

INTRODUCTORY **S**TATISTICS

NEIL A. WEISS

**DATA
DISK
ENCLOSED**

3.5" floppy disk contains
data files for exercises,
examples, case studies,
and database

4TH
EDITION

ALTERNATE VERSION
Featuring
MINITAB® for Windows®

Introductory Statistics

• • • • •

Fourth Edition

Alternate Version

Featuring MINITAB® for Windows®

Neil A. Weiss

Arizona State University

Biographies by Carol A. Weiss



Addison-Wesley Publishing Company, Inc.

Reading, Massachusetts Menlo Park, California New York
Don Mills, Ontario Harlow, United Kingdom Amsterdam Bonn
Sydney Singapore Tokyo Madrid San Juan Milan Paris

Senior Sponsoring Editor: Julia Berrisford
Associate Editor: Jennifer Albanese
Editorial Assistant: Sara Peterson
Marketing Manager: Michelle Babinec
Production Supervisor: Patsy DuMoulin
Art Supervisor: Joseph Vetere
Technical Art Consultant: Susan London-Payne
Text Designer: Rebecca Lemna
Cover Designer: Marshall Henrichs

Cover Design Supervisor: Eileen Hoff
Copy Editor: Sally Stickney
Compositor: Carol and Neil Weiss
PostScript Consultant: Arthur Ogawa
Illustrator: Scientific Illustrators
Proofreader: Phyllis Coyne
Prepress Services Buyer: Caroline Fell
Electronic Prepress: Pure Imaging
Manufacturing Coordinator: Evelyn Beaton

Photographs:

Florence Nightingale, page 2, courtesy of The Bettmann Archive.
Adolphe Quetelet, page 42. Copyright Bibliothèque royale Albert I^{er},
Cabinet des Estampes, Bruxelles, J. Odevaëre.
John Tukey, page 100, courtesy of John W. Tukey.
Andrei Kolmogorov, page 180, courtesy of The Bettmann Archive.
James Bernoulli, page 264, courtesy of The Bettmann Archive.
Carl Friedrich Gauss, page 334, courtesy of The Bettmann Archive.
Pierre-Simon Laplace, page 402, courtesy of The Bettmann Archive.
William Gosset, page 438, courtesy of The Granger Collection, New York.
Jerzy Neyman, page 480, courtesy of the Department of Statistics, University of
California, Berkeley.
Gertrude Cox, page 572, courtesy of Research Triangle Institute.
Abraham de Moivre, page 652, courtesy of The Granger Collection, New York.
Karl Pearson, page 686, courtesy of Brown Brothers.
Adrien Legendre, page 734, courtesy of The Bettmann Archive.
Sir Francis Galton, page 792, courtesy of Stock Montage.
Sir Ronald Fisher, page 864, courtesy of The Bettmann Archive.

Library of Congress Cataloging-in-Publication Data

Weiss, N. A. (Neil A.)
Introductory statistics / Neil A. Weiss ; biographies by Carol A.
Weiss. — 4th alternate ed.
p. cm.
Includes index.
ISBN 0-201-54567-5
I. Statistics. I. Title.
QA276.12.W45 1997
519.5—dc20 96-15749
CIP

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and Addison-Wesley was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Copyright © 1997 by Addison-Wesley Publishing Company, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America.

2 3 4 5 6 7 8 9 10 DOC 00 99 98 97

Reprinted with corrections, May 1997.

Preface

.

Statistics has become an indispensable tool in business, government, and virtually every academic discipline. Some familiarity with statistics is essential for all of us if we are to comprehend the world around us.

The purpose of this book is to provide a clear understanding of basic statistical concepts and techniques and to present well-organized procedures for applying them. Introductory high-school algebra is a sufficient prerequisite.

The text is designed to be flexible. It can be used in a one-quarter, one-semester, two-quarter, two-semester, or three-quarter course. The amount of time devoted to the book can be varied by both choice of topics and depth of coverage.

Technological advances and ever increasing calls for new approaches to presenting statistics have made this an exciting time to learn, practice, and teach statistics. In writing the fourth edition of *Introductory Statistics*, we have incorporated many of the techniques and attitudes that reflect recent developments in statistics.

Features

The text contains the following features that will provide valuable assistance for the reader in learning introductory statistics.

Emphasis on application. We have concentrated on the application of statistical techniques to the analysis of data. Although statistical theory has been kept to a minimum, we have provided a thorough explanation of the rationale for using each statistical procedure.

Data analysis and exploration. We agree wholeheartedly with the trend of including more exploratory and confirmatory data analysis in statistics courses and have incorporated an extensive amount into the text and exercises. We also recognize, however, that not all readers will have access to computers and therefore have presented data analysis in a way that does not require using a computer even though access to one is recommended.

Detailed and careful explanations. We have included every step of explanation we think a typical reader might need. Our guiding principle is to avoid “cognitive jumps” and thereby make the learning process smooth and enjoyable. We believe detailed and careful explanations result in better understanding.

Data sets. In most examples and exercises, we have presented raw data in addition to summary statistics. This gives a more realistic view of statistics and provides an opportunity for the problems to be solved by computer, if so desired.

Procedure boxes. To help the reader learn to apply statistical procedures, we have developed easy-to-follow, step-by-step methods for carrying out those procedures. For ease in locating, each procedure is displayed with a color background. A unique feature of this book is that when a procedure is illustrated by an example, each step in the procedure is presented again within the example. This serves a twofold purpose: It shows how the procedure is applied and helps the reader master the steps in the procedure.

Procedure index. Given the numerous statistical procedures, it is sometimes difficult to find a specific one, especially when the book is being used for reference purposes. Consequently, we have included a *procedure index*. Located on the last page of the book, the procedure index provides a quick and easy way to find the required procedure for performing any particular statistical analysis.

Computer simulations. We have incorporated many computer simulations in both the text and the exercises. These serve as pedagogical aids for understanding complex concepts such as sampling distributions. Readers can benefit from this material even if they do not have access to a computer.

Computer usage. Today, virtually all professional applications of statistics are done by computer. It is therefore important that every student of statistics have at least some familiarity with statistical software. Although we have chosen Minitab[†] to illustrate the use of statistical software, the text has been written so that the instructor is free to select other statistical software packages. The Minitab discussion in this book is self-contained.[‡] All computer material is *optional*, but recommended.

The computer sections are integrated as optional subsections occurring immediately following the particular statistical concept under consideration. In each subsection we explain how Minitab can be used to solve problems that were solved by hand earlier in the section. Each solution consists of introducing the required commands, displaying the computer output, and interpreting the results.

Additionally, computer exercises (clearly marked as such) are incorporated into the exercise sets. Three types of computer exercises have been included. The first type asks the reader to interpret computer printouts; no knowledge of or access to statistical software is necessary for these exercises. The second type asks the reader to use Minitab or some other statistical software to solve exercises that were presented previously for hand solution; all Minitab commands required for these computer exercises will have already been discussed in the text. The third type of computer exercise asks the reader to use statistical software to perform a computer simulation; these exercises are designed to provide concrete illustrations of some of the more complex concepts (e.g., sampling distributions) and to show the reader how the computer can be used to reveal statistical facts.

[†] Minitab is a registered trademark of Minitab Inc., 3081 Enterprise Drive, State College, PA 16801. Telephone: 814-238-3280. Fax: 814-238-4383. World Wide Web: <http://www.minitab.com>. We would like to thank Minitab Inc., for their assistance.

[‡] Additional details and further topics, such as writing and using macros, are provided in the *Minitab Supplement* to this book.

For maximum flexibility we have allowed for three options in computer coverage: (1) no coverage, (2) coverage omitting the Minitab-command discussions but including the interpretation of computer printouts, and (3) coverage of Minitab commands (or those for some other statistical software) and the interpretation of computer printouts.

Minitab® for Windows® menu interface. We feature Release 10 of the professional version of Minitab for Windows. Only minor differences exist between the professional versions of Minitab for Windows and Macintosh® and, aside from worksheet size, the same is true for *The Student Edition of MINITAB for Windows*.

Traditionally, Minitab has been applied by typing (session) commands at the MTB > prompt. But in response to the strong emergence of graphical user interface (GUI), Minitab now provides a menu interface as well, so that commands can also be executed by choosing them from menus and completing dialog boxes. In keeping with the trend toward GUI, we present only the menu interface in this book. Those wishing to use the session interface are referred to the original version of *Introductory Statistics, Fourth Edition*, or to *Elementary Statistics, Third Edition*, both by Neil A. Weiss and published by Addison-Wesley.

Minitab Quick Reference. For quick access and reference to the Minitab menu instructions discussed in this book, we have provided a Minitab Quick Reference (MQR). Located opposite the last page of the book, the MQR includes the procedures, their menu instructions, and page-number references for more details.

Biographical sketches. Each chapter begins with a brief biography of a famous statistician. Besides being of general interest, these biographies help the reader obtain a perspective on how the science of statistics developed.

Chapter introductions and chapter outlines. Also included at the beginning of each chapter is a description of the chapter and an explanation of how the chapter relates to the text as a whole. As a further aid, a chapter outline, which follows the chapter introduction, lists the sections in the chapter.

Case studies. We have presented case studies ranging from the classic to the current. At the beginning of each chapter, a case study is described briefly; at the end of the chapter, when the reader has studied the required concepts, the case study is considered in detail. Exercises are provided for all case studies.

Definitions, formulas, key facts. As an aid to learning and for reference, we have prominently displayed all definitions, formulas, and key facts. These items are enclosed by a color rectangle to make them easy to locate.

Real examples. Because we believe that the majority of students learn by example, every concept discussed in the book is illustrated by at least one detailed example. The examples are, for the most part, based on real-life situations and have been chosen for their interest as well as for their illustrative value.

Extensive and diverse exercise sets. We have constructed exercise sets that are both extensive and diverse. Most of the exercises are based on information found in newspapers, magazines, statistical abstracts, and journal articles; sources are explicitly cited. The exercises are designed not only to help the reader learn the material but also to show that statistics is a lively and relevant discipline.

Since students in introductory statistics courses often have different mathematical backgrounds, we have included three levels of exercises: basic, intermediate, and advanced.

Every exercise set contains several basic exercises. The *basic exercises* provide applications of material presented in the text, and every reader should master these. We have organized the basic exercises so that each concept is covered by at least two problems; for each odd-numbered basic exercise that involves a particular concept, there is also an even-numbered basic exercise that involves that same concept. A single color bullet (•) preceding an exercise number identifies the exercise as a basic one. The answers to the odd-numbered basic exercises are presented in Appendix C; the answers to the even-numbered basic exercises are in the *Instructor's Solutions Manual*.

Most exercise sets also include intermediate and advanced problems. The *intermediate exercises* contain supplementary material that is not necessarily covered in the text but that may interest some of the more highly motivated students. Two color bullets (• •) preceding an exercise number identify the exercise as an intermediate one. The *advanced exercises* cover abstract concepts, theory, and algebraic derivations. Those exercises are intended for students with special mathematical background and aptitude. Three color bullets (• • •) preceding an exercise number identify the exercise as an advanced one. The solutions to all intermediate and advanced exercises are in the *Instructor's Solutions Manual*.

Chapter reviews. Students in introductory statistics courses often experience some anxiety and confusion about how to study and review. To help the student, a chapter-review section appears at the end of every chapter. The chapter reviews include (1) a list of key terms with page references, (2) formulas with page references, (3) chapter objectives, and (4) a review test. These pedagogical aids provide the student with an organized method for reviewing and studying. The answers to the review tests are given in Appendix C.

Database exercises. Appendix B contains a printout of a database obtained by randomly selecting 500 Arizona State University sophomores. Seven variables are considered for each student: sex, high-school GPA, SAT math score, cumulative GPA, SAT verbal score, age, and total hours. At the end of each chapter review, a section entitled “Using the Focus Database” presents several exercises about the database. These exercises are optional and are to be done by computer. We assume that the student knows or will be taught any additional commands required to carry out these exercises.

Formula/table card. A detachable formula/table card is provided with the book. This card contains all of the formulas and many of the tables that appear in *Introductory Statistics*. The formula/table card is useful for quick-reference purposes; many instructors also find it convenient for use with examinations.

Organization

The text offers a great deal of flexibility in choosing material to cover. Chapter 1 introduces the nature of statistics, sampling designs, observational studies, and designed experiments; an optional introduction to Minitab is presented as well. Chapters 2 and 3 provide the fundamentals of descriptive statistics.

Chapters 4–6 present probability, discrete random variables, and the normal distribution. Chapter 7 introduces the concept of sampling distributions and provides a detailed discussion of the sampling distribution of the mean. Following that, Chapters 8 and 9 ex-

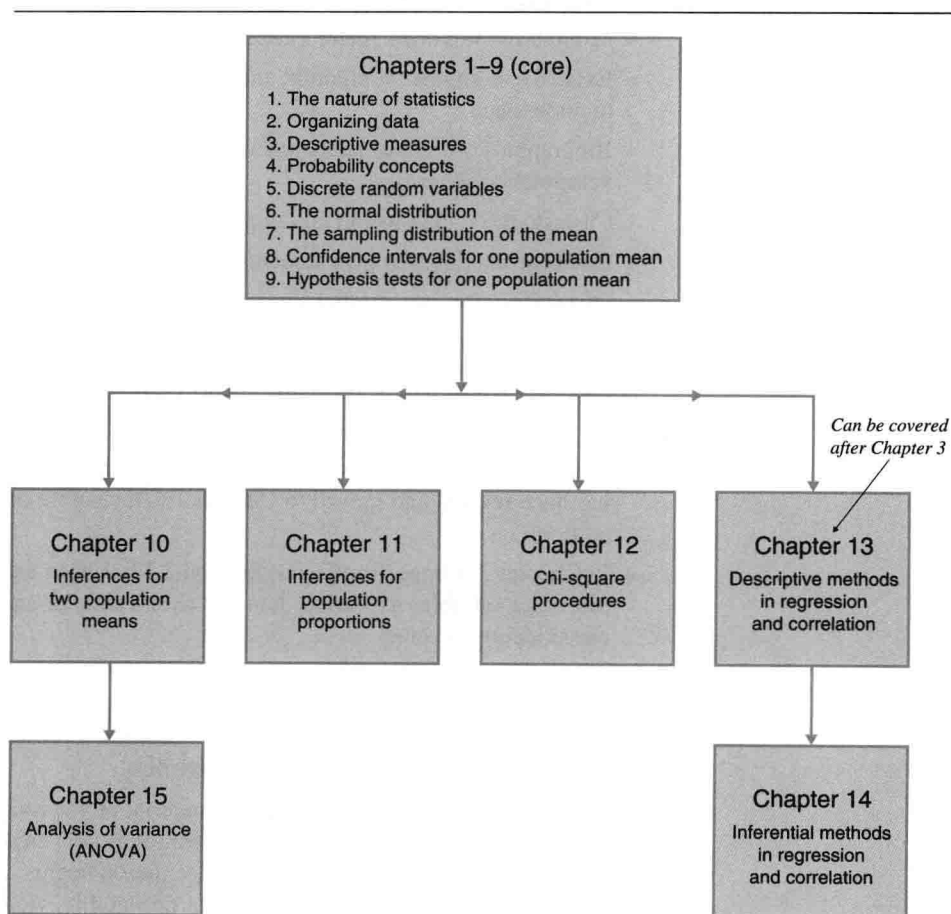
amine confidence intervals and hypothesis tests for one population mean. We consider Chapters 1–9 the core of an introductory statistics course.

Chapter 10 presents inferences for two population means, Chapter 11 inferences for population proportions, and Chapter 12 chi-square procedures (goodness-of-fit test, independence test, and inferences for a population standard deviation).

We have divided the traditional material on regression and correlation into two chapters. Chapter 13 examines descriptive methods in regression and correlation and can be covered at any time after Chapter 3. Chapter 14 presents inferential methods in regression and correlation and can be covered once Chapters 9 and 13 have been completed.

Chapter 15 examines one-way ANOVA and two-way ANOVA with interaction. Also included are optional sections on multiple comparisons and the Kruskal–Wallis test.

The following flowchart summarizes the preceding discussion and shows the interdependence among chapters. In the flowchart, the prerequisites for a given chapter consist of all chapters having a path leading to that chapter.



Changes in the fourth edition

We have made significant changes in the fourth edition of *Introductory Statistics*. All chapters have been rewritten for the purpose of updating and expanding the material, fine-tuning the organization, and adding new sections where appropriate. The entire manuscript has been edited to provide a smoother and more succinct presentation.

Among the global enhancements to the fourth edition of *Introductory Statistics* are:

- An increase in the number of exercises from approximately 1500 in the third edition to almost 1700 (not including parts) in the fourth edition.
- Extensive use of graphics to illustrate and explain concepts.
- Expanded coverage of data analysis, both exploratory and confirmatory.
- Increased interpretation of concepts in text and in exercises.
- Optional integration of nonparametric methods with the corresponding parametric methods.
- Fully expanded treatment of P -values with all hypothesis tests.
- Minitab for Windows menu interface.
- Extensive computer simulation and resampling to make complex concepts easier to understand.
- Biographies of famous statisticians to help students gain a perspective of the development of statistics.
- Case studies—classical to contemporary—with exercises.
- Database exercises to give students practice with very large data sets.
- Discussion of observational studies and designed experiments.
- Flowcharts for choosing the correct statistical procedure.

A complete list of the local enhancements to the fourth edition of *Introductory Statistics* is too long to present here; but a list of some of the most important ones follows.

- Material on sampling designs, which occurred in Chapter 16 of the third edition, has been revised and moved to Chapter 1 (The Nature of Statistics) in the fourth edition.
- In Chapter 2 (Organizing Data), Section 2.1 has been extensively revised: it now discusses variables as well as data and uses a simpler and more modern data-classification scheme.
- Chapter 3 (Descriptive Measures) has been revised throughout to provide a more data analytic tone. Section 3.6 contains the major changes: a thorough discussion of outliers, a general definition of quartiles, a presentation of the interquartile range, and the inclusion of modified boxplots.
- In Chapter 4 (Probability Concepts), material on the relative-frequency interpretation of probability has been inserted to provide an intuitive description of probability before the more formal presentation begins. Also, Venn diagrams have been used to introduce the rules of probability, yielding an informal and easy-to-understand approach.

- Probability histograms have been included throughout Chapter 5 (Discrete Random Variables) to provide visual displays of discrete probability distributions. A new and optional Section 5.6 covers the Poisson distribution.
- In Chapter 6 (The Normal Distribution), we have gone to a cumulative normal table, areas between $-\infty$ and z , because it is simpler to learn and more efficient to use than the normal table giving areas between 0 and z ; the cumulative table is also consistent with Minitab's CDF and INVCDF commands, which are discussed in a new and optional subsection of Section 6.1. A new Section 6.5 presents normal probability plots.
- To give the reader an intuitive understanding of the sampling distribution of the mean, computer simulations in both the normal and nonnormal cases have been incorporated into Chapter 7 (The Sampling Distribution of the Mean).
- In Chapter 8 (Confidence Intervals for One Population Mean), explanations have been included on how to deal with outliers when obtaining confidence intervals. In Section 8.4 a computer simulation has been used to compare the distributions of the z -statistic (standard normal) and the t -statistic (Student's t).
- In Chapter 9 (Hypothesis Tests for One Population Mean), data analytic techniques have been employed to help decide on the correct procedure to apply. An explanation of the difference between statistical and practical significance has been included in Section 9.3. An optional Section 9.4 discusses Type II error probabilities and power; an optional Section 9.7 covers the Wilcoxon one-sample signed-rank test; and a new and optional Section 9.8 summarizes the hypothesis-testing procedures and presents a flowchart for deciding which procedure to use.
- In Chapter 10 (Inferences for Two Population Means), an optional Section 10.4 provides a discussion of the Mann–Whitney test; a new and optional subsection of Section 10.5 presents the Wilcoxon paired-sample signed-rank test; and an optional Section 10.6 summarizes the hypothesis-testing procedures and presents a flowchart for deciding which procedure to use.
- Chapter 11 (Inferences for Population Proportions) is new; it is comprised roughly of the proportion material found in Chapters 8, 9, and 10 of the third edition. New topics include margin of error and sample-size determination for one and two proportions.
- In Chapter 12 (Chi-Square Procedures), segmented bar graphs have been included in Section 12.3 to provide a visual interpretation of statistical independence for two characteristics of a population. Section 12.4 contains a simulation of a chi-square distribution, obtained by resampling a normal distribution and calculating $(n - 1)s^2/\sigma^2$ for each sample.
- Chapter 13 (Descriptive Methods in Regression and Correlation) now contains a discussion of influential observations and outliers. Also, the terms *predictor variable* and *response variable* have been introduced explicitly in Section 13.2.
- Chapter 14 (Inferential Methods in Regression and Correlation) now includes a discussion of analysis of residuals for checking the assumptions for regression

inferences. A new and optional Section 14.6 presents a test for normality based on the correlation between the sample data and its normal scores; the test is essentially equivalent to the powerful Shapiro–Wilk test.

- In Chapter 15 (Analysis of Variance), the material on one-way ANOVA has been revised and split into two sections, one presenting the logic and one the procedure. A new and optional Section 15.4 covers multiple comparisons; a new and optional Section 15.5 examines the Kruskal–Wallis test; and two new sections on two-way ANOVA with interaction present the logic (Section 15.6) and the procedure (Section 15.7).

Supplements and other support

The following supplements have been prepared to accompany the alternate fourth edition of *Introductory Statistics*.

Minitab Supplement. This supplement, written by Professor Peter W. Zehna, provides in-depth coverage of Minitab, augmenting that given in the book. It is designed to be used in conjunction with *Introductory Statistics* and is keyed to the book. No prerequisite knowledge of computers or statistical software is presumed.

Instructor's Solutions Manual. This manual, prepared by Professors Bernard Morzuch, Chris Franklin, and David Lund, contains detailed, worked-out solutions to all exercises in *Introductory Statistics*.

Student's Solutions Manual. This manual, also by Professors Bernard Morzuch, Chris Franklin, and David Lund, presents detailed, worked-out solutions to every fourth exercise in *Introductory Statistics*.

OmniTest II Computerized Testing. This unique testing software offers a virtually endless supply of quizzes, tests, final examinations, and additional instructional exercises. Features include multiple-exam versions, customized question editing, on-screen preview and edit functions, and pull-down menus. *OmniTest II* is available for use with IBM PCs or compatibles.

Printed Test Bank. This supplement provides several printed examinations for each chapter of *Introductory Statistics*.

DataDisk with Instruction Sheet. This supplement is included with the book and can be found inside the back cover. DataDisk is a floppy disk that contains text files for the data sets appearing in the (1) computer-section examples, (2) odd-numbered basic exercises, (3) Focus database, and (4) case studies. DataDisk makes it possible to store and analyze these data sets without having to enter them manually. The instruction sheet explains how to import data from DataDisk into Minitab.

Transparency Masters. Many of the text's figures, tables, and procedure boxes have been reproduced on transparency masters for classroom use.

Additional support is available with the following items.

STAT101. This inexpensive statistical software package is based on Release 6 of Minitab and is distributed by Addison-Wesley. Working with a spreadsheet-like Data Editor that allows up to 2000 data points, STAT101 provides an excellent foundation for progression to other, more powerful Minitab software products.

The Student Editions of Minitab. These statistical software packages, also distributed by Addison-Wesley, are available in three versions: for Windows, Release 8 for DOS, and Release 8 for the Macintosh. The student editions of Minitab allow the user access to 3500 data points and to a full range of Minitab functionalities and graphic capabilities, as well as a wide range of data sets drawn from business, the social sciences, and the physical sciences. Also accompanying the software is a user's manual that includes case studies and hands-on tutorials for using the software.

Acknowledgments

It is our pleasure to thank the following reviewers, whose comments and suggestions most certainly improved the fourth edition of *Introductory Statistics*:

Jasper Adams	Eugene Enneking
Stephen F. Austin State University	Portland State University
Larry Ammann	Ruby Evans
University of Texas, Dallas	Sante Fe Community College
Jerald T. Ball	Dale O. Everson
Las Positas College	University of Idaho
Mary Sue Beersman	Charles M. Farmer
Northeast Missouri State University	James Madison University
Gary B. Beus	Chris Franklin
Brigham Young University	University of Georgia
Jerry Bloomberg	Jeff Frost
Essex Community College	Johnson County Community College
Patricia Buchanan	Joe Fred Gonzalez, Jr.
Penn State University	University of Maryland
Curtis Church	Shu-Ping Hodgson
Middle Tennessee State University	Central Michigan University
Gabie Church	Mark E. Johnson
Louisiana State University	University of Central Florida
Patti Collings	Debra Landre
Brigham Young University	San Joaquin Delta College
Rickie J. Domangue	Benny Lo
James Madison University	Ohlone College

(continued)

David R. Lund
University of Wisconsin, Eau Claire
Rhonda Magel
North Dakota State University
Linda Malone
University of Central Florida
Bernard J. Morzuch
University of Massachusetts, Amherst
Tom Ribley
Valencia Community College
Larry Ringer
Texas A&M University
Gaspard Rizzuto
University of Southwestern Louisiana
C. Bradley Russell
Clemson University
Leroy Sathre
Valencia Community College

Robert L. Schaefer
Miami University
Franklin Sheehan
San Francisco State University
Donald Sisson
Utah State University
Duane Steffey
San Diego State University
Larry Stephens
University of Nebraska, Omaha
Ram C. Tiwari
University of North Carolina
Joseph J. Walker
Georgia State University
Lyndon C. Weberg
University of Wisconsin, River Falls
Calvin L. Williams
Clemson University

We also wish to thank Professor Michael Driscoll for helping us select the statisticians for the biographical sketches; and Professors Charles Kaufman, Sharon Lohr, Walter Reid, and Dennis Young, with whom we have had several illuminating discussions. Our special thanks go to Professors Matthew Hassett and Ronald Jacobowitz for their many helpful comments and suggestions.

To Professor Larry Griffey, we express our appreciation for his formula/table card. We are grateful to Professor Peter Zehna for preparing the *Minitab Supplement* and for many interesting and informative conversations. Our thanks go as well to Professors Bernard Morzuch, David Lund, and Chris Franklin and her assistants, Kristie Ball and Amy Deitrick, for writing the *Instructor's Solutions Manual* and *Student's Solutions Manual*. In addition, we thank Professor Bernard Morzuch for also overseeing the development of the OmniTest computerized-testing software. Professors Joe Fred Gonzalez, Jr., and Larry Stephens did an outstanding job verifying the solutions to the examples and the answers to the exercises; our sincere thanks to them.

We are grateful to Dr. William Feldman and Mr. Frank Crosswhite for providing the data from their study on the Golden Torch Cactus; to Professor Thomas A. Ryan, Jr., for his correspondence concerning the correlation test for normality; and to Dr. George McManus and Mr. Gregory Weiss for supplying the data from their study of zooplankton nutrition in the Gulf of Mexico.

We also extend our appreciation to Professor John Tukey for taking the time to provide us with autobiographical information and a photo of himself; to Alison Stern-Dunyak of the American Statistical Association for supplying biographical information on Gertrude Cox; to Maureen Quinn of The Gallup Organization for providing the information used in the case study of Chapter 12; to Cathy Akritas and Rob Kelly of Minitab Technical Services; to Pam McHugh, Susan Shirey, Beth Solt, and Melissa Meehan of Minitab Author's

Assistance Program; and to Jennifer Wang for obtaining and updating much of the data used in the examples and exercises in the book. Our appreciation also goes to Howard Blaut and Rick Hanna for providing data on real estate; to Jeffrey Jirele for supplying data on automobile insurance; and to Mary Neary for furnishing the Focus database.

Our thanks to our cover designer, Marshall Henrichs, and our cover design supervisor, Eileen Hoff; to our text designer, Rebecca Lemna; to our technical art consultant, Susan London-Payne; to our art supervisor, Joe Vetere; to our proofreader, Phyllis Coyne et al.; to our prepress buyer, Caroline Fell; to our PostScript consultant, Arthur Ogawa; and to our manufacturing coordinator, Evelyn Beaton. Thanks also to Berthold Horn, Louis Vosloo, and Jordan Backler at Y&Y, Inc.; to George Morris, Brian Morris, and the others at Scientific Illustrators; and to John Smioldo at Pure Imaging.

Special thanks to our editors, Jennifer Albanese and Julia Berrisford; to our production supervisor, Patsy DuMoulin; to our editorial assistant, Sara Peterson; to our marketing manager, Michelle Babinec; and to our copy editor, Sally Stickney.

Finally, we would like to thank Carol Weiss. Apart from writing the text, she was involved in every aspect of development and production. Moreover, Carol researched and wrote the biographies and took on the task of typesetter using the T_EX typesetting system.

Tempe, Arizona

N.A.W.

Contents

.

Part I Introduction 1

. . .

Chapter 1 The Nature of Statistics 3

.

- 1.1 Two Kinds of Statistics 4
- 1.2 Classifying Statistical Studies 6
- 1.3 The Development of Statistics 9
- 1.4 Using the Computer (Optional) 10
- 1.5 Is a Study Necessary? 18
- 1.6 Simple Random Sampling 18
- 1.7 Other Sampling Procedures 26
- 1.8 Observational Studies and Designed Experiments 31
 - Chapter Review 34
 - *Case Study: Are Americans Lazy?* 38

Part II Descriptive Statistics 41

. . .

Chapter 2 Organizing Data 43

.

- 2.1 Variables and Data 44
- 2.2 Grouping Data 49
- 2.3 Graphs and Charts 62
- 2.4 Stem-and-Leaf Diagrams 76

- 2.5 Distribution Shapes; Symmetry and Skewness 83
- 2.6 Misleading Graphs 90
 - Chapter Review 94
 - *Case Study: Infant Mortality in Developed Nations* 98

Chapter 3 Descriptive Measures 101

-
- 3.1 Measures of Central Tendency 102
- 3.2 Summation Notation; the Sample Mean 113
- 3.3 Measures of Dispersion; the Sample Standard Deviation 117
- 3.4 Interpretation of the Standard Deviation; z-Scores 129
- 3.5 Computing \bar{x} and s for Grouped Data 138
- 3.6 The Five-Number Summary; Boxplots 146
- 3.7 Descriptive Measures for Populations; Use of Samples 160
 - Chapter Review 172
 - *Case Study: Per Capita Income by State* 177

Part III Probability, Random Variables, and Sampling Distributions 179

• • •

Chapter 4 Probability Concepts 181

-
- 4.1 Introduction; Classical Probability 182
- 4.2 Events 190
- 4.3 Some Rules of Probability 200
- 4.4 Contingency Tables; Joint and Marginal Probabilities 209
- 4.5 Conditional Probability 215
- 4.6 The Multiplication Rule; Independence 224
- 4.7 Bayes's Rule (Optional) 235
- 4.8 Counting Rules (Optional) 243
 - Chapter Review 256
 - *Case Study: Lotto—The Arizona State Lottery* 262

Chapter 5	Discrete Random Variables	265
•		
5.1	Discrete Random Variables; Probability Distributions	266
5.2	The Mean and Standard Deviation of a Discrete Random Variable	277
5.3	Binomial Coefficients; Bernoulli Trials	289
5.4	The Binomial Distribution	297
5.5	The Mean and Standard Deviation of a Binomial Random Variable	313
5.6	The Poisson Distribution (Optional)	317
• Chapter Review		328
• <i>Case Study: Aces Wild on the Sixth at Oak Hill</i>		333
Chapter 6	The Normal Distribution	335
•		
6.1	The Standard Normal Curve	336
6.2	Normal Curves	348
6.3	Normally Distributed Populations	358
6.4	Normally Distributed Random Variables	369
6.5	Normal Probability Plots	380
6.6	The Normal Approximation to the Binomial Distribution	388
• Chapter Review		397
• <i>Case Study: Chest Sizes of Scottish Militiamen</i>		401
Chapter 7	The Sampling Distribution of the Mean	403
•		
7.1	Sampling Error; the Need for Sampling Distributions	404
7.2	The Mean and Standard Deviation of \bar{x}	411
7.3	The Sampling Distribution of the Mean	420
• Chapter Review		431
• <i>Case Study: The Chesapeake and Ohio Freight Study</i>		435