

PACE and STATLEV - Two Examples of Software Packages Developed with the Particular Purpose of Teaching Introductory Business Statistics

Robin G Boyle - Melbourne, Australia
Peter S Mortensen - Aarhus, Denmark

1. Introduction

We first met at a seminar on teaching introductory business statistics at the Aarhus Business School, Denmark, in 1984. It soon became clear that we shared the same views on a number of matters critical to the teaching of our subjects. In particular, we were not happy with the standard business statistics course as promoted by many American texts, we were willing to admit that statistics can be very unpopular with students and misunderstood and under-utilised by colleagues teaching other subjects, and we recognised that if computers were to be used in teaching business statistics they must be introduced in a way where students concentrate on learning statistics rather than computing.

We have both been involved in the formulation of syllabuses tailored specifically for the types of courses and students at our respective business schools and have written extensive accompanying material and texts. And both of us have overseen the development of statistical computer packages for use in our courses. In developing these packages, a basic criterion has been to produce software covering all the statistical functions in the syllabus but requiring no special computing or programming skills. The aim has been to ensure that students can get on and off the computer as quickly and painlessly as possible while still appreciating the power of computers. Most commercial packages available ten years ago required teachers and students to spend excessive time on computing, with consequent loss of valuable time for understanding statistics. The packages were certainly not suitable for the courses in introductory

statistics or quantitative methods which all business students have to study; for advanced statistics and research they were excellent. The situation has hardly changed.

Writing our own software has been far from an easy and rewarding task. Numerous problems in data handling, in input and output design, and in writing adequate documentation have been encountered. This paper presents our views on the use of computer packages in teaching business statistics and discusses some of our experiences, problems, and successes in developing our software.

2. Computers in business statistics

Today, when an electronic calculator is almost as important as pen and paper, a course in business statistics without computers ought to be unthinkable. Students need to be familiar with statistical computing as, in the future, the bulk of statistical calculations will be performed by computer, and employers will expect competence in statistical computing. But, more importantly, computers can significantly improve the teaching and understanding of statistics and increase motivation of students since larger and more realistic data sets can be analysed, statistical concepts can be illustrated in class, and students have a convenient tool for completing or checking assignment work.

But some warnings have to be given. Clearly the students must have adequate computing facilities, the software used must be appropriate for a course, care must be taken not to turn a course in statistics into one in computing, and it is important not to let the software determine what statistical concepts are taught.

Personal computers are to be recommended since students and staff can be independent of mainframe computing facilities and can use the software on PCs at home or at work. More and more students will buy their own PC and will be further motivated to do so if reasonably priced statistical software is made available.

3. Selecting statistical software

There is no single answer to the question of which statistical package to choose since it depends on so many factors: the type of course, the capabilities of the students, the computing facilities available, and, most importantly, the desired syllabus.

The standard introductory business statistics course text is not necessarily appropriate. Account should be taken of the specific needs of your students, the courses they are studying, and their eventual employers, perhaps devising a quite radical syllabus and developing your own course material, and continually reviewing and revising the course content; this is not an easy step to take, given the convenience of working from a readily available text with accompanying workbooks and support material.

For your particular course, describe the desired computing functions, both for illustrative purposes and for "real life" analysis. The computer package you currently use may not be adequate and it may be time to change. Information on existing software is available from several sources including journals such as *Applied Statistics*, directories, and computer periodicals. Booklets on specific software, such as MINITAB and SPSS, are available and some texts have handbooks for these. If you cannot find adequate software, you can either choose the best one available and supplement it with

other software or develop your own routines or package. However, this last option requires a great deal of courage, effort, and funding.

A frequent review of the statistics syllabus, with consequences for the software used, is essential. Demands on statistical knowledge by employers and colleagues teaching other subjects change and new packages become available, but there is a tendency for courses not to change (and for software to become entrenched) as the effort required can be significant.

Clearly, the software chosen must have statistical functions that fit the syllabus; however, other features are also desirable.

It is important to remember we are teaching statistics not computing, and that many students might not be comfortable with computers. Hence, the package should require little in the way of computing skills and it should be easy to demonstrate in class and easy for students to use by themselves in assignment work. A menu-driven system (both computer and package) is the simplest and quickest for novices to learn and use. On-screen or on-line help is desirable. The package should add little to the time demands on students. With more complicated packages there is significant decay of knowledge unless they are constantly used.

Even if you use a commercial package you may still need to develop for your own students a simplified user guide that is specifically related to the syllabus.

Data entry, editing, and manipulation must be straightforward and comprehensive. It should be possible to key in data at the screen or retrieve it from a file, to make corrections or changes, and perform mathematical transformations. It should be possible to save the data as a file for later use. The ability to convert files for use by packages such as SAS, SPSS or Lotus, or to retrieve files set up by these packages is an extra benefit, as are being able to handle both numerical and alpha data and missing values. Graphics and simulation capabilities are particularly useful for demonstration purposes in lectures.

Finally, it would be a significant advantage if the package were not only suitable for teaching purposes but also for the workplace so that students can take their knowledge of the package with them when leaving business school.

4. Developing our own packages

When we began using computers in teaching our statistics courses we chose not to use any of the commercially available packages because, even though many of these packages were extremely powerful, they were generally not easy to use and only parts of our syllabuses were covered. They were, and still can be, expensive to buy and maintain, considering there may be hundreds of students involved.

The development of PACE began in 1977 and the development of STATLEV dates from 1984. Both packages began with fairly modest aims but changed and grew as computing facilities and courses changed. Both are still in use and will continue to be used in the foreseeable future as we are still not satisfied that the commercial packages currently available, including spreadsheet packages, meet essential criteria. We both intend to continue development of our packages, even though there is immense commitment involved in terms of time and funds, and there could be a wonderful package just "around the corner". There is, of course, a lot of invested effort which we are loath to

give up but we are happy that the packages work for us. Neither package is considered to be ideal but they certainly have their advantages. As they are user-friendly and based around the course syllabuses, little time is spent by lecturers or students on "computing", leaving adequate time for analysing and understanding the results. A significant advantage is being able to add new routines and options when courses need them, although this may mean days or weeks of concentrated work.

Writing the computer algorithms to perform the necessary statistical calculations is relatively straightforward. The most difficult and time-consuming aspects involve setting up logical and systematic ways of interacting with the user, making data entry as straightforward and comprehensive as possible, designing well-arranged and informative output on both screen and paper, building in "traps" so that users cannot "blow-up" the program, and debugging and testing. These "design" tasks have comprised about 80% of the programming effort. Quality graphics do not form part of PACE but they do form part of STATLEV; however, they have required expert programming.

We strongly believe that an understanding of statistics is more important than knowing how a particular package works. It cannot be guaranteed that an employer will have the same package that a student has studied or even have one at all, and if a student is faced with using a new package at work it can be learnt on the job.

5. PACE

PACE stands for Program for Accounting, Commerce and Economics. Robin Boyle commenced its development in 1977 after experience with some small programs at Monash University (Melbourne). Development continued in close cooperation with their Department of Econometrics and Operations Research which uses the package in teaching first-year economic statistics. In 1987 a version was developed for operation on IBM PCs and since then versions for the VAX and PC have been developed in tandem. Raising funds for development has been a problem.

PACE is fully interactive and menu-based. It has been written solely for teaching purposes and has been designed to achieve a number of aims: in particular, to be as easy, enjoyable, and interesting as possible to use for both staff and students, and to impact as little as possible on students' time. It is not nearly as powerful as commercially available packages but it was never meant to be used for work or research purposes. It is used at only three other institutes, including Monash University, and even though it is relatively rudimentary, lecturers are quite satisfied and feel it meets their teaching needs.

There are two compulsory units in quantitative methods in the Faculty of Business at Victoria College. Each unit lasts one semester of 13 weeks with one two-hour lecture and one two-hour tutorial per week. Quantitative Methods I covers descriptive statistics, index numbers, time series, and financial mathematics. Quantitative Methods II covers probability and probability distributions, sampling, estimation and hypothesis testing, and simple and multiple regression.

PACE has routines covering most of these topics, including financial mathematics. Most routines are designed so that students are led through a series of calculations or operations, thereby highlighting the types of analysis the student should consider for particular types of data. Statistical options include analysis of ungrouped

and grouped data, a frequency count routine to generate a frequency distribution or contingency table, a correlation-matrix routine, simple and multiple regression, Chi-square analysis, index number calculation, trend fitting and seasonal adjustment. The Chi-square, correlation, and regression routines include calculation of p-values for hypothesis tests. Under development are routines for calculating probabilities from various distributions (binomial, Poisson, normal and t-distribution), confidence intervals and appropriate sample sizes, and performing single and two sample hypothesis tests.

Data are handled in matrix form with columns representing variables and rows representing observations or cases. Only numeric data can be entered - alpha data must be transformed. Columns (or variables) can be given labels. There is a range of options for performing mathematical transformations on variables or for generating specific types of variables. Data matrices can be saved as disk files and can also be converted to/from operation with Lotus 1-2-3. A major problem with any package is that users will not read printed documentation properly, if at all. To overcome this, PACE has the option of full or brief prompts at the screen.

The package requires little instruction from lecturers. At Victoria College, students are given appropriate documentation and offered voluntary one-hour demonstration classes at the beginning of the semester; after that they are expected to teach themselves or each other. Students in Quantitative Methods I and II have one compulsory assignment each week. They are required to understand most of the statistical calculations in the course and all assignments include "long-hand" problems, although students may check their results on the computer. Most assignments also require some computer work involving more difficult data or calculations, and a report. The computer is not used just for computing's sake: computer-based assignment work is designed to be effective and to add to understanding but not to the time students spend on the subject. Assignments normally include some computing guidelines to minimise computing time. Examinations may include computer printout for interpretation.

PACE is sold through Victoria College at a more-or-less nominal price. It is available to other institutes for \$US400 for either the VAX version or the PC version. The current PC version is sold on floppy disks to students at a total cost of \$US30 but a new PC version, making more efficient use of memory and disk space, is being developed and it is hoped that its price will be no more than \$US25. The aim is to keep the cost of the package as reasonable as possible for students but at the same time generate some funds for future development.

6. STATLEV

Peter Mortensen has been the project leader for the development of STATLEV, coordinating graduate students hired by the Business School's Computer Centre to do the programming. This has caused some complications as the development has been partly out of his control. Also, students can only be hired for two years and it has been expensive to pay them.

STATLEV is fully menu-driven and has a comprehensive range of data entry and data manipulation modules. Data can be keyed in by variable or by observation and can be saved in various formats. Many kinds of manipulations on variables, observations, and elements may be performed, and observations can be generated from 16 probability

distributions. Variable and value labels are fully supported.

Sampling functions include the ability to draw a sample from an existing data set (for teaching purposes). The samples can be drawn as simple random, stratified, cluster or systematic samples, and the estimations will take into account how the sample has been drawn. Special modules are "Description of a data set as a sample" and "Calculating the sample size" (given a specified margin of error and confidence level for simple random or stratified samples). Other inferential routines include confidence intervals and prob-value calculations for a mean, a proportion, and differences between two means or proportions, and hypothesis testing for independence/homogeneity and goodness-of-fit to various distributions.

Descriptive features include calculating summary measures and frequency distributions, cross-tabulation and correlation matrix generation. The most commonly used graphics are available in high resolution, including bar charts, histograms, polygons, ogives, box-plots, cross-plots and fractile diagrams (Q-Q plots).

STATLEV is used in Statistics I; this and Statistics II are compulsory to all business students at The Aarhus Business School.

Semester	Course Title	Content	Software
1-2	Statistics I (20 lectures 15 tutorials)	Survey Sampling including Descriptive Statistics Probability Theory Estimation and Inference	STATLEV
3-4	Statistics II (27 lectures 22 tutorials)	Quality Control Econometrics (regression) Experiments (ANOVA) Analysis of Categorical Data	SAS

The theme of Statistics I - and thus STATLEV - is survey sampling. The calculations in Statistics II are still covered by SAS.

STATLEV is introduced to the students in the first tutorial in Semester 1. This takes place in a PC lab - every 2-3 students having access to a microcomputer. In the rest of the tutorials students work on their own. There will be at least one exercise where STATLEV is needed for solution. Output from STATLEV is presented in lectures, and written examinations include output and questions requiring knowledge of STATLEV. In the middle of Semester 2 students are set a project involving an analysis of some real business data using STATLEV.

STATLEV is generally the first example of large-scale software students have met, and overall The Aarhus Business School has been fairly satisfied with the success. Students encounter only minor problems and learn quickly how to use STATLEV; some students even wish to use and buy the package in the later part of their studies and after they have left business school.

However, we can improve even more: plans for the immediate future include writing an improved manual detailing the more complicated data-manipulations that may

be required, particularly with assignment work.

There is a lack of PC capacity for classroom and student use, so the intention is to make STATLEV available for the 30% - by 1990 - of students having (access to) a PC and commercially available to others. Also, an English version will be available, as all text in STATLEV is organised in separate files ready for translation into any language.

7. Conclusions

We are not satisfied that the statistics and spreadsheet packages available commercially are satisfactory for teaching *our* courses in introductory business statistics. For this reason we have developed our own and intend to continue development even though it is a most demanding task.

We would not recommend anyone to start writing their own package as commercial packages appear to be improving for teaching statistics. In the future there will be easy-to-use spreadsheet packages with statistical, graphical, and simulation capabilities suitable for teaching all levels of business statistics and it is likely that knowledge of such a package will become a prerequisite for incoming business students, not just for statistical work, but for other areas of study as well.

In the meantime we recommend against allowing a particular package to dictate the syllabus and how business statistics is taught. It is more important for students to understand statistics than the workings of a particular package.

References

- Boyle, R G (1984) *PACE Statistics Reference Manual*. Victoria College Press.
- Boyle, R G (1986) *The Use of Computer Packages in Teaching Business and Economics Statistics*. Contributed paper at the Second International Conference on Teaching Statistics, University of Victoria, Canada.
- Mortensen, P S (1984) The art of planning a course in applied statistics for business and economics. In: R C Brown (ed) *Quantity and Quality in Economic Research I*. University Press, New York.
- Mortensen, P S (1986) On a new approach to an introductory course in statistics for business and economics including some experience. In: R Davidson and J Swift (eds) *Proceedings of the Second International Conference on Teaching Statistics*. University of Victoria, Canada.
- Pedersen, H K N (1989) *STATLEV Grafik Introduktion*. Aarhus Business School. (In Danish)
- Pedersen, H K N and Lysgaard, J (1989) *STATLEV Introduktion*. Aarhus Business School. (In Danish)
- Woodward, W A, Elliot, A C and Gray, H L (1985) *Directory of Statistical Microcomputer Software*. Marcel Dekker, New York.