teaching, since the syllabus was largely covered in physics which their pupils were taking anyway. We proposed that this subject should be re-named Applied

Mathematics and consist of statistics and mechanics equally. The statistics could be used in association with the mechanics practical work when appropriate.

I was able to give a completely practical blueprint for the course based on the Melbourne and Sheffield experiences, the technical school teachers liked the idea anyway, and the other mathematics teachers could not decently object since they were not obliged to teach the subject until they were ready. I was asked by the Mathematical Association to give one-week courses for teachers at the main centres. There were a few enthusiastic people willing to help with tutoring, and the first task was to show them a school-level and difficulty-avoiding approach to the subject as a mathematical discipline, and to convince them of how and why it was suitable. Otherwise the teachers, most of whom knew no mathematical statistics at all, would have been bewildered and confused by different stories from different people. I was reminded of the words of the poet:

> "Mine is the higher aim, the longer reach:
> "To teach men how to teach men how to teach!"

Each half-day of a typical course started with a talk presenting a piece of theory, hints on handling difficulties, and an analysis of the structure and marking of a relevant examination question. The course then split into groups to try problems and discuss them with tutors or with me as I circulated. Each evening there was a practical session, with an experiment to reinforce the day's ideas and to raise questions for the next day. The teachers soon began to feel at home with the subject (though they expected to be on top of it in the week!) and much enthusiasm was generated. So it got off to a flying start, and by the time the main-stream mathematics syllabuses were ready for restructuring there was a general willingness for statistics to be included.

Sadly, the satisfactory implementation of this was held up because of the weather. On the day of an every-few-years meeting of the syllabus-reviewing committee my solo statistical voice was absent because all transport to the North Island had been put out of action (in fact the tragic sinking of the ferry *Wahine* occurred on the very day of the meeting, with another member of the committee having to be rescued from the water). The committee made a decision without us - the seventh-form syllabus was to be "the sixth-form syllabus treated more fully and with derivation of formulae" and we were stuck with it, since it was a fundamental policy to let new syllabuses run for a while before changing them again. Fortunately there was a major restructuring of the mathematics syllabuses before too long, and when the job was finally done a lot of effort went into incorporating statistics into the main-stream sixth and seventh-form subjects in what I felt was a very satisfactory way.

Getting a good school statistics syllabus adopted by a group of people who have no experience of statistics at this level, but who are involved because they did a course at university, is a hazardous business. This is particularly so when there are no custom-designed textbooks. A school mathematics subject should be self-contained, not full of "it can be shown that ..." topics and approximations of unverifiable validity, and there should be a clear path through it without blind alleys. The people who had done a mathematical statistics course tended to think that topics had to be handled in the far-too-advanced-for-pupils way that they themselves had experienced without realising that a lot of the difficulties could be avoided by handling topics in a different way, and those who had done the statistical methods courses were inclined to campaign unknowingly for approaches that bristled with difficulties, or to throw out something vital because it involved an unfamiliar word. For example, significance in terms of standard errors rather than precise probabilities gets rid of a whole host of difficulties, and regression as a property of means does not inevitably require a choice between quoted formulae and impracticable calculus.

My other involvement with school children was giving a course for "specially gifted" pre-high-school children at the request of the senior inspector of mathematics. This was fun for both sides. We blew up a lot of balloons with bicycle pumps till they popped, counting the pump strokes, then did a little operational research job that involved a cumulative frequency graph, a little Fortran program, and a little consumer preference survey to find the number of pump strokes giving the most appealing compromise between the equal-cost alternatives of a small soggy balloon with several sweets and a large tense balloon with no sweets.

4. Practical experimentation : a shared experience

This has always been a bee in my bonnet since my first lectures to mathematics students, when it became clear to me that I had to take steps to give them some inner feeling for the nature of statistical variation if I was ever to generate any interest in the problems statistics attempts to solve. Tossing coins and throwing dice only goes a little way towards this, and what was really needed was some cheap, simple and rapid data-generating system, suitable in its simplest form for a first (e.g. homework) assignment, that could be developed in step with the teaching so as to provide a shared experience to which reference could be made.

I made and used various devices, and ultimately settled on the "shove-halfpenny" experiment which I have used ever since. The apparatus is simply a 10cm square of hardboard notched to hold a thin elastic band round the middle which is used to shoot a coin along the table. In its first application the band is pulled back to a mark and the variate is the distance travelled by the coin, giving a distribution which is single-humped and positively skew enough not to look too exactly like a normal distribution. It can be used directly or in modified form to raise questions about the cumulative graph, occurrence of discrete events, censoring of outliers, serial independence, regression (with a scale instead of a mark), ANOVA (different operators, bands, and so on), predictability by mechanics and a host of other statistical matters. There are two drawbacks: one is that the situation may be regarded as trivial by the unimaginative, the other that questions may be raised that fox the teacher with no "right answer" in the back of a book. To stimulate the imagination other apparatus can be brought in later, preferably simulating something obviously important in real life.

The most ambitious experiment that I have used was a "model factory" where the class was organised in a production line cutting up bits of wire ("rodlets") to an adjustable length using edge-cutting pliers mounted in a block, making up matchbox lots, and running control charts, acceptance sampling schemes and evolutionary operation experiments in as lifelike a way as could be organised. This was the climax of a course for students, many of whom had never set foot in an engineering factory.

Sometimes it has been suggested to me that practical experiments along these lines are not necessary for teaching people who have day-to-day involvement with data of their own which can be used in illustration. However, other non-statistical matters are apt to be distracting. For example, in teaching dairy executives I found that butter moistures were too inextricably tangled with legal considerations to keep red herrings out of the thinking, but amazingly enough the class came to life statistically when we did an experiment counting dunks-to-disintegration of ginger biscuits in cold or hot tea with or without milk; likewise the dairy scientists were happy to get deeply involved in studying the statistical properties of a somewhat bizarre test of cheese by dropping darts into it which paralleled their own test but which they knew no more about than I did.

Experimental work of this kind is apt to be just as valuable for the teachers as for the taught, particularly in expanding the experience of raw graduate tutors.

5. Presentation

The nature of the group, their experience, their mathematical confidence, what they want to know, what they need to know, how to help them to grasp an idea, to remember it, to use it, how much time you have with them, how much you can expect of them in concentration and follow-up; all these affect presentation, and all I can do is to offer a few sample experiences.

In presentation to senior mathematics students it helped to bring in other branches of mathematics. Multivariate analysis, for example, was made much more intuitive by relating it to n-dimensional coordinate geometry with oblique axes. When possible long stretches of algebraic manipulation were broken into small parts with interesting interpretations at each stage; this was done with maximum likelihood theory, where log-likelihood variates have properties which are interesting in themselves.

In the three-lecture statistics contribution to the medical curriculum the problem was to make something in this tiny drop in the ocean stick. Medicals traditionally use mnemonics to help them remember anatomy, so for statistical significance, along with due explanation, I offered them:

(1) Minority frequencies can probably be trusted to within plus or minus twice their square roots.
(2) No need for terror of the standard error. Take sigma and then divide by root n.

My hope was that they would apply (1) to cures or non-cures of their patients; I certainly find it useful myself for a first appraisal. They were given a bizarre analogy to make the explanation of (2) stick: an invisible man and his visible dog. The dog is the sample mean, the man the true mean. They are connected by an elastic lead, one standard error in unstretched length. The lead is more likely than not to be loose, but the dog can readily stretch it to up to twice its length, and by nearly choking itself up to as much as three times its length. I was so pleased with these that I have used them since in other situations, and even wrote a song "The mean mean" around (2), in desperation to make the standard error and its interpretation stick.

Requests for isolated lectures to societies, management courses, or to groups such as Rotary, school classes, parents and so on have been common throughout my teaching life, and success can be measured roughly by whether or not one is asked for more. One hopes that they will lead to people enquiring further, or encouraging acquaintances or families to do so when the occasion arises. It is advisable to be entertaining in putting the message over, and operational research applications can be very useful for this purpose; for example an extract from a simulated game of cricket with spinners:

> "... Spin the bowler: a leg break. Spin the batsman: attack towards first slip (if you can call it that!). Spin first slip: caught ..."

For Rotary it seemed appropriate to go round the occupations with a statistical application in each: the bishop was much amused by the idea of a Latin Square experiment to maximise the collection with factors: sermon (long/medium/short), hymns (English Hymnal/Moody and Sankey/Ancient and Modern) and so on.

One public lecture consisted of showing how statistics could be used at different stages in producing a can of baked beans, e.g. agricultural yield trials, tenderness tests, canning faults.

An analogy that I have often found useful in discussing the rationale of statistical tests of significance is a criminal trial in a court of law; most people have watched enough of these on television to have some idea of the principles. The null hypothesis is that the defendant is innocent; if it were believed to be true he wouldn't be there, but the onus is on the evidence to prove that it is false. Errors of the first and second kinds have obvious parallels, and some tests/lawyers are more powerful than others. And so on; there are differences too.

Acknowledgements

Along the way many people have helped me, too many to mention by name. In New Zealand outstanding help came from Mr H S Roberts (DSIR) in the schools area and from Mr T G Robertson (Otago and MAF) and Mr G F Spears (Otago) in the tertiary area.

Reference

Jowett, G H and Davies, H M (1960) Practical experimentation as a teaching method in statistics. *J.R.Statist.Soc.* *A* **123**, 10-27.