CLASSICAL AND BAYESIAN PARADIGMS: CAN WE TEACH BOTH?

Dalene K. Stangl, Institute of Statistics and Decision Sciences, Duke University, USA

This paper describes my experiences and attitudes toward teaching both the Classical and Bayesian paradigms in an introductory statistics class. At Duke University, I teach two first-course classes, one at the undergraduate level for social-science majors and one at the graduate level for professional students in public policy. While covering both perspectives is not easy, it is well worth the effort. By introducing the Bayesian paradigm students are better able to interpret an observed significance level correctly. By introducing the Classical paradigm students are able to understand that subjectivity is not reserved for the Bayesian paradigm. Many other advantages that arise from teaching both perspectives will also be discussed.

INTRODUCTION

In a recent paper (Moore, 1997) David Moore argues that it is, at best, premature to teach the ideas and methods of Bayesian inference in a first statistics course for general students. He argues that: 1) Bayesian techniques are little used, 2) Bayesians have not yet agreed on standard approaches to standard problem settings, 3) Bayesian reasoning requires a grasp of conditional probability, a concept confusing to beginners, and 4) an emphasis on Bayesian inference might impede the trend toward experience with real data and a better balance among data analysis, data production, and inference in first statistics courses. These arguments were striking to me, not because I vehemently agreed or disagreed with them, but because they were remarkably similar to the reasons why, for the last five years, I have been slowly incorporating Bayesian inference into my introductory statistics classes.

Moore argues that Bayesian techniques are used little, and he uses the published medical literature and a survey of Department of Energy statisticians as evidence. Since coming to Duke as an assistant professor in 1992, I have collaborated with researchers trained in medicine and the social sciences. These collaborators collect, analyze, and publish their research with very little aid from trained statisticians. In these collaborations I have observed one thing in particular. While the statistical analyses they present in publications is nearly 100% classical, the statistical interpretations made in their day-to-day work is not. In daily conversations, debates, and statistical analyses, they rarely follow classical prescriptions for ‘legitimate’ data analyses or give classical interpretations to their inference. In their day-to-day activities their thinking and the decisions they make based on this thinking are nearly 100% Bayesian. What appears on paper is not indicative of what goes on in their heads. I think it is important that consumers of our statistical methods understand that this discrepancy exists and make a conscious choice to live with the split reality or to work towards congruency between thinking and publication. The choice for my collaborators is an easy one, they will live with the split reality. Resources give them no alternative. This brings me to my students. My work at Duke also requires that I teach students interested primarily in social science who aspire to be the doctors, lawyers, policy makers, and researchers. I want them to be conscious of the split reality too. I take it as part of my job as a responsible educator to do what I can to make sure that they understand both perspectives. My ultimate goal is to teach students to learn to think analytically about applied problems. Teaching standard templates of statistics, whether Classical or Bayesian, is counterproductive to this effort. Teaching them to ask questions from both a classical perspective and from a Bayesian perspective and to examine the differences requires intense analytic thinking. Because conditional probability is one of the most important analytic concepts I teach, and because mastery of this concept is required for interpreting both p-values and posterior probabilities correctly, teaching both classical and Bayesian paradigms gives my students a double dose of this concept and a double dose of analytic thinking. So, how can a professor teach both paradigms in a single semester?

METHODS

Because of the breadth and pace of the course, I use what Schau and Mattern (1997) call a map technique. Students receive a visual aid depicting the topics and interrelation between topics to be covered in the next 14 weeks. Every few weeks we review a low-resolution map to refresh their image of how the details just covered fit into the ‘big-picture’ and we preview a high-resolution map that tells where we will be going in the next few weeks. At the lowest resolution, topics are grouped into 3 categories: 1) descriptive statistics, 2) probability, and 3) inference. During the 14 weeks of class, the proportion of time spent on each of the categories is roughly 3/7, 1/7, and 3/7 respectively. Because it is unsurpassed in cogently presenting the basics of descriptive statistics, probability, and classical inference, I use the book *Statistics* by Freedman, Pisani, and Purves (1997). I supplement the probability section of the book with my own segments on conditional probability and Bayes theorem, and the inference section with my own segments on Bayesian inference that borrow heavily from *Statistics: A Bayesian Perspective* by Berry (1996). Students spend 2.5 hours/week in lecture and 50 minutes/week in a supplemental session used for ‘real-data’ computer assignments and discussion of the use of ‘real’ statistics in articles related to their fields of interest.

The first half of the semester probably looks like an introductory statistics class at most any university, just a slightly faster pace. Students are taught 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 the basics of probability. The pace is quick, but every year more and more students enter this course having seen this material in high school mathematics classes. I find the quicker pace is also made manageable by the vastly improved computing technology of the last few years. The new menu-driven statistical software allows students to begin analyzing data instantly, and the new multi-media teaching tools save an enormous amount of chalkboard time.

The second half of the semester probably does not look like any other introductory statistics class at any university. Using Freedman et al. (1997) I cover classical inference showing students $\left(\frac{\text{Observed} - \text{Expected}}{\text{Standard Error}}\right)$ in a half-dozen different contexts. Then we return to probability and cover subjective probability, conditional probability, and Bayes theorem. Finally, we get to Bayesian inference. I teach two contexts only. First I demonstrate simple binomial-data examples with a discrete parameter space and hence discrete prior, and then I demonstrate simple normal-data examples with a continuous parameter space and a normal prior. In the latter case, they learn to calculate the posterior mean and standard deviation of the population average. The emphasis is on thinking through Bayes theorem, updating beliefs, and making predictions about future observations.

Students must pull together the entire course and compare and contrast the Bayesian and Classical paradigms in three assignments, an oral presentation of a journal article, a written data-analysis project, and a role-playing exercise. The oral presentation requires students to select an article from a list. Popular choices are: “Should Pregnant Women Move? Linking Risks for Birth Defects with Proximity to Toxic Waste Sites”, (Geschwind, 1992), “Lesson Learned From Challenger: A Statistical Perspective”, (Dalal et al., 1989), “Small Cars, Big Cars: What is the Safety Difference?”, (Evans, 1994), "When batterer turns murderer", (Good, 1995), "DNA, Statistics, and the Simpson Case", (Berry, 1994), "How Birth Order Influences Individual Characteristics", (Moore et al., 1995), "Statistical Evidence of Cheating on Multiple-Choice Tests", (Klein, 1992), "Techno-Thriller' Statistics: Chance in the Fiction of Michael Crichton", (Rossman, 1994), and "Who's Number 1 in College Football? ... And How Might We Decide?", (Stern, 1995). During discussion sections students present the article relating it to the concepts learned in class, placing it in context of the Classical or Bayesian paradigm, and challenging, if possible, the statistician's methods.

For the data analysis project, students must find a data set on a topic of personal interest. Most students conduct surveys, surf the internet, or skim through my own collection to find a data set of interest. The data set must contain nominal and continuous variables. The students must analyze the data set demonstrating mastery of each technique (graphical and numerical) learned throughout the semester. They must hand in not only the analysis, but a written description and interpretation of each piece of output for a person who has not yet taken statistics. Their analysis must include at least one question addressed from both the Classical and Bayesian viewpoint.

The role-playing assignment has taken two forms, a written exercise or a mock legal trial, both based on information from the 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 the Classical and Bayesian paradigm and highlights the controversies that can arise in interpretation using the published results from a clinical trial testing the superiority of tissue-plasminogen activator over streptokinase in the treatment of acute myocardial infarction. They are also given excerpts from the original articles that inspired the *Discover* article (Brophy and Joseph, 1995 and The Gusto Investigators, 1993).

In the written exercise, students are asked to role-play 3 individuals: 1) a government policy maker deciding whether Medicare will pay for the more expensive treatment, 2) an insurance company officer deciding whether their company will pay for the more expensive drug, and 3) an individual who is trying to convince the insurance company that they should pay for the more expensive drug. They must present a written statistical argument (Bayesian or Classical) to defend each position. In the mock legal trial, students are given roles of plaintiff, defendant, prosecuting attorney, defense attorney, or expert witness (statistical). The case they must act out is a malpractice suit against a doctor who prescribes the cheaper drug and the patient dies. Both the written exercise and the mock trial have worked well.

What do these final assignments that require students to compare and contrast the Bayesian and Classical paradigms teach the students? These final assignments bring to light the advantages and disadvantages of each inferential paradigm and highlight the nuances that distinguish them. 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 more than hypothesis testing: they are a tool for decision making. Students learn that statistics will be useful in the future not just for testing null hypotheses, but rather 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.

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.

Berry, D. A. (1994). DNA, statistics, and the Simpson case. *Chance, 7(4).*

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.

Dalal, S. R., Fowlkes, E. B., and Hoadley, B. (1989). Lesson learned from challenger: A statistical perspective. *STATS*, ASA, Fall(2), 3-6,20-21.

Evans, L. (1994). Small cars, big cars: What is the safety difference? *Chance*, 7(3):9-16.

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

Geschwind, S. (1992). Should pregnant women move? Linking risks for birth defects with proximity to toxic waste sites. *Chance, 5(3-4),* 40-45,86.

Good, I. J. (1995). When batterer turns murderer. *Nature, 375,* 541.

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

Klein, S. P. (1992). Statistical evidence of cheating on multiple-choice tests. *Chance, 5((3-4),* 23-7.

Moore, D. S. (1997). Bayes for beginners? Some reason to hesitate. *The American Statistician*, *51(3),* 254-261.

Moore, K., Trower, J. and Borowick, K. (1995). How birth order influences individual characteristics. *Chance, 8(4),* 20-4.

Rossman, A.J. (1994). Techno-thriller' statistics: Chance in the fiction of Michael Crichton. *Chance, 7(1),* 32-4.

Schau, C. and Mattern, N. (1997). Use of map techniques in teaching applied statistics courses. *The American Statistician*, 51(2):171-175.

Stern, H. S. (1995). Who's number 1 in college football? ... And how might we decide? *Chance, 8(3),* 7-14.

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