Contents

Preface v

PART I
Principles of Estimation for Finite Populations and Important Sampling Designs

CHAPTER 1
Survey Sampling in Theory and Practice 3

1.1 Surveys in Society 3
1.2 Skeleton Outline of a Survey 4
1.3 Probability Sampling 8
1.4 Sampling Frame 9
1.5 Area Frames and Similar Devices 12
1.6 Target Population and Frame Population 13
1.7 Survey Operations and Associated Sources of Error 14
1.8 Planning a Survey and the Need for Total Survey Design 17
1.9 Total Survey Design 19
1.10 The Role of Statistical Theory in Survey Sampling 20
Exercises 22

CHAPTER 2
Basic Ideas in Estimation from Probability Samples 24

2.1 Introduction 24
2.2 Population, Sample, and Sample Selection 24
2.3 Sampling Design 27
2.4 Inclusion Probabilities 30
2.5 The Notion of a Statistic 33
2.6 The Sample Membership Indicators 36
2.7 Estimators and Their Basic Statistical Properties 38

Bibliografische Informationen
http://d-nb.info/920992218

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

2.8 The π Estimator and Its Properties 42
2.9 With-Replacement Sampling 48
2.10 The Design Effect 53
2.11 Confidence Intervals 55
Exercises 58

CHAPTER 3
Unbiased Estimation for Element Sampling Designs 61

3.1 Introduction 61
3.2 Bernoulli Sampling 62
3.3 Simple Random Sampling 66
3.3.1 Simple Random Sampling without Replacement 66
3.3.2 Simple Random Sampling with Replacement 72
3.4 Systematic Sampling 73
3.4.1 Definitions and Main Result 73
3.4.2 Controlling the Sample Size 76
3.4.3 The Efficiency of Systematic Sampling 78
3.4.4 Estimating the Variance 83
3.5 Poisson Sampling 85
3.6 Probability Proportional-to-Size Sampling 87
3.6.1 Introduction 87
3.6.2 *πps* Sampling 90
3.6.3 *pps* Sampling 97
3.6.4 Selection from Randomly Formed Groups 99
3.7 Stratified Sampling 100
3.7.1 Introduction 100
3.7.2 Notation, Definitions, and Estimation 101
3.7.3 Optimum Sample Allocation 104
3.7.4 Alternative Allocations under *STSI* Sampling 106
3.8 Sampling without Replacement versus Sampling with Replacement 110
3.8.1 Alternative Estimators for Simple Random Sampling with Replacement 110
3.8.2 The Design Effect of Simple Random Sampling with Replacement 112
Exercises 114

CHAPTER 4
Unbiased Estimation for Cluster Sampling and Sampling in Two or More Stages 124

4.1 Introduction 124
4.2 Single-Stage Cluster Sampling 126
4.2.1 Introduction 126
4.2.2 Simple Random Cluster Sampling 129
4.3 Two-Stage Sampling 133
4.3.1 Introduction 133
4.3.2 Two-Stage Element Sampling 135
4.4 Multistage Sampling 144
4.4.1 Introduction and a General Result 144
4.4.2 Three-Stage Element Sampling 146
4.5 With-Replacement Sampling of PSUs 150

4.6 Comparing Simplified Variance Estimators in Multistage Sampling 153
Exercises 154

CHAPTER 5
Introduction to More Complex Estimation Problems 162

5.1 Introduction 162
5.2 The Effect of Bias on Confidence Statements 163
5.3 Consistency and Asymptotic Unbiasedness 166
5.4 π Estimators for Several Variables of Study 169
5.5 The Taylor Linearization Technique for Variance Estimation 172
5.6 Estimation of a Ratio 176
5.7 Estimation of a Population Mean 181
5.8 Estimation of a Domain Mean 184
5.9 Estimation of Variances and Covariances in a Finite Population 186
5.10 Estimation of Regression Coefficients 190
5.10.1 The Parameters of Interest 190
5.10.2 Estimation of the Regression Coefficients 192
5.11 Estimation of a Population Median 197
5.12 Demonstration of Result 5.10.1 205
Exercises 207

PART II
Estimation through Linear Modeling, Using Auxiliary Variables

CHAPTER 6
The Regression Estimator 219

6.1 Introduction 219
6.2 Auxiliary Variables 219
6.3 The Difference Estimator 221
6.4 Introducing the Regression Estimator 225
6.5 Alternative Expressions for the Regression Estimator 230
6.6 The Variance of the Regression Estimator 234
6.7 Comments on the Role of the Model 238
6.8 Optimal Coefficients for the Difference Estimator 239
Exercises 242

CHAPTER 7
Regression Estimators for Element Sampling Designs 245

7.1 Introduction 245
7.2 Preliminary Considerations 245
7.3 The Common Ratio Model and the Ratio Estimator 247
7.3.1 The Ratio Estimator under *SI* Sampling 249
7.3.2 The Ratio Estimator under Other Designs 252
7.3.3 Optimal Sampling Design for the π Weighted Ratio Estimator 253
7.3.4 Alternative Ratio Models 255
7.4 The Common Mean Model 258
7.5 Models Involving Population Groups 260
7.6 The Group Mean Model and the Poststratified Estimator 264
7.7 The Group Ratio Model and the Separate Ratio Estimator 269

7.8 Simple Regression Models and Simple Regression Estimators 272
7.9 Estimators Based on Multiple Regression Models 275
7.9.1 Multiple Regression Models 276
7.9.2 Analysis of Variance Models 281
7.10 Conditional Confidence Intervals 283
7.10.1 Conditional Analysis for *BE* Sampling 284
7.10.2 Conditional Analysis for the Poststratification Estimator 287
7.11 Regression Estimators for Variable-Size Sampling Designs 289
7.12 A Class of Regression Estimators 291
7.13 Regression Estimation of a Ratio of Population Totals 294
Exercises 297

CHAPTER 8
Regression Estimators for Cluster Sampling and Two-Stage Sampling 303

8.1 Introduction 303
8.2 The Nature of the Auxiliary Information When Clusters of Elements Are Selected 304
8.3 Comments on Variance and Variance Estimation in Two-Stage Sampling 307
8.4 Regression Estimators Arising Out of Modeling at the Cluster Level 308
8.5 The Common Ratio Model for Cluster Totals 312
8.6 Estimation of the Population Mean When Clusters Are Sampled 314
8.7 Design Effects for Single-Stage Cluster Sampling 315
8.8 Stratified Clusters and Poststratified Clusters 319
8.9 Regression Estimators Arising Out of Modeling at the Element Level 322
8.10 Ratio Models for Elements 327
8.11 The Group Ratio Model for Elements 330
8.12 The Ratio Model Applied within a Single PSU 332
Exercises 333

PART III
Further Questions in Design and Analysis of Surveys

CHAPTER 9
Two-Phase Sampling 343

9.1 Introduction 343
9.2 Notation and Choice of Estimator 345
9.3 The π^* Estimator 347
9.4 Two-Phase Sampling for Stratification 350
9.5 Auxiliary Variables for Selection in Two Phases 354
9.6 Difference Estimators 356
9.7 Regression Estimators for Two-Phase Sampling 359
9.8 Stratified Bernoulli Sampling in Phase Two 366
9.9 Sampling on Two Occasions 368
9.9.1 Estimating the Current Total 370
9.9.2 Estimating the Previous Total 376
9.9.3 Estimating the Absolute Change and the Sum of the Totals 377
Exercises 379

CHAPTER 10
Estimation for Domains 386

10.1 Introduction 386
10.2 The Background for Domain Estimation 387
10.3 The Basic Estimation Methods for Domains 390
10.4 Conditioning on the Domain Sample Size 396
10.5 Regression Estimators for Domains 397
10.6 A Ratio Model for Each Domain 403
10.7 Group Models for Domains 405
10.8 Problems Arising for Small Domains; Synthetic Estimation 408
10.9 More on the Comparison of Two Domains 412
Exercises 413

CHAPTER 11
Variance Estimation 418

11.1 Introduction 418
11.2 A Simplified Variance Estimator under Sampling without Replacement 421
11.3 The Random Groups Technique 423
11.3.1 Independent Random Groups 423
11.3.2 Dependent Random Groups 426
11.4 Balanced Half-Samples 430
11.5 The Jackknife Technique 437
11.6 The Bootstrap 442
11.7 Concluding Remarks 444
Exercises 445

CHAPTER 12
Searching for Optimal Sampling Designs 447

12.1 Introduction 447
12.2 Model-Based Optimal Design for the General Regression Estimator 448
12.3 Model-Based Optimal Design for the Group Mean Model 455
12.4 Model-Based Stratified Sampling 456
12.5 Applications of Model-Based Stratification 461
12.6 Other Approaches to Efficient Stratification 462
12.7 Allocation Problems in Stratified Random Sampling 465
12.8 Allocation Problems in Two-Stage Sampling 471
12.8.1 The π Estimator of the Population Total 471
12.8.2 Estimation of the Population Mean 475
12.9 Allocation in Two-Phase Sampling for Stratification 478
12.10 A Further Comment on Mathematical Programming 480
12.11 Sampling Design and Experimental Design 481
Exercises 481

CHAPTER 13
Further Statistical Techniques for Survey Data 485

13.1 Introduction 485
13.2 Finite Population Parameters in Multivariate Regression and Correlation Analysis 486

13.3 The Effect of Sampling Design on a Statistical Analysis 491
13.4 Variances and Estimated Variances for Complex Analyses 494
13.5 Analysis of Categorical Data for Finite Populations 500
13.5.1 Test of Homogeneity for Two Populations 500
13.5.2 Testing Homogeneity for More than Two Finite Populations 507
13.5.3 Discussion of Categorical Data Tests for Finite Populations 510
13.6 Types of Inference When a Finite Population Is Sampled 513
Exercises 520

PART IV
A Broader View of Errors in Surveys

CHAPTER 14
Nonsampling Errors and Extensions of Probability Sampling Theory 525

14.1 Introduction 525
14.2 Historic Notes: The Evolution of the Probability Sampling Approach 525
14.3 Measurable Sampling Designs 527
14.4 Some Nonprobability Sampling Methods 529
14.5 Model-Based Inference from Survey Samples 533
14.6 Imperfections in the Survey Operations 537
14.6.1 Ideal Conditions for the Probability Sampling Approach 537
14.6.2 Extension of the Probability Sampling Approach 538
14.7 Sampling Frames 540
14.7.1 Frame Imperfections 540
14.7.2 Estimation in the Presence of Frame Imperfections 543
14.7.3 Multiple Frames 545
14.7.4 Frame Construction and Maintenance 545
14.8 Measurement and Data Collection 546
14.9 Data Processing 548
14.10 Nonresponse 551
Exercises 553

CHAPTER 15
Nonresponse 556

15.1 Introduction 556
15.2 Characteristics of Nonresponse 556
15.2.1 Definition of Nonresponse 556
15.2.2 Response Sets 557
15.2.3 Lack of Unbiased Estimators 558
15.3 Measuring Nonresponse 559
15.4 Dealing with Nonresponse 563
15.4.1 Planning of the Survey 564
15.4.2 Callbacks and Follow-Ups 564
15.4.3 Subsampling of Nonrespondents 566
15.4.4 Randomized Response 570
15.5 Perspectives on Nonresponse 573
15.6 Estimation in the Presence of Unit Nonresponse 575
15.6.1 Response Modeling 575

15.6.2 A Useful Response Model 577
15.6.3 Estimators That Use Weighting Only 580
15.6.4 Estimators That Use Weighting as Well as Auxiliary Variables 583
15.7 Imputation 589
Exercises 595

CHAPTER 16
Measurement Errors 601

16.1 Introduction 601
16.2 On the Nature of Measurement Errors 602
16.3 The Simple Measurement Model 605
16.4 Decomposition of the Mean Square Error 608
16.5 The Risk of Underestimating the Total Variance 612
16.6 Repeated Measurements as a Tool in Variance Estimation 614
16.7 Measurement Models Taking Interviewer Effects into Account 617
16.8 Deterministic Assignment of Interviewers 618
16.9 Random Assignment of Interviewers to Groups 622
16.10 Interpenetrating Subsamples 627
16.11 A Measurement Model with Sample-Dependent Moments 630
Exercises 634

CHAPTER 17
Quality Declarations for Survey Data 637

17.1 Introduction 637
17.2 Policies Concerning Information on Data Quality 638
17.3 Statistics Canada's Policy on Informing Users of Data Quality and Methodology 641
Exercise 648

APPENDIX A
Principles of Notation 649

APPENDIX B
The MU284 Population 652

APPENDIX C
The Clustered MU284 Population 660

APPENDIX D
The CO124 Population 662

References 666

Answers to Selected Exercises 680

Author Index 684

Subject Index 688