A CASE STUDY FOR TEACHING BAYESIAN METHODS

Dalene K. Stangl, Institute of Statistics and Decision Sciences,
Box 90251, Duke University, Durham, NC 27708

Key Words: Pedagogy, GUSTO, Role-playing

Abstract

This paper presents a case study that introduces Bayesian inference in introductory undergraduate statistics courses. The case study uses clinical trial data analyzed and published in the New England Journal of Medicine (GUSTO Investigators, 1993) using a frequentist analysis and the Journal of the American Medical Association (Brophy and Joseph, 1995) using a Bayesian analysis. An article from the popular science magazine, Discover (Hively, 1996) is also included in teaching the case. This article discusses, in laymen's terms, the controversy between the frequentist and Bayesian analysis. Prerequisite concepts, where the case-study fits into the course schedule, and classroom exercises that use court-room role playing and health-policy decision-making will be presented. While a calculus-based course makes the case study easier to explain, the case has been taught successfully in non-calculus based introductory courses as well.

Introduction

Over the past 8 years, I have been teaching Bayesian methods in undergraduate service courses, i.e. courses designed for non-statistics majors. At Duke, we've learned that you can't teach primarily the Bayesian paradigm without upsetting other departments. What you can do is present both frequentist and Bayesian paradigms with equal emphasis, and encourage students to ponder the differences. My strategy follows.

The first half of the course looks like many other introductory statistics classes. The textbook is *Statistics* by Freedman, Pisani, and Purves (1998). Students learn the differences between observational studies and controlled experiments, how to describe the distribution of a single variable and the relationship between two variables using graphical and numerical techniques, and how to use the basics of probability. The probability section of the book is supplemented with my own 'souped-up' segments on conditional probability and Bayes theorem.

The second half of the semester does not look like other introductory statistics classes. The topic covered first is frequentist inference, using confidence intervals, $\left(\frac{\text{Observed} - \text{Expected}}{\text{Standard Error}}\right)$ and hypothesis tests in a half-dozen different context. Then Bayesian inference is introduced defining subjective probability and reviewing Bayes theorem. This segment borrows heavily from *Statistics: A Bayesian Perspective* by Berry (1996) and a supplement to the Freedman et al. text written by Michael Lavine and myself. The students see binomial and normal data examples. In the binomial-data examples both discrete and continuous parameter spaces are explored. Conjugate beta priors are used for the latter. Then normal-data examples with a continuous parameter space for the mean are covered. Again conjugate priors are used. Students learn to calculate the posterior and predictive distributions. Throughout, the emphasis is on thinking through Bayes theorem, updating beliefs, and making predictions about future observations.

Time allows 3-4 formal lectures on Bayesian inference followed by 2-3 lecture periods that cover a case study. The utility of using cases in teaching has been argued by many authors. (Barnes, Christiansen and Hansen, 1994, Bryant and Smith, 1995, Chatterjee Handcock and Simonoff, 1995, Nolan and Speed, 1999, and Parr and Smith, 1998) The case presented here involves the GUSTO clinical trial, a trial comparing tissue plasminogen activator (t-PA) and streptokinase (SK) for the treatment of myocardial infarction. The results of the trial were first presented in the *New England Journal of Medicine*, (The Gusto Investigators, 1993) and were subsequently reanalyzed by Brophy and Joseph in the *Journal of the American Medical Association* (1995).

The statistical argument in the *NEJM* paper uses confidence intervals and tests of significance. Finding an increased survival of 1% and rejecting the null hypothesis of no difference between treatments, the GUSTO investigators conclude that t-PA is clinically superior. In the *JAMA* paper, Brophy and Joseph use Bayesian statistical arguments to argue that the jury is still out. They find that the posterior probability that survival on t-PA is greater than survival on SK by at least 1% ranges only from 0% to 36% depending on how much weight is placed on previous trials.

A third source for the case is an article, "The Mathematics of Making up Your Mind", by W. Hively. The article appeared in the popular science magazine *Discover* in May, 1996. It covers the differences between inferential paradigms and highlights the controversies that can arise between them. The article uses the GUSTO trial as their primary example.

After introducing and discussing the case, there are two student exercises both based on role-playing. One is a written exercise, the other a mock legal trial. Students are expected to use the information from the three articles. In the written role-playing exercise students are asked to role-play 3 individuals: 1) a government policy maker deciding whether Medicare will pay for t-PA, the more expensive treatment, 2) an insurance company deciding whether their company will pay for the more expensive drug, and 3) a son/daughter who's parent was given the more expensive drug, and the insurance company is refusing to pay. They must present a written statistical argument (Bayesian or Frequentist) to defend each position.

In the second role-playing exercise, a mock legal trial, students are given roles of plaintiff, defendant, prosecuting attorney, defense attorney, or expert (statistical) witness (one for each side). The case they must act out is a malpractice suit against a doctor who prescribes the cheaper drug (SK) and the patient dies. Students are encouraged to present clear statistical arguments rather than arguments based on emotion, although the latter invariably appear.

Both the written exercise and the mock trial have worked well. The trial works best when you have pre-law and pre-med students in the class. It can be hilarious when you have students with acting experience. Regardless, it is wonderful to see the wheels churn as students sift through the inference issues while presenting their arguments.

The case has been used successfully with undergraduates that do not know calculus as well as those that do. While teaching the former group is more challenging and requires staying with discrete parameter spaces or conjugate analysis, the students have been able to master the conceptual differences between the two paradigms. This mastery not only enriches their learning by teaching Bayesian methods, but their understanding of frequentist methods improves. Students do not leave the class misinterpreting confidence intervals and p-values. This was not the case before I began teaching Bayesian methods alongside frequentist ones.

What does this case study teach the students? Students love what George Cobb calls "authentic play." That is, they love to imitate what they will actually be doing as professionals. This case is an excellent example of authentic play. At the same time it brings to light the advantages and disadvantages of each paradigm. Students learn to make persuasive statistical arguments and are better able to critique others' statistical arguments. Students learn that there are alternative ways of thinking and publishing, and it is their choice. Students learn that statistics are a tool for decision making. Students learn that statistics will be useful for most everything they do and read for the rest of their lives.

References

Albert, J. (1997). Teaching Bayes' rule: A data-oriented approach. *The American Statistician*, 51(3): 247-253.

Barnes, L.B., Christensen, C.R. and Hansen, A.J. (1994), Teaching and the Case Method: Text, Cases, and Readings, Boston: Harvard Business School Press.

Berry, D. A. (1997). Teaching elementary Bayesian statistics with real applications in science. *The American Statistician*, 51(3):241-246.

Berry, D.A. (1996). *Statistics: A Bayesian Perspective*. Duxbury Press.

Brophy J.M., and Joseph, L. (1995). Placing trials in context using Bayesian analysis: Gusto revisited by Reverend Bayes. *Journal of the American Medical Association* **273**(11): 871-875.

Bryant, P.G., and Smith, M.A. (1995), Practical Data Analysis: Case Studies in Business Statistics, vols. I and II, Chicago, IL: Richard D. Irwin, Inc.

Chatterjee, S., Handcock, M.S., and Simonoff, J.S. (1995). A Casebook for A First course in Statistics and Data Analysis, New York: Wiley.

Cobb, G, and Moore, D. (1997), Mathematics, Statistics and Teaching, *Mathematical Monthly*, 104, 9, 801-823.

Freedman, D., Pisani, R., Purves, R. (1997). *Statistics*, 3rd Edition, Norton.

Hively, W. (1996). The mathematics of making up your mind. *Discover*, May: 90-97.

Nolan, D., and Speed, T.P. (1999), Teaching Statistics Theory Through Applications, *The American Statistician*, 53, 370-375.

Parr, W., Smith, Marlene, (1998), Developing Case-Based Business Statistics Courses, *The American Statistician*, 52, 330-337.

Stangl, D., (1998). "Classical and Bayesian Paradigms: Can We Teach Both," In Proceedings of the Fifth International Conference on Teaching Statistics, edited by L. Pereira-Mendoza, L.S. Kea, T.W. Kee, W. Wong, International Statistics Institute, **1**, 251-258.

The GUSTO Investigators, (1993). An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *New England Journal of Medicine* 329:673-682.

Powerpoint slides for teaching this case study are available from the author at dalene@stat.duke.edu.