CONTENTS

Acknowledgments ix

1. Introduction 1
- 1.1 Can You Repeat That Please? 2
- 1.2 Simulation and Resampling Methods 4
  - 1.2.1 Simulations as Experiments 4
  - 1.2.2 Simulations Help Develop Intuition 5
  - 1.2.3 An Overview of Simulation 6
  - 1.2.4 Resampling Methods as Simulation 7
- 1.3 OLS as a Motivating Example 8
- 1.4 Two Brief Examples 12
  - 1.4.1 Example 1: A Statistical Simulation 13
  - 1.4.2 Example 2: A Substantive Theory Simulation 15
- 1.5 Looking Ahead 15
  - 1.5.1 Assumed Knowledge 16
  - 1.5.2 A Preview of the Book 16
- 1.6 R Packages 17

2. Probability 19
- 2.1 Introduction 19
- 2.2 Some Basic Rules of Probability 20
  - 2.2.1 Introduction to Set Theory 20
  - 2.2.2 Properties of Probability 22
  - 2.2.3 Conditional Probability 22
  - 2.2.4 Simple Math With Probabilities 23
- 2.3 Random Variables and Probability Distributions 24
- 2.4 Discrete Random Variables 29
  - 2.4.1 Some Common Discrete Distributions 30
- 2.5 Continuous Random Variables 33
  - 2.5.1 Two Common Continuous Distributions 36
  - 2.5.2 Other Continuous Distributions 39
- 2.6 Conclusions 43

3. Introduction to R 45
- 3.1 Introduction 45
- 3.2 What Is R? 45
  - 3.2.1 Resources 46
- 3.3 Using R With a Text Editor 46

3.4 First Steps 47
3.4.1 Creating Objects 47
3.5 Basic Manipulation of Objects 48
3.5.1 Vectors and Sequences 48
3.5.2 Matrices 49
3.6 Functions 50
3.6.1 Matrix Algebra Functions 51
3.6.2 Creating New Functions 51
3.7 Working With Data 52
3.7.1 Loading Data 52
3.7.2 Exploring the Data 53
3.7.3 Statistical Models 54
3.7.4 Generalized Linear Models 57
3.8 Basic Graphics 59
3.9 Conclusions 61

4. Random Number Generation 63
4.1 Introduction 63
4.2 Probability Distributions 63
4.2.1 Drawing Random Numbers 65
4.2.2 Creating Your Own Distribution Functions 67
4.3 Systematic and Stochastic 68
4.3.1 The Systematic Component 69
4.3.2 The Stochastic Component 70
4.3.3 Repeating the Process 71
4.4 Programming in R 72
4.4.1 `for` Loops 73
4.4.2 Efficient Programming 74
4.4.3 If-Else 76
4.5 Completing the OLS Simulation 77
4.5.1 Anatomy of a Script File 80

5. Statistical Simulation of the Linear Model 83
5.1 Introduction 83
5.2 Evaluating Statistical Estimators 84
5.2.1 Bias, Efficiency, and Consistency 84
5.2.2 Measuring Estimator Performance in R 87
5.3 Simulations as Experiments 96
5.3.1 Heteroskedasticity 96
5.3.2 Multicollinearity 103
5.3.3 Measurement Error 105
5.3.4 Omitted Variable 109
5.3.5 Serial Correlation 112
5.3.6 Clustered Data 114
5.3.7 Heavy-Tailed Errors 118
5.4 Conclusions 125

6. Simulating Generalized Linear Models **127**
6.1 Introduction 127
6.2 Simulating OLS as a Probability Model 128
6.3 Simulating GLMs 130
6.3.1 Binary Models 130
6.3.2 Ordered Models 135
6.3.3 Multinomial Models 141
6.4 Extended Examples 145
6.4.1 Ordered or Multinomial? 145
6.4.2 Count Models 150
6.4.3 Duration Models 157
6.5 Computational Issues for Simulations 162
6.5.1 Research Computing 162
6.5.2 Parallel Processing 163
6.6 Conclusions 167

7. Testing Theory Using Simulation **169**
7.1 Introduction 169
7.2 What Is a Theory? 169
7.3 Zipf's Law 171
7.3.1 Testing Zipf's Law With *Frankenstein* 171
7.3.2 From Patterns to Explanations 174
7.4 Punctuated Equilibrium and Policy Responsiveness 181
7.4.1 Testing Punctuated Equilibrium Theory 183
7.4.2 From Patterns to Explanations 185
7.5 Dynamic Learning 190
7.5.1 Reward and Punishment 193
7.5.2 Damned If You Do, Damned If You Don't 195
7.5.3 The Midas Touch 197
7.6 Conclusions 200

8. Resampling Methods **201**
8.1 Introduction 201
8.2 Permutation and Randomization Tests 202
8.2.1 A Basic Permutation Test 203
8.2.2 Randomization Tests 205
8.2.3 Permutation/Randomization and Multiple Regression Models 208
8.3 Jackknifing 209
8.3.1 An Example 210
8.3.2 An Application: Simulating Heteroskedasticity 213
8.3.3 Pros and Cons of Jackknifing 214
8.4 Bootstrapping 215
8.4.1 Bootstrapping Basics 217
8.4.2 Bootstrapping With Multiple Regression Models 220
8.4.3 Adding Complexity: Clustered Bootstrapping 225
8.5 Conclusions 228

9. Other Simulation-Based Methods **231**
9.1 Introduction 231
9.2 QI Simulation 232
9.2.1 Statistical Overview 232
9.2.2 Examples 235
9.2.3 Simulating QI With `Zelig` 245
9.2.4 Average Case Versus Observed Values 249
9.2.5 The Benefits of QI Simulation 254
9.3 Cross-Validation 255
9.3.1 How CV Can Help 257
9.3.2 An Example 258
9.3.3 Using `R` Functions for CV 266
9.4 Conclusions 267

10. Final Thoughts **269**
10.1 A Summary of the Book 270
10.2 Going Forward 271
10.3 Conclusions 272

References **275**

Index **283**