

Contents

List of Figures	xv
List of Tables	xix
Preface	xxiii
1 Experiments	1
1.1 Examples	1
1.1.1 Spinning a Penny	2
1.1.2 The Speed of Light	3
1.1.3 Termite Foraging Behavior	5
1.2 Randomization	7
1.3 The Importance of Probability	12
1.4 Games of Chance	14
1.5 Exercises	19
2 Mathematical Preliminaries	23
2.1 Sets	23
2.2 Counting	27
2.3 Functions	34
2.4 Limits	35
2.5 Exercises	36
3 Probability	43
3.1 Interpretations of Probability	43
3.2 Axioms of Probability	44
3.3 Finite Sample Spaces	52
3.4 Conditional Probability	58
3.5 Random Variables	69

3.6	Case Study: Padrolling in Milton Murayama’s <i>All I asking for is my body</i>	77
3.7	Exercises	82
4	Discrete Random Variables	89
4.1	Basic Concepts	89
4.2	Examples	90
4.3	Expectation	94
4.4	Binomial Distributions	105
4.5	Exercises	110
5	Continuous Random Variables	117
5.1	A Motivating Example	117
5.2	Basic Concepts	120
5.3	Elementary Examples	124
5.4	Normal Distributions	128
5.5	Normal Sampling Distributions	132
5.6	Exercises	136
6	Quantifying Population Attributes	141
6.1	Symmetry	141
6.2	Quantiles	143
6.2.1	The Median of a Population	146
6.2.2	The Interquartile Range of a Population	147
6.3	The Method of Least Squares	148
6.3.1	The Mean of a Population	148
6.3.2	The Standard Deviation of a Population	149
6.4	Exercises	150
7	Data	153
7.1	The Plug-In Principle	154
7.2	Plug-In Estimates of Mean and Variance	156
7.3	Plug-In Estimates of Quantiles	158
7.3.1	Box Plots	160
7.3.2	Normal Probability Plots	163
7.4	Kernel Density Estimates	164
7.5	Case Study: Forearm Lengths	167
7.6	Transformations	173
7.7	Exercises	175

8	Lots of Data	181
8.1	Averaging Decreases Variation	183
8.2	The Weak Law of Large Numbers	185
8.3	The Central Limit Theorem	187
8.4	Exercises	194
9	Inference	197
9.1	A Motivating Example	198
9.2	Point Estimation	200
9.2.1	Estimating a Population Mean	200
9.2.2	Estimating a Population Variance	201
9.3	Heuristics of Hypothesis Testing	202
9.4	Testing Hypotheses about a Population Mean	212
9.4.1	One-Sided Hypotheses	216
9.4.2	Formulating Suitable Hypotheses	217
9.4.3	Statistical Significance and Material Significance . . .	221
9.5	Set Estimation	222
9.5.1	Sample Size	225
9.5.2	One-Sided Confidence Intervals	226
9.6	Exercises	227
10	1-Sample Location Problems	233
10.1	The Normal 1-Sample Location Problem	236
10.1.1	Point Estimation	236
10.1.2	Hypothesis Testing	237
10.1.3	Set Estimation	241
10.2	The General 1-Sample Location Problem	243
10.2.1	Hypothesis Testing	243
10.2.2	Point Estimation	246
10.2.3	Set Estimation	247
10.3	The Symmetric 1-Sample Location Problem	248
10.3.1	Hypothesis Testing	249
10.3.2	Point Estimation	254
10.3.3	Set Estimation	256
10.4	Case Study: Deficit Unawareness	257
10.5	Exercises	262
11	2-Sample Location Problems	269
11.1	The Normal 2-Sample Location Problem	271
11.1.1	Known Variances	274

11.1.2	Unknown Common Variance	275
11.1.3	Unknown Variances	278
11.2	The Case of a General Shift Family	282
11.2.1	Hypothesis Testing	283
11.2.2	Point Estimation	287
11.2.3	Set Estimation	289
11.3	Case Study: Etruscan versus Italian Head Breadth	290
11.4	Exercises	294
12	The Analysis of Variance	305
12.1	The Fundamental Null Hypothesis	306
12.2	Testing the Fundamental Null Hypothesis	307
12.2.1	Known Population Variance	308
12.2.2	Unknown Population Variance	309
12.3	Planned Comparisons	313
12.3.1	Orthogonal Contrasts	318
12.3.2	Bonferroni t -Tests	321
12.4	Post Hoc Comparisons	323
12.4.1	Bonferroni t -Tests	323
12.4.2	Scheffé F -Tests	324
12.5	Case Study: Treatments of Anorexia	325
12.6	Exercises	328
13	Goodness-of-Fit	337
13.1	Partitions	337
13.2	Test Statistics	338
13.3	Testing Independence	341
13.4	Exercises	344
14	Association	349
14.1	Bivariate Distributions	350
14.2	Normal Random Variables	351
14.2.1	Bivariate Normal Samples	353
14.2.2	Inferences about Correlation	357
14.3	Monotonic Association	362
14.4	Explaining Association	368
14.5	Case Study: Anorexia Treatments Revisited	369
14.6	Exercises	372

5 Simple Linear Regression	379
15.1 The Regression Line	380
15.2 The Method of Least Squares	385
15.3 Computation	393
15.4 The Simple Linear Regression Model	395
15.5 Assessing Linearity	400
15.6 Case Study: Are Thick Books More Valuable?	406
15.7 Exercises	408
6 Simulation-Based Inference	417
16.1 Termite Foraging Revisited	418
16.2 The Bootstrap	423
16.3 Case Study: Adventure Racing	427
16.4 Exercises	431
R A Statistical Programming Language	435
R.1 Introduction	435
R.1.1 What Is R?	435
R.1.2 Why Use R?	435
R.1.3 Installing R	436
R.1.4 Learning about R	437
R.2 Using R	437
R.2.1 Vectors	438
R.2.2 R Is a Calculator!	440
R.2.3 Some Statistics Functions	440
R.2.4 Matrices	440
R.2.5 Creating New Functions	443
R.3 Functions That Accompany This Book	445
R.3.1 Inferences about a Center of Symmetry	446
R.3.2 Inferences about a Shift Parameter	449
R.3.3 Inferences about Monotonic Association	452
R.3.4 Exploring Bivariate Normal Data	455
R.3.5 Simulating Random Termite Foraging	459
Index	463