·0242·2
S948
E.3

8660823

Sampling Theory of Surveys with Applications

THIRD EDITION

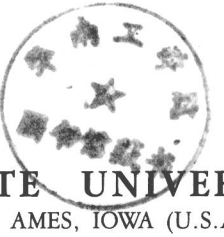
PANDURANG V. SUKHATME, *Ph.D., D.Sc.*
Maharashtra Association for the Cultivation of Science
Pune (India)

(Late) BALKRISHNA V. SUKHATME, *Ph.D.*
Formerly at Iowa State University, Ames (U.S.A.)

SHASHIKALA SUKHATME, *Ph.D.*
Iowa State University, Ames (U.S.A.)

AND

C. ASOK, *Ph.D.*
Maharashtra Association for the Cultivation of Science
Pune (India)



IOWA STATE UNIVERSITY PRESS
AMES, IOWA (U.S.A.)

AND

Indian Society of Agricultural Statistics
NEW DELHI-110 012 (India)

E8660823

© 1954 by the Indian Society of Agricultural Statistics, New Delhi India, and the Iowa State University Press, Ames, Iowa, USA.

© 1970 by Pandurang V. Sukhatme and Balkrishna V. Sukhatme. *All Rights Reserved. Printed in the U.S.A.*

© 1984 by Pandurang V. Sukhatme, Balkrishna V. Sukhatme, Shashikala Sukhatme and C. Asok. *All Rights Reserved.*

Published by the Indian Society of Agricultural Statistics, New Delhi, India, and the Iowa State University Press, Ames, Iowa, U.S.A.

First Edition, 1954
Second Printing, 1958
Third Printing, 1960
Second Revised Edition, 1970
Second Printing, 1976
Third Revised Edition, 1984

Library of Congress Catalog Card No. 84-082161
ISBN 0-8138-1617-7

SAMPLING THEORY OF SURVEYS WITH APPLICATIONS

THIRD EDITION

8660823

PREFACE TO THE THIRD EDITION

"I do not think that I would be able to write any more with my hand, but as this letter will show, I can type and I am now sufficiently recovered to feel confident that I will be able to complete the revision of our book in another few months."

This is how Bal (B.V. Sukhatme) wrote to me on April 23, 1979. Bal passed away within a week after writing this. Actually, I had rushed from India to see him at Mayo Clinic, Rochester. He recognised my voice and was so full of emotions that the doctor in charge of the Emergency Room was forced to warn him to restrain movement. It was only after I returned to India that I got his letter of April 23.

That Bal was working hard on the third Edition was known to his students and to his colleagues, particularly Dr. H.A. David, Director, Statistical Laboratory, Iowa State University, Ames, Iowa.

Bal had spent a month with Prof. V.P. Godambe, discussing with him the foundations of the Sampling Theory of Surveys and trying to think how to go about in reflecting them in the new edition and yet maintaining essentially the same character as in the old Editions. Bal even began teaching a modified course in sampling theory of Surveys and as he taught, he completed the manuscript of the first five Chapters. He was busy finalising the text of the 6th and 7th Chapters, when he was taken ill in December, 1978.

Normally, I should have had little difficulty in completing the revision of the book, but I have been out of touch with the subject too long. Bal's death also raised for Dr. H.A. David the question of making temporary arrangement to guide Ph. D. students and completing his course of lectures, pedding new appointment in his

此为试读,需要完整PDF请访问: www.ertongbook.com

place. It was agreed that I should depute Dr. C. Asok, Bal's student, to help in these tasks. However, the load of work proved too heavy for Dr. Asok to complete the revision of the book.

Meanwhile, Shashikala was feeling her way in her lonely life. As she began putting together Bal's notes, I found her gradually interested in the subject of sampling surveys. I suggested to her that she might try her hand to complete the job which Bal left unfinished.

And to my delight, the suggestion worked. Night and day she worked on the manuscript and completed the revision by May, 1981. Dr. Asok on his part went through the Chapters and satisfied himself that the revision was well done.

As this was going on, I was toying with the idea to refer the manuscript to some well-known authority on the subject. But I found it difficult to get anyone sufficiently senior to spare the time to look through the manuscript. I wanted the book to be published in the joint names of Bal, Shashi and Asok, but Dr David did not like me to drop out. Finally, it was decided that we all should be shown as co-authors. Asok's association as co-author has the additional advantage of holding assurance that he would be able to help in future revisions as and when necessary.

The one satisfaction I have is that this revision helped Shashi to rehabilitate herself sooner than I had thought. She is a lady of tremendous will power and once she takes a thing in hand, she will never leave it until she completes it.

To Dr H.A. David I am most grateful for all the assistance and encouragement he gave. I feel indebted to his secretarial staff for the enormous work they put in typing. I would also like to express my thanks to Mr. Merritt E. Bailey, Director of the Iowa State University Press for his continued help and encouragement in the publication of the book.

I am also indebted to Bal's students, particularly Shriram Biyani, Chand Midha, James Drew, Rick Auer and Eric Bondy, for reading parts of the manuscript.

I am also grateful to Dr. Prem Narain, Secretary and Drs. B.B.P.S. Goel and Shivtar Singh, Joint Secretaries of the Indian Society of Agricultural Statistics for their continued interest in the revision of the book and for their help in reading through proofs.

The edition has become much larger than I would have liked to see. To include all the principal developments in theory, illustrate them with examples and exercises and yet not allow the size to increase, proved too large a task for us. Only Bal amongst us could have attempted it. I feel guilty that I have not been able to play a more active part. I sincerely hope that the new edition will continue to meet the needs of students in India and abroad.

P.V. SUKHATME

CONTENTS

I. INTRODUCTION AND BASIC CONCEPTS 1

1.1 Sampling Concept 1
1.2 Why Sample ? 2
1.3 Some Uses of Sampling 3
1.4 Planning and Organization of a Survey 4
1.5 Sampling Design 6
1.6 Probability Sampling Procedure 9
1.7 Estimation 11
1.8 Properties of Estimators 13
1.9 Sampling Strategy 15
1.10 Cost Considerations in the Choice of a Sampling Strategy 16
Exercises 17
References 18
Appendix I Table of Random Numbers 19

II. SIMPLE RANDOM SAMPLING WITHOUT REPLACEMENT 21

2.1 Introduction 21
2.2 Some Properties of Simple Random Sampling Without Replacement 23
2.3 Notation 25
2.4 Unbiasedness of the Sample Mean $\bar{y}$ and the Sample Mean Square s^2 26
2.5 Variance of the Sample Mean $\bar{y}$ 29
2.6 Estimation of the Standard Error of $\bar{y}$ 30
2.7 Confidence Limits for $\bar{Y}$ 32
2.8 Some Optimal Properties of the Sample Mean 38
2.9 Estimation of the Sample Size 40
2.10 Estimation of Proportions 42

2.11 Confidence Limits and Sample Size for Estimating Proportions 45
2.12 Estimation of Proportions in Domains 48
2.13 Inverse Sampling 51
2.14 Quantitative and Qualitative Characters 53
Exercises 56
References 60

III. SAMPLING WITH VARYING PROBABILITIES 62

3.1 Introduction 62
3.2 Procedure of Selecting a Sample with Varying Probabilities 62
3.3 Estimation of the Population Mean $\bar{Y}$ 64
3.4 An Estimator of Population Mean Based on Distinct Units 69
3.5 Sampling with Varying Probabilities and Without Replacement 73
3.6 Ordered Estimates 74
3.7 Unordered Estimates 77
3.8 The Horvitz-Thompson Estimator 78
3.9 Optimal Properties of the Horvitz-Thompson Estimator 81
3.10 Estimation of Variance of $\hat{\bar{Y}}_{HT}$ 85
3.11 Some I.P.P.S. Sampling Without Replacement Procedures 87
3.12 Rao, Hartley and Cochran's Procedure 95
Exercises 99
References 101

IV. STRATIFIED SAMPLING 105

4.1 Introduction 105
4.2 Notation 106
4.3 Estimation of the Population Mean $\bar{Y}$ 106
4.4 Choice of Sample Sizes in Different Strata 107

4.5 Comparison of Stratified Sampling with Simple Random Sampling Without Replacement in Terms of Precision 113
4.6 Comparison of Stratified Sampling with Simple Random Sampling Without Replacement in Terms of Cost 121
4.7 Practical Difficulties in Adopting Optimum Allocation 123
4.8 Estimation of the Gain in Precision due to Stratification 128
4.9 Post Stratification 133
4.10 Effects of Inaccuracies in Strata Sizes 135
4.11 Deep Stratification 142
4.12 Construction of Strata 148
4.13 Determination of Number of Strata 154
4.14 Allocation to Strata with several Characteristics 158
4.15 Estimation of $V(\bar{y}_{st})$ with one unit per Stratum 169
4.16 Controlled Selection 170
Varying Probabilities of Selection 171
4.17 Estimation of the Population Mean 171
4.18 Choice of Sample Sizes in Different Strata 173
4.19 Efficiency of Stratified Sampling and Estimation of Gain in Precision due to Stratification 174
Exercises 178
References 180

V. RATIO TYPE METHODS OF ESTIMATION 184
5.1 Introduction 184
5.2 The Ratio Estimator 184
5.3 Bias and Mean Square Error of the Ratio Estimator 185
5.4 Approximations to the Bias and Mean Square Error of the Ratio Estimator 188
5.5 Accuracy of the Approximation to the Bias and Mean Square Error of $\bar{y}_r$ 190
5.6 Optimality of the Ratio Estimator under a Super-Population Probability Model 194

5.7 Estimation of Bias and Mean Square Error of the Ratio Estimator 198
5.8 Confidence Limits 202
5.9 Unbiased or Almost Unbiased Ratio-Type Estimators 203
5.10 Double Sampling : $\bar{X}$ not known 211
5.11 Product Estimators 216
5.12 Multivariate Ratio-Type Estimators 217
5.13 Estimation in Stratified Sampling 223
Exercises 228
References 229
Appendix II Expected Values of Certain Higher Order Product Moments 232

VI. REGRESSION METHODS OF ESTIMATION 236
6.1 Introduction 236
6.2 The Regression Estimator 236
6.3 Bias and Mean Square Error of the Regression Estimator 238
6.4 Efficiency of the Regression Estimator 240
6.5 Optimality of the Regression Estimator under a Super-Population Probability Model 241
6.6 Double Sampling : $\bar{X}$ not known 244
6.7 Sampling on Two or More Occasions 248
6.8 Estimation in Stratified Sampling 259
6.9 Multivariate Regression-Type Estimators 262
Exercises 267
References 268

VII. CHOICE OF SAMPLING UNIT 270
7.1 Cluster Sampling 270
7.2 Estimation of the Population Mean 271
7.3 Efficiency of Cluster Sampling 272
7.4 Estimation from the Sample of the Efficiency of Cluster Sampling 278

此为试读,需要完整PDF请访问: www.ertongbook.com

7.5 Relationship between the Variance of the mean of a Single Cluster and its Size 280
7.6 Optimal Unit of Sampling and Multipurpose Surveys 284
7.7 Use of Supplementary Information in improving the Efficiency of Cluster Sampling 290
7.8 Unequal Clusters : Estimation of the Population Mean 290
7.9 Unequal Clusters : Comparison of Sampling Strategies Under a Super-Population Model 294
Exercises 300
References 302

VIII. SUB-SAMPLING 304

8.1 Introduction 304
8.2 Two-Stage Sampling, Equal First-Stage Units : Estimation of the Population Mean 304
8.3 Allocation of Sample to the Two Stages : Equal First-Stage Units 308
8.4 Comparison of Two-Stage with One-Stage Sampling 314
8.5 Effect of change in Size of First-Stage units on the Variance 316
8.6 Three-Stage Sampling, Equal First-Stage and Second-Stage Units 317
8.7 Allocation of the Sample to the Three Stages 322
8.8 Two-Stage Sampling, Unequal First-Stage Units 323
8.9 Two-Stage Sampling, Unequal First-Stage Units: Allocation of Sample 337
8.10 Three-Stage Sampling, Unequal First and Second-Stage Units 346
8.11 Stratified Sub-Sampling 352
8.12 Optimum Allocation in Stratified Sub-Sampling 354
8.13 Efficiency of Stratification in Sub-Sampling 356
Exercises 360
References 363

IX. SUB-SAMPLING (*Continued*) 364

9.1 Introduction 364
9.2 Estimation of the Population Mean 364
9.3 Allocation of Sample 370
9.4 Determination of Optimum Probabilities 374
9.5 Relative Efficiency of the Two Sub-Sampling Designs 378
9.6 Sub-Sampling without Replacement 379
9.7 Stratification and Gain due to it 385
9.8 Self-Weighting Designs 392
9.9 Sub-Sampling with Varying Probabilities of Selection at Each Stage : Sampling With Replacement 396
9.10 Sampling Without Replacement at Each tage 399
Exercises 401
References 402

X. SYSTEMATIC SAMPLING 404

10.1 Introduction 404
10.2 Systematic Sampling in Two Dimensions 406
10.3 Estimation of the Population Mean 407
10.4 Comparison of Systematic Sampling with Simple Random Sampling Without Replacement 410
10.5 Comparison of Systematic Sampling with Stratified Sampling 412
10.6 Comparison of Systematic Sampling with Simple Random Sampling and Stratified Sampling for Some Specified Populations 414
10.7 Estimation of the Variance 426
10.8 Two-Stage Sample : Equal First-Stage Units : Systematic Sampling of Second-Stage Units 429
10.9 Two-Stage Sample : Unequal First-Stage Units : Systematic Sampling of Second-Stage Units 431
Exercises 436
References 439

XI. NON-SAMPLING ERRORS 441

Part A—Errors in Surveys 441
11.1 Introduction 441
11.2 Types of Errors 441

Part B—Observational Errors 450

11.3 Mathematical Mcdel for the Measurement of Observational Errors 450

11.4 The Sample Mean and its Variance 451

11.5 Estimation of the Different Components of Variance 456

11.6 Stratified Sampling in which Enumerators are Assigned the Units in Their Respective Strata 470

11.7 Unstratified Sampling in which Enumerators are Assigned Neighbouring Units 473

11.8 Determination of the Optimum Number of Enumerators 474

11.9 Some Studies on Enumerator Variability and Respondent Bias 476

11.10 Limitations of the Methcd of Replicated Samples in Surveys 479

Part C—Incomplete Samples 484

11.11 The Problem 484

11.12 Effects of Non-Response 485

11.13 Hansen and Hurwitz Technique 487

11.14 Deming's Model 491

11.15 Politz and Simmons Technique 494

11.16 Randomized Response Techniques 496

Part D—Some Experiences in Sample Censuses of Agriculture 502

11.17 Introduction 502

11.18 The Frame 502

11.19 Practical Difficulties of the "Open Segment" Concept 507

11.20 Response Errors 508

11.21 Relative Magnitude of Sampling and Non-Sampling Errors 510

Exercises 510

References 516

Index 521

CHAPTER I

INTRODUCTION AND BASIC CONCEPTS

1.1 Sampling Concept

Sampling, which is the selection of part of an aggregate to represent the whole, is frequently used in everyday life in all kinds of investigations. Almost instinctively, before deciding to buy a lot, we examine a few articles preferably from different parts of the lot. Another example is provided by a handful of grain taken from a sack to determine the quality of the grain. These are examples where inference is drawn on the basis of the results obtained from a sample.

Sampling is used most frequently in surveys. The purpose of a sample survey is to obtain information about population. By 'population' we understand a group of units defined according to the aims and objects of the survey. Thus the population may consist of all the fields under a specified crop as in area and yield surveys, all the agricultural holdings larger than a specified size as in agricultural survey, or all the households having four or more children as in socio-economic surveys. Of course, the population may also refer to human beings of the whole population of a country or a particular sector of the country. The information that we seek about the population is usually the total number of units such as the number of farms in a State growing corn, aggregate values of the various characteristics such as the total acreage under corn, averages of the various characteristics per unit such as the average size of a household and proportions of units possessing specified attributes such as the proportion of households having income above a certain level.

It is clear that it hardly matters how we draw a sample, provided the population is homogeneous with respect to the characteristic under study. If, however, the units constituting the population vary considerably in respect of the characteristic under study, the method by which we draw a sample plays a critical role and it is

necessary to pay due attention to the selection of the sample and drawing inference therefrom.

1.2 Why Sample ?

A sampling method is a scientific and objective procedure of selecting units from a population and provides a sample that is expected to be representative of the population as a whole. It also provides procedures for the estimation of results that would be obtained if a comparable survey was taken on all the units in the population. No matter how good a sampling method is used to draw a sample, it is clear that a sample can never reproduce exactly the various characteristics of the population unless the population itself is taken as the sample and a *census* (a complete enumeration of the population) is carried out. The resulting discrepancies between the sample estimates, and the population values that would be obtained by enumerating all the units in the population in the same manner in which the sample is enumerated, are termed *sampling errors*. Their average magnitude will naturally depend on the population under study, the size of the sample, the manner in which the sample is drawn and the method of estimation. However, the fact that sampling errors cannot be avoided should not be disturbing, as long as these errors are sufficiently small so that the validity of the results obtained and the conclusions drawn are not affected.

Because of limited resources by way of time, trained personnel and money, it is generally not feasible to collect data from more than a fraction of a population. In a sample survey, the data can be collected and analyzed with much greater speed and the results made available at relatively short notice which would not be possible in the case of a census. In certain investigations, it may be essential to use specialized equipment or highly trained field staff for data collection, making it impossible to carry out such investigations except on a sampling basis. For the same reason, it is possible to employ efficient supervisory staff and exercise greater supervision over the field work to ensure completeness of returns and reliability of the data collected. Errors other than sampling errors such as those arising through non-response, incompleteness and inaccuracy of returns are termed *non-sampling errors* and are likely to be more wide-spread and important in a census than in a sample survey. Moreover, there is no way to assess these errors unlike in sample surveys where it is possible to assess the average magnitude of the

sampling errors. Sampling methods also provide the means of fixing in advance the details of the survey design, such as the procedure for selecting the sample and for choosing the sample size, in such a manner that with a preassigned probability the average magnitude of the sampling errors does not exceed the specified limit. In other words, sampling methods enable us to control the precision of sample estimates within limits fixed in advance.

1.3 Some Uses of Sampling

Sampling can be used in a variety of ways. However, it is mostly used in all kinds of surveys all over the world. Depending upon the objectives of the survey and the purposes for which the data may be used, sample surveys can be broadly classified into three categories: descriptive, analytical or both descriptive and analytical.

In descriptive surveys, the object is usually to obtain some descriptive measures with respect to the characteristics of the entire population under study. Such surveys are very common and are required for national planning and socio-economic development, to collect data on agricultural production and utilization of land and water resources, industrial production, unemployment and size of labor force, wholesale and retail prices, income and expenditure per household, numbers of literate persons and school-going children and so on. On the other hand, the object in analytical surveys is to obtain descriptive information for different sub-groups of the population in order to test hypotheses concerning possible relationships between the sub-groups. For example, in labor force surveys one would be interested not only in knowing the average number of hours worked per day and the wages paid but also whether men work longer hours than women and whether they receive higher wages than women for the same type of work.

Sampling methods are also used in population census. In fact, except for certain basic information required in respect of every individual, data on various items such as occupation, parentage, marriage fertility, income, migration, housing, is collected on a sampling basis. Sampling methods are used to provide counter checks and speed up tabulation and publication of results.

Sampling methods are used extensively in business and industry to increase operational efficiency. They play an important role in problems encountered in market research such as estimating the size of readership of news-magazines and newspapers or finding the reactions of

此为试读,需要完整PDF请访问: www.ertongbook.com