Investigating Statistical Concepts, Applications, and Methods

Beth L. Chance and Allan J. Rossman

Preface to the First Edition

To the Student

Statistics is a mathematical science.

While this is a very short sentence, perhaps a self-evident one, and certainly one of the shortest that you will find in this book, we want to draw your attention to several things about it:

- We use the singular "is" and not the plural "are." It is certainly grammatically correct and more common usage to say "statistics are...", but that use of the term refers to statistics as numerical values. In this sentence we mean statistics as a field of study, one that has its own concepts and techniques, and one that can be exciting to study and practice.
- We use "mathematical" as an adjective. Statistics certainly makes use of much mathematics, but it is a separate discipline and not a branch of mathematics. Many, perhaps most, of the concepts and methods in statistics are mathematical in nature, but there are also many that do not involve mathematics. You will see an example of this early in the book as you study the difference between observational studies and controlled experiments. You will find that even in cases where the mathematical aspects of two situations may be identical, the scope of one's conclusions depends crucially on how the data were collected, a statistical rather than a mathematical consideration.
- We use the noun "science." Statistics is the science of gaining insight from data. Data are (notice the plural here) pieces of information (often but not always numerical) gathered on people or objects or processes. The science of statistics involves all aspects of inquiry about data. Well-designed studies begin with a research question or hypothesis, devise a plan for collecting data to address that issue, proceed to gather the data and analyze them, and then often make inferences about how the findings generalize beyond the particular group being studied. Statistics concerns itself with all phases of this process and therefore encompasses the scientific method.

In these materials, our goal is to introduce you to this practice of statistics, to help you think about the applications of statistics and to study the mathematical underpinnings of the statistical methods. While you will only scratch the surface of the statistical methods used in practice, you will learn fundamental concepts (such as variability, randomness, confidence, and significance) that are an integral part of many statistical analyses. A distinct emphasis will be the focus on how the data are collected and how this determines the scope of conclusions that you can draw from the data.

One of the first features you will notice about these materials is that you will play the active role of investigator. You will read about an actual study and consider the research question, and then we will lead you to discover and apply the appropriate tools for carrying out the analysis. Almost all of the investigations in this book are based on actual scientific studies. At the end of each investigation is a “study conclusion” that allows you to confirm your analysis as well as to see examples of how to properly word conclusions to your studies, for the effective communication of statistical results is as important as the analysis. There are also numerous “explorations” where the primary goal is for you to delve deeper into a particular method or statistical concept.   A primary reason for the investigative nature of these materials is that we strongly believe that you will better understand and retain the concepts if you build your own knowledge and are engaged in the context. We don’t leave you without support and reference materials: be sure to read the expository passages, especially those appearing in boxes, and the section and chapter summaries. You may find this approach rather frustrating at first, but we also hope you will appreciate developing problem solving skills that will increase in utility as you progress through this and other courses.

We will ask you to use the computer extensively, both to analyze genuine data and also to investigate statistical concepts. Modern data exploration and analysis make heavy use of the computer. We have chosen the statistical package Minitab as the primary tool for analyzing data. Minitab is increasingly used in industry, but our choice mostly centers on its ease of use. After using Minitab for this course, you will have sufficient background to use most standard statistical packages.

This book will also make heavy use of *simulation* to help you focus on the central question behind many statistical procedures: “How often would this happen in the long run?” Often, the simulation results will direct us to a mathematical model that we can then use as a short-cut; at other times we will only be able to use simulation to obtain an approximation to the answer we are interested in finding. You will make frequent use of several technological tools, namely Minitab, Excel, and Java applets, to carry out these simulations and explorations. We have included instructions for how to use these tools throughout the text so that you may proceed through the investigations with minimal computer and programming background. Still, it will be important that you remember and build on the computing skills that you will develop as the course progresses. All data files and Java applets can be accessed from www.rossmanchance.com/iscam/.

We have also included a series of “practice problems” throughout the book. We envision these exercises as short, initial reviews of the terminology and concepts presented in the preceding investigations. Strategies for learning statistics often mirror those of learning a foreign language – you need to continually practice and refine your use of the terminology and to continually check that your meaning is understood. We hope that you will use these practice problems as a way of quickly assessing your knowledge. Some of the practice problems are in the nature of “further explorations,” and a few introduce new ideas or techniques based on what you just learned. Not all of the practice problems have simple correct answers but can also be used to spur debate and discussion in your class. Your instructor will inform you about obtaining access to the solutions to these practice problems.

Most of all, we hope you will find fun and engaging examples. Statistics is a vitally important subject, and also fun to study and practice, largely because it brings you into contact with all kinds of interesting questions. You will analyze data from medical studies, legal cases, psychology experiments, sociological studies, and many other contexts. To paraphrase the late statistician John Tukey, “the best thing about statistics is that it allows you to play in everyone else’s backyard.” You never know what you might learn in a statistics class!

To the Instructor

Motivation/Audience

The statistics education reform movement has revolutionized the teaching of introductory statistics. Key features of this movement include the use of data, activities, and technology to help students understand fundamental concepts and the nature of statistical thinking. Many innovative and effective materials have been developed to support this shift in teaching, and many instructors have changed their approach to teaching introductory statistics. So, where does this book fit in? What problem does it address?

The vast majority of these statistics education reform efforts have been aimed at “Stat 101,” the algebra-based service course. As a result, the movement has largely ignored mathematically inclined students, the very students who might be attracted into the field of statistics and into the teaching of statistics. Our goal in this book is to support an introductory course at the post-calculus level around the best features of statistics education reform: data, activities, concepts, and technology.

Mathematically inclined students have typically had to choose between taking “Stat 101” or taking a course in probability and mathematical statistics. The first option neither challenges them mathematically nor takes advantage of their mathematical abilities. The second option often devotes an entire course to probability before turning to statistics, which could delay capturing the interest of students who would find data analysis appealing. We offer this book as an alternative that we hope provides a balanced introduction to the discipline of statistics, emphasizing issues of data collection and data analysis as well as statistical inference.

While we believe that this type of introduction to statistics is appropriate for many students, we want to draw particular attention to two student audiences: potential statisticians and teachers. Mathematically inclined students who take “Stat 101” may not recognize the mathematical richness of the material, and those who begin with a mathematical statistics course may not appreciate the wide applicability of statistics. We hope that this book reveals the great appeal of statistics by asking students to investigate both the applicability of statistics and also some of its mathematical underpinnings. We also hope to entice students to study statistics further. This book can lay the foundation for several types of follow-up courses, such as a course in regression analysis, design of experiments, mathematical statistics, or probability models.

Prospective teachers of statistics are also an important audience, all the more so due to the growth of the Advanced Placement (AP) program in Statistics and the emphasis on data analysis throughout the K-12 program in the NCTM Standards. Not only is the content of this book in line with the AP course and the NCTM Standards, but so too is the pedagogical approach that emphasizes students’ active construction of their knowledge and the use of technology for developing conceptual understanding. We hope that both the content and pedagogy presented here will prepare future teachers to implement similar approaches in their own teaching.

Principles
Some of the principles that have guided the development of these materials are:

- *Motivate with real studies and genuine data.*

Almost all of the investigations in this book center on real studies and genuine data. The contexts come from a variety of scientific disciplines, including some historically important studies. With all of these studies we provide ample background information without overwhelming students or expecting them to know much about the field of application. We also take some examples from popular media, aiming to appeal to diverse student interests.

- *Emphasize connections among study design, inference technique, and scope of conclusion.*

Issues of study design come up early and recur throughout the book. From the opening chapter we emphasize the distinction between observational studies and controlled experiments, stressing the different types of conclusions that can be drawn from each. We also highlight the importance of randomness and its crucial role in drawing inferences, paying attention to the difference between *randomization* of subjects to treatments and *random sampling* of objects from a population. The connection between study design and inference technique depends heavily on the concept of a sampling/randomization distribution, which we also emphasize throughout.

- *Conduct simulations often.*

We make frequent use of simulations, both as a problem-solving tool and as a pedagogical device. (One challenge with this approach is helping students to recognize the difference between these two uses.) These simulations address the fundamental question underlying many statistical inference procedures: "How often would this happen in the long run?" We often start with tactile simulations before proceeding to technological ones, so that students can better understand the random process being simulated. Technology-based simulations often involve using a Java applet, but we also ask students to write their own small-scale macros in Minitab to conduct simulations. Developing this skill can help students to apply simulation as a general problem-solving tool to other situations.

- *Use variety of computational tools.*

We expect that students will have frequent access to computer software as they work through this book. We ask students to use technology both to analyze data and to explore statistical concepts. Our guiding philosophy is to choose the appropriate software tool for the task at hand. When the task is to analyze data, the appropriate tool is a statistical analysis package. We've chosen Minitab in this book for its ease of use, but other packages could be used as well. When the task is to develop understanding of a concept, the tool is often a Java applet specifically designed for that purpose, typically with a premium on interactivity and visualization. For a few tasks, such as examining the effect of changing a parameter value, the appropriate tool might be a spreadsheet package. We've chosen Excel as the spreadsheet package for this book, but its use is minimal.

- *Investigate mathematical underpinnings.*

The primary contrast between this book and a "Stat 101" book is that we often ask students to use their mathematical training to investigate some of the underpinnings behind statistical procedures. An example is that students examine the principle of least squares and other minimization criteria in both univariate and bivariate settings. Students also examine functions symbolically and numerically to investigate issues such as sample size effects. Many of these more mathematical aspects emerge after the ideas have been motivated through student investigations of an application.

- *Introduce probability "just in time."*

We don't see probability as the goal of an introductory statistics course, so we introduce probability ideas whenever they are needed to address a statistical issue. Often a probability analysis follows a simulation analysis as a way to obtain exact answers to the simulation's approximation. For example, the hypergeometric distribution is introduced after simulating a randomization test with a 2×2 table, and the binomial distribution arises after using simulation to analyze data from a Bernoulli process. Later probability models are introduced as another type of approximate analysis. Examples include the normal approximation to the binomial for the sampling distribution of a sample proportion and *t*-distributions as approximations to randomization distributions.

- *Foster active explorations.*

This book consists mostly of investigations that lead students to construct their own knowledge and develop their own understanding of statistical concepts and methods. These investigations contain directed questions that lead students to those discoveries. We expect that this pedagogical approach leads to deeper understanding, better retention, and more interest in the material. (See our "To The Student" comments at the beginning of this preface.)

- *Experience the entire statistical process over and over again.*

From the outset we ask students to consider issues of data collection, produce graphical and numerical summaries, consider whether inference procedures apply to the situation, apply inference procedures when appropriate, and communicate their findings in the context of the original research question. This pattern is repeated over and over as students encounter new situations, for example moving from categorical to quantitative responses or from two to several comparison groups. We hope that this frequent repetition helps students to see the entire story, to appreciate the "big picture" of the statistical process, and to develop a feel for "doing statistics." We also emphasize students' development of communication skills so that they can complete the last phase of the statistical process successfully.

Content and Sequencing

Much of the content here is standard for an introductory statistics course, but you will find some less typical inclusions as well. In addition to the early emphasis on study design and scope of conclusions, Chapter 1 concentrates on comparisons in the context of categorical response variables, including topics such as relative risk, odds ratio, and Fisher's exact test. Concepts introduced here include variability, confounding, randomization, probability, and significance. Then Chapter 2 repeats the themes of Chapter 1 in the context of quantitative response variables and randomization distributions, and also introduces concepts such as resistance. Chapter 3 moves from comparisons to drawing samples from a population, turning again to categorical variables and focusing on hypergeometric and binomial models. Concepts of bias, precision, confidence and types of errors are introduced here. This univariate analysis continues with quantitative variables in Chapter 4, where students study more probability models and sampling distributions. They also encounter the *t*-distribution and a discussion of bootstrapping in this chapter. Chapter 5 then to the theme of comparisons between two groups, focusing on large-sample approximations. The final chapter considers comparisons among several groups and associations between variables, including chi-square tests, analysis of variance, and simple linear regression.

Our primary goal with this distinctive sequencing is to help students to understand the "big picture" by experiencing the entire process of statistical investigations from start to finish, over and over again in new settings. We do not expect students to fully understand the concept of significance, or the principle that the conclusion of a study depends crucially on how the data were collected, after their first exposure to that concept and principle. But we do expect their understanding to deepen each time they encounter the same concept and principle. For this reason we present issues of data collection, data description, and statistical inference all in the first chapter. We first present these ideas in the setting of a randomized comparative experiment, because that type of study highlights the role of randomization both for neutralizing confounding variables and for assessing statistical significance. Finally, we begin with a categorical response variable to simplify the analysis and the probability calculations, which led us to introduce combinations, hypergeometric probabilities, and Fisher's exact test in Chapter 1.

We believe that most instructors will find that there is too much material here for a one-semester course, so some picking and choosing is necessary. We have generally found that each chapter takes 2-3 weeks to study. We strongly recommend that Chapter 1 be covered in its entirety, because it introduces many concepts and principles that permeate the rest of the course and the practice of statistics. One option after that is to stick with the analysis of categorical variables by covering Chapter 3 and then Sections 5.1, 5.2, and 6.1. A more conventional option would be to next cover Chapter 2 in its entirety for an introduction to quantitative variables, and then pick and choose among the remaining four chapters. An instructor may also choose to supplement the probability detours in this book with more development of the probability ideas. We think it is important to keep in mind that even a course that does not go beyond the first three chapters has included the most fundamental statistical concepts, illustrated to students the application of probability models to statistical analyses, and focused on the statistical process in several settings. Several sample syllabi can be found among the instructor resources on the web.

Another decision concerns the level of mathematical depth to present. We believe this course can appeal to science students who are less interested in the mathematical detail as well as to mathematics and statistics majors. For that reason, we have aimed for flexibility in the level of mathematical aspects covered. Most of the mathematical detail occurs in explorations and exercises. For example, the text briefly highlights the distinction between relative risk and odds ratio but then offers the option of having students work through an Excel exploration that analyzes these calculations in more detail and highlights the invariance property of the odds ratio. Instructors also have the option of assigning homework problems that ask students to derive the mathematical relationship between the odds ratio and the relative risk formulas. The key is that students will be motivated to understand the deeper mathematical ideas after seeing the applications of the ideas in practice, and it is often their own curiosity that drives this follow-up study.

Structure

Most of this book consists of "investigations" that ask students to discover and apply statistical concepts and methods needed for analyzing data gathered to address a particular research question. The investigations contain a series of directed questions, with space provided for students to record their responses, which then serve as their class notes. At the end of each investigation are "study conclusions" that summarize what the student should have discovered about the study and model effective communication of the analysis, often followed by a "discussion" of the statistical issues that emerged. Interspersed among the investigations are optional "explorations" that ask students to delve deeper into a statistical concept, often involving the use of technology. You will also find "detours" that introduce terminology or technology hints; we set these apart so as not to interrupt the flow of the study investigation. We hope that these detours will help students to not become bogged down in these details but instead maintain focus on the larger ideas and the applications/interpretations.

One common concern that students express about this discovery-oriented pedagogical approach is that they want to see worked-out examples. We resist providing such examples, because we worry that students learn to mimic examples mindlessly rather than to develop their own understanding. We do include some solved examples in this book, but they come at the end of the chapters, and even there we leave space for students to write their own answers before we provide ours.

In addition to the investigations, we provide two types of assessment questions through which students can develop their understanding and skills, and through which instructors can provide feedback on that development. One type is "practice problems" that assess students' level of understanding of what the preceding investigations were meant to convey. These are interspersed throughout the investigations, and their goal is to provide students and instructor alike with an immediate indication of how well students are learning essential concepts and skills. The other type of assessment is more conventional homework exercises that appear at the end of each chapter. We present over 300 exercises, some of which focus on students' abilities to apply the techniques that they have learned to real data from genuine studies. Other exercises test students' conceptual understanding, while many ask students to further investigate mathematical underpinnings or extensions of what they have learned.

So, the overall structure of a chapter is a collection of investigations focused on a common problem structure (e.g., comparing two groups on a categorical variable), discussed from the data collection process through inferential procedures, with general definitions and terminology embedded in the context and practice problems interspersed for quick checks of understanding, and then a complete example in detail at the end of the chapter, followed by exercises. At the end of each chapter we also summarize the entire chapter and the technological tools discussed within the chapter. Then the next chapter repeats this structure for a new problem structure (e.g., comparing two groups on a quantitative variable). The sections are numbered within a chapter (e.g., Section 1.2) and then the investigations and practice problems are numbered within a section (e.g., Investigation 1.2.1). To facilitate navigation through the book, these section and investigation numbers appear in the page headers. Often, a section will correspond to one class period, but see the instructor resources for more detailed suggestions on timing. The instructor notes also provide guidance on which exercises can follow which sections in the chapter.

Pedagogy

This text is designed for use in an active learning environment where students can work collaboratively. Our view is that students learn more from what they *do* than from what they *hear* from their instructors or *read* from their authors. As the authors of these materials, we have tried to provide engaging contexts and to ask questions that will lead students to develop and apply their knowledge. As much as possible, we have based these investigations on research findings about how students learn statistics, as well as on our own extensive class-testing experiences. We encourage you as instructor to act as the intellectual manager of the students' learning activities.

We believe that the ideal classroom environment provides constant computer availability for students and that the instructor ideally gives students ample time to struggle with the questions before confirming the answers. However, these materials are flexible enough that an instructor can use the investigations as examples through which to lead students in a lecture setting without a computer lab. The "study conclusions" and "discussion" sections provide exposition to help students make sure that they are discovering what they are expected to in the investigations. Students may need to be encouraged to pay close attention to these discussions and to truly appreciate that there is substance and assistance in these summaries for them to consider.

We also believe this course can be supplemented with additional examples of the instructor's choosing, having students carry out data collection and analysis projects, additional data collection on students in the class, as well as quizzes and exams. Sample project assignments, quizzes and exams are available among the instructor resources. It is crucial that the assessments parallel the materials in their focus on applications and interpretations, in order to highlight to students that these ideas are of utmost importance in the course.

We also contend that it is important to provide frequent checks and feedback on the students' understanding of the material. We typically use the practice problems (see the earlier section on "Structure") for this purpose. Students can be asked to submit their solutions in between class periods (including via email or an electronic course delivery system like Blackboard and WebCT). This provides the instructor with feedback on what the students may be struggling with, as well as providing an opportunity for the instructor to supply individual feedback to the students and/or begin the next class session by focusing on the most prevalent misconceptions. We tell the students that one set of practice problems should only take about 30 minutes and that these do not always have black-and-white answers, as these problems often serve to raise issues and motivate questions as well.

Pre-requisites

This book provides an introduction to statistics for students who have completed a course in one-variable calculus. We do not make frequent use of calculus, but we do assume that students are comfortable with basic mathematical ideas such as functions, and we do call on students to use derivatives and integrals on occasion. We do ask students to use technology heavily, including some small-scale programming, but we do not assume prior knowledge of programming ideas or of any particular software. No prior knowledge of statistics is assumed, although we do not devote much time to ideas (such as mean and median, histograms, and scatterplots) that they are likely to have encountered before.

Instructor Resources

You can find a plethora of instructor resources to accompany this book and support your teaching on the web at www.rossmanchance.com/iscam. Resources available there include the many Java applets and data files, available in a variety of formats, used in the book. Additional resources include an extensive instructor's guide with specific teaching suggestions for each investigation, frequently asked questions, errata, Minitab hints, solutions to investigations, practice problems and exercises, and sample syllabi and exams. Many of these components are password protected so they are not accessible to students. We have also established an email discussion list for adopters of the materials.

Acknowledgements

We gratefully acknowledge the National Science Foundation for supporting the development of these curricular materials through grants #9950746 and #0321973. This grant has enabled us to do extensive development and testing of these materials. We especially thank Joan Garfield for her evaluation work on this project and for the inspiration that she has provided. We thank our Cal Poly colleagues Ulric Lund and Karen McGaughey, who have helped with this grant project by reviewing early drafts of materials. We especially appreciate the help that we have received from students who have reviewed the materials and helped with a variety of tasks related to this project. Many thanks to Laurel Koester, Tierra Stimson, Nicole Walterman, and Rebecca Russ. We are also very appreciative of the faculty members who attended the NSF-sponsored workshop that we conducted in June 2004; the feedback from these workshop participants led to substantial improvements in these materials.

We are very grateful to colleagues who reviewed early drafts of these materials. We especially appreciate the suggestions of:

Robin Lock, Saint Lawrence University
Jackie Dietz, Meredith College
John Holcomb, Cleveland State University
Julie Legler, Saint Olaf College
Ginger Rowell, Middle Tennessee State University
Chris Franklin, University of Georgia
Julie Clark, Hollins University
Tom Short, Indiana University of Pennsylvania.

Robin and Jackie provided invaluable advice and encouragement on this project, and in other collaborations over the years. John, Julie, Ginger, Chris, and Julie all class-tested the materials and provided very helpful, extensive feedback. Tom served not only as a class-tester but also as a very effective accuracy checker. We appreciate the time and expertise that these exceptional teachers offered to us and this project.

We extend our thanks to those who have class-tested the preliminary edition of this book and provided us with helpful feedback:

Nancy Boynton, SUNY Fredonia
Greg Cicconetti, Muhlenberg College
Stephanie Fitchett, Florida Atlantic University
Mary Fowler, Worcester State University
David Gurney, Southeastern Louisiana University
Johanna Hardin, Pomona College
Gary Kader, Applachian State University
John Kern, Duquesne University
Bessie Kirkwood, Sweet Briar College
Shonda Kuiper, Grinnell College
Andre Lubecke, Lander University
Leigh Lunsford, Longwood University
Michael Orrison, Harvey Mudd College
Michael Wylder, Woodward Academy

We thank Carolyn Crockett, Dan Geller, and their editorial team at Duxbury for their support, guidance, and patience throughout the development process, and we thank Kelsey McKee and Merrill Peterson for their work on the production end.

We are very pleased to have the opportunity to teach and develop innovative courses at Cal Poly, whose "learn by doing" motto provides a perfect match for our teaching philosophy. We especially appreciate the leadership and support provided by Dean Phil Bailey, Associate Dean Roxy Peck, and Department Chair Jay Devore. Special thanks to Carol Erickson for all that she does to support our work, far above and beyond the call. We thank the faculty and students at St. Olaf College for their hospitality during a sabbatical leave.

Most of all, we thank all of our students who have served as our pedagogical "experimental units" over the years, for their patience and good cheer in helping us to class-test drafts of these materials, and for inspiring us to strive constantly to become better teachers.

Beth Chance and Allan Rossman
bchance@calpoly.edu, arossman@calpoly.edu
San Luis Obispo, California
April 2005