
Contents

Preface to the Fifth Edition	XI
Prefaces to the Previous Editions	XIII
1 Basic Concepts	1
1.1 Introduction	1
1.2 Random Variables, Probability Distributions and Some Standard Notation	2
1.3 Characteristics of a Distribution: Mean, Median and Variance	7
1.4 Statistical Calculations Using the Software Package Called R	9
2 Tests of Significance	11
2.1 Introduction	11
2.2 An Example	12
2.3 Common Features of Significance Tests	16
3 Fisher’s Test for 2 × 2 Contingency Tables	19
3.1 Introduction	19
3.2 Details of the Test	19
3.3 Additional Examples of Fisher’s Test	23
3.4 Sample R Code for Chapter 3	27
4 Approximate Significance Tests for Contingency Tables	29
4.1 Introduction	29
4.2 The χ^2 Test for 2 × 2 Tables	29
4.3 The χ^2 Test for Rectangular Contingency Tables	35
4.4 Sample R Code for Chapter 4	41
5 Some Warnings concerning 2 × 2 Tables	43
5.1 Introduction	43
5.2 Combining 2 × 2 Tables	43
5.3 Matched Pairs Binary Data	48
5.4 Multiple Comparisons and False Discovery Rates	50
5.5 Sample R Code for Chapter 5	52

6 Kaplan-Meier or ‘Actuarial’ Survival Curves 55

6.1 Introduction 55

6.2 General Features of the K-M Estimate 57

6.3 A Novel Use of the K-M Estimator 62

6.4 Confidence Bands for the K-M Estimator 64

6.5 Sample R Code for Chapter 6 67

7 The Log-Rank or Mantel-Haenszel Test for Comparing Survival Curves 69

7.1 Introduction 69

7.2 Details of the Test 70

7.3 Several Examples of the Log-Rank Test 72

7.4 Sample R Code for Chapter 7 75

8 An Introduction to the Normal Distribution 76

8.1 Introduction 76

8.2 Basic Features of the Normal Distribution 77

8.3 The Normal Distribution and Significance Testing 81

8.4 The Normal Distribution and Confidence Intervals 83

8.5 Sample R Code for Chapter 8 87

9 Analyzing Normally Distributed Data 89

9.1 Introduction 89

9.2 Some Preliminary Considerations 90

9.3 Analyzing a Single Sample 95

9.4 Comparisons Based on the Normal Distribution 97

9.5 Testing the Equality of Variances 102

9.6 Sample R Code for Chapter 9 104

10 Linear Regression Models for Medical Data 108

10.1 Introduction 108

10.2 A Historical Note 109

10.3 Multiple Linear Regression 110

10.4 Graphical Tools for Model Checking 115

10.5 Correlation 119

10.6 The Analysis of Variance 124

10.7 Sample R Code for Chapter 10 130

11 Binary Logistic Regression 133

11.1 Introduction 133

11.2 Logistic Regression 134

11.3 Estimation in 2×2 Tables 136

11.4 Reanalysis of a Previous Example 140

11.5 The Analysis of Dose-Response Data 141

11.6 Global Tests and a Previous Example 145

11.7 Sample R Code for Chapter 11 147

12 Regression Models for Count Data 151

12.1 Introduction 151

12.2 The Model for Poisson Regression 152

12.3 An Experimental Study of Cellular Differentiation 155

12.4 Overdispersion 157

12.5 Zero-Inflated Poisson Models 159

12.6 Sample R Code for Chapter 12 161

13 Proportional Hazards Regression 163

13.1 Introduction 163

13.2 A Statistical Model for the Death Rate 164

13.3 The Lymphoma Example 166

13.4 The Use of Time-Dependent Covariates 169

13.5 Sample R Code for Chapter 13 174

14 The Analysis of Longitudinal Data 176

14.1 Introduction 176

14.2 Liang-Zeger Regression Models 176

14.3 Random Effects Models 180

14.4 Multi-State Models 183

14.5 Sample R Code for Chapter 14 194

15 Analysis of Variance 196

15.1 Introduction 196

15.2 Representing Categorical Information in Regression Models 197

15.3 Understanding Two-Factor Interactions 202

15.4 Revisiting the INR Study 204

15.5 Sample R Code for Chapter 15 210

16 Data Analysis 212

16.1 Introduction 212

16.2 Quality Data 212

16.3 Initial or Exploratory Analysis 214

16.4 Primary Analysis 216

16.5 Secondary Analyses 220

16.6 Sample R Code for Chapter 16 222

17 The Question of Sample Size 225

17.1 Introduction 225

17.2 General Aspects of Sample Size Calculations 225

17.3 Two Examples of Sample Size Calculations 229

17.4 Some Hazards of Small Studies 231

17.5 Sample R Code for Chapter 17 234

18 The Design of Clinical Trials	236
18.1 Introduction	236
18.2 General Considerations	237
18.3 Trial Organization	239
18.4 Randomized versus Historical Controls	241
18.5 Intention to Treat	244
18.6 Factorial Designs	245
18.7 Repeated Significance Testing	245
18.8 Sequential Analysis	247
19 Further Comments regarding Clinical Trials	250
19.1 Introduction	250
19.2 Surrogate Endpoints	250
19.3 Active Control or Equivalence Trials	252
19.4 Other Designs	253
19.5 Multiple Outcomes	255
19.6 Multiple Treatment Arms	257
19.7 Stochastic Curtailment	259
19.8 Adaptive Trials	261
20 Meta-Analysis	263
20.1 Introduction	263
20.2 Background	264
20.3 Study Heterogeneity	265
20.4 An Illustrative Example	266
20.5 Graphical Displays	269
20.6 Using Funnel Plots to Detect Publication Bias	269
20.7 Sensitivity	272
20.8 Sample R Code for Chapter 20	273
21 Epidemiological Applications	276
21.1 Introduction	276
21.2 Epidemiological Studies	276
21.3 Relative Risk Models	278
21.4 Odds Ratio Models	281
21.5 Confounding and Effect Modification	283
21.6 Mantel-Haenszel Methodology	286
21.7 Poisson Regression Modelling of Cohort Studies	287
21.8 Clinical Epidemiology	292
21.9 Sample R Code for Chapter 21	293
22 Diagnostic Tests	295
22.1 Introduction	295
22.2 Some General Considerations	295
22.3 Sensitivity, Specificity, and Post-Test Probabilities	299
22.4 Likelihood Ratios and Related Issues	303

23 Agreement and Reliability	309
23.1 Introduction	309
23.2 Intraclass Correlation Coefficient	310
23.3 Assessing Agreement	313
23.4 Bland-Altman Plots.....	314
23.5 The Kappa Coefficient	316
23.6 Weighted κ	318
23.7 Measures of Agreement for Discrete Data.....	319
23.8 The Dependence of κ on Prevalence.....	320
23.9 Sample R Code for Chapter 23	322
References	325
Subject Index	329