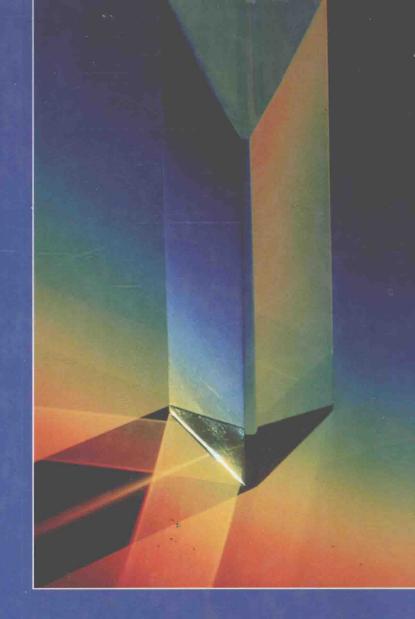
INTRODUCTION TO

CONCEPTS AND APPLICATIONS



THIRD EDITION

Anderson
Sweeney
Williams

INTRODUCTION TO STATISTICS

CONCEPTS AND APPLICATIONS

THIRD EDITION

DAVID R. ANDERSON University of Cincinnati

DENNIS J. SWEENEY
University of Cincinnati

THOMAS A. WILLIAMS
Rochester Institute of Technology

Copyediting Design

Luana Richards John Rokusek

Artwork

Rolin Graphics

Composition

Carlisle Communications

Proofreading

Jerrold Moore

Cover Design John Rokusek Cover Image

The Stock Market/Peter Saloutos © 1991

Indexing

Sandi Schroeder

Prepress, Printing, and Binding by West Publishing Company

Minitab is a registered trademark of Minitab, Inc., 3081 Enterprise Drive, State College, PA, 16801 (telephone 814/238-3280; telex 881612; fax 814/238-4383).

West's Commitment to the Environment

In 1906, West Publishing Company began recycling materials left over from the production of books. This began a tradition of efficient and responsible use of resources. Today, up to 95 percent of our legal books and 70 percent of our college texts are printed on recycled, acid-free stock. West also recycles nearly 22 million pounds of scrap paper annually—the equivalent of 181,717 trees. Since the 1960s, West has devised ways to capture and recycle waste inks, solvents, oils, and vapors created in the printing process. We also recycle plastics of all kinds, wood, glass, corrugated cardboard, and batteries, and have eliminated the use of styrofoam book packaging. We at West are proud of the longevity and the scope of our commitment to our environment.

COPYRIGHT © 1986, 1991

By WEST PUBLISHING COMPANY

COPYRIGHT © 1994 By WEST PUBLISHING COMPANY

> 610 Opperman Drive P.O. Box 64526

St. Paul, MN 55164-0526

All rights reserved Printed in the United States of America 00 99 98 97 96 95 94

87654321

Library of Congress Cataloging-in-Publication Data

Anderson, David Ray, 1941-

Introduction to statistics: concepts and applications / David R. Anderson, Dennis J. Sweeney, Thomas A. Williams.—3rd ed.

Includes bibliographical references and index.

ISBN 0-314-02813-7

1. Statistics. F. I. Sweeney, Dennis J. II. Williams, Thomas

Arthur, 1944-. III. Title.

QA276.12.A44 1993

519.5-dc20

93-30031





INTRODUCTION TO STATISTICS CONCEPTS AND APPLICATIONS

THIRD EDITION

PREFACE

The purpose of this book is to provide a comprehensive treatment of introductory statistics for students from a wide variety of academic backgrounds. The text is applications oriented and has been written with the needs of the nonmathematician in mind. The mathematical prerequisite is a course in college algebra.

APPLICATIONS AND METHODOLOGY INTEGRATED

Applications of the statistical methodology are an integral part of the organization and presentation of the material. Each chapter begins by motivating the student with a general interest writeup involving statistics in practice that demonstrates a use of the statistical procedures that will be introduced in the chapter. Statistical techniques are then introduced using examples where the techniques have been successfully applied. The discussion and development of each technique is centered around an application setting, with the statistical results providing information helpful in solving the underlying problem.

In addition, we have taken care to provide a sound methodological development. Throughout the text we have utilized notation that is generally accepted for the topic being covered. Students will thus find that the text provides good preparation for the study of more advanced statistical material. A bibliography that should prove useful as a guide to further study has been included as an appendix.

CHANGES IN THE THIRD EDITION

In making modifications for this new edition, we have maintained the presentation style of the previous edition. However, this edition includes a more complete and detailed topical coverage than the previous editions, including the addition of counting rules for combinations and permutations and expanded binomial probability tables. We have also extended our treatment of experimental design, revised the three-chapter sequence on regression analysis, and included a new chapter on sample survey methods. The more significant changes in the third edition are summarized below.

MORE EXAMPLES AND PROBLEMS BASED ON REAL DATA

We have continued our emphasis on helping students understand the wide range of statistical applications by expanding the number of examples and problems based on real statistical studies. Sources such as *The New York Times, USA Today, Psychology Today, New England Journal of Medicine,* and *The Wall Street Journal* provide referenced

applications and problems that demonstrate wide-ranging uses of statistics. The use of real data means that students not only learn about the statistical methodology but also learn about the content of applications encountered in practice. Approximately 300 new problems have been added in this edition.

METHODS EXERCISES AND APPLICATIONS EXERCISES

The end-of-section exercises have been split into two parts. With the Methods exercises, students are required to make sure they can handle the formulas and make the necessary computations. In the Applications exercises, students are given problems that require using the chapter material in real-world situations. With this approach, students first focus on the computational "nuts and bolts," then move on to the subtleties of statistical application and interpretation.

SELF-TEST EXERCISES

Certain exercises are identified as self-test exercises. Completely worked-out solutions for these exercises are provided in an appendix at the end of the text. Students can attempt the self-test exercises and immediately check the solution to evaluate their understanding of the concepts presented in the chapter.

COMPUTER CASES

Many chapters have computer cases, which contain problem scenarios accompanied by modest-sized data sets. Computer solution by Minitab, The Data Analyst, or another statistical software package is required. Each case outlines a report that the student prepares to summarize statistical results as well as present interpretations and recommendations. The data sets for all computer cases are available on data disks formatted for Minitab, The Data Analyst, or MYSTAT.

A NEW CHAPTER ON SAMPLING AND SURVEY METHODOLOGY

Chapter 18 treats sampling and survey methodology. A discussion of issues involved in designing and conducting survey research is presented. The probability sampling methods of simple random sampling, stratified simple random sampling, cluster sampling, and systematic sampling are covered. The nature and control of sampling and nonsampling error are treated.

ANALYSIS OF VARIANCE AND EXPERIMENTAL DESIGN

The presentation of analysis of variance and experimental design (Chapter 12) has been revised with a new *Statistics in Practice* used to introduce students to the concepts. The chapter begins with analysis of variance and its use in testing for the equality of means with three or more populations. Procedures for making multiple comparisons have been expanded and now appear immediately following the material on testing for the equality of means. Experimental design concepts are presented after the introduction to analysis of variance. Completely randomized, randomized block, and factorial experimental designs complete the coverage. Formulas for the analysis of variance computations appear in the chapter appendix.

REGRESSION ANALYSIS

The three-chapter sequence on regression analysis (Chapters 13, 14, and 15) has been enhanced with an expanded discussion of the use and interpretation of dummy variables and interaction terms. The best-subset selection technique and the use of normal probability plots are now included, and the discussion of residual analysis has been expanded.

FEATURES

We have continued many of the features that appeared in earlier editions. Some of the more important ones include the following.

LEARNING OBJECTIVES

Each chapter begins with a statement of learning objectives which falls under the heading of "What You Will Learn in This Chapter." This list contains the concepts that the student will be expected to master and should help guide the student's study of the material.

STATISTICS IN PRACTICE

Each chapter opens with a general interest news article that demonstrates a use of the statistical procedures about to be introduced. These *Statistics in Practice* applications are based on actual articles appearing in journals, magazines, and newspapers such as *The New England Journal of Medicine, Time, USA Today, The New York Times,* and so on. The *Statistics in Practice* is a condensed version of the original article and specifically focuses on the use of statistics as reported in the publication. Topical selections include the cost of a college education, marriage statistics, pay differentials for men and women, and professional sports; these selections should capture the student's interest and show them that the statistical procedures that he or she is about to learn have some interesting applications.

CHAPTER PEDAGOGY

Each chapter introduces statistical methodology in the context of examples that demonstrate the use of the methodology in a wide variety of general interest applications. Problems are provided after each section to enable the student to check his or her progress. Answers to the even-numbered exercises and self-test exercises are provided at the back of the book. Each chapter concludes with a review of key concepts and topics introduced therein. A glossary of statistical terms found in the chapter follows, and a key formula section itemizes the important equations that the student should know how to apply. A review quiz is then included to reinforce the key concepts presented. Supplementary exercises, based on the material throughout the chapter, then provide additional opportunities to practice applying the methodology presented. Where appropriate, computer printouts are included; these print-outs demonstrate how computer packages can be used in statistical computation and summaries.

REVIEW QUIZZES

The review quizzes consist of true-false and multiple choice questions. Each review quiz provides the student with an opportunity to evaluate his or her progress after the chapter material has been covered. Answers for the review quiz questions are included at the back of the book.

NOTES AND COMMENTS

At the end of many sections, we have provided *Notes and Comments* designed to give the student additional insights about the statistical methodology presented in the section. They include warnings and/or limitations about the methodology, recommendations for application, brief descriptions of additional technical considerations, and so on. It is hoped that this feature will expand the students' understanding of statistics and their ability to use the material.

COMPUTER SOFTWARE

The text contains numerous examples and discussions of the important role statistical software packages play in the computation and presentation of statistical results. More examples and problems based on the Minitab statistical package have been added. Use and interpretation of the information provided by Minitab printouts are emphasized. Available to adopters is The Data Analyst version 2.0, an IBM-compatible microcomputer software package developed by the authors. The Data Analyst or MYSTAT may be ordered shrink-wrapped with the text. The Data Analyst, MYSTAT, Minitab, or other statistical packages can be used to solve problems and computer cases appearing in the text.

DATA DISK

Data sets for text examples, problems, and computer cases are available on special Data Disks, that can be ordered with the text. The Data Disks contain the data sets in a format acceptable to Minitab, MYSTAT, or The Data Analyst software packages.

ANCILLARIES

Accompanying the text is a complete package of support materials. A solutions manual, prepared by the authors, contains complete solutions to all exercises in the text. A test bank contains multiple choice questions followed by problems, and WESTTEST Computerized Testing also accompanies the package. Transparency masters display worked-out demonstration problems, and a student workbook (prepared by Meredith Many and Charlotte Lewis, University of New Orleans) provides chapter summaries and glossaries, formula references, and additional practice problems with solutions. The Data Analyst microcomputer software package, developed by the authors, is available free to adopters, and students may purchase the software shrinkwrapped with the text for a small additional charge. In addition, the MYSTAT statistical software package may also be purchased shrinkwrapped with the text for a small additional charge.

ACKNOWLEDGMENTS

We owe a debt to many of our colleagues and friends for their helpful comments and suggestions during the development of this manuscript. Among these are Paul B. Berger, Ben P. Bockstege, John M. Burns, Curtis K. Church, Louis J. Cote, Robert H. Cranford, Henry Crouch, Carl Cuneo, Shirley Dowdy, Ruby A. Evans, David Gillman, Penelope Greene, Terry H. Hughes, Peter Kerwin, Basil P. Korin, Robert L. Lacher, Stanley M. Lukawecki, David R. Lund, Daniel E. McNamara, Robert Mee, Jeff Mock, Alex Papadopoulos, Charles D. Reinauer, Franklin D. Rich, Arnold L. Schroeder, A.K. Shah, Robert K. Smidt, W. Robert Stephenson, Bill Stines, Glen H. Swindle, Chin-Chyuan Tai, Barbara Treadwell, David L. Turner, and Vasant B. Waikar.

We are also indebted to our editor, Mary Schiller, for her counsel and support during the preparation of this text. We would like to express our appreciation to Tom Hilt, our production editor, for the outstanding job he performed in pulling all the pieces together to produce this book. Finally, a sincere thank you to everyone else at West Publishing Company that played a part in helping us complete this project.

David R. Anderson
Dennis J. Sweeney
Thomas A. Williams

CONTENTS

CHAPTER 1 DATA, MEASUREMENT, AND STATISITCS, 1

STATISTICS IN PRACTICE: Statistics Provide Facts About Life, 1

1.1 Data, 2

Elements, Variables, and Observations, 2

1.2 Scales of Measurement, 2

Nominal Scale, 4

Ordinal Scale, 4

Interval Scale, 5

Ratio Scale, 6

Qualitative and Quantitative Data, 7

1.3 Data Acquisition, 8

Existing Sources of Data, 8

Statistical Studies, 8

Possible Data-Acquisition Errors, 8

- 1.4 Descriptive Statistics, 9
- 1.5 Statistical Inference, 12

Summary, 14

Glossary, 14

Review Quiz, 15

Exercises, 15

CHAPTER 2 DESCRIPTIVE STATISTICS I: TABULAR AND GRAPHICAL METHODS, 20

STATISTICS IN PRACTICE: What is America's Favorite Beverage?, 21

2.1 Summarizing Qualitative Data, 22

Frequency Distribution, 22

Relative Frequency Distribution, 22

Bar Graphs and Pie Charts, 24

Exercises, 27

2.2 Summarizing Quantitative Data, 30

Frequency Distribution, 30

Relative Frequency Distribution, 31

Cumulative Frequency and Cumulative Relative Frequency Distributions, 32

Dot Plots, 33

Histograms, 34

Ogives, 35

Exercises, 35

2.3 The Role of the Computer, 39

2.4 Exploratory Data Analysis (Optional), 42

Exercises, 45

2.5 Summarizing Bivariate Data (Optional), 46

Exercises, 49

Summary, 49

Glossary, 50

Key Formulas, 51

Review Quiz, 51

Supplementary Exercises, 53

Computer Case: Consolidated Foods, Inc., 58

CHAPTER 3 DESCRIPTIVE STATISTICS II: MEASURES OF LOCATION AND DISPERSION, 60

STATISTICS IN PRACTICE: The American Dream—Is It Still Affordable?, 61

3.1 Measures of Location, 62

Mean, 62

Trimmed Mean, 63

Median, 64

Mode, 65

Percentiles, 66

Quartiles and Hinges, 67

Exercises, 69

3.2 Measures of Dispersion, 72

Range, 73

Interquartile Range, 74

Variance, 74

Standard Deviation, 77

Coefficient of Variation, 78

Exercises, 79

3.3 Some Uses of the Mean and the Standard Deviation, 80

z-Scores, 81

Chebyshev's Theorem, 81

The Empirical Rule, 82

Detecting Outliers, 83

Exercises, 84

3.4 Exploratory Data Analysis, 86

Five-Number Summary, 86

Box Plots, 87

Fences and the Detection of Outliers, 87

Exercises, 89

3.5 The Role of the Computer, 91

A Note on the Computation of Quartiles, 93

3.6 Measures of Location and Dispersion for Grouped Data (Optional), 93

Mean, 94

Variance, 95

Standard Deviation, 95

Population Summaries for Grouped Data, 96

Exercises, 96

Summary, 98

Glossary, 99

Key Formulas, 100

Review Quiz, 101

Supplementary Exercises, 102

Computer Case 1: Consolidated Foods, Inc., 106

Computer Case 2: National Health Care Association, 107

CHAPTER 4 INTRODUCTION TO PROBABILITY, 110

STATISTICS IN PRACTICE: Probability—What It Tells About the Typical American, 111

4.1 Experiments, Counting Rules, and Assigning Probabilities, 112

Counting Rules, Combinations, and Permutations, 113

Assigning Probabilities, 116

Exercises, 119

4.2 Events and Their Probabilities, 122

Exercises, 124

4.3 Rules for Computing Event Probabilities, 126

Complement of an Event, 126

Union and Intersection of Events, 127

Addition Rule, 127

Mutually Exclusive Events, 130

Exercises, 131

4.4 Conditional Probability, Independence, and the Multiplication Rule, 133

Conditional Probability, 133

Independent Events, 133

Multiplication Rule, 133

Exercises, 137

4.5 Bayes' Theorem, 141

The Tabular Approach, 143

Exercises, 144

Summary, 145

Glossary, 146

G1055a1 y, 140

Key Formulas, 147

Review Quiz, 147

Supplementary Exercises, 149

CHAPTER 5 RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS, 156

STATISTICS IN PRACTICE: Pete Rose Versus Joe DiMaggio—Baseball Greats, 157

5.1 Random Variables, 158

Exercises, 159

5.2 Discrete Probability Distributions, 160

Exercises, 162

5.3 Expected Value and Variance, 165

Expected Value, 165

Variance, 166

Exercises, 167

5.4 The Binomial Probability Distribution, 170

Using Tables of Binomial Probabilities, 177

The Expected Value and Variance for the Binomial Probability Distribution, 180

Exercises, 181

5.5 The Poisson Probability Distribution (Optional), 184

Exercises, 187

5.6 The Hypergeometric Probability Distribution (Optional), 188

Exercises, 190

Summary, 191

Glossary, 192

Key Formulas, 192

Review Quiz, 193

Supplementary Exercises, 194

CHAPTER 6 CONTINUOUS PROBABILITY DISTRIBUTIONS, 198

STATISTICS IN PRACTICE: IQ Scores—Are You Normal?, 199

6.1 The Uniform Probability Distribution, 200

Area as a Measure of Probability, 201

Exercises, 203

6.2 The Normal Probability Distribution, 204

The Standard Normal Probability Distribution, 206

Computing Probabilities for Any Normal Distribution, 212

Exercises, 215

6.3 Normal Approximation of Binomial Probabilities, 218

Exercises, 220

6.4 The Exponential Probability Distribution (Optional), 221

Computing Probabilities for the Exponential Distribution, 222

Exercises, 223

Summary, 224

Glossary, 225

Key Formulas, 226

Review Quiz, 226

Supplementary Exercises, 227

CHAPTER 7 SAMPLING AND SAMPLING DISTRIBUTIONS, 232

STATISTICS IN PRACTICE: Charitable Contributions—The People Who Give, 233

7.1 Introductory Sampling Concepts, 233

Populations and Samples, 233

Reasons for Sampling, 234

Exercises, 235

7.2 Simple Random Sampling, 236

Sampling from a Finite Population, 236

Sampling from an Infinite Population, 239

Exercises, 240

7.3 Sampling Distribution of \bar{x} , 243

Expected Value of \bar{x} , 247

Standard Deviation of \bar{x} , 248

Exercises, 250

7.4 The Central Limit Theorem, 252

Exercises, 255

7.5 Computing Probabilities Using the Sampling Distribution of \bar{x} , 256

Exercises, 258

Summary, 260

Glossary, 260

Key Formulas, 260

Review Quiz, 261

Supplementary Exercises, 262

Chapter 8 Inferences about a Population Mean, 266

STATISTICS IN PRACTICE: Radon Poses Health Hazard, 267

8.1 Interval Estimation of a Population Mean: Large-Sample Case, 268

Probability Statements About the Sampling Error, 269

Calculating an Interval Estimate, 272

Exercises, 276

8.2 Interval Estimation of a Population Mean: Small-Sample Case, 278

Computer-Generated Confidence Intervals, 282

Exercises, 284

8.3 Determining the Sample Size, 286

Exercises, 288

8.4 Hypothesis Tests About a Population Mean, 289

Testing Research Hypotheses, 290

Testing the Validity of a Claim, 290

Testing in Decision-Making Situations, 291

A Summary of Forms for Null and Alternative Hypotheses, 291

Exercises, 292

8.5 Type I and Type II Errors, 293

Exercises, 294

8.6 One-Tailed Hypothesis Tests About a Population Mean: Large-Sample Case, 295

Summary: One-Tailed Tests About a Population Mean, 300

The Use of p-Values, 301

The Steps of Hypothesis Testing, 302

Exercises, 303

8.7 Two-Tailed Hypothesis Tests About a Population Mean: Large-Sample Case, 305

Summary: Two-Tailed Tests About a Population Mean, 307

p-Values for Two-Tailed Tests, 308

Computer Software and Hypothesis Testing, 308

The Relationship Between Interval Estimation and Hypothesis Testing, 309

Exercises, 311

8.8 Hypothesis Tests About a Population Mean: Small-Sample Case, 313

p-Values and the t Distribution, 314

A Two-Tailed Test, 315

Exercises, 316

8.9 Calculating the Probability of Type II Errors (Optional), 318

Exercises, 322

8.10 Determining the Sample Size for a Hypothesis Test About a Population Mean (Optional), 323

Summary, 327

Glossary, 331

Key Formulas, 331

Review Quiz, 332

Supplementary Exercises, 334

Computer Case: Metropolitan Research, Inc., 337

Computer Case: Quality Associates, Inc., 337

CHAPTER 9 INFERENCES ABOUT A POPULATION PROPORTION, 340

STATISTICS IN PRACTICE: Cost Is the Major Reason for Not Going to College, 341

9.1 The Sampling Distribution of \bar{p} , 343

Expected Value of \bar{p} , 343

Standard Deviation of \bar{p} , 343

Form of the Sampling Distribution of \bar{p} , 344

Exercises, 347

9.2 Interval Estimation of a Population Proportion, 350

Determining the Size of the Sample, 351

Exercises, 354

9.3 Hypothesis Tests About a Population Proportion, 356

Exercises, 359

Summary, 361

Glossary, 361

Key Formulas, 361

Review Quiz, 362

Supplementary Exercises, 363

CHAPTER 10 INFERENCES ABOUT MEANS AND PROPORTIONS WITH TWO POPULATIONS, 366

STATISTICS IN PRACTICE: Nicotine Chewing Gum Helps Stop Smoking, 367

10.1 Estimation of the Difference Between the Means of Two Populations: Independent Samples, 368

Sampling Distribution of $\bar{x}_1 - \bar{x}_2$, 369

Large-Sample Case, 370

Interval Estimation: Small-Sample Case, 372

Exercises, 376

10.2 Hypothesis Tests About the Difference Between the Means of Two Populations: Independent Samples, 378

Large-Sample Case, 379

Small-Sample Case, 382

Computer Solution, 384

Exercises, 386

10.3 Inferences About the Difference Between the Means of Two Populations: Matched Samples, 388

Exercises, 391

10.4 Inferences About the Difference Between the Proportions of Two Populations, 395

Sampling Distribution of $\bar{p}_1 - \bar{p}_2$, 395

Interval Estimation of $\bar{p}_1 - \bar{p}_2$, 397

Hypothesis Tests About $\bar{p}_1 - \bar{p}_2$, 398

Exercises, 400

Summary, 403

Glossary, 403

Key Formulas, 404

Review Quiz, 405

Supplementary Exercises, 406

Computer Case: Par, Inc., 409

CHAPTER 11 INFERENCES ABOUT POPULATION VARIANCES, 410

STATISTICS IN PRACTICE: GAO Audits Pollution Control, 411

11.1 Inferences About a Population Variance, 412

The Sampling Distribution of $(n-1)s^2/\sigma^2$, 412

The Chi-Square Distribution, 413

Interval Estimation of σ^2 , 416

Hypothesis Tests About σ^2 , 418

Exercises, 419

11.2 Inferences About the Variances of Two Populations, 421

The Sampling Distribution of s_1^2/s_2^2 , 421

The F Distribution, 422

Hypothesis Tests About σ_1^2 and σ_2^2 , 425

Exercises, 427

Summary, 429

Key Formulas, 429

Review Quiz, 430

Supplementary Exercises, 431

Computer Case: Air Force Training Program, 432

CHAPTER 12 ANALYSIS OF VARIANCE AND EXPERIMENTAL DESIGN, 434

STATISTICS IN PRACTICE: High Self-Esteem: A Key to Academic Achievement, 435

12.1 An Introduction to Analysis of Variance, 436

Assumptions for Analysis of Variance, 437

A Conceptual Overview, 437

12.2 Analysis of Variance: Testing the Equality of k Population Means, 441

Between-Samples Estimate of Population Variance, 442

Within-Samples Estimate of Population Variance, 443

Comparing the Variance Estimates: The F Test, 443

The ANOVA Table, 445

Computer Results for Analysis of Variance, 445

Exercises, 447

12.3 Multiple Comparison Procedures, 451

Fisher's LSD, 451

Type I Error Rates, 453

Bonferroni Adjustment, 455

Tukey's Procedure, 456

Recommendations for Performing Multiple Comparisons, 458

Exercises, 459

12.4 An Introduction to Experimental Design, 460

Data Collection, 461

12.5 Completely Randomized Designs, 463

Between-Treatments Estimate of Population Variance, 463

Within-Treatments Estimate of Population Variance, 464

Comparing the Variance Estimates: The F Test, 465

The ANOVA Table for Completely Randomized Designs, 465

Pairwise Comparisons, 465

Exercises, 466

12.6 Randomized Block Designs, 470

The Analysis of Variance for a Randomized Block Design, 472

Computations and Conclusions, 473

Exercises, 475

12.7 Factorial Experiments, 476

The Analysis of Variance for a Factorial Experiment, 478

Computations and Conclusions, 479

Exercises, 482

Summary, 484

Glossary, 484

Key Formulas, 485

Review Quiz, 488

Supplementary Exercises, 489

Computer Case: Wentworth Medical Center, 494

Appendix 12.1: Computational Procedure for a Completely Randomized Design, 495

Appendix 12.2: Computational Procedure for a Randomized Block Design, 496

Appendix 12.3: Computational Procedure for a Two-Factor Factorial Design, 497