# STATISTICAL INVESTIGATIONS – DRAWING IT ALL TOGETHER ®

Flavia Jolliffe
University of Greenwich
UK

*The author's approach to teaching an integrative unit to a small group of master's level Applied Statistics students in 2000-2001 is described. Details of the various activities such as data analysis, reading and discussion of papers, and training in consultancy skills are given, as also are details of the assessment. The students' and lecturer's views of the unit are discussed.*

## INTRODUCTION

Statisticians work in many different application areas and need to interact with people whose expertise and backgrounds are different from their own. In educating future statisticians it is therefore important to expose students to a wide selection of real problems and to give them experience of a range of communication skills. The unit *Statistical Investigations* discussed in this paper was designed to equip students for work as a practical statistician. It was taught by the author in 2000-2001 when a variety of teaching methods was used in line with recommendations to change from teaching statistics in a traditional lecture format to making more use of problem-based and active learning methods, and placing more emphasis on data analysis (for example see Boyle, 1999; Moore, 1997; Snee, 1993).

A background to the unit is given in the second section, a description of the syllabus experienced by the students when taught by the author in the third, and the assessment is outlined in the fourth. Evaluations made by the students are discussed in the fifth section, and in the final section the author reflects on the unit and makes some recommendations.

## BACKGROUND

The unit *Statistical Investigations* formed part of the MSc degree in Applied Statistics which was given at the University of Greenwich in the UK between 1995 and 2001. This degree was designed to convert graduates in other disciplines into statisticians and evolved from an intensive two-term course given to three cohorts of trainee UK government statisticians from 1995 to 1998 (see Jolliffe, 2001, for more details). It has been run in both part-time and full-time modes from the 1997-8 academic year.

The degree was re-validated in 2001 in a different form - the conversion aspects (introductory units assuming no prior knowledge of statistics) were dropped and replaced by more advanced material – but the *Statistical Investigations* unit has remained. This unit is assessed entirely by course-work and accounts for an eighth of the taught part of the degree. There was no entry to the degree in either form in the 2001-2002 academic year. The *Statistical Investigations* unit aims to integrate the contents of the other units in the degree and to show how they relate to professional practice. The aims, as stated in the latest specification, are:

1. To provide students with knowledge, experience and understanding of the techniques and good practice of data display and interpretation.
2. To give students experience of a range of communication skills in a statistical consultancy context.
3. To enable the student to understand and apply a range of problem solving techniques in practical contexts.
4. To provide the student with training for the conduct of a MSc project.

The author's additional aim was to give the students an overview of statistics and to help them see statistics as a whole rather than as sets of techniques packaged in short and largely independent modules. The indicative content of the unit when given by the author was "Integrative case studies; descriptive statistics; EDA; simulation and Monte Carlo methods; project design, management and reporting". At one time the intention was to revise, and if necessary teach, descriptive statistics in *Statistical Investigations* rather than in other units, and this was the case in some years, but this only makes sense when the unit is given at the start of the

degree. However, use of descriptive statistics is an integral part of data display and interpretation and therefore underpins much of the unit.

Ideally *Statistical Investigations* would perhaps be taught concurrently with other units throughout the student's time on the degree course. In practice time-tabling constraints, in particular the fact that in many years there were both full-time and part-time students who were taught together and the latter attended for only two evenings or one day each week (depending on the year), prevented this. Its position in the course was varied over the years, but by its nature its content is restricted by the rate at which the other units are delivered. When it was taught entirely in semester 1 it was not possible to draw on methods and techniques taught in semester 2, and new students with no statistical background found the unit difficult. The sequencing required by some of the other units prevented it being taught in its entirety in semester 2, so in both 1999-2000 and in 2000-2001 the unit was time-tabled as two hours a week in semester 1 and one hour a week in semester 2. In 1999-2000 the unit was taken by full-time students only. The colleague who taught it started with an introduction to computing and word processing, and was initially very constrained as to what could be covered as some students had little statistical background when they enrolled on the degree. In contrast, in 2000-2001 it was taken only by three second year part-time students who had already taken half of the MSc taught programme at the start of the unit, and who had more or less completed the taught programme by the end of the unit.

## THE EXPERIENCED SYLLABUS

In 2000-2001 weeks 1 to 13 of *Statistical Investigations* were 2-hour sessions and weeks 14-24 were 1-hour sessions. Although mainly an experiential unit in which students were given some responsibility for their own learning, some lectures to extend their knowledge and awareness of the range of tools available in a statistician's work-box were given. Some weeks the students were set work to do on their own, such as reading or activities in the computing laboratory. The type of activity varied from week to week as shown in Figure 1 which also indicates the proportion of time spent on each type. In weeks 2, 8 and 21 more than one type of activity was involved.

| Activity | Weeks | Approximate % of time |
|---|---|---|
| Lectures on statistical topics | 3,4,6,9,11,17,23 | 32 |
| Case studies/data analysis | 7,8,10,13,18,24 | 24 |
| Reading/discussion | 2,8,14,16,20,21 | 16 |
| Investigation of Web resources | 1,12,15 | 12 |
| Consultancy | 5,21,22 | 10 |
| Presentations | 2,19 | 5 |

*Figure 1.* Summary of the Unit by Type of Activity

With such a very small group of students the approach taken was very informal, but even so, in some sessions the lecturer took the lead and the students were relatively passive. This was so particularly in the weeks when there were lectures, although the students were encouraged to participate. Two of the lectures were on general issues in data analysis such as initial data analysis (IDA) (Chatfield, 1995, Chapter 6), EDA, editing and transformations. Another two lectures, supplemented by notes written by a colleague for psychology students, gave an overview of the design of experiments, a topic which was not covered in other units. One lecture on each of simulation, observational studies, and econometrics, were given with the idea of making students aware of the scope of these areas.

Case studies and data analysis were an important part of the course and many of the assessments (see section on assessment) involved analysis of data. In week 7 the students were asked to analyse data on a random sample of two hundred married men and their wives drawn from a study of the heights and weights of the adult population of Great Britain in 1980 (Marsh, 1988). Nine pieces of information are given for each couple. The students were set five research questions to investigate, and their analysis and findings were discussed in class the following week. We discussed three published studies. These were concerned with the loss of sales in a store due to a fire (Manly, 1992, pp. 120-131), comparison of athletic performances (Grubb,

1998), and sensory evaluation of cooked pork (Avery & Masters, 1999) in weeks 13, 18, and 24 respectively. Data were published for the first of these, and data for the third were made available by one of the authors and the students were encouraged to analyse the data themselves in their own time. The papers were chosen to be within reach of the students' statistical stage at the time, readily understandable, with mention of some techniques new to the students.

The students were set several papers or selected passages from books to read and some weeks were set aside for discussion of reading or of data analysis. For example, in week 19 they were given a selection of photocopied material relating to inference issues to read in week 20 and were questioned about it in week 21 to focus the discussion. In week 21 we also discussed what qualities are needed for consultancy to be effective. Some of the issues involved in undertaking an MSc project were covered at the start of semester 2, and this was followed by a coursework (see assessment section). Towards the end of the unit we discussed in class their own proposals of project topics.

Three weeks were devoted to investigation of Web resources. In the first week of the unit the students looked at the American Statistical Association (ASA) publications on social surveys (http://www.stat.ncsu.edu/info/srms.html) and the Centre for Applied Social Surveys (CASS) question bank (http://qb.soc.surrey.ac.uk). In the second week each student made a short presentation on one of the ASA booklets. Towards the end of semester 1 they were given some references to web resources for learning statistics, and early in semester 2 references to sources of data on the web. We did not meet formally these weeks, freeing time for their explorations.

1. A dental researcher
*I have some gingivitis index scores for a number of patients who are on different treatments and I want to compare the effectiveness of the different treatments. I am going to find the mean GI scores per tooth per patient and do t-tests between every pair of treatments. I just want your agreement that this is the right approach so that I can quote your name in a funding application which has to be in tomorrow.*

2. A medical researcher
*I have been studying a group of patients. This study has ended but some patients have agreed to take part in a small extension of the study. I'd like to test with a 5% error rate that patients in the extension are comparable to all patients in the earlier study. I need to confirm that patients in the extension were the same as I am giving a paper at a conference tomorrow in which I say this. The paper has already been published in the conference proceedings.*

3. A researcher in forestry
*I have measurements on a number of trees which have been cut down, and have used SPSS to fit a linear regression. I want to use this equation to predict the biomass of trees from simple measurements that can be taken without cutting the trees down as I I got a high value of $R^2$. Are there any statistical problems in doing this?*

*Figure 2***.** The Consultation Exercises

Preparation for acting as consultants was an important part of the unit, and several different kinds of activities focussed on consultancy. A scientist in Ethiopia had asked me to comment on her notes on sample surveys. With her permission this became a group activity early in the unit and I prepared a response in conjunction with the students. One of my colleagues ran consultancy sessions for final year psychology students needing statistical help with their projects and she allowed my students to sit in as observers one week. We had hoped that when they were more experienced the MSc students would act as consultants to the psychology students, but time-table constraints prevented this. However, I prepared three consultancy exercises (one for each student) based on questions asked by researchers on Allstat (a UK based world-wide email broadcast system for the statistical community). The students worked in three different pairs, with each pair watched by the non-participating student and myself. One in each pair took the role of the statistician and the other the role of consultee. The consultee was given a few lines, as shown

in Figure 2, to give in response to the statistician's initial question as to how (s)he might help, plus some further information, not shown here, to help in answering the statistician's questions. They were told that this was the initial consultation, and that they had not brought any details of the project with them, but should use their imagination to answer any questions they felt the researcher should be able to answer immediately. They had five minutes to prepare for acting as a consultee, and ten minutes for the consultation. We finished by discussing the exercise and how consultees sometimes have unreasonable expectations of the statistician. They were reminded of references, for example Chatfield (1995), giving further comments on consulting.

ASSESSMENT

The unit specification stated that the course-work was to consist of project based course-works and presentations, and two two-hour open book tests in a computer laboratory requiring students to "demonstrate their ability to work through the whole consultancy process from problem definition to final report". There was no formal examination. All the learning outcomes shown in Figure 3 had to be assessed. At the start of the unit the author and students discussed the weighting to give to the tests, the number and weighting of other pieces of course-work, and the approximate dates of all assessments. We agreed that there would be a test worth 20% at the end of each semester, and six other pieces of work each worth 10%, with four in semester 1 and two in semester 2. Students were given three weeks to work on each take-home assignment. Any data sets needed were made available to the students in electronic form.

At the end of the course the student will be able to:

| | |
|---|---|
| A | define problems and select and apply appropriate problem solving techniques, |
| B | define an approach to an analysis of data presented by a client and apply appropriate techniques to illuminate and resolve related questions, |
| C | prepare and present oral and written reports on statistical analyses and interpretations, |
| D | use mathematical and computational tools as an aid to the problem analysis/solving and communication process. |

*Figure 3***.** Learning Outcomes

In the first coursework the students were asked to design a short questionnaire about lifestyle and health, including questions about self-esteem, blood pressure, quality of life, and risk taking behaviour. They were asked to incorporate questions from the CASS question bank (see experienced syllabus section) and to comment on whether or not they had found this a useful resource. The second coursework was based on a paper on comparison of students' marks across subjects (Manly, 1988) The methods used had been partly discussed in class. The students were asked to reproduce some of the results and to describe the computer procedures they had used to enable someone else with little experience of these to get the same results. They were also asked to read and summarise a paper referenced by Manly.

Coursework 3 involved analysis of a data set on crime in the USA (Hand et al., 1994, pp 101-3). There were observations on 14 variables for 47 states, and the students were asked to investigate how the crime rate depended on the other variables. This coursework was to help students prepare for the first test, and they were given extended typed comments on the investigation as well as feed-back on their individual attempts. Coursework 4 involved investigating simulation facilities in Excel and Minitab, and using simulation methods to investigate approximation methods for finding confidence limits for an unknown proportion.

The fifth coursework took the students through the initial thinking of a project. I selected three topics where I knew there was plenty of scope for investigation and which would appeal to individual students, and as hoped they all took a different topic. They had to formulate and state objectives, perform literature and data searches, outline a method suitable for collecting data relevant to their objectives and the way they would analyse these data, and give an oral presentation in class on their thinking as well as write it up. The last coursework was a further data analysis exercise. The data set contained information on 186 physicians in Tampa Bay and there were ten variables, six of them categorical (Sincich, 1993, Appendix C).

In both two-hour open book tests in a computer laboratory the students were given a data set from Hand et al. (1994). In the first the data consisted of observations on the average outside temperature and of gas consumption at a house in the UK for 26 weeks before and 30 weeks after the installation of cavity-wall insulation (p. 69) and students were given three questions to examine. In the second test the data consisted of the weights of young girls before and after receiving one of three treatments for anorexia (p.229) and the students were asked to compare the methods of treatment. In both tests they were asked to hand in a general summary of their findings supported by results of their statistical analyses. The two tests assessed all four learning outcomes (see Figure 3), the first course-work was considered to assess outcomes A and B, the second C and D, the third, fourth and sixth B, C and D, and the fifth A, B and C.

## STUDENT EVALUATIONS

At the end of the unit the students were asked to complete a questionnaire. This started with a reminder of the aims of the unit and the content of each session. After some open-ended questions about the unit as a whole, for each type of activity (see Figure 1) they were asked three questions – "How easy or difficult did you find these (sessions)?", "How interesting did you find them?" and "How useful did you find them?". Questions were asked in the form of a 5-point scale going from "Very difficult (1)" to "Very easy (5)", "Very interesting (1)" to "Not at all interesting (5)", and "Very useful (1)" to "Not at all useful (5)" respectively. The same three questions were asked about each of the take-home course-works. The students were also asked if they had any comments about specific activities or course-works, and approximately how many hours they had spent on each course-work.

As there were only three students I knew which questionnaire came from which student, and with such a small number, one of whom had missed several weeks at the start of semester 2 and made few comments, tabulations of the results are not sensible. They all found the unit demanding in terms of time and intellectual effort, particularly in comparison with their other units where, to quote one of them, "you know which method you are to use". One student felt that aim 2 (see the section on background) had not been met and one that aim 4 had not been met. There was a feeling that too many topics had been covered, and a suggestion that more sessions should have had follow-up discussion sessions. On the whole they would have liked more on data analysis and fewer theoretical lectures, although one of them thought that more time should have been devoted to simulation. One felt that some of the papers looked at were rather old.

The ratings and comments made by the students tend to reflect their backgrounds. For example, two were very familiar with using web resources so found this activity easy and not particularly useful or interesting, whereas the third student found this very interesting and very useful and rated it 3 on difficulty. One student with good presentation skills thought that more time should have been spent on presentation but found this activity easier and less interesting than the other two students and commented that tips on improving presentations would have been helpful. This student also found the consultancy very easy, possibly because of work experience, whereas the others found it more difficult. On most of the other activities the ratings of all three students were very similar. They found activities difficult but interesting and useful.

On the whole the three students also rated the six course-works in a similar way to one another, giving ratings towards the very difficult, very interesting, and very useful ends of the scales. Two students gave almost identical ratings, but the third, who also tended to be different in ratings of activities, gave ratings more towards the easy, not interesting, and not useful ends of the scales. The two students who estimated how long each piece of course-work had taken gave figures greatly in excess of the time given as a guideline (lecturers are required to give students estimates of how long each course-work should take an average student to do).

## REFLECTION AND RECOMMENDATIONS

The unit was indeed packed and as it was the first time the author had taught it a lot of time was spent in planning and preparation. Originally my intention was to give the students more time working on their own, and more effort was made to do this in semester 2. It was too ambitious to try to cover so many extra topics. Finding suitable case studies was not easy.

Theoretical papers were not suitable as the degree is in applied, not mathematical, statistics. Searching for statistical papers in journals in application areas might have been more fruitful.

In retrospect there probably was too much assessed work. In weighting the unit is equivalent to a one-semester unit assessed by a two-hour formal examination counting for 80% of the assessment and either one or two course-works. Marking was no burden with only three students, but finding suitable tasks and preparing questions and mark schemes was very time consuming. As the assignments were open-ended it was hard to estimate the time they would take and these students were of high standard and aiming to do well. The comment that they found this unit more demanding than the others could mean that the other units did not stretch them enough.

There was not time for detailed discussion of tasks the students were left to do on their own, such as reading or computer work. In fact even when the plan was to have discussion the session was only partially successful as on several occasions at least one student was absent for some or all of the session, and those present had not always done the work. In their defence they were all in full-time employment taking the degree as day-release, and sometimes work pressures conflicted with study. This would have been less of a problem with a larger group.

It was encouraging that the students had found most of what we did useful. The real evaluation of the unit is whether they realise its value in five or ten years time. It would have been a worry if they had all found much of it easy as this could mean that they were not learning much. Some activities worked better than others with these students, but there were none which were not relevant to the aims and learning outcomes of the unit. Manly (1992), Hand et al. (1994), and Chatfield (1995) are all useful sources of material. If I were to teach the unit again I would cut down on the lectures and have more sessions on data analysis and case studies plus more discussion. Fruitful discussion does depend on student participation, however. I would have less time in the classroom thus giving the students more time to work on assessment tasks, which are a learning experience for the students. Such units are well worth teaching, but some of the elements of this unit could usefully be incorporated into many units taught in a more traditional manner. It is the drawing together of material from different units which makes this unit unique.

REFERENCES

Avery, P.J., & Masters, G.A. (1999). Advice on the design and analysis of sensory evaluation experiments using a case-study with cooked pork. *The Statistician*, *48*, 349-359.

Boyle, C.R. (1999). A problem-based learning approach to teaching biostatistics. *Journal of Statistical Education, 7*, 1.

Chatfield, C. (1995). *Problem solving: A statistician's guide* (2nd edn). London: Chapman & Hall.

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J., & Ostrowski, E. (1994) *A handbook of small data sets.* London: Chapman & Hall.

Grubb, H.J. (1998). Models for comparing athletic performance. *The Statistician, 47*, 509-521.

Jolliffe, F. (2001). Learning from experience. In C. Batanero (Ed.), *Training researchers in the use of statistics, IASE Round Table Conference, Tokyo 2000*, (pp.355-370). Voorburg, The Netherlands: International Association for Statistical Education and International Statistical Institute.

Manly, B. (1988). The comparison and scaling of student assessment marks in several subjects. *Applied Statistics*, *37*, 385-395.

Manly, B. (1992). *The design and analysis of research studies.* Cambridge: Cambridge University Press.

Marsh, C. (1988). *Exploring data: An introduction to data analysis for social scientists.* Cambridge: Polity Press.

Moore, D.S. (1997) *New pedagogy and new content: the case of statistics.* International Statistical Review, *65*, 123-165.

Sincich, T. (1993). S*tatistics by example (5th edn).* New York: Dellen.

Snee, R.D. (1993). What's missing in statistical education? *The American Statistician*, *47*, 149-154.