Unmathematical Statistics

Ruth Hubbard - Brisbane, Australia

1. Introduction

This paper is concerned with the process of designing and implementing a new statistics course. In the past, university-level statistics courses have just evolved. Frequently, academics use the courses they were taught and modify them from time to time to reflect changing circumstances. I have done this myself several times.

The approach I have taken in this case is quite the opposite. I began by specifying:

(i) the client group and their attributes;
(ii) the institutional constraints within which the course would have to operate;
(iii) objectives for the course which meet the clients' needs;
(iv) the results of research on how students learn.

In addition, I considered the factors that make learning statistics difficult for students, for each such factor asking, "Is this factor an essential part of understanding the methodology of statistics or is it peripheral? Should it be left out altogether or should it be taught in a new way?". I will begin by describing the specifications for this course.

The clients are undergraduates who have weak mathematics backgrounds, do not like mathematics, or just do not want to study a mathematical subject. They need to learn something about statistics because it is now an important part of most disciplines and professional courses. This may be the only formal statistics course they take.

The objectives are to interest the students in the investigation of data, to enable them to understand technical and media reports which contain statistical information, and to make them aware of what statistics can and cannot do.

The constraints are not very different from those applying at other Australian universities. The course is allotted three contact hours a week for a fourteen-week semester. For one hour per week the students can be divided into tutorial groups with a ratio of about 20 students per tutor. The enrolment is about 200 students. There are computer laboratories with 60 machines available for this course. The tutorials can be scheduled so that all students can take part in a one-hour computer tutorial each week and can have access to the laboratories at other times.

2. Student learning

Meaningful learning only takes place when students can incorporate new concepts into their existing cognitive structure, thereby modifying the structure (Ausubel, 1963). Whether or not they studied probability and statistics at school, most students have had experiences which have enabled them to develop some concepts about these subjects. The students' concepts may be correct or incorrect but, either way, they should not be ignored. By beginning a course with a series of abstract definitions quite unrelated to students' previous experiences we almost ensure that no learning will take place. Instead of being understood the definitions are memorised (in case they are asked for in the examination). The definitions are not used to do exercises. The exercises are done by copying very similar worked examples in lecture notes or texts. In extreme cases this gets down to rules such as "If it is a two-way table, do a chi-square". If an exercise deviates in any way from worked examples it is a big problem.

If the objective of a course is to enable students to answer a very narrow range of standard exercises the above approach is quite successful. My course has quite different objectives so student learning has to be different too. The preferred learning style has to be established right at the beginning and reinforced throughout the course.

Students have many other problems with standard courses, such as the following:

(i) Understanding definitions expressed in concise mathematical language, for example, the definition of a random variable.

(ii) Understanding formulas with lots of symbols in them. A good example is any formula for variance; this conveys no meaning to many students about *the* most important concept in statistics.

(iii) Coping with all the technical jargon (for example, terms such as mutually exclusive).

(iv) Having to reformulate problems to suit the presentation of various statistical tables. This seems to become the main focus of many exercises and distracts students from the main concept being illustrated.

(v) Indirect or mathematical questions which students find baffling. Here is an example: "A normal distribution has the mean $\mu = 63.4$. Find its standard deviation if 20% of the area under the curve lies to the right of 79.2."

(vi) Following difficult logical arguments as in hypothesis testing. Very few students grasp this method of reasoning at their first encounter.

(vii) Absorbing the large amount of detail that is presented and extracting the important concepts from it.

3. A new course

In my new course I have tried to overcome the above problems and to incorporate several strategies for improving learning.

MINITAB: MINITAB is used as a tool in the exposition of the subject matter so that in a sense it becomes the expert, the authority, throughout the course. The students do not have to "learn" MINITAB. Each time I make use of it in the text I include the commands with the printout so that students can make use of them later. They gradually learn to operate in the MINITAB environment quite unconsciously and use MINITAB-speak like normal language. The approach to probability and sampling distributions is based entirely on students' previous experiences and simulation. This kind of approach to probability was taken by Moore (1985), but without the advantages of a simulation package. These simple simulations are quite sufficient to equip students to use the concepts of probability in estimation.

MINITAB also provides the values of the standard statistical distributions. No tables are needed, just the graphs of the distributions and the CDF command. A related bonus is that it is not necessary to convert to z or t scores. Since the clients of this course are not very comfortable with formulae, using MINITAB to do most of the calculations means that many formulae can simply be left out. Necessary formulae are written in words.

Language: It is well-known that students have great difficulty reading textbooks written in conventional mathematical language. In previous courses in which I used a standard text (Bhattacharyya and Johnson, 1977) I tried to teach the students "how to read the book" as we progressed through the course (Hubbard, 1990). This was of some help but it seems to me now that we should try to meet the students half-way and make the text easier to read as well. Since I could not find a text for the present course that addressed the difficulties which I listed at the beginning of this paper, I have had to write my own. This has given me the opportunity to write in a language which students *can read.* I set myself the following guidelines for writing this course material:

(i) To make it personal I use "I" and "you" rather than the formal "we".
(ii) To make it seem alive I use the active rather than the passive voice.
(iii) To make it easy to understand I use short sentences, with one idea per sentence.
(iv) I only use technical terms that are really necessary, and these are used frequently so that they become part of the students' vocabulary.
(v) I insert simple questions at critical points through the text so that students can check their comprehension.
(vi) I deliberately make the text redundant. This is very "unmathematical" but contributes to readability and understanding. After all, when we are explaining something verbally to a student and the student appears not to understand, we rephrase the explanation.

Course content: Many educators have commented on the deep chasm that exists between what the teacher teaches and what the student learns. Teachers either want or feel it is their duty "to tell it all". I have presented the students with an amount of material that I think they can actually absorb each week. Some topics have been treated in great detail, others just by way of examples. The topics treated in detail are those that I will use in the development of further concepts. Some of the usual topics in elementary courses are left out altogether. Formally the syllabus reads much like any other for an elementary, terminal course in statistics. The course begins with the investigation of data and finishes with regression and experimental design.

Other pedagogical issues: The course is based on a spiral model of learning. A concept is introduced in one chapter at an elementary or intuitive level and later the concept is discussed again in more depth or more analytically. For example, confidence intervals for proportions appear in the first chapter. At this stage students construct them using a table. Later when all the necessary background has been dealt with they construct them analytically. Later still the idea of confidence intervals is extended to predicted values in regression. By this time confidence intervals seem like old friends rather than mysterious objects that come from very complicated formulae. New concepts are introduced through examples rather than definitions. Once the example is thoroughly understood, generalisation is sometimes attempted. A number of examples are used over and over to remind students of the important ideas that those examples embody.

Data: Statistics is concerned with the collection and interpretation of data. It is important for the student to be conscious of this throughout the course. Students start by discussing and describing data sets about everyday life. Then they learn the rights and wrongs of collecting data from samples and experiments. At every stage of the course students are required to interpret the result they have obtained, whether it is a graph or a single number or a complete ANOVA table.

Motivation: Personal data from the students are collected and used for data investigation exercises. To demonstrate the widespread use of statistics, students go to the library and find what proportion of articles in technical journals contain statistics. Right at the start of the course students are given an opportunity to experiment with MINITAB. This helps arouse their curiosity.

5. Student reactions

The course has only run for one complete semester with a small group of students (25). Unfortunately they are not a typical group, quite a few are not first year students and several have failed previous statistics courses. However, I have learned a great deal by carefully observing their reactions to the course. At regular intervals the students filled in a brief Reactions form which gave them the opportunity to criticise and make suggestions for improvements. This is a very unconventional course so my main concern was "will the whole thing be a meaningful learning experience or will it leave the students in total confusion"? Overall, the students' reactions were encouraging and the quality of their weekly assignments improved as the semester progressed.

I was grateful to find that there were no complaints about the readability of the material or the time required to do the reading. The students were not aware of my strategies to make reading easier but they did not complain about the material and sometimes they even said it was interesting.

The students loved using MINITAB to do their exercises. Many commented on this as being the best part of the course. Producing their own results on the computer seemed to give them a sense of power and mastery over the subject. If they made an error it was easy to rectify. They were intrigued by unexpected results in their printouts and were keen to have them explained.

This course is the culmination of many years of experimenting with methods of teaching statistics. However, it is a greater leap into the unknown than anything I have attempted before. Producing the course and helping the trial group through it has been exciting and challenging for me. I hope it will have the same effect on future students.

References

Ausubel, D P (1963) *The Psychology of Meaningful Verbal Learning*. Grune and Stratton, New York.

Bhattacharyya, G K and Johnson, R A (1977) *Statistical Concepts and Methods*. Wiley, New York.

Hubbard, R F (1990) Teaching mathematics reading and study skills. *Int.J.Math.Educ.Sci. Technol.* **21(2)**.

Moore, D S (1985) *Statistics Concepts and Controversies*. W H Freeman, New York.