

TABLE OF CONTENTS

1.1: Why does this book exist?

1.2: Why R?

Licensing

1.3: The golden age of data

1.4: An open source book

1.5: Acknowledgements

1: Introduction

- 1.1: What Is Statistical Thinking?
- 1.2: Dealing with Statistics Anxiety
- 1.3: What Can Statistics Do for Us?
- 1.4: The Big Ideas of Statistics
- 1.5: Causality and Statistics
- 1.6: Suggested Readings

2: Working with Data

- 2.1: What Are Data?
- 2.2: Discrete Versus Continuous Measurements
- 2.3: Suggested Readings
- 2.4: Appendix
- 2.5: What Makes a Good Measurement?

3: Introduction to R

- 3.1: Why Programming Is Hard to Learn
- 3.2: Using RStudio
- 3.3: Getting Started with R
- 3.4: Variables
- 3.5: Functions
- 3.6: Vectors
- 3.7: Math with Vectors
- 3.8: Data Frames
- 3.9: Using R Libraries
- 3.10: Working with Data Files
- 3.11: Suggested Readings and Videos

4: Summarizing Data

- 4.1: Why Summarize Data?
- 4.2: Summarizing Data Using Tables
- 4.3: Idealized Representations of Distributions
- 4.4: Suggested Readings

5: Summarizing Data with R (with Lucy King)

- 5.1: Introduction to the Tidyverse
- 5.2: Creating or Modifying Variables Using Mutate()
- 5.3: Tidyverse in Action
- 5.4: Looking at Individual Variables Using Pull() and Head()
- 5.5: Computing a Frequency Distribution (Section 4.2.1)
- 5.6: Computing a Cumulative Distribution (Section 4.2.2)
- 5.7: Data Cleaning and Tidying with R

6: Data Visualization

- 6.1: Anatomy of a Plot
- 6.2: Principles of Good Visualization
- 6.3: Accommodating Human Limitations
- 6.4: Correcting for Other Factors
- 6.5: Suggested readings and videos

7: Data Visualization with R

- 7.1: The Grammar of Graphics
- 7.2: Getting Started
- 7.3: Let's Think Through a Visualization
- 7.4: Plotting the Distribution of a Single Variable
- 7.5: Plots with Two Variables
- 7.6: Creating a More Complex Plot
- 7.7: Additional Reading and Resources

8: Fitting Models to Data

- 8.1: Appendix
- 8.2: What Is a Model?
- 8.3: Statistical Modeling- An Example
- 8.4: What Makes a Model "Good"?
- 8.5: Can a Model Be Too Good?
- 8.6: The Simplest Model- The Mean
- 8.7: The Mode
- 8.8: Variability- How Well Does the Mean Fit the Data?
- 8.9: 8.8 Using Simulations to Understand Statistics
- 8.10: Z-scores

9: Fitting Simple Models with R

- 9.1: Mean
- 9.2: Median
- 9.3: Mode
- 9.4: Variability
- 9.5: Z-scores

10: Probability

- 10.1: What Do Probabilities Mean?
- 10.2: Suggested Readings
- 10.3: Appendix
- 10.4: What Is Probability?

- 10.5: How Do We Determine Probabilities?
- 10.6: Probability Distributions
- 10.7: Conditional Probability
- 10.8: Computing Conditional Probabilities from Data
- 10.9: Independence
- 10.10: Reversing a Conditional Probability- Bayes' Rule
- 10.11: Learning from Data
- 10.12: Odds and Odds Ratios

11: Probability in R

- 11.1: Basic Probability Calculations
 - 11.1.1: Empirical Frequency (Section 10.2.2)
- 11.2: Conditional Probability (Section 10.4)

12: Sampling

- 12.1: How Do We Sample?
- 12.2: Sampling Error
- 12.3: Standard Error of the Mean
- 12.4: The Central Limit Theorem
- 12.5: Confidence Intervals
- 12.6: Suggested Readings

13: Sampling in R

- 13.1: Sampling Error (Section @ref{samplingerror})
- 13.2: Central Limit Theorem
- 13.3: Confidence Intervals (Section @ref{confidence-intervals})

14: Resampling and Simulation

- 14.1: Monte Carlo Simulation
- 14.2: Randomness in Statistics
- 14.3: Generating Random Numbers
- 14.4: Using Monte Carlo Simulation
- 14.5: Using Simulation for Statistics- The Bootstrap
- 14.6: Suggested Readings

15: Resampling and Simulation in R

- 15.1: Generating Random Samples (Section @ref{generating-random-numbers})
- 15.2: Simulating the Maximum Finishing Time
- 15.3: The Bootstrap

16: Hypothesis Testing

- 16.1: Null Hypothesis Statistical Testing (NHST)
- 16.2: Null Hypothesis Statistical Testing- An Example
- 16.3: The Process of Null Hypothesis Testing
- 16.4: NHST in a Modern Context- Multiple Testing
- 16.5: Suggested Readings

17: Hypothesis Testing in R

- 17.1: Simple Example- Coin-flipping (Section 16.3.5.1)
- 17.2: Simulating p-values

18: Quantifying Effects and Designing Studies

- 18.1: Confidence Intervals
- 18.2: Effect Sizes
- 18.3: Statistical Power
- 18.4: Suggested Readings

19: Statistical Power in R

- 19.1: Power Analysis
- 19.2: Power Curves
- 19.3: Simulating Statistical Power

20: Bayesian Statistics

- 20.1: Generative Models
- 20.2: Bayes' Theorem and Inverse Inference
- 20.3: Doing Bayesian Estimation
- 20.4: Estimating Posterior Distributions
- 20.5: Choosing a Prior
- 20.6: Bayesian Hypothesis Testing
- 20.7: Suggested Readings
- 20.8: Appendix-

21: Bayesian Statistics in R

- 21.1: A Simple Example (Section 20.3)
- 21.2: Estimating Posterior Distributions (Section 20.4)
- 21.3: Bayes Factors (Section 20.6.1)

22: Modeling Categorical Relationships

- 22.1: Example- Candy Colors
- 22.2: Pearson's chi-squared Test
- 22.3: Contingency Tables and the Two-way Test
- 22.4: Standardized Residuals
- 22.5: Odds Ratios
- 22.6: Bayes Factor
- 22.7: Categorical Analysis Beyond the 2 X 2 Table
- 22.8: Beware of Simpson's Paradox
- 22.9: Additional Readings

23: Modeling Categorical Relationships in R

24: Modeling Continuous Relationships

- 24.1: An Example- Hate Crimes and Income Inequality
- 24.2: Is income Inequality Related to Hate Crimes?
- 24.3: Covariance and Correlation
- 24.4: Correlation and Causation

- 24.5: Suggested Readings
- 24.6: Appendix-

25: Modeling Continuous Relationships in R

- 25.1: Computing Covariance and Correlation (Section 24.3)
- 25.2: Hate Crime Example
- 25.3: Robust Correlations (24.3.2)

26: The General Linear Model

- 26.1: Linear Regression
- 26.2: Fitting More Complex Models
- 26.3: Interactions Between Variables
- 26.4: Beyond Linear Predictors and Outcomes
- 26.5: Criticizing Our Model and Checking Assumptions
- 26.6: What Does “Predict” Really Mean?
- 26.7: Suggested Readings
- 26.8: Appendix

27: The General Linear Model in R

- 27.1: Linear Regression (Section 26.1)
- 27.2: Model Criticism and Diagnostics (Section 26.5)
- 27.3: Examples of Problematic Model Fit
- 27.4: Extending Regression to Binary Outcomes.
- 27.5: Cross-validation (Section 26.6.1)

28: Comparing Means

- 28.1: Testing the Value of a Single Mean
- 28.2: Comparing Two Means
- 28.3: The t-test as a Linear Model
- 28.4: Bayes Factor for Mean Differences
- 28.5: Comparing Paired Observations
- 28.6: Comparing More Than Two Means
- 28.7: Appendix

29: Comparing Means in R

- 29.1: Testing the Value of a Single Mean (Section 28.1)
- 29.2: Comparing Two Means (Section 28.2)
- 29.3: The t-test as a Linear Model (Section 28.3)
- 29.4: Comparing Paired Observations (Section 28.5)
- 29.5: Analysis of Variance (Section 28.6.1)

30: Practical statistical modeling

- 30.1: The Process of Statistical Modeling

31: Practical Statistical Modeling in R

32: Doing Reproducible Research

- [32.1: How We Think Science Should Work](#)
- [32.2: How Science \(Sometimes\) Actually Works](#)
- [32.3: The Reproducibility Crisis in Science](#)
- [32.4: Questionable Research Practices](#)
- [32.5: Doing Reproducible Research](#)
- [32.6: Doing Reproducible Data Analysis](#)
- [32.7: Conclusion- Doing Better Science](#)
- [32.8: Suggested Readings](#)

33: References

[Index](#)

[Index](#)

[Glossary](#)

[Detailed Licensing](#)