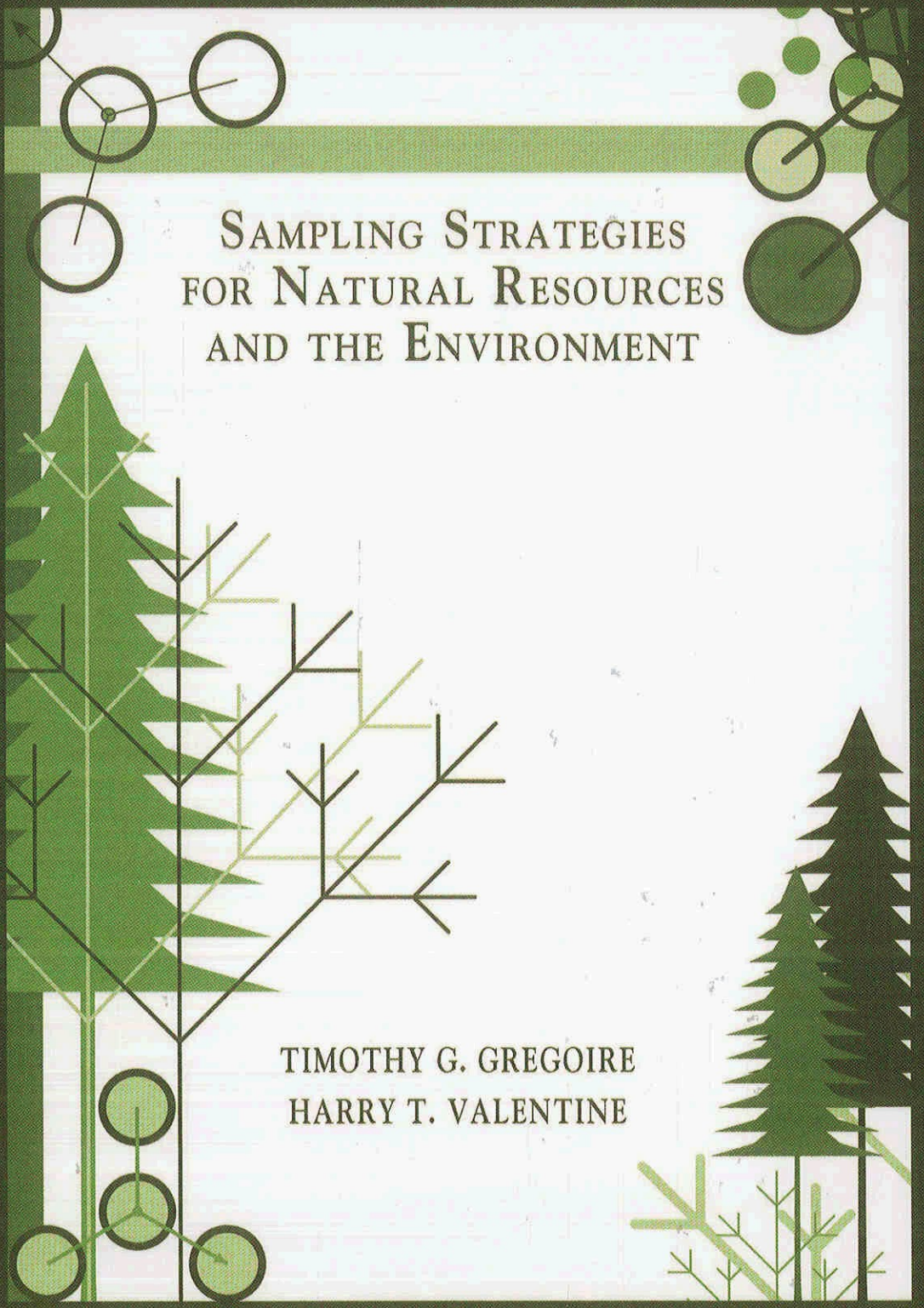


CHAPMAN & HALL/CRC
APPLIED ENVIRONMENTAL STATISTICS



SAMPLING STRATEGIES
FOR NATURAL RESOURCES
AND THE ENVIRONMENT

TIMOTHY G. GREGOIRE
HARRY T. VALENTINE

 Chapman & Hall/CRC
Taylor & Francis Group

Contents

Preface

xvii

1	Introduction	1
1.1	The need for sampling strategies	1
1.2	A medley of sampling scenarios	3
1.2.1	Sampling for biomass estimation	3
1.2.2	Sampling to estimate composition	3
1.2.3	Sampling for cover information	4
1.2.4	How many trees in Sweden?	4
1.3	Probability sample	5
1.3.1	Selection probability of a sample	5
1.3.2	Inclusion probability	7
1.3.3	Sampling frame	8
1.4	Inference	9
1.5	Population descriptive parameters	10
1.5.1	Discrete populations	10
1.5.2	Continuous populations	12
1.6	Historical note	13
1.7	Terms to remember	14
1.8	Appendix	14
1.8.1	Alternative expressions for σ_y^2	14
2	Sampling Distribution of an Estimator	15
2.1	Distribution of values	15
2.2	Estimation	18
2.2.1	Consistency	19
2.2.2	Expected value	21
2.2.3	Bias	24
2.2.4	Variance	25
2.2.5	Standard error	27
2.2.6	Mean square error	27
2.3	Interval estimation	28
2.4	The role of simulated sampling	32
2.5	Other considerations	32
2.6	Terms to remember	33
2.7	Exercises	33

2.8	Appendix	34
2.8.1	Derivation of the relationship between mean square error, variance, and squared bias	34
3	Sampling Designs for Discrete Populations	35
3.1	Introduction	35
3.2	Equal probability designs	35
3.2.1	Simple random sampling	35
	Simple random sampling without replacement	36
	Selecting a SRSwoR sample	42
	Simple random sampling with replacement	44
	Selecting a SRSwR sample	48
3.2.2	Systematic sampling	49
	Circular systematic sampling	56
3.2.3	Bernoulli sampling	57
3.3	Unequal probability designs	61
3.3.1	List sampling	62
	Manipulating selection probabilities to increase precision	63
3.3.2	Poisson sampling	68
3.3.3	Unequal probability systematic sampling	71
3.3.4	Rao, Hartley, Cochran sampling strategy	74
3.4	Terms to remember	77
3.5	Exercises	77
3.6	Appendix	79
3.6.1	Factorial and combinatorial notation	79
3.6.2	Derivation of the inclusion probability of an element for SRSwoR.	79
3.6.3	Proof of unbiasedness of $\hat{\tau}_{y\pi}$ as an estimator of τ_y	80
3.6.4	Derivation of $V[N\bar{y}]$ in (3.6) following SRSwoR	81
3.6.5	Proof of unbiasedness of s_y^2 as estimator of σ_y^2 following SRSwoR	83
3.6.6	Derivation of the inclusion probability of an element for SRSwR	84
3.6.7	Derivation of $E[n]$ and $V[n]$ following Bernoulli sampling	84
3.6.8	Derivation of the expected value and variance of $\hat{\tau}_{yp}$	85
3.6.9	Proof of the unbiasedness of $\hat{v}[\hat{\tau}_{yp}]$ as an estimator of $V[\hat{\tau}_{yp}]$	89
3.6.10	Variance of estimated proportions	89
3.6.11	Nearly exact confidence intervals for proportions	90
3.6.12	Expected sample size	91
4	Sampling Designs for Continuous Populations	93
4.1	Introduction	93
4.2	Crude Monte Carlo	94

4.2.1	Definitions	95
4.2.2	Selection	97
4.2.3	Estimation	98
4.2.4	Antithetic variates	102
4.2.5	Systematic selection	104
4.3	Importance sampling	106
4.3.1	Proxy function	107
4.3.2	Selection by the inverse-transform method	108
4.3.3	Selection by the acceptance-rejection method	109
4.3.4	Estimation	110
4.4	Control variate estimation	114
4.5	Sampling in two or three dimensions	116
4.5.1	Selection of a sample point	117
4.5.2	Estimation	118
4.5.3	Systematic selection	119
4.5.4	Three dimensions	120
4.6	General notation	121
4.7	Terms to remember	121
4.8	Exercises and projects	122
4.9	Appendix	123
4.9.1	Proof of the unbiasedness of $\hat{\tau}_\rho$ as an estimator of τ_ρ	124
4.9.2	Derivation of $V[\hat{\tau}_\rho]$	125
4.9.3	The Horvitz-Thompson estimator of τ_ρ and its variance	126
5	Stratified Sampling Designs	127
5.1	Introduction	127
5.2	Rationale for stratified sampling	127
5.3	Estimation with stratified sampling	129
5.3.1	Notation	129
5.3.2	HT estimation	130
5.3.3	More general estimation	134
5.3.4	Stratified random sampling	137
5.4	Sample allocation among strata	140
5.4.1	Equal allocation	140
5.4.2	Proportional allocation	141
5.4.3	x-Proportional allocation	143
5.4.4	Optimal allocation, equal sampling costs	143
5.4.5	Optimal allocation, unequal sampling costs	145
5.4.6	Power allocation	145
5.4.7	Allocation for multiresource surveys	146
5.4.8	Comparison of allocation rules	147
5.5	Incorrect assignment of population elements into strata	148
5.6	Double sampling for stratification	149
5.7	Poststratification	152
5.7.1	Preliminary details about poststratification	152

5.7.2	Poststratification of a SRSwoR sample	153
5.7.3	Poststratification of samples other than SRSwoR	155
5.8	Stratified sampling of a continuous population	156
5.8.1	Probability densities	158
5.8.2	Estimation	158
5.8.3	Sample allocation	161
5.9	Terms to Remember	162
5.10	Exercises	162
5.11	Appendix	162
5.11.1	Proof of design-unbiasedness of $\hat{\tau}_{y,pst,h}$.	162
5.11.2	Double sampling for stratification from Resource Inventory Notes (Frayer 1978)	163
6	Using Auxiliary Information to Improve Estimation	165
6.1	Generalized ratio estimator	166
6.2	Bias of the generalized ratio estimator	169
6.3	Variance of the generalized ratio estimator	171
6.4	Estimated variance of the generalized ratio estimator	176
6.5	Confidence interval estimation	179
6.6	Ratio estimation with systematic sampling design	179
6.7	Generalized ratio estimation with stratified sampling	180
6.7.1	Ratio estimation of population and strata totals	180
6.7.2	Estimating the variances of $\hat{\tau}_{y\pi,st, strat}$ and $\hat{\tau}_{y\pi,st, crat}$	183
6.7.3	Ratio estimation of population and strata means	184
6.8	Generalized regression estimator	185
6.8.1	Regression estimation following SRSwoR	186
6.8.2	Regression estimation following any sampling design	189
6.9	Double sampling with ratio and regression estimation	192
6.10	Terms to Remember	198
6.11	Exercises	198
6.12	Appendix	201
6.12.1	Linear correlation coefficient	201
6.12.2	The jackknife estimator of $R_{y x}$.	203
6.12.3	Equivalent expression of B_{π}	204
6.12.4	Joint inclusion probability under unequal probability systematic sampling	204
7	Sampling with Fixed Area Plots	207
7.1	Introduction	207
7.2	Notation	208
7.2.1	Selection and installation	209
7.3	Sampling protocol	209
7.4	Estimation	210
7.4.1	Inclusion zone of an element	210
7.4.2	Estimation following plot sampling	215

7.4.3	Prorating estimates to unit area values	219
7.4.4	Estimating the mean attribute per element	222
7.4.5	Stratification	222
7.4.6	Sampling intensity	223
7.4.7	Elements near the plot border	223
7.5	Edge effect	223
7.5.1	External peripheral zone	224
7.5.2	Pullback method	225
7.5.3	Grosenbaugh's method	226
7.5.4	Mirage method	227
7.5.5	Walkthrough method	229
7.5.6	Edge corrections for plot clusters	231
	Vectorwalk method	232
	Reflection method	233
7.6	Plot size and shape	234
7.7	Estimating change	237
7.8	Terms to remember	242
7.9	Exercises	242
7.10	Appendix	246
	7.10.1 Joint inclusion zones	246
8	Bitterlich Sampling	247
8.1	Introduction	247
8.2	Fundamental concepts	247
8.2.1	Limiting distance for inclusion	248
8.2.2	Basal area factor	249
8.2.3	Plot radius factor	249
8.2.4	Sampling protocol	250
8.2.5	Bitterlich sampling with English units	253
8.2.6	Noncircular tree boles	253
8.2.7	Expected number of trees selected at a point	254
8.3	Estimation following Bitterlich sampling	254
8.3.1	Estimating τ_y	254
8.3.2	Estimating basal area	255
8.3.3	Estimating population size	257
8.3.4	Estimating μ_y	257
8.3.5	Variances and variance estimators	257
8.3.6	Product estimator	258
8.4	Edge effect	262
8.5	Double sampling	262
8.5.1	Double sampling with second-phase subset of sample points	263
8.5.2	Optimal allocation	268
8.5.3	Double sampling with a second-phase subset of trees	268
	Big BAF sampling	268

	Subsampling trees with probability proportional to height	269
8.6	Sampling to estimate change in stock	270
	8.6.1 Conventional components of change	271
	8.6.2 Eriksson's components of change	273
8.7	Terms to remember	275
8.8	Exercises	276
9	Line Intersect Sampling	279
9.1	Introduction	279
9.2	LIS with straight-line transects	280
	9.2.1 Sampling protocol	280
	9.2.2 Estimation conditioned on transect orientation	283
	9.2.3 Unconditional estimation	285
	9.2.4 Estimator variance and variance estimators	288
	9.2.5 Interval estimation	289
	9.2.6 Conditional vs. unconditional estimation	289
9.3	Unit area estimators	290
9.4	Estimation with an auxiliary variate	293
9.5	Estimating the mean attribute	298
9.6	Nesting transects of different lengths	299
9.7	Dealing with edge effect in LIS	299
	9.7.1 Walkback method	299
	9.7.2 Reflection method	300
9.8	Transects with multiple segments	301
	9.8.1 Radial and polygonal transects	302
	9.8.2 Sampling protocol for transects having multiple segments	303
	9.8.3 Estimation from transects having multiple segments	304
	9.8.4 Averaging estimators for straight transects	307
9.9	Parallel transects of uneven length	307
9.10	Terms to remember	309
9.11	Exercises	309
9.12	Appendix	310
	9.12.1 Joint inclusion zones	310
	9.12.2 Derivation of $E\{t_{ks}q_k(\theta_s) \theta_s\}$	310
	9.12.3 Variance of conditional estimator of τ_y with auxiliary variate using transects with a single segment	311
	9.12.4 Variance of unconditional estimator of τ_y with auxiliary variate using transects with a single segment	311
	9.12.5 Conditional expected value and variance of $t_{ks\bullet}$ with transects having multiple segments	311
	9.12.6 Unconditional expected value and variance of $t_{ks\bullet}$ with transects having multiple segments	314
	9.12.7 Variance of \check{t}_{yms}^c	315
	9.12.8 Variance of \check{t}_{yms}^u	317
	9.12.9 Variance of \check{t}_{yqs}^c	318

10 A Monte Carlo Integration Approach to Areal Sampling	327
10.1 Areal sampling	328
10.1.1 Selection	329
10.1.2 Estimation	330
10.2 Plot sampling	331
10.2.1 Cluster plots	332
10.3 Bitterlich sampling	333
10.4 Point relascope sampling	334
10.5 Line intersect sampling	336
10.5.1 Unconditional estimation	338
10.6 Perpendicular distance sampling	339
10.6.1 Omnibus PDS	341
Coverage area	342
10.7 Edge correction	343
10.7.1 Buffer method	343
10.7.2 Direct measurement of inclusion area	344
10.7.3 Mirage method	344
10.7.4 Reflection method for transects	346
10.7.5 Walkthrough method	348
10.8 Redux: Continuous versus discrete	349
10.9 Terms to Remember	351
10.10 Exercises	351
10.11 Appendix	351
10.11.1 Variance of $\hat{\tau}_{\rho,rep}$	351
10.11.2 Variance with edge correction by walkthrough or mirage	353
10.11.3 LIS with segmented transects and correction of edge effect	354
11 Miscellaneous Methods	357
11.1 Introduction	357
11.2 Point relascope sampling	357
11.3 Horizontal line sampling	359
11.3.1 Sausage sampling	361
11.3.2 Modified sausage sampling	362
11.3.3 Redux: LIS	363
11.4 Transect relascope sampling	364
11.4.1 Ståhl's method	366
11.5 Ranked set sampling	367
11.5.1 Overview	368
11.5.2 Traditional applications	369
Estimation	370
11.5.3 Discrete applications	371
Efficiency	372

11.6	Adaptive cluster sampling	373
11.6.1	Sampling tessellated landscapes	373
11.6.2	Selection and Estimation	376
	Probability proportional to area	376
	Probability proportional to number of plots	377
	Selection of plots without replacement	377
11.6.3	Sampling landscapes without tessellation	378
	Estimation with network inclusion probabilities	379
	Estimation with element inclusion probabilities	380
	Edge effect	381
11.7	3P sampling	382
11.7.1	Selection	382
11.7.2	Variance estimation following 3P sampling	385
11.7.3	Regression adjustment of inclusion probabilities	387
11.7.4	Composite and calibration estimation	388
11.8	Terms to Remember	388
11.9	Exercises	389
11.10	Appendix	390
11.10.1	Details on 3P sampling	390
12	Two-stage Sampling	393
12.1	Customary two-stage sampling	394
12.1.1	Estimating cluster totals	394
12.1.2	Estimating the population total from cluster totals	395
12.1.3	SRSwoR special case	397
12.1.4	Self-weighting SRSwoR designs	398
12.1.5	Ppswr special case	398
12.1.6	Examples	398
12.1.7	Alternative estimators	399
12.2	General two-stage sampling	400
12.3	Three-stage sampling	402
12.4	Terms to Remember	404
13	Randomized Branch Sampling	405
13.1	Terminology	405
13.2	Path selection	406
13.3	Estimation	409
13.3.1	Average stem length	411
13.4	Selection probabilities	412
13.5	Tools and tricks of the trade	414
13.6	Subsampling a path	415
13.6.1	Subsampling strategy	415
13.6.2	Crude Monte Carlo	416
13.6.3	<i>Importance sampling</i>	417
13.6.4	Constructing a proxy function	417

13.6.5	Sampling with a control variate	419
13.6.6	Choice of method	420
13.7	Terms to remember	420
13.8	Exercises and projects	421
13.9	Appendix 13	421
13.9.1	Proof of the unbiasedness of $\hat{\tau}_{yQ_i}$	421
14	Sampling with Partial Replacement	423
14.1	Introduction	423
14.2	Estimation with partially replaced sampling units	424
14.2.1	Optimal number of remeasured units	429
14.2.2	Estimation of $\tau_y(t_2)$ with estimated parameter values	429
14.2.3	Contribution of remeasured sample units	431
14.3	Estimation of change	431
14.3.1	Optimal estimator	432
14.3.2	Additive estimator	432
14.3.3	Change estimator based on remeasured sampling locations	432
14.4	SPR with stratification	433
14.5	SPR for three occasions	433
14.6	Concluding remarks	434
14.7	Terms to remember	435
14.8	Appendix	435
14.8.1	Derivation of the composite estimator	435
14.8.2	Variance of the composite estimator	436
14.8.3	Components of the composite estimator $\hat{\tau}_{y,SPR}(t_3:t_1)$	437
	Bibliography	439
	Author Index	459
	Subject Index	467