Contents

Preface to the Third Edition v

Preface to the Second Edition vii

Preface to the First Edition ix

1 A Wide Range of Applications 1
- 1.1 Basic Concepts 1
  - 1.1.1 Stochastic Phenomena 1
  - 1.1.2 Distribution Functions 3
  - 1.1.3 Hypotheses 3
- 1.2 Applications 5
- 1.3 Testing a Hypothesis 7
  - 1.3.1 Five Steps to a Test 8
  - 1.3.2 Analyze the Experiment 8
  - 1.3.3 Choose a Test Statistic 8
  - 1.3.4 Compute the Test Statistic 9
  - 1.3.5 Determine the Frequency Distribution of the Test Statistic 9
  - 1.3.6 Make a Decision 10
  - 1.3.7 Variations on a Theme 10
- 1.4 A Brief History of Statistics in Decision-Making 10
- 1.5 Exercises 12

2 Optimal Procedures 13
- 2.1 Defining Optimal 13
  - 2.1.1 Trustworthy 13
  - 2.1.2 Two Types of Error 14
  - 2.1.3 Losses and Risk 16
  - 2.1.4 Significance Level and Power 17

2.1.4.1 Power and the Magnitude of the Effect ... 18
2.1.4.2 Power and Sample Size ... 18
2.1.4.3 Power and the Alternative ... 20
2.1.5 Exact, Unbiased, Conservative ... 21
2.1.6 Impartial ... 22
2.1.7 Most Stringent Tests ... 23
2.2 Basic Assumptions ... 23
2.2.1 Independent Observations ... 23
2.2.2 Exchangeable Observations ... 24
2.3 Decision Theory ... 25
2.3.1 Bayes' Risk ... 26
2.3.2 Mini-Max ... 27
2.3.3 Generalized Decisions ... 28
2.4 Exercises ... 29

3 Testing Hypotheses ... 33
3.1 Testing a Simple Hypothesis ... 33
3.2 One-Sample Tests for a Location Parameter ... 34
3.2.1 A Permutation Test ... 34
3.2.2 A Parametric Test ... 36
3.2.3 Properties of the Parametric Test ... 37
3.2.4 Student's t ... 38
3.2.5 Properties of the Permutation Test ... 39
3.2.6 Exact Significance Levels: A Digression ... 39
3.3 Confidence Intervals ... 40
3.3.1 Confidence Intervals Based on Permutation Tests ... 41
3.3.2 Confidence Intervals Based on Parametric Tests ... 42
3.3.3 Confidence Intervals Based on the Bootstrap ... 43
3.3.4 Parametric Bootstrap ... 45
3.3.5 Better Confidence Intervals ... 46
3.4 Comparison Among the Test Procedures ... 46
3.5 One-Sample Tests for a Scale Parameter ... 48
3.5.1 Semiparametric Tests ... 48
3.5.2 Parametric Tests: Sufficiency ... 48
3.5.3 Unbiased Tests ... 50
3.5.4 Comparison Among the Test Procedures ... 50
3.6 Comparing the Location Parameters of Two Populations ... 51
3.6.1 A UMPU Parametric Test: Student's t ... 51
3.6.2 A UMPU Semiparametric Procedure ... 51
3.6.3 An Example ... 54
3.6.4 Comparison of the Tests: The Behrens–Fisher Problem ... 54

3.7 Comparing the Dispersions of Two Populations 57
3.7.1 The Parametric Approach 57
3.7.2 The Permutation Approach 58
3.7.3 The Bootstrap Approach 61
3.8 Bivariate Correlation 62
3.9 Which Test? 63
3.10 Exercises 63

4 Distributions 67
4.1 Properties of Independent Observations 67
4.2 Binomial Distribution 67
4.3 Poisson: Events Rare in Time and Space 68
4.3.1 Applying the Poisson 69
4.3.2 A Poisson Distribution of Poisson Distributions 70
4.3.3 Comparing Two Poissons 70
4.4 Time Between Events 71
4.5 The Uniform Distribution 71
4.6 The Exponential Family of Distributions 72
4.6.1 Proofs of the Properties 73
4.6.2 Normal Distribution 74
4.7 Which Distribution? 75
4.8 Exercises 75

5 Multiple Tests 79
5.1 Controlling the Overall Error Rate 79
5.1.1 Standardized Statistics 80
5.1.2 Paired Sample Tests 81
5.2 Combination of Independent Tests 81
5.2.1 Omnibus Statistics 82
5.2.2 Binomial Random Variables 82
5.2.3 Bayes' Factor 83
5.3 Exercises 84

6 Experimental Designs 85
6.1 Invariance 85
6.1.1 Some Examples 86
6.2 k-Sample Comparisons—Least-Squares Loss Function 87
6.2.1 Linear Hypotheses 87
6.2.2 Large and Small Sample Properties of the F-ratio Test 89
6.2.3 Discrete Data and Time-to-Event Data 90
6.3 k-Sample Comparisons—Other Loss Functions 91
6.3.1 F-ratio 91
6.3.2 Pitman Correlation 92

6.3.3 Effect of Ties 95
6.3.4 Cochran–Armitage Test 96
6.3.5 Linear Estimation 96
6.3.6 A Unifying Theory 97
6.4 Four Ways to Control Variation 97
6.4.1 Control the Environment 98
6.4.2 Block the Experiment 98
6.4.2.1 Using Ranks 99
6.4.2.2 Matched Pairs 100
6.4.3 Measure Factors That Cannot Be Controlled 101
6.4.3.1 Eliminate the Functional Relationship 101
6.4.3.2 Selecting Variables 102
6.4.3.3 Restricted Randomization 102
6.4.4 Randomize 103
6.5 Latin Square 104
6.6 Very Large Samples 106
6.7 Sequential Analysis 107
6.7.1 A Vaccine Trial 107
6.7.2 Determining the Boundary Values 110
6.7.3 Power of a Sequential Analysis 110
6.7.4 Expected Sample Size 111
6.7.5 Curtailed Inspection 112
6.7.6 Restricted Sequential Sampling Schemes 112
6.8 Sequentially Adaptive Treatment Allocation 113
6.8.1 Group Sequential Trials 113
6.8.2 Determining the Sampling Ratio 113
6.8.3 Exact Random Allocation Tests 114
6.9 Exercises 115

7 Multifactor Designs 119
7.1 Multifactor Models 119
7.2 Analysis of Variance 120
7.3 Permutation Methods: Main Effects 124
7.3.1 An Example 125
7.4 Permutation Methods: Interactions 126
7.5 Synchronized Rearrangements 127
7.5.1 Exchangeable and Weakly Exchangeable Variables 128
7.5.2 Two Factors 129
7.5.3 Three or More Factors 132
7.5.4 Similarities 133
7.5.5 Test for Interaction 135
7.6 Unbalanced Designs 137
7.6.1 Missing Combinations 138
7.6.2 The Boot-Perm Test 139

7.7 Which Test Should You Use? . . . 140
7.8 Exercises . . . 140

8 Categorical Data . . . 143
8.1 Fisher's Exact Test . . . 143
8.1.1 Hypergeometric Distribution . . . 145
8.1.2 One-Tailed and Two-Tailed Tests . . . 145
8.1.3 The Two-Tailed Test . . . 146
8.1.4 Determining the p-Value . . . 146
8.1.5 What is the Alternative? . . . 148
8.1.6 Increasing the Power . . . 148
8.1.7 Ongoing Controversy . . . 149
8.2 Odds Ratio . . . 150
8.2.1 Stratified 2×2's . . . 151
8.3 Exact Significance Levels . . . 152
8.4 Unordered $r \times c$ Contingency Tables . . . 154
8.4.1 Agreement Between Observers . . . 156
8.4.2 What Should We Randomize? . . . 157
8.4.3 Underlying Assumptions . . . 158
8.4.4 Symmetric Contingency Tables . . . 158
8.5 Ordered Contingency Tables . . . 160
8.5.1 Ordered $2 \times c$ Tables . . . 160
8.5.1.1 Alternative Hypotheses . . . 161
8.5.1.2 Back-up Statistics . . . 162
8.5.1.3 Directed Chi-Square . . . 162
8.5.2 More Than Two Rows and Two Columns . . . 163
8.5.2.1 Singly Ordered Tables . . . 163
8.5.2.2 Doubly Ordered Tables . . . 163
8.6 Covariates . . . 164
8.6.1 Bross' Method . . . 164
8.6.2 Blocking . . . 165
8.7 Exercises . . . 166

9 Multivariate Analysis . . . 169
9.1 Nonparametric Combination of Univariate Tests . . . 169
9.2 Parametric Approach . . . 171
9.2.1 Canonical Form . . . 171
9.2.2 Hotelling's T^2 . . . 172
9.2.3 Multivariate Analysis of Variance (MANOVA) . . . 173
9.3 Permutation Methods . . . 173
9.3.1 Which Test—Parametric or Permutation? . . . 175
9.3.2 Interpreting the Results . . . 176

9.4 Alternative Statistics 177
9.4.1 Maximum-t 177
9.4.2 Block Effects 177
9.4.3 Runs Test 178
9.4.4 Which Statistic? 180
9.5 Repeated Measures 181
9.5.1 An Example 181
9.5.2 Matched Pairs 182
9.5.3 Response Profiles 183
9.5.4 Missing Data 183
9.5.5 Bioequivalence 184
9.6 Exercises 185

10 Clustering in Time and Space 189
10.1 The Generalized Quadratic Form 189
10.1.1 Mantel's U 189
10.1.2 An Example 189
10.2 Applications 190
10.2.1 The MRPP Statistic 190
10.2.2 The BW Statistic of Cliff and Ord [1973] 191
10.2.3 Equivalances 192
10.2.4 Extensions 192
10.2.5 Another Dimension 192
10.3 Alternate Approaches 193
10.3.1 Quadrant Density 193
10.3.2 Nearest-Neighbor Analysis 193
10.3.3 Comparing Two Spatial Distributions 193
10.4 Exercises 194

11 Coping with Disaster 195
11.1 Missing Data 195
11.2 Covariates After the Fact 197
11.2.1 Observational Studies 197
11.3 Outliers 198
11.3.1 Original Data 199
11.3.2 Ranks 199
11.3.3 Scores 200
11.3.4 Robust Transformations 201
11.3.5 Use an L_1 Test 201
11.3.6 Censoring 201
11.3.7 Discarding 202
11.4 Censored Data 202
11.4.1 GAMP Tests 202
11.4.2 Fishery and Animal Counts 204

11.5 Censored Match Pairs . . . 204
11.5.1 GAMP Test for Matched Pairs . . . 205
11.5.2 Ranks . . . 206
11.5.3 One-Sample: Bootstrap Estimates . . . 206
11.6 Adaptive Tests . . . 207
11.7 Exercises . . . 208

12 Solving the Unsolved and the Insolvable . . . 209
12.1 Key Criteria . . . 209
12.1.1 Sufficient Statistics . . . 209
12.1.2 Three Stratagems . . . 210
12.1.3 Restrict the Alternatives . . . 210
12.1.4 Consider the Loss Function . . . 212
12.1.5 Impartiality . . . 213
12.2 The Permutation Distribution . . . 213
12.2.1 Ensuring Exchangeability . . . 213
12.2.1.1 Test for Parallelism . . . 214
12.2.1.2 Linear Transforms That Preserve Exchangeability . . . 215
12.3 New Statistics . . . 216
12.3.1 Nonresponders . . . 216
12.3.1.1 Extension to K-samples . . . 217
12.3.2 Animal Movement . . . 217
12.3.3 The Building Blocks of Life . . . 217
12.3.4 Structured Exploratory Data Analysis . . . 218
12.3.5 Comparing Multiple Methods of Assessment . . . 219
12.4 Model Validation . . . 221
12.4.1 Regression Models . . . 221
12.4.1.1 Via the Bootstrap . . . 221
12.4.1.2 Via Permutation Tests . . . 221
12.4.2 Models With a Metric . . . 222
12.5 Bootstrap Confidence Intervals . . . 223
12.5.1 Hall–Wilson Criteria . . . 224
12.5.2 Bias-Corrected Percentile . . . 225
12.6 Exercises . . . 226

13 Publishing Your Results . . . 229
13.1 Design Methodology . . . 229
13.1.1 Randomization in Assignment . . . 229
13.1.2 Choosing the Experimental Unit . . . 230
13.1.3 Determining Sample Size . . . 231
13.1.4 Power Comparisons . . . 231
13.2 Preparing Manuscripts for Publication . . . 231
13.2.1 Reportable Elements . . . 232
13.2.2 Details of the Analysis . . . 232

14 Increasing Computational Efficiency 233
14.1 Seven Techniques 233
14.2 Monte Carlo 233
14.2.1 Stopping Rules 234
14.2.2 Variance of the Result 235
14.2.3 Cutting the Computation Time 235
14.3 Rapid Enumeration and Selection Algorithms 236
14.3.1 Matched Pairs 236
14.4 Recursive Relationships 236
14.5 Focus on the Tails 237
14.5.1 Contingency Tables 239
14.5.1.1 Network Representation 239
14.5.1.2 The Network Algorithm 241
14.5.2 Play the Winner Allocation 242
14.5.3 Directed Vertex Peeling 242
14.6 Gibbs Sampling 243
14.6.1 Metropolis–Hastings Sampling Methods 244
14.7 Characteristic Functions 245
14.8 Asymptotic Approximations 246
14.8.1 A Central Limit Theorem 246
14.8.2 Edgeworth Expansions 246
14.8.3 Generalized Correlation 247
14.9 Confidence Intervals 247
14.10 Sample Size and Power 248
14.10.1 Simulations 248
14.10.2 Network Algorithms 249
14.11 Some Conclusions 250
14.12 Software 251
14.12.1 Do-It-Yourself 251
14.12.2 Complete Packages 252
14.12.2.1 Freeware 252
14.12.2.2 Shareware 252
14.12.2.3 $$$$ 252
14.13 Exercises 253

Appendix: Theory of Testing Hypotheses 255
A.1 Probability 255
A.2 The Fundamental Lemma 257
A.3 Two-Sided Tests 258
A.3.1 One-Parameter Exponential Families 259
A.4 Tests for Multiparameter Families 262
A.4.1 Basu's Theorem 262
A.4.2 Conditional Probability and Expectation 263
A.4.3 Multiparameter Exponential Families 263

A.5 Exchangeable Observations 268
A.5.1 Order Statistics 269
A.5.2 Transformably Exchangeable 270
A.5.3 Exchangeability-Preserving Transforms 271
A.6 Confidence Intervals 272
A.7 Asymptotic Behavior 273
A.7.1 A Theorem on Linear Forms 273
A.7.2 Monte Carlo 274
A.7.3 Asymptotic Efficiency 274
A.7.4 Exchangeability 275
A.7.5 Improved Bootstrap Confidence Intervals 276
A.8 Exercises 276

Bibliography 279

Author Index 303

Subject Index 309