DEVELOPING CASE STUDIES THROUGH CONSULTANCIES

Stephen Olejnik, University of Georgia, USA

It has often been recommended that when teaching an applied statistical methods course, an instructor use data from actual research studies for both instruction and assignments. Although readily available from several sources, a serious problem with using many data sets is the instructor's limited knowledge of how the data were obtained. That is, the instructor generally lacks an understanding of the research context from which the data were gathered. Consequently, these data simply are numbers to be "crunched". Consultancies, on the other hand, provide a rich source of data that an instructor can use to increase understanding and interest in the analysis process by providing the context from which the data were obtained. The art as well as the science of data analysis can more readily be presented. Three data examples are presented along with student views regarding the analysis of actual data sets. Problems encountered and solutions tried are discussed.

Instructors of applied statistical methods classes often encounter a classroom of students who may be characterized as anxious, unmotivated, poorly prepared, and maybe even a bit hostile. Applied statistics is generally not the number one priority of many students. This tends to be true at both the undergraduate and graduate levels of study. Students enroll in these courses primarily to meet degree requirements, not because of an intrinsic interest in the subject matter. Many do not see the relevancy of the subject and have no intention of ever using what they are taught.

To increase interest and to motivate students, applied statistical methods instructors have tried a variety of techniques including using videos (Moore, 1993), projects (Tanner, 1985), real-world problem solving (Derry, Levin, and Schauble, 1995) and technology to enhance instruction. One suggestion that has been repeatedly advocated for at least two decades has been the use of data sets taken from actual research studies (Becker, 1996; Kempthorne, 1980; Singer and Willett,, 1990). Advocates for the use of real data sets have argued that real data sets can motivate students, excite them with real problems, and provide a more realistic exposure to the world of a scientist. Further, actual data sets can expose students to real problems in data analysis like outliers, missing values, and violations of data assumptions. There is no shortage of real data sets. We live in the information age with data collected, analyzed, and reported in virtually every facet of today's society. Singer and Willett (1990) provide an extensive list of resources from which data sets may be obtained.. Access to these data are frequently available in

libraries, government reports, and of course the internet (see Becker, 1996 for Websites). Data can also be obtained from students in courses through student projects (Stallings, 1993; Thompson, 1994), or classroom demonstrations (Tanner, 1985). But in spite of the abundance of actual data sets, Harwell, Herrick, Curtis, Mundfrom, and Gold (1996) found very few statistical methods texts used in education that actually included realistic examples or problems. The purpose of the present paper is to discuss consultancies as still another data source. Consultancies provide a rich source of data (Wilson, 1992) that have some benefits not available from some other sources.

CONSULTANCIES

My opportunities to work with applied researchers comes from a variety of avenues. As a professor at a research institution there are many opportunities to work with colleagues. As an instructor of applied statistical methods at a doctoral granting institution, service on student research committees is expected. And a third type of consultancy is provided though activities outside the University. In addition to working with school districts, State agencies like Roosevelt Warm Springs Institute for Rehabilitation in Warm Springs Georgia (RWSIR), have frequently sought statistical assistance in their research projects. RWSIR is a comprehensive rehabilitation center in the rural southwestern corner of Georgia. It houses both medical and vocational units that in 1996 provided services to over 3,000 people with disabilities. In the past five years over 100 research projects have been completed at the Institute by staff and visiting scholars. All three types of consultancy provide numerous opportunities for gathering data that can be used as an instructional aid in the applied statistical methods courses that are taught.

STUDIES

Each of the consultancies provide a different type of data base. Consultancies with colleagues tend to be theory driven, based on experimental design, and generally complex. involving multiple outcomes and multiple explanatory variables. Consultancies with doctoral students, on the other hand, tend to be more application oriented, quasi-experimental, and while multiple outcomes are investigated the design is simpler often with a single explanatory variable. Consultancies with RWSIR are workplace problem centered, ex post facto designs, with one or two primary outcomes

and generally a single explanatory variable. Often the data available at RWSIR are past client records collected originally for purposes other than research. Data from student consultancies tend to be the smallest while the RWSIR databases are the largest.

An example of a data set obtained from a colleague involved an investigation of vocabulary acquisition of college students. The theoretical issue focused on the role of two sources of vocabulary development: dictionary use and word context. The researcher was interested in comparing weak and strong dictionary definitions and weak and strong context definitions, as well as the interaction between the two factors. Four ways of "knowing" a word were studied: a multiple choice test identifying the correct definition of the target word; an identification of the correct use of a word; a word application task; and sentence completion task. Each test consisted of ten items. The available data set can be analyzed in several ways. Originally the study was designed as a 2x2 completely randomized factorial design. But it could also be analyzed as a single factor design with four levels - a cell means model. The data could also be analyzed as a mixed model with words as a repeated measures factor and dictionary and context as between groups factors. The later approach has been the recommended strategy in previous vocabulary studies.

An example of a student consultancy is a study investigating the merits of viewing a video on story comprehension when learning a second language. In this study five classrooms participated in the investigation. Two classes (one honors section and one regular section) viewed a movie based on the novel before reading the novel, one class (honors section) read the novel before viewing the movie, and two classes (one honors section and one regular section) read the novel but did not see the movie. Two pre-measures were available and five post-intervention measures were available. The data set can be used for an analysis of covariance and a set of interesting contrasts can be examined as well as a close examination of data assumptions.

An example of a RWSIR consultancy is an investigation on the changing knowledge levels and attitudes of the adult working population toward individuals infected with the HIV virus and AIDS. Over a seven year period the Institute had been providing training sessions for new employees on HIV/AIDS. Before instruction, participants were asked to complete a 20 item knowledge/awareness measure and a scale designed to assess attitude toward individuals with AIDS or infected with the HIV virus. The data provide an application of a single factor ANOVA but contrasts and tests of trend can also be examined.

CASE DISCUSSIONS

Given the context of a research study and the investigators' initial questions the research and analysis process can be simulated. A series of questions can be provided to guide the discussion of a case. While questions vary by case, some example questions include the following: How would you design a study to answer the researchers' questions? What are the strengths and weaknesses of your design? What are some alternatives? What analyses would you recommend? How would they address the researcher's questions? Would you anticipate any difficulties with these analyses? How would you determine whether these difficulties are actually present? How would resolve these difficulties? These questions generally generate responses regarding data collection procedures, sampling, instrumentation, distributional properties, data assumptions, and displaying data. After thinking through the problem, depending on time constraints, either the raw data are provided in file format or the analyses that were actually conducted are shared with the students. Given the output students are asked additional questions such as: How would you interpret the results of these analyses? Are there additional analyses that should be conducted? What would the additional analysis tell you? How do these results relate to the researchers' questions? Finally, students might be asked to write a brief interpretation of the results. Discussions following the analyses often focus on the limitations of the study and suggestions for further inquiry.

COURSES

Data sets like the ones described above have been used in several of the graduate level research courses that I teach in the College of Education. Specifically, I try to use these data and the contexts from which they were obtained in an analysis of variance course, an applied correlation and regression course, and a research design course. Students who enroll in these courses are extremely diverse in their backgrounds and interests. Most come from the College of Education but other Colleges are represented as well. Most students do not have a strong mathematics background, many are part time students. For the most part the students described by Belli and Seaver (1989) reflect my students with the only difference being the classes are offered on campus rather than off campus.

STUDENT VIEWS

A survey of 74 students enrolled in the applied statistical methods course sequence indicated that only 58% of the students had access to their own data and only 54% would prefer analyzing their own data. Over 65% of the students felt that actual data would be more interesting, motivating, informative and fun. But an equal percentage felt that actual data sets would be more complicated and difficult. Faculty research was viewed as being more interesting and student projects the least interesting. Finally, students expressed an interest in the use of hypothetical data sets equal to that of actual data sets. From a student's perspective there is a need for both actual and hypothetical data sets in applied statistical methods courses.

CASE PROBLEMS

A major problem with using consultancy data sets is the amount of time that is needed to present the context. To make the analysis meaningful and interesting requires that students be given sufficient information on the study's background. This background information is, of course, one of the reasons for using consultancies as a data source. In the case of the vocabulary study the theories on vocabulary acquisition must be presented. To reduce this difficulty portions of an article are shared with the students that provide some of the necessary background to understand the variables under investigation.

A second problem with the use of cases is that they may become distracting. Given the background for some studies, students become more interested and involved with the context and not with the analysis. Monitoring discussions can limit the diversions.

A third problem with some cases is that the data can be more complex than appropriate for the students' level of understanding. Non-orthogonal ANOVA, outliers, and violations of data assumptions are often encountered but dealing with these issues must wait until students have an adequate background. Data may be modified to meet the students' backgrounds.

A fourth problem is that not all cases are interesting to all students. The vocabulary study appeals to reading education, educational psychology, and language education students. Other students may not find the case worth pursuing. Multiple data sets may help here.

FINAL COMMENTS

Using real data obtained from consultancies is not always easy, but students generally do enjoy working with them. The real life context of the problems make the study of statistical methods more relevant and interesting but not all students agree. Students also find it interesting to learn about their instructor's involvement in research projects. As an instructor, discussing real studies with students is more interesting than hypothetical studies. Analyzing actual data is always a learning experience and sharing what is learned during consultancies has great rewards.

REFERENCES

Becker, B. J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics , 21*, 71-90.

Belli, G. M. and Seaver, W. L. (1989). Graduate statistics service courses in part-time off-campus programs. *The American Statistician, 43*, 86-90.

Derry, S., Levin, J. and Schauble, L. (1995). Simulating statistical thinking through situated simulations. *Teaching Psychology, 22*, 51-57.

Harwell, M. R., Herrick, M. L., Curtis, D., Mundfrom, D., & Gold, K. (1996). Evaluating statistics texts used in education. *Journal of Educational and Behavioral Statistics, 21*, 3-34.

Kempthorne, O. (1980). The teaching of statistics: Content versus form. *The American Statistician, 34*, 17-21.

Moore, D. S. (1993). The place of video in new styles of teaching and learning statistics. *The American Statistician, 47*, 172-176.

Singer, J. D. and Willett, J. B. (1990). Improving the teaching of applied statistics: Putting data back in data analysis. *The American Statistician, 44*, 223-230.

Stallings, W. M. (1993). Return to our roots: Raising radishes to teach experimental design. *Teaching Psychology, 20*, 165-167.

Tanner, M. A. (1985). The use of investigations in the introductory statistics course. *The American Statistician, 39*, 306-310.

Thompson, W. B. (1994). Making data analysis realistic: Incorporating research into statistics courses. *Teaching Psychology, 21*, 41-43.

Wilson, W. J. (1992). Statistical consulting is scholarship. *The American Statistician, 46*, 295-298.