

INTERNATIONAL STUDENT EDITION

Roxy Peck

Chris Olsen

Jay Devore

3rd
Edition

Introduction to

Statistics

&

Data
Analysis

Not for Sale in the
United States

Contents



1 The Role of Statistics and the Data Analysis Process 1

1.1 Three Reasons to Study Statistics 1

1.2 The Nature and Role of Variability 4

1.3 Statistics and the Data Analysis Process 7

1.4 Types of Data and Some Simple Graphical Displays 12

Activity 1.1 Head Sizes: Understanding Variability 22

Activity 1.2 Estimating Sizes 23

Activity 1.3 A Meaningful Paragraph 24



2 Collecting Data Sensibly 27

2.1 Statistical Studies: Observation and Experimentation 27

2.2 Sampling 32

2.3 Simple Comparative Experiments 42

2.4 More on Experimental Design 51

2.5 More on Observational Studies: Designing Surveys (Optional) 56

2.6 Interpreting and Communicating the Results of Statistical Analyses 61

Activity 2.1 Designing a Sampling Plan 63

Activity 2.2 An Experiment to Test for the Stroop Effect 64

Activity 2.3 McDonald's and the Next 100 Billion Burgers 64

Activity 2.4 Video Games and Pain Management 65

Graphing Calculator Explorations 69



3 Graphical Methods for Describing Data 75

3.1 Displaying Categorical Data: Comparative Bar Charts and Pie Charts 76

3.2 Displaying Numerical Data: Stem-and-Leaf Displays 87

3.3 Displaying Numerical Data: Frequency Distributions and Histograms **97**

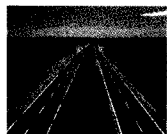
3.4 Displaying Bivariate Numerical Data **117**

3.5 Interpreting and Communicating the Results of Statistical Analyses **127**

Activity 3.1 Locating States **134**

Activity 3.2 Bean Counters! **134**

Graphing Calculator Explorations **141**



4 Numerical Methods for Describing Data 147

4.1 Describing the Center of a Data Set **148**

4.2 Describing Variability in a Data Set **159**

4.3 Summarizing a Data Set: Boxplots **169**

4.4 Interpreting Center and Variability: Chebyshev's Rule, the Empirical Rule, and z Scores **176**

4.5 Interpreting and Communicating the Results of Statistical Analyses **186**

Activity 4.1 Collecting and Summarizing Numerical Data **190**

Activity 4.2 Airline Passenger Weights **190**

Activity 4.3 Boxplot Shapes **190**

Graphing Calculator Explorations **195**



5 Summarizing Bivariate Data 199

5.1 Correlation **200**

5.2 Linear Regression: Fitting a Line to Bivariate Data **210**

5.3 Assessing the Fit of a Line **221**

5.4 Nonlinear Relationships and Transformations **238**

5.5 Logistic Regression (Optional) **255**

5.6 Interpreting and Communicating the Results of Statistical Analyses **264**

Activity 5.1 Exploring Correlation and Regression **267**

Activity 5.2 Age and Flexibility **268**

Graphing Calculator Explorations **272**

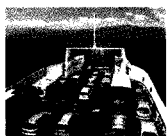


6 Probability 279

6.1 Chance Experiments and Events **279**

6.2 Definition of Probability **288**

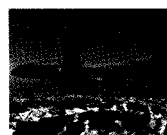
- 6.3 Basic Properties of Probability 295**
- 6.4 Conditional Probability 302**
- 6.5 Independence 313**
- 6.6 Some General Probability Rules 323**
- 6.7 Estimating Probabilities Empirically Using Simulation 335**
 - Activity 6.1 Kisses 347**
 - Activity 6.2 A Crisis for European Sports Fans? 347**
 - Activity 6.3 The “Hot Hand” in Basketball 347**
 - Graphing Calculator Explorations 351**



- 7 Random Variables and Probability Distributions 357**
 - 7.1 Random Variables 358**
 - 7.2 Probability Distributions for Discrete Random Variables 361**
 - 7.3 Probability Distributions for Continuous Random Variables 367**
 - 7.4 Mean and Standard Deviation of a Random Variable 372**
 - 7.5 Binomial and Geometric Distributions 386**
 - 7.6 Normal Distributions 397**
 - 7.7 Checking for Normality and Normalizing Transformations 414**
 - 7.8 Using the Normal Distribution to Approximate a Discrete Distribution 425**
 - Activity 7.1 Rotten Eggs? 429**
 - Graphing Calculator Explorations 434**

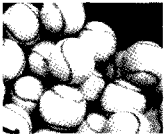


- 8 Sampling Variability and Sampling Distributions 445**
 - 8.1 Statistics and Sampling Variability 446**
 - 8.2 The Sampling Distribution of a Sample Mean 450**
 - 8.3 The Sampling Distribution of a Sample Proportion 461**
 - Activity 8.1 Do Students Who Take the SATs Multiple Times Have an Advantage in College Admissions? 468**
 - Graphing Calculator Explorations 471**



- 9 Estimation Using a Single Sample 475**
 - 9.1 Point Estimation 476**
 - 9.2 Large-Sample Confidence Interval for a Population Proportion 482**
 - 9.3 Confidence Interval for a Population Mean 495**

- 9.4** Interpreting and Communicating the Results of Statistical Analyses **508**
- Activity 9.1** Getting a Feel for Confidence Level **514**
- Activity 9.2** An Alternative Confidence Interval for a Population Proportion **515**
- Activity 9.3** Verifying Signatures on a Recall Petition **516**
- Activity 9.4** A Meaningful Paragraph **516**
- Graphing Calculator Explorations **521**



- 10** Hypothesis Testing Using a Single Sample **525**
- 10.1** Hypotheses and Test Procedures **526**
- 10.2** Errors in Hypotheses Testing **531**
- 10.3** Large-Sample Hypothesis Tests for a Population Proportion **537**
- 10.4** Hypotheses Tests for a Population Mean **550**
- 10.5** Power and Probability of Type II Error **562**
- 10.6** Interpreting and Communicating the Results of Statistical Analyses **571**
- Activity 10.1** Comparing the t and z Distributions **574**
- Activity 10.2** A Meaningful Paragraph **575**
- Graphing Calculator Explorations **580**



- 11** Comparing Two Populations or Treatments **583**
- 11.1** Inferences Concerning the Difference Between Two Population or Treatment Means Using Independent Samples **583**
- 11.2** Inferences Concerning the Difference Between Two Population or Treatment Means Using Paired Samples **606**
- 11.3** Large Sample Inferences Concerning a Difference Between Two Population or Treatment Proportions **619**
- 11.4** Interpreting and Communicating the Results of Statistical Analyses **629**
- Activity 11.1** Helium-Filled Footballs **632**
- Activity 11.2** Thinking About Data Collection **633**
- Activity 11.3** A Meaningful Paragraph **633**
- Graphing Calculator Explorations **641**



12 The Analysis of Categorical Data and Goodness-of-Fit Tests 647

12.1 Chi-Square Tests for Univariate Data 647

12.2 Tests for Homogeneity and Independence in a Two-way Table 660

12.3 Interpreting and Communicating the Results of Statistical Analyses 677

Activity 12.1 Pick a Number, Any Number . . . 680

Activity 12.2 Color and Perceived Taste 680

Graphing Calculator Explorations 685



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

13.4 Inferences Based on the Estimated Regression Line (Optional) 725

13.5 Inferences About the Population Correlation Coefficient (Optional) 734

13.6 Interpreting and Communicating the Results of Statistical Analyses 737

Activity 13.1 Are Tall Women from “Big” Families? 739

Graphing Calculator Exploration 746



14 Multiple Regression Analysis 749

14.1 Multiple Regression Models 750

14.2 Fitting a Model and Assessing Its Utility 763

14.3 Inferences Based on an Estimated Model 14-1

14.4 Other Issues in Multiple Regression 14-13

14.5 Interpreting and Communicating the Results of Statistical Analyses 14-26

Activity 14.1 Exploring the Relationship Between Number of Predictors and Sample Size 780



15 Analysis of Variance 783

15.1 Single-Factor ANOVA and the F Test 784

15.2 Multiple Comparisons 800

15.3 The F Test for a Randomized Block Experiment 15-1

15.4 Two-Factor ANOVA 15-9

**15.5 Interpreting and Communicating the Results of
Statistical Analyses 15-22**

Activity 15.1 Exploring Single-Factor ANOVA 808

Graphing Calculator Exploration 811



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 (Optional) 16-1**

**16.2 Distribution-Free Procedures for Inferences About a Difference
Between Two Population or Treatment Means Using
Paired Samples 16-10**

16.3 Distribution-Free ANOVA 16-23

Appendix A: Statistical Tables 813

Appendix B: References 833

Answers to Selected Odd-Numbered Exercises 835

Index I-1