Evaluating Student Performance in Applied Statistics Courses which Emphasise Data Analysis

C Mitchell Dayton - College Park, Maryland, USA

1.   Introduction

In professional fields such as education, psychology, sociology, etc., applied statistics courses emphasise developing skills in planning quantitative research studies, properly analysing data, and correctly interpreting the results of analysis. In general, students lack the background which would be required to deal with mathematical derivations. Furthermore, it is doubtful that this background would materially benefit these students in the professional roles for which they are preparing. The vast majority of researchers in the behavioural sciences are able to conduct their data analyses using the sophisticated statistical packages that are readily available. Thus, it becomes critical that applied statistics courses realistically prepare them for their role as data analysts.

Each academic year, our department enrols about 1000 students in undergraduate and graduate applied statistics courses. Approximately 50% of these students are enrolled in an undergraduate elementary statistics course and the other 50% of the students are enrolled in a series of four graduate courses offered, primarily, for students in the College of Education.

2.   Data sets

All of our applied statistics courses have a planned orientation toward realistic data analysis utilising real-world databases. The characteristics which we consider to be desirable for sets of data used for instructional purposes include the following items.

(i)   The sample size and number of variables should be sufficiently large that students learn to appreciate the data handling capabilities of statistical packages. If, alternatively, the instructor bases problems on the small data sets usually found in textbooks, students may fail to recognise the importance of computer-based analysis.

(ii)   There should be missing values for some of the variables so that students become familiar with the missing-data conventions of the statistical packages being used and so that analytical problems can be realistically treated during the course.

(iii)   There should be a mixture of types of variables. For example, databases with both continuous and categorical variables permit regression modelling with coded variables, etc.

(iv)   It is desirable that some of the variables have outliers or other suspicious values so that students can learn the importance of data screening prior to undertaking preplanned statistical analyses.

In addition to the above characteristics, the databases should have some motivating properties for the type of students taking the courses. Related to this, there should be sufficient diversity among the types of variables that a variety of plausible and interesting research hypotheses can be formulated and evaluated. Some databases which I have found useful in our courses are:

QUALITY - A set of 17 variables abstracted from the 1979 *Information Please Almanac*. The unit is nations of the world and there are 113 cases. The variables comprise demographics, economics, health and education. Two of the demographic variables are almost perfectly collinear and the patterning of missing values results in the loss of about half of the cases if all variables are analysed together.

SESAME - A set of 26 variables for 240 pre-school children involved in a national assessment of the educational television show, *Sesame Street* (abstracted from Glasnapp and Poggio, 1985). The variables include an ability measure, several pre- and post-test achievement measures, measures of viewing behaviour and parental and home environment variables. There are missing data and it is substantial for one of the parental variables.

FAC86 - A set of 13 variables reported for the 140 tenure-track faculty members in our College of Education for the year 1986 (with no individual identifying information, of course). The variables include department affiliation, sex, rank, length-of-service measures, teaching-load measures, and productivity measures. There are missing values for only five of the cases.

GSS84 - A set of 16 variables supplied with the SPSS Studentware microcomputer package. The variables come from a 1984 social science survey and represent nearly 1500 respondents.

3. Statistical packages

Microcomputer statistical packages are used for data analysis in all of our graduate applied statistics courses, as well as in our undergraduate beginning statistics course. Our selection of statistical software has been guided by the principle that students should become familiar with statistical analysis systems which are widely available and used in real-world settings. Thus, mastering the software in our courses can have a broader outcome in terms of future use of the package in the workplace. We have selected software from SPSS for use in all of the graduate applied statistics courses. However, in our undergraduate beginning statistics course, instructors have adopted MYSTAT, which is a limited-feature version of SYSTAT.

4. Student self assessment

In each graduate applied statistics course, we require that students complete a series of data analyses utilising databases of the types described above. These analyses are integrated into the course curriculum and considerable in-class time is devoted to reviewing and explaining components of the computer printouts. Since this review is of greatest value to students when they have successfully completed the analysis, we provide one or more "check values" in connection with each required analysis. Obtaining the correct check value for an assignment provides the student with immediate reinforcement as they interact with the microcomputer program and avoids the frustration of waiting for in-class review in order to determine whether or not some error has been made. After the in-class review, the assignments are collected and scored on a 5-point scale.

A different form of self-assessment is provided by means of a self-testing system which we have developed for use on Macintosh computers. For most topics, we have created a depository of test items from past quizzes administered in these courses. This depository is updated on a regular basis using quiz items from current course offerings. The locally-developed self-testing program allows students to select a topic and then be presented with randomly selected items on that topic. The student's response to an item (which may be multiple choice or completion) is evaluated and feedback is presented concerning the correctness of the response. At the end of the testing period, the student is supplied with a record of their percent of correct responses for each topic. Students with access to Macintosh or Apple II-series computers may copy item disks and programs for use at their own convenience.

5. Formal evaluation

The formal evaluation and grading of students in our graduate applied statistics courses are relatively conventional, although we do emphasise frequent feedback concerning student performance. Thus, in addition to the usual midterm and final examinations, each course has a series of four quizzes spaced out over the semester. Each quiz is relatively short (e.g. 10 completion items) and the time allowed is about 20 minutes. Consistent with our efforts to integrate the interpretation of statistical

analyses of databases into these courses, the quizzes and examinations often contain items based on segments of computer printouts which are supplied to students as part of the tests. Thus, a typical quiz in the Intermediate Statistics course might include five items based on textbook readings and lectures along with five items based on aspects of an SPSS REGRESSION printout.

Since the problem assignments focusing on the analysis of databases consume a significant portion of the students' study and preparation time for our courses, it is appropriate that they receive some weight in the determination of course grades. As mentioned above, we usually grade each of these assignments on a 5-point scale which can be accumulated over the semester. In addition to the actual printout, each assignment asks students to supply certain other written material such as appropriate null hypotheses for statistical tests, hand computation of certain quantities not supplied by the microcomputer program, and interpretation of statistical tests. While virtually all students obtain the correct printout due to the availability of check values for each assignment, there is variability in completion of the additional requirements so that the overall score from the set of assignments is useful in the grading process.

6.   Conclusion

Applied statistics courses for students in professional fields, other than the field of statistics itself, pose a unique set of instructional problems. The design of such courses must recognise the roles which these professionals will assume and should attempt to provide training that emphasises the professional development of the students. Many of these students will become successful and influential researchers who will utilise relatively sophisticated statistical methods. However, because of limited backgrounds in formal mathematics, it is unrealistic to present these students with proofs, derivations or exercises involving elaborate algebraic manipulations. On the other hand, virtually all of the practical aspects of design, analysis and interpretation can be appreciated by these students, especially if the instructional situation is made as realistic as possible. Toward this end, our sequence of graduate applied statistics courses has a strong orientation toward data analysis using relatively large-scale, real-world data sets which have some motivational properties for our students. Data analysis assignments based on these data sets utilising popular microcomputer statistical packages are integrated into the classroom instruction. In addition to conventional testing procedures, each course has significant self-assessment opportunities so that students can be continually aware of their progress in the courses. While test results continue to dominate course grading, a significant amount of weight is given to success on the data analysis assignments.

References

Glasnapp, D R and Poggio, J P (1985) *Essentials of Statistical Analysis for the Behavioral Sciences.* Charles E Merrill Publishing Co., Columbus, Ohio.