

Contents

Preface	xvii
Preface to the Second Edition	xxi
Preface to the First Edition	xxv
1. An Introduction to Applied Probability	1
1.1. Notation and Definitions, 1	
1.2. The Rule of Total Probability, 4	
1.3. The Evaluation of a Screening Test, 5	
1.4. Biases Resulting from the Study of Selected Samples, 9	
Problems, 13	
References, 15	
2. Statistical Inference for a Single Proportion	17
2.1. Exact Inference for a Single Proportion: Hypothesis Tests, 18	
2.2. Exact Inference for a Single Proportion: Interval Estimation, 22	
2.2.1. Definition of an Exact Confidence Interval, 22	
2.2.2. A Fundamental Property of Confidence Intervals, 24	
2.3. Using the F Distribution, 25	
2.4. Approximate Inference for a Single Proportion, 26	
2.4.1. Hypothesis Tests, 27	
2.4.2. Confidence Intervals, 28	

- 2.5. Sample Size for a One-Sample Study, 29
 - 2.5.1. Sample Size for Hypothesis Tests, 30
 - 2.5.2. Sample Size for Confidence Intervals, 34
- 2.6.* Standard Errors by the Delta Method, 36
- 2.7.* Alternative Definitions of Two-Sided P-Values and Confidence Intervals, 37
 - 2.7.1. The Point Probability Method, 37
 - 2.7.2. The Tail Probability Method, 39
 - 2.7.3. The Likelihood Ratio Method, 40
 - 2.7.4. Some Concluding Remarks, 41
- Problems, 43
- References, 49

3. Assessing Significance in a Fourfold Table 50

- 3.1. Methods for Generating a Fourfold Table, 51
- 3.2. “Exact” Analysis of a Fourfold Table, 55
- 3.3. Yates’ Correction for Continuity, 57
- 3.4. One-Tailed versus Two-Tailed Tests, 58
- 3.5. A Simple Confidence Interval for the Difference between Two Independent Proportions, 60
- 3.6. An Alternative Critical Ratio Test, 61
- Problems, 62
- References, 63

4. Determining Sample Sizes Needed to Detect a Difference between Two Proportions 64

- 4.1. Specifying a Difference Worth Detecting, 65
- 4.2. The Mathematics of Sample Size Determination, 69
- 4.3. Using the Sample Size Tables, 73
- 4.4. Unequal Sample Sizes, 75
- 4.5. Some Additional Uses of the Tables, 77
- 4.6. Some Additional Comments, 80
- Problems, 80
- References, 83

5. How to Randomize 86

- 5.1. Selecting a Simple Random Sample, 87
- 5.2. Randomization in a Clinical Trial, 89

5.3.	Variations on Simple Randomization, 93	
	References, 94	
6.	Comparative Studies: Cross-sectional, Naturalistic, or Multinomial Sampling	95
6.1.	Some Hypothetical Data, 96	
6.2.	Measures of Association Derived from χ^2 , 97	
6.3.	The Odds Ratio and Its Logarithm, 100	
6.4.	Exact Inference for an Odds Ratio: Hypothesis Tests, 106	
6.5.	Exact Inference for an Odds Ratio: Confidence Intervals, 111	
6.6.	Approximate Inference for an Odds Ratio, 112	
6.6.1.	Hypothesis Tests, 113	
6.6.2.	Confidence Intervals, 116	
6.6.3.*	A Confidence Interval Method to Be Avoided, 119	
6.7.	Criticisms of the Odds Ratio, 122	
6.8.	Attributable Risk, 125	
6.9.*	Standard Errors for Measures of Association, 129	
	Problems, 135	
	References, 140	
7.	Comparative Studies: Prospective and Retrospective Sampling	144
7.1.	Prospective Studies, 144	
7.2.	Retrospective Studies, 147	
7.3.	Estimating Attributable Risk from Retrospective Studies, 151	
7.4.	The Retrospective Approach versus the Prospective Approach, 153	
	Problems, 154	
	References, 157	
8.	Randomized Controlled Trials	159
8.1.	The Simple Comparative Trial, 161	
8.2.	The Two-Period Crossover Design, 163	
8.3.	Factors Affecting Power in a Randomized Controlled Trial, 165	
8.3.1.	The Intent-to-Treat Principle, 165	

- 10.9.2. Fixed and Random Effects Meta-analysis, 265
- 10.9.3.* Tests of Odds Ratio Homogeneity in the Large Sparse Case, 269

Problems, 276

References, 279

11. Logistic Regression

284

- 11.1. Introduction, 284
- 11.2. The Log Odds Transformation Revisited, 286
- 11.3. A Closer Look at Some Logistic Regression Models, 292
 - 11.3.1. Simple Binary Logistic Regression, 292
 - 11.3.2. Multiple Binary Logistic Regression: Additive and Interactive Models, 296
 - 11.3.3. Generalized Likelihood Ratio Tests, 300
 - 11.3.4. Log-Likelihood Goodness-of-Fit Tests and the Analysis of Information, 304
- 11.4. Polytomous Logistic Regression, 308
 - 11.4.1. An Example: The Double Dichotomy, 308
 - 11.4.2. A General Framework for Multinomial Responses: The Exponential Family of Distributions, 314
 - 11.4.3.* General Logistic Regression, 325

Problems, 335

References, 338

12. Poisson Regression

340

- 12.1. Poisson Random Variables, 340
- 12.2. Poisson Regression, 345
 - 12.2.1. Simple Poisson Regression, 345
 - 12.2.2. Multiple Poisson Regression, 348
- 12.3.* Overdispersion, 358

Problems, 366

References, 371

13. The Analysis of Data from Matched Samples

373

- 13.1. Matched Pairs: Dichotomous Outcome, 374
- 13.2. Matched Pairs: Polytomous Outcome, 380

13.3.	Multiple Matched Controls per Case, 384	
13.4.	The Comparison of Matched Samples with m Distinct Types, 387	
13.5.	Sample Size Determination for Matched Samples, 394	
13.6.	Advantages and Disadvantages of Matching, 398	
	Problems, 400	
	References, 403	
14.	Regression Models for Matched Samples	407
14.1.	Direct and Indirect Parametric Modeling of Matched-Sample Data, 409	
14.2.	Conditional Logistic Regression, 416	
	14.2.1. Matched Case-Control Studies: One Case Matched to One or More Controls, 416	
	14.2.2. Matched Prospective Studies, 423	
14.3.	Extensions, 427	
	14.3.1. Matched Studies with Varying Numbers of Cases and Controls, 427	
	14.3.2. Matched Studies with Polytomous Outcomes, 428	
14.4.	An Example, 431	
14.5.	Other Issues, 434	
	Problems, 436	
	References, 437	
15.	Analysis of Correlated Binary Data	440
15.1.	Inference for a Single Proportion, 441	
15.2.	Inference for Two Proportions, 444	
15.3.	Design Considerations, 448	
15.4.	$2 \times 2 \times 2$ Tables, 449	
	15.4.1. Hypothesis Testing, 449	
	15.4.2. Inference for the Mantel-Haenszel Odds Ratio with Clustered Data, 460	
15.5.*	Extensions of Logistic Regression for Correlated Outcomes, 462	
	15.5.1. Generalized Estimating Equations, 462	
	15.5.2. Random Effects Models, 475	

17.2.	The Algebra of Misclassification, 565	
17.3.	The Algebra of Misclassification: Both Variables in Error, 569	
17.4.	Statistical Control for Error, 571	
17.5.	Probabilistic Control for Error, 573	
17.6.	Experimental Control of Error, 575	
17.7.*	Misclassification in Logistic Regression Models, 578	
	Problems, 589	
	References, 592	
18.	The Measurement of Interrater Agreement	598
18.1.	The Same Pair of Raters per Subject, 599	
18.2.	Weighted Kappa, 608	
18.3.	Multiple Ratings per Subject with Different Raters, 610	
18.4.	Further Applications, 617	
18.5.*	Interrater Agreement as Association in a Multivariate Binary Vector, 618	
	Problems, 622	
	References, 623	
19.	The Standardization of Rates	627
19.1.	Reasons for and Warnings against Standardization, 629	
19.2.	Two Types of Standardization: Direct and Indirect, 630	
19.3.	Indirect Standardization, 632	
19.4.	A Feature of Indirect Standardization, 634	
19.5.	Direct Standardization, 636	
19.6.	Some Other Summary Indices, 638	
19.7.	Adjustment for Two Factors, 640	
	Problems, 645	
	References, 646	
Appendix A.	Numerical Tables	648
A.1.	Critical Values of the Normal Distribution, 649	
A.2.	Critical Values of the Chi Squared Distribution, 650	
A.3.	Percentage Points of the F Distribution, 652	
A.4.	Sample Sizes per Group for a Two-Tailed Test of Two Proportions, 660	

A.5.	20,000 Random Digits, 684	
A.6.	Natural Logarithms of $x(\ln x)$, 692	
A.7.	Percentage Points of Bartholomew's Test for Order When $m = 3$ Proportions Are Compared, 694	
A.8.	Percentage Points of Bartholomew's Test for Order When $m = 4$ Proportions Are Compared, 695	
A.9.	Percentage Points of Bartholomew's Test for Order When Up to $m = 12$ Proportions Based on Equal Sample Sizes Are Compared, 696	
A.10.	The Logit Transformation, 697	
Appendix B.	The Basic Theory of Maximum Likelihood Estimation	698
Appendix C.	Answers to Selected Problems	705
Author Index		727
Subject Index		737