CONTENTS

Online Resources		xi
About the Author		xiii
Acknowledgements		xv
Preface		xvii
1	Introduction and R Instructions	1
	Basic Terminology	2
	Data: Qualitative or Quantitative	5
	Data: Cross-Sectional or Longitudinal	7
	Descriptive Statistics	9
	Probability	9
	Statistics: Estimation and Inference	10
	Summary	11
	Appendix: An Introduction to R	14
	Our Approach to Writing This Book	26
2	Descriptive Statistics: Tabular and Graphical Methods	27
	Methods of Presenting Qualitative Data	28
	Methods of Presenting Quantitative Data	33
	Cross-tabulations and Scatter Plots	36
	Summary	44
	Appendix: Basic Data Manipulation	47
3	Descriptive Statistics: Numerical Methods	55
	Measures of Central Tendency	56
	Measures of Location	59
	Exploratory Data Analysis: The Box Plot Display	60
	Measures of Variability	62
	The z-Score: A Measure of Relative Location	65
	Measures of Association: The Bivariate Case	69
	The Geometric Mean	74
	Summary	77
	Appendix: Vectorization	80
4	Introduction to Probability	83
	Some Important Definitions	84
	Counting Rules	85

Assigning Probabilities 89
Events and Probabilities 91
Probabilities of Unions and Intersections of Events 94
Conditional Probability 98
Bayes' Theorem 102
Summary 104

5 Discrete Probability Distributions 109

Some Important Definitions 110
The Discrete Uniform Probability Distribution 112
The Expected Value and Standard Deviation of a Discrete Random Variable 114
The Binomial Probability Distribution 115
The Poisson Probability Distribution 121
The Hypergeometric Probability Distribution 126
The Hypergeometric Probability Distribution: The General Case 129
Summary 131
Appendix: Using R to Find Probabilities—Binomial, Poisson, and Hypergeometric 134

6 Continuous Probability Distributions 137

Continuous Uniform Probability Distribution 138
Normal Probability Distribution 142
Exponential Probability Distribution 152
Optional Material: A Derivation of the Cumulative Exponential Probability Function 157
Summary 158
Appendix: Using R to Find Probabilities for Continuous Probability Distributions—Uniform, Standard Normal, General Normal, and Exponential 160

7 Point Estimation and Sampling Distributions 165

Populations and Samples 167
The Simple Random Sample 168
The Sample Statistic: $\bar{x}$, s, and $\bar{p}$ 171
The Sampling Distribution of $\bar{x}$ 171
The Sampling Distribution of $\bar{p}$ 178
Some Other Commonly Used Sampling Methods 182
Summary 185
Appendix: Using R to Select a Random Sample from a Data Set 188

8 Confidence Interval Estimation 191

Interval Estimate of μ When σ is Known 193
Interval Estimate of μ When σ is Unknown 198
Sample Size Determination in the Case of μ 205
Interval Estimate of p 206

	Sample Size Determination in the Case of p	208
	Summary	210
9	Hypothesis Tests: Introduction, Basic Concepts, and an Example	215
	Summary	226
10	Hypothesis Tests about Means and Proportions: Applications	231
	The Lower-Tail Hypothesis Test about μ: σ is Known	232
	The Two-Tail Hypothesis Test about μ: σ is Known	235
	The Upper-Tail Hypothesis Test about μ: σ is Unknown	237
	The Two-Tail Hypothesis Test about μ: σ is Unknown	240
	Hypothesis Tests about p	242
	Calculating the Probability of a Type II Error: β	247
	Adjusting the Sample Size to Control the Size of β	250
	Summary	252
11	Comparisons of Means and Proportions	257
	The Difference between μ_1 and μ_2: Independent Samples	258
	The Difference between μ_1 and μ_2: Paired Samples	264
	The Difference between p_1 and p_2: Independent Samples	271
	Summary	276
12	Simple Linear Regression	283
	Simple Linear Regression: The Model	284
	The Estimated Regression Equation	285
	Goodness of Fit: The Coefficient of Determination, r^2	291
	The Hypothesis Test about β_1	296
	Alternative Approaches to Testing Significance	302
	So Far, We Have Tested Only b_1; Will We Also Test b_0?	303
	Assumptions: What Are They?	304
	Assumptions: How Are They Validated?	305
	Optional Material: Derivation of the Expressions for the Least Square Estimates of β_0 and β_1	309
	Summary	311
13	Multiple Regression	319
	Simple Linear Regression: A Reprise	320
	Multiple Regression: The Model	324
	Multiple Regression: The Multiple Regression Equation	325
	The Estimated Multiple Regression Equation	325
	Multiple Regression: Two Independent Variables	326
	Assumptions: What Are They? Can We Validate Them?	335
	Tests of Significance: The Overall Regression Model	336

Tests of Significance: The Independent Variables 340
There Must Be An Easier Way Than This, Right? 344
Using the Estimated Regression Equation for Prediction 345
Independent Variable Selection: The Best-Subsets Method 346
Summary 353

Index 363