Contents

Tables xvi
Boxes xx
Figures xxii
Getting Files from the Wiley ftp and Internet Sites xxiii
List of Data Sites Provided on Web Site xxiv
Preface to the Fourth Edition xxv

PART 1 BASIC CONCEPTS 1

1. Uses of Sample Surveys 3
1.1 Why Sample Surveys Are Used, 3
1.2 Designing Sample Surveys, 5
   1.2.1 Sample Design, 5
   1.2.2 Survey Measurements, 6
   1.2.3 Survey Operations, 6
   1.2.4 Statistical Analysis and Report Writing, 7
1.3 Preliminary Planning of a Sample Survey, 7
Exercises, 8
Bibliography, 9

2. The Population and the Sample 11
2.1 The Population, 11
   2.1.1 Elementary Units, 13
   2.1.2 Population Parameters, 13
2.2 The Sample, 18
   2.2.1 Probability and Nonprobability Sampling, 18
   2.2.2 Sampling Frames, Sampling Units, and Enumeration Units, 19
2.3 Sampling Measurements and Summary Statistics, 20
   2.2.4 Estimation of Population Characteristics, 22
2.3 Sampling Distributions, 25
2.4 Characteristics of Estimates of Population Parameters, 30
  2.4.1 Bias, 31
  2.4.2 Mean Square Error, 32
  2.4.3 Validity, Reliability, and Accuracy, 35
2.5 Criteria for a Good Sample Design, 36
2.6 Summary, 37
Exercises, 37
Bibliography, 42

PART 2 MAJOR SAMPLING DESIGNS AND ESTIMATION PROCEDURES 43

3. Simple Random Sampling 45
  3.1 What Is a Simple Random Sample?, 45
    3.1.1 How to Take a Simple Random Sample, 46
    3.1.2 Probability of an Element Being Selected, 47
  3.2 Estimation of Population Characteristics Under Simple Random Sampling, 49
    3.2.1 Estimation Formulas, 49
    3.2.2 Numerical Computation of Estimates and Their Standard Errors, 50
  3.3 Sampling Distributions of Estimated Population Characteristics, 55
  3.4 Coefficients of Variation of Estimated Population Parameters, 58
  3.5 Reliability of Estimates, 61
  3.6 Estimation of Parameters for Subdomains, 64
  3.7 How Large a Sample Do We Need?, 70
  3.8 Why Simple Random Sampling Is Rarely Used, 75
  3.9 Summary, 76
Exercises, 76
Bibliography, 79

4. Systematic Sampling 83
  4.1 How To Take a Systematic Sample, 83
  4.2 Estimation of Population Characteristics, 84
  4.3 Sampling Distribution of Estimates, 86
  4.4 Variance of Estimates, 90
  4.5 A Modification That Always Yields Unbiased Estimates, 92
  4.6 Estimation of Variances, 100
  4.7 Repeated Systematic Sampling, 103
CONTENTS

4.7.1 Use of Stata For Estimation In Repeated Systematic Sampling, 107
4.7.2 Use of SUDAAN for Estimation in Repeated Systematic Sampling, 109
4.8 How Large a Sample Do We Need?, 111
4.9 Using Frames That Are Not Lists, 113
4.10 Summary, 114
Exercises, 114
Bibliography, 120

5. Stratification and Stratified Random Sampling 121
5.1 What Is a Stratified Random Sample?, 121
5.2 How to Take a Stratified Random Sample, 122
5.3 Why Stratified Sampling? 122
5.4 Population Parameters for Strata, 128
5.5 Sample Statistics for Strata, 133
5.6 Estimation of Population Parameters from Stratified Random Sampling, 134
5.7 Summary, 139
Exercises, 139
Bibliography, 142

6. Stratified Random Sampling: Further Issues 143
6.1 Estimation of Population Parameters, 143
6.2 Sampling Distributions of Estimates, 144
6.3 Estimation of Standard Errors, 146
6.4 Estimation of Characteristics of Subgroups, 148
6.5 Allocation of Sample to Strata, 150
   6.5.1 Equal Allocation, 151
   6.5.2 Proportional Allocation: Self-Weighting Samples, 151
   6.5.3 Optimal Allocation, 158
   6.5.4 Optimal Allocation and Economics, 160
6.6 Stratification After Sampling, 168
6.7 How Large a Sample Is Needed?, 175
6.8 Construction of Stratum Boundaries and Desired Number of Strata, 179
6.9 Summary, 183
Exercises, 184
Bibliography, 188
10.1.1 How to Take a Simple Two-Stage Cluster Sample, 270
10.1.2 Estimation of Population Characteristics, 271
10.1.3 Estimation of Standard Errors, 273
10.1.4 Sampling Distribution of Estimates, 284
10.1.5 How Large a Sample Is Needed? 289
10.1.6 Choosing the Optimal Cluster Size $n$ Considering Costs, 292
10.1.7 Some Shortcut Formulas for Determining the Optimal Number $n$, 295
10.2 Situation in Which Not All Clusters Have the Same Number $N_i$ of Enumeration Units, 300
10.2.1 How to Take a Simple Two-Stage Cluster Sample for This Design, 300
10.2.2 Estimation of Population Characteristics, 301
10.2.3 Estimation of Standard Errors of Estimates, 301
10.2.4 Sampling Distribution of Estimates, 311
10.2.5 How Large a Sample Do We Need? 314
10.2.6 Choosing the Optimal Cluster Size $n$ Considering Costs, 317
10.3 Systematic Sampling as Cluster Sampling, 319
10.4 Summary, 320
Exercises, 321
Bibliography, 330

11. Cluster Sampling in Which Clusters Are Sampled with Unequal Probability: Probability Proportional to Size Sampling

11.1 Motivation for Not Sampling Clusters with Equal Probability, 332
11.2 Two General Classes of Estimators Valid for Sample Designs in Which Units Are Selected with Unequal Probability, 336
11.2.1 The Horvitz-Thompson Estimator, 336
11.2.2 The Hansen-Hurwitz Estimator, 337
11.3 Probability Proportional to Size Sampling, 340
11.3.1 Probability Proportional to Size Sampling with Replacement: Use of the Hansen-Hurwitz Estimator, 342
11.3.2 PPS Sampling When the Measure of Size Variable Is not the Number of Enumeration Units, 350
11.3.3 How to Take a PPS Sample with Replacement, 353
11.3.4 Sequential Methods of PPS Sampling with Replacement—Chromy’s Probability with Minimum Replacement (PMR) Method, 353
11.3.5 How Large a Sample Is Needed for a Two-Stage Sample in Which Clusters Are Selected PPS with Replacement? 356
11.3.6 Telephone PPS Sampling: The Mitofsky-Waksberg Method of Random Digit Dialing, 359
11.4 Further Comment on PPS Sampling, 361
11.5 Summary, 361
Exercises, 362
Bibliography, 364

12. Variance Estimation in Complex Sample Surveys 367
12.1 Linearization, 368
12.2 Replication Methods, 373
   12.2.1 The Balanced Repeated Replication Method, 373
   12.2.2 Jackknife Estimation, 381
   12.2.3 Estimation of Interviewer Variability by Use of Replicated Sampling (Interpenetrating Samples), 384
12.3 Summary, 386
Exercises, 387
Technical Appendix, 390
Bibliography, 392

PART 3 SELECTED TOPICS IN SAMPLE SURVEY METHODOLOGY 395

13. Nonresponse and Missing Data in Sample Surveys 397
13.1 Effect of Nonresponse on Accuracy of Estimates, 397
13.2 Methods of Increasing the Response Rate in Sample Surveys, 399
   13.2.1 Increasing the Number of Households Contacted Successfully, 399
   13.2.2 Increasing the Completion Rate in Mail Questionnaires, 400
   13.2.3 Decreasing the Number of Refusals in Face-to-Face Telephone Interviews, 400
   13.2.4 Using Endorsements, 401
13.3 Mail Surveys Combined with Interviews of Nonrespondents, 402
   13.3.1 Determination of Optimal Fraction of Initial Nonrespondents to Subsample for Intensive Effort, 403
   13.3.2 Determination of Sample Size Needed for a Two-Stage Mail Survey, 405
13.4 Other Uses of Double (or Two-Phase) Sampling Methodology, 406
13.5 Item Nonresponse: Methods of Imputation, 408
13.5.1 Mechanisms by Which Missing Values Arise, 408
13.5.2 Some Methods for Analyzing Data in the Presence of Missing Values, 411
13.5.3 Some Imputation Methods, 412
13.6 Multiple Imputation, 416
13.7 Summary, 419
Exercises, 419
Bibliography, 425

14. Selected Topics in Sample Design and Estimation Methodology
14.1 World Health Organization EPI Surveys: A Modification of PPS Sampling for Use in Developing Countries, 427
14.2 Quality Assurance Sampling, 429
14.3 Sample Sizes for Longitudinal Studies, 432
14.3.1 Simple Random Sampling, 433
14.3.2 Simple One-Stage Cluster Sampling, 434
14.3.3 Cluster Sampling with More Than One Domain, 435
14.4 Estimation of Prevalence of Diseases from Screening Studies, 436
14.5 Estimation of Rare Events: Network Sampling, 440
14.6 Estimation of Rare Events: Dual Samples, 444
14.7 Estimation of Characteristics for Local Areas: Synthetic Estimation, 446
14.8 Extraction of Sensitive Information: Randomized Response Techniques, 449
14.9 Summary, 451
Exercises, 451
Bibliography, 452

15. Telephone Survey Sampling
Michael W. Link and Mansour Fahimi
15.1 Introduction, 455
15.1.1 The Twentieth Century, 455
15.1.2 The Twenty-First Century, 456
15.2 History of Telephone Sampling in the United States, 456
15.2.1 Early Design of Telephone Surveys, 457
15.2.2 Random Digit Dialing, 457
15.2.3 Mitofsky-Waksberg Sampling Method, 458
15.2.4 List-Assisted Random Digit Dialing Methods, 458
15.3 Within-Household Selection Techniques, 459
15.3.1 Probability-Based Methods, 460
15.3.2 Quasi-Probability Methods, 460
15.3.3 Nonprobability Methods, 461
15.3.4 Minimally Intrusive Method, 461
15.4 Steps in the Telephone Survey Process, 461
15.4.1 Computer-Assisted Telephone Interviewing, 462
15.4.2 Quality Control in Telephone Surveys, 463
15.5 Drawing and Managing a Telephone Survey Sample, 463
15.5.1 Drawing the Sample, 463
15.5.2 Managing the Sample, 464
15.5.3 Developing an Analysis File, 465
15.5.4 Data Weighting and Adjustment, 466
15.6 Post-Survey Data Enhancement Procedures, 466
15.6.1 Data Weighting, 466
15.6.2 Steps in the Weighting Process, 466
15.6.3 Compensation for Exclusion of Nontelephone Households, 474
15.7 Imputation of Missing Data, 476
15.8 Declining Coverage and Response Rates, 477
15.9 Addressing the Problems with Cell Phones, 478
15.9.1 Research on Cell Phone Surveys, 479
15.9.2 Sampling from the Cell Phone Frame, 479
15.10 Address-Based Sampling, 482
Exercises, 483
Bibliography, 483

16. Constructing the Survey Weights
Paul P. Biemer and Sharon L. Christ
16.1 Introduction, 489
16.2 Objectives of Weighting, 492
16.2.1 Basic Concepts, 492
16.2.2 Weighting to Reduce Frame Bias, 494
16.2.3 Weighting to Reduce Nonresponse Bias, 495
16.2.4 Weighting to Reduce Sampling Variance, 496
16.3 Constructing the Sampling Weights, 498
16.3.1 Base Weights, 499
16.3.2 Nonresponse Adjustments, 501
16.3.3 Frame Coverage Adjustments, 505
16.3.4 Constructing the Final Weights, 510
16.4 Estimation and Analysis Issues, 511
16.4.1 Effect of Weighting on the Variance, 511
16.4.2 Using Weights in Analysis, 513
16.5 Summary, 514
Bibliography, 515