QUESTIONING STRATEGIES AND SAMPLE PROBLEMS FOR A COURSE IN STATISTICAL LITERACY *

Eleanor W. Jordan
University of Texas at Austin

ABSTRACT

At the end of a first undergraduate course in statistics, most students and their teacher would agree that the student has had a very limited exposure to a few statistical techniques. At the end of an introductory graduate course or second undergraduate course, the students and teacher may believe that the student has acquired a working knowledge of a number of techniques and literacy in several more. If the teacher has used a typical textbook approach, it is doubtful that this optimistic appraisal of student accomplishment is justified.

Questioning strategies and sample problems are presented in this paper to allow the teacher to test the statistical[1] literacy of their students. Classroom presentation methods and specific library resources are suggested to reduce functional illiteracy. However, the best method of increasing the ability of the student to intelligently read and think about applications of statistics is considered to be well-designed opportunities to make stupid mistakes.

1. INTRODUCTION

Learning, defined as a long lasting change in behavior, is difficult to achieve in a beginning statistics course. If the target behavior is performance in a second statistics course, the standard textbook approach and mathematical lecture method will probably be effective. If the target behavior is improved problem solving, intelligent newspaper or journal reading, thoughtful statistical package computing, or all of the above; heavy reliance on a textbook will probably be ineffective.

The students who account for a large proportion of statistics credit hours are taking a first and last course in statistics. The motivation for taking the course is a requirement for an undergraduate degree in business, psychology, sociology, whatever, or a masters degree in almost any area. The non-major's goal is likely to be survival with a grade that won't dramatically hurt their grade point average. For some students this may be their normal style, but for many their survival attitude is caused by fear of statistics. (In two required undergraduate business statistics sections at U.T. Austin last fall, less than thirty percent of the 79 students reported no fear of statistics early in the semester. I suspect that even fewer education and social science majors enter statistics courses fearlessly.)

[1] *The title in the American Statistical Association 1981 conference program was "Is there a Relationship Between a 'Textbook' Statistics Course and Statistical Literacy?" A reader suggested the title change, since the paper is based on the assumption that the answer to the original title question is negative. The title above better reflects the contents of the paper.

The student who is afraid of statistics is likely to focus on getting the right answer in a very narrow sense. This focus can be a considerable handicap for a teacher who is concerned about statistical literacy '(critical, intelligent reading of reported statistics) or correct use of statistical packages. The problem of different goals of instructors and students -- learning versus survival with an acceptable grade -- can be approached from two directions. One is to try to focus the students' attention more appropriately. The ether is to test in such a way that a good grade directly reflects an understanding of the fundamental nature of statistics.

F'or the non-major the most effective course orientation is on problem solving and decision making rather than on intimidating numbers and their manual generation.

Two recent issues of The American Statistician have included a series of articles on teaching statistics from an applications standpoint (February, 1980) and using computer packages in a beginning course (February, 1979). Therefore coverage in my paper on the issue of course orientation and strategies to focus students' attention_ more broadly will be mainly a brief review of the arguments and suggestions in these articles and other relevant publications. The remainder of my paper will be sample problems and testing strategies for one-semester statistics courses for non-majors.

2. COURSE FOCUS

Frederick Mosteller (1983) argues that the use of class time for presenting both large-scale and small-scale applications is worthwhile because "students like to know that their studies do relate to important activities in the practical world" and the "intriguing applications' maybe all that some of them remember from the course. Oscar Kempthorne (1980) argues that the beginning of statistics must be "in problems, questions, and data" because "statistics is not mathematics" and "the ultimate content must be philosophical."

John Tukey (1980) agrees with Kempthorne about the importance of looking at the questions; he goes further to state that the beginning question is 'NOT...TIDY" and is probably better identified as just an "idea".

If prominent statisticians are so supportive of focusing on questions rather than methodology, why is there still a feeling that the point must be argued? One reason is suggested by Margaret Martin (1981) when she states that "persons trained in statistical methodology think of statistics as a scientific discipline, but most of the rest of the world uses the term to mean orderly arrays of data, an end product of statistical activity." This attitude and the resulting class focus on the scientific discipline work fine for mathematics and statistics majors, but a large proportion of the non-majors will find a numerically oriented class intimidating. To get the typical non-major to listen and make the relevant connections, it is necessary to refer to the logic of a problem.

Martin stresses that the practice of statistics is "an art based on a mixture of intuition, experience, and judgement, as well as scientific evidence of procedures." If the students can be convinced that their intuition and judgment are relevant to statistics, indeed necessary even to understand the numbers in newspapers and annual reports, then they may actually learn something from a single statistics class.

Mosteller (1980) recommends starting with a large-scale application to engage students' interest. Kempthorne (1980) provides a list of questions that he likes. I use questions from Statistics: A Guide to the Unknown (Tanur, 1978) because I can also present the answers--eventually.

A favorite question of mine from the Guide is: "Will reducing the police force cause an increase in crime?" (Press, 1978). Proposition 13 makes this a question of current interest. The question is effective in class because it is difficult to conceive of an experiment that is socially acceptable, ethical, and still provides a nontrivial answer to the question. Once an acceptable experiment is agreed upon, choosing an adequate statistical strategy is relatively easy. Since the Guide is used as a resource and not a student text, the discussion of how to design the study is usually very interesting. Eventually a simplified version of the experimental design is provided for the New York City study in the Guide.

The class reading for the day is then reviewed. The topic could be anything from confidence intervals to seasonal adjustments for this question. The police size problem and the statistical concerns presented in the text are then alternated in an effort to minimize the glazed eye effect that statistics has for many normal people.

Another idea for engaging the attention of reluctant students is to use physical demonstrations. Mosteller (1980) describes several and provides a general guide on how to manage physical demonstrations in class. Two rather intriguing ones are provided by James Hansen (1980) to illustrate t-tests and factorial experiments; one requires a bathroom scale, while the other requires toilet tissue.

Applications and physical demonstrations take time. If statistical literacy is the goal, then discussions of applications are more important than covering a large number of techniques. Still, if one wishes to cover a reasonable number of logical concerns, then replacing manual work with statistical computing is a necessity.

Besides saving time statistical computing is an effective way to increase interest in statistics and give a sense of the reality of practical statistics. If exercises are designed properly, the future manager or newspaper reader will remember well why it is important not to rely too naively on computer generated numbers of a UPI reporter's interpretation of numbers.

An excellent reference on how to use computer packages in a first statistics course is provided by William DuMouchel (1980). My own comments will be on questioning strategies and designing data sets.

3. PROBLEM DESIGN

Inclusion of statistical packages in the very first course may be motivated by the concern that it will be the only chance to prevent alot of foolish statistical computing (Thisted, 1979; Welsch, 1979). If this is the motivation or if literacy is the goal, then the exercises designed for the class must provide sufficient opportunity for foolishness.

3.1 The Matching Trap

One problem with relying on textbook exercises is that they mainly test the student's ability to match a limited number of techniques to a limited number of problems. If the chapter is on testing the relationship of means from two samples, a student who had learned very little statistics can probably figure out when to compute a 2-test and when to compute a t-test. A student who has had fourteen years or more of speeded tests may be very right to apply a matching strategy to save time. But well-designed exercises, at least after the first third of the course, should not be solvable by a matching strategy. Exercises should enforce and reward broader thinking.

The solution to this is fairly simple. if a computer package is provided. If manual work is required, then it is difficult to require enough problems to avoid rewarding matching strategies. Asking thoughtful questions that require choosing strategies, but not completing the arithmetic, is one possibility. Still the generation of numbers that often look so good, even if they are irrelevent, seems much more effective in preventing a matching strategy.

Given that a computer package is provided, there are two prevention strategies. First, ask realistically, messy questions. There are few, if any, managers who could ask: "Is there a statistically significant interaction effect of the magazines in which we advertise and geographical region on the sales of our product?" A realistic question would be "Why are sales down?" Second, provide exercises that require techniques covered earlier in the class. It is horrifying to see students use ANOVA where chi-square is appropriate, but the value of the exercise is proven by the fact that some will use ANOVA just because ANOVA is more recent than chi--square and seems more sophisticated. Just as the education process encourages the use of long words when shorter words will do, later and harder techniques are seen to be better than simpler ones. This occurs even if a little thinking would reveal that the simpler technique is correct.

3.2 Data Enticement

The best way to teach an appreciation of statistics is to have students design and conduct an experiment to answer a question that they themselves pose. Projects are often impossible in a first course because there are too many students or too little time or both. Computer packages and clean data sets make a large number of exercises possible.

In the design of clean data sets, it is important not to have the variables so limited that the student can guess the strategy by looking at the variables. Another data design strategy is to include variables that are related to the question of interest but are not the appropriate type for the most recently presented statistical technique. Including irrelevant variables tests their understanding of the relationship of statistical tools to questions of interest. Including nominal and numeric variables tests their understanding of the relevance of data levels and various statistical assumptions.

The data for a sample problem with extra variables and unsuitable variables is presented in Table 1. This is a revision of a textbook exercise in the multiple regression chapter of Statistical Methods for Business Decisions by Clark and Schkade (1974). The variable of interest is the change in Population for cities in the United States. The sample data includes five possible predictor variables for thirty cities. Each city is numbered 1 through 30, but there is no other identification of the city. The data set includes the city identification number as well as the predictor variables and the change in population. Codes for regions of the U.S. (1=northeast, 2=southeast, .etc.) and an ordinal city size variable-(1=small, 2=medium, and 3=large) were added to the text problem so that it would not be a straight forward regression problem. The class handout for the problem does not specify the level of each variable or whether a linear relationship can be assumed between any predictor variable and the variable of interest, Earlier exams indicated that data level and the concept of linearity were at least memorized. A large portion of students will include all variables-in the regression analysis. The identification codes 1 through 30 are, after all, numbers. Earlier lectures included examples of this fallacy but the idea that all the data provided by the teacher must be relevant overwhelms thinking for some students.

3.3 Regression Without Fear or Thought

Mark Twain (1911) has provided some thoughts on statistics that many students probably cheer:

> In the space of one hundred and seventy-six years the Lower Mississippi has shortened itself two hundred and forty-two miles. That is an average of a trifle over one mile and a third per year. Therefore, any calm person, who is not blind or idiotic, can see that in the old Oolitic Silurian Period, just a million years ago next November, the Lower Mississippi River was upward of one million three hundred thousand miles long, and stuck out over the Gulf of Mexico like a fishing-rod. And by the same token any person can see that seven hundred and forty-two years from now the Lower Mississippi will be only a mile and three-quarters long, and Cairo and New Orleans will have joined their streets together, and be plodding comfortably along under a single mayor and a mutual board of aldermen. There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact.

The class will usually laugh, if not cheer when I read Twain's words. However a number of them are still likely to project a stick across the Gulf if given data that seems to encourage them to do so.

This encouragement comes, in my class, from an exercise that is assigned after multiple regression has been discussed. The data is presented in Table 2. In this exercise, real data has been gathered by an accountant who wishes to estimate the cost underruns of construction projects in a state agency. The accountant believes that a good estimate of underruns will allow the agency to spend a larger percentage of their construction budget, thus reducing the amount of money that must be returned to the state at the end of every fiscal year. The problem sheet and class discussion make it very clear that the foremen who make the estimates are under no pressure to be accurate. The accountant is sure that their main concern is to give themselves plenty of room to complete their projects under the estimate amount. In spite of this knowledge most students will make cost projections that are greater than the budgeted amount. The data "encourages" this, because there is a highly significant trend in the last few years of increasingly high ratios of actual expenses to budgeted expenses. None of the sample projects go over the budget, but the students predict that several new projects will go over the estimated budget amount. A closer examination of the data suggests that the increasingly ratio is an effect of recent inflation rates that foreman have not adjusted for properly.

There are a number of statistical strategies to improve the estimate over simple means, but linear regression is not one of them in this case. The projections made from the regression model are foolish because there is ample reason to believe that the foremen will adjust their estimates for inflation on future jobs. Once again the educational point is to look closely at the question and broadly at the problem, before throwing the data at the computer.

4. COMMENTS

Experience as a statistical computing consultant for graduate students has led me to believe hat the matching strategy, the misuse of regression, and misconceptions about issues as simple as data levels are related to the traditional approach to teaching statistics in second and third courses, not just the first course. If the course is the student's first and statistics, it seems very worthwhile to combine a logical approach to statistics that appeals to the general intelligence and interest of the student with a strategy to demonstrate how wrong the computer (with a little help) can be.

The main justifications given for exercises presented here have been increasing learning and preventing the foolish use of statistical packages. Another potential benefit to the business major or social science major is knowing when to call in a statistician. Spending class time on defining a strategy to answer messy questions should make it clear that the statistician should be involved in designing the study, rather than be called in after the data has already been collected.

Another benefit is that future managers (of marketing or the food stamp program or whatever) should be able to better define their real question. They also may learn that they need to hire another consultant if the statistician they are dealing with speaks only in statistical jargon and doesn't attempt to understand the functional question.

Exercises with messy questions may be valuable to statistics majors as well as non-majors. The educators who claim that the proofs and exercises that they assign adequately teach functional skills can have their students demonstrate this claim by assigning problems of the type discussed in this paper. Stupid mistakes on exercises provide an opportunity for learning that few lectures can surpass.

References

Clark, Charles T. and Schkade, Lawrence L. (1974), Statistical Analysis for Administrative Decisions (2nd ed.) Cincinnati: South-Western.

DuMouchel, William H. (1979) "Comment on Teaching Statistical Computing Using Computer Packages," The American Statistician, 33, 30-31.

Hansen, James L. (1980), "Using Physical Demonstrations When Teaching Data Analysis," Proceedings of the Statistical Education Section, Houston: American Statistical Association, 61-62.

Kempthorne, Oscar (1980), "The Teaching of Statistics: Content Versus Form," The American Statistician, 34, 17-21.

Martin, Margaret E. (1981), "Statistical Practice in Bureaucracies," Journal of the American Statistical Association, 76, 1-8.

Mosteller, Frederick (1980), "Classroom and Platform Performance," The American Statistician, 34, 11-16.

Press, S. James (1978), "Police Manpower Versus Crime," in Tanur, et. al (eds.) Statistics: A Guide to the Unknown (2nd ed.), San Franscisco: Holden-Day.

Tanur, Judith M., Mosteller, Frederick, Kruskal, William H., Link, Richard F., Pieters, Richard S., Rising, Gerald R., and Lehmann, Erich L. (eds.) (1978), Statistics: A Guide to the Unknown (2nd ed.), San Francisco: Holden-Day.

Thisted, Ronald A. (1979) "Teaching Statistical Computing Using Computer Packages," The American Statistician, 33, 27-30 and 35.

Tukey, John W. (1980), "We Need Both Exploratory and Confirmatory," The American Statistician, 34, 23-25.

Twain, Mark (1911), Life on the Mississippi, New York: Harper Brothers

Welsch, Roy. E. (1979), "Comment on Teaching Statistical Computing Using Computer Packages," The American Statistician, 33, 34.

Note: Card decks of the data presented can be obtained from the author. The address is Department of General Business, University of Texas at Austin, Austin, Texas 78712.

[*Note: Tables not attached*]