Two Projects from the British "Computers in Teaching Initiative"

Adrian Bowman - Glasgow, Scotland

1. The Computers in Teaching Initiative

In 1983 the Nelson report on "Computer facilities for teaching in universities" was published in the UK, commissioned by the Computer Board for Universities and Research Councils. This report looked ahead to 1992 and foresaw extensive student use of computer workstations in the university teaching of many subjects. In an attempt to prepare the academic community for the changes ahead, an injection of funds was made with the aim of developing the use of computers in teaching. The Computers in Teaching Initiative has since funded 139 projects covering virtually all academic subjects, including those such as statistics, where intensive use of computing resources was standard practice, and those which might largely be regarded as virgin soil, such as Greek and English. At least one project was based in nearly every university in Britain.

The nature of these projects is enormously diverse, ranging from the development of completely new software to the integration of existing packages with lecture and laboratory classes. Five projects were pursued in statistics.

1. The use of Statgraphics in statistics courses.
   (University College of Wales)
2. The teaching of APL to statistics students.
   (University of Warwick)
3. Statistics teaching using graphics workstations.
   (University of Lancaster)
4. A help and macro system for statistical packages.
   (London School of Economics)
5. Computer illustrated texts in statistics.
   (Universities of Glasgow and Sussex)

Some of these projects were funded as part of larger activities in the mathematical area. Together they illustrate some of the roles which computers can play in the teaching of statistics.

Projects 1, 2 and 3 have taken existing software and addressed the important issue of how a statistical computing environment should be integrated within the teaching programme of a department. This issue is faced by all university teachers of statistics, but the additional resources given to these projects has allowed more ambitious solutions than have been available to most departments. For example, the project at Lancaster has been able to use powerful workstations to run GLIM and GENSTAT in an attractive windows environment with the facility to display and manipulate worksheets, data, and graphics, while running these packages.

Project 4 developed new software to provide a flexible and programmable facility for adding exercises, worksheets, or helpful information on top of an existing package running on IBM compatible microcomputers. The "memory resident" program allows pull-down menus to be superimposed on the screen at any point. The contents of the menus and underlying pages of text can be created by the teacher. In the sense that it takes an existing statistical computing environment and enhances it for use in teaching, this project has a similar aim to the first three.

Project 5 also developed new software but here the aims are rather different, with emphasis placed on making use of the graphical facilities of microcomputers to aid the understanding of basic concepts in statistics. This project is described in greater detail in Section 2, where the relative merits of purpose-built menu-based software and standard statistical computing environments in achieving these aims are contrasted.

In Section 3, a new phase of the Computers in Teaching Initiative is described and the role of an information centre on the use of computers in teaching statistics is discussed. Personal experiences within the Computers in Teaching Initiative so far are used to give some pointers to the future in Section 4.

2. Computer illustrated texts

Considerable attention has rightly been paid to the role of the computer in supplying a friendly environment within which students can explore data, put into practice the modelling tools which they have been taught in lectures, and gradually acquire the skills of data analysis. Much less attention has been paid to the potential role of the computer in helping to communicate basic ideas, techniques, and theoretical results of the subject. An example of this is provided by a type of teaching material called "computer illustrated texts", which has been developed by the author in conjunction with Dr D R Robinson of Sussex University, and other colleagues in the mathematical area. The aim of this integrated text and software approach is to develop the intuition and understanding of a student by the use of computer graphics.

The simplest use of such software is as an "electronic blackboard", where pictures can be drawn quickly, easily, and accurately. Figure 1 gives an example where Normal approximations to Binomial distributions are displayed. The theoretical approximation result, and rules of thumb on when it is appropriate, can then be explored in a simple way. In other types of picture a gradual step-by-step construction can be more helpful than the direct presentation of the completed picture. Animation is a

technique which can be put to particularly good effect in the illustration of techniques such as the construction of a stem-and-leaf plot or the calculation of a two-sample non-parametric test statistic. The aims and techniques of this type of software are described at greater length in Robinson and Bowman (1990).

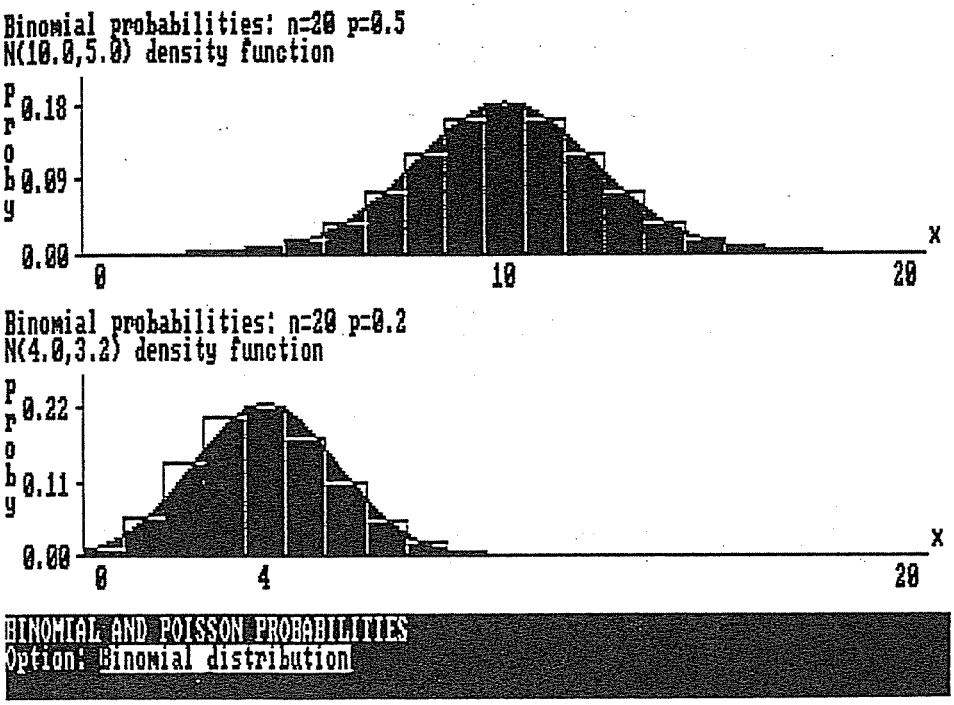

FIGURE 1
Normal approximations to Binomial distributions from the computer illustrated texts software

It is important to identify the relative merits of this approach, using specially designed menu-driven software, and standard statistical computing packages. Although the latter are not designed as tools to explain, they can sometimes be used in this way with some thought. An example is provided by an illustration of the meaning of a confidence interval. Bowman and Robinson (1988) provide a program which simulates samples of data and displays the resulting confidence intervals as illustrated in Figure 2. This communicates the fact that the intervals have random endpoints and that "confidence" is associated with the proportion of times which the intervals capture the true value. A limited choice of distributions allows the effects of non-Normality to be explored.

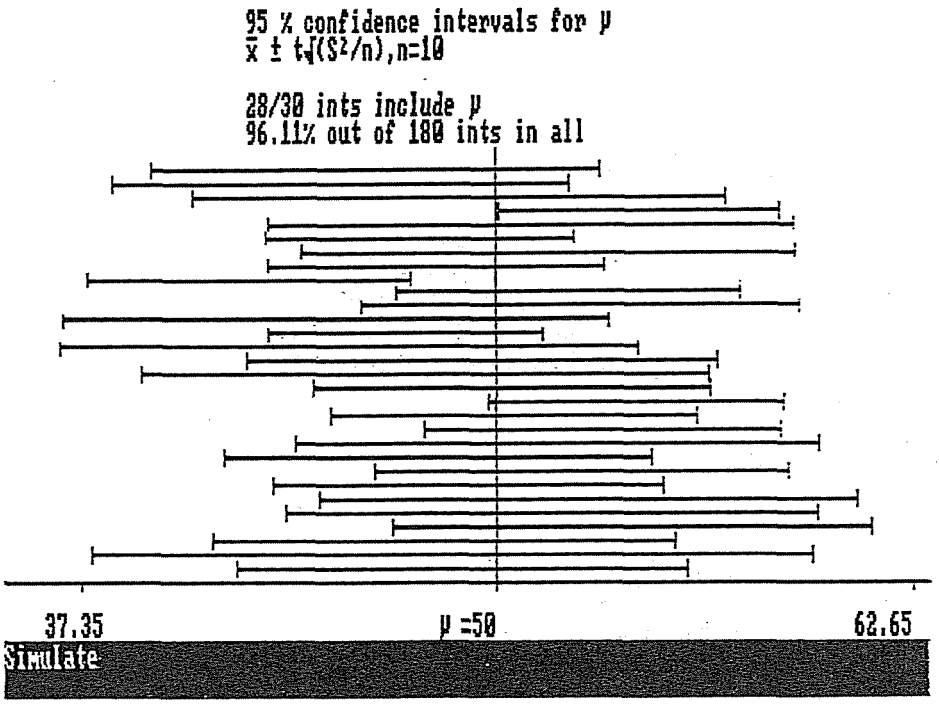

FIGURE 2
Simulated confidence intervals from the computer illustrated texts software

Alternatively, the following MINITAB macro achieves similar aims in a less packaged form.

```
store 'ci'
random 10  C1:
     normal 0   1.
let k2=mean(c1)-2.26*stdev(c1)/sqrt(10)
let k3=mean(c1)+2.26*stdev(c1)/sqrt(10)
let k1=k1+1
print k1
let c2(k1)=k1
let c3(k1)=k2
let c4(k1)=k3
end
```

Repeated execution can be carried out through the commands

```
let k1=0
noecho
exec 'ci'  15
```

and graphical output obtained through

```
mplot    c2    c3    c2    c4
```

The resulting scatterplot displaces the end points of simulated confidence intervals in the same form as Figure 2.

The numerical verification of the meaning of confidence is more difficult to organise but there can be benefits in understanding if a student has to think through the exercise of constructing a program even as simple as this one, and with some there is the motivation of viewing the output of one's own creation. The contrast is between a tailor-made piece of software which targets a specific concept and makes a small number of points in a forceful way, or a more general-purpose environment which requires more time and guided effort by the student but encourages other skills. The issue of time and facilities may often be a deciding factor between these two approaches, but as approaches to teaching they should be seen as complementary rather than competitors.

Of course, there are many areas where the use of a standard statistical computing environment does not allow the graphical approach being advocated here. Consider Figure 3 where the simple concept of a two variable regression model as a fitted plane in three dimensions can be helpfully illustrated to students who meet this for the first time. Even those who are familiar with the mathematics may benefit from a visual display which can make the effects of outliers, and collinearity, easier to understand.

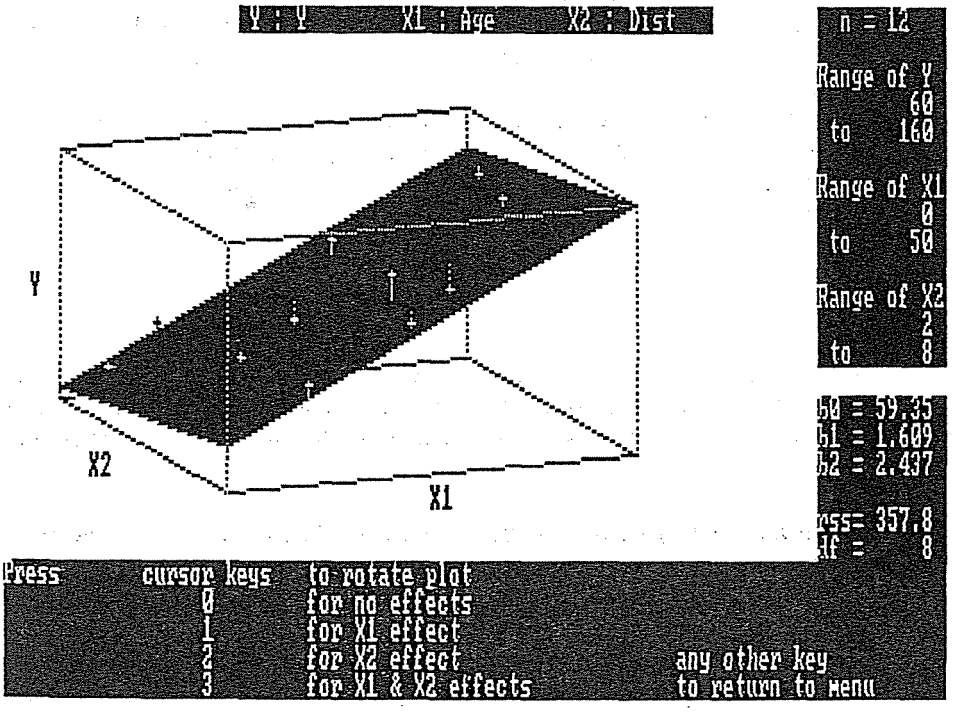

FIGURE 3

A pseudo three-dimensional plot of a regression model with two explanatory variables from the computer illustrated texts software

3. The CTI Centre for Statistics

The development stage of the Computers in Teaching Initiative, intended as a "pump-priming" exercise to allow a few projects in each subject to explore possibilities, is now over. A second stage has now been launched with the funding of twenty-two subject centres throughout the country. The aim of these centres is to collect, digest, and disseminate information on the use of computers in teaching in particular areas, and not just on previously funded CTI projects. The CTI Centre for Mathematics and Statistics has recently been established with two bases, one in Birmingham (Mathematics) and one in Glasgow (Statistics).

The brief to "collate" has been fulfilled initially by surveying the departments which appear in the UK Academic Directory of Statisticians. A detailed description of the results of this exercise will be described elsewhere when returns are more complete. Some simple and unsurprising, but nonetheless useful, results are already clear in the frequencies of occurrence of particular names in the lists of standard packages used, with MINITAB and GLIM, in that order, appearing at the top of the list. This kind of information will be used to guide the future compilation of auxiliary material, such as sources of macros, linked books, and general advice, which the centre will seek to provide. A database of software useful in teaching is under construction, and the first edition of an annual software guide has already been produced.

The brief to "digest" is being fulfilled through reviewing. A list of packages and technical descriptions would be of little value without informed comment on their effectiveness in teaching. This is difficult to achieve but is being built up slowly through the commissioning of reviews and the permitted reproduction of reviews from other sources.

The brief to "disseminate" is being fulfilled through a software guide, regular newsletters and, in the near future, a bulletin board. A programme of workshops has also been initiated as a very effective means of allowing smaller numbers of people to make their own evaluation of software.

4. The future

The development projects funded under the CTI scheme, and the activities of the CTI centre, have underlined the diversity of ways in which computers can be used to assist in teaching statistics. Some personal comments on the future are given under the headings below.

(i) *Statistical computing environments:* Some standard packages have now adapted well to the interactive and often graphical environment of the PC. With increasing moves to distributed computing systems, of increasing power, it is to be hoped that this trend will accelerate.

(ii) *Enhancements to statistical computing environments:* The window environments of workstations and PCs offer great potential for increasing the friendliness of software and for integrating it with coursework. Two of the projects mentioned in Section 1 have worked along these lines and it is to be hoped that others will follow suit.

(iii) *Software for specific tasks:* As standard packages become more powerful, and include more flexible programming facilities, the need for separate software for specialised tasks, such as nonlinear regression or cluster analysis, will reduce. It is to be hoped that the providers of software packages will give careful thought to the programming tools they provide and to the need for standard interfaces with Fortran and other routines.

(iv) *Software to explain:* The role of the computer in assisting explanation, as in the computer illustrated texts, is a narrower but very useful one. The ability to create this type of software within a standard package would be extremely helpful but is still some way off.

(v) *Computer assisted learning packages:* Some question-and-answer-based systems have been developed in the past but few of these seem to be in current use. The advent of expert systems software, and of hypertext, may lead to a renaissance.

(vi) *Database management systems:* As the world at large makes increasing use of this type of software, database concepts and software are likely to continue to find their way into university courses.

I look forward with expectation to the continued vigorous development of computers in the teaching of statistics throughout the 1990s.

References

Bowman, A W and Robinson, D R (1986) *Introduction to Statistics : A Computer Illustrated Text.* IOP Publishing Ltd., Bristol.

Bowman, A W and Robinson, D R (1990) *Introduction to Regression and Analysis of Variance : A Computer Illustrated Text.* IOP Publishing Ltd., Bristol.

Francis, B and Whittaker, J (1986) Teaching computational statistics on workstations. In: Lawo et al. (eds) *Proceedings of Compstat 86.* Physica Verlag, Heidelberg.

Jones, A (1986) The use of microcomputers in teaching time series analysis. *Proceedings of the Second International Conference on Teaching Statistics.* University of Victoria, Canada, 158-165.

Robinson, D R and Bowman, A W (1986) *Introduction to Probability : A Computer Illustrated Text.* IOP Publishing Ltd., Bristol.

Robinson, D R and Bowman, A W (1990) Computer illustrated texts in statistics. *Collegiate Microcomputer* **8**, 9-13.