

Teacher's Corner

Sequencing Topics in Introductory Statistics: A Debate on What to Teach When

Beth L. CHANCE and Allan J. ROSSMAN

We discuss various perspectives on the sequencing of topics to be studied in an introductory statistics course, debating the merits and drawbacks of different options. We focus on the introduction of data collection issues; the study of descriptive statistics for bivariate data; the presentation order of inference for means and proportions; and the placement of tests of significance and confidence intervals. Our goal is not to declare final resolution on these issues, but to stimulate instructors' thinking about this important aspect of course design. We conclude by identifying a set of core recommendations emerging from our points of agreement.

KEY WORDS: Course design; Statistics education.

1. INTRODUCTION

The past decade has seen the emergence of a reform movement in statistics education that has advocated teaching statistical thinking by engaging students in the active exploration of genuine data with the help of technology. Moore (1997a) recently summarized these emerging trends with regard to issues of content and pedagogy for introductory statistics courses. In response, Scheaffer (1997) commented that consensus about the content of the introductory course is stronger now than at any time in his career. Although generally agreeing with Moore's analysis, industrial statisticians Hoerl, Hahn, and Doganaksoy (1997) responded that "a major point not made . . . is that the sequencing of topics within the course needs to be rethought."

In this article our goal is to stimulate instructors' thoughts concerning the sequencing of topics in a reform introductory statistics course by presenting a "debate" of four propositions. Each proposition examines the relative placement in which to present two specific topics typically covered in an introductory course. For each proposition we present an argument in its favor and then respond with a rebuttal against the proposition. Our intent is not to declare an ultimate "winner" for each proposition or even to present all possible positions on the issue. Rather, our

primary aim is to generate reflection and discussion about these important decisions. We expect our arguments to provoke many differing reactions, but in the end we recognize several common goals that appear throughout both sets of arguments. From these points of agreement we identify several central principles related to course design. We hope that starting with these propositions will help instructors to identify what they feel is most important in their courses and to think through how sequencing impacts these goals as they make their individual decisions.

The propositions to be debated are:

1. that issues of data analysis should be studied prior to issues of data collection;
2. that descriptive analyses for bivariate data should come before inference procedures for one variable;
3. that inference for proportions should be studied before inference for means; and
4. that tests of significance should be studied prior to confidence intervals.

2. DEBATING THE PROPOSITIONS

2.1 Resolved, That Issues of Data Analysis Should be Studied Prior to Issues of Data Collection

Argument. Cobb and Moore (1997) argued strongly that the introductory course should begin with exploratory data analysis and descriptive statistics. They pointed out that this practice builds on students' motivation to analyze interesting data. Furthermore, since descriptive methods can be simple at first, students can gain confidence and good habits that will serve them well throughout the course. Exploratory analyses also introduce students early and often to the omnipresence of variability, a key theme of the entire course.

The distinction between population and sample need not be made at the beginning of the course, as meaningful analyses can be applied to, and interesting conclusions drawn from, available data. Examples of data that are readily available and highlight the drastic consequences of not properly using statistical methods include the 1970 draft lottery (Fienberg 1971) and the preliminary NASA analysis of space shuttle data (Dalal, Folkes, and Hoadley 1989). Calculating monthly medians of the draft numbers reveals a pattern indicating that random selection was not achieved in the lottery. Examining a scatterplot of O-ring failures versus temperature suggests a negative association that was missed by analysts and could have helped to prevent the tragic launching of the Challenger shuttle at such a low temperature. Towards preparing them to be consumers of quantitative information, students quickly learn valuable lessons from analyzing and drawing conclusions from existing information. Students

Beth Chance is Assistant Professor, Department of Statistics, California Polytechnic State University, San Luis Obispo, CA 93407 (E-mail: bchance@calpoly.edu). Allan Rossman is Associate Professor of Mathematics and Statistics, Department of Mathematics and Computer Science, Dickinson College, Carlisle, PA 17013 (E-mail: rossman@dickinson.edu). We originally presented this "debate" at the Western Statistics Teachers Conference held in Colorado Springs in March 1998. We thank the reviewers of this article for their helpful suggestions.

can also collect and analyze data about themselves from the first day of class using simple summaries and graphs, a natural way to establish students' personal identification and interest in the material.

Issues of data production are indeed essential for students to grasp, as the data collection method determines the scope of interpretation permissible from the data, but these need not be studied first. Having gained some experience with data analysis, students can be asked questions about interpretation that help them to realize the importance of considering the data collection plan in order to draw conclusions beyond merely the data analyzed. Moreover, the confidence and skills that students acquire by studying descriptive methods can enhance their learning of data production concepts such as bias, precision, and randomization. Data production issues can be studied after exploratory data analysis, thereby providing an effective bridge for linking exploratory methods with inferential ones studied later in the course, for inference procedures are appropriately applied precisely when randomization has been deliberately introduced into the data production process.

Rebuttal. The best habit we can teach students is to be conscious of proper data collection techniques before they begin any analysis of data, whether they are data found in the popular media or data they collect on themselves. Instead of blindly performing descriptive analyses at the instructor's direction, these analyses should always be preceded by questions such as what was being measured, how subjects were chosen, what was being asked, how the question was worded, and which type of study was implemented. The subsequent data analysis is much more meaningful when students fully understand how the data were obtained and whether they appropriately and meaningfully address the question posed. Then students can decide for themselves the appropriate level of analysis and conclusions. For example, students need to recognize the scope and legitimacy of conclusions that can be drawn from different types of studies (e.g., anecdotal, observational, experimental). They also need to be exposed to the sometimes dramatic consequences of improperly conducted studies [e.g., the *Literary Digest* wrongly predicting Landon would defeat Roosevelt in the 1936 presidential election based on biased sampling, see pp. 334–336 in Freedman, Pisani, and Purves (1998)], and to experience the difficulties and variability inherent in apparently simple measurements (e.g., diameter of a tennis ball). By presenting these issues at the very beginning of the course, students immediately become more intelligent consumers of quantitative information and learn to always question the source of the data before they interpret results.

Beginning with data collection issues from day one also mirrors the practice of statistics and allows students to (properly) begin their own data collection projects early in the course. What better way to give students an introduction to the nature of statistics than by examining the critical questions surrounding data production and immersing them immediately in examples of genuine usage? The first thing students learn, corresponding to what should always be their first step in practice, is how to formulate a question. It is too tempting for students who find data, often on the Web, to analyze them without having a question

in mind. By starting with the question and deciding what data will best address it, students learn the tools and good habits of statistical thinking in the same direct, logical order in which they use them. This starting point emphasizes to students the crucial role data collection issues play in analysis and interpretation and helps ensure they will always apply these principles when they collect their own data

Ideas of data collection are also easily absorbed by beginning students. Concepts of bias, precision, representative samples, and legitimacy of conclusions are often intuitive to students. Starting the course with these concepts allows students to build on their prior knowledge and to enhance their confidence and critical thinking skills, important goals considering the trepidation with which most students enter the course. They also appreciate that the course does not immediately plunge into calculations. Discussion of these ideas can still form a bridge to inference, but now that bridge is review, built on existing foundations. Furthermore, students begin immediately using terminology and descriptions, such as variability and randomization, which they will need throughout the course. These ideas are also very motivational. Recent textbooks aimed at consumers of statistics, such as those by Utts (1999), Moore (1997b), and Freedman, Pisani and Purves (1998) begin here. Even the most math-phobic students are drawn to the prospect of debunking a published result, increasing their interest in the course material and their pride in their abilities.

2.2 Resolved, That Descriptive Analyses for Bivariate Data Should Come Before Inference Procedures for One Variable

Argument. Examining relationships between variables is a fundamental idea that can serve as a unifying theme in a first course. It therefore warrants early attention and frequent repetition. Furthermore, studying bivariate analyses early in the course enables students to recognize the fundamental distinction between causation and association. Students in the first course encounter no more important idea than this, so they should study variations on this principle throughout the course. One illustration of how students can recognize this principle themselves is to ask whether the strong negative association that exists between a country's life expectancy and its ratio of people per television implies that sending televisions to impoverished nations would cause their life expectancy to rise (Rossman 1994).

Proceeding directly to descriptive bivariate analyses from univariate ones also highlights the parallel structure of descriptive analyses in both settings. In each, one begins with graphical displays, moves on to numerical summaries, and then produces mathematical models to summarize the data. This important process can be reinforced in students' minds by not detouring to a study of inference before considering bivariate relationships. By including categorical as well as quantitative variables in these analyses, students come to see that such disparate techniques as comparative boxplots, segmented bar graphs, and scatterplots all fit together under the framework of examining relationships between pairs of variables.

Although the study of regression can be daunting, it need not be. Students can benefit from an early study of regression that uses a descriptive perspective without a high degree of computational burden placed upon them. For example, formulas for the

least squares slope and intercept can be presented in terms of the means and standard deviations of the two variables and the correlation coefficient between them (also presented with minimal formulaic detail). When regression is presented at this level, students' understanding of it requires less maturity than does their understanding of inference, which should therefore wait until later in the course. Devoting the first third of the course to descriptive statistics for univariate and bivariate data also reinforces the idea that exploratory analyses are important to perform first and that inference is not necessarily the goal of every statistical analysis.

Rebuttal. If highlighting parallel structures is the goal, why not complete the process—graphical and numerical descriptions, specifying a model for the data, and then making decisions about the data through appropriate inferential procedures—before moving to a new setting? This is not to say that all questions lead to statistical inference, but this sequencing allows students to learn all the potentially relevant tools in one complete package. Students can carry a question all the way through, instead of learning some tools for univariate questions, then some for bivariate questions, then returning to univariate questions, then finally addressing bivariate questions again, a process that often requires numerous reminders of forgotten information. Students better learn the material when it is presented in a complete, coherent manner. The complete package can be modeled in the univariate setting and then *reinforced* in the bivariate setting. With this approach, all the stages of statistical analysis are sewn together, instead of appearing as disjointed pieces of a puzzle.

Second, too often students enter a course in statistics believing the focus will be on manipulating and memorizing formulas, frequently to the point where they can become intimidated by, and fixated on, the formulas. This attitude is reinforced when a course begins by introducing many formulas and can be detrimental to learning. By delaying even the simple expressions $\hat{y} = a + bx$ or $b = r(s_y/s_x)$ until later in the course, students are less likely to feel overwhelmed by the equations when they do appear. Instead, time is spent helping students focus on the general concepts in the course and building on their intuition. By the time regression is introduced, students have gained confidence in their statistical abilities and are better able to see the role of formulas as tools in a larger process.

Perhaps the largest benefit of delaying bivariate analyses is that treatment of inference comes earlier in the course. Clearly, inference is a difficult concept. Starting discussion earlier in the course and then repeating this process in different settings provides students with more time to absorb and practice with the ideas, instead of rushing through numerous inference procedures at the end of the course. Concepts that require complex reasoning should be addressed early in the course to have the best chance of being resolved in students' minds by the end. However, complex mathematical manipulations can be delayed until students have developed more confidence and trust.

The distinction between association and causation should indeed be visited early and often. However, this idea can also be explored early by beginning the course with discussion on the

distinction between experiments and observational studies as outlined in the first proposition. This focuses on the principle conceptually and intuitively instead of mathematically.

2.3 Resolved, That Inference for Proportions Should be Studied Before Inference for Means

Argument. One reason for studying inference for proportions prior to inference for means is that the setting is conceptually simpler. With binary variables, the proportion parameter uniquely describes the entire population. In contrast, with quantitative data the mean is merely one parameter that summarizes the center of the distribution. Other measures of center should be considered, and center might not even be the most interesting feature of the population. Wardrop (1994) took this recommendation to the extreme by devoting the first two-thirds of his innovative textbook to analysis of categorical variables, diving right into issues of experimental design in the first chapter and those of inference in the second.

Furthermore, simulations involving binary data are more straightforward to implement and to interpret than with quantitative data. To conduct a simulation with binary data, one does not need to specify a shape for the population distribution or other characteristics apart from the value of one parameter. In addition, one can start with real data by examining, for example, the proportion of brown candies in a sample, and then proceeding to use dice or playing cards to simulate random binary and multinomial processes. Such simulations can lead students to develop an intuitive understanding of fundamental concepts of inference.

A third argument in support of this proposition is that students encounter proportions frequently in the popular media. For example, most issues of *USA Today* report a plethora of statistics in the form of proportions. Students also tend to be drawn toward project topics that involve binary variables and therefore proportions. Studying inference for these parameters prior to inference for means allows for early consideration of design and inference components of statistical analysis, facilitating students' ability to perform substantive project work early in the course. Wardrop's book contains excellent examples of such student projects, involving such topics as wording of survey questions, dating habits, temperature forecasts, and throwing popcorn for a dog to catch.

Working first with proportions rather than means enables students to focus on the fundamental and difficult ideas of confidence and significance. Important but peripheral concerns associated with inference for means should wait until students have an understanding of basic inferential principles. Studying proportions first also allows for exact calculations of p values and power from the binomial distribution.

Rebuttal. Instead of arguing in favor of presenting means before proportions, this rebuttal proposes presenting inference for proportions and means (with variance known) concurrently. Students can learn properties and consequences of sampling distributions for proportions at the same point in the course as for means. Then students can learn how these ideas relate to the concepts of confidence and significance; they can also be introduced to the relevant formulas side-by-side. This highlights for students the common structure of the sampling distributions

(normal shape, centered at parameter, variation decreases with larger sample sizes) and helps them to focus on one overall idea, for example, (statistic-hypothesized value)/(standard deviation of the statistic), instead of several isolated formulas. They learn to apply these general properties independent of a particular setting, providing them with a much more powerful tool.

The propensity of students and media to focus on binary data can also be used as an argument to discuss quantitative data earlier. This broadens the scope of problems students examine, allowing more flexibility in examples and questions explored. Students also learn to report the sample standard deviation in contrast to many media examples. Typically, students suggest project topics that are split between categorical and quantitative measurements (e.g., time, GPA, speed, exam scores, heights, weights, heart rate), allowing much more variety than simple yes/no questions. This allows instructors to focus on helping students distinguish between variable types, and therefore which graphs and formulas are appropriate. This is important as students typically have tremendous difficulty when they initially encounter a research question with no contextual clues as to the proper analysis.

Hands-on measurements can also be easily implemented with quantitative variables. For example, weighing candy bars and recording the mint date of a penny are interactive, interesting examples, and can be used to draw repeated samples. Through these examples students also learn to conceptualize and deal with variability, a core concept of the course. Furthermore, recent technological tools easily give the user the ability to conveniently specify different population shapes and parameters. Instead of ignoring these properties, students can now explore them through intuitive, visual simulations. In fact, by visually displaying both the population and sample, these representations may even give students a better grasp of their distinction.

Assuming knowledge of the population standard deviation can be artificial, but students realize this and strive on their own to understand how to correct this assumption. By addressing this concern, students learn early on the crucial role of questioning the assumptions of each inferential procedure (including the binomial process upon which inference for proportions is based). Otherwise, the underlying assumptions can be too easily glossed over.

2.4 Resolved, That Tests of Significance Should be Studied Prior to Confidence Intervals

Argument. After studying sampling distributions through physical and technology simulations, the concept of significance provides the logical next step. Physical simulations include shuffling and dealing playing cards to simulate a randomization test for a question of sex discrimination (as in Scheaffer, Gnanadesikan, Watkins, and Witmer 1996). With these types of simulations students concentrate on the concept of rare event and on an intuitive understanding of p value. Although the ideas are closely related, the concept of significance arises more naturally from this treatment than does confidence. One need only ask how many of the simulated samples produced a result as extreme as that in the observed data. With the concept of confidence, after one asks how many simulated sample statistics fall within two standard deviations of the parameter, one must go on to invert the process and ask in how many of the simulated sam-

ples the parameter falls within two standard deviations of the sample statistic. Studying confidence intervals immediately after simulating sampling distributions can be a detour that diverts students' attention and causes them to miss the connection.

Putting significance before confidence also better models the process of scientific inquiry. It is natural to start with the question "Is there an effect?" and then to ask "If so, how much of an effect?" or to start with "Do the groups differ?" and then continue with "If so, by how much do they differ?"

Presenting significance before confidence can also provide an opportunity to emphasize that confidence intervals should accompany tests of significance whenever possible. Presenting confidence intervals second can also emphasize the complementary relationship between tests and intervals. Indeed, the confidence interval can be presented as containing the parameter values for which the null hypothesis would not be rejected.

Rebuttal. Too often, introducing inference with tests of significance requires a cumbersome detour into new terminology and notation. However, moving to confidence intervals from sampling distributions starts with application of the previously learned empirical rule: The confidence interval formula can be viewed as a rearrangement of the "within two standard deviations" expression allowing students to get their inferential feet wet more directly. For example, Rossman and Chance (2001) introduce the concept of confidence by having students take samples of Reese's Pieces candies, first directly, then with technology. After simulating the process many times using technology, students can state that 95% of the observations are within two standard deviations of the mean and build to the statement that for 95% of sample proportions, the population proportion is within two standard deviations.

This idea of confidence can then be generalized to discussion of which parameter values are and are not consistent with the sample data, as an introduction to the reasoning of significance tests. This approach steps students through the material changing only one dimension at a time.

Furthermore, since confidence intervals should accompany every analysis and are often used in practice in place of tests of significance, instructors can promote their importance by presenting them first, instead of as an afterthought. Students become comfortable performing and interpreting confidence intervals and thinking of effect sizes rather than just significance levels, so that confidence intervals truly are an automatic tool of any analysis.

3. POINTS OF AGREEMENT

Upon examining the preceding arguments, we find that several important points of agreement emerge:

- *Data production issues warrant serious attention.* Although we disagree about the point at which to introduce concepts of measurement, sampling, and experimentation, we agree strongly that these issues deserve considerable attention throughout an introductory statistics course.

- *Fundamental ideas should be introduced early and revisited often.* We agree that instructors should identify the central ideas that they want students to take away from the course (e.g., variability, relationships between variables, reasoning of inference). These ideas should be presented early and then repeated in

a variety of contexts and levels of complexity to enrich students' understanding, to help them build connections among different course components, and to develop their capacity to combine different statistical tools. For example, both sets of arguments agreed that the distinction between association and causation was a key concept and that such fundamental ideas should be introduced early in the course and emphasized throughout. One example would be to return to the *Literary Digest* prediction later in the course and realize that even though Landon's lead was highly statistically significant, inference is at best meaningless and at worst highly misleading when applied to data gathered with a biased sampling procedure.

- *Minimize distractions to allow students to concentrate on fundamental ideas.* In keeping with this emphasis on fundamental ideas, we recommend that instructors not devote substantial time to finer points that can distract students' attention from the larger issues. Our propositions concern the sequencing of general topics, not specific individual statistical techniques, because we contend that decisions about which techniques to cover are much less important than helping students to understand fundamental concepts as we have discussed above. For example, we would prefer to help students to acquire a firm understanding of the concepts of confidence and significance and an awareness of their roles and limitations, even if such a focus deemphasizes some mathematical details.

- *Emphasize common elements of analysis that arise in different situations.* It is important for students to see that several principles permeate much of statistical analysis. By helping them to understand these principles, the introductory course can better prepare students to comprehend subsequent techniques that they may encounter beyond that course. For example, instructors can stress that the approach of progressing from graphical displays to numerical summaries to mathematical models to formal inferences holds for both univariate and bivariate analyses. Instructors should also emphasize that the interpretation of p values and confidence intervals remains unchanged in all situations. The common structure of the test statistic and confidence interval formulas in the introductory course should also be emphasized.

- *Simulations are the way to study randomness, with tactile simulations preceding technology ones.* While we have presented different viewpoints on whether to start with means or proportions and on whether to begin with confidence intervals or tests of significance, we are in complete agreement that students should be introduced to the concept of randomness through the use of simulations. Moreover, we feel that it is important for students to perform physical simulations with hands-on manipulatives (using candies, dice, cards, . . .) before turning to the computer or calculator. While the technology simulations are efficient and potentially effective, we worry that students will fail to relate the output to the process being simulated unless they have engaged in the physical simulation first. We feel that these simulation exercises should be designed to introduce students to statistical issues such as confidence and significance, not to study probability for its own sake.

- *Understanding sampling distributions is crucial for understanding concepts of inference.* An instructor should be wary

not to treat the ideas of sampling distributions too quickly. These ideas are not simple, but are prerequisite knowledge for any true understanding of inference.

4. CONCLUSIONS

Our goal has been to focus attention and generate discussion about the important, but often overlooked, issue of sequencing in the introductory statistics course. Every instructor of introductory statistics must make a decision about each of the propositions that we have debated whenever he or she teaches the course. While we have simplified the discussion by presenting only two options for each proposition, we do recognize that there are other options. For example, an instructor might present inference for a population mean prior to or even instead of inference for a proportion, as opposed to the "proportions first" or "both concurrently" options that we have debated. It is also important to remember that these four propositions are not independent choices, for one must consider the impact that the choices have on each other. For example, if an instructor decides to delay regression, he or she may still want to introduce ideas of association earlier in the course by covering scatterplots or experiments.

While decisions about these propositions have important implications for facilitating students' learning and for sending cues about the relative importance of topics, we feel strongly that our points of agreement are much more central to course design than the specific resolution of these propositions. Instructors need to concentrate on their course goals and audience with these principles in mind, rather than automatically committing to the sequence presented in their text. With careful planning and management, instructors can sequence topics in a manner that most effectively accomplishes these larger goals.

[Received June 1999. Revised May 2000.]

REFERENCES

- Cobb, G., and Moore, D. (1997), "Mathematics, Statistics, and Teaching," *American Mathematical Monthly*, 104, 801–824.
- Dalal, S., Folkles, E., and Hoadley, B. (1989), "Lessons Learned from Challenger: A Statistical Perspective," *STATS: The Magazine for Students of Statistics*, 2, 14–18.
- Fienberg, S. (1971), "Randomization and Social Affairs: the 1970 Draft Lottery," *Science*, 171, 255–261.
- Freedman, D., Pisani, R., and Purves, R. (1998), *Statistics* (3rd ed.), New York: W. W. Norton.
- Hoerl, R., Hahn, G., and Doganaksoy, N. (1997), Comment on "New Pedagogy and New Content: The Case of Statistics" by D. Moore, *International Statistical Review*, 65, 147–153.
- Moore, D. (1997a), "New Pedagogy and New Content: The Case of Statistics," *International Statistical Review*, 65, 123–127.
- (1997b), *Statistics: Concepts and Controversies* (4th ed.), New York: W. H. Freeman.
- Rossmann, A. (1994), "Televisions, Physicians, and Life Expectancy," *Journal of Statistics Education*, 2, (www.amstat.org/publications/jse)
- Rossmann, A., and Chance, B. (2001), *Workshop Statistics: Discovery with Data*, 2nd edition, Emeryville, CA: Key College Publishing.
- Scheaffer, R. (1997), Comment on "New Pedagogy and New Content: The Case of Statistics," by D. Moore, *International Statistical Review*, 65, 156–158.
- Scheaffer, R. L., Gnanadesikan, M., Watkins, A., and Witmer, J. A. (1996), *Activity-Based Statistics*, New York: Springer Verlag.
- Utts, J. (1999), *Seeing Through Statistics* (2nd ed.), Belmont: Duxbury Press.
- Wardrop, R. (1994), *Statistics: Learning in the Presence of Variation*, Dubuque, Iowa: Wm C. Brown Publishers. Available from the author directly, wardrop@stat.wisc.edu