Teaching applied statistics courses using computer laboratory final examinations

Tania Prvan and **Margaret Mackisack**, Macquarie University, Australia
tprvan@efs.mq.edu.au mmackisack@efs.mq.edu.au

Abstract: *Courses taught in Applied Statistics, such as regression or multivariate analysis, tend to have the examination component based on a final written paper, either with computer output attached for interpretation or with summary statistics given so the calculator can be used to evaluate test statistics and hence make inference, or a mixture of both. Assignments may be based on students using a statistical package to do the analysis. The authors found this approach unsatisfactory. The first author trialled at the University of Canberra in 2001 both a mid semester computer based examination and final computer based examination with the students being allowed to use a variety of statistical packages. Student feedback was so favourable that she taught the course the following year in the same manner as well as a regression course. When she moved to Macquarie University she split the final examination for the third year regression course she was in charge of in 2004, 2005 and 2006 into a computer laboratory examination and a separate written paper. We have been unable to find more than a small number of other examples of this approach, which seems to have considerable promise as a way of implementing authentic assessment in applied courses. As well as case studies, issues associated with setting, running and marking such examinations are discussed.*

Introduction

Applied statistics courses lend themself to computer based assessment. The statistical procedures are tedious to do by calculator and most statistical packages implement a wide range of procedures. Emphasis in such a course should be placed on statistical analysis and not procedural calculations. As noted by Smith (1998) "statistical reasoning should take precedence over statistical methods.'' In any case, for many procedures including regression, obvious hand calculation algorithms are not numerically reliable and students should not be led to believe in their usefulness.

For statistical procedures that are not implemented in the available software package the computer can still be used to assist in the calculations. Furthermore, with a computer real-sized problems can be used. Students with poor mathematical skills are not disadvantaged. Students using statistics in their later working life are more likely to need to know what procedures to use, how to implement them using a statistical package, how to interpret their output and to assess whether their analysis is valid.

Traditional assessment does not totally address the problem that students need to utilise statistical packages in solving real life statistical problems. We want authentic assessment techniques (see Chance, 1997) that assess whether a student can do the statistical analysis including the mechanics of using a statistical package to implement the procedure/s needed.

Computer based assessment, where the computer is a management tool for the teacher, has been common for many years now. Multiple choice type questions and short answer/single number questions particularly lend themselves to this type of approach, and tools such as *WebCT* have been developed to facilitate this use of the technology. However, such management approaches have not usually been developed with a view to giving the students a real-world analysis experience in their examination. It is this which is the focus of this paper.

Assignments may allow students to use statistical packages but most assignment work by students, in the authors' experience, are group efforts whether they are meant to be or not. The lecturer cannot judge from such unsupervised work whether the student can either use the software, or interpret the resulting output. The obvious next step therefore is to conduct formal examinations in the computer laboratory, presenting the students with an analysis problem and requiring a solution in a short report in a limited time frame.

Such an examination could be a part or the whole of the assessment for an applied course. Setting, conducting and marking laboratory based examinations raises some issues that are not present in traditional examination formats. In the next section the paper discusses some of the practical issues associated with using laboratory based examinations, preparing for the examination (setting the paper and arranging the physical environment) and conducting the examination. In section 3 two cases are presented where such examinations have been implemented by the authors.

Conducting laboratory based examinations

Preparing for the examination: setting the paper

Setting a paper for a practical examination conducted in a computer laboratory requires firstly the acquisition of a suitable set of data: this is also required for a paper-based examination where the students are presented with the results of an analysis for discussion and interpretation.

Questions of scope then arise: how closely are the students to be directed in carrying out the analysis? This will depend firstly on the level of experience of the students: for the courses described in this paper they were all in the later years of undergraduate statistical majors. The authors have not considered issues associated with using such examinations for beginning students. Secondly, the amount of direction required will depend on the range of methods covered in the course. If linear models are the only topic then the students will assume that they are required to fit a linear model of some type, and one then examines how well they apply the software to fitting a suitable model. If the course covers for example linear discrimination, cluster analysis and principal component analysis then the examiner has options: the students may be directed to carry out a particular analysis (and one discovers how well they can apply the software to carrying out that analysis): or the students could be directed to answer a substantive question about the data, and left to decide for themselves which analysis to perform. This allows scope for much more searching investigation of the students' learning, but also opens the possibility of making the examination too hard: we want to learn what the students know, not what they do not know.

The length of the paper also needs to be carefully considered, particularly if the students are required to produce a report based on their analysis, using some word processing software. Will there be time for the mechanics of cutting and pasting etc? Choosing the appropriate analysis, carrying it out and reporting on it will be likely to take longer than simply reviewing an analysis presented by the examiner so taking a 2-hour paper with package output for comment, removing the output and asking the students to both produce and comment on it will result in a paper that needs more time.

Another scope issue arises if the students have access to more than one package: are they to be directed which software to use for a particular problem, or required to make a choice?

Preparing for the examination: other issues

It is highly desirable that all students taking such an examination should be in the same room at the same time so that all may be equally exposed to the vagaries of the hardware system so far as that is possible. The local system will influence the ease or difficulty of making the examination system secure, so that a student cannot engage an outsider to carry out the examination and send in results. Arranging a suitable location, making data and software available to the students may be simple or complex depending on local conditions. Whatever these are, the students should be fully briefed about them and allowed trial access to the system (although not, of course, to the examination data) beforehand: this is as much a pilot test for the examiner as familiarisation for students.

Our experience is that soft copy of solutions is preferable to students printing a hard copy. Depending on the local system this may require handing in some form of disk, or sending a file to a secure location. Examining work presented totally in soft copy has the advantage of not having to

interpret handwriting, but the disadvantage that a student is unlikely to have time to give highly lucid interpretations and descriptions of output. One way of overcoming this is having spaces on the examination paper for students to write their answers and have them only submit relevant output on soft copy.

Another issue is that the students' positions are dictated by the layout of the machines and it may not be feasible to prevent individuals seeing others' work on screen and hence getting hints about what to do. If possible students should be seated so that they do not have direct views of others' screens: it would not be reasonable to insist that students never look up during the exam, and if another screen is directly in their line of sight it may not be possible to ignore it.

As with paper-based examinations it must be decided what students are allowed to bring to the examination room in the way of books, notes or reference material. With a computer available the scope of possibilities is increased to potentially include material on CD-ROMs or accessible over the Internet. Managing the use of such material may be complicated depending on local system configuration, and may need to be controlled by examination invigilators.

Conducting the examination

Laboratory based examinations differ from normal paper-based examinations in several ways. The most obvious is that they depend completely on the functioning of the hardware and software for the duration of the examination. Laboratory based examinations are tinged with a faint anxiety not associated with paper examinations .

There is also more scope for students to accidentally delete work they intended to hand in, to erase important files, or to have other operating disasters that do not affect paper examinations. Students should be advised beforehand what provision is made for taking such events into account. To some extent the examination assesses the students' general facility with the system as well as with the statistical package and the course content, and it should be made clear beforehand what level of skill is required. The need to develop skill for the final examination should help to ensure that the students actually do carry out the assignment work themselves.

Invigilation should probably be carried out by someone very familiar with the examination so that the use of forbidden materials can be detected.

Case studies

Described below are two experiences of the authors in running an applied multivariate analysis course and an advanced linear models course in which computer based assessment under examination conditions was used. Two other case studies, not taught by the authors, are described briefly.

Applied Multivariate Analysis - taught at University of Canberra in 2001 and 2002

This course deals with the application of multivariate methods. Enrolment in this third year subject is low: three to eight students. Both mid-semester and final examinations were carried out in the computer laboratory. In 2001 class contact consisted of three hours of lectures and one one-hour tutorial a week. In 2002 the tutorials were increased to two hours while reducing the weekly lectures by one hour as a result of student feedback. To counteract student reluctance for a final examination on the computer in 2001, the mid-semester test was held in the computer laboratory and all tutorials were held in the computer laboratory using a variety of statistical packages as well as *Excel*. *MINITAB 13* was used to reduce the reliance on programming skills.

Mid semester test

The author distributed the previous year's mid session test as well as a computer version two weeks before the test. Selected solutions were distributed in the week before the test.

In 2001 the test was 110 minutes long instead of the usual 90 minutes in case of any glitches. A separate disk for each student with the data needed for the test was distributed. Output produced by the student was meant to be saved to this disk. Output produced for each part of a question was meant to be saved by the students in a separate file for ease of marking later by the lecturer. There were also spaces on the test paper to answer questions about the analyses performed. Students were allowed to bring in lecture notes, course material handouts, and other hand written material into the examination but no text books. For the mid session test students used their own computer accounts. One student looking up an electronic text book on the web prompted the lecturer to hold the final examination on computers that were not on the network so there was no access to the web or their individual computer accounts. This particular student's performance on the final examination was significantly worse when deprived of electronic textbook access. The same format was used for the mid semester examination in 2002 and no major difficulties were encountered by the students. No students in either year reported finding the test more stressful than a conventional test.

The most onerous part of writing the mid session test was finding data sets to use. One way of getting around the fact that a lot of data sets do not meet the model assumptions for the analyses performed was to get the students to first perform the analyses and then to get them to discuss whether the analyses were valid. Students did not have to then perform the more appropriate analyses for the test.

Marking the computer mid session test was quite straightforward. The time consuming part was preparing the solutions including computer output. These were distributed to the students when their tests were returned to them. Spaces on the examination paper to answer specific questions speeded up the process. One whole test at a time was marked while looking at the student's output on the disk that was handed back with the test paper. Marking the test was no more onerous than marking a standard paper based test with hand calculations and computer output attached.

Final examination

Before the final examination in both years two sample computer final examinations were distributed. The author used mainly data from private consulting problems. Selected solutions were handed out approximately two weeks before the final examination. The last week of lectures was spent revising.

The three hour final examination in 2001 was extended by an extra fifty minutes since this was the first time any of these students had done a final at a computer terminal. The lecturer supervised the examination herself. All machines used were off the network. Again space was provided on the examination paper to answer questions. Explicit instructions were given regarding saving of output. In fact each part of a question that required the use of a package required the output to be saved to a file of the form q-number-part, for example the output produced for question 3b) would have file stem q3b. This made marking the questions much quicker since the lecturer did not have to search for the output on the disk. All examination data sets were on disks which each student was given a copy of at the beginning of the examination. The students also saved their output to the same disk. Feedback immediately after the examination from the students was that it was no more traumatic than doing a paper based final examination.

In 2001 all of the examination papers and disks bar one student were quite straightforward to mark and no more time consuming than marking a conventional paper based examination. The only difference was being in front of a computer to check the output produced. The only examination that was time consuming to mark was that by the student who had ignored the instructions on how to save

any output produced. He had all the output in two files that the lecturer had to search for the relevant output. In 2002 all examinations were straightforward to mark. This time the computers were still connected to the network but otherwise examination conditions were similar to 2001.

Computer based assessment for the final examination enabled the students to at least check the model assumptions for some of the techniques used which is something that is difficult to do successfully in a conventional examination because the output provided usually gives the students clues to what they should do.

Linear Models - taught at Macquarie University in 2004, 2005 and 2006
This case study presents a course in which the final examination had both a paper-based and a laboratory component. The computer laboratory final examination was worth 25% of the linear model assessment. In 2004 forty-five students sat the final examinations, in 2005 forty-six students and in 2006 forty-four students. The computer laboratory based examination was held in the final week of lectures during the three fifty minute long tutorials. Having access to the internet blocked was not problematic. Students were not given any choice about taking the practical examination.

Practical examination
Students were given a description of the examination procedure two weeks beforehand, explaining the location and requirements for the exam. The IT support people blocked access to the network for the duration of the examinations. No major problems were encountered. Instructions for saving output were the same as the course described above and space was left on the paper to write answers.

Students were allowed to take their lecture notes, hand written summaries, tutorials and tutorial solutions into the examination. Students saved their data to the disk that contained the data.

The written examination also included questions presenting computer output for comment and discussion. There was a moderate positive correlation between student results in the two components of the final ($r = 0.55$ in 2005 and in $r = 0.62$ in 2006), although in the second year in particular there were several students who did much better in the practical component than in the theory. There were no students who did well on the theory but poorly on the practical examination, so it was felt that holding the practical examination allowed some students to demonstrate skills and knowledge that were not reflected in their results in the paper-based examination.

Other case studies
Between 1995 and 2000, the second first year science service course Statistical Techniques 2, taught at the Australian National University (enrolment around 100), had a final examination with a laboratory component. The practical examination was held in more than one room. On each occasion there were several students who were disadvantaged by hardware problems, or issues arose due to slight differences in operation between the rooms. This raised ethical issues which, together with course reorganisation, led to the termination of practical examinations for this large class.

At the University of Technology, Sydney, the fourth year Time Series course was taught eight times with a final practical examination (enrolment around 6). In that period only three students did not cope well. These students all had a background in pure mathematics and were unready for practical data analysis to be the focus of an examination.

Making computer based assessment work

The most obvious way to make this form of computer based assessment work is to make sure that the students become comfortable with using the statistical packages to perform the statistical analyses. The easiest way to do this is to hold all tutorials in the computer laboratory and make most of the

tutorial work requiring the use of a statistical package. Solutions, not necessarily complete, should be provided a week or two after the tutorial so students can see where they have gone wrong. As they advance, students should also become familiar with the idea that there is not necessarily "a single correct answer" for a real-world data set, nor a single path to finding a good model.

Students need some sample mid semester examinations and final examinations to become familiar with what is expected of them. Selected solutions also need to be provided to give them an idea of what is expected of them.

Finding real data sets that could be used was the most time consuming part of these assessment tasks, as there is a constant need for unseen problems to present. A surprising number of examples used in multivariate analysis text books violated model assumptions yet no mention of this was made in the text, just the technique illustrated. Data sets were more readily available for the regression course.

For the mid session test and final examination explicit instructions on how to save output needs to be verbally reinforced before the students commence the test or final examination. From a marking perspective saving each part separately, labelled according to question and part where relevant, facilitated ease of checking that students had used the correct statistical procedure with the right data. Leaving blank spaces on the examination paper for questions requiring written answers also sped up the marking process.

Conclusion

For applied statistics courses with enrolments of less than fifty students computer based assessment can be used to make the learning experience a more enjoyable one and one that they perceive to be more relevant. On a much smaller scale it is good preparation for what they are likely to do if they end up using statistics in their job. It also removes the doubts of the lecturer as to whether or not the student can actually use the statistical package, explore the data set, perform the analysis, check (if appropriate) whether the model assumptions hold, and interpret the output produced.

References

Chance, B.L. (1997) Experiences with Authentic Assessment Techniques in an Introductory Statistics Course, *Journal of Statistics Education* [Online] **5**(3).

Smith, G. (1998) Learning Statistics By Doing Statistics, *Journal of Statistics Education* [Online] **6**(3).