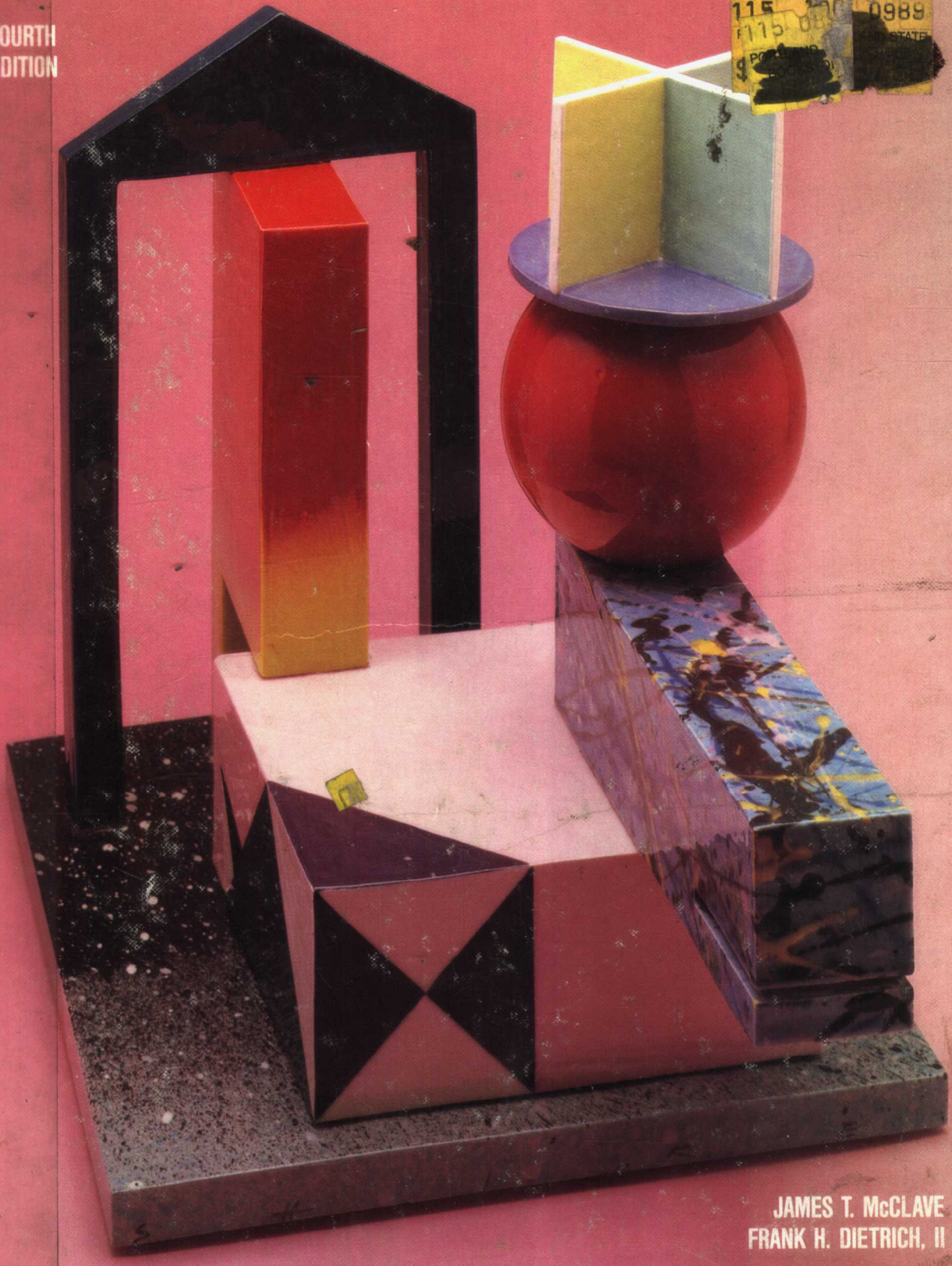
STATISTICS

FOURTH EDITION



JAMES T. McCLAVE
FRANK H. DIETRICH, II

FOURTH EDITION

STATISTICS

JAMES T. McCLAVE
University of Florida

FRANK H. DIETRICH, II
Northern Kentucky University

DELLEN PUBLISHING COMPANY
San Francisco

COLLIER MACMILLAN PUBLISHERS
London

divisions of Macmillan, Inc.

PREFACE

The fourth edition of *Statistics* finds the teaching of statistics in the midst of a number of significant changes. The availability of microcomputers with a wide range of statistical software has dramatically affected the amount of time we spend teaching computational formulas. The use of statistical quality control in American industry is perhaps the fastest-growing application of statistics in any area. Academic professional organizations are paying increased attention to the content of statistics courses, and the philosophy and manner in which they are taught. The American Statistical Association's Section on Statistical Education plays an increasingly important role in the development and enhancement of materials and techniques for teaching statistics.

In light of these and other developments, we have decided that the time has come to acknowledge that the text we wrote 10 years ago is not the same one we would write today. For example, 10 years ago the calculator had taken the "arithmetic" out of analysis of variance, but not the formulas. The computer gives us an additional option: We can use precious time teaching the interpretation of ANOVA results rather than cookbook formulas that produce those results. This fourth edition exercises that option (and many others like it).

Before detailing the specifics of the fourth edition, we stress that the underlying philosophy of the text is unchanged, despite the many new features and additional options. Our original intent was to write a text *for the students* that stressed *inference making*. If anything, the rapid evolution of applied statistics is making that easier to accomplish.

1. **Using the Computer.** A new feature has been added at the end of most chapters to encourage the use of computers in the analysis of real data. A demographic data base, consisting of 1,000 observations on 15 variables, has been described in Appendix C and is available on diskette from the publisher. Most chapters include a suggested computer application, "Using the Computer . . . ," that provides one or more computer exercises that utilize the data in Appendix C and enhance the new material covered in the chapter.
2. **Exploratory Data Analysis.** The computer permits the analyst to look at the data from every angle in order to extract all the information they have to offer. The use of microcomputer packages to generate stem and leaf displays, box plots, and the more traditional histograms is introduced in Chapter 2, and utilized throughout the text. We are careful to delineate the difference between descriptive and inferential statistics, and the student is constantly reminded of the necessity to consider the reliability of any statistical procedure used.
3. **Computers in Regression and Residual Analysis.** Ten years ago we set out to include the most complete treatment of regression analysis on the market, because we believed that modeling was one of the most important applications of statistics. Our view has not changed, and the computer has enabled aspects of modeling to be realized that were only theoretical constructs when we wrote

the original text. In keeping with our goal of a thorough, modern treatment of regression analysis, the fourth edition introduces computer printouts in the simple linear regression chapter, adds significantly to the number of examples and exercises that use the computer for regression analyses, and incorporates residual analyses into the model-building process. Although our first three editions made substantial use of the computer in the treatment of regression modeling, this edition adds significantly to that base. Microcomputer printouts are included, more exercises show printouts instead of requiring the student to perform the calculations, and more exercises with real data requiring the student to use the computer have been added. In addition, Chapter 12 contains a new section showing how to use the computer to perform residual analyses, and, more importantly, how to integrate them into the model-building process.

4. **Tests of Hypotheses.** Much current debate centers on how and what, if anything, to teach about testing hypotheses. Our view is that a modern text must present the subject, because much research continues to make use of statistical testing procedures. In this fourth edition, we have expanded the explanatory material about testing, including a new (optional) section on the computation of Type II error probabilities (β) and on the power of tests. We have also added exercises that require the student to confront the issue of which inferential procedure—testing or confidence intervals—provides more information about a parameter. Thus, the fourth edition offers an option as to the depth of treatment the instructor selects for testing hypotheses, while at the same time encouraging the student to contrast and to evaluate the statistical procedures employed.
5. **Normal Distribution and Sampling Distributions.** The normal distribution remains the cornerstone of applied statistics. Although robustness rightfully receives increased attention, the Central Limit Theorem ensures the continued importance of the normal distribution. We have added more explanatory examples to the use of the normal distribution in Chapter 5, and we have replaced the more difficult two-sample material in the sampling distribution chapter (Chapter 6) with an expanded treatment of the one-sample case, including more computer simulations of sampling distributions and an introduction to the sampling distribution of the sample proportion for a binomial random variable.
6. **Analysis of Variance.** The fourth edition adopts a computerized rather than a "cookbook" approach to ANOVA. All calculation formulas are relegated to Appendix B, and the emphasis is changed to the understanding and interpretation of designed experiments. The terminology of designed experiments is defined and utilized throughout the chapter, with constant reinforcement by example. Computer printouts are presented for each type of analysis covered, with space that was formerly occupied by tedious calculations now devoted to interpretation of the statistical results produced by the software. The Bonferroni multiple comparisons procedure is introduced early and is utilized during each analysis, where it would naturally occur, rather than in a separate section at the end of the chapter.

7. **Quality Control.** Numerous case studies, examples, and exercises that focus on statistical quality control applications have been added throughout the text. Since the application of statistical ideas to quality control has become so pervasive, we decided that the subject needed to be stressed throughout, as the relevant statistical concepts are covered, rather than concentrated in a few pages or chapter devoted to the subject.
8. **Determination of Sample Size.** To accompany our increased emphasis on designed experiments, we have included separate sections for determining the sample size for estimating a population mean and a binomial probability. Sample size formulas in terms of both the bound on the error of estimation and the total width of the confidence interval are provided.
9. **Normal Approximation to Binomial Probabilities.** The rationale behind and the conditions under which the normal distribution can be used for approximating binomial probabilities are significantly expanded in Chapter 5.
10. **Real Applications of Statistics.** From the start we have attempted to include many examples and exercises that draw on the extensive (and growing) literature containing real applications of statistics. The fourth edition adds to and updates the many examples, exercises, and case studies that represent applications of statistics to real problems. It should be noted, however, that constructed exercises and examples often better serve our pedagogic objectives, and we therefore continue to make use of them where appropriate.

In spite of the fact that we consider the fourth edition a major revision, the flexibility of past editions is maintained. Sections that are not prerequisite to succeeding sections and chapters are marked "(Optional)." For example, an instructor who wishes to devote significant time to exploratory data analysis might cover all topics in Chapter 2. In contrast, an instructor who wishes to move rapidly into inferential procedures might omit the optional section on box plots and devote only several lectures to this chapter.

We have maintained the features of this text that we believe make it unique among introductory statistics texts. These features, which assist the student in achieving an overview of statistics and an understanding of its relevance in the sciences, business, and everyday life, are as follows:

1. **Case Studies.** (See the list of case studies on page xix.) Many important concepts are emphasized by the inclusion of case studies, which consist of brief summaries of actual applications of statistical concepts and are often drawn directly from the research literature. These case studies allow the student to see applications of important statistical concepts immediately after their introduction. The case studies also help to answer by example the often-asked questions, "Why should I study statistics? Of what relevance is statistics to my program?" Finally, the case studies constantly remind the student that each concept is related to the dominant theme—statistical inference.
2. **The Use of Examples as a Teaching Device.** We have introduced and illustrated almost all new ideas by examples. Our belief is that most students will better understand definitions, generalizations, and abstractions *after* seeing an

application. In most sections, an introductory example is followed by a general discussion of the procedures and techniques, and then a second example is presented to solidify the understanding of the concepts.

3. **A Simple, Clear Style.** We have tried to achieve a simple and clear writing style. Subjects that are tangential to our objective have been avoided, even though some may be of academic interest to those well-versed in statistics. We have not taken an encyclopedic approach in the presentation of material.
4. **Many Exercises—Labeled by Type.** The text has a large number (almost 1,200) of exercises illustrating applications in almost all areas of research. However, we believe that many students have trouble learning the mechanics of statistical techniques when problems are all couched in terms of realistic applications—the concept becomes lost in the words. Thus, the exercises at the ends of all sections are divided into two parts:
   a. **Learning the Mechanics.** These exercises are intended to be straightforward applications of the new concepts. They are introduced in a few words and are unhampered by a barrage of background information designed to make them "practical," but which often detracts from instructional objectives. Thus, with a minimum of labor, the student can recheck his or her ability to comprehend a concept or a definition.
   b. **Applying the Concepts.** The mechanical exercises described above are followed by realistic exercises that allow the student to see applications of statistics across a broad spectrum. Once the mechanics are mastered, these exercises develop the student's skills at comprehending realistic problems that describe situations to which the techniques may be applied.
5. **On Your Own . . .** Each chapter ends with an exercise entitled "On Your Own . . . ." The intent of this exercise is to give the student some hands-on experience with an application of the statistical concepts introduced in the chapter. In most cases, the student is required to collect, analyze, and interpret data relating to some real application.
6. **A Choice in Level of Coverage of Probability.** One of the most troublesome aspects of an introductory statistics course is the study of probability. Probability is troublesome for instructors because they must decide on the level of presentation, and it is troublesome for students because they (often) find it difficult at any level. We believe that one cause for these problems is the mixture of probability and counting rules that occurs in most introductory texts. We have included the counting rules in a separate and optional section at the end of the chapter on probability. In addition, all exercises that require the use of counting rules are marked with an asterisk (*) to indicate this. Thus, the instructor can control the level of coverage of probability.

   A word should be added about the length of the probability chapter. Although more space is devoted to probability than in many introductory texts, there are three simple explanations for this: more examples, more exercises, and the optional counting rule section mentioned above. We have included the usual die and coin examples to introduce concepts, but we also work many practical examples so that the connection between probability and statistics is

clearly made. This same pattern is followed in the exercise sections. Thus, the 34 examples and 121 exercises account for the length of the chapter. We trust that these will make this troublesome subject easier to learn and to teach.

7. **Where We've Been . . . Where We're Going . . .** The first page of each chapter is a "unification" page. Our purpose is to allow the student to see how the chapter fits into the scheme of statistical inference. First, we briefly show how the material presented in previous chapters helps us to achieve our goal (Where We've Been). Then, we indicate what the next chapter (or chapters) contributes to the overall objective (Where We're Going). This feature allows us to point out that we are constructing the foundation block by block, with each chapter an important component in the structure of statistical inference. Furthermore, this feature provides a series of brief resumés of the material covered as well as glimpses of future topics.
8. **An Extensive Coverage of Multiple Regression Analysis and Model Building.** This topic represents one of the most useful statistical tools for the solution of applied problems. Although an entire text could be devoted to regression modeling, we feel that we have presented a coverage that is understandable, usable, and much more comprehensive than the presentations in other introductory statistics texts. We devote three chapters to discussing the major types of inferences that can be derived from a regression analysis, showing how these results appear in computer printouts and, most important, selecting multiple regression models to be used in an analysis. Thus, the instructor has the choice of a one-chapter coverage of simple regression, a two-chapter treatment of simple and multiple regression, or a complete three-chapter coverage of simple regression, multiple regression, and model building. This extensive coverage of such useful statistical tools will provide added evidence to the student of the relevance of statistics to the solution of applied problems.
9. **Footnotes.** Although the text is designed for students with a noncalculus background, footnotes explain the role of calculus in various derivations. Footnotes are also used to inform the student about some of the theory underlying certain results. The footnotes allow additional flexibility in the mathematical and theoretical level at which the material is presented.
10. **Supplementary Material.** A solutions manual, a study guide, a Minitab supplement, an integrated companion software system, a computer-generated test system, and a 1,000-observation demographic data base are available.
    a. **Solutions Manual** (by Nancy Shafer). The solutions manual presents the solutions to most odd-numbered exercises in the text. Many points are clarified and expanded to provide maximum insight into and benefit from each exercise.
    b. **Study Guide** (by Susan L. Reiland). For each chapter, the study guide includes (1) a brief summary that highlights the concepts and terms introduced in the textbook; (2) section-by-section examples with detailed solutions; and (3) exercises (with answers provided at the end of the study guide) that allow the student to check mastery of the material in each section.

c. Minitab Supplement (by David D. Krueger and Ruth K. Meyer). The Minitab computer supplement was developed to be used with Minitab Release 5.1, a general-purpose statistical computing system. The supplement, which was written especially for the student with no previous experience with computers, provides step-by-step descriptions of how to use Minitab effectively as an aid in data analysis. Each chapter begins with a list of new commands introduced in the chapter. Brief examples are then given to explain new commands, followed by examples from the text illustrating the new and previously learned commands. Where appropriate, simulation examples are included. Exercises, many of which are drawn from the text, conclude each chapter.

   A special feature of the supplement is a chapter describing a survey sampling project. The objectives of the project are to illustrate the evaluation of a questionnaire, provide a review of statistical techniques, and illustrate the use of Minitab for questionnaire evaluation.

d. DellenStat (by Michael Conlon). DellenStat is an integrated statistics package consisting of a workbook and an IBM PC floppy diskette with software and example sets of data. The system contains a file creation and management facility, a statistics facility, and a presentation facility. The software is menu-driven and has an extensive help facility. It is completely compatible with the text.

   The DellenStat workbook describes the operation of the software and uses examples from the text. After an introductory chapter for new computer users, the remaining chapters follow the outline of the text. Additional chapters show how to create new sets of data. Technical appendices cover material for advanced users and programmers.

   DellenStat runs on any IBM PC or close compatible with at least 256K of memory and at least one floppy disk drive.

e. DellenTest. This unique computer-generated random test system is available to instructors without cost. Utilizing an IBM PC computer and a number of commonly used dot-matrix printers, the system will generate an almost unlimited number of quizzes, chapter tests, final examinations, and drill exercises. At the same time, the system produces an answer key and student worksheet with an answer column that exactly matches the column on the answer key.

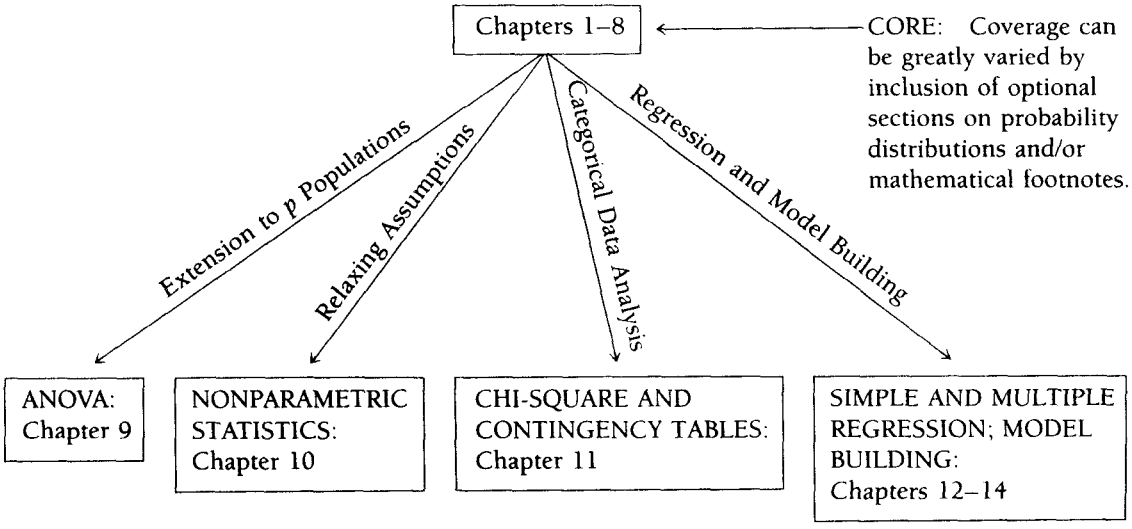
f. Data Base. A demographic data set was assembled based on a systematic random sample of 1,000 U.S. zip codes. Demographic data for each zip code area selected were supplied by CACI, an international demographic and market information firm. Fifteen demographic measurements (including population, number of households, median age, median household income, variables related to the cost of housing, educational levels, the work force, and purchasing potential indexes based on the Bureau of the Census Consumer Expenditure Surveys) are presented for each zip code area.

   Some of the data are referenced in the "Using the Computer" sections. The objectives are to enable the student to analyze real data in a relatively

large sample using the computer, and to gain experience using the statistical techniques and concepts on real data.

We have organized the text in what we believe is a logical sequence. For example, rather than place analysis of variance and nonparametric statistics at the end of the book, we have placed them immediately following the chapters on making inferences about one and two populations. The analysis of variance extends these methods to p populations, and nonparametric statistics gives procedures for making the inferences when the assumptions necessary for parametric methods are in doubt. The three regression and model building chapters, which we consider an essential and unique component, are the final three chapters, because regression represents a greater change in direction than the other inference-making techniques.

Because many introductory courses in statistics are of one-term duration, we show several possible sequences of coverage in the diagram below:



ACKNOWLEDGMENTS

Thanks are due to many individuals who helped in the preparation of this text. Among them are the reviewers, whose names are listed below. Special thanks to John Dirkse, California State University at Bakersfield. Susan Reiland has our appreciation and admiration for editing and producing this book. Her work defies explanation; you have to see to believe the care and professionalism with which she works. Finally, we thank the thousands of students who have helped us to form our ideas about teaching statistics. Their most common complaint seems to be that texts are written for the instructor rather than the student. We hope that this book is an exception.

William H. Beyer
University of Akron

Michael S. Broida
Miami University, Oxford

Edward Carlstein
University of North Carolina, Chapel Hill

Ron L. Coccari
Cleveland State University

Rudy Gideon
University of Montana

Joseph Glaz
University of Connecticut

Jean L. Holton
Virginia Commonwealth University

Iris B. Ibrahim
Clemson University

John H. Kellermeier
Northern Illinois University

Timothy J. Killeen
University of Connecticut

William G. Koellner
Montclair State University

David D. Krueger
St. Cloud State University

James R. Lackritz
San Diego State University

Diane Lambert
Carnegie–Mellon University

James Lang
Valencia Junior College

Pi-Erh Lin
Florida State University

Paul I. Nelson
Kansas State University

William B. Owen
Central Washington University

Won J. Park
Wright State University

Charles W. Sinclair
Portland State University

Craig W. Slinkman
University of Texas at Arlington

Vasanth B. Solomon
Drake University

Augustin Vukov
University of Toronto

CONTENTS

Preface ix

CHAPTER 1 WHAT IS STATISTICS? 1

1.1 Statistics: What Is It? 2
1.2 The Elements of Statistics 5
1.3 Statistics: Witchcraft or Science? 7
1.4 Why Study Statistics? 9

CHAPTER 2 METHODS FOR DESCRIBING SETS OF DATA 13

2.1 Types of Data 14
2.2 Graphic Methods for Describing Quantitative Data: Stem and Leaf Displays 15
2.3 Graphic Methods for Describing Quantitative Data: Histograms 22
2.4 Numerical Measures of Central Tendency 33
2.5 Numerical Measures of Variability 42
2.6 Interpreting the Standard Deviation 49
2.7 Measures of Relative Standing 58
2.8 Box Plots: Graphic Descriptions Based on Quartiles (Optional) 65
2.9 Distorting the Truth with Descriptive Techniques 75

CHAPTER 3 PROBABILITY 91

3.1 Events, Sample Spaces, and Probability 92
3.2 Compound Events 107
3.3 Complementary Events 111
3.4 Conditional Probability 117
3.5 Probabilities of Unions and Intersections 125
3.6 Probability and Statistics: An Example 138
3.7 Random Sampling 139
3.8 Some Counting Rules (Optional) 148

CHAPTER 4 DISCRETE RANDOM VARIABLES 169

4.1 Two Types of Random Variables 171
4.2 Probability Distributions for Discrete Random Variables 174
4.3 Expected Values of Discrete Random Variables 179
4.4 The Binomial Random Variable 187
4.5 The Poisson Random Variable (Optional) 202
4.6 The Hypergeometric Random Variable (Optional) 209
4.7 The Geometric Random Variable (Optional) 214

CHAPTER 5

CONTINUOUS RANDOM VARIABLES 227

5.1 Continuous Probability Distributions 228
5.2 The Uniform Distribution 229
5.3 The Normal Distribution 235
5.4 Approximating a Binomial Distribution with a Normal Distribution 253
5.5 The Exponential Distribution (Optional) 261

CHAPTER 6

SAMPLING DISTRIBUTIONS 277

6.1 What Is a Sampling Distribution? 279
6.2 Properties of Sampling Distributions: Unbiasedness and Minimum Variance 287
6.3 The Central Limit Theorem 292
6.4 The Relation Between Sample Size and a Sampling Distribution 301

CHAPTER 7

ESTIMATION AND TESTS OF HYPOTHESES: SINGLE SAMPLE 317

7.1 Large-Sample Estimation of a Population Mean 318
7.2 Determining the Sample Size Necessary for Making Inferences About a Population Mean 327
7.3 A Large-Sample Test of Hypothesis About a Population Mean 331
7.4 Calculating Type II Error Probabilities: More About β (Optional) 348
7.5 Observed Significance Levels; p-Values 358
7.6 Small-Sample Inferences About a Population Mean 363
7.7 Large-Sample Inferences About a Binomial Population Proportion 375
7.8 Determining the Sample Size Necessary to Make Inferences About a Binomial Probability 383
7.9 Inferences About a Population Variance (Optional) 387

CHAPTER 8

ESTIMATION AND TESTS OF HYPOTHESES: TWO SAMPLES 403

8.1 Large-Sample Inferences About the Difference Between Two Population Means: Independent Sampling 404
8.2 Small-Sample Inferences About the Difference Between Two Population Means: Independent Sampling 416
8.3 Inferences About the Difference Between Two Population Means: Paired Difference Experiments 427
8.4 Inferences About the Difference Between Population Proportions: Independent Binomial Experiments 441
8.5 Determining the Sample Size 450
8.6 Comparing Two Population Variances: Independent Random Samples 455

CHAPTER 9

ANALYSIS OF VARIANCE: COMPARING MORE THAN TWO MEANS 479

9.1 Elements of a Designed Experiment 480
9.2 The Completely Randomized Design 486
9.3 The Randomized Block Design 509
9.4 Factorial Experiments 527

CHAPTER 10

NONPARAMETRIC STATISTICS 567

10.1 Comparing Two Populations: Wilcoxon Rank Sum Test for Independent Samples 569
10.2 Comparing Two Populations: Wilcoxon Signed Rank Test for the Paired Difference Experiment 579
10.3 Kruskal–Wallis H-Test for a Completely Randomized Design 589
10.4 The Friedman F_r-Test for a Randomized Block Design 599
10.5 Spearman's Rank Correlation Coefficient 608

CHAPTER 11

THE CHI-SQUARE TEST AND THE ANALYSIS OF CONTINGENCY TABLES 637

11.1 One-Dimensional Count Data: The Multinomial Distribution 639
11.2 Contingency Tables 647
11.3 Contingency Tables with Fixed Marginal Totals 662
11.4 Caution 669

CHAPTER 12

SIMPLE LINEAR REGRESSION 681

12.1 Probabilistic Models 683
12.2 Fitting the Model: Least Squares Approach 686
12.3 Model Assumptions 697
12.4 An Estimator of σ^2 698
12.5 Assessing the Usefulness of the Model: Making Inferences About the Slope β_1 702
12.6 Correlation: A Measure of the Usefulness of the Model 710
12.7 The Coefficient of Determination 715
12.8 Using the Model for Estimation and Prediction 722
12.9 Simple Linear Regression: An Example 733
12.10 Using the Computer for Simple Linear Regression 737

CHAPTER 13

MULTIPLE REGRESSION 757

13.1 Model Assumptions 759
13.2 Fitting the Model: Least Squares Approach 760
13.3 Estimation of σ^2, the Variance of ε 763
13.4 Estimating and Testing Hypotheses About the β-Parameters 764
13.5 Checking the Usefulness of a Model: R^2 and the Analysis of Variance F-Test 777
13.6 Using the Model for Estimation and Prediction 793
13.7 Multiple Regression: An Example 796
13.8 Statistical Computer Programs 799
13.9 Residual Analysis: Checking the Regression Assumptions 803
13.10 Some Pitfalls: Estimability, Multicollinearity, and Extrapolation 822

CHAPTER 14 MODEL BUILDING 847

14.1 The Two Types of Independent Variables: Quantitative and Qualitative 849
14.2 Models with a Single Quantitative Independent Variable 851
14.3 Models with Two or More Quantitative Independent Variables 861
14.4 Model Building: Testing Portions of a Model 872
14.5 Models with One Qualitative Independent Variable 887
14.6 Comparing the Slopes of Two or More Lines 898
14.7 Comparing Two or More Response Curves 913
14.8 Model Building: Stepwise Regression 921

APPENDIX A TABLES 943

Table I Random Numbers 945
Table II Binomial Probabilities 948
Table III Poisson Probabilities 953
Table IV Normal Curve Areas 958
Table V Exponentials 959
Table VI Critical Values of t 961
Table VII Critical Values of χ^2 962
Table VIII Percentage Points of the F-Distribution, $\alpha = .10$ 964
Table IX Percentage Points of the F-Distribution, $\alpha = .05$ 966
Table X Percentage Points of the F-Distribution, $\alpha = .025$ 968
Table XI Percentage Points of the F-Distribution, $\alpha = .01$ 970
Table XII Critical Values of T_L and T_U for the Wilcoxon Rank Sum Test: Independent Samples 972
Table XIII Critical Values of T_0 in the Wilcoxon Paired Difference Signed Rank Test 973
Table XIV Critical Values of Spearman's Rank Correlation Coefficient 974

APPENDIX B CALCULATION FORMULAS FOR ANALYSIS OF VARIANCE 975

APPENDIX C DEMOGRAPHIC DATA SET 979

Answers to Selected Exercises 983

Index 1011

CASE STUDIES

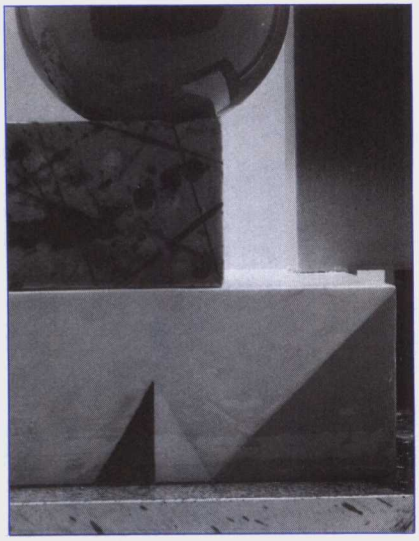
1.1	A Survey: Where "Women's Work" Is Done by Men	2
1.2	An Experiment: Investigating an Effect of Smoking During Pregnancy	2
1.3	Does Judicial Action Affect the Probability of Conviction?	4
1.4	Taste Preference for Beer: Brand Image or Physical Characteristics of the Beer?	4
2.1	Mercury Poisoning and the Dental Profession	28
2.2	Hotels: A Rational Method for Overbooking	35
2.3	Statistics and Air Quality Standards	62
2.4	*Children Out of School in America:* Making an Ugly Picture Look Worse	79
3.1	Bloom County Probabilities	95
3.2	Comparing Subjective Probability Assessments with Relative Frequencies	100
3.3	Purchase Patterns and the Conditional Probability of Purchasing	120
3.4	The 1970 Draft Lottery	143
4.1	A Restaurant Chain Fights Sales Tax Claim	183
4.2	The Space Shuttle *Challenger*: Catastrophe in Space	191
4.3	A Survey of Children's Political Knowledge	197
5.1	Grading on the Curve	246
5.2	Assessing the Reliability of Computer Software	266
7.1	Dancing to the Customer's Tune: The Need to Assess Customer Preferences	322
7.2	Statistics Is Murder!	336
7.3	Hypothesis Tests in Computer Security Systems	343
8.1	Productivity and Mobility	410
8.2	Matched Pairing in Studying the Mentally Retarded	435
9.1	Comparing the Strengths of Women Athletes	498
9.2	Anxiety Levels and Mathematical Achievement	539
10.1	Consumer Rankings of Products	602
10.2	Does Perceived Prestige of Occupation Depend on Titles of Jobs?	615
10.3	The Problem of Nonresponse Bias in Mail Surveys	616
11.1	Evaluating a New Method for Treating Cancer	655
11.2	Does Aid to a Crime Victim Depend on Commitment of a Bystander?	664

12.1	Predicting the United States Crime Index	716
13.1	Predicting Corporate Executive Compensation	759
13.2	An Apparent Inconsistency: The Culprit, Multicollinearity	825
14.1	Modeling Strike Activity	867
14.2	A Statistical Method for Land Appraisal	928

CHAPTER 1

WHERE WE'RE GOING . . .

Statistics? Is it a field of study, a group of numbers that summarize the state of our national economy, the performance of a football team, the social conditions in a particular locale, or, as the title of a popular book (Tanur et al., 1978) suggests, "a guide to the unknown"? We will attempt to answer this question in Chapter 1. Throughout the remainder of the text, we will show you how statistics can be used to interpret experimental and sample survey data. Since many jobs in government, industry, medicine, and other fields require this facility, you will see how statistics can be beneficial to you.



WHAT IS STATISTICS?

CONTENTS

1.1 Statistics: What Is It?

1.2 The Elements of Statistics

1.3 Statistics: Witchcraft or Science?

1.4 Why Study Statistics?

Case Study 1.1 A Survey: Where "Women's Work" Is Done by Men

Case Study 1.2 An Experiment: Investigating an Effect of Smoking During Pregnancy

Case Study 1.3 Does Judicial Action Affect the Probability of Conviction?

Case Study 1.4 Taste Preference for Beer: Brand Image or Physical Characteristics of the Beer?