

Satisfying the Need for Applied Statisticians

Ken Sharpe - Melbourne, Australia

1. Introduction

During the 60s and 70s students were encouraged to become statisticians because of the intrinsic interest of the discipline and because of the predicted future need for statisticians. Many people who appreciated the wide applicability of statistics felt that it was only a matter of time before employers would be crying out for statisticians and that newspapers would be full of advertisements for interesting and well-paid jobs for statisticians. The ensuing campaign to attract students was reasonably successful and the number of students majoring in statistics showed slow but steady growth. Unfortunately, the demand for statisticians did not grow at the predicted rate. One of the reasons for this was that the people making the employment decisions tended to have little understanding of statistics and would be far more inclined to employ an engineer or an economist, even when the job was primarily statistical in nature. For a number of years I have given a course, "Statistics for Research Workers", and have been disturbed by the number of people attending this introductory course who have been working as statisticians.

The last decade has seen a dramatic change due, primarily, to the computer. Many more people now have the opportunity and the need to work with data, and the ability to readily carry out, though not necessarily understand, quite complex statistical procedures. There has also been a substantial increase in the statistical training of people for a wide range of professions so that there is, overall, a much greater awareness of the need for and the benefits to be gained from statistical expertise. As a consequence, the number of jobs being advertised which specify statistical expertise has grown to the point where there is now a recognised shortage of well trained statisticians.

At the same time that computers were starting to affect the practice of statistics and increase the demand for statisticians, computer science was growing in popularity and there was a corresponding drop in the number of students majoring in statistics. In this age of "economic realism" students are attracted to those disciplines which are likely to result in the highest income so that, at the University of Melbourne, it is considerably easier to get into science than into economics and of those students who come to university with an interest in the "Mathematical Sciences", by far the majority

start out with the intention of majoring in computer science.

During the 60s and 70s, students taking Honours courses in statistics at the University of Melbourne were required to take half mathematics and half statistics in their Honours year. The mathematics courses covered topics such as complex analysis, measure theory, topology and abstract algebra whilst the statistics courses also tended to be quite theoretical with a heavy emphasis on probability theory and stochastic processes. The effect of this was to discourage many students, who were quite competent with the applied aspects of statistics but either struggled with or disliked the more abstract theoretical aspects.

By the early 80s Honours courses were starting to change to reflect the students' interests and in 1984 Melbourne, Monash and LaTrobe Universities started to explore the possibility of mounting a joint MSc by Coursework programme in order to try to satisfy the need for applied statisticians. In next to no time (18 months or so) all of the administrative difficulties had been sorted out and the programme started in 1986. By this time the number of participating institutions had grown to four (RMIT, The Royal Melbourne Institute of Technology, was included) and the programme was made available to both Honours and Masters students.

2. The MSc by Coursework programme

In setting up the programme a conscious decision was made to emphasise applied statistics and encourage students with more theoretical interests to pursue their postgraduate studies through the usual MSc by research and PhD channels. The programme has remained reasonably stable over the five years it has been operating and consists, this year, of the following components:

CF	Computer Familiarity
CAS	Consulting and Applied Statistics
ACD	Analysis of Categorical Data
ADS	Actuarial and Demographic Statistics
AFSD	Analysis of Failure and Survival Data
AMD	Analysis of Medical Data
APM	Applied Probability Models
APT	Advanced Probability Theory
DAE	Design and Analysis of Experiments
DAS	Design and Analysis of Surveys
DFI	Distribution Free Inference
EDA	Exploratory Data Analysis
F	Forecasting
FSI	Foundations of Statistical Inference
MVA	Multivariate Analysis
OR	Operations Research
PDT	Probability and Distribution Theory
RA	Regression Analysis
SC	Statistical Computing
SI	Statistical Inference
SQPI	Statistics for Quality and Productivity in Industry

SP	Stochastic Processes
TSA	Time Series Analysis

Most, but not all, of these components are very applied in nature and some, such as ADS, AFSD, AMD and SQPI, relate to specific subject areas. Many institutions offer introductory statistics courses for engineers, economists, etc., as well as advanced, context non-specific, courses for statisticians; what is often lacking is a synthesis of these two types of courses to train statisticians in the specialised techniques that are required in particular disciplines.

Course structure: The components, other than CAS, can be taken as part of an Honours or an MSc degree. The requirements for Honours differs slightly between institutions but consists of (about) eight coursework components, CF, and a reading project. For the MSc, the requirements are five coursework components, CAS, and a minor thesis, plus CF if not previously taken.

Each coursework component consists of one two-hour session per week for 13 weeks. All components given by an institution are given on the one day so as to minimise the amount of travelling required of students and each institution offers sessions in the 4-6 and/or the 6-8 timeslot in order to make the programme more accessible to part-time students.

Course content: The CF component is given over two weeks prior to the start of the other components and provides an introduction to the packages MINITAB and GLIM and to FORTRAN. The purpose of this component is simply to bring all students up to minimal level of competence in computing so that they can reasonably be expected to cope with the computing demands of the other components. Many of the students have adequate prior experience with computers but are required to do the course anyway as a means of refreshing their knowledge after the summer break.

The CAS component is a cornerstone of the MSc programme and constitutes 25% of the work. Part of CAS is presented as formal classes in one institution; the rest consists of actual "consulting" and is carried out in the student's "home institution".

For the remaining components, the syllabuses were developed in 1985 by small teams of staff, generally one lecturer from each of the four institutions. Since then the lecturer presenting a component has had a reasonably free hand so that the actual course content has tended to vary, sometimes quite substantially, from year to year. As might be expected, some components are more popular among staff than others, and there is an informal agreement that a lecturer has the right to give a component for three consecutive years but that, after three years, it is desirable that someone else give the component, preferably someone from a different institution.

3. Teaching

The range of teaching approaches used is very wide and reflects the attitudes and interests of the large number of staff involved. There is, however, a noticeable emphasis on computer use and on real applications and, for many of the components, students are required to undertake a substantive practical project. Some lecturers invite outside "experts" to come and give special lectures.

The one course that stands out as being the most innovative and the one I wish to comment on in more detail is CAS. The formal part of CAS has been presented by

people with considerable consulting experience and has attempted to cover a wide range of areas not usually encountered in statistics courses. These include: report writing; consultant-client relationships; consulting ethics; working in groups; library use; critique of published papers.

The other aspect of the component is the actual consulting. The aim here is to provide students with a mini-apprenticeship. The provision of meaningful consulting experiences depends critically upon the type and extent of the statistical consulting actually carried out in the institution. At the University of Melbourne we are especially well placed as the department has a Statistical Consulting Centre which started in 1984, and is operated on a commercial fee-for-service basis. The Centre has, at present, the equivalent of nine full-time staff. The other institutions also do a reasonable amount of consulting and are in various stages of trying to establish similar, though not as extensive, consulting groups as we have at the University of Melbourne.

Ideally, training in consulting should proceed through three distinct stages, namely: observing experienced consultants at work; working for an experienced consultant; working with an experienced consultant.

At the University of Melbourne we try, with varying degrees of success, to take students through each of these stages. The first stage is addressed by having students attend the free lunch-time consulting service offered by the Statistical Consulting Centre. Sessions are held once a week with two consultants "on duty" and provide for up to four people to have a 3/4 hour meeting with a consultant. These sessions were introduced primarily as a way of dealing with (internal) jobs which were too short to warrant charging for, and to cater for those potential clients who wanted to discuss their project with a statistician before making a commitment to seek paid assistance.

CAS students are required to sit in on a number of these lunch-time sessions; this enables them to come into contact with a wide range of projects and to see how different consultants approach the task. One thing they pick up very quickly is an appreciation of the fact that what a client wants is not necessarily what they need. At first students are expected to be passive observers but, as the year proceeds, and provided the consultant considers the situation to be appropriate, students are expected to take a more active role.

The second and to some extent the third stages are addressed by having students do an amount of actual consulting working under the direction of an experienced consultant. This tends to be a high risk activity; some students can make good progress on some projects with minimal supervision whereas others need to be helped every step of the way. The preference is to have students work on current projects, however suitable projects are not always available. Also, students need to fit in this work around their other commitments which can create difficulties with the scheduling of meetings with clients and with meeting deadlines so that it has, at times, been necessary to have students work on projects which have already been completed. A few students have worked directly for people from outside the department and then simply discussed their work with a member of staff. This approach can pose problems but has proven to be useful in catering for part-time students, some of whom have been able to combine their consulting experiences with their normal jobs.

A side benefit of having students work on consulting projects is that it has, on occasions, enabled the Consulting Centre to provide consulting for clients, usually postgraduate students from other departments, who are short of funds. In these cases the

CAS students are required to do most of the work and the client is only charged for the supervision time.

Where possible, students are encouraged to work in pairs; this not only reduces the amount of supervision required but enables the students to try out their ideas on someone other than their supervisor. Finally there is the question of how much consulting the students should be expected to do. The fairest approach seems to be to expect them to put in a certain number of hours and to require them to keep a diary of what they have done; we expect students to do about 100 hours.

It is recognised that this amount of training only scratches the surface of what is required to train applied statisticians and, on completion of the programme, students are encouraged to seek jobs where they can work with experienced statisticians. To this end the Statistical Consulting Centre has a policy of employing recent graduates and currently is employing, either part-time or full-time, six people who have been through the programme.

4. Advantages and disadvantages

Advantages are that the programme has:

- (i) enabled the institutions to offer an MSc by coursework programme. None of the institutions is really large enough to be able to offer a viable programme on its own.
- (ii) enabled the institutions to offer a very wide range of components so that, even at Honours level, students have a lot more choice than previously was the case;
- (iii) resulted in larger class sizes. Some components have attracted more than 20 students whereas, prior to 1986, some Honours courses were given to as few as a single student. Educationally it is highly desirable that students have other students to work with, in addition to which departments are funded, at least in part, according to the number of students taking its courses, so that larger classes produce both educational and financial benefits.
- (iv) enabled students, for the first time in Melbourne, to study part-time at this level;
- (v) decreased disruption by staff taking leave, or leaving, since there is often more than one person who is willing and able to present a particular component;
- (vi) encouraged students to stay on beyond third year;
- (vii) provided students with a wider range of educational experiences than would otherwise be the case;
- (viii) fostered greater interaction between the four institutions;
- (ix) provided a stimulus for consulting.

Among the disadvantages of the programme are that:

- (i) there is much greater diversity in students' backgrounds than would have been the case had they all taken their undergraduate studies at the same institution. The consequence of this is that some courses repeat material which some students have covered in their undergraduate courses.
- (ii) students to travel to different institutions which are up to 20km apart;
- (iii) courses often need to be concentrated into one two-hour session per week;

- (iv) some students have had difficulty consulting some of their lecturers, especially those from institutions other than their home institution;
- (v) there is a need to coordinate the programme across four institutions.

5. Results

Table 1 gives the numbers of students who have taken at least some components and the numbers of students who actually completed their Honours or Masters degrees. The reasons for the large differences between the numbers of students taking components and the numbers completing degrees are that many of the part-time students have been counted in more than one year, there has been a moderate drop-out rate among part-time students and there have been some students who have taken just one or two components either out of interest or as part of some other programme. We have been very pleased with the total numbers of students but somewhat disappointed that more students have not gone on to complete the entire programme.

TABLE 1
Student enrolments and numbers completed

Year	Number enrolled	Number completed	
		Honours	MSc*
1986	35	16	
1987	39	14	3
1988	48	25	9
1989	52	13	9
1990	61		

* includes the minor thesis which invariably delays completion beyond the year in which the last coursework component has been completed.

To the best of my knowledge, all of the students who have completed the MSc by Coursework programme, and many who left after the Honours year, are currently working in statistical jobs; I am aware of some who have been asked to apply for positions and others who have been offered positions which have not been advertised. This is in stark contrast with our BSc graduates, many if not most of whom end up in non-statistical jobs.

In conclusion, combining to offer the joint MSc by Coursework programme, the four participating institutions have, for the first time in Melbourne, been able to offer something approaching an adequate training in applied statistics. To date the programme has been quite successful in terms of student numbers and encouraging students to pursue a career in statistics. However, it is not without its share of problems due mainly to having to coordinate the programme across four institutions.

It is likely to be a long time before we satisfy the need for applied statisticians, but we have at least made a start.