
Advanced Sampling Theory with Applications

How Michael 'selected' Amy

VOLUME I

by
Sarjinder Singh

Kluwer Academic Publishers

TABLE OF CONTENTS

PREFACE

xxi

1 BASIC CONCEPTS AND MATHEMATICAL NOTATION

1.0	Introduction	1
1.1	Population	1
1.1.1	Finite population	1
1.1.2	Infinite population	1
1.1.3	Target population	1
1.1.4	Study population	1
1.2	Sample	2
1.3	Examples of populations and samples	2
1.4	Census	2
1.5	Relative aspects of sampling versus census	2
1.6	Study variable	2
1.7	Auxiliary variable	3
1.8	Difference between study variable and auxiliary variable	3
1.9	Parameter	3
1.10	Statistic	3
1.11	Statistics	4
1.12	Sample selection	4
1.12.1	Chit method or Lottery method	4
1.12.1.1	With replacement sampling	4
1.12.1.2	Without replacement sampling	5
1.12.2	Random number table method	5
1.12.2.1	Remainder method	6
1.13	Probability sampling	7
1.14	Probability of selecting a sample	7
1.15	Population mean/total	8
1.16	Population moments	8
1.17	Population standard deviation	8
1.18	Population coefficient of variation	8
1.19	Relative mean square error	9
1.20	Sample mean	9
1.21	Sample variance	9
1.22	Estimator	10
1.23	Estimate	10
1.24	Sample space	10
1.25	Univariate random variable	11
1.25.1	Qualitative random variables	11

1.25.2	Quantitative random variables	11
1.25.2.1	Discrete random variable	11
1.25.2.2	Continuous random variable	11
1.26	Probability mass function (p.m.f.) of a univariate discrete random variable	12
1.27	Probability density function (p.d.f.) of a univariate continuous random variable	12
1.28	Expected value and variance of a univariate random variable	13
1.29	Distribution function of a univariate random variable	13
1.29.1	Discrete distribution function	14
1.29.2	Continuous distribution function	14
1.30	Selection of a sample using known univariate distribution function	15
1.30.1	Discrete random variable	15
1.30.2	Continuous random variable	17
1.31	Discrete bivariate random variable	19
1.32	Joint probability distribution function of bivariate discrete random variables	20
1.33	Joint cumulative distribution function of bivariate discrete random variables	20
1.34	Marginal distributions of a bivariate discrete random variable	20
1.35	Selection of a sample using known discrete bivariate distribution function	20
1.36	Continuous bivariate random variable	21
1.37	Joint probability distribution function of bivariate continuous random variable	21
1.38	Joint cumulative distribution function of a bivariate continuous random variable	22
1.39	Marginal cumulative distributions of bivariate continuous random variable	22
1.40	Selection of a sample using known bivariate continuous distribution function	22
1.41	Properties of a best estimator	24
1.41.1	Unbiasedness	24
1.41.1.1	Bias	28
1.41.2	Consistency	28
1.41.3	Sufficiency	28
1.41.4	Efficiency	29
1.41.4.1	Variance	29
1.41.4.2	Mean square error	29
1.42	Relative efficiency	29
1.43	Relative bias	29
1.44	Variance estimation through splitting	30
1.45	Loss function	31
1.46	Admissible estimator	31
1.47	Sample survey	31
1.48	Sampling distribution	32
1.49	Sampling frame	33

1.50	Sample survey design	33
1.51	Errors in the estimators	33
	1.51.1 Sampling errors	34
	1.51.2 Non-sampling errors	34
	1.51.2.1 Non-response errors	35
	1.51.2.2 Measurement errors	35
	1.51.2.3 Tabulation errors	35
	1.51.2.4 Computational errors	35
1.52	Point estimator	35
1.53	Interval estimator	35
1.54	Confidence interval	35
1.55	Population proportion	38
1.56	Sample proportion	38
1.57	Variance of sample proportion and confidence interval estimates	39
1.58	Relative standard error	50
1.59	Auxiliary information	50
1.60	Some useful mathematical formulae	56
1.61	Ordered statistics	57
	1.61.1 Population median	57
	1.61.2 Population quartiles	58
	1.61.3 Population percentiles	59
	1.61.4 Population mode	59
1.62	Definition(s) of statistics	59
1.63	Limitations of statistics	60
1.64	Lack of confidence in statistics	60
1.65	Scope of statistics	60
	Exercises	60
	Practical problems	63

2 SIMPLE RANDOM SAMPLING

2.0	Introduction	71
2.1	Simple random sampling with replacement	71
2.2	Simple random sampling without replacement	79
2.3	Estimation of population proportion	94
2.4	Searls'estimator of population mean	103
2.5	Use of distinct units in the WR sample at the estimation stage	106
	2.5.1 Estimation of mean	107
	2.5.2 Estimation of finite population variance	113
2.6	Estimation of total or mean of a subgroup (domain) of a population	118
2.7	Dealing with a rare attribute using inverse sampling	123
2.8	Controlled sampling	125
2.9	Determinant sampling	127
	Exercises	128
	Practical problems	132

3 USE OF AUXILIARY INFORMATION: SIMPLE RANDOM SAMPLING

3.0	Introduction	137
3.1	Notation and expected values	137
3.2	Estimation of population mean	138
3.2.1	Ratio estimator	138
3.2.2	Product estimator	145
3.2.3	Regression estimator	149
3.2.4	Power transformation estimator	160
3.2.5	A dual of ratio estimator	161
3.2.6	General class of estimators	164
3.2.7	Wider class of estimators	166
3.2.8	Use of known variance of auxiliary variable at estimation stage of population mean	167
3.2.8.1	A class of estimators	167
3.2.8.2	A wider class of estimators	169
3.2.9	Methods to remove bias from ratio and product type estimators	173
3.2.9.1	Quenouille's method	173
3.2.9.2	Interpenetrating sampling method	175
3.2.9.3	Exactly unbiased ratio type estimator	180
3.2.9.4	Unbiased product type estimator	183
3.2.9.5	Class of almost unbiased estimators of population ratio and product	185
3.2.9.6	Filtration of bias	187
3.3	Estimation of finite population variance	191
3.3.1	Ratio type estimator	192
3.3.2	Difference type estimator	197
3.3.3	Power transformation type estimator	198
3.3.4	General class of estimators	199
3.4	Estimation of regression coefficient	203
3.4.1	Usual estimator	203
3.4.2	Unbiased estimator	204
3.4.3	Improved estimators of regression coefficient	207
3.5	Estimation of finite population correlation coefficient	209
3.6	Superpopulation model approach	214
3.6.1	Relationship between linear model and regression estimator	214
3.6.2	Improved estimator of variance of linear regression estimator	217
3.6.3	Relationship between linear model and ratio estimator	221
3.7	Jackknife variance estimator	223
3.7.1	Ratio estimator	223
3.7.2	Regression estimator	226

3.8	Estimation of population mean using more than one auxiliary variable	229
3.8.1	Multivariate ratio estimator	230
3.8.2	Multivariate regression type estimators	231
3.8.3	General class of estimators	239
3.9	General class of estimators to estimate any population parameter	245
3.10	Estimation of ratio or product of two population means	248
3.11	Median estimation in survey sampling	250
	Exercises	257
	Practical problems	281

4 USE OF AUXILIARY INFORMATION: PROBABILITY PROPORTIONAL TO SIZE AND WITH REPLACEMENT (PPSWR) SAMPLING

4.0	Introduction	295
4.1	What is PPSWR sampling?	295
4.1.1	Cumulative total method	300
4.1.2	Lahiri's method	303
4.2	Estimation of population total	306
4.3	Relative efficiency of PPSWR sampling with respect to SRS WR sampling	312
4.3.1	Superpopulation model approach	312
4.3.2	Cost aspect	315
4.4	PPSWR sampling: More than one auxiliary variable is available	317
4.4.1	Notation and expectations	318
4.4.2	Class of estimators	319
4.4.3	Wider class of estimators	320
4.4.4	PPSWR sampling with negatively correlated variables	324
4.5	Multi-character survey	326
4.5.1	Study variables have poor positive correlation with the selection probabilities.	326
4.5.1.1	General class of estimators	335
4.5.2	Study variables have poor positive as well as poor negative correlation with the selection probabilities	336
4.6	Concept of revised selection probabilities	339
4.7	Estimation of correlation coefficient using PPSWR sampling	340
	Exercises	341
	Practical problems	345

5 USE OF AUXILIARY INFORMATION: PROBABILITY PROPORTIONAL TO SIZE AND WITHOUT REPLACEMENT (PPSWOR) SAMPLING

5.0	Introduction	349
	5.0.1 Useful symbols	349
	5.0.2 Some mathematical relations	349
5.1	Horvitz and Thompson estimator and related topics	351
5.2	General class of estimators	373
5.3	Model based estimation strategies	375
	5.3.1 A brief history of the superpopulation model	377
	5.3.2 Scott, Brewer and Ho's robust estimation strategy	378
	5.3.3 Design variance and anticipated variance of linear regression type estimator	383
5.4	Construction and optimal choice of inclusion probabilities	385
	5.4.1 Pareto π ps sampling estimation scheme	386
	5.4.2 Hanurav's method	387
	5.4.3 Brewer's method	388
	5.4.4 Sampford's method	389
	5.4.5 Narain's method	390
	5.4.6 Midzuno—Sen method	390
	5.4.7 Kumar-Gupta-Nigam scheme	391
	5.4.8 Dey and Srivastava scheme for even sample size	392
	5.4.9 SSS sampling scheme	393
	5.4.10 Optimal choice of first order inclusion probabilities	394
5.5	Calibration approach	399
5.6	Calibrated estimator of the variance of the estimator of population total	409
5.7	Estimation of variance of GREG	413
5.8	Improved estimator of variance of the GREG: The higher level calibration approach	419
	5.8.1 Recalibrated estimator of the variance of GREG	424
	5.8.2 Recalibration using optimal designs for the GREG	426
5.9	Calibrated estimators of variance of estimator of total and distribution function	428
	5.9.1 Unified setup	430
5.10	Calibration of estimator of variance of regression predictor	431
	5.10.1 Chaudhuri and Roy's results	433
	5.10.2 Calibrated estimators of variance of regression predictor	436
	5.10.2.1 Model assisted calibration	436
	5.10.2.2 Calibration estimators when variance of auxiliary variable is known	440
	5.10.2.2.1 Each component of V_x is known	441
	5.10.2.2.2 Compromized calibration	442
	5.10.2.3 Prediction variance	444

5.11	Ordered and unordered estimators	444
5.11.1	Ordered estimators	445
5.11.2	Unordered estimators	449
5.12	Rao--Hartley--Cochran (RHC) sampling strategy	452
5.13	Unbiased strategies using IPPS sampling schemes	462
5.13.1	Estimation of population mean using a ratio estimator	462
5.13.2	Estimation of finite population variance	464
5.14	Godambe's strategy: Estimation of parameters in survey sampling	465
5.14.1	Optimal estimating function	470
5.14.2	Regression type estimators	472
5.14.3	Singh's strategy in two-dimensional space	473
5.14.4	Godambe's strategy for linear Bayes and optimal estimation	476
5.15	Unified theory of survey sampling	479
5.15.1	Class of admissible estimators	479
5.15.2	Estimator	479
5.15.3	Admissible estimator	479
5.15.4	Strictly admissible estimator	479
5.15.5	Linear estimators of population total	483
5.15.6	Admissible estimators of variances of estimators of total	485
5.15.6.1	Condition for the unbiased estimator of variance	485
5.15.6.2	Admissible and unbiased estimator of variance	485
5.15.6.3	Fixed size sampling design	485
5.15.6.4	Horvitz and Thompson estimator and its variance in two forms	485
5.15.7	Polynomial type estimators	489
5.15.8	Alternative optimality criterion	490
5.15.9	Sufficient statistic in survey sampling	491
5.16	Estimators based on conditional inclusion probabilities	493
5.17	Current topics in survey sampling	494
5.17.1	Survey design	495
5.17.2	Data collection and processing	495
5.17.3	Estimation and analysis of data	496
5.18	Miscellaneous discussions/topics	497
5.18.1	Generalized IPPS designs	497
5.18.2	Tam's optimal strategies	498
5.18.3	Use of ranks in sample selection	498
5.18.4	Prediction approach	498
5.18.5	Total of bottom (or top) percentiles of a finite population	499
5.18.6	General form of estimator of variance	499
5.18.7	Poisson sampling	499
5.18.8	Cosmetic calibration	500
5.18.9	Mixing of non-parametric models in survey sampling	501
5.19	Golden Jubilee Year 2003 of the linear regression estimator	504
	Exercises	507
	Practical Problems	520

6 USE OF AUXILIARY INFORMATION: MULTI-PHASE SAMPLING

6.0	Introduction	529
6.1	SRS WOR scheme at the first as well as at the second phases of the sample selection	530
6.1.0	Notation and expected values	530
6.1.1	Ratio estimator	532
6.1.1.1	Cost function	535
6.1.2	Difference estimator	539
6.1.3	Regression estimator	540
6.1.4	General class of estimators of population mean	541
6.1.5	Estimation of finite population variance	544
6.1.6	Calibration approach in two-phase sampling	545
6.2	Two-phase sampling using two auxiliary variables	549
6.3	Chain ratio type estimators	554
6.4	Calibration using two auxiliary variables	555
6.5	Estimation of variance of calibrated estimator in two-phase sampling: low and higher level calibration	560
6.6	Two-phase sampling using multi-auxiliary variables	563
6.7	Unified approach in two-phase sampling	563
6.8	Concept of three-phase sampling	565
6.9	Estimation of variance of regression estimator under two-phase sampling	567
6.10	Two-phase sampling using PPSWR sampling	572
6.11	Concept of dual frame surveys	576
6.11.1	Common variables used for further calibration of weights	576
6.11.2	Estimation of variance using dual frame surveys	577
6.12	Estimation of median using two-phase sampling	578
6.12.1	General class of estimators	578
6.12.2	Regression type estimator	579
6.12.3	Position estimator	581
6.12.4	Stratification estimator	582
6.12.5	Optimum first and second phase samples for median estimation	584
6.12.5.1	Cost is fixed	584
6.12.5.2	Variance is fixed	584
6.12.6	Kuk and Mak's technique in two-phase sampling	584
6.12.7	Chen and Qin technique in two-phase sampling	586
6.13	Distribution function with two-phase sampling	588
6.14	Improved version of two-phase calibration approach	590
6.14.1	Improved first phase calibration	590
6.14.2	Improved second phase calibration	592
	Exercises	594
	Practical problems	612

6 USE OF AUXILIARY INFORMATION: MULTI-PHASE SAMPLING

6.0	Introduction	529
6.1	SRS WOR scheme at the first as well as at the second phases of the sample selection	530
6.1.0	Notation and expected values	530
6.1.1	Ratio estimator	532
6.1.1.1	Cost function	535
6.1.2	Difference estimator	539
6.1.3	Regression estimator	540
6.1.4	General class of estimators of population mean	541
6.1.5	Estimation of finite population variance	544
6.1.6	Calibration approach in two-phase sampling	545
6.2	Two-phase sampling using two auxiliary variables	549
6.3	Chain ratio type estimators	554
6.4	Calibration using two auxiliary variables	555
6.5	Estimation of variance of calibrated estimator in two-phase sampling: low and higher level calibration	560
6.6	Two-phase sampling using multi-auxiliary variables	563
6.7	Unified approach in two-phase sampling	563
6.8	Concept of three-phase sampling	565
6.9	Estimation of variance of regression estimator under two-phase sampling	567
6.10	Two-phase sampling using PPSWR sampling	572
6.11	Concept of dual frame surveys	576
6.11.1	Common variables used for further calibration of weights	576
6.11.2	Estimation of variance using dual frame surveys	577
6.12	Estimation of median using two-phase sampling	578
6.12.1	General class of estimators	578
6.12.2	Regression type estimator	579
6.12.3	Position estimator	581
6.12.4	Stratification estimator	582
6.12.5	Optimum first and second phase samples for median estimation	584
6.12.5.1	Cost is fixed	584
6.12.5.2	Variance is fixed	584
6.12.6	Kuk and Mak's technique in two-phase sampling	584
6.12.7	Chen and Qin technique in two-phase sampling	586
6.13	Distribution function with two-phase sampling	588
6.14	Improved version of two-phase calibration approach	590
6.14.1	Improved first phase calibration	590
6.14.2	Improved second phase calibration	592
	Exercises	594
	Practical problems	612

Advanced Sampling Theory with Applications

How Michael 'selected' Amy
Volume II

by

Sarjinder Singh

*St. Cloud State University,
Department of Statistics,
St. Cloud, MN, U.S.A.*

KLUWER ACADEMIC PUBLISHERS

DORDRECHT / BOSTON / LONDON

TABLE OF CONTENTS

PREFACE

xxi

1 BASIC CONCEPTS AND MATHEMATICAL NOTATION

1.0	Introduction	1
1.1	Population	1
1.1.1	Finite population	1
1.1.2	Infinite population	1
1.1.3	Target population	1
1.1.4	Study population	1
1.2	Sample	2
1.3	Examples of populations and samples	2
1.4	Census	2
1.5	Relative aspects of sampling versus census	2
1.6	Study variable	2
1.7	Auxiliary variable	3
1.8	Difference between study variable and auxiliary variable	3
1.9	Parameter	3
1.10	Statistic	3
1.11	Statistics	4
1.12	Sample selection	4
1.12.1	Chit method or Lottery method	4
1.12.1.1	With replacement sampling	4
1.12.1.2	Without replacement sampling	5
1.12.2	Random number table method	5
1.12.2.1	Remainder method	6
1.13	Probability sampling	7
1.14	Probability of selecting a sample	7
1.15	Population mean/total	8
1.16	Population moments	8
1.17	Population standard deviation	8
1.18	Population coefficient of variation	8
1.19	Relative mean square error	9
1.20	Sample mean	9
1.21	Sample variance	9
1.22	Estimator	10
1.23	Estimate	10
1.24	Sample space	10
1.25	Univariate random variable	11
1.25.1	Qualitative random variables	11

1.25.2	Quantitative random variables	11
1.25.2.1	Discrete random variable	11
1.25.2.2	Continuous random variable	11
1.26	Probability mass function (p.m.f.) of a univariate discrete random variable	12
1.27	Probability density function (p.d.f.) of a univariate continuous random variable	12
1.28	Expected value and variance of a univariate random variable	13
1.29	Distribution function of a univariate random variable	13
1.29.1	Discrete distribution function	14
1.29.2	Continuous distribution function	14
1.30	Selection of a sample using known univariate distribution function	15
1.30.1	Discrete random variable	15
1.30.2	Continuous random variable	17
1.31	Discrete bivariate random variable	19
1.32	Joint probability distribution function of bivariate discrete random variables	20
1.33	Joint cumulative distribution function of bivariate discrete random variables	20
1.34	Marginal distributions of a bivariate discrete random variable	20
1.35	Selection of a sample using known discrete bivariate distribution function	20
1.36	Continuous bivariate random variable	21
1.37	Joint probability distribution function of bivariate continuous random variable	21
1.38	Joint cumulative distribution function of a bivariate continuous random variable	22
1.39	Marginal cumulative distributions of bivariate continuous random variable	22
1.40	Selection of a sample using known bivariate continuous distribution function	22
1.41	Properties of a best estimator	24
1.41.1	Unbiasedness	24
1.41.1.1	Bias	28
1.41.2	Consistency	28
1.41.3	Sufficiency	28
1.41.4	Efficiency	29
1.41.4.1	Variance	29
1.41.4.2	Mean square error	29
1.42	Relative efficiency	29
1.43	Relative bias	29
1.44	Variance estimation through splitting	30
1.45	Loss function	31
1.46	Admissible estimator	31
1.47	Sample survey	31
1.48	Sampling distribution	32
1.49	Sampling frame	33

1.50	Sample survey design	33
1.51	Errors in the estimators	33
1.51.1	Sampling errors	34
1.51.2	Non-sampling errors	34
1.51.2.1	Non-response errors	35
1.51.2.2	Measurement errors	35
1.51.2.3	Tabulation errors	35
1.51.2.4	Computational errors	35
1.52	Point estimator	35
1.53	Interval estimator	35
1.54	Confidence interval	35
1.55	Population proportion	38
1.56	Sample proportion	38
1.57	Variance of sample proportion and confidence interval estimates	39
1.58	Relative standard error	50
1.59	Auxiliary information	50
1.60	Some useful mathematical formulae	56
1.61	Ordered statistics	57
1.61.1	Population median	57
1.61.2	Population quartiles	58
1.61.3	Population percentiles	59
1.61.4	Population mode	59
1.62	Definition(s) of statistics	59
1.63	Limitations of statistics	60
1.64	Lack of confidence in statistics	60
1.65	Scope of statistics	60
	Exercises	60
	Practical problems	63

2 SIMPLE RANDOM SAMPLING

2.0	Introduction	71
2.1	Simple random sampling with replacement	71
2.2	Simple random sampling without replacement	79
2.3	Estimation of population proportion	94
2.4	Searls'estimator of population mean	103
2.5	Use of distinct units in the WR sample at the estimation stage	106
2.5.1	Estimation of mean	107
2.5.2	Estimation of finite population variance	113
2.6	Estimation of total or mean of a subgroup (domain) of a population	118
2.7	Dealing with a rare attribute using inverse sampling	123
2.8	Controlled sampling	125
2.9	Determinant sampling	127
	Exercises	128
	Practical problems	132

3 USE OF AUXILIARY INFORMATION: SIMPLE RANDOM SAMPLING

3.0	Introduction	137
3.1	Notation and expected values	137
3.2	Estimation of population mean	138
3.2.1	Ratio estimator	138
3.2.2	Product estimator	145
3.2.3	Regression estimator	149
3.2.4	Power transformation estimator	160
3.2.5	A dual of ratio estimator	161
3.2.6	General class of estimators	164
3.2.7	Wider class of estimators	166
3.2.8	Use of known variance of auxiliary variable at estimation stage of population mean	167
3.2.8.1	A class of estimators	167
3.2.8.2	A wider class of estimators	169
3.2.9	Methods to remove bias from ratio and product type estimators	173
3.2.9.1	Quenouille's method	173
3.2.9.2	Interpenetrating sampling method	175
3.2.9.3	Exactly unbiased ratio type estimator	180
3.2.9.4	Unbiased product type estimator	183
3.2.9.5	Class of almost unbiased estimators of population ratio and product	185
3.2.9.6	Filtration of bias	187
3.3	Estimation of finite population variance	191
3.3.1	Ratio type estimator	192
3.3.2	Difference type estimator	197
3.3.3	Power transformation type estimator	198
3.3.4	General class of estimators	199
3.4	Estimation of regression coefficient	203
3.4.1	Usual estimator	203
3.4.2	Unbiased estimator	204
3.4.3	Improved estimators of regression coefficient	207
3.5	Estimation of finite population correlation coefficient	209
3.6	Superpopulation model approach	214
3.6.1	Relationship between linear model and regression estimator	214
3.6.2	Improved estimator of variance of linear regression estimator	217
3.6.3	Relationship between linear model and ratio estimator	221
3.7	Jackknife variance estimator	223
3.7.1	Ratio estimator	223
3.7.2	Regression estimator	226

3.8	Estimation of population mean using more than one auxiliary variable	229
3.8.1	Multivariate ratio estimator	230
3.8.2	Multivariate regression type estimators	231
3.8.3	General class of estimators	239
3.9	General class of estimators to estimate any population parameter	245
3.10	Estimation of ratio or product of two population means	248
3.11	Median estimation in survey sampling	250
	Exercises	257
	Practical problems	281

4 USE OF AUXILIARY INFORMATION: PROBABILITY PROPORTIONAL TO SIZE AND WITH REPLACEMENT (PPSWR) SAMPLING

4.0	Introduction	295
4.1	What is PPSWR sampling?	295
4.1.1	Cumulative total method	300
4.1.2	Lahiri's method	303
4.2	Estimation of population total	306
4.3	Relative efficiency of PPSWR sampling with respect to SRS WR sampling	312
4.3.1	Superpopulation model approach	312
4.3.2	Cost aspect	315
4.4	PPSWR sampling: More than one auxiliary variable is available	317
4.4.1	Notation and expectations	318
4.4.2	Class of estimators	319
4.4.3	Wider class of estimators	320
4.4.4	PPSWR sampling with negatively correlated variables	324
4.5	Multi-character survey	326
4.5.1	Study variables have poor positive correlation with the selection probabilities.	326
4.5.1.1	General class of estimators	335
4.5.2	Study variables have poor positive as well as poor negative correlation with the selection probabilities	336
4.6	Concept of revised selection probabilities	339
4.7	Estimation of correlation coefficient using PPSWR sampling	340
	Exercises	341
	Practical problems	345

5 USE OF AUXILIARY INFORMATION: PROBABILITY PROPORTIONAL TO SIZE AND WITHOUT REPLACEMENT (PPSWOR) SAMPLING

5.0	Introduction	349
	5.0.1 Useful symbols	349
	5.0.2 Some mathematical relations	349
5.1	Horvitz and Thompson estimator and related topics	351
5.2	General class of estimators	373
5.3	Model based estimation strategies	375
	5.3.1 A brief history of the superpopulation model	377
	5.3.2 Scott, Brewer and Ho's robust estimation strategy	378
	5.3.3 Design variance and anticipated variance of linear regression type estimator	383
5.4	Construction and optimal choice of inclusion probabilities	385
	5.4.1 Pareto π ps sampling estimation scheme	386
	5.4.2 Hanurav's method	387
	5.4.3 Brewer's method	388
	5.4.4 Sampford's method	389
	5.4.5 Narain's method	390
	5.4.6 Midzuno--Sen method	390
	5.4.7 Kumar-Gupta-Nigam scheme	391
	5.4.8 Dey and Srivastava scheme for even sample size	392
	5.4.9 SSS sampling scheme	393
	5.4.10 Optimal choice of first order inclusion probabilities	394
5.5	Calibration approach	399
5.6	Calibrated estimator of the variance of the estimator of population total	409
5.7	Estimation of variance of GREG	413
5.8	Improved estimator of variance of the GREG: The higher level calibration approach	419
	5.8.1 Recalibrated estimator of the variance of GREG	424
	5.8.2 Recalibration using optimal designs for the GREG	426
5.9	Calibrated estimators of variance of estimator of total and distribution function	428
	5.9.1 Unified setup	430
5.10	Calibration of estimator of variance of regression predictor	431
	5.10.1 Chaudhuri and Roy's results	433
	5.10.2 Calibrated estimators of variance of regression predictor	436
	5.10.2.1 Model assisted calibration	436
	5.10.2.2 Calibration estimators when variance of auxiliary variable is known	440
	5.10.2.2.1 Each component of V_x is known	441
	5.10.2.2.2 Compromized calibration	442
	5.10.2.3 Prediction variance	444

5.11	Ordered and unordered estimators	444
	5.11.1 Ordered estimators	445
	5.11.2 Unordered estimators	449
5.12	Rao~Hartley~Cochran (RHC) sampling strategy	452
5.13	Unbiased strategies using IPPS sampling schemes	462
	5.13.1 Estimation of population mean using a ratio estimator	462
	5.13.2 Estimation of finite population variance	464
5.14	Godambe's strategy: Estimation of parameters in survey sampling	465
	5.14.1 Optimal estimating function	470
	5.14.2 Regression type estimators	472
	5.14.3 Singh's strategy in two-dimensional space	473
	5.14.4 Godambe's strategy for linear Bayes and optimal estimation	476
5.15	Unified theory of survey sampling	479
	5.15.1 Class of admissible estimators	479
	5.15.2 Estimator	479
	5.15.3 Admissible estimator	479
	5.15.4 Strictly admissible estimator	479
	5.15.5 Linear estimators of population total	483
	5.15.6 Admissible estimators of variances of estimators of total	485
	5.15.6.1 Condition for the unbiased estimator of variance	485
	5.15.6.2 Admissible and unbiased estimator of variance	485
	5.15.6.3 Fixed size sampling design	485
	5.15.6.4 Horvitz and Thompson estimator and its variance in two forms	485
	5.15.7 Polynomial type estimators	489
	5.15.8 Alternative optimality criterion	490
	5.15.9 Sufficient statistic in survey sampling	491
5.16	Estimators based on conditional inclusion probabilities	493
5.17	Current topics in survey sampling	494
	5.17.1 Survey design	495
	5.17.2 Data collection and processing	495
	5.17.3 Estimation and analysis of data	496
5.18	Miscellaneous discussions/topics	497
	5.18.1 Generalized IPPS designs	497
	5.18.2 Tarn's optimal strategies	498
	5.18.3 Use of ranks in sample selection	498
	5.18.4 Prediction approach	498
	5.18.5 Total of bottom (or top) percentiles of a finite population	499
	5.18.6 General form of estimator of variance	499
	5.18.7 Poisson sampling	499
	5.18.8 Cosmetic calibration	500
	5.18.9 Mixing of non-parametric models in survey sampling	501
5.19	Golden Jubilee Year 2003 of the linear regression estimator	504
	Exercises	507
	Practical Problems	520

6 USE OF AUXILIARY INFORMATION: MULTI-PHASE SAMPLING

6.0	Introduction	529
6.1	SRS WOR scheme at the first as well as at the second phases of the sample selection	530
6.1.0	Notation and expected values	530
6.1.1	Ratio estimator	532
6.1.1.1	Cost function	535
6.1.2	Difference estimator	539
6.1.3	Regression estimator	540
6.1.4	General class of estimators of population mean	541
6.1.5	Estimation of finite population variance	544
6.1.6	Calibration approach in two-phase sampling	545
6.2	Two-phase sampling using two auxiliary variables	549
6.3	Chain ratio type estimators	554
6.4	Calibration using two auxiliary variables	555
6.5	Estimation of variance of calibrated estimator in two-phase sampling: low and higher level calibration	560
6.6	Two-phase sampling using multi-auxiliary variables	563
6.7	Unified approach in two-phase sampling	563
6.8	Concept of three-phase sampling	565
6.9	Estimation of variance of regression estimator under two-phase sampling	567
6.10	Two-phase sampling using PPSWR sampling	572
6.11	Concept of dual frame surveys	576
6.11.1	Common variables used for further calibration of weights	576
6.11.2	Estimation of variance using dual frame surveys	577
6.12	Estimation of median using two-phase sampling	578
6.12.1	General class of estimators	578
6.12.2	Regression type estimator	579
6.12.3	Position estimator	581
6.12.4	Stratification estimator	582
6.12.5	Optimum first and second phase samples for median estimation	584
6.12.5.1	Cost is fixed	584
6.12.5.2	Variance is fixed	584
6.12.6	Kuk and Mak's technique in two-phase sampling	584
6.12.7	Chen and Qin technique in two-phase sampling	586
6.13	Distribution function with two-phase sampling	588
6.14	Improved version of two-phase calibration approach	590
6.14.1	Improved first phase calibration	590
6.14.2	Improved second phase calibration	592
	Exercises	594
	Practical problems	612

VOLUME II

7 SYSTEMATIC SAMPLING

7.0	Introduction	615
7.1	Systematic sampling	615
7.2	Modified systematic sampling	620
7.3	Circular systematic sampling	621
7.4	PPS circular systematic sampling	623
7.5	Estimation of variance under systematic sampling	624
	7.5.1 Sub-sampling or replicated sub-sampling scheme	625
	7.5.2 Successive differences	626
	7.5.3 Variance of circular systematic sampling	627
7.6	Systematic sampling in population with linear trend	627
	7.6.1 Estimators with linear trend	627
	7.6.2 Modification of estimates	629
	7.6.3 Estimators based on centrally located samples	631
	7.6.4 Estimators based on balanced systematic sampling	633
7.7	Singh and Singh's systematic sampling scheme	635
7.8	Zinger strategy in systematic sampling	637
7.9	Populations with cyclic or periodic trends	638
7.10	Multi-dimensional systematic sampling	639
	Exercises	642
	Practical problems	646

8 STRATIFIED AND POST-STRATIFIED SAMPLING

8.0	Introduction	649
8.1	Stratified sampling	650
8.2	Different methods of sample allocation	659
	8.2.1 Equal allocation	659
	8.2.2 Proportional allocation	659
	8.2.3 Optimum allocation method	662
8.3	Use of auxiliary information at estimation stage	676
	8.3.1 Separate ratio estimator	677
	8.3.2 Separate regression estimator	681
	8.3.3 Combined ratio estimator	684
	8.3.4 Combined regression estimator	688
	8.3.5 On degree of freedom in stratified random sampling	693
8.4	Calibration approach for stratified sampling design	696
	8.4.1 Exact combined linear regression using calibration	700
8.5	Construction of strata boundaries	701
	8.5.1 Strata boundaries for proportional allocation	702
	8.5.2 Strata boundaries for Neyman allocation	703
	8.5.3 Stratification using auxiliary information	708
8.6	Superpopulation model approach	712
8.7	Multi-way stratification	713

8.8	Stratum boundaries for multi-variate populations	718
8.9	Optimum allocation in multi-variate stratified sampling	723
8.10	Stratification using two-phase sampling	726
8.11	Post-stratified sampling	729
	8.11.1 Conditional post-stratification	730
	8.11.2 Unconditional post-stratification	731
8.12	Estimation of proportion using stratified random sampling	735
	Exercises	738
	Practical problems	748

9 NON-OVERLAPPING, OVERLAPPING, POST, AND ADAPTIVE CLUSTER SAMPLING

9.0	Introduction	765
9.1	Non-overlapping clusters of equal size	766
9.2	Optimum value of non-overlapping cluster size	790
9.3	Estimation of proportion using non-overlapping cluster sampling	792
9.4	Non-overlapping clusters of different sizes	796
9.5	Selection of non-overlapping clusters with unequal probability sampling	805
9.6	Optimal and robust strategies for non-overlapping cluster sampling	808
9.7	Overlapping cluster sampling	812
	9.7.1 Population size is known	812
	9.7.2 Population size is unknown	814
9.8	Post-cluster sampling	817
9.9	Adaptive cluster sampling	819
	Exercises	820
	Practical problems	822

10 MULTI-STAGE, SUCCESSIVE, AND RE-SAMPLING STRATEGIES

10.0	Introduction	829
10.1	Notation	830
10.2	Procedure for construction of estimators of the total	831
10.3	Method of calculating the variance of the estimators	833
	10.3.1 Selection of first and second stage units using SRS WOR sampling	834
	10.3.2 Optimum allocation in two-stage sampling	836
10.4	Optimum allocation of sample in three-stage sampling	837
10.5	Modified three-stage sampling	838
10.6	General class of estimators in two-stage sampling	839
10.7	Prediction estimator under two-stage sampling	842
10.8	Prediction approach to robust variance estimation in two-stage cluster sampling	844

10.8.1	Royall's technique of variance estimation	846
10.9	Two-stage sampling with successive occasions	847
10.9.1	Arnab's successive sampling scheme	848
10.10	Estimation strategies in supplemented panels	865
10.11	Re-sampling methods	866
10.11.1	Jackknife variance estimator	867
10.11.2	Balanced half sample (BHS) method	871
10.11.3	Bootstrap variance estimator	873
	Exercises	873
	Practical problems	887

11 RANDOMIZED RESPONSE SAMPLING: TOOLS FOR SOCIAL SURVEYS

11.0	Introduction	889
11.1	Pioneer model	889
11.2	Franklin's model	892
11.3	Unrelated question model and related issues	897
11.3.1	When proportion of unrelated character is known	897
11.3.2	When proportion of unrelated character is unknown	898
11.4	Regression analysis	903
11.4.1	Ridge regression estimator	905
11.5	Hidden gangs in finite populations	907
11.5.1	Two sample method	907
11.5.2	One sample method	911
11.5.3	Estimation of correlation coefficient between two characters of a hidden gang	912
11.6	Unified approach for hidden gangs	916
11.7	Randomized response technique for a quantitative variable	920
11.8	GREG using scrambled responses	924
11.8.1	Calibration of scrambled responses	925
11.8.2	Higher order calibration of the estimators of variance under scrambled responses	928
11.8.3	General class of estimators	930
11.9	On respondent's protection: Qualitative characters	930
11.9.1	Leysieffer and Warner's measure	930
11.9.2	Lanke's measure	932
11.9.3	Mangat and Singh's two-stage model	933
11.9.4	Mangat and Singh's two-stage and Warner's model at equal level of protection	935
11.9.5	Mangat's model	939
11.9.6	Mangat's and Warner's model at equal of protection	940
11.10	On respondent's protection: Quantitative characters	942
11.10.1	Unrelated question model for quantitative data	942
11.10.2	The additive model	943
11.10.3	The multiplicative model	943

11.10.4	Measure of privacy protection	944
11.10.5	Comparison between additive and multiplicative models	945
11.11	Test for detecting untruthful answering	949
11.12	Stochastic randomized response technique	951
	Exercises	954
	Practical problems	972

12 NON-RESPONSE AND ITS TREATMENTS

12.0	Introduction	975
12.1	Hansen and Hurwitz pioneer model	976
12.2	Politz and Simmons model	980
12.3	Horvitz and Thompson estimator under non-response	984
12.4	Ratio and regression type estimators	986
	12.4.1 Distribution and some expected values	987
	12.4.2 Estimation of population mean	987
	12.4.3 Estimation of finite population variance	993
12.5	Calibrated estimators of total and variance in the presence of non-response	1000
	12.5.1 Estimation of population total and variance	1000
	12.5.2 Calibration estimator for the total	1002
	12.5.3 Calibration of the estimators of variance	1003
	12.5.3.1 PPSWOR Sampling	1005
	12.5.3.2 SRSWOR Sampling	1007
12.6	Different treatments of non-response	1009
	12.6.1 Ratio method of imputation	1010
	12.6.2 Mean method of imputation	1010
	12.6.3 Hot deck (HD) method of imputation	1010
	12.6.4 Nearest neighbor (NN) method of imputation	1011
12.7	Superpopulation model approach	1013
	12.7.1 Different components of variance	1014
12.8	Jackknife technique	1016
12.9	Hot deck imputation for multi-stage designs	1017
12.10	Multiple imputation	1021
	12.10.1 Degree of freedom with multiple imputation for small samples	1024
12.11	Compromised imputation	1025
	12.11.1 Practicability of compromised imputation	1027
	12.11.2 Recommendations of compromised imputation	1027
	12.11.3 Warm deck imputation	1028
	12.11.4 Mean cum NN imputation	1028
12.12	Estimation of response probabilities	1031
12.13	Estimators based on estimated response probabilities	1033
	12.13.1 Estimators based on response probabilities	1035
	12.13.2 Calibration of response probabilities	1037
	12.13.2.1 Calibrated estimator and its variance	1038

12.13.2.2 Estimation of variance of the calibrated estimator	1039
Exercises	1041
Practical problems	1058

13 MISCELLANEOUS TOPICS

13.0 Introduction	1065
13.1 Estimation of measurement errors	1065
13.1.1 Estimation of measurement error using a single measurement per element	1066
13.1.1.1 Model and notation	1066
13.1.1.2 Grubbs'estimators	1066
13.1.2 Bhatia, Mangat, and Morrison's (BMM) repeated measurement estimators	1068
13.1.2.1 Model and notation	1069
13.2 Raking ratio using contingency tables	1073
13.3 Continuous populations	1077
13.4 Small area estimation	1081
13.4.1 Symptomatic accounting techniques	1081
13.4.2 Vital rates method (VRM)	1081
13.4.3 Census component method (CCM)	1082
13.4.4 Housing unit method (HUM)	1083
13.4.5 Synthetic estimator	1083
13.4.6 Composite estimator	1086
13.4.7 Model based techniques	1090
13.4.7.1 Henderson's model	1090
13.4.7.2 Nested error regression model	1093
13.4.7.3 Random regression coefficient model	1095
13.4.7.4 Fay and Herriot model	1097
13.4.8 Further generalizations	1097
13.4.9 Estimation of proportion of a characteristic in small areas of a population	1099
Exercises	1101
Practical problems	1101

APPENDIX

TABLES

1 Pseudo-Random Numbers (PRN)	1105
2 Critical values based on t distribution	1107
3 Area under the standard normal curve	1109

POPULATIONS

All operating banks: Amount (in \$000) of agricultural loans outstanding in different states in 1997	1111
Hypothetical situation of a small village having only 30 older persons (age more than 50 years): Approximate duration of sleep (in minutes) and age (in years) of the persons	1113
Apples, commercial crop: Season average price (in \$) per pound, by States, 1994—1996	1114
Fish caught: Estimated number of fish caught by marine recreational fishermen by species group and year, Atlantic and Gulf coasts, 1992—1995	1116
Tobacco: Area (hectares), yield and production (metric tons) in specified countries during 1998	1119
Age specific death rates from 1990 to 2065 (Number per 100,000 births)	1123
State population projections, 1995 and 2000 (Number in thousands)	1124
Projected vital statistics by country or area during 2000	1126
Number of immigrants admitted to the USA	1129

BIBLIOGRAPHY

1131

AUTHOR INDEX

1193

HANDY SUBJECT INDEX

1215

ADDITIONAL INFORMATION

1219