

An Introduction to

# Statistical Methods & Data Analysis

SEVENTH EDITION

R. Lyman Ott  
Michael Longnecker  
**Texas A&M University**



# CONTENTS

Preface xi

## PART 1 INTRODUCTION 1

### CHAPTER 1 **Statistics and the Scientific Method 2**

- 1.1 Introduction 2
- 1.2 Why Study Statistics? 6
- 1.3 Some Current Applications of Statistics 9
- 1.4 A Note to the Student 13
- 1.5 Summary 13
- 1.6 Exercises 14

## PART 2 COLLECTING DATA 17

### CHAPTER 2 **Using Surveys and Experimental Studies to Gather Data 18**

- 2.1 Introduction and Abstract of Research Study 18
- 2.2 Observational Studies 20
- 2.3 Sampling Designs for Surveys 26
- 2.4 Experimental Studies 32
- 2.5 Designs for Experimental Studies 38
- 2.6 Research Study: Exit Polls Versus Election Results 48
- 2.7 Summary 50
- 2.8 Exercises 50

## PART 3 SUMMARIZING DATA 59

### CHAPTER 3 **Data Description 60**

- 3.1 Introduction and Abstract of Research Study 60
- 3.2 Calculators, Computers, and Software Systems 65
- 3.3 Describing Data on a Single Variable: Graphical Methods 66
- 3.4 Describing Data on a Single Variable: Measures of Central Tendency 82
- 3.5 Describing Data on a Single Variable: Measures of Variability 90
- 3.6 The Boxplot 104
- 3.7 Summarizing Data from More Than One Variable: Graphs and Correlation 109

- 3.8 Research Study: Controlling for Student Background in the Assessment of Teaching 119
- 3.9 R Instructions 124
- 3.10 Summary and Key Formulas 124
- 3.11 Exercises 125

## CHAPTER 4

### Probability and Probability Distributions 149

- 4.1 Introduction and Abstract of Research Study 149
- 4.2 Finding the Probability of an Event 153
- 4.3 Basic Event Relations and Probability Laws 155
- 4.4 Conditional Probability and Independence 158
- 4.5 Bayes' Formula 161
- 4.6 Variables: Discrete and Continuous 164
- 4.7 Probability Distributions for Discrete Random Variables 166
- 4.8 Two Discrete Random Variables: The Binomial and the Poisson 167
- 4.9 Probability Distributions for Continuous Random Variables 177
- 4.10 A Continuous Probability Distribution: The Normal Distribution 180
- 4.11 Random Sampling 187
- 4.12 Sampling Distributions 190
- 4.13 Normal Approximation to the Binomial 200
- 4.14 Evaluating Whether or Not a Population Distribution Is Normal 203
- 4.15 Research Study: Inferences About Performance-Enhancing Drugs Among Athletes 208
- 4.16 R Instructions 211
- 4.17 Summary and Key Formulas 212
- 4.18 Exercises 214

### PART 4 ANALYZING THE DATA, INTERPRETING THE ANALYSES, AND COMMUNICATING THE RESULTS 231

## CHAPTER 5

### Inferences About Population Central Values 232

- 5.1 Introduction and Abstract of Research Study 232
- 5.2 Estimation of  $\mu$  235
- 5.3 Choosing the Sample Size for Estimating  $\mu$  240
- 5.4 A Statistical Test for  $\mu$  242
- 5.5 Choosing the Sample Size for Testing  $\mu$  255
- 5.6 The Level of Significance of a Statistical Test 257
- 5.7 Inferences About  $\mu$  for a Normal Population,  $\sigma$  Unknown 260
- 5.8 Inferences About  $\mu$  When the Population Is Nonnormal and  $n$  Is Small: Bootstrap Methods 269
- 5.9 Inferences About the Median 275
- 5.10 Research Study: Percentage of Calories from Fat 280
- 5.11 Summary and Key Formulas 283
- 5.12 Exercises 285

## CHAPTER 6

### Inferences Comparing Two Population Central Values 300

- 6.1 Introduction and Abstract of Research Study 300
- 6.2 Inferences About  $\mu_1 - \mu_2$ : Independent Samples 303

- 6.3 A Nonparametric Alternative:  
The Wilcoxon Rank Sum Test 315
- 6.4 Inferences About  $\mu_1 - \mu_2$ : Paired Data 325
- 6.5 A Nonparametric Alternative:  
The Wilcoxon Signed-Rank Test 329
- 6.6 Choosing Sample Sizes for Inferences About  $\mu_1 - \mu_2$  334
- 6.7 Research Study: Effects of an Oil Spill on Plant Growth 336
- 6.8 Summary and Key Formulas 341
- 6.9 Exercises 344

## CHAPTER 7

### Inferences About Population Variances 366

- 7.1 Introduction and Abstract of Research Study 366
- 7.2 Estimation and Tests for a Population Variance 368
- 7.3 Estimation and Tests for Comparing Two Population Variances 376
- 7.4 Tests for Comparing  $t > 2$  Population Variances 382
- 7.5 Research Study: Evaluation of Methods for Detecting *E. coli* 385
- 7.6 Summary and Key Formulas 390
- 7.7 Exercises 391

## CHAPTER 8

### Inferences About More Than Two Population Central Values 400

- 8.1 Introduction and Abstract of Research Study 400
- 8.2 A Statistical Test About More Than Two Population Means:  
An Analysis of Variance 403
- 8.3 The Model for Observations in a Completely Randomized Design 412
- 8.4 Checking on the AOV Conditions 414
- 8.5 An Alternative Analysis: Transformations of the Data 418
- 8.6 A Nonparametric Alternative: The Kruskal–Wallis Test 425
- 8.7 Research Study: Effect of Timing on the Treatment  
of Port-Wine Stains with Lasers 428
- 8.8 Summary and Key Formulas 433
- 8.9 Exercises 435

## CHAPTER 9

### Multiple Comparisons 445

- 9.1 Introduction and Abstract of Research Study 445
- 9.2 Linear Contrasts 447
- 9.3 Which Error Rate Is Controlled? 454
- 9.4 Scheffé's *S* Method 456
- 9.5 Tukey's *W* Procedure 458
- 9.6 Dunnett's Procedure: Comparison of Treatments to a Control 462
- 9.7 A Nonparametric Multiple-Comparison Procedure 464
- 9.8 Research Study: Are Interviewers' Decisions Affected by Different  
Handicap Types? 467
- 9.9 Summary and Key Formulas 474
- 9.10 Exercises 475

**Categorical Data 482**

- 10.1 Introduction and Abstract of Research Study 482
- 10.2 Inferences About a Population Proportion  $\pi$  483
- 10.3 Inferences About the Difference Between Two Population Proportions,  $\pi_1 - \pi_2$  491
- 10.4 Inferences About Several Proportions: Chi-Square Goodness-of-Fit Test 501
- 10.5 Contingency Tables: Tests for Independence and Homogeneity 508
- 10.6 Measuring Strength of Relation 515
- 10.7 Odds and Odds Ratios 517
- 10.8 Combining Sets of  $2 \times 2$  Contingency Tables 522
- 10.9 Research Study: Does Gender Bias Exist in the Selection of Students for Vocational Education? 525
- 10.10 Summary and Key Formulas 531
- 10.11 Exercises 533

**Linear Regression and Correlation 555**

- 11.1 Introduction and Abstract of Research Study 555
- 11.2 Estimating Model Parameters 564
- 11.3 Inferences About Regression Parameters 574
- 11.4 Predicting New  $y$ -Values Using Regression 577
- 11.5 Examining Lack of Fit in Linear Regression 581
- 11.6 Correlation 587
- 11.7 Research Study: Two Methods for Detecting *E. coli* 598
- 11.8 Summary and Key Formulas 602
- 11.9 Exercises 604

**Multiple Regression and the General Linear Model 625**

- 12.1 Introduction and Abstract of Research Study 625
- 12.2 The General Linear Model 635
- 12.3 Estimating Multiple Regression Coefficients 636
- 12.4 Inferences in Multiple Regression 644
- 12.5 Testing a Subset of Regression Coefficients 652
- 12.6 Forecasting Using Multiple Regression 656
- 12.7 Comparing the Slopes of Several Regression Lines 658
- 12.8 Logistic Regression 662
- 12.9 Some Multiple Regression Theory (Optional) 669
- 12.10 Research Study: Evaluation of the Performance of an Electric Drill 676
- 12.11 Summary and Key Formulas 683
- 12.12 Exercises 685

**Further Regression Topics 711**

- 13.1 Introduction and Abstract of Research Study 711
- 13.2 Selecting the Variables (Step 1) 712
- 13.3 Formulating the Model (Step 2) 729
- 13.4 Checking Model Assumptions (Step 3) 745

- 13.5 Research Study: Construction Costs for Nuclear Power Plants 765
- 13.6 Summary and Key Formulas 772
- 13.7 Exercises 773

**CHAPTER 14**

**Analysis of Variance for Completely Randomized Designs 798**

- 14.1 Introduction and Abstract of Research Study 798
- 14.2 Completely Randomized Design with a Single Factor 800
- 14.3 Factorial Treatment Structure 805
- 14.4 Factorial Treatment Structures with an Unequal Number of Replications 830
- 14.5 Estimation of Treatment Differences and Comparisons of Treatment Means 837
- 14.6 Determining the Number of Replications 841
- 14.7 Research Study: Development of a Low-Fat Processed Meat 846
- 14.8 Summary and Key Formulas 851
- 14.9 Exercises 852

**CHAPTER 15**

**Analysis of Variance for Blocked Designs 865**

- 15.1 Introduction and Abstract of Research Study 865
- 15.2 Randomized Complete Block Design 866
- 15.3 Latin Square Design 878
- 15.4 Factorial Treatment Structure in a Randomized Complete Block Design 889
- 15.5 A Nonparametric Alternative—Friedman’s Test 893
- 15.6 Research Study: Control of Leatherjackets 897
- 15.7 Summary and Key Formulas 902
- 15.8 Exercises 904

**CHAPTER 16**

**The Analysis of Covariance 917**

- 16.1 Introduction and Abstract of Research Study 917
- 16.2 A Completely Randomized Design with One Covariate 920
- 16.3 The Extrapolation Problem 931
- 16.4 Multiple Covariates and More Complicated Designs 934
- 16.5 Research Study: Evaluation of Cool-Season Grasses for Putting Greens 936
- 16.6 Summary 942
- 16.7 Exercises 942

**CHAPTER 17**

**Analysis of Variance for Some Fixed-, Random-, and Mixed-Effects Models 952**

- 17.1 Introduction and Abstract of Research Study 952
- 17.2 A One-Factor Experiment with Random Treatment Effects 955
- 17.3 Extensions of Random-Effects Models 959
- 17.4 Mixed-Effects Models 967
- 17.5 Rules for Obtaining Expected Mean Squares 971

- 17.6 Nested Factors 981
- 17.7 Research Study: Factors Affecting Pressure Drops  
Across Expansion Joints 986
- 17.8 Summary 991
- 17.9 Exercises 992

## CHAPTER 18

### Split-Plot, Repeated Measures, and Crossover Designs 1004

- 18.1 Introduction and Abstract of Research Study 1004
- 18.2 Split-Plot Designed Experiments 1008
- 18.3 Single-Factor Experiments with Repeated Measures 1014
- 18.4 Two-Factor Experiments with Repeated Measures on  
One of the Factors 1018
- 18.5 Crossover Designs 1025
- 18.6 Research Study: Effects of an Oil Spill on Plant Growth 1033
- 18.7 Summary 1035
- 18.8 Exercises 1035

## CHAPTER 19

### Analysis of Variance for Some Unbalanced Designs 1050

- 19.1 Introduction and Abstract of Research Study 1050
- 19.2 A Randomized Block Design with One or More  
Missing Observations 1052
- 19.3 A Latin Square Design with Missing Data 1058
- 19.4 Balanced Incomplete Block (BIB) Designs 1063
- 19.5 Research Study: Evaluation of the Consistency  
of Property Assessors 1070
- 19.6 Summary and Key Formulas 1074
- 19.7 Exercises 1075

### Appendix: Statistical Tables 1085

### Answers to Selected Exercises 1125

### References 1151

### Index 1157