SURVEY SAMPLING
THEORY AND METHODS

DANKIT K. NASSIUMA
Table of Contents

PREFACE ix

1 INTRODUCTION 1
1.1 Overview of research methods 1
1.2 Surveys and sampling 7
1.3 Definitions 10
1.4 Properties of estimators 12
1.5 Sampling methods 14
1.6 Survey design and planning 15
1.7 Data collection methods 15
1.8 Sources of error in sampling 22
1.9 Pilot surveys (pre-surveys) 23
Practice problems 24

2 SIMPLE RANDOM SAMPLING 27
2.1 Introduction 27
2.2 Simple random sampling without replacement (SRSWOR) 27
2.3 Estimating the finite population variance $S^2$ 35
2.4 Estimation of the population total 39
2.5 Estimation of the population proportion 40
2.6 Simple random sampling with replacement (SRSWR) 43
2.7 Subpopulations (domains) 50
2.8 Pooling of independent estimates in SRS 52
Practice problems 54

3 DETERMINATION OF SAMPLE SIZE 59
3.1 Introduction 59
3.2 Use of coefficient of variation 59
3.3 Use of probability statements 61
3.4 Determination of $n$ for unknown $\sigma$ and fixed c.i. length 63
3.5 Sample size based on sampling cost 64
Practice problems 67
4  SAMPLING WITH UNEQUAL PROBABILITY  71
4.1  Introduction  71
4.2  Sample selection by PSS with replacement  74
4.3  Estimation of the population total in PPS WR  76
4.4  Estimation of the \( \text{var}(\hat{Y}) \)  79
4.5  Comparison of PPS WR with SRS WR  82
4.6  Selection of samples by PPS WOR  84
4.7  Comparison of PPS WOR and PPS WR  96
4.8  Combination of PPS and SRS schemes  98
Practice problems  99

5  SYSTEMATIC SAMPLING  103
5.1  Introduction  103
5.2  Linear systematic sampling (LSS)  103
5.3  Circular systematic sampling (CSS)  106
5.4  Variance of \( \bar{y}_{sys} \)  107
5.5  Comparison of systematic sampling and simple random sampling SRS  109
5.6  Estimation of the variance of a systematic sample mean  112
5.7  Super populations  112
5.8  Systematic sampling using unequal probabilities  117
5.9  Repeated sampling in SYS  122
Practice problems  123

6  STRATIFIED RANDOM SAMPLING  125
6.1  Introduction  125
6.2  Estimation of the population mean  126
6.3  Estimation of the variance of \( \text{var}(\bar{y}_{st}) \)  129
6.4  Allocation of sample size  129
6.5  Comparison of stratified sampling with SRS  135
6.6  Allocations which need more than 100% sampling  137
6.7  Stratified sampling for proportions  138
6.8  Post-stratification  140
6.9  Stratified sampling with unequal probabilities  142
Practice problems  144
7  **RATIO AND REGRESSION ESTIMATORS**  147
   7.1  Introduction  147
   7.2  Ratio estimators  148
   7.3  Ratio estimators in stratified sampling  160
   7.4  Regression estimators  164
   7.5  Regression estimators for stratified sampling  168
   Practice problems  172

8  **CLUSTER SAMPLING**  177
   8.1  Introduction  177
   8.2  Single stage cluster sampling  181
   8.3  Multistage cluster sampling  196
   8.4  Two stage cluster sampling  197
   8.5  Stratification in cluster sampling  210
   Practice problems  213

9  **FURTHER TOPICS**  215
   9.1  Two phase sampling  216
   9.2  Successive sampling  220
   9.3  Estimation of population size  221
   Practice problems  227

REFERENCES  229

APPENDIX  231

INDEX  233
Preface

Surveys are inseparable from research and planning. This necessitates the teaching of sampling methods as well as their application not only at college and university level, but also to applied researchers.

It is the objective of this book to introduce the language, methods and application of sampling from a practical, mathematical perspective. It is expected that the reader will be enabled to plan and execute surveys and also be capable of evaluating estimates of various parameters especially the location and scale parameters as well as their standard errors.

The book is divided into nine chapters with the first chapter introducing the language of sampling and an overview of research projects, proposal writing and experimentation. The material in this chapter has a social science and educational research flair and is thus easy to apply in a wide range of situations including market research and opinion polls. The second chapter introduces the simple random sampling procedure which is the most basic sampling technique. This is followed by a study on methods of sample size determination in chapter three. This chapter is based mainly on the assumption that a simple random sampling procedure is used but it can easily be extended to other sampling procedures. Chapters four to eight focus on the unequal probability, systematic sampling, stratified sampling, ratio and regression estimation, and cluster sampling procedures. Elaborate proofs for various procedures are given in these chapters. In chapter nine, three topics in sampling are briefly discussed. These are aspects which are usually ignored in most books on sampling and include multiphase sampling, successive sampling and the estimation of population size.

In chapters two to nine, a cook-book kind of option is availed for those who may be interested in applications of the various methods.
Important functions for the estimates and their standard errors are highlighted in boxes so that the rest of the formulae can be skipped.

The material presented in this text when used for teaching at undergraduate level should be covered in two semesters. It is expected that the student has already taken some basic first and second year undergraduate mathematics courses. In the case of researchers interested in the applications aspect, a basic knowledge and appreciation of statistical inference is assumed.

It is hoped that this book will be friendly and invaluable to the researcher, lecturer and the student for the understanding and application of sampling procedures to a wide range of problems.