

# Contents

## Preface ix

## Chapter 1 Introduction 1

**STATISTICS IN ACTION** DDT Contamination of Fish in the Tennessee River 2

1.1 Statistics: The Science of Data 2

1.2 Fundamental Elements of Statistics 3

1.3 Types of Data 6

1.4 Collecting Data: Sampling 8

1.5 The Role of Statistics in Critical Thinking 16

1.6 A Guide to Statistical Methods Presented in This Text 16

**STATISTICS IN ACTION REVISITED** DDT Contamination of Fish in the Tennessee River—Identifying the Data Collection Method, Population, Sample, and Types of Data 18

## Chapter 2 Descriptive Statistics 21

**STATISTICS IN ACTION** Characteristics of Contaminated Fish in the Tennessee River, Alabama 22

2.1 Graphical and Numerical Methods for Describing Qualitative Data 22

2.2 Graphical Methods for Describing Quantitative Data 29

2.3 Numerical Methods for Describing Quantitative Data 39

2.4 Measures of Central Tendency 39

2.5 Measures of Variation 46

2.6 Measures of Relative Standing 52

2.7 Methods for Detecting Outliers 55

2.8 Distorting the Truth with Descriptive Statistics 60

**STATISTICS IN ACTION REVISITED** Characteristics of Contaminated Fish in the Tennessee River, Alabama 65

## Chapter 3 Probability 76

**STATISTICS IN ACTION** Assessing Predictors of Software Defects in NASA Spacecraft Instrument Code 77

3.1 The Role of Probability in Statistics 78

3.2 Events, Sample Spaces, and Probability 78

3.3 Compound Events 88

3.4 Complementary Events 90

3.5 Conditional Probability 94

3.6 Probability Rules for Unions and Intersections 99

3.7 Bayes' Rule (Optional) 109

3.8 Some Counting Rules 112

3.9 Probability and Statistics: An Example 123

**STATISTICS IN ACTION REVISITED** Assessing Predictors of Software Defects in NASA Spacecraft Instrument Code 125

## Chapter 4 Discrete Random Variables 133

- STATISTICS IN ACTION* The Reliability of a "One-Shot" Device 134
- 4.1 Discrete Random Variables 134
  - 4.2 The Probability Distribution for a Discrete Random Variable 135
  - 4.3 Expected Values for Random Variables 140
  - 4.4 Some Useful Expectation Theorems 144
  - 4.5 Bernoulli Trials 146
  - 4.6 The Binomial Probability Distribution 147
  - 4.7 The Multinomial Probability Distribution 154
  - 4.8 The Negative Binomial and the Geometric Probability Distributions 159
  - 4.9 The Hypergeometric Probability Distribution 164
  - 4.10 The Poisson Probability Distribution 168
  - 4.11 Moments and Moment Generating Functions (Optional) 175
- STATISTICS IN ACTION REVISITED* The Reliability of a "One-Shot" Device 178

## Chapter 5 Continuous Random Variables 186

- STATISTICS IN ACTION* Super Weapons Development—Optimizing the Hit Ratio 187
- 5.1 Continuous Random Variables 187
  - 5.2 The Density Function for a Continuous Random Variable 189
  - 5.3 Expected Values for Continuous Random Variables 192
  - 5.4 The Uniform Probability Distribution 197
  - 5.5 The Normal Probability Distribution 200
  - 5.6 Descriptive Methods for Assessing Normality 206
  - 5.7 Gamma-Type Probability Distributions 212
  - 5.8 The Weibull Probability Distribution 216
  - 5.9 Beta-Type Probability Distributions 220
  - 5.10 Moments and Moment Generating Functions (Optional) 223
- STATISTICS IN ACTION REVISITED* Super Weapons Development—Optimizing the Hit Ratio 225

## Chapter 6 Bivariate Probability Distributions and Sampling Distributions 234

- STATISTICS IN ACTION* Availability of an Up/Down Maintained System 235
- 6.1 Bivariate Probability Distributions for Discrete Random Variables 235
  - 6.2 Bivariate Probability Distributions for Continuous Random Variables 241
  - 6.3 The Expected Value of Functions of Two Random Variables 245
  - 6.4 Independence 247
  - 6.5 The Covariance and Correlation of Two Random Variables 250
  - 6.6 Probability Distributions and Expected Values of Functions of Random Variables (Optional) 253
  - 6.7 Sampling Distributions 261
  - 6.8 Approximating a Sampling Distribution by Monte Carlo Simulation 262
  - 6.9 The Sampling Distributions of Means and Sums 265
  - 6.10 Normal Approximation to the Binomial Distribution 271
  - 6.11 Sampling Distributions Related to the Normal Distribution 274
- STATISTICS IN ACTION REVISITED* Availability of an Up/Down Maintained System 280

## Chapter 7 Estimation Using Confidence Intervals 288

<b>STATISTICS IN ACTION</b>	Bursting Strength of PET Beverage Bottles	289
7.1	Point Estimators and their Properties	289
7.2	Finding Point Estimators: Classical Methods of Estimation	294
7.3	Finding Interval Estimators: The Pivotal Method	301
7.4	Estimation of a Population Mean	308
7.5	Estimation of the Difference Between Two Population Means: Independent Samples	314
7.6	Estimation of the Difference Between Two Population Means: Matched Pairs	322
7.7	Estimation of a Population Proportion	329
7.8	Estimation of the Difference Between Two Population Proportions	331
7.9	Estimation of a Population Variance	336
7.10	Estimation of the Ratio of Two Population Variances	340
7.11	Choosing the Sample Size	346
7.12	Alternative Interval Estimation Methods: Bootstrapping and Bayesian Methods (Optional)	350
<b>STATISTICS IN ACTION REVISITED</b>	Bursting Strength of PET Beverage Bottles	355

## Chapter 8 Tests of Hypotheses 368

<b>STATISTICS IN ACTION</b>	Comparing Methods for Dissolving Drug Tablets—Dissolution Method Equivalence Testing	369
8.1	The Relationship Between Statistical Tests of Hypotheses and Confidence Intervals	370
8.2	Elements and Properties of a Statistical Test	370
8.3	Finding Statistical Tests: Classical Methods	376
8.4	Choosing the Null and Alternative Hypotheses	381
8.5	The Observed Significance Level for a Test	382
8.6	Testing a Population Mean	386
8.7	Testing the Difference Between Two Population Means: Independent Samples	393
8.8	Testing the Difference Between Two Population Means: Matched Pairs	402
8.9	Testing a Population Proportion	408
8.10	Testing the Difference Between Two Population Proportions	411
8.11	Testing a Population Variance	416
8.12	Testing the Ratio of Two Population Variances	420
8.13	Alternative Testing Procedures: Bootstrapping and Bayesian Methods (Optional)	426
<b>STATISTICS IN ACTION REVISITED</b>	Comparing Methods for Dissolving Drug Tablets—Dissolution Method Equivalence Testing	431

## Chapter 9 Categorical Data Analysis 442

<b>STATISTICS IN ACTION</b>	The Case of the Ghoulish Transplant Tissue – Who is Responsible for Paying Damages?	443
9.1	Categorical Data and Multinomial Probabilities	444
9.2	Estimating Category Probabilities in a One-Way Table	444
9.3	Testing Category Probabilities in a One-Way Table	448
9.4	Inferences About Category Probabilities in a Two-Way (Contingency) Table	453
9.5	Contingency Tables with Fixed Marginal Totals	462
9.6	Exact Tests for Independence in a Contingency Table Analysis (Optional)	467
<b>STATISTICS IN ACTION REVISITED</b>	The Case of the Ghoulish Transplant Tissue	473

**Chapter 10 Simple Linear Regression 482**

- STATISTICS IN ACTION** Can Dowders Really Detect Water? 483
- 10.1 Regression Models 484
  - 10.2 Model Assumptions 485
  - 10.3 Estimating  $\beta_0$  and  $\beta_1$ : The Method of Least Squares 488
  - 10.4 Properties of the Least-Squares Estimators 500
  - 10.5 An Estimator of  $\sigma^2$  503
  - 10.6 Assessing the Utility of the Model: Making Inferences About the Slope  $\beta_1$  507
  - 10.7 The Coefficients of Correlation and Determination 513
  - 10.8 Using the Model for Estimation and Prediction 521
  - 10.9 Checking the Assumptions: Residual Analysis 530
  - 10.10 A Complete Example 541
  - 10.11 A Summary of the Steps to Follow in Simple Linear Regression 546
- STATISTICS IN ACTION REVISITED** Can Dowders Really Detect Water? 546

**Chapter 11 Multiple Regression Analysis 556**

- STATISTICS IN ACTION** Bid-Rigging in the Highway Construction Industry 557
- 11.1 General Form of a Multiple Regression Model 558
  - 11.2 Model Assumptions 559
  - 11.3 Fitting the Model: The Method of Least Squares 560
  - 11.4 Computations Using Matrix Algebra: Estimating and Making Inferences About the Individual  $\beta$  Parameters 561
  - 11.5 Assessing Overall Model Adequacy 568
  - 11.6 A Confidence Interval for  $E(y)$  and a Prediction Interval for a Future Value of  $y$  572
  - 11.7 A First-Order Model with Quantitative Predictors 582
  - 11.8 An Interaction Model with Quantitative Predictors 592
  - 11.9 A Quadratic (Second-Order) Model with a Quantitative Predictor 597
  - 11.10 Regression Residuals and Outliers 605
  - 11.11 Some Pitfalls: Estimability, Multicollinearity, and Extrapolation 617
  - 11.12 A Summary of the Steps to Follow in a Multiple Regression Analysis 626
- STATISTICS IN ACTION REVISITED** Building a Model for Road Construction Costs in a Sealed Bid Market 627

**Chapter 12 Model Building 642**

- STATISTICS IN ACTION** Deregulation of the Intrastate Trucking Industry 643
- 12.1 Introduction: Why Model Building Is Important 644
  - 12.2 The Two Types of Independent Variables: Quantitative and Qualitative 645
  - 12.3 Models with a Single Quantitative Independent Variable 647
  - 12.4 Models with Two or More Quantitative Independent Variables 654
  - 12.5 Coding Quantitative Independent Variables (Optional) 662
  - 12.6 Models with One Qualitative Independent Variable 667
  - 12.7 Models with Both Quantitative and Qualitative Independent Variables 674
  - 12.8 Tests for Comparing Nested Models 685
  - 12.9 External Model Validation (Optional) 692
  - 12.10 Stepwise Regression 694
- STATISTICS IN ACTION REVISITED** Deregulation in the Intrastate Trucking Industry 701

## Chapter 13 Principles of Experimental Design 716

**STATISTICS IN ACTION** Anti-corrosive Behavior of Epoxy Coatings Augmented with Zinc 717

- 13.1 Introduction 718
- 13.2 Experimental Design Terminology 718
- 13.3 Controlling the Information in an Experiment 720
- 13.4 Noise-Reducing Designs 721
- 13.5 Volume-Increasing Designs 728
- 13.6 Selecting the Sample Size 733
- 13.7 The Importance of Randomization 736

**STATISTICS IN ACTION REVISITED** Anti-Corrosive Behavior of Epoxy Coatings Augmented with Zinc 736

## Chapter 14 The Analysis of Variance for Designed Experiments 742

**STATISTICS IN ACTION** Pollutants at a Housing Development—A Case of Mishandling Small Samples 743

- 14.1 Introduction 743
- 14.2 The Logic Behind an Analysis of Variance 744
- 14.3 One-Factor Completely Randomized Designs 746
- 14.4 Randomized Block Designs 758
- 14.5 Two-Factor Factorial Experiments 772
- 14.6 More Complex Factorial Designs (Optional) 791
- 14.7 Nested Sampling Designs (Optional) 799
- 14.8 Multiple Comparisons of Treatment Means 810
- 14.9 Checking ANOVA Assumptions 817

**STATISTICS IN ACTION REVISITED** Pollutants at a Housing Development—A Case of Mishandling Small Samples 821

## Chapter 15 Nonparametric Statistics 837

**STATISTICS IN ACTION** How Vulnerable Are New Hampshire Wells to Groundwater Contamination? 838

- 15.1 Introduction: Distribution-Free Tests 839
- 15.2 Testing for Location of a Single Population 840
- 15.3 Comparing Two Populations: Independent Random Samples 845
- 15.4 Comparing Two Populations: Matched-Pairs Design 853
- 15.5 Comparing Three or More Populations: Completely Randomized Design 859
- 15.6 Comparing Three or More Populations: Randomized Block Design 864
- 15.7 Nonparametric Regression 869

**STATISTICS IN ACTION REVISITED** How Vulnerable are New Hampshire Wells to Groundwater Contamination? 876

## Chapter 16 Statistical Process and Quality Control 888

**STATISTICS IN ACTION** Testing Jet Fuel Additive for Safety 889

- 16.1 Total Quality Management 890
- 16.2 Variable Control Charts 890
- 16.3 Control Chart for Means:  $\bar{x}$ -Chart 896
- 16.4 Control Chart for Process Variation: R-Chart 904

16.5	Detecting Trends in a Control Chart: Runs Analysis	910
16.6	Control Chart for Percent Defectives: $p$ -Chart	912
16.7	Control Chart for the Number of Defects per Item: $c$ -Chart	917
16.8	Tolerance Limits	921
16.9	Capability Analysis (Optional)	925
16.10	Acceptance Sampling for Defectives	933
16.11	Other Sampling Plans (Optional)	937
16.12	Evolutionary Operations (Optional)	938
STATISTICS IN ACTION REVISITED Testing Jet Fuel Additive for Safety		939

## Chapter 17 Product and System Reliability 951

STATISTICS IN ACTION Modeling the Hazard Rate of Reinforced Concrete Bridge Deck Deterioration		952
17.1	Introduction	952
17.2	Failure Time Distributions	952
17.3	Hazard Rates	954
17.4	Life Testing: Censored Sampling	958
17.5	Estimating the Parameters of an Exponential Failure Time Distribution	959
17.6	Estimating the Parameters of a Weibull Failure Time Distribution	962
17.7	System Reliability	967
STATISTICS IN ACTION REVISITED Modeling the Hazard Rate of Reinforced Concrete Bridge Deck Deterioration		971

## Appendix A Matrix Algebra 977

A.1	Matrices and Matrix Multiplication	977
A.2	Identity Matrices and Matrix Inversion	981
A.3	Solving Systems of Simultaneous Linear Equations	984
A.4	A Procedure for Inverting a Matrix	986

## Appendix B Useful Statistical Tables 991

TABLE 1	Random Numbers	992
TABLE 2	Cumulative Binomial Probabilities	996
TABLE 3	Exponentials	1000
TABLE 4	Cumulative Poisson Probabilities	1001
TABLE 5	Normal Curve Areas	1003
TABLE 6	Gamma Function	1004
TABLE 7	Critical Values for Student's $T$	1005
TABLE 8	Critical Values of $\chi^2$	1006
TABLE 9	Percentage Points of the $F$ Distribution, $\alpha = .10$	1008
TABLE 10	Percentage Points of the $F$ Distribution, $\alpha = .05$	1010
TABLE 11	Percentage Points of the $F$ Distribution, $\alpha = .025$	1012
TABLE 12	Percentage Points of the $F$ Distribution, $\alpha = .01$	1014
TABLE 13	Percentage Points of the Studentized Range $q(p, \nu)$ , $\alpha = .05$	1016
TABLE 14	Percentage Points of the Studentized Range $q(p, \nu)$ , $\alpha = .01$	1018
TABLE 15	Critical Values of $T_L$ and $T_U$ for the Wilcoxon Rank Sum Test: Independent Samples	1020

TABLE 16 Critical Values of $T_0$ in the Wilcoxon Matched-Pairs Signed Rank Test	1021
TABLE 17 Critical Values of Spearman's Rank Correlation Coefficient	1022
TABLE 18 Critical Values of $C$ for the Theil Zero-Slope Test	1023
TABLE 19 Factors Used When Constructing Control Charts	1027
TABLE 20 Values of $K$ for Tolerance Limits for Normal Distributions	1028
TABLE 22 Sample Size Code Letters: MIL-STD-105D	1029
TABLE 21 Sample Size $n$ for Nonparametric Tolerance Limits	1029
TABLE 23 A Portion of the Master Table for Normal Inspection (Single Sampling): MIL-STD-105D	1030

## Appendix C SAS for Windows Tutorial 1031

## APPENDIX D MINITAB for Windows Tutorial 1062

## APPENDIX E SPSS for Windows Tutorial 1094

### References 1125

### Selected Short Answers 1133

### Credits 1147

### Index 1152