

# Contents

<i>Figures</i>	<i>xv</i>
<i>Tables</i>	<i>xvi</i>
<i>Foreword</i>	<i>xvii</i>
<i>Preface</i>	<i>xix</i>
<i>Acknowledgments</i>	<i>xxiii</i>
<b>1. Introduction to Sample Survey Designs</b>	
1.1 Introduction	1
1.2 Population, Units and Sampling Units	1
1.3 Sampling Design	3
1.4 Probability and Purposive Sampling	3
1.4.1 Probability sampling	3
1.4.2 Purposive sampling	4
1.5 Frame	5
1.6 Bias and Error	6
1.7 Few Guidelines for a Desirable Sampling Design	11
<b>2. Basic Sampling Designs</b>	
2.1 Introduction	13
2.2 Simple Random Sampling	14
2.2.1 Description	14
2.2.2 Methods of selection	15
2.2.3 Estimation of mean, total and need for weights	16
2.2.3.1 Normalization of weights	19
2.2.3.2 Role of weights	21
2.2.4 Estimation of proportion	22
2.2.5 Subclass estimates	22
2.2.6 Sampling variance of estimates	23
2.2.6.1 Sampling variance of a sample mean	23
2.2.6.2 Sampling variance of estimated population total	25
2.2.6.3 Sampling variance of proportion	25

2.2.6.4	Sampling variance of subclass estimates	25
2.2.7	Determination of sample size	28
2.3	Stratified Sampling	31
2.3.1	Description	31
2.3.2	Estimation of parameters	33
2.3.2.1	Estimation of mean	34
2.3.2.2	Estimation of total	35
2.3.2.3	Estimation of proportion	35
2.3.3	Weighting and its similarity with standardization	35
2.3.4	Sampling variance of estimates	37
2.3.4.1	Sampling variance of mean	38
2.3.4.2	Sampling variance of total	38
2.3.4.3	Sampling variance of proportion	38
2.3.5	Allocation and selection of units	38
2.3.5.1	Proportional allocation	39
2.3.5.2	Optimum allocation	40
2.3.5.3	Practical guidelines for allocation	41
2.3.6	Some advantages of stratification	42
2.3.7	Post-stratification	43
2.4	Systematic Sampling	45
2.4.1	Description	45
2.4.2	Method of selection	45
2.4.2.1	Decimal interval method	46
2.4.3	Advantages of systematic sampling	46
2.4.4	Disadvantages of systematic sampling	47
2.4.4.1	Monotonic trend	47
2.4.4.2	Periodicity	47
2.4.5	Estimation of parameters and their sampling variances	48
2.4.5.1	Two consecutive units per stratum	49
2.5	Probability Proportional to Size Sampling	50
2.5.1	Description	50
2.6	Cluster Sampling	54
2.6.1	Description	54
2.6.1.1	Preparation of artificial clusters	55
2.6.2	Method of selection	56
2.6.3	Estimation of parameters and sampling variances	57
2.6.3.1	Clusters of equal size	57
2.6.3.2	Clusters of unequal size	58
2.7	Key Points	60
<b>3.</b>	<b>Multi-stage Designs</b>	
3.1	Introduction	62
3.2	Two-stage Design with Equal Size Clusters	63
3.2.1	Components of overall variation	64
3.2.2	Two-stage design for selection of units with equal probability	69
3.2.2.1	Procedure of selection in a two-stage EPSEM design	69

3.2.2.2	Estimation of mean and variance	70
3.2.2.3	Clustering, design effect and choice of number of PSUs/cluster	73
3.3	Two-Stage Design with Unequal Cluster Size	79
3.3.1	Estimation of mean and sampling variance	80
3.3.2	Two desirable properties of the design	84
3.3.3	Guidelines for attaining desired property	85
3.3.4	Ways to control variations in cluster size	87
3.3.4.1	Controlling size of clusters	87
3.3.4.2	Alternative selection procedure	87
3.3.4.2.1	Alternatives when information on size of PSUs refers to a past period	89
3.3.4.3	Stratification of PSUs to reduce variations in cluster size	96
3.4	Stratification in Multistage Design	96
3.4.1	Estimation of parameters in unequal cluster size	97
3.4.2	Estimation of parameters in equal cluster size	100
3.5	Selection of Sampling Units at Different Stages	103
3.5.1	Selection of PSUs	103
3.5.2	Selection of second-stage units	104
3.5.3	Selection of individuals within a household	105
3.6	Key Points	107
<b>4.</b>	<b>Probability Sampling under Imperfect Frame</b>	
4.1	Introduction	109
4.2	Sampling Populations Having Specific Attributes	110
4.2.1	Sampling when target population is not rare	111
4.2.1.1	Sampling without screening	111
4.2.1.2	Sampling with screening	112
4.2.1.3	Relative advantages with screening and without screening	112
4.2.1.4	Facilitating screening	113
4.2.2	Sampling for rare attributes	114
4.2.2.1	Household-based sampling of rare population	114
4.3	Defective Frame	117
4.3.1	Duplications	118
4.3.1.1	Estimation in presence of duplications	118
4.3.1.2	Procedure to deal with duplicate listing	119
4.3.1.3	Incompleteness or omissions in a frame	121
4.4	Sampling in Absence of a Frame	122
4.4.1	Facilitating a cluster design	123
4.4.2	Selection, data collection and estimation	125
4.5	Household Listing	128
4.5.1	Two alternatives if listing is to be avoided	129
4.6	Key Points	131



7.4.1.1	Description and estimation of effect	168
7.4.1.2	Biasing effects and remedies	169
7.4.1.3	Estimation of standard error of estimated impact	173
7.4.2	One sample each from $T$ and $C$ observed at one point of time (treatment-control design)	174
7.4.2.1	Description and estimation of effect	174
7.4.2.2	Biasing effects and remedies	175
7.4.2.3	Estimation of standard error of estimated impact	175
7.4.3	Two samples each from $T$ and $C$ observed at two points in time (before-after and treatment-control design)	175
7.4.3.1	Description and estimation of effect	175
7.4.3.2	Biasing effects and remedies	177
7.4.3.3	Estimation of standard error of estimated impact	178
7.5	Output and Its Timing	179
7.6	Key Points	179
<b>8.</b>	<b>Designs for Causal Effects: Allocation of Study Units</b>	
8.1	Introduction	181
8.2	Alternative Tools to Attain Balance	181
8.2.1	Randomization	182
8.2.2	Stratification	183
8.2.3	Pair matching	184
8.3	Advantages and Disadvantages of Three Tools	184
8.3.1	Randomization	185
8.3.2	Matching	185
8.3.2.1	Stratification	185
8.3.2.2	Pair matching	186
8.4	Choice of Study Units	187
8.4.1	Procedure of allocation of units/clusters	188
8.4.1.1	Randomization	188
8.4.1.1.1	Restricted randomization	189
8.4.1.2	Stratification	191
8.4.1.3	Pair matching	191
8.5	Potential Outcome Framework	191
8.5.1	Propensity score matching	192
8.6	Choice of a Design	196
8.7	Key Points	198
<b>9.</b>	<b>Statistical Tests for Measuring Impact</b>	
9.1	Introduction	199
9.1.1	Two different ways to estimate impact	200
9.2	Impact when Units are Allocated Randomly	201
9.2.1	Testing difference between two means	202
9.2.1.1	Testing means from two different populations	203
9.2.1.2	Large-sample $z$ -test	204
9.2.1.3	Testing several means: Application of ANOVA	205

9.2.1.4	Non-parametric tests	207
9.2.2	Testing difference between two proportions	213
9.2.2.1	Chi-square test of independence	213
9.2.2.2	Testing odds ratio	215
9.3	Impact when Clusters are Allocated Randomly	216
9.3.1	Analysis at cluster level	216
9.3.2	Analysis at individual level	217
9.4	Impact when Stratification is used before allocation	218
9.4.1	When units are allocated	218
9.4.1.1	Two-way ANOVA test	219
9.4.2	When clusters are allocated	222
9.5	Impact in Pair Matching	222
9.5.1	Variables measured in interval scale	223
9.5.2	Dichotomous variable	223
9.5.2.1	Exact binomial test	224
9.5.3	Non-parametric test	226
9.6	Model-Based Analysis	227
9.6.1	Multiple regression analysis	228
9.6.1.1	Modifications in the case of cluster sampling	229
9.6.2	Logistic regression	230
9.6.3	Assumptions in regressions	232
9.7	Key Points	233
<b>10.</b>	<b>Case Studies</b>	
10.1	Introduction	234
	<i>Part I: Sample Survey Designs</i>	234
10.2	National Family Health Surveys, India (NFHS, India)	234
10.3	Sampling Design of NFHS	234
10.3.1	Sample size	234
10.3.2	Choice of PSU	236
10.3.3	Design for rural area	236
10.3.3.1	Merging of small villages	236
10.3.3.2	Stratification	237
10.3.3.3	Selection of sampling units	239
10.3.3.4	An alternative two-stage selection	243
10.3.4	Design for urban area	243
10.3.5	Estimation	244
10.3.5.1	Computation of weights	244
10.3.5.2	Estimation of parameters	245
10.4	Other Global Large-Scale Surveys	247
10.4.1	Use of master sample in survey designs	247
10.4.2	Example of GATS sample design in Nigeria	249
10.5	National Sample Surveys (NSS) in India	249
	<i>Part II: Evaluation Design</i>	250
10.6	Illustration of Evaluation Designs	250

---

10.6.1 Impact evaluation of a life skills education intervention on adolescent girls' empowerment	250
10.6.2 Fisher's tea test	252
10.6.3 Impact of an intervention to promote use of Intra Uterine Device in a population	253
10.6.4 An experiment to test effectiveness of a medicine	253
10.6.5 An intervention to reduce post-partum haemorrhage	254
<i>References</i>	255
<i>Index</i>	260