PREFACE

MOTIVATION

The idea for this book grew out of discussions between the statistics faculty and the engineering faculty at the Colorado School of Mines regarding our introductory statistics course for engineers. Our engineering faculty felt that the students needed substantial coverage of propagation of error, as well as more emphasis on model-fitting skills. The statistics faculty believed that students needed to become more aware of some important practical statistical issues such as the checking of model assumptions and the use of simulation.

My view is that an introductory statistics text for students in engineering and science should offer all these topics in some depth. In addition, it should be flexible enough to allow for a variety of choices to be made regarding coverage, because there are many different ways to design a successful introductory statistics course. Finally, it should provide examples that present important ideas in realistic settings. Accordingly, the book has the following features:

- The book is flexible in its presentation of probability, allowing instructors wide latitude in choosing the depth and extent of their coverage of this topic.
- The book contains many examples that feature real, contemporary data sets, both to motivate students and to show connections to industry and scientific research.
- The book contains many examples of computer output and exercises suitable for solving with computer software.
- The book provides extensive coverage of propagation of error.
- The book presents a solid introduction to simulation methods and the bootstrap, including applications to verifying normality assumptions, computing probabilities, estimating bias, computing confidence intervals, and testing hypotheses.
- The book provides more extensive coverage of linear model diagnostic procedures than is found in most introductory texts. This includes material on examination of residual plots, transformations of variables, and principles of variable selection in multivariate models.
- The book covers the standard introductory topics, including descriptive statistics, probability, confidence intervals, hypothesis tests, linear regression, factorial experiments, and statistical quality control.

MATHEMATICAL LEVEL

Most of the book will be mathematically accessible to those whose background includes one semester of calculus. The exceptions are multivariate propagation of error, which requires partial derivatives, and joint probability distributions, which require multiple integration. These topics may be skipped on first reading, if desired.

COMPUTER USE

Over the past 35 years, the development of fast and cheap computing has revolutionized statistical practice; indeed, this is one of the main reasons that statistical methods have been penetrating ever more deeply into scientific work. Scientists and engineers today must not only be adept with computer software packages, they must also have the skill to draw conclusions from computer output and to state those conclusions in words. Accordingly, the book contains exercises and examples that involve interpreting, as well as generating, computer output, especially in the chapters on linear models and factorial experiments. Many statistical software packages are available for instructors who wish to integrate their use into their courses, and this book can be used effectively with any of these packages.

The modern availability of computers and statistical software has produced an important educational benefit as well, by making simulation methods accessible to introductory students. Simulation makes the fundamental principles of statistics come alive. The material on simulation presented here is designed to reinforce some basic statistical ideas, and to introduce students to some of the uses of this powerful tool.

CONTENT

Chapter 1 covers sampling and descriptive statistics. The reason that statistical methods work is that samples, when properly drawn, are likely to resemble their populations. Therefore Chapter 1 begins by describing some ways to draw valid samples. The second part of the chapter discusses descriptive statistics.

Chapter 2 is about probability. There is a wide divergence in preferences of instructors regarding how much and how deeply to cover this subject. Accordingly, I have tried to make this chapter as flexible as possible. The major results are derived from axioms, with proofs given for most of them. This should enable instructors to take a mathematically rigorous approach. On the other hand, I have attempted to illustrate each result with an example or two, in a scientific context where possible, that is designed to present the intuition behind the result. Instructors who prefer a more informal approach may therefore focus on the examples rather than the proofs.

Chapter 3 covers propagation of error, which is sometimes called "error analysis" or, by statisticians, "the delta method." The coverage is more extensive than in most texts, but because the topic is so important to many engineers I thought it was worthwhile. The presentation is designed to enable instructors to adjust the amount of coverage to fit the needs of of the course. In particular, Sections 3.2 through 3.4 can be omitted without loss of continuity.

Chapter 4 presents many of the probability distribution functions commonly used in practice. Point estimation, probability plots and the Central Limit Theorem are also covered. The final section introduces simulation methods to assess normality assumptions, compute probabilities, and estimate bias.

Chapters 5 and 6 cover confidence intervals and hypothesis testing, respectively. The *P*-value approach to hypothesis testing is emphasized, but fixed-level testing and power calculations are also covered. The multiple testing problem is covered in some depth. Simulation methods to compute confidence intervals and to test hypotheses are introduced as well.

Chapter 7 covers correlation and simple linear regression. I have worked hard to emphasize that linear models are appropriate only when the relationship between the variables is linear. This point is all the more important since it is often overlooked in practice by engineers and scientists (not to mention statisticians). It is not hard to find in the scientific literature straight-line fits and correlation coefficient summaries for plots that show obvious curvature or for which the slope of the line is determined by a few influential points. Therefore this chapter includes a lengthy section on checking model assumptions and transforming variables.

Chapter 8 covers multiple regression. Model selection methods are given particular emphasis, because choosing the variables to include in a model is an essential step in many real-life analyses. The topic of confounding is given careful treatment as well.

Chapter 9 discusses some commonly used experimental designs and the methods by which their data are analyzed. One-way and two-way analysis of variance methods, along with randomized complete block designs and 2^p factorial designs, are covered fairly extensively.

Chapter 10 presents the topic of statistical quality control, discussing control charts, CUSUM charts, and process capability; and concluding with a brief discussion of six-sigma quality.

NEW FOR THIS EDITION

The fourth edition of this book is intended to extend the strengths of the third. Some of the changes are:

- A large number of new exercises have been included, many of which involve real data from recently published sources.
- A new section on confidence intervals for a population variance has been added to Chapter 5.
- Chapter 6 now contains material on tests for a population variance.
- The material on goodness-of-fit tests has been expanded.
- The exposition has been improved in a number of places.

RECOMMENDED COVERAGE

The book contains enough material for a year-long course. For a one-semester course, there are a number of options. In our three-hour course at the Colorado School of Mines, we cover all of the first four chapters, except for joint distributions, the more theoretical aspects of point estimation, and the exponential, gamma, and Weibull distributions. We then cover the material on confidence intervals and hypothesis testing in Chapters 5 and 6, going quickly over the two-sample methods and power calculations and omitting distribution-free methods and the chi-square and F tests. We finish by covering as much of the material on correlation and simple linear regression in Chapter 7 as time permits.

A course with a somewhat different emphasis can be fashioned by including more material on probability, spending more time on two-sample methods and power, and reducing coverage of propagation of error, simulation, or regression. Many other options

are available; for example, one may choose to include material on factorial experiments in place of some of the preceding topics. Sample syllabi, emphasizing a variety of approaches and course lengths, can be found on the book website www.mhhe.com/navidi.

McGRAW-HILL CONNECT® ENGINEERING

The online resources for this edition include McGraw-Hill Connect Engineering, a web-based assignment and assessment platform that can help students to perform better in their coursework and to master important concepts. With Connect Engineering, instructors can deliver assignments, quizzes, and tests easily online. Students can practice important skills at their own pace and on their own schedule. Ask your McGraw-Hill Representative for more detail and check it out at www.mcgrawhillconnect.com/engineering.

In addition, the website for *Statistics for Engineers and Scientists*, 4e, features data sets for students, as well as solutions, PowerPoint lecture notes for each chapter, an image library, and suggested syllabi for instructors. The website can be accessed at www.mhhe.com/navidi.

McGRAW-HILL LEARNSMART®

McGraw-Hill LearnSmart® is an adaptive learning system designed to help students learn faster, study more efficiently, and retain more knowledge for greater success. Through a series of adaptive questions, Learnsmart pinpoints concepts the student does not understand and maps out a personalized study plan for success. It also lets instructors see exactly what students have accomplished, and it features a built-in assessment tool for graded assignments. Ask your McGraw-Hill Representative for more information, and visit www.mhlearnsmart.com for a demonstration.

ELECTRONIC TEXTBOOK OPTION

This text may be purchased in electronic form through an online resource known as CourseSmart. Students can access the complete text online at a lower cost than the traditional text. In addition, purchasing the eTextbook allows students to use CourseSmart's web tools, which include full text search, notes, and highlighting, and email tools for sharing notes among classmates. More information can be found at www.CourseSmart.com.

ACKNOWLEDGMENTS

I am indebted to many people for contributions at every stage of development. I received valuable suggestions from my colleagues Barbara Moskal, Gus Greivel, Ashlyn Munson, and Melissa Laeser at the Colorado School of Mines. Mike Colagrosso developed some excellent applets, and Lesley Strawderman developed PowerPoint slides to supplement the text. I am particularly grateful to Jackie Miller of The Ohio State University, who has corrected many errors and made many valuable suggestions for improvement.

The staff at McGraw-Hill has been extremely capable and supportive. In particular, I would like to express my thanks to Developmental Editors Kathryn Neubauer, and Vincent Bradshaw, and Global Publisher Raghu Srinivasan for their patience and guidance in the preparation of this edition.

William Navidi