PREFACE

It is the authors' strongly held belief that the most effective way for beginning statistics students to learn statistics well is for them to directly experience its central concepts and standard procedures through closely observing and working with simulated and real data sets. The particular advantage of the use of simulated data is that the student knows precisely the underlying mechanism—that is, the probability model—that is producing the data. Thus the student can see how close the data come, when studied and analyzed with statistical expertise, to revealing their underlying model-producing truths. The learning of statistics experientially by direct contact with data is accomplished in this textbook by the presentation and demonstration of statistical concepts using the five- and six-step simulation methods throughout. The same experiential pedagogical approach was previously developed into a precollege textbook by two of the authors and used in a five-year National Science Foundation Educational Directorate national teacher training grant to William Stout and Kenneth Travers.

A vital complement to the experiential learning of statistical concepts is a clear and understandable verbal and formulaic presentation of the formal, logical, and deductive aspects of the subject, shorn of all extraneous mathematical baggage. This is unfailingly carried out throughout this text. In most cases this theoretical component is presented after a thorough experiential immersion in and discussion of the topic of interest. For example, the basic aspects of probability—often a difficult set of principles for beginning students to master—are first presented as empirical phenomena in Chapters 4 and 5 and only then are abstracted and expanded into a deductive body of mathematical propositions and statistical models in Chapters 8 and 13. The textbook is very deliberate in thoroughly verbally explaining all concepts and methods, based in the authors' belief that verbal explanations are a key component of an in-depth understanding of statistics.

To be most effective, an introductory statistics textbook must strike a balance between helping the student learn to reason statistically and

exposing the student to the standard body of statistical procedures that are widely used in practice and are judged to constitute statistical literacy. This balance, a central goal of this book, equips a student with both real statistical discernment and the needed familiarity with those statistical procedures so widely used in science, business, and government and so often reported in the media.

We believe that an introductory statistics textbook should both be able to stand alone in educational settings where calculator and computer technology is not being used and also be equipped to make profound and meaningful use of available technology when the instructor wants. This book is designed with these principles in mind. The book is totally self-contained for classrooms in which technology is not available or not desired. However, for those wishing to use it, such technology—in this case Texas Instruments graphing calculators and computers—can greatly enhance students' understanding of statistics. A major effort has therefore gone into developing instructional programs for use with TI calculators and computers. In particular, the computer software is an enhancement and a customization of software developed over several years by John Marden. A precursor of the textbook software is currently being used in the multiple-section Statistics 100 course at the University of Illinois, which is taken by six hundred to eight hundred students each year. The role of the software provided as a computer supplement to the book is to add to a lecture-based statistics course a very efficient laboratory component in which the central concepts are rapidly and clearly experienced first-hand as the student carries out the computer exercises or even just plays with the software. The computer exercises in the supplement constitute a sort of laboratory with manual that guides the student through various data-driven experiences of important statistical concepts. The TI calculator programs and exercises, provided in a TI calculator supplement, play an important similar but somewhat more modest role. Throughout the technology exercises, the central role of the technology component is instructional rather than computational. Interestingly, with the computer software, theoretical quantities of interest can be estimated rapidly with great accuracy using the five- and six-step simulation methods. To a lesser extent, this is true of the TI programs too.

The textbook breaks new ground in content through its inclusion of bootstrap-based inferential procedures as an important component. The bootstrap approach is one of the most important statistical advances of the past 25 years and is rapidly becoming an essential component of the practicing statistician's toolbox. When presented appropriately, the bootstrap approach is intuitive and hence easily understandable by the beginning statistics student. In keeping with the instructional approach of the book, the bootstrap method is simulation-based, and it seamlessly meshes with the five- and six-step methods used throughout. Most important, it

provides the student with a powerful cutting-edge method of statistical inference that, together with small-sample normal-population sampling inferential approaches and large-sample inferential approaches based on the central limit theorem and the analogous large-sample result for chi-square statistics—another emphasis of the book—provides the student with enormous statistical inferential power.

The book is written to be accessible to all students having at least a modest exposure to algebra: "intermediate algebra" suffices. Although formulas and graphs appear frequently and students are encouraged to think deeply about what they are learning, the amount of formal mathematical background needed is minimal. The book should work well in any non-calculus-based introductory statistics course at any two-year or four-year college or university. It can also be effectively used at the precollege level, especially for Advanced Placement statistics courses. It can be taught to science and math majors and to humanities and social science majors with equal ease. It presumes no prior exposure to statistics.

Just as our experiential approach to the teaching of statistics makes learning statistics easier and more rewarding for the student, this approach plays the same role for an instructor who is teaching statistics while having a limited background in the subject. In fact, the book makes it very easy for an instructor of mathematics or science with limited *statistical* training to learn statistics while teaching the course.

Several influential nationally circulated reports have stressed the need for new emphases in the teaching of statistics. For example, the widely influential American Statistical Association/Mathematical Association of America Cobb report recommends that the teaching of statistics be heavily *data based* and proposes that more emphasis be placed on statistical concepts than on abstract theory. Further, it stresses *active learning*. This textbook is tailor-made to address these valid recommendations because of the reasons stated above, and moreover because of the numerous exercises provided, both after each section and as review exercises at the end of each chapter. There are both conventional pencil-and-paper exercises and optional TI and computer exercises. Thus the student has ample opportunity to practice using the concepts learned in the text.

The book begins with three chapters describing how one summarizes data—data being the focus of statistics—graphically and by means of statistical indices. Chapters 4 and 5 then provide a heavily empirical introduction to probability through emphasis on expected value and on probability itself. This is appropriate because probability is the logical underpinning of inferential statistics. Chapters 6 and 7 introduce the process of statistical inference, with the widely used and fascinating topic of chi-square testing appearing unusually early, as contrasted with the typical statistics textbook,

in Chapter 7. Then Chapter 8 deepens the student's understanding of probability, which is the source of statistical modeling. The topic of how to achieve accurate measurement in science and industry, and in particular the role of statistics in measurement, is covered in Chapter 9. This important topic of measurement is too often ignored in introductory statistics textbooks. Chapter 10 tells the student how one obtains good data for statistical analyses through approximate sampling of populations and well-designed statistical experiments. Then Chapters 11, 12, and 13 cover statistical inference in depth, stressing confidence-interval estimation and hypothesis testing. Chapter 14 provides a formal and rigorous introduction to probability and some useful modeling applications, while providing some more inferential techniques. Finally, Chapter 15 provides a very accessible introduction to Analysis of Variance and multiple regression, which together provide the backbone to much of applied statistical practice.

The entire book can easily be covered in a two-semester (or three-quarter) course. If desired, the instructor can omit Chapter 9, Chapter 4, or Chapter 15 without loss of continuity. In a one-semester course, Chapters 1–8 and 11–12 can be covered, in which case the instructor is encouraged to monitor the ongoing cognitive progress of the class and set the pace accordingly. Coverage of Chapter 10 is also highly desirable. In the one-semester course, to avoid undesirable time pressure, the optional sections should likely be omitted, as well as all or selected sections from Sections 1.3, 1.6, 2.5, 6.3, 11.7–11.9, and 12.4–12.6. Moreover, Chapter 7 can be omitted if the instructor wishes greater class time for the important Chapters 11 and 12.

This second edition has been heavily influenced by instructor and student input based on classroom use of the first edition. In particular, Chapter 10 on data collection, Chapter 15 on analysis of variance and multiple regression, and an annotated compilation of Formulas and Techniques were added in response to such input. We are confident that you will find teaching statistics from the textbook to be a rewarding experience. We look forward to interacting with instructors by means of the Möbius Web site for this textbook. We welcome queries and are eager to assist in teaching statistics in this very exciting, effective, and experientially based way.

It is worth noting that in addition to the college setting it was designed for, the second edition is also ideal for high school Advanced Placement statistics courses. Its coverage of the College Board's outline of topics covered by the AP Examination in Statistics is remarkably complete. Moreover, the College Board's recommended emphasis on taking a simulation approach, heavily using technology, and having students construct their own statistics knowledge base (through individual and group projects, etc.) mesh seamlessly with the second edition and its optional graphing calculator and computer supplements.