

Statistics in a Nutshell?

Linda A. Tappin
Montclair State University

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

Copyright (c) 2000 by Linda A. Tappin, 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: Binary data; Deductive reasoning; Inductive reasoning; Probability vs. statistics; Statistical thinking.

Abstract

The paper reports on a two-year investigation into the feasibility of allocating three weeks of an undergraduate calculus-based probability course to statistics. This brief introduction to statistics would take the place of a course, thus constituting the students' only exposure to statistical science. At first glance, the request seemed quite reasonable. Statistical inference is based on probability, and statistical inference could be presented as an application of probability. Besides introducing some statistical concepts, it was hoped to enhance understanding of probability by highlighting this connection. However, it was not possible for the students to learn anything meaningful about statistical science in three weeks. In addition, any enhancements to the learning of probability were not significant enough to warrant the omission of material from that course.

1. The Challenge

1 Consider the following challenge: Can you include a three-week introduction to statistical science in an undergraduate calculus-based probability course? An external review board had recommended that our computer science curriculum include statistics, yet the course requirements of the program did not allow time for a statistics course. The mathematics requirements for the computer science program were calculus I and II, followed by probability and linear algebra, with calculus II the prerequisite for the probability course. Since probability was required for our statistical methods course, it was not possible to add the statistics course to the curriculum without adding three credits to the load. One solution under consideration was to add some statistical content to the probability course. Over a period of four semesters, I had the opportunity to evaluate the feasibility of this request, and if feasible, to recommend how the probability course could be modified.

2 At first glance, the request seemed quite reasonable. After all, statistical inference is based on probability. Topics in statistical inference could be included as applications of the probability theory, thus enhancing appreciation of the probability topics.

3 This task was not simple, but the exercise of considering the possibility has been extremely valuable and can be enlightening for all teachers of statistics and probability. Some questions to address are "Where does probability end and statistics begin?," "How does thinking differ in probability and statistics?," "What aspects of statistics should be covered?," "How should the material be presented?," and "Is it reasonable to expect to present any meaningful introduction to statistics in three weeks?"

4 Adding three weeks of statistics to the probability course necessitated omitting other material. I omitted the unit on bivariate distributions, the final unit of the course. This omission could be justified if the time spent highlighting the differences between statistics and probability served to enhance understanding of the probability.

2. Probability vs. Statistics

5 Probabilistic models relate to phenomena with random or chance outcomes. In introductory probability, we study the properties of certain of these models. Here we assume that the model is completely specified, in the sense that all parameter values are known explicitly. In statistics, on the other hand, while we consider a certain underlying model, we do not assume that all parameter values are known, and we use data to estimate or test hypotheses about them. Traditionally, we study the underlying probabilistic models and their characteristics before learning how to use data to make inferences about them.

6 [Scheaffer \(1995\)](#) distinguishes the reasoning involved in traditional probability from the reasoning involved in applications of probability to statistical science by describing the following two "uses" of probability theory: "Both involve an underlying probabilistic model, but the first hypothesizes a model and then uses this model for practical purposes, whereas the second deals with the more basic question of whether the hypothesized model is in fact a correct one ... Based on the sample data, inferences can be drawn about the nature of the underlying probabilistic mechanism -- a type of application known as *Statistical inference*" (pp. 3, 4). [Cobb and Moore \(1997\)](#) note that "probability provides the chance models that describe the variability in observed data" (p. 820).

7 Our probability course is a typical course at the undergraduate level, beginning with counting rules and the axiomatic definition of probability and covering the common discrete and continuous probability distributions and expectation. The final units usually include bivariate probability distributions and the central limit theorem. Simulation is used whenever feasible to illustrate the random nature of the phenomena being modeled and the variability inherent in these models.

8 Undergraduate students taking a probability course think in much the same way they have in all of the mathematics courses prior to probability: they use deductive reasoning. They prove theorems, make calculations, and draw conclusions based on a given underlying probability distribution.

9 In statistical inference, on the other hand, students use inductive reasoning, generalizing from the observed information in the sample to uncertain conclusions regarding the population or process. [Moore \(1992\)](#) states that "the subject matter of statistics is reasoning from uncertain empirical data" (p. 15). Statistics is usually the first mathematics course where students are required to use inductive reasoning extensively, and they find this to be a very difficult transition. Students like to plug numbers into formulas, and they like certainty in their answers. Interestingly, at the beginning of an introductory statistics course, the students with more mathematical training and experience with deductive reasoning have a more difficult time with this transition than students with a more limited background in mathematics.

10 This simplified view is not meant to suggest that students learn probability in exactly the same way that they learned calculus. Discussions of the importance of learning to think probabilistically and the difficulties in teaching students to do so can be found in [Falk and Konold \(1992\)](#), [Konold \(1995\)](#) and [Pfannkuch and Brown \(1996\)](#). I emphasize inductive vs. deductive reasoning to highlight the difficult transition from thinking probabilistically to thinking statistically, a transition that usually takes two to three weeks at the beginning of an introductory statistics course. The challenge here is that we had only three weeks available for the entire statistics unit.

3. First Attempts

11 Our external review board did not specify what it meant by "statistics," and perhaps it would have been satisfied with evidence that the students had learned how to calculate sample means and variances. However, this

approach would in no way reflect the spirit of statistical science, nor would it enhance understanding of the probability being studied.

12 The current trend in statistics education is to first cover exploratory data analysis (EDA) and data production, and then follow with the more formal statistical inference. See the discussion in [Cobb and Moore \(1997\)](#). EDA would lend itself well as a brief introduction to statistical science. There is no underlying theory necessary, much of the work is graphical, and students can very quickly begin making discoveries with their data. We did not take this approach because of the time required to resolve data management and technology issues. It was also felt that the statistical content should include some calculation of probabilities, to allow a smooth transition between the two subjects. The most logical approach seemed to be to present statistical inference as an application of the probability models already in the course. It is in this area where the link between probability and statistical science is the strongest.

13 My first attempt at revising the course was to introduce the sampling distribution of the mean and significance testing for a single mean as the introduction to statistics. Because the normal distribution was the last distribution studied, this unit appeared at the end of the course. The topics were presented as applications of the normal distribution, with p -values calculated as normal probabilities. The mean and variance of the sampling distribution were derived using the properties of expectation and independence. The simplest approach was to work with small samples assumed to come from a normal distribution with known variance. Although we studied the central limit theorem, working with large samples added complications to the statistics unit. The data handling effort was greater, and time had to be spent discussing when a sample is "large enough" to assume normality. I also did not want to have to estimate the standard deviation from the data. With such a limited time frame, I wanted to focus attention on the ideas of inference, and not on the properties of sampling distributions other than the normal distribution. Simulation was done to illustrate these ideas, and significance testing was restricted to one-sided tests of the mean.

14 This approach yielded disappointing results. Most students could calculate a p -value but could not interpret it. In a typical assignment, students might be given a small dataset and asked to calculate the p -value and use it to reach a conclusion as to whether or not the data suggested that the population mean was significantly larger than some value. Although most students could calculate p , they did not know how to use it to reach a conclusion. Even the better students admitted that they would guess the answer, or they reasoned that if the sample mean were larger, then the parameter must be significantly larger than the hypothesized value. They had no notion of how this question related to the sampling distribution of the mean or to the p -value.

15 The students were discouraged and frustrated, as they realized that I was not satisfied with the results. I thought the difficulty might be due to the abstract nature of significance testing, so I repeated this approach during a subsequent semester using confidence intervals. The idea of obtaining a "ballpark" estimate of the population mean is a lot easier to motivate. This is more intuitive than asking students to assess whether or not the evidence in a sample is sufficient to reject a certain null hypothesis regarding a population mean. Students were able to calculate a confidence interval, but they really did not understand what it represented. They could memorize the statement, "I am 95% confident that the population mean lies between a and b ." Or they could state that " c " is a viable estimate for the population mean because $a < c < b$. Morale was better when working with confidence intervals, and there were more correct responses on exams, but it did not seem that the students had any deeper understanding of confidence intervals than of p -values. Few students seemed to understand the difference between the sample and the population, or that their intervals represented estimates of a population mean.

16 These difficulties are no different from those frequently encountered by students in a general statistics course when they are first introduced to the abstraction of statistical inference. In a full course, however, there is time to develop the ideas more fully, and to work with many datasets having a context. Suppose that, instead of just working textbook inference problems, the students are given the problem of estimating the mean grade point average (GPA) of all students at Montclair State University (MSU). If they obtain the GPA's of a sample of 30 students, they certainly understand the difference between the mean GPA of the 30 students and the mean GPA of the entire student body. They may not fully understand the theory, but they can understand that the methods of

statistical inference allow them to use the information in their samples to make statements about the population mean.

17 The exercises used in the first two offerings of the course attempted to link statistics to probability but did not expose students to the practice of statistics. The p -value or confidence interval may be useful at a certain stage of solving a statistical problem, but their calculation is not the focal point of the problem-solving process. Without understanding the motivation of the original problem, having knowledge of the data collection process, and being involved in assessing the quality and characteristics of the data, calculating a p -value or confidence interval is simply an exercise in arithmetic. These results have no context. As [Moore \(1992\)](#) points out, "Statistics is the science of data ... they [data] are numbers with a context" (p. 15).

18 At this point I realized that whatever else was included in the course, the students should experience statistical thinking.

19 [Kinney's \(1997\)](#) recent textbook, *Probability: An Introduction With Statistical Applications*, proceeds in much the same way as described above, exposing students to sampling distributions, significance testing, and confidence intervals with a mathematical, rather than a statistical, treatment. If Kinney's course is followed with a course in statistics then the exposure to some of these concepts may be useful. But this type of treatment should not be confused with statistical thinking, and students and instructors alike should realize that statistical practice does not resemble this kind of theoretical treatment.

20 [Snee \(1993\)](#) states that "the 'content side' of statistics education should move away from the mathematical and probabilistic approach and place greater emphasis on data collection, understanding and modeling variation, graphical display of data, design of experiments, surveys, problem solving, and process improvement." He adds, "The goal should be to integrate the statistical thinking into the subject matter on which the student is working. Personal interest in the subject matter to which the statistical methodology is being applied and personal participation in the data collection and problem-solving processes are essential for developing value for statistical thinking" (p. 151).

4. Improvements

21 In later offerings of the course, two major changes were made to the original approach: the students collected their own data, and the data were binary. In statistical practice the normal distribution is prevalent, but the derivation of the sampling distribution of the mean requires considerable time. In analyzing binary data, we used the total number of successes, rather than the proportion of successes, as the test statistic. This allowed the students to work with exactly the same random variable they had been studying in the probability course, thus eliminating the need to derive a sampling distribution.

22 The need for experiential learning in statistics is well-documented, and as [Snee \(1993\)](#) states, "Collection and analysis of data is at the heart of statistical thinking. Data collection promotes learning by experience and connects the learning process to reality" (p. 152). The exercises described below are simple, yet they illustrate many of the important issues of proper survey and experimental methodology. Students are required to think about random sampling, wording of survey questions, and nonresponse in surveys, as well as randomization and consistent measurement in conducting experiments. Because the students collect the data themselves, the data have a context. Because no time must be spent deriving the sampling distribution, attention can be focused on statistical thinking and interpretation of the data.

23 Note that "significance testing" is done in only the most informal way. Assuming a certain value of p , the binomial probability of success, the graphing calculator is used to display a histogram of the distribution of X , the total number of successes. The appearance of the histogram is used to initially classify values as "likely" or "unlikely." The tail probability is then calculated and judgment is based upon this value. No distinction is made between one- and two-sided significance tests.

5. Specific Exercises

24 The following four exercises can be used to introduce ideas of statistical thinking. The exercises are very simple, but the issues addressed and the discussions generated are extremely rich in content.

25 In the first exercise, students perform an experiment where the parameter value is known. For example, each student is asked to flip a coin six times and record the total number of heads observed. They are asked, "Do your results seem consistent with the hypothesis that $p = .5$, i.e., that the coin is fair?" Ten students then combine their results and are asked the same question. They continue by comparing the combined result with each individual result. This is a very important exercise where students can see that even though the coin was fair, the data from a single sample may have suggested otherwise. They also see the increased accuracy in using a larger sample. Discussions also follow regarding the possibility that different students showed different results due to variations in flipping technique.

26 The second exercise has students predict the value of an unknown parameter, and then conduct an experiment to see if the data support their hypothesis. Each student is asked to begin by estimating the probability that a thumbtack will land point up when tossed. Each student then flips a thumbtack six times and records the total number of "ups." They address the same issues mentioned above in the coin tossing experiment, but here they are made uncomfortable by the fact that the true probability is unknown. They begin to see that the statements we make from our data can never be made with certainty.

27 The third exercise involves a survey where responses are binary -- for example, "Do you believe that abortion should be legal?" First, each student predicts the proportion of MSU students who favor legalized abortion. They survey ten "randomly selected" MSU students, use the data to test their hypothesis, and, as above, combine their results with those of other students. Here additional discussion topics include "Whom did you include in your sample?," "What is a random sample?," "How did you approach each respondent?," "Did each respondent understand the question?," "Did some people refuse to respond?," "Did you falsify any of your data?" Most of these questions are generated by the students themselves as they attempt to produce accurate data. They realize firsthand how the conclusions they can draw from their data are limited by the reliability of their collection methods.

28 In the fourth exercise, each student formulates a survey question or plans an experiment with a binary response. They must consider all of the issues brought out in the previous exercise, as well as select a sample size that they think will be adequate to accurately estimate the unknown parameter. The sample size question generates interesting discussion, as students have realized by now that a larger sample size will produce a more accurate estimate. However, faced with the prospect of actually performing the survey or experiment, they realize that sample sizes must be limited, and that there is a tradeoff between economy and precision.

29 I present the first two exercises at the beginning of any probability course. They make an excellent introduction to the definition of a probability. For the current investigation, I revisited all four exercises at the end of the unit on the binomial random variable, the first discrete random variable studied. We then added calculator simulations of binomial random variables with a variety of parameter values. Presenting this connection between real-world and theoretical probability models relatively early in the course allows the instructor to refer to this relationship as other probability models are studied.

30 This approach was used during two semesters, and the results were excellent. The students were excited by their work and showed evidence of truly understanding the concepts. Their questions and concerns regarding data acquisition and interpretation demonstrated an awareness of many of the issues involved in the practice of statistical science.

31 The exercise of having the students generate their own survey questions was particularly revealing. Many who had performed very well on standard binomial problems were not able to generate their own binary survey questions without some assistance from the instructor. This came as quite a surprise. I had expected this fourth exercise to be straightforward for students at the end of a unit on the binomial random variable. Some did not

know how to begin, some suggested open-ended questions, and others gave questions whose responses would not be binary. Upon completion of this exercise, however, I did feel that the students had a deeper understanding of the binomial random variable and of the notion that the other probability models we would study also have practical applications.

6. Conclusions

32 Have we achieved our goal? Yes and no.

33 Yes, the above approach using binary variables does successfully introduce the essence of statistical science in three weeks, and preserves a connection to the material in the probability course. The students collect their own data and are asked to use these data to make inferences. They discuss proper sampling and the significance of sample size. They have a better understanding of the binomial random variable from having to formulate their own binary survey questions. They also apply this experience to the other probability distributions they study and realize that those distributions represent different models for random behavior.

34 Although the basics of statistical inference are presented, this treatment is much too brief and far too limited in scope to convey the power and subtlety of statistical science. Students never work with data from a continuous distribution, nor do they experience exploratory data analysis. They do not see linear regression, or consider the issues involved in more formal significance testing or parameter estimation. They have no exposure to the testing of assumptions for a variety of methods.

35 Although some of the proposed exercises may be useful to enhance understanding of the probability topics, an effort to specifically introduce statistical ideas into the probability course is not recommended. Probability is important in its own right, not only as preparation for statistics. Our mathematics majors benefit greatly from this required course. Besides honing their problem-solving skills and applying the calculus, they receive preparation for work in such areas as mathematical modeling. [Cobb and Moore \(1997\)](#) note, "the domain of determinism in natural and social phenomena is limited, so that the mathematical description of random behavior must play a large role in describing the world" (p. 820). They are "hesitant to move any strictly statistical ideas into the probability semester" (p. 822).

7. Recommendations

36 My recommendation (which has been adopted) was to remove the probability prerequisite from the statistical methods course and leave the traditional probability course unchanged. Thus, students could take either probability, statistical methods, or both. The statistics course would change slightly, with more emphasis on data and less on theory. For students taking both courses, I would recommend they take statistical methods before probability. I believe that the theory will make more sense to students after they have been exposed to the practice of statistics.

37 [Cobb and Moore \(1997\)](#) also believe that "first courses in statistics should contain essentially no formal probability theory" (p. 820). They reason that "informal probability is sufficient for a conceptual grasp of inference." They feel that "we faculty imagine that formal probability illumines" the essential ideas of statistics and that this is "simply not true for almost all of our students" (p. 821).

38 When I began this project, I thought of probability and statistics as inexorably linked. I had not given much thought to where one ended and the other began. Now I realize their distinctness -- each taught and learned in very different ways, yet still connected. These differences and dependencies should be carefully considered when developing either of these courses or the programs that depend on them.

References

- Cobb, G., and Moore, D. S. (1997), "Mathematics, Statistics and Teaching," *The American Mathematical Monthly*, 104(11), 801-823.
- Falk, R., and Konold, C. (1992), "The Psychology of Learning Probability," in *Statistics for the Twenty-First Century*, eds. Florence Gordon and Sheldon Gordon, MAA Notes No. 26, Washington: Mathematical Association of America, pp. 151-164.
- Kinney, J. (1997), *Probability: An Introduction With Statistical Applications*, New York: John Wiley and Sons.
- Konold, C. (1995), "[Issues in Assessing Conceptual Understanding in Probability and Statistics.](http://www.amstat.org/publications/jse/v3n1/konold.html)" *Journal of Statistics Education* [Online], 3(1). (<http://www.amstat.org/publications/jse/v3n1/konold.html>)
- Moore, D. S. (1992), "Teaching Statistics as a Respectable Subject," in *Statistics for the Twenty-First Century*, eds. Florence Gordon and Sheldon Gordon, MAA Notes No. 26, Washington: Mathematical Association of America, pp. 14-25.
- Pfannkuch, M., and Brown, C. (1996), "[Building On and Challenging Students' Intuitions About Probability: Can We Improve Undergraduate Learning?](http://www.amstat.org/publications/jse/v4n1/pfannkuch.html)," *Journal of Statistics Education* [Online], 4(1). (<http://www.amstat.org/publications/jse/v4n1/pfannkuch.html>)
- Scheaffer, R., (1995), *Introduction to Probability and Its Applications*, Belmont, CA: Wadsworth Publishing Co.
- Snee, R. (1993), "What's Missing in Statistical Education?," *The American Statistician*, 47(2), 149-154.

Linda A. Tappin
Department of Mathematical Sciences
Montclair State University
Upper Montclair, NJ 07043

tappinL@mail.montclair.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](#)