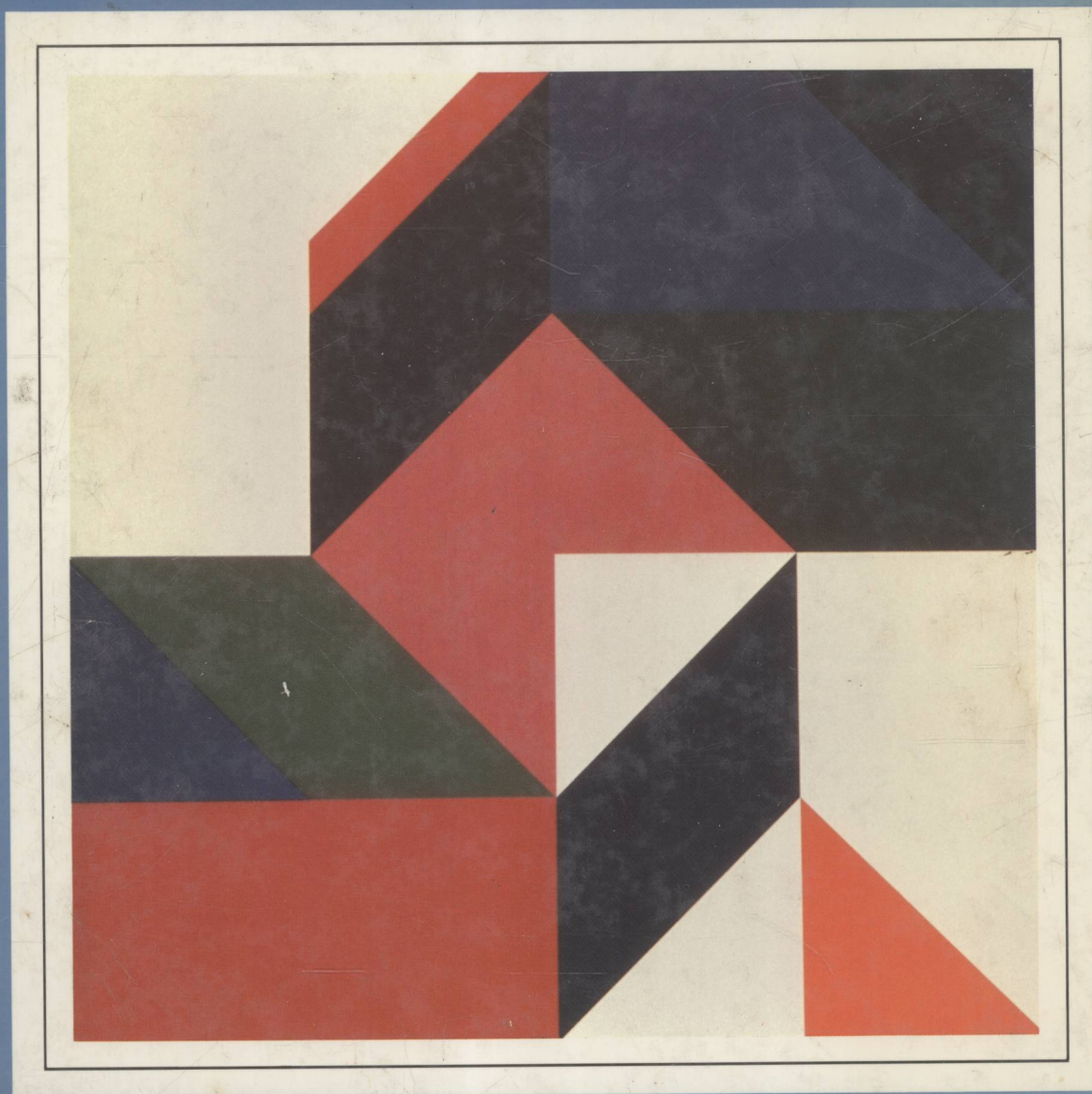


UNDERSTANDING STATISTICS

Fifth Edition

BRIDGEWATER COLLEGE
BOOKSTORE



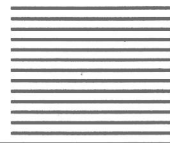
OTT ■ MENDENHALL

0212

089
E.5

9761149

FIFTH EDITION



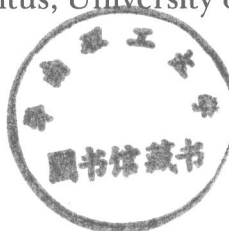
UNDERSTANDING STATISTICS

Lyman Ott

Merrell Dow Research Institute

William Mendenhall

Professor Emeritus, University of Florida



E9761149



PWS-KENT Publishing Company
Boston



PWS-KENT
Publishing Company

20 Park Plaza
Boston, Massachusetts 02116

Sponsoring Editor: Michael Payne
Assistant Editor: Marcia Cole
Production Editor: Susan L. Krikorian
Production Services: Sara Hunsaker, Ex Libris
Text and Cover Designer: Susan L. Krikorian
Compositor: Weimer Typesetting, Inc.
Manufacturing Coordinator: Margaret Sullivan Higgins
Cover Printer: Henry N. Sawyer Company, Inc.
Text Printer/Binder: R. R. Donnelley & Sons Company

Cover art: *Ariel's Conation* by Michael Lasuchin. Reprinted by permission of the artist.

© 1990 by PWS-KENT Publishing Company. All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transcribed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, PWS-KENT Publishing Company.

PWS-KENT Publishing Company is a division of Wadsworth, Inc.

Printed in the United States of America

2 3 4 5 6 7 8 9 — 94 93 92 91 90

Library of Congress Cataloging-in-Publication Data

Ott, Lyman.

Understanding statistics / Lyman Ott, William Mendenhall. — 5th ed.

p. cm.

Includes bibliographical references.

1. Statistics. I. Mendenhall, William. II. Title.

QA276.12.087 1990






001.4'22—dc20






89-16136





CIP

A partially programmed study guide with self-correcting exercises has been developed to assist you in mastering difficult statistical concepts you may encounter in this text. It is available from your local bookstore under the name *Study Guide for Understanding Statistics*, 5th Edition, by Barbara and Robert Beaver.

Key to Exercise Symbols

	environmental studies
	engineering/technical
	business/economics
	political science/ government
	agriculture

	psychology
	education
	sociology/ public opinion
	chemistry
	medicine

	biology
	geology
	physics
	law

Population mean

$$\mu$$

Sample mean

$$\bar{x} = \frac{\sum x}{n}$$

Population variance

$$\sigma^2$$

Sample variance

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \left[\sum x^2 - \frac{(\sum x)^2}{n} \right]$$

Population standard deviation

$$\sigma$$

Sample standard deviation

$$s$$

Population proportion
(binomial distribution)

$$\pi$$

Sample proportion

$$\hat{\pi} = \frac{x}{n}$$

Confidence interval for μ ,
 σ known

$$\bar{x} \pm z\sigma_{\bar{x}}$$

Statistical test of $H_0: \mu = \mu_0$,
 σ known

$$z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}$$

Confidence interval for μ ,
 σ unknown

$$\bar{x} \pm t \frac{s}{\sqrt{n}}$$

Statistical test of $H_0: \mu = \mu_0$,
 σ unknown

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Confidence interval for
 $\mu_1 - \mu_2$

$$\bar{x}_1 - \bar{x}_2 \pm t_{s_p} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Statistical test of H_0 :
 $\mu_1 - \mu_2 = 0$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Confidence interval for π

$$\hat{\pi} \pm z\sigma_{\hat{\pi}}$$

Statistical test of $H_0: \pi = \pi_0$

$$z = \frac{\hat{\pi} - \pi_0}{\sigma_{\hat{\pi}}}$$

Confidence interval for $\pi_1 - \pi_2$

$$\hat{\pi}_1 - \hat{\pi}_2 \pm z\sigma_{\hat{\pi}_1 - \hat{\pi}_2}$$

Statistical test of H_0 :
 $\pi_1 - \pi_2 = 0$

$$z = \frac{\hat{\pi}_1 - \hat{\pi}_2}{\sigma_{\hat{\pi}_1 - \hat{\pi}_2}}$$

Test of $H_0: \sigma^2 = \sigma_0^2$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Test of $H_0: \sigma_1^2 = \sigma_2^2$

$$F = \frac{s_1^2}{s_2^2}$$

Least squares line

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Slope of least squares line

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$

Intercept of least squares line

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Test of $H_0: \beta_1 = 0$

$$t = \frac{\hat{\beta}_1}{\sqrt{s_e^2/S_{xx}}}$$

Population correlation coefficient ρ

Sample correlation coefficient

$$\hat{\rho} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Paired-difference test

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

Chi-square test of independence

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Spearman's rank correlation
coefficient

$$\hat{\rho}_s = 1 - \frac{6(\sum d^2)}{n(n^2 - 1)}$$

UNDERSTANDING STATISTICS

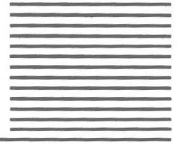
The Duxbury Advanced Series in Statistics and Decision Sciences

Applied Nonparametric Statistics, Second Edition, Daniel
Applied Regression Analysis and Other Multivariable Methods, Second Edition,
Kleinbaum, Kupper, and Muller
Classical and Modern Regression with Applications, Second Edition, Myers
Elementary Survey Sampling, Fourth Edition, Scheaffer, Mendenhall, and Ott
Introduction to Contemporary Statistical Methods, Second Edition, Koopmans
Introduction to Probability and Its Applications, Scheaffer
Introduction to Probability and Mathematical Statistics, Bain and Engelhardt
Linear Statistical Models, Bowerman and O'Connell
Probability Modeling and Computer Simulation, Matloff
Quantitative Forecasting Methods, Farnum and Stanton
Time Series Analysis, Cryer
Time Series Forecasting: Unified Concepts and Computer Implementation, Second Edition,
Bowerman and O'Connell

The Duxbury Series in Statistics and Decision Sciences

Applications, Basics, and Computing of Exploratory Data Analysis, Velleman and Hoaglin
A Course in Business Statistics, Second Edition, Mendenhall
Elementary Statistics, Fifth Edition, Johnson
Elementary Statistics for Business, Second Edition, Johnson and Siskin
Essential Business Statistics: A Minitab Framework, Bond and Scott
Fundamentals of Biostatistics, Third Edition, Rosner
Fundamentals of Statistics in the Biological, Medical, and Health Sciences, Runyon
Fundamental Statistics for the Behavioral Sciences, Second Edition, Howell
Introduction to Probability and Statistics, Seventh Edition, Mendenhall
An Introduction to Statistical Methods and Data Analysis, Third Edition, Ott
Introductory Business Statistics with Microcomputer Applications, Shiffler and Adams
Introductory Statistics for Management and Economics, Third Edition, Kenkel
Mathematical Statistics with Applications, Fourth Edition, Mendenhall, Wackerly, and
Scheaffer
Minitab Handbook, Second Edition, Ryan, Joiner, and Ryan
Minitab Handbook for Business and Economics, Miller
Operations Research: Applications and Algorithms, Winston
Probability and Statistics for Engineers, Third Edition, Scheaffer and McClave
Probability and Statistics for Modern Engineering, Second Edition, Lapin
Statistical Experiments Using BASIC, Dowdy
Statistical Methods for Psychology, Second Edition, Howell
Statistical Thinking for Behavioral Scientists, Hildebrand
Statistical Thinking for Managers, Second Edition, Hildebrand and Ott
Statistics: A Tool for the Social Sciences, Fourth Edition, Ott, Larson, and Mendenhall
Statistics for Business and Economics, Bechtold and Johnson
Statistics for Management and Economics, Sixth Edition, Mendenhall, Reinmuth, and
Beaver
Understanding Statistics, Fourth Edition, Ott and Mendenhall

PREFACE



This text is designed for a one-quarter or one-semester introductory course in statistics. It can be used in those colleges and universities that teach a general course in statistics appropriate for students majoring in, or intending to major in, many different areas. The approach, examples, and exercises provide a basic knowledge of statistical concepts that will be useful in business and in the biological, social, and physical sciences.

The focus of this edition is similar to that for the fourth edition. One of the primary objectives of an introductory statistics course is to develop in the student an understanding and appreciation of the role of statistics in society. The fifth edition of *Understanding Statistics* emphasizes that statistics, as a subject, is the study of making sense of data. Thus, after completing an introductory course in statistics, students should have a *basic understanding* of how to make sense of data. In this text, we approach the study of statistics by considering four steps in making sense of data: (1) gathering methods, (2) methods for summarizing data, (3) methods for analyzing data, and (4) ways to report the results of analyses. Most texts focus on the summarization and analysis steps; we emphasize that all problems involve the collection, summarization, analysis, and reporting steps of making sense of data. We do not, however, spend a great deal of time on teaching statistical calculations. Rather, with calculators and computers available to do the calculations required in analyzing data, we try to focus on the inferences that are made after the calculations have been done.

The use of this text does not require that students have access to a computer and standard statistical software. However, we do discuss the use of calculators and computers in statistics (Chapter 2) and, in an optional section at the end of each chapter, we give reusable programs and output to show how two software systems (SAS and Minitab) can be used to perform some of the calculations of that chapter. Exercises with computer output also provide opportunities for students to understand and interpret the results of a statistical analysis presented as computer output.

To summarize, the fifth edition provides an introduction to statistics with strong emphasis on making sense of data. From this text, students gain an understanding and appreciation of the role of statistics in society and of the steps and methods used in making sense of data. Since computers are becoming so much a part of our lives, an optional section is included in each chapter to show where computer software can be of assistance in performing the calculations of statistics. Important features of this edition include the following:

- Organizational and textual changes to emphasize the four steps in making sense of data: collecting data, summarizing data, analyzing data, and reporting data.
- A new chapter on methods for gathering data (Chapter 3). This chapter discusses some of the survey methods and designs of scientific studies used in collecting data.
- A new chapter on the reporting of results of statistical analyses (Chapter 18).
- A brief discussion of observational studies (Section 5.7).
- An early discussion of methods for summarizing data from more than one variable (Chapter 5).
- New sections on discussing assumptions (Sections 8.7, 9.5, and 15.4).
- Discussion of inferences about β_0 (Section 13.2).
- A new section on predicting y for a given value of x (Section 13.4).
- Updated and expanded computer output for examples and exercises.
- Expanded and improved exercise sets and examples.

A special note of appreciation is extended to Susan Reiland, who provided an extensive prerevision review of this edition. We are also indebted to Will Sullivan for redoing the computer output and to Jim Stegeman for preparation of the solutions manual. Special thanks are also due to Barbara and Robert Beaver for preparing the study guide.

Many students and professors who used previous editions of this text have contacted us with suggestions, and these are greatly appreciated. The authors especially want to acknowledge the reviewers for this edition:

Maria Betkowski
Middlesex City College

James R. Case
Hiram College

Ester Guerin
Seton Hall University

Harold Nemer
Riverside City College

Franklin D. Shobe
St. Andrews Presbyterian College

Joseph Walker
Georgia State University

Karl Zilm
Lewis and Clark Community College

Howard A. Bird
St. Cloud State University

James Finch
University of San Francisco

Donald Hotchkiss
Iowa State University

Franklin Sheehan
San Francisco State University

Karen H. Smith
West Georgia College

Mary Woods
University of Louisville

The success of this book through the years is in many ways a team effort. For this edition, we gratefully thank our secretarial assistant, Phyllis Switzer, who transformed many drafts into polished word-processing documents. Thanks are also due A. Hald, E. S. Pearson, the Biometrika Trustees, and the Chemical Rubber Company for permission to reproduce tables, and to the SAS Institute, and Minitab, Inc., for allowing us to include output from their software systems.

We also want to acknowledge the cooperative efforts of our editor, Michael Payne, his assistant editor, Marcia Cole, his editorial assistant, Susan Hankinson, and the production staff of PWS-KENT. Finally, we mention the ongoing support and encouragement of our families.

Lyman Ott
William Mendenhall

CONTENTS



PART 1 STATISTICS: MAKING SENSE OF DATA 1

CHAPTER 1	WHAT IS STATISTICS?	3
1.1	Introduction	3
1.2	Why Study Statistics?	6
1.3	Some Current Applications of Statistics	6
1.4	What Do Statisticians Do?	10
1.5	A Note to the Student	11
	Summary	12
	Key Terms	12
	Exercises	12

PART 2 STEP ONE: METHODS FOR 15 **GATHERING DATA**

CHAPTER 2	USING CALCULATORS AND COMPUTERS	17
	IN STATISTICS	
2.1	Introduction	17
2.2	Calculators	17
2.3	Computers and Statistical Software	18
2.4	A Brief Introduction to Minitab and SAS	19
	Summary	22
	Key Terms	22

CHAPTER 3	USING SURVEYS AND SCIENTIFIC STUDIES	23
	TO GATHER DATA	
3.1	Introduction	23
3.2	Surveys	24
3.3	Scientific Studies	29
3.4	Observational Studies	37
	Summary	38
	Key Terms	38

PART 3 STEP TWO: METHODS FOR SUMMARIZING DATA **39**

CHAPTER 4 GRAPHICAL METHODS OF DATA DESCRIPTION **41**

- 4.1** Introduction 41
- 4.2** Organizing the Data 42
- 4.3** Pie Charts 42
- 4.4** Bar Charts 44
- 4.5** Frequency Histograms 50
- 4.6** Frequency Polygons 60
- 4.7** Comments Concerning Histograms 60
- 4.8** Stem-and-Leaf Plots 63
- 4.9** Using Computers 68
- Summary 76
- Key Terms 76
- Supplementary Exercises 76

CHAPTER 5 NUMERICAL METHODS OF DATA DESCRIPTION **90**

- 5.1** Introduction 90
- 5.2** Measures of Central Tendency 92
- 5.3** Measures of Variability 103
- 5.4** Shortcut Method for Calculating the Variance
and Standard Deviation 116
- 5.5** How to Guess the Standard Deviation of Sample Data 121
- 5.6** The Box Plot 124
- 5.7** Summarizing Data from More Than One Variable 127
- 5.8** Using Computers 132
- Summary 136
- Key Terms 137
- Key Formulas 138
- Supplementary Exercises 138

PART 4 STEP THREE: BACKGROUND FOR ANALYZING DATA **153**

CHAPTER 6 PROBABILITY AND PROBABILITY DISTRIBUTIONS **155**

- 6.1** Introduction 155
- 6.2** Interpretations of Probability 156
- 6.3** Finding the Probability of an Event 159

6.4	Event Relations	161
6.5	Conditional Probability and Independence	163
6.6	Random Variables	167
6.7	Probability Distributions for Discrete Random Variables	169
6.8	A Useful Discrete Random Variable: The Binomial	170
6.9	Probability Distributions for Continuous Random Variables	181
6.10	A Useful Continuous Random Variable: The Normal	184
	Summary	194
	Key Terms	195
	Key Formulas	196
	Supplementary Exercises	196
	Exercises from the Data Base	208

CHAPTER 7 SAMPLING DISTRIBUTIONS 209

7.1	Introduction	209
7.2	Random Sampling	209
7.3	The Sampling Distribution for \bar{x} and the Central Limit Theorem	214
7.4	Normal Approximation to the Binomial	226
7.5	Sampling Distribution for $\hat{\pi}$	230
7.6	Using Computers	234
	Summary	236
	Key Terms	236
	Key Formulas	237
	Supplementary Exercises	237

PART 5 STEP THREE: METHODS FOR ANALYZING DATA 245

CHAPTER 8 INFERENCES ABOUT μ 247

8.1	Introduction	247
8.2	Estimation of μ	250
8.3	Elements of a Statistical Test	256
8.4	A Statistical Test About μ	257
8.5	The Level of Significance of a Statistical Test	267
8.6	Inferences About μ , with σ Unknown	270
8.7	Assumptions Underlying the Analysis Methods of This Chapter	282
8.8	Using Computers	284
	Summary	285
	Key Terms	285

Key Formulas	286
Supplementary Exercises	286
Exercises from the Data Base	299

CHAPTER 9 INFERENCES ABOUT $\mu_1 - \mu_2$ 300

9.1	Introduction	300
9.2	The Sampling Distribution for the Difference Between Two Sample Means	301
9.3	Estimation of $\mu_1 - \mu_2$	304
9.4	A Statistical Test About $\mu_1 - \mu_2$	309
9.5	Assumptions	316
9.6	Using Computers	318
	Summary	321
	Key Terms	322
	Key Formulas	322
	Supplementary Exercises	323
	Exercises from the Data Base	339

CHAPTER 10 INFERENCES ABOUT π AND $\pi_1 - \pi_2$ 340

10.1	Introduction	340
10.2	Estimation of π	340
10.3	A Statistical Test About π	347
10.4	Comparing π_1 and π_2	350
	Summary	357
	Key Terms	358
	Key Formulas	358
	Supplementary Exercises	359
	Exercises from the Data Base	373

CHAPTER 11 INFERENCES ABOUT σ^2 AND σ_1^2 / σ_2^2 374

11.1	Introduction	374
11.2	Estimation and Tests for σ^2	374
11.3	Estimation and Tests for Comparing σ_1^2 and σ_2^2	380
	Summary	386
	Key Terms	386
	Key Formulas	387
	Supplementary Exercises	388
	Exercises from the Data Base	395

CHAPTER 12 REGRESSION AND CORRELATION 396

- 12.1** Introduction 396
- 12.2** Scatter Plots and the Freehand Regression Line 398
- 12.3** Method of Least Squares 406
- 12.4** Correlation 417
- 12.5** Multiple Regression (Optional) 424
- 12.6** Using Computers 427
- Summary 433
- Key Terms 433
- Key Formulas 434
- Supplementary Exercises 434
- Exercises from the Data Base 455

CHAPTER 13 INFERENCES RELATED TO LINEAR REGRESSION AND CORRELATION 456

- 13.1** Introduction 456
- 13.2** Inferences About β_0 and β_1 457
- 13.3** Inferences About $E(y)$ 471
- 13.4** Predicting y for a Given Value of x 477
- 13.5** A Statistical Test About ρ 482
- 13.6** Using Computers 482
- Summary 486
- Key Terms 486
- Key Formulas 486
- Supplementary Exercises 488
- Exercises from the Data Base 501

CHAPTER 14 THE DESIGN OF AN EXPERIMENT: GETTING MORE INFORMATION FOR YOUR MONEY 502

- 14.1** Introduction 502
- 14.2** Choosing the Sample Size to Estimate μ , $\mu_1 - \mu_2$, or π 504
- 14.3** The Paired-Difference Experiment: An Example of a Designed Experiment 511
- 14.4** Using Computers 519
- Summary 521
- Key Terms 522
- Key Formulas 522
- Supplementary Exercises 523

CHAPTER 15 ANALYSIS OF VARIANCE 534

- 15.1 Introduction 534
- 15.2 The Logic Behind an Analysis of Variance 535
- 15.3 A Statistical Test About More Than Two Population Means:
An Example of an Analysis of Variance 537
- 15.4 Checking on the Equal Variance Assumption 544
- 15.5 Using Computers 546
 - Summary 548
 - Key Terms 549
 - Key Formulas 549
 - Supplementary Exercises 549
 - Exercises from the Data Base 565

CHAPTER 16 ANALYZING COUNT DATA 566

- 16.1 Introduction 566
- 16.2 The Chi-Square Goodness-of-Fit Test 567
- 16.3 The Chi-Square Test of Independence 572
- 16.4 Using Computers 583
 - Summary 587
 - Key Terms 587
 - Key Formulas 588
 - Supplementary Exercises 588
 - Exercises from the Data Base 602

CHAPTER 17 NONPARAMETRIC STATISTICS 604

- 17.1 Introduction 604
- 17.2 A Simple Comparative Test: The Sign Test 605
- 17.3 Wilcoxon's Signed-Rank Test 613
- 17.4 Wilcoxon's Rank Sum Test 617
- 17.5 Spearman's Rank Correlation Coefficient 623
- 17.6 Using Computers 628
 - Summary 633
 - Key Terms 634
 - Key Formulas 634
 - Supplementary Exercises 635
 - Exercise from the Data Base 640

PART 6 STEP FOUR: REPORTING RESULTS 641

CHAPTER 18	COMMUNICATING THE RESULTS OF STATISTICAL ANALYSES	643
18.1	Introduction	643
18.2	Good Communication Is Not Easy	643
18.3	Communication Hurdles: Graphical Distortions	645
18.4	Communication Hurdles: Biased Samples	647
18.5	Communication Hurdles: What's the Sample Size?	648
18.6	The Statistical Report	649
	Summary	651

APPENDIX 1	DATA BASE: CLINICAL TRIAL DATA	653
-------------------	---------------------------------------	------------

APPENDIX 2	USEFUL STATISTICAL TESTS AND CONFIDENCE INTERVALS	664
-------------------	--	------------

APPENDIX 3	STATISTICAL TABLES	672
Table 1	Binomial Probabilities, $P(x)$ for $n \leq 20$	
Table 2	Normal Curve Areas	
Table 3	Critical Values of t	
Table 4	Critical Values of χ^2	
Table 5	Upper-Tail Values of F , $\alpha = .10$	
Table 6	Upper-Tail Values of F , $\alpha = .05$	
Table 7	Upper-Tail Values of F , $\alpha = .025$	
Table 8	Upper-Tail Values of F , $\alpha = .01$	
Table 9	Upper-Tail Values of F , $\alpha = .005$	
Table 10	Wilcoxon Signed-Rank Test ($n = 50$)	
Table 11	Random Numbers	
Table 12	$F_{\max} = s_{\max}^2 / s_{\min}^2$	

GLOSSARY OF COMMON STATISTICAL TERMS	695
---	------------

REFERENCES	702
-------------------	------------

ANSWERS TO SELECTED EXERCISES	704
--------------------------------------	------------

INDEX	710
--------------	------------