Table of Contents

Preface to the First Edition xiii
Preface to the Second Edition xvii
About the Companion Website xix

1 Descriptive Methods for Categorical Data 1
1.1 Proportions 1
1.1.1 Comparative Studies 2
1.1.2 Screening Tests 5
1.1.3 Displaying Proportions 7
1.2 Rates 10
1.2.1 Changes 11
1.2.2 Measures of Morbidity and Mortality 13
1.2.3 Standardization of Rates 15
1.3 Ratios 18
1.3.1 Relative Risk 18
1.3.2 Odds and Odds Ratio 18
1.3.3 Generalized Odds for Ordered 2 × k Tables 21
1.3.4 Mantel–Haenszel Method 25
1.3.5 Standardized Mortality Ratio 28
1.4 Notes on Computations 30
Exercises 32

2 Descriptive Methods for Continuous Data 55
2.1 Tabular and Graphical Methods 55
2.1.1 One]way Scatter Plots 55
2.1.2 Frequency Distribution 56
2.1.3 Histogram and Frequency Polygon 60
2.1.4 Cumulative Frequency Graph and Percentiles 64
2.1.5 Stem and Leaf Diagrams 68
2.2 Numerical Methods 69
2.2.1 Mean 69
2.2.2 Other Measures of Location 72
2.2.3 Measures of Dispersion 73
2.2.4 Box Plots 76
2.3 Special Case of Binary Data 77
2.4 Coefficients of Correlation 78
2.4.1 Pearson’s Correlation Coefficient 80
2.4.2 Nonparametric Correlation Coefficients 83
2.5 Notes on Computations 85
Exercises 87

3 Probability and Probability Models 103
3.1 Probability 103
3.1.1 Certainty of Uncertainty 104
3.1.2 Probability 104
3.1.3 Statistical Relationship 106
3.1.4 Using Screening Tests 109
3.1.5 Measuring Agreement 112
3.2 Normal Distribution 114
3.2.1 Shape of the Normal Curve 114

3.2.2 Areas Under the Standard Normal Curve 116
3.2.3 Normal Distribution as a Probability Model 122
3.3 Probability Models for Continuous Data 124
3.4 Probability Models for Discrete Data 125
3.4.1 Binomial Distribution 126
3.4.2 Poisson Distribution 128
3.5 Brief Notes on the Fundamentals 130
3.5.1 Mean and Variance 130
3.5.2 Pair]Matched Case–Control Study 130
3.6 Notes on Computations 132
Exercises 134

4 Estimation of Parameters 141
4.1 Basic Concepts 142
4.1.1 Statistics as Variables 143
4.1.2 Sampling Distributions 143
4.1.3 Introduction to Confidence Estimation 145
4.2 Estimation of Means 146
4.2.1 Confidence Intervals for a Mean 147
4.2.2 Uses of Small Samples 149
4.2.3 Evaluation of Interventions 151
4.3 Estimation of Proportions 153
4.4 Estimation of Odds Ratios 157
4.5 Estimation of Correlation Coefficients 160
4.6 Brief Notes on the Fundamentals 163
4.7 Notes on Computations 165
Exercises 166

5 Introduction to Statistical tests of Significance 179
5.1 Basic Concepts 180
5.1.1 Hypothesis Tests 181
5.1.2 Statistical Evidence 182
5.1.3 Errors 182
5.2 Analogies 185
5.2.1 Trials by Jury 185
5.2.2 Medical Screening Tests 186
5.2.3 Common Expectations 186
5.3 Summaries and Conclusions 187
5.3.1 Rejection Region 187
5.3.2 p Values 189
5.3.3 Relationship to Confidence Intervals 191
5.4 Brief Notes on the Fundamentals 193
5.4.1 Type I and Type II Errors 193
5.4.2 More about Errors and p Values 194
Exercises 194

6 Comparison of Population Proportions 197
6.1 One]Sample Problem with Binary Data 197
6.2 Analysis of Pair]Matched Data 199
6.3 Comparison of Two Proportions 202
6.4 Mantel–Haenszel Method 206
6.5 Inferences for General Two]Way Tables 211
6.6 Fisher’s Exact Test 217

6.7 Ordered 2 × K Contingency Tables 219
6.8 Notes on Computations 222
Exercises 222

7 Comparison of Population Means 235
7.1 One]Sample Problem with Continuous Data 235
7.2 Analysis of Pair]Matched Data 237
7.3 Comparison of Two Means 242
7.4 Nonparametric Methods 246
7.4.1 Wilcoxon Rank]Sum Test 246
7.4.2 Wilcoxon Signed]Rank Test 250
7.5 One]Way Analysis of Variance 252
7.5.1 One]way Analysis of Variance Model 253
7.5.2 Group Comparisons 258
7.6 Brief Notes on the Fundamentals 259
7.7 Notes on Computations 260
Exercises 260

8 Analysis of Variance 273
8.1 Factorial Studies 273
8.1.1 Two Crossed Factors 273
8.1.2 Extensions to More Than Two Factors 278
8.2 Block Designs 280
8.2.1 Purpose 280
8.2.2 Fixed Block Designs 281
8.2.3 Random Block Designs 284
8.3 Diagnostics 287
Exercises 291

9 Regression Analysis 297
9.1 Simple Regression Analysis 298
9.1.1 Correlation and Regression 298
9.1.2 Simple Linear Regression Model 301
9.1.3 Scatter Diagram 302
9.1.4 Meaning of Regression Parameters 302
9.1.5 Estimation of Parameters and Prediction 303
9.1.6 Testing for Independence 307
9.1.7 Analysis of Variance Approach 309
9.1.8 Some Biomedical Applications 311
9.2 Multiple Regression Analysis 317
9.2.1 Regression Model with Several Independent Variables 318
9.2.2 Meaning of Regression Parameters 318
9.2.3 Effect Modifications 319
9.2.4 Polynomial Regression 319
9.2.5 Estimation of Parameters and Prediction 320
9.2.6 Analysis of Variance Approach 321
9.2.7 Testing Hypotheses in Multiple Linear Regression 322
9.2.8 Some Biomedical Applications 330
9.3 Graphical and Computational Aids 334
Exercises 336

10 Logistic Regression 351
10.1 Simple Regression Analysis 353

10.1.1 Simple Logistic Regression Model 353
10.1.2 Measure of Association 355
10.1.3 Effect of Measurement Scale 356
10.1.4 Tests of Association 358
10.1.5 Use of the Logistic Model for Different Designs 358
10.1.6 Overdispersion 359
10.2 Multiple Regression Analysis 362
10.2.1 Logistic Regression Model with Several Covariates 363
10.2.2 Effect Modifications 364
10.2.3 Polynomial Regression 365
10.2.4 Testing Hypotheses in Multiple Logistic Regression 365
10.2.5 Receiver Operating Characteristic Curve 372
10.2.6 ROC Curve and Logistic Regression 374
10.3 Brief Notes on the Fundamentals 375
10.4 Notes on Computing 377
Exercises 377

11 Methods for Count Data 383
11.1 Poisson Distribution 383
11.2 Testing Goodness of Fit 387
11.3 Poisson Regression Model 389
11.3.1 Simple Regression Analysis 389
11.3.2 Multiple Regression Analysis 393
11.3.3 Overdispersion 402
11.3.4 Stepwise Regression 404
Exercises 406

12 Methods for Repeatedly Measured Responses 409
12.1 Extending Regression Methods Beyond Independent Data 409
12.2 Continuous Responses 410
12.2.1 Extending Regression using the Linear Mixed Model 410
12.2.2 Testing and Inference 414
12.2.3 Comparing Models 417
12.2.4 Special Cases: Random Block Designs and Multi]level Sampling 418
12.3 Binary Responses 423
12.3.1 Extending Logistic Regression using Generalized Estimating Equations 423
12.3.2 Testing and Inference 425
12.4 Count Responses 427
12.4.1 Extending Poisson Regression using Generalized Estimating Equations 427
12.4.2 Testing and Inference 428
12.5 Computational Notes 431
Exercises 432

13 Analysis of Survival Data and Data from Matched Studies 439
13.1 Survival Data 440
13.2 Introductory Survival Analyses 443
13.2.1 Kaplan–Meier Curve 444
13.2.2 Comparison of Survival Distributions 446
13.3 Simple Regression and Correlation 450
13.3.1 Model and Approach 451
13.3.2 Measures of Association 452
13.3.3 Tests of Association 455
13.4 Multiple Regression and Correlation 456

13.4.1 Proportional Hazards Model with Several Covariates 456
13.4.2 Testing Hypotheses in Multiple Regression 457
13.4.3 Time]Dependent Covariates and Applications 461
13.5 Pair]Matched Case–Control Studies 464
13.5.1 Model 465
13.5.2 Analysis 466
13.6 Multiple Matching 468
13.6.1 Conditional Approach 469
13.6.2 Estimation of the Odds Ratio 469
13.6.3 Testing for Exposure Effect 470
13.7 Conditional Logistic Regression 472
13.7.1 Simple Regression Analysis 473
13.7.2 Multiple Regression Analysis 478
Exercises 484

14 Study Designs 493
14.1 Types of Study Designs 494
14.2 Classification of Clinical Trials 495
14.3 Designing Phase I Cancer Trials 497
14.4 Sample Size Determination for Phase II Trials and Surveys 499
14.5 Sample Sizes for Other Phase II Trials 501
14.5.1 Continuous Endpoints 501
14.5.2 Correlation Endpoints 503
14.6 About Simon's Two]Stage Phase II Design 503
14.7 Phase II Designs for Selection 504
14.7.1 Continuous Endpoints 505
14.7.2 Binary Endpoints 505
14.8 Toxicity Monitoring in Phase II Trials 506
14.9 Sample Size Determination for Phase III Trials 508
14.9.1 Comparison of Two Means 509
14.9.2 Comparison of Two Proportions 511
14.9.3 Survival Time as the Endpoint 513
14.10 Sample Size Determination for Case–Control Studies 515
14.10.1 Unmatched Designs for a Binary Exposure 516
14.10.2 Matched Designs for a Binary Exposure 518
14.10.3 Unmatched Designs for a Continuous Exposure 520

Exercises 522
References 529
Appendices 535
Answers to Selected Exercises 541
Index 581