Contents

List of figures xi

List of tables xiii

Preface xvii

Glossary xix

1 Introduction **1**

1.1 About This Book 1
1.2 Why Sequential Methods 3
1.3 A Short History of Sequential and Group Sequential Methods 5
1.4 Some Examples 11
1.5 Chapter Organization: A Roadmap 15
1.6 Bibliography and Notes 18

2 Two-Sided Tests: Introduction **21**

2.1 Two-Sided Tests for Comparing Two Treatments with Normal Response of Known Variance 21
2.2 A Fixed Sample Test 22
2.3 Group Sequential Tests 23
2.4 Pocock's Test 24
2.5 O'Brien & Fleming's Test 29
2.6 Properties of Pocock and O'Brien & Fleming Tests 31
2.7 Other Tests 39
2.8 Conclusions 47

3 Two-Sided Tests: General Applications **49**

3.1 A Unified Formulation 49
3.2 Applying the Tests with Equal Group Sizes 53
3.3 Applying the Tests with Unequal Increments in Information 57
3.4 Normal Linear Models 62
3.5 Other Parametric Models 70
3.6 Binary Data: Group Sequential Tests for Proportions 74
3.7 The Group Sequential Log-Rank Test for Survival Data 77
3.8 Group Sequential t-Tests 79

4 One-Sided Tests **87**
4.1 Introduction 87
4.2 The Power Family of One-Sided Group Sequential Tests 89
4.3 Adapting Power Family Tests to Unequal Increments in Information 96
4.4 Group Sequential One-Sided t-Tests 99
4.5 Whitehead's Triangular Test 105

5 Two-Sided Tests with Early Stopping Under the Null Hypothesis **111**
5.1 Introduction 111
5.2 The Power Family of Two-Sided Inner Wedge Tests 112
5.3 Whitehead's Double Triangular Test 123

6 Equivalence Tests **129**
6.1 Introduction 129
6.2 One-Sided Tests of Equivalence 130
6.3 Two-Sided Tests of Equivalence: Application to Comparative Bioavailability Studies 130
6.4 Individual Bioequivalence: A One-Sided Test for Proportions 141
6.5 Bibliography and Notes 143

7 Flexible Monitoring: The Error Spending Approach **145**
7.1 Unpredictable Information Sequences 145
7.2 Two-Sided Tests 146
7.3 One-Sided Tests 161
7.4 Data Dependent Timing of Analyses 166
7.5 Computations for Error Spending Tests 169

8 Analysis Following a Sequential Test **171**
8.1 Introduction 171
8.2 Distribution Theory 171
8.3 Point Estimation 175
8.4 P-values 179
8.5 Confidence Intervals 181

9 Repeated Confidence Intervals **189**
9.1 Introduction 189
9.2 Example: Difference of Normal Means 193
9.3 Derived Tests: Use of RCIs to Aid Early Stopping Decisions 194
9.4 Repeated P-values 202
9.5 Discussion 202

10 Stochastic Curtailment **205**
10.1 Introduction 205
10.2 The Conditional Power Approach 205
10.3 The Predictive Power Approach 210

10.4 A Parameter-Free Approach 213
10.5 A Case Study with Survival Data 215
10.6 Bibliography and Notes 219

11 General Group Sequential Distribution Theory 221
11.1 Introduction 221
11.2 A Standard Joint Distribution for Successive Estimates of a Parameter Vector 221
11.3 Normal Linear Models 222
11.4 Normal Linear Models with Unknown Variance: Group Sequential t-Tests 224
11.5 Example: An Exact One-Sample Group Sequential t-Test 226
11.6 General Parametric Models: Generalized Linear Models 228
11.7 Bibliography and Notes 232

12 Binary Data 235
12.1 A Single Bernoulli Probability 235
12.2 Two Bernoulli Probabilities 244
12.3 The Odds Ratio and Multiple 2×2 Tables 251
12.4 Case-Control and Matched Pair Analyses 254
12.5 Logistic Regression: Adjusting for Covariates 256
12.6 Connection with Survival Analysis 257

13 Survival Data 259
13.1 Introduction 259
13.2 The Log-Rank Test 260
13.3 The Stratified Log-Rank Test 261
13.4 Group Sequential Methods for Survival Data with Covariates 262
13.5 Repeated Confidence Intervals for a Hazard Ratio 265
13.6 Example: A Clinical Trial for Carcinoma of the Oropharynx 267
13.7 Survival Probabilities and Quantiles 274
13.8 Bibliography and Notes 276

14 Internal Pilot Studies: Sample Size Re-estimation 279
14.1 The Role of an Internal Pilot Phase 279
14.2 Sample Size Re-estimation for a Fixed Sample Test 281
14.3 Sample Size Re-estimation in Group Sequential Tests 293

15 Multiple Endpoints 299
15.1 Introduction 299
15.2 The Bonferroni Procedure 300
15.3 A Group Sequential Hotelling Test 302
15.4 A Group Sequential Version of O'Brien's Test 306
15.5 Tests Based on Other Global Statistics 310
15.6 Tests Based on Marginal Criteria 311
15.7 Bibliography and Notes 314

16 Multi-Armed Trials **317**
16.1 Introduction 317
16.2 Global Tests 317
16.3 Monitoring Pairwise Comparisons 321
16.4 Bibliography and Notes 324

17 Adaptive Treatment Assignment **327**
17.1 A Multi-Stage Adaptive Design 327
17.2 A Multi-Stage Adaptive Design with Time Trends 331
17.3 Validity of Adaptive Multi-stage Procedures 333
17.4 Bibliography and Notes 335

18 Bayesian Approaches **337**
18.1 The Bayesian Paradigm 337
18.2 Stopping Rules 338
18.3 Choice of Prior Distribution 341
18.4 Discussion 343

19 Numerical Computations for Group Sequential Tests **345**
19.1 Introduction 345
19.2 The Basic Calculation 346
19.3 Error Probabilities and Sample Size Distributions 351
19.4 Tests Defined by Error Spending Functions 353
19.5 Analysis Following a Group Sequential Test 355
19.6 Further Applications of Numerical Computation 357
19.7 Computer Software 360

References **363**

Index **387**

List of figures

2.1	A Pocock test for five groups of observations	28
2.2	An O'Brien & Fleming test for five groups of observations	32
2.3	Boundaries for Pocock, O'Brien & Fleming, Haybittle-Peto, and Wang & Tsiatis tests	47
3.1	Pocock and O'Brien & Fleming tests for four groups of observations expressed in terms of Z_k and S_k	73
4.1	Two power family one-sided tests for four groups of observations	94
4.2	One-sided test with four groups of observations	100
4.3	One-sided triangular tests for four groups of observations	107
5.1	A power family inner wedge test for five groups of observations	117
5.2	Comparison of two inner wedge tests	126
8.1	Sub-densities $p(k, z; \theta)$ when $\theta = 0$ for a four-stage O'Brien & Fleming test	174
8.2	Sub-densities $p(k, z; \theta)$ when $\theta = \delta$ for a four-stage O'Brien & Fleming test	175
8.3	Sampling densities of $\hat{\theta}$ when $\theta = 0$ for a four-stage O'Brien & Fleming test	176
8.4	Sampling densities of $\hat{\theta}$ when $\theta = 1$ for a four-stage O'Brien & Fleming test	177
8.5	Stopping boundaries where the consistency condition does and does not hold for the MLE ordering	185
9.1	RCIs, parent test and derived tests	195
10.1	Conditional and unconditional power curves for a one-sided test	208
10.2	Stopping boundary for a stochastically curtailed one-sided test using the conditional power approach	209
10.3	Stopping boundary for a stochastically curtailed two-sided test using the conditional power approach	210
10.4	Stopping boundary for a stochastically curtailed one-sided test using the predictive power approach with a uniform prior	212

10.5 Stopping boundary for a stochastically curtailed two-sided test using the predictive power approach 213
10.6 Stopping boundary for a stochastically curtailed two-sided test using the parameter-free approach 215
10.7 Power and conditional power curves for the case study 217

List of tables

2.1 Pocock tests: constants $C_P(K, \alpha)$ 26
2.2 Pocock tests: constants $R_P(K, \alpha, \beta)$ 27
2.3 O'Brien & Fleming tests: constants $C_B(K, \alpha)$ 29
2.4 O'Brien & Fleming tests: constants $R_B(K, \alpha, \beta)$ 30
2.5 Pocock and O'Brien & Fleming tests: an example 33
2.6 Sample size distributions for Pocock and O'Brien & Fleming tests 34
2.7 Properties of Pocock tests 37
2.8 Properties of O'Brien & Fleming tests 38
2.9 Wang & Tsiatis tests: constants $C_{WT}(K, \alpha, \Delta)$ 40
2.10 Wang & Tsiatis tests: constants $R_{WT}(K, \alpha, \beta, \Delta)$ 41
2.11 Properties of Wang & Tsiatis tests for $\alpha = 0.05$ and $1 - \beta = 0.8$ 42
2.12 Properties of Wang & Tsiatis tests for $\alpha = 0.05$ and $1 - \beta = 0.9$ 43
2.13 Haybittle-Peto tests: constants $C_{HP}(K, \alpha)$ 45
2.14 Properties of Haybittle-Peto tests 46

3.1 Properties of two-sided tests for unequal group sizes 59
3.2 Properties of two-sided tests for various information sequences 60
3.3 Properties of two-sided group sequential t-tests 82

4.1 Power family one-sided tests: constants and properties for $\alpha = 0.05$ and $1 - \beta = 0.8$ 90
4.2 Power family one-sided tests: constants and properties for $\alpha = 0.05$ and $1 - \beta = 0.9$ 91
4.3 Power family one-sided tests: constants and properties for $\alpha = 0.05$ and $1 - \beta = 0.95$ 92
4.4 One-sided tests: minimum expected sample sizes 95
4.5 Properties of power family one-sided tests for various information sequences 98
4.6 Properties of group sequential one-sided t-tests 103
4.7 Whitehead's triangular tests: constants $R_{Wh}(K, \alpha, \beta)$ and properties 108
4.8 Properties of Whitehead's triangular tests for various information sequences 110

5.1 Power family two-sided inner wedge tests: constants and properties for $\alpha = 0.05$ and $1 - \beta = 0.8$ 114

5.2 Power family two-sided inner wedge tests: constants and properties for $\alpha = 0.05$ and $1 - \beta = 0.9$ 115
5.3 Power family two-sided inner wedge tests: constants and properties for $\alpha = 0.05$ and $1 - \beta = 0.95$ 116
5.4 Properties of power family inner wedge tests for various information sequences 119
5.5 Properties of power family inner wedge t-tests 121
5.6 Whitehead's double triangular tests: constants $R_{DT}(K, \alpha, \beta)$ and properties 125

6.1 Properties of power family equivalence tests for various information sequences 136
6.2 Properties of power family t-tests for equivalence 139
6.3 AUC equivalence test data 140

7.1 Lan & DeMets error spending tests: constants $R_{LD}(K, \alpha, \beta, \rho)$ 150
7.2 Properties of Lan & DeMets error spending tests with $\alpha = 0.05$ and $1 - \beta = 0.8$ 151
7.3 Properties of Lan & DeMets error spending tests with $\alpha = 0.05$ and $1 - \beta = 0.9$ 152
7.4 Power attained by Lan & DeMets error spending tests for various information sequences 155
7.5 Power attained by Lan & DeMets error spending tests when the number of analyses differs from that planned 156
7.6 One-sided error spending tests: constants $R_{OS}(K, \alpha, \beta, \rho)$ 164
7.7 Properties of one-sided error spending tests with $\alpha = 0.05$ and $1 - \beta = 0.8$ 165
7.8 Properties of one-sided error spending tests with $\alpha = 0.05$ and $1 - \beta = 0.9$ 166
7.9 Properties of one-sided error spending tests with $\alpha = 0.05$ and $1 - \beta = 0.95$ 167

9.1 Simultaneous coverage probabilities for naive 95% confidence intervals 190
9.2 Ratios of widths of 95% RCIs to unadjusted 95% confidence intervals 193
9.3 Properties of one-sided tests derived from RCIs 198

10.1 Interim mortality data for the case study 218
10.2 Additional person-years under three possible scenarios for continuing the case study 219

11.1 Repeated t-tests: constants $Z_P(K, m, \alpha)$ and $Z_B(K, m, \alpha)$ 227

12.1 Selected two-stage and three-stage one-sided binomial tests 238
12.2 90% confidence intervals following a three-stage sampling plan 240

12.3 90% confidence intervals following a three-stage sampling plan (continued) 241
12.4 Continuation and acceptance regions for a three-stage test of $p_A = p_B$ 247
12.5 Acceptance region of a modified three-stage test of $p_A = p_B$ 248
12.6 Hypothetical interim data for the Ille-et-Vilaine study 253
12.7 Interim results for the data of Table 12.6 254
12.8 Hypothetical interim data for the Leisure World study 255
12.9 Interim results for the data of Table 12.8 255

13.1 Summary data and critical values for a group sequential stratified log-rank test in a trial for carcinoma of the oropharynx 270
13.2 Repeated confidence intervals for the log hazard ratio for the oropharynx trial data 271
13.3 Group sequential estimates of Cox model parameters for the oropharynx trial data 272
13.4 Test statistics and critical values for a group sequential test based on a Cox regression model for the oropharynx trial data 273

14.1 Properties of "internal pilot" tests for the difference of two binary response probabilities using Herson & Wittes' approach 284
14.2 Properties of "internal pilot" tests for the difference of two binary response probabilities using Gould's approach 285
14.3 Properties of "internal pilot" tests for the ratio of two binary response probabilities 287
14.4 Properties of "internal pilot" tests of Wittes & Brittain and of Birkett & Day for two normal distributions of unknown variance 289
14.5 Properties of "internal pilot" tests of Denne & Jennison for two normal distributions of unknown variance 292

15.1 Bonferroni procedures with p endpoints: constants $C_P(K, \alpha)$ for Pocock tests 301
15.2 Bonferroni procedures with p endpoints: constants $C_B(K, \alpha)$ for O'Brien & Fleming tests 302
15.3 Worksheet for a group sequential test using Hotelling's $\mathcal{T}^2$ statistic 305
15.4 Worksheet for a group sequential test using O'Brien's GLS statistic 310

16.1 Repeated χ^2 tests of homogeneity of J normal means: constants $C_P(p, K, \alpha)$ and $C_B(p, K, \alpha)$ 319

17.1 Properties of power family one-sided tests using adaptive sampling 330

18.1 False positive rates for stopping rules based on Bayesian posterior probabilities 339