

EDITION

6

# Introduction to Statistics and Data Analysis



**Roxy Peck**

California Polytechnic State University, San Luis Obispo

**Tom Short**

West Chester University of Pennsylvania

**Chris Olsen**

Grinnell College



Australia • Brazil • Mexico • Singapore • United Kingdom • United States

# Brief Contents

- 1** The Role of Statistics and the Data Analysis Process 1
  - 2** Collecting Data Sensibly 28
  - 3** Graphical Methods for Describing Data 77
  - 4** Numerical Methods for Describing Data 148
  - 5** Summarizing Bivariate Data 197
  - 6** Probability 277
  - 7** Random Variables and Probability Distributions 343
  - 8** Sampling Variability and Sampling Distributions 427
  - 9** Estimation Using a Single Sample 453
  - 10** Hypothesis Testing Using a Single Sample 507
  - 11** Comparing Two Populations or Treatments 576
  - 12** The Analysis of Categorical Data and Goodness-of-Fit Tests 654
  - 13** Simple Linear Regression and Correlation: Inferential Methods 689
  - 14** Multiple Regression Analysis 730
  - 15** Analysis of Variance 759
  - 16** Nonparametric (Distribution-Free) Statistical Methods 16-1
- Appendix: Statistical Tables 785
- Answers to Selected Odd-Numbered Exercises 805
- Index 845

Sections and/or chapter numbers shaded in color can be found at  
[www.cengage.com](http://www.cengage.com)

# Contents

## **CHAPTER 1 The Role of Statistics and the Data Analysis Process 1**

- 1.1** Why Study Statistics? 2
- 1.2** The Nature and Role of Variability 3
- 1.3** Statistics and the Data Analysis Process 5
- 1.4** Types of Data and Some Simple Graphical Displays 9
  - Activity 1.1** Head Sizes: Understanding Variability 22
  - Activity 1.2** Estimating Shape Sizes 23
  - Activity 1.3** A Meaningful Paragraph 24
  - Summary: Key Concepts and Formulas 24
  - Chapter Review 24
  - Technology Notes 26

## **CHAPTER 2 Collecting Data Sensibly 28**

- 2.1** Statistical Studies: Observation and Experimentation 29
- 2.2** Sampling 34
- 2.3** Simple Comparative Experiments 45
- 2.4** More on Experimental Design 60
- 2.5** Interpreting and Communicating the Results of Statistical Analyses 65
  - Activity 2.1** Facebook Friending 68
  - Activity 2.2** An Experiment to Test for the Stroop Effect 68
  - Activity 2.3** McDonald's and the Next 100 Billion Burgers 69
  - Activity 2.4** Video Games and Pain Management 69
  - Activity 2.5** Be Careful with Random Assignment! 70
  - Summary: Key Concepts and Formulas 70
  - Chapter Review 71
  - Technology Notes 73

See Chapter 2 online materials for More on Observational Studies: Designing Surveys.

## **CHAPTER 3 Graphical Methods for Describing Data 77**

- 3.1** Displaying Categorical Data: Comparative Bar Charts and Pie Charts 78
- 3.2** Displaying Numerical Data: Stem-and-Leaf Displays 88
- 3.3** Displaying Numerical Data: Frequency Distributions and Histograms 97
- 3.4** Displaying Bivariate Numerical Data 116
- 3.5** Interpreting and Communicating the Results of Statistical Analyses 125
  - Activity 3.1** Locating States 134
  - Activity 3.2** Bean Counters! 134
  - Summary: Key Concepts and Formulas 135
  - Chapter Review 135

Technology Notes 139  
Cumulative Review Exercises 144

## **CHAPTER 4 Numerical Methods for Describing Data 148**

- 4.1** Describing the Center of a Data Set 149
- 4.2** Describing Variability in a Data Set 159
- 4.3** Summarizing a Data Set: Boxplots 168
- 4.4** Interpreting Center and Variability: Chebyshev's Rule, the Empirical Rule, and  $z$  Scores 175
- 4.5** Interpreting and Communicating the Results of Statistical Analyses 183
  - Activity 4.1** Collecting and Summarizing Numerical Data 188
  - Activity 4.2** Airline Passenger Weights 188
  - Activity 4.3** Boxplot Shapes 188
- Summary: Key Concepts and Formulas 189
- Chapter Review 189
- Technology Notes 191

## **CHAPTER 5 Summarizing Bivariate Data 197**

- 5.1** Correlation 198
- 5.2** Linear Regression: Fitting a Line to Bivariate Data 209
- 5.3** Assessing the Fit of a Line 221
- 5.4** Nonlinear Relationships and Transformations 241
- 5.5** Interpreting and Communicating the Results of Statistical Analyses 259
  - Activity 5.1** Age and Flexibility 265
- Summary: Key Concepts and Formulas 265
- Chapter Review 266
- Technology Notes 269
- Cumulative Review Exercises 273

See Chapter 5 online materials for coverage of Logistic Regression.

## **CHAPTER 6 Probability 277**

- 6.1** Chance Experiments and Events 278
- 6.2** Definition of Probability 285
- 6.3** Basic Properties of Probability 290
- 6.4** Conditional Probability 297
- 6.5** Independence 307
- 6.6** Some General Probability Rules 315
- 6.7** Estimating Probabilities Empirically and Using Simulation 327
  - Activity 6.1** Kisses 337
  - Activity 6.2** A Crisis for European Sports Fans? 338
  - Activity 6.3** The "Hot Hand" in Basketball 338
- Summary: Key Concepts and Formulas 339
- Chapter Review 339

## **CHAPTER 7 Random Variables and Probability Distributions 343**

- 7.1** Random Variables 344
- 7.2** Probability Distributions for Discrete Random Variables 347
- 7.3** Probability Distributions for Continuous Random Variables 353
- 7.4** Mean and Standard Deviation of a Random Variable 358

- 7.5 Binomial and Geometric Distributions 371
- 7.6 Normal Distributions 383
- 7.7 Checking for Normality and Normalizing Transformations 400
- 7.8 Using the Normal Distribution to Approximate a Discrete Distribution (Optional) 410
  - Activity 7.1 Is It Real? 415
  - Activity 7.2 Rotten Eggs? 416
  - Summary: Key Concepts and Formulas 416
  - Chapter Review 417
  - Technology Notes 420
  - Cumulative Review Exercises 423

## CHAPTER 8 Sampling Variability and Sampling Distributions 427

- 8.1 Statistics and Sampling Variability 428
- 8.2 The Sampling Distribution of a Sample Mean 432
- 8.3 The Sampling Distribution of a Sample Proportion 441
  - Activity 8.1 Sampling Distribution of the Sample Mean 447
  - Activity 8.2 Sampling Distribution of the Sample Proportion 449
  - Activity 8.3 Do Students Who Take the SATs Multiple Times Have an Advantage in College Admissions? 450
  - Summary: Key Concepts and Formulas 452
  - Chapter Review 452

## CHAPTER 9 Estimation Using a Single Sample 453

- 9.1 Point Estimation 454
- 9.2 Large-Sample Confidence Interval for a Population Proportion 459
- 9.3 Confidence Interval for a Population Mean 472
- 9.4 Interpreting and Communicating the Results of Statistical Analyses 484
- 9.5 Bootstrap Confidence Intervals for a Population Proportion (Optional) 489
- 9.6 Bootstrap Confidence Intervals for a Population Mean (Optional) 496
  - Activity 9.1 Getting a Feel for Confidence Level 500
  - Activity 9.2 An Alternative Confidence Interval for a Population Proportion 501
  - Activity 9.3 Verifying Signatures on a Recall Petition 502
  - Activity 9.4 A Meaningful Paragraph 502
  - Summary: Key Concepts and Formulas 502
  - Chapter Review 503
  - Technology Notes 504

## CHAPTER 10 Hypothesis Testing Using a Single Sample 507

- 10.1 Hypotheses and Test Procedures 508
- 10.2 Errors in Hypothesis Testing 512
- 10.3 Large-Sample Hypothesis Tests for a Population Proportion 517
- 10.4 Hypothesis Tests for a Population Mean 530
- 10.5 Power and Probability of Type II Error 541
- 10.6 Interpreting and Communicating the Results of Statistical Analyses 549
- 10.7 Randomization Test and Exact Binomial Test for a Population Proportion (Optional) 552

- 10.8** Randomization Test for a Population Mean (Optional) 562
  - Activity 10.1** Comparing the  $t$  and  $z$  Distributions 567
  - Activity 10.2** A Meaningful Paragraph 568
  - Summary: Key Concepts and Formulas 568
  - Chapter Review 569
  - Technology Notes 571
  - Cumulative Review Exercises 573

## **CHAPTER 11** Comparing Two Populations or Treatments 576

- 11.1** Inferences Concerning the Difference Between Two Population or Treatment Means Using Independent Samples 577
- 11.2** Inferences Concerning the Difference Between Two Population or Treatment Means Using Paired Samples 595
- 11.3** Large-Sample Inferences Concerning the Difference Between Two Population or Treatment Proportions 608
- 11.4** Interpreting and Communicating the Results of Statistical Analyses 619
- 11.5** Simulation-Based Inference for Two Means (Optional) 623
- 11.6** Simulation-Based Inference for Two Proportions (Optional) 633
  - Activity 11.1** Helium-Filled Footballs? 641
  - Activity 11.2** Thinking About Data Collection 642
  - Activity 11.3** A Meaningful Paragraph 642
  - Summary: Key Concepts and Formulas 642
  - Chapter Review 643
  - Technology Notes 646

## **CHAPTER 12** The Analysis of Categorical Data and Goodness-of-Fit Tests 654

- 12.1** Chi-Square Tests for Univariate Data 655
- 12.2** Tests for Homogeneity and Independence in a Two-way Table 665
- 12.3** Interpreting and Communicating the Results of Statistical Analyses 679
  - Activity 12.1** Pick a Number, Any Number ... 683
  - Activity 12.2** Color and Perceived Taste 683
  - Summary: Key Concepts and Formulas 684
  - Chapter Review 684
  - Technology Notes 685

## **CHAPTER 13** Simple Linear Regression and Correlation: Inferential Methods 689

- 13.1** Simple Linear Regression Model 690
- 13.2** Inferences About the Slope of the Population Regression Line 702
- 13.3** Checking Model Adequacy 713
  - Activity 13.1** Are Tall Women from “Big” Families? 724
  - Summary: Key Concepts and Formulas 725
  - Technology Notes 725
  - Cumulative Review Exercises 726
- 13.4** Inferences Based on the Estimated Regression Line 13-1
- 13.5** Inferences About the Population Correlation Coefficient 13-8
- 13.6** Interpreting and Communicating the Results of Statistical Analyses 13-11

## CHAPTER 14 Multiple Regression Analysis 730

- 14.1 Multiple Regression Models 731
- 14.2 Fitting a Model and Assessing Its Utility 742
  - Activity 14.1 Exploring the Relationship Between Number of Predictors and Sample Size 758
  - Summary: Key Concepts and Formulas 758
- 14.3 Inferences Based on an Estimated Model 14-1
- 14.4 Other Issues in Multiple Regression 14-12
- 14.5 Interpreting and Communicating the Results of Statistical Analyses 14-22
- Chapter Review 14-23

## CHAPTER 15 Analysis of Variance 759

- 15.1 Single-Factor ANOVA and the  $F$  Test 760
- 15.2 Multiple Comparisons 772
  - Activity 15.1 Exploring Single-Factor ANOVA 780
  - Summary: Key Concepts and Formulas 782
  - Technology Notes 782
- 15.3 The  $F$  Test for a Randomized Block Experiment 15-1
- 15.4 Two-Factor ANOVA 15-7
- 15.5 Interpreting and Communicating the Results of Statistical Analyses 15-17
- Chapter Review 15-21

## CHAPTER 16 Nonparametric (Distribution-Free) Statistical Methods 16-1

- 16.1 Distribution-Free Procedures for Inferences About a Difference Between Two Population or Treatment Means Using Independent Samples 16-2
- 16.2 Distribution-Free Procedures for Inferences About a Difference Between Two Population or Treatment Means Using Paired Samples 16-9
- 16.3 Distribution-Free ANOVA 16-19
  - Summary: Key Concepts and Formulas 16-26
  - Appendix: Tables 16-27

Appendix: Statistical Tables 785

Answers to Selected Odd-Numbered Exercises 805

Index 845

Sections and/or chapter numbers shaded in color can be found at [www.cengage.com](http://www.cengage.com)