

Using a Sample Survey Project to Assess the Teaching of Statistical Inference

Jim Albert
Bowling Green State University

Journal of Statistics Education v.8, n.1 (2000)

Copyright (c) 2000 by Jim Albert, all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

Key Words: Bayesian inference; Interval estimation; Prior distributions; Proportion inference.

Abstract

This article describes the evaluation of the teaching of statistical inference in a first statistics class. A sample survey project is described as a means of assessing the effectiveness of a Bayesian approach in communicating the basic tenets of inference. There are several advantages of the Bayes viewpoint in performing this survey project, including the explicit modeling of one's prior opinion by means of a probability distribution and the relative ease in reporting statistical conclusions. Some evidence is presented to show that students with sufficient knowledge can accurately specify probability distributions. The success of the survey project is evaluated, and changes to the structure of the project are described that facilitate the interaction of the instructor with the students.

1. Introduction

1.1 An Introductory Statistics Class

1 This article focuses on the introduction of statistical inference from a Bayesian perspective and the evaluation of this teaching of inference in the college-level one-semester introductory statistics class. This particular class at Bowling Green State University is taught to satisfy the mathematics elective requirement for students majoring in the liberal arts. The course is also taken as a recommended course by students in other colleges, such as the College of Health and Human Resources. For most of the students, this is the only course in statistics that they will take. For this audience, the class is designed to introduce the students to the use of statistical reasoning in the real world. The goals of this course are well stated by the participants of a recent workshop on statistical education ([Hogg 1992](#), p. 8):

Our aim in a first course is to develop critical reasoning skills necessary to understand our quantitative world. The focus of the course is the process of learning how to ask appropriate questions, how to collect data effectively, how to summarize and interpret that information, and how to understand the limitations of statistical inferences. Statistical thinking is central to education.

2 To achieve these general goals, the class focuses on the topics of data analysis, the production of data through designed experiments and sample surveys, and drawing conclusions from data. In data analysis, basic methods are presented for displaying and summarizing categorical and measurement data on one or two variables. The

data collection part of the course introduces several good methods of collecting data, such as the simple random sample and the completely randomized designed experiment.

1.2 Statistical Inference: What Do We Want Our Students to Learn?

3 In this "liberal arts" statistics class, it is not necessary to teach particular inferential methods such as a t test or an analysis of variance procedure. However, it is desirable to communicate some basic tenets of statistical inference. It is important that the students distinguish samples and their summaries (statistics) from populations and their summaries (parameters). Students should be made aware of the inherent variability in data, and that sample data provide an incomplete description of a population. The validity of any statistical procedure rests on the assumptions of the underlying model. The data that are collected may not be very informative if the assumptions of the model are violated. For example, there is often little information contained in a nonrandom sample chosen by convenience. If a good sampling process is used, one gets more accurate information from a larger sample.

4 Students should be able to distinguish different types of inference problems. Estimation problems in which one wishes to learn about the value of a parameter are fundamentally different from testing problems in which one has to make a decision about the parameter. The student should understand the correct interpretation of statistical "confidence" which underlies interval estimation and tests of hypotheses.

1.3 A Bayesian Class Using Collaborative Learning

5 The introductory statistics class was taught from a Bayesian viewpoint using [Rossman and Albert \(2000\)](#). The necessary probability background was presented for Bayes' rule, including the subjective interpretation of probability, the interpretation and summarization of probability distributions, and an introduction to conditional probability by means of a two-way probability table. This particular course was activity-driven. The students worked on directed activities from [Rossman and Albert \(2000\)](#) in small groups during class. The instructor's role was primarily to provide assistance to particular students working on the activities during class. Occasionally the instructor in this class would give short lectures on topics that were generally confusing in the working of the activities.

1.4 How Do We Assess Learning of Statistical Inference?

6 If one wishes the students to understand the basic tenets of statistical inference as described above, then it is important to develop assessment methods that specifically test the understanding of the basic concepts. [Hubbard \(1997\)](#) describes the pitfalls of standard assessment procedures that mimic the exercises in many elementary statistics texts. Students may memorize procedures for correctly solving questions of different types without really understanding the underlying statistical concepts. Hubbard describes how one can construct questions that test understanding of concepts. [Garfield \(1994\)](#) discusses the importance of using assessment methods other than traditional tests and quizzes. These methods include projects, portfolios, attitude surveys, written reports, and open-ended questions. The goal of these alternative assessment methods is to improve learning. They can provide information to students on how well they have learned a particular topic and provide diagnostic information to instructors about individual students' understanding of new material.

7 A number of authors, including [Fillebrown \(1994\)](#), [Roberts \(1992\)](#), [Sevin \(1995\)](#), and [Chance \(1997\)](#), describe the use of student-conducted projects in statistics courses. Chance describes a student project as an "authentic" assessment technique that obtains information about students' understanding in contexts that reflect realistic situations. Specifically, Chance outlines the goals of a student project. The project gauges the students' understanding of the entire statistical process (formulation of questions, data collection, description, and analysis), it judges the students' ability to interpret statistical arguments and computer output, it assesses the students' ability to work with others and communicate results, and it increases their interest in statistics.

1.5 Outline

8 This article describes the use of a sample survey project to assess the understanding of statistical inference. Section 2 outlines the Bayesian methodology based on discrete priors that is used to learn about a population proportion. Section 3 discusses the implementation of the project, including the written assignment, the specification of priors, the implementation of the Bayesian methodology, and the interpretation of the interval estimation procedure. In the evaluation of the first year's class, several problems were noted in the completed projects, and Section 3.4 describes changes to the project design for the second year's class that appeared to improve the quality of the final reports. Section 4 summarizes the benefits of the projects in assessing the learning of statistical inference.

2. A Bayesian Approach to Learning About a Proportion

9 The Bayesian approach to statistical inference is based on the subjective notion of probability. One is uncertain about the value of a parameter θ , and one models this uncertainty by means of a subjective probability distribution. This distribution reflects one's personal opinion about the location of the parameter. This probability distribution is called the prior because it reflects a person's knowledge about the location of θ before or prior to collecting data. Once a random sample is taken and data are observed, one's knowledge about the parameter will change. The new probability distribution for θ , called the posterior, is computed by use of a conditional probability computation called Bayes' rule. All statistical inferences about the parameter are based on this probability distribution. For example, a 95% Bayesian interval estimate for θ is an interval that contains 95% of the posterior probability distribution.

10 I describe the Bayesian methodology in the setting of the sample survey project that will be discussed in Section 3. Suppose that a student is interested in the proportion of undergraduate students p who smoke regularly. She takes a random sample of n students by a phone survey and records the number s who smoke; let $f = n - s$ denote the number who don't regularly smoke. She is interested in constructing a 95% interval estimate for the population proportion p .

11 A Bayesian interval estimate for the proportion p can be constructed by use of a discrete prior ([Berry 1995](#), [Albert 1995](#)). Suppose that the student can specify a collection of plausible values of p , call them p_1, \dots, p_k , with respective prior probabilities q_1, \dots, q_k . In teaching it is convenient to start with a set of equally spaced proportion values, say 0, .1, .2, ..., 1, and then give the student guidance on the choice of prior probabilities. One relatively easy method of assigning probabilities is based on relative likelihoods. For example, the student can assign a large value, say 10, to the proportion value that is most likely, assign 5 to values of p that are half as likely as the most likely value, and so on. At the end of the process, the student can divide the assigned numbers by their sum to obtain a proper probability distribution. Using this algorithm, the student can obtain a prior that is a rough approximation to her initial opinion about the location of the population proportion.

12 After this prior has been constructed, data are observed, and the student updates her probabilities by means of Bayes' rule. [Table 1](#) illustrates the calculation of the posterior probabilities. The "Proportion values" line lists the k possible proportion values, and the "Prior" line the corresponding prior probabilities. If s smokers and f nonsmokers are sampled, then the likelihood of a particular proportion value p is $p^s(1 - p)^f$. The likelihoods of the different proportion values are shown in the "Likelihood" row. The posterior probabilities of the proportions are proportional to the respective products of the prior probabilities and likelihoods -- these products are shown in the "Product" row. The posterior probabilities, shown in the "Posterior" row, are obtained by dividing each product by the sum of the products (denoted by SP in Table 1).

Table 1. Illustration of Bayes' Rule Calculations for a Proportion With a Discrete Prior and a Binomial Sample of s Successes and f Failures

					Sum
Proportion values	p_1	p_2	...	p_k	
Prior	q_1	q_2	...	q_k	
Likelihood	$p_1^s (1 - p_1)^f$	$p_2^s (1 - p_2)^f$...	$p_k^s (1 - p_k)^f$	
Product	$q_1 p_1^s (1 - p_1)^f$	$q_2 p_2^s (1 - p_2)^f$...	$q_k p_k^s (1 - p_k)^f$	SP
Posterior	$q_1 p_1^s (1 - p_1)^f / SP$	$q_2 p_2^s (1 - p_2)^f / SP$...	$q_k p_k^s (1 - p_k)^f / SP$	

13 The computation of the posterior probabilities can be tedious for the student, and software is needed to facilitate this task. I have written a Javascript Web-based program http://www.amstat.org/publications/jse/secure/v8n1/p_discrete.html to simplify this calculation. The student enters his or her prior by typing numbers in the "PRIOR" column of a spreadsheet, values of s and f are placed in "DATA" boxes on the web page, and all of the calculations are displayed in the spreadsheet when an "UPDATE" button is pressed.

14 The posterior probabilities can be interpreted as the actual probabilities that the proportion p takes on the k individual values. An interval estimate for p can be constructed by first ordering the values of p by their posterior probabilities, most likely values first, and then putting values of p into a confidence set until the total probability of the set exceeds .95. Suppose that the confidence set turns out to be $S = \{.1, .2, .3, .4\}$. The probability that the proportion of interest p is contained in the set S is at least .95.

3. The Sample Survey Project

3.1 The Assignment

15 In the sample survey project, students were asked to think of two questions of interest to the undergraduate student body, each question having two possible answers, yes or no. The goal of the survey was to learn about population proportions, where the population of interest is the undergraduate student body, and the proportions are the fractions of the students who would answer yes to each of the two questions. To learn about the proportions, the students took a single random sample of at least 40 students and asked each sampled student the two questions. Based on the observed data, the students were asked to construct interval estimates for each of the two proportions of interest and to interpret what they learned from this survey.

16 The students worked on these projects in groups of size three or smaller. A major component of this project was the writing of a final report that gave all details of the statistical investigation. Specific guidelines were given on the contents of this report. The project report was to contain the following components:

1. **Introduction.** The students were to describe how they decided on their particular inference problems. What motivated them to choose the particular two questions? Were they based on some personal experience, something they had read in the newspaper, or something else?
2. **Prior opinion.** Before they took their sample, what did the students think they would find out? Could the students make an intelligent guess at each population proportion value? If there were differences in

opinion between members of the group, they were asked to discuss those differences and the rationale for each opinion.

3. **Sampling.** The students were supposed to explain in detail how they took their random sample. What particular sampling method was used and how many students were called? Were there any difficulties encountered in collecting the data?
4. **The data.** The results of the survey should be explained in detail. Specifically, students were asked to list the data in a raw form that displayed the responses to each question for each student in the sample.
5. **The inference.** The students were asked to write down all of the work involved in constructing the 95% probability intervals. This included a reference (page number in the book) for the methodology that was used and all of the details of the computation.
6. **Interpretation.** The students were asked to explain what 95% confidence means in the context of their individual survey problems. This interpretation should be made using language understandable by a layman who is not familiar with statistics jargon.
7. **Wrap-up.** At the end of the report, the students were asked to discuss what they learned from their project. Were the results surprising? Were any problems encountered in doing the project? If the students had to do the project again, what would be done differently?

3.2 Project Results

17 A total of 12 projects were completed by the student groups. I first summarize the contents of the project reports. I next discuss the students' specifications of their priors and their interpretations of their interval estimates. The results of the projects motivated changes to the project implementation that were followed in the second year.

3.2.1 Questions Posed

18 Generally, the students asked survey questions on topics that interested them. Some representative survey questions were:

- Do you regularly attend university sporting events?
- Do you own a car?
- Do you drink regularly on weekends?
- Do you enjoy living on campus?
- Do you have a TV in your room?

Note that these questions are distinctively different from many sample survey questions that are posed in elementary statistics textbooks. Most of the questions deal with issues in the students' daily lives, including their social behavior, their living conditions, what they do on weekends, and so on. There were very few questions on political issues, such as their feelings about abortion or their agreement with the policies of the President of the United States. The important thing to note is that the students asked questions on subjects that genuinely interested them -- they wanted to learn about the feelings of fellow students on these subjects.

3.2.2 Prior Opinions

19 One distinctive feature of the Bayesian viewpoint is that the prior probability distribution provides a mechanism for the student to input his or her opinion about the proportion value. Initially the students expressed an opinion by means of a guess at the proportion value. A statement such as

"We assumed that around 30% of students had cars on campus"

was typical. An alternative and equally valid method of expressing this prior opinion guessed at the number of students in the sample who would say yes to the particular question:

"We estimated that about 36 out of our 40 sampled people would say yes for the weekend."

This statement indicates a preference for making a prediction of the future sample result s instead of guessing at the population parameter value p . Suppose a student wishes to learn about the proportion of students on campus who own cars. It can be difficult for the student to directly guess at the proportion of all students who own cars since this population parameter is an abstract characteristic of a model whose value will never be known with certainty. In contrast, it can be easier to predict the number of students owning cars in a future sample of size 40. This sample result is a more tangible quantity whose value will be known once the survey is completed.

20 After a guess at the proportion value was made, the students followed the general advice on constructing a prior provided in class. They set up a grid of values for p and assigned the largest prior probability to their best guess of this proportion. Most of the prior distributions constructed were relatively diffused about the most likely value. These distributions possibly reflected the students' lack of knowledge about the results of their survey. More likely, since the construction of a prior distribution is not easy to perform in practice, the students may have imitated examples of prior distributions that were discussed in class.

3.2.3 Can the Students Accurately Specify Priors?

21 One criticism of the Bayesian viewpoint is the subjective nature of the inference. The posterior distribution depends on both the data and the prior which reflects one's belief about the location of the proportion. If a student does a poor job in specifying the prior, that is, the prior does not accurately reflect his or her subjective beliefs, then the inference based on the posterior distribution may be suspect.

22 That raises the interesting question: are students able to accurately specify priors for unknown quantities about which they have a good amount of information? A simple experiment was performed to address this question. The first item on their final test asked each student to give a probability distribution (see Table 2) for his or her letter grade on the test.

Table 2. Probability Distribution Table for Students to Fill Out Regarding Their Grade on Their Final Test

GRADE	A	B	C	D	F
PROBABILITY					

23 The students had not yet been tested on the material on the final test. However, each student probably had a good idea how he or she was performing in the class and likely could make an educated guess at his or her grade on this test. Is a student able to specify a probability distribution on the grades A, B, C, D, F that reflects this knowledge?

24 It was found that the students appeared to give reasonable prior distributions on the letter grades. Most of the students assigned their most likely grade a probability value between .4 and .6, and then spread out the remaining probability on neighboring grades. All of the students specified proper probability distributions where the values were nonnegative and summed to one. Two students assigned the "unreasonable" uniform probability distribution where each grade was assigned a probability of .2.

25 One way of assessing the accuracy of the students' probability distributions is to relate them to the actual scores on this final test. For each student's prior, a mean prior grade was computed by assigning the usual four-point scale to the letter grades ($A = 4$, $F = 0$), and then computing the mean of the probability distribution of the scores.

26 [Figure 1](#) displays a scatterplot of the students' actual test grades and their mean prior grades. One sees a strong positive trend -- if one removes the two outlying points which correspond to the uniform priors, the correlation between the actual and prior mean grades is .875. The fitted regression line to the points (with the two outliers removed) is

$$\text{ACTUAL_GRADE} = 11.3 + 22.5 \times \text{PRIOR_GRADE}$$

The students, on average, were on target in predicting their test grades. Using the above regression line, a student with a prior grade of 3.5 would receive a test grade of 90, a student with a prior grade of 2.5 would receive a test grade of 68, and so on.



Figure 1. Scatterplot of the Students' Actual Test Grades and Their Mean Prior Grades.

27 The results of this experiment are encouraging. The students appear able to specify reasonably accurate probability distributions for an unknown quantity (their grade) about which they know something.

3.2.4 The Impact of the Prior on the Posterior Inference

28 The previous section questioned the accuracy of the priors specified by the students. This raises a second issue. Given that students will possess different prior beliefs about the proportion of interest and therefore specify different prior distributions, how sensitive is the posterior inference to the choice of prior?

29 I address this concern by focusing on one survey project. One group was interested in learning about the proportion of campus students who believed they got enough sleep at night. Three students, Arica, Sara, and Caroline, each specified a prior for this proportion over the values $p = 0, .1, .2, \dots, 1$. Looking at [Table 3](#), one sees different prior beliefs reflected in these distributions. Arica's prior is quite spread out with a mode of .4, Sara's prior is less diffuse with a mode of .3, and Caroline places a large probability on the value $p = .2$. The group took a sample of 40 students, and 19 students said that they got enough sleep. The posterior probabilities for each of the three students are also shown in the table. Although the students had substantially different prior beliefs, note that their posterior distributions are very similar. For each student, most of the probability (82-88%) is concentrated on the values $p = .4, .5$. This example illustrates that the posterior distribution is relatively insensitive to the choice of prior. Moreover, it illustrates the convergence of beliefs of people when they have obtained more information.

Table 3. Prior and Posterior Probabilities for Proportion p for Three Students

Prior				Posterior			
p	Arica	Sara	Caroline	p	Arica	Sara	Caroline

.0	.000	.000	.000	.0	.000	.000	.000
.1	.075	.133	.125	.1	.000	.000	.000
.2	.100	.267	.450	.2	.000	.001	.002
.3	.150	.333	.250	.3	.032	.119	.126
.4	.250	.200	.100	.4	.499	.664	.468
.5	.125	.033	.050	.5	.376	.167	.353
.6	.100	.033	.025	.6	.089	.049	.052
.7	.075	.000	.000	.7	.003	.000	.000
.8	.050	.000	.000	.8	.000	.000	.000
.9	.050	.000	.000	.9	.000	.000	.000
1.0	.025	.000	.000	1.0	.000	.000	.000

3.2.5 Sampling

30 Students may have learned the most in the sampling phase of the project. Although they had been exposed to the notion of a random sample earlier in the course, they had not really thought about drawing such a sample in a real situation. Some students thought that various types of convenience samples were reasonable to use, and this thinking motivated a class discussion of the types of biases that could be introduced by the use of these nonrandom samples.

3.2.6 Inference and Interpretation

31 For most of the student groups, the actual implementation of the statistical methodology was relatively easy. Only a few groups calculated the interval incorrectly. Currently my students use a Javascript program http://www.amstat.org/publications/jse/secure/v8n1/p_discrete.html to compute the posterior probabilities.

32 An important component of the project report was the interpretation of the survey results. Although the Bayesian interval estimate was easy to compute, it was more difficult to provide the correct interpretation of the procedure, especially using non-statistical language that could be understood by the person on the street.

33 What is the correct interpretation of 95% confidence from a Bayesian viewpoint? If the computed interval estimate for p is $(.23, .42)$, then it is correct to say, "The probability that the proportion p is contained in $(.23, .42)$ is .95." The interpretations from ten of the projects are listed below.

1. The estimates that we were given tell us that 95% of the undergraduates surveyed fall between the probabilities of 50% and 60% when answering the first question. It was more likely for 50-60% percent of the people interviewed to enjoy going out on the weekends. In other words, if you asked forty people if they liked to go out on the weekends or weekdays, 20-24 of them would answer the weekends.
2. The 95% interval was $(.2, .7)$ which means that between 20% and 70% of the students fall in the category of living more than 100 miles from campus and go home more than twice a semester.
3. 95% of the data plus or minus .05 will lie between .033-.029.

4. 95% means that the probability that the proportion of students who smoke at least once a day is in the set of $p = .2, .3, .4$ is 95% -- a very very good chance. In other words, there is a 95% chance that $p = .2, .3$ or $.4$. That means that it is very unlikely that the p value would be anything other than $.2, .3, .4$.
5. The 95% interval means that 95% of the sample lies between $.4$ and $.6$ for the question of students going home on the weekends and for the question of not enough weekend activities, the sample results lie between $.6$ and $.8$.
6. What the 95% means is that of our prior data, we were correct to 95% between $.1$ and $.3$. Basically, that between these two percentiles is where you will find the most girls on campus with boyfriends.
7. The 95% interval estimate for question 1 says that the answer is between $.4$ and $.6$. That means that there is a 95% chance that between 40 to 60 percent of people prefer basketball to football.
8. The 95% interval estimate means that between the priors of $.1$ -. $.5$ is the set of proportion values for the probability set. That means that the proportion is in this set of 95%.
9. The 95% interval means that .95 of the response fell between $.3$ and $.5$, which is also where the majority is located on the graph. So the majority of undergraduates do not skip class regularly and do not stay out late during the week.
10. 95% of the probabilities fall between $.7$ and $.9$. This means that 70 to 90% of the students sampled think that the food on campus is overpriced.

34 Projects numbered 4, 7, and 10 gave essentially correct interpretations -- they all talked about the probability of the proportion p falling in the confidence set. A common mistake (responses 3, 5, 9) was to confuse probabilities of the proportion p with probabilities of different data results. For example, project 5 stated that "95% of the sample lies between $.4$ and $.6$ " and project 9 said that ".95 of the response fell between $.3$ and $.5$." These responses are incorrect because one knows with certainty the proportion of yes's and no's in the sample. Project 1 incorrectly stated that 95% of the population falls within the interval limits. The remaining three responses were incomplete but harder to classify. Responses 2, 6, and 8 made statements about the interval estimate that don't refer to probabilities. For example, it is difficult to understand what project 6 meant by "correct to 95% between $.1$ and $.3$."

3.3 Evaluation -- How Were the Projects Deficient?

35 Although the survey project was generally successful, there were particular components of the project that needed improvement.

1. **Priors.** The construction of the prior distribution is likely a difficult task for most of the students. It was not clear that each student formed his or her own prior -- the prior that a group used for their project may have been constructed by a single student. Moreover, from reading the project reports, it appeared that the construction of the prior and the data may have been confounded. The prior should reflect one's prior beliefs *before* any data has been collected. If the students worked on their project the last week before the assignment date, they might have simply based their prior beliefs on some of the sample survey results.
2. **Sampling.** Although the students were taught basic principles of random sampling, they were not taught specific methods for selecting a random sample from the student body. The standard method of taking a simple random sample by means of a random digit table was not applicable for their problems, and alternative sampling schemes were required. It was unclear from reading the project reports if valid samples were indeed taken.
3. **Interpretation.** As demonstrated above, the students had difficulties in correctly interpreting their interval estimates. It was clear that there should be more instruction on the meaning of a Bayesian probability

interval.

4. **Writing of report.** The project reports were generally too brief. The students often did not explain why they were interested in their particular survey questions or how the prior distribution reflected their prior beliefs. Some of the students seemed to think that the numerical work (the displaying of the prior and posterior probabilities and the computation of the interval estimate) "spoke for itself." Some of the reports were deficient in explaining how the students' prior beliefs were changed by the data and what they learned from the survey project.

3.4 Revision of the Project Assignment

36 In response to the concerns stated above, changes were made to the project assignment in the following year's class. (The revised project assignment, including a prior distribution worksheet, can be found at http://www.amstat.org/publications/jse/secure/v8n1/project_assignment.html.) The project was divided into three stages, and each group had to turn in a written report when each stage was completed.

37 The first stage of the report included the choice of survey questions and the construction of the prior probability distributions. Every student was asked to complete a "[Prior Distribution Worksheet](#)" for each survey question. This worksheet included blanks for the question to be asked, the meaning of the proportion p in this setting, the student's guess at the value of p , and a table of prior probabilities. This report also included a description of the proposed methodology for selecting a random sample from the student body.

38 The second stage of the report described the "data phase" of the investigation. In this report, the individual survey responses would be listed, and the group would summarize these data. This summary would include the proportions of students who answered "yes" and "no" to each question and bar charts that displayed these proportions.

39 The final stage of the report focused on the statistical inference. The projects would contain a description of the calculation of the interval estimates, an interpretation of these intervals, and a summary of what the students learned in this project.

40 The main rationale behind these stages was to give the instructor more opportunity to evaluate the students' progress on the project. Since there were multiple opportunities for feedback, the instructor would be able to suggest changes that would help in completing a successful project. The instructor could suggest the rewording of a confusing survey question, question the construction of a prior distribution, or criticize the sampling methodology. These interactions helped to reinforce ideas in probability and statistics that were discussed in class.

3.5 The Projects in the Second Year

41 The students in the second year's class generally did well on their survey projects. Due to the inclusion of the prior worksheets, there was more evidence that the students worked on constructing their prior distributions. There was more care in selecting their random samples. In addition, due to the class discussion time, the students were more successful in correctly interpreting their interval estimates.

42 The following interpretation was provided by a group that was interested in the proportion of campus students who wore a watch:

The 70% probability set for question one fell between .7 and .8. This means that there is a 70% chance that between 70-80% of the undergraduate students wear a watch.

Of the eleven group project reports, nine included an interpretation similar to the one above. In each of these nine reports, there was a statement on the chance or probability that the computed interval contained a

proportion. However, only four of these nine reports explicitly referred to the proportion of the population of undergraduate students. Three reports talked vaguely about the "proportion of people," one report incorrectly referred to the proportion of survey participants, and one report didn't explicitly say what proportion they were talking about. So, although most all of the reports used the right phrases in interpreting the probability interval, it is not clear that all of the students understood that they were constructing an interval for a population proportion.

43 A few groups went beyond the basic interpretation of the probability interval. These groups were able to discuss what they learned about the proportion of interest by comparing the prior and posterior distributions. One group reported:

This interval set (.5 to .6) is lower than our prior probability range. Originally, we had the set of $p = .4, .5, .6$, and $.7$ to be the 90% interval set. However, the actual probability set has a shorter range with lesser possibilities.

This group observed the basic fact that, as one collects more data, the length of the probability interval for the proportion will decrease. The Bayesian approach allows one to see the impact of the data by comparing the prior and posterior probability distributions. For future classes it would be beneficial to spend more class time on the comparison of probability distributions so that students can relate the prior to the posterior in their project write-ups. It would be instructive, for example, to have the students compare two posterior distributions -- one for the full dataset and one for half of the dataset.

44 Generally, the quality of the projects improved in the second year's class. However, based on the above analysis, there is a concern that the students were relying too much on the instructor's guidance in completing the reports. There was additional time devoted in class to the proper interpretation of the confidence intervals, and some groups may have mimicked the instructor's wording of "confidence" without really understanding the interpretation of a posterior probability distribution.

4. Concluding Remarks

45 The sample survey project used in this Bayesian class has several desirable features. First, the students actually perform a complete statistical analysis on problems of genuine interest to them. It is easy to spend more time on the relatively glamorous topics of data analysis and inference without discussing the motivation for the particular statistical study or how the data were collected. Often statistical studies are useless because the wrong questions were asked or the data were collected by means of a nonrandom sample or poor experimental design. The students begin to recognize these problems when they do their own study. Second, the prior is a useful construct to get the students thinking about the parameter of interest. Believe it or not, students in an elementary statistics class often do not understand the difference between a parameter and a statistic, and the Bayesian prior helps in overcoming this confusion. Last, the Bayesian interpretations of statistical conclusions are somewhat intuitive. The Bayesian student can talk about the probability that his interval contains the unknown proportion -- this is a valid Bayesian statement because the parameter is a random variable.

46 However, this article has demonstrated some potential problems in the implementation of this project. The students can be careless in specifying their priors, and they can make mistakes in taking their sample survey or interpreting their inferential results. In the second year's class, changes were made to the assignment that enabled the instructor to provide constructive feedback on the different aspects of the project. This feedback helped to reinforce many of the key ideas in the course and also helped the students to write more complete reports. It seems that a continuous dialog between the instructor and students is necessary to ensure the success of the project.

Acknowledgments

Partial support for this research was provided by the National Science Foundation's Division of Undergraduate Education through grant DUE #9752428.

References

- Albert, J. (1995), "[Teaching Inference About Proportions Using Bayes and Discrete Models](http://www.amstat.org/publications/jse/v3n3/albert.html)," *Journal of Statistics Education* [Online], 3(3). (<http://www.amstat.org/publications/jse/v3n3/albert.html>)
- Berry, D. A. (1995), *Basic Statistics: A Bayesian Perspective*, Belmont, CA: Wadsworth.
- Chance, B. (1997), "[Experiences With Authentic Assessment Techniques in an Introductory Statistics Course](http://www.amstat.org/publications/jse/v5n3/chance.html)," *Journal of Statistics Education* [Online], 5(3). (<http://www.amstat.org/publications/jse/v5n3/chance.html>)
- Fillebrown, S. (1994), "[Using Projects in an Elementary Statistics Course for Non-Science Majors](http://www.amstat.org/publications/jse/v2n2/fillebrown.html)," *Journal of Statistics Education* [Online], 2(2). (<http://www.amstat.org/publications/jse/v2n2/fillebrown.html>)
- Garfield, J. B. (1994), "[Beyond Testing and Grading: Using Assessment to Improve Student Learning](http://www.amstat.org/publications/jse/v2n1/garfield.html)," *Journal of Statistics Education* [Online], 2(1). (<http://www.amstat.org/publications/jse/v2n1/garfield.html>)
- Hogg, R. V. (1992), "Towards Lean and Lively Courses in Statistics," in *Statistics for the Twenty-First Century*, eds. F. Gordon and S. Gordon, MAA Notes No. 26, Washington, DC: Mathematical Association of America, pp. 3-13.
- Hubbard, R. (1997), "[Assessment and the Process of Learning Statistics](http://www.amstat.org/publications/jse/v5n1/hubbard.html)," *Journal of Statistics Education* [Online], 5(1). (<http://www.amstat.org/publications/jse/v5n1/hubbard.html>)
- Roberts, H. V. (1992), "Student-Conducted Projects in Introductory Statistics Courses," in *Statistics for the Twenty-First Century*, eds. F. Gordon and S. Gordon, MAA Notes No. 26, Washington, DC: Mathematical Association of America, pp. 109-121.
- Rossman, A., and Albert, J. (2000), *An Introduction to Statistics: Data, Probability, and Learning from Data*, New York: Springer-Verlag.
- Sevin, A. (1995), "Some Tips for Helping Students in Introductory Statistics Classes Carry Out Successful Data Analysis Projects," in *Proceedings of the Section on Statistical Education, American Statistical Association*, pp. 159-164.
-

Jim Albert
Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, OH 43403

albert@bgnet.bgsu.edu

[JSE Homepage](#) | [Subscription Information](#) | [Current Issue](#) | [JSE Archive \(1993-1998\)](#) | [Data Archive](#) | [Index](#) | [Search JSE](#) | [JSE Information Service](#) | [Editorial Board](#) | [Information for Authors](#) | [Contact JSE](#) | [ASA Publications](#)