

A Manual on Sampling Techniques for Surveys

Dr. Sheikh Sirajul Islam



Center for Research and Publication
International Islamic University Chittagong

A MANUAL ON SAMPLING TECHNIQUES FOR SURVEYS

(APPLICABLE IN VILLAGE FOREST AND OTHER SURVEYS)

DR. SHEIKH SIRAJUL ISLAM

Professor of Statistics
Department of Business Administration
International Islamic University Chittagong
Chittagong, Bangladesh.

June, 2015

A Manual on Sampling Techniques for Surveys

(applicable in village forest and other surveys)

Dr. Sheikh Sirajul Islam

First Edition

June 2015

Cover Design

Dipak Dutta

Printed by

The Ad Communication

65, Jamal Khan Road, Chittagong

Phone: 031-611711

Price: Taka 400.00

Published by

Center for Research and Publication (CRP)

International Islamic University Chittagong

Kumira, Sitakunda, Chittagong

www.iiuc.ac.bd

City Liaison Office:

154/A, College Road, Chittagong-4203

Email: crp@iiuc.ac.bd

Tel: 88-031-638656-7, 610085, 639981 Ext. 143

© Center for Research and Publication, IIUC

All rights reserved. No part of this book may be reproduced or transmitted in any forms by any means, electronic, mechanical, or otherwise, whether now or hereafter devised, including photocopying, recording, or by any information storage and retrieval system without prior written permission from the publisher.

DEDICATION

This book is dedicated to my dear parents;
Late Sheikh Estab Ali
and
Late Maryam Khatun.

FOREWORD

The Center for Research and Publication (CRP) of International Islamic University Chittagong (IIUC) has been functioning for research, publication and Islamization of knowledge for a couple of years. CRP aims at creating new knowledge in different branches of disciplines, Islamization of existing disciplines and dissemination of this knowledge to the learners, academicians and practical users. Proper initiatives are being taken by the Governing Body of CRP to create an environment conducive for conducting research works for all sorts of researchers in general and for IIUC academics in particular.

In order to facilitate the teaching and research activities and motivate the authors, CRP has formulated policies to publish text and research oriented books at tertiary level. The main features of the policies are:

- a. A manuscript will be considered for publication by CRP subject to positive recommendation by two-expert reviewers in the relevant field.
- b. The author will have to resubmit the soft copy to the office of CRP incorporating the suggestions and remarks of the reviewers (if any).
- c. Initially a total of 500 copies will be printed by CRP.
- d. On behalf of IIUC, CRP will be the proprietor of the printed copies.
- e. Only 10 copies of the book will be given to the author free of cost.
- f. In view of the cost involved in printing process, the selling price of the book will be determined.
- g. The author will get 50% of the sale proceeds of the printed copies

'A Manual on Sampling Techniques for Surveys' is the first book under the above policies of CRP. The book is authored by Dr. Sheikh Sirajul Islam, Professor of Statistics in the Department of Business Administration at IIUC. Writing a book is a thoughtful and tedious work. It takes huge amount of time and in many cases couple of years to publish a good book. As informed, it was a dream of Dr. Islam to publish a book on applied sampling technique while he was pursuing his PhD study in 1992. It is our pleasure to see that his dream has become the reality for the encouragement of his teachers, friends, colleagues and students and finally by the patronization of IIUC.

The book contains twelve chapters on sampling designs. Each chapter includes definition, discussion, parameters estimating formulas, criteria of reliabilities, advantages, disadvantages and examples of all typical sampling designs. As informed, the variables and parameters are defined in terms of village forest surveys.

It is expected that this book will be highly beneficial for the researchers, academicians and students of IIUC and other universities in general. May Allah accept our endeavors.

Prof. Dr. Farid A. Sobhani

Director

Center for Research and Publication
International Islamic University Chittagong
Email: sobhani_farid@iiuc.ac.bd

Chittagong,
Bangladesh,
June, 2015

PREFACE

My first dream was to prepare a book of applied sampling technique when I was working on my Ph.D thesis in 1992. The title of the thesis is “Sampling design for assessing village wood and bamboo resources in Bangladesh – An exploratory study”. At that time I was Divisional Officer (Principal Scientific Officer) of Forest Economics Division under Bangladesh Forest Research Institute (BFRI), Chittagong. Later on I was promoted to Chief Research Officer (CRO) and to Director of BFRI. Because of research and administrative activities, I could not produce my dream in practice. After retirement now as a Professor of Statistics I have been teaching undergraduate and graduate students of Business Administration Department in International Islamic University Chittagong, Bangladesh. A large number of colleagues and students requested me to present academic contribution in the form of a book for their benefit and for the society as well. I was encouraged and inspired because of this request. I started to work in September, 2014 and completed all activities in June, 2015. Finally, “A Manual on Sampling Techniques for Surveys” is published by the grace of Almighty Allah. I have entertained almost all typical sampling techniques in this manual applicable in different surveys including village forest surveys. This is a manual for both undergraduate and graduate students of various disciplines including Statistics. This manual is also helpful to research when assessment is necessary. It is important for both the teachers and students. The students will find keen interest in preparing their course works. They will be able to solve their sampling problems with the help of this manual.

This manual primarily consists of 12 Chapters with definition, discussion, parameter estimating formulas, criteria of reliabilities, advantages, disadvantages and examples of all typical sampling designs. The typical sampling designs such as Simple Random Sampling, Stratified Random Sampling, Cluster Sampling, Two Stage Sampling, Stratified Two Stage Sampling, Double Sampling, Systematic Sampling, and their Probability Proportional to Size Sampling techniques are presented. A Chapter of Determination of Sample Size is also included. The variables and the parameters are defined in terms of village forest surveys. The terms and terminology, general remarks, wood and bamboo resources are also mentioned in village forest population. The general formulas are applied in village forest sampling so that researchers and general users can apply easily these formulas in their interested fields. A large number of examples is provided for estimating various parameters with reliabilities. The reliabilities such as percentage standard error, percentage accuracy, percentage bias and confidence intervals are provided for each of the estimates. Some estimates show reasonable reliabilities, but some do not. It depends on the extent of variation of

data. Data used in each example are collected during the study of Ph.D thesis in Chittagong district in Bangladesh. Variation of data related to auxiliary and main characters among the sampling units (households) is small in some areas and large in some other areas. The variables and the parameters are defined among the sampling units of the households. The criteria of reliabilities show a little high because of more variation of the characters among the households. The true values of the characteristics of the population are gathered from previous knowledge mentioned in examples. This manual is prepared consulting many related books, journals and reports shown in Bibliography.

Finally I hope and believe that this manual will be of great help to students, teachers and researchers working in the field of sampling techniques for assessment in respect of estimation of the parameters of different populations under consideration. Honest suggestions regarding improvement of this manual is certainly welcome. This manual may be revised in future with the valuable suggestions if there are any.

Chittagong,
Bangladesh.
June, 2015

Dr. Sheikh Sirajul Islam

ACKNOWLEDGEMENTS

First of all I am grateful to al-mighty Allah who has bestowed me knowledge and patience to complete this manual on Sampling Techniques for surveys.

I am indebted to Prof. Dr. A. K. M. Azharul Islam, hon'ble Vice-Chancellor of IIUC for his kind suport and patronization in publishing this book. I am also thankful to Prof. Dr. F. A. Sobhani, Director, Centre for Research and Publication (CRP) and Dean, Faculty of Business Studies, IIUC for his initiation and consideration to get the book published under the publication policy of CRP, IIUC. May Allah grant them the best reward.

I express my gratefulness to Late Dr. B.V. Sukhatme, Professor of Statistics, Iowa State University, Ames, Iowa, USA. He was my supervisor for M.S. degree in Statistics with major in Sampling Technique.

I am also grateful to Dr. Sultan Ahmad, Professor of Statistics, Chittagong University, Bangladesh. He was my supervisor for Ph.D. Thesis in Statistics based on Sampling Technique.

I acknowledge gratefully to Mr. Mohammad Nazmul Hoque, Assistant Professor (Statistics), Mrs. Israth Sultana, Lecturer (Statistics), Department of Business Administration, International Islamic University Chittagong (IIUC), Bangladesh and Engr. Md. Shahidul Islam, Maintenance Engineer, Bangladesh Forest Research Institute, Chittagong for their active help in calculating some examples in the manual.

I also acknowledge gratefully to Mr. Muhammad Sadek Shah, Calculator Assistant, Bangladesh Forest Research Institute, Chittagong and to Mr. Mohammad Fardu, Student of BBA 8th Semester Program, IIUC for their sincere efforts in calculating a few examples in the manual.

Thanks are due to a large number of colleagues and students of IIUC for their encouragement in completing the manual.

I acknowledge with thanks to Mr. Dibyendu Dasgupta, Stenographer and Mr. Ranjan Datta, Steno-Typist, BFRI, Chittagong for word processing of this manual.

Lastly, I remember inspiration of my wife, Dr. Nazma Begum, Medical Officer (Retired), Bangladesh Forest Research Institute, Chittagong to complete this manual.

CONTENTS

CHAPTER		Page
CHAPTER 1	INTRODUCTION	15
	1.1 Village wood and bamboo resources	15
	1.2 General remarks on village forest population	16
	1.3 Advantages of sampling over complete enumeration	17
	1.4 Sampling error and non-sampling error	18
	1.5 Sampling with replacement and without replacement	18
	1.6 Random sampling and non-random sampling	19
	1.7 Planning and organization of sample survey	20
CHAPTER 2	TERMINOLOGY AND CONCEPT	22
	2.1 Population and sample	22
	2.2 Sampling unit and frame	22
	2.3 Point estimation and interval estimation	22
	2.4 Village forest terminology	24
	2.5 The parameters and the variables	27
	2.6 Choice of sampling design	28
CHAPTER 3	SIMPLE RANDOM SAMPLING (SRS):	29
	3.1 Definition and discussion	29
	3.2 Estimation of parameters	29
	3.3 Advantages and disadvantages	31
	3.4 Example	32
CHAPTER 4	STRATIFIED RANDOM SAMPLING	34
	4.1 Definition and discussion	34
	4.2 Sampling Scheme	34
	4.3 Estimation of parameters	36
	4.4 Advantages	37
	4.5 Example	38
CHAPTER 5	RATIO METHOD OF ESTIMATION	40
	5.1 Ratio method	40
	5.2 Estimation of parameters	40
	5.3 Example	45
	5.4 Ratio method in stratified random sampling	46
	5.5 Separate ratio estimate	46
	5.6 Advantages	49
	5.7 Example	49

CHAPTER 6	CLUSTER SAMPLING	51
6.1	Definition and discussion	51
6.2	Estimation of parameters	51
6.3	Example	55
CHAPTER 7	TWO STAGE SAMPLING	59
7.1	Definition and discussion	59
7.2	Estimation of parameters	59
7.3	Example	64
CHAPTER 8	STRATIFIED TWO STAGE SAMPLING (Self-Weighted)	70
8.1	Definition and discussion	70
8.2	Estimation of parameters	70
8.3	Example	74
CHAPTER 9	DOUBLE SAMPLING	78
9.1	Definition and discussion	78
9.2	Estimation of parameters	79
9.3	Example	83
CHAPTER 10	SYSTEMATIC SAMPLING	86
10.1	Definition and discussion	86
10.2	Estimation of parameters	86
10.3	Approximation to standard error using simple random sample formula	87
10.4	Approximation to standard error using stratified random sample formula	88
10.5	Approximation to standard error using method of first difference	88
10.6	Approximation to standard error using method of second difference	89
10.7	Advantages and disadvantages	90
10.8	Example	91
CHAPTER 11	PROBABILITY PROPORTIONAL TO SIZE (PPS) SAMPLING	99
11.1	PPS Sampling (PPS of households with replacement)	100
11.1.1	Example	102
11.2	PPS Stratified Sampling (PPS of households with replacement)	105
11.2.1	Example	107

11.3	PPS Cluster Sampling (PPS of cluster with replacement)	110
11.3.1	Example	111
11.4	PPS Two Stage Sampling (PPS of first stage units with replacement and SRS of second stage units without replacement)	113
11.4.1	Example	114
11.5	PPS Stratified Two Stage Sampling (PPS of first stage units with replacement and SRS of second stage units without replacement)	116
11.5.1	Example	118
11.6	PPS Double Sampling (PPS of households with replacement at second phase)	122
11.6.1	Example	125
CHAPTER 12	DETERMINATION OF SAMPLE SIZE	128
12.1	Simple random sample	128
12.2	Stratified random sample	129
12.3	Two stage sample	131
12.1.1	Example	132
12.1.2	Example	134
12.2.1	Example	135
12.3.1	Example	136

LIST OF THE TABLES

	Page
Table 1. A simple random sample of 15 households.	32
Table 2. A sample of 34 stratified households.	38
Table 3. A sample of 15 households for ratio estimate	45
Table 4. A sample of 34 stratified households for separate ratio estimate	49
Table 5. A sample of 67 households at second stage of two stage sampling	64
Table 6. Self-weighting nature.	74
Table 7. A sample of 40 households at second phase of double sampling.	83
Table 8. A population of 64 households of systematic sampling.	91
Table 9. Two independent systematic samples each of size 16.	92
Table 10. A systematic sample of size 16 considered as simple random sample.	94
Table 11. A systematic sample of size 16 considered as stratified random sample.	95
Table 12. A systematic sample of size 16 analyzed by first difference method	96
Table 13. A systematic sample of size 16 analyzed by second difference method	97
Table 14. A PPS sample of 50 households with data of y_3 and x_2	102
Table 15. A PPS sample of 34 households allocated to 3 strata with data of y_3 and x_2	107
Table 16. A PPS sample of 3 clusters of 241 households with data of y_3	111
Table 17. A sample of 50 households at second stage with data of y_3	114
Table 18. A sample of stratified 50 households at second stage with data of y_3	118
Table 19. A PPS sample of 40 households with data of y_3 and x_2	125
Table 20. Data of volume of trees of a simple random sample of 10 households collected through pilot survey of 50 households.	132
Table 21. Data of volume of trees of a simple random sample of 10 households collected through pilot survey of 250 households.	134

LIST OF APPENDICES

	Page
Appendix – 1 : The field sheet	140
Appendix – 2 : The population of 50 households	145
Appendix – 3 : The population of 112 households stratified into 3 strata	146
Appendix – 4 : The sample of 3 clusters (villages) with 94 households	149
Appendix – 5 : The sample of 105 households at second stage in 3 strata	152
Appendix – 6 : Areas under the normal probability curve	155
Appendix – 7 : Percentage points of the t – distribution	157
Appendix – 8 : Percentage points of the χ^2 – distribution	158
Appendix – 9 : 95% Significance points of the F-distribution*	160
Appendix – 10 : 99% Significance points of the F-distribution*	162
Appendix – 11 : Random numbers	164

CHAPTER 1

INTRODUCTION

1.1 Village wood and bamboo resources

The area of village groves is very limited. The homestead land for tree and bamboo plantation is decreasing day by day. The village people have been conserving village grove resources for their own benefit for many years. The homestead which is located mainly on raised land contains a typical farming system. This farming system is composed of tree components, kitchen gardens, medicinal plants, ornamentals, pisciculture, livestock, duck and poultry raising etc. Village groves is a mixture of woodlots and other multipurpose fast growing trees, bamboos, canes, shrubs, etc. (Aleem, 1981). The groves grow mostly on homestead land around dwelling houses in the villages. The majority of the farmers prefer fruit and timber tree species for homestead plantation. They do not grow trees primarily for fuel despite fuel crisis. They prefer to grow fruit trees for fruit and fuel (also fodder in some cases). Less shady species are preferred as they do not hamper vegetable production, seedling raising and other homestead based activities (Abedin and Quddus, 1988). The exact area covering bamboo groves in the homestead is not known. It is assumed that 20% of the crown covered area of the homestead are occupied by bamboo. The primary utilities of the trees grown on the homestead are fruit juice, timber and building materials and the secondary utilities are fuel, fodder, manure, handicrafts, fencing, wind break, etc. The wood based industries such as match, tea chest, lime burning, pulp, brick burning, etc., are largely dependent on village wood and bamboos. Homesteads are more reliable and suitable for growing trees from physical and socio-economic view points. It is less vulnerable to flood hazards and thus more stable than crop fields. Moreover, the landless and marginal farmers have their own homestead and they consider it as fixed asset to meet various needs of the family. Food and fuel are of equal importance because most food needs fuel for cooking to make them edible. Purpose of the tree is that it is used as emergency resource. It is sold during financial hardship of a family. It provides cash for ceremonies and many other occasions such as marriages, school expenses of children and buying land and other assets (Abedin, et al., 1988). Wood, bamboo and cane together

support a host of cottage industries which not only offer greater opportunities for employment and income but also supply essential inputs which accelerate farm productivity. Any improvement of the resources will bring tangible benefits to both the economy of the rural population and the country's wood based industries through supply of raw materials.

1.2 General remarks on village forest population

Rural poverty and deforestation are two critical and interrelated problems in a country. The ever increasing population is creating growing demand for more food, fuel, timber and bamboo resulting in rapid deforestation. Deforestation process has a lot of adverse effect on environment. Droughts, floods, cyclone, tidal bore and other climatic hazards are believed to be the consequence of the deforestation process. The homestead land is being rapidly converted into agricultural land for food production. Consequently, the scarcity of homestead land for production of trees and bamboos is acute. Moreover, the large scale wood based cottage industries have tremendous pressure on homestead wood and bamboo consumption. The large number of landless and marginal poor farmers are cutting trees and bamboos for sale during their financial hardship. The forestry sector must recognize the regional disparities and should serve the real needs of the majority of the people by providing fuel wood, timber, bamboo and other forest products (Davidson, 1984). So, the homestead woodlot is of concern and proper management mechanism is to be adopted. People are burning cow dung and crop residues during acute shortage of fuel wood. There is a gap between consumption of fuel wood and consumption of other domestic energies (Byron, 1984). It means that the major source of domestic energy is not fuel wood but agricultural and other biomass residues (Ohlsson, 1984). It is because of shortage of fuel wood in the country. Bamboo-"the poor man's timber" is extremely important for the country (Byron, 1981). It is used as building materials, fuel, furniture, implements, scaffolding, ladders, tool handles, rafts and hundreds of various products. It also provides food, fodder and pulp. It is a fast growing plant and easy to propagate. It makes a useful hedge or windbreak and can prevent soil erosion. The national forest supplies bamboos to the paper mills, but the homestead supplies larger quantity of bamboos to the country for various utilities mentioned earlier.

Therefore, a necessary step is to be taken for more bamboo production quickly and economically in the homestead garden. It is important to get the information on present stocking, species composition, cropping pattern, planting and harvesting rate, land use system, etc., of homestead woodlot before undertaking extensive rural forestry sector planning. Without reliable information, proper planning as well as monitoring of these valuable resources are not possible. It is, therefore, necessary to apply a suitable sampling design so that reliable data could be generated for estimating the necessary parameters.

1.3 Advantages of sampling over complete enumeration

Sampling method: A sampling method is a scientific and objective procedure of selecting units from the population and provides a sample that is expected to be representative of the population. It is advisable to get as many information as possible.

Sample survey: A survey carried out on a properly selected representative sample is called a sample survey.

Complete enumeration: Complete enumeration is a method of enumerating or surveying whole population.

Advantages:

- (i) **Greater speed:** Since sample survey needs small part or fraction of a population, data can be collected and analyzed more quickly than complete enumeration.
- (ii) **Reduced cost:** For the same reason mentioned in (i), total cost and manpower of sample survey is less than complete enumeration.
- (iii) **Greater scope:** If a survey requires highly trained personnel and more sophisticated equipment for the collection and analysis of data, complete enumeration is impracticable. Sample survey has greater scope as compared to complete enumeration.
- (iv) **Greater accuracy:** The results of a sample survey are more reliable than those obtained from complete enumeration due to following reasons:
 - (a) It is possible to determine the extent of the sampling error in sample survey.

- (b) It is possible to control non-sampling error more effectively by employing qualified and trained personnel, better supervision and equipment for processing and analyzing limited data in sample survey.

1.4 Sampling error and non-sampling error

Sampling error: Sampling error is a type of error which arises in the sample survey. Sampling error is obvious, because only a part of a population is used to estimate parameter and draw inferences about a population. Even there will be some variation of values of different samples drawn from same population. This error is absent in complete enumeration. Sampling error is primarily due to the following reasons:

- i) Faulty selection (purposive or judgment) of a sample.
- ii) Substitution of a convenient sampling unit instead of trouble maker one while selection of a sample.
- iii) Faulty demarcation of sampling units.
- iv) Improper choice of the statistics for estimating parameters.

Non-sampling error: Non-sampling error is a type of error which arises in both sample survey and complete enumeration. Non-sampling error is due to following reasons:

- i) Negligence, forgetfulness and lack of knowledge both for investigator and respondent.
- ii) Faulty planning.
- iii) Response errors (Prestige bias, self-interest, memory failure, etc.)
- iv) Compiling and publication error.

1.5 Sampling with replacement and without replacement

Sampling with replacement: If we draw a random sample such that there is a chance of being included of a sampling unit more than once in a sample, it is said to be sampling with replacement. The units are drawn one by one with known probability of selection assigned to the unit at the first and each subsequent draw.

Example : (i) Let a population size be N.

N	N	N	
draw	draw	draw	...Continue unless the required
1 unit	1 unit	1 unit	sample size is selected.

Example: (ii) Let a population be $N = (a, b, c)$

No. of possible samples each of size 2 with replacement $= 3p_2 + 3 = 9$

Possible samples each of size 2: (a,a), (a,b), (a,c), (b,a), (b,b), (b,c), (c,a), (c,b), (c,c).

Sampling without replacement: If we draw a random sample such that no sampling unit occurs more than once in the sample, it is said to be sampling without replacement. The units are drawn one by one with known probability of selection assigned to the unit at the first and each subsequent draw.

Example: (i) Let a populations size be N

N	N-1	N-2	
draw	draw	draw	---continue unless the required
1 Unit	1 Unit	1 Unit	sample size is drawn

Example (ii): Let a population be $N = \{a, b, c\}$

No of possible samples each of size 2 without replacement ${}^3C_2 = 3$

Possible samples each of size 2: (a,b), (a,c), (b,c)

1.6 Random sampling and non-random sampling

Random sampling (or probability sampling): When sampling is performed by using scientific rules such that every unit in a population has some chance of being selected in a sample. The method of sampling is called random sampling. The precision and reliability of the estimates obtained from a random sample can be evaluated only for a probability sample. Thus the errors of sampling can be controlled in this case.

Non-random sampling: When a sample of units is selected not at

random rather purposively from a population is called non-random sampling. Such method of selection is likely to be biased and the results will thus be biased and unreliable. This method does not permit the evaluation of the precision of the estimates calculated from the sample. It is statistically unsound and should be discouraged.

Purposive sampling: Purposive method is a method which does not require a frame and selection procedure. A series of characteristics of the sampling units may be established as per requirement.

Example: The farmers (units) are selected purposively on the basis of a series of characteristics, viz. who produces maize and beans, had more than four children, two cows and 15 chickens from a same village.

Accident sampling: Accidental sampling is a method in which researchers select and interview the sampling units in a haphazard way.

Example: A accidental sample of first 10 individuals the researchers met when walking along a path to a village.

Quota sampling: Quota sampling is a method which is used to decide (stratum sample size) in advance that are wanted from each stratum and to instruct the enumerator to continue sampling until the necessary "quota" is obtained in each stratum. In this method, the enumerator does not select the units at random rather he selects them enables the quota to be filled quickly.

1.7 Planning and organization of sample survey

1. **Objectives of the survey:** A statement of the objectives is necessary. The characteristics on which information is to be collected and the desired accuracy of the estimates must be decided at the outset of the survey.
2. **Construction of a frame:** A frame is first requirement of probability sampling. It is to be constructed with the sampling units. Without this, no probability sampling is possible. The construction of a frame requires experience and clear knowledge.
3. **Choice of sampling units:** The proper choice of a sampling unit depends on a number of factors; the purpose of the survey, the characteristics, the variability among sampling units, the

sampling design, the field work plan and the total cost of the survey. The choice is also determined by practical convenience.

4. **Choice of sampling strategy:** To obtain unbiased estimates, a probability sample can be drawn such that every unit in population has a known probability of being included in the sample. There are several procedures for a probability sample each of which has several estimation methods. A sampling procedure together with an estimate is called a sampling strategy. We should choose such a strategy for which mean square error of the estimate is minimum for a given cost.
5. **Organization of the field work:** The entire success of a sample survey depends on reliability of field work. Even with the best sampling design without proper organization, the sample results may be incomplete and misleading. Proper selection of personnel, intensive training, clear instructions and proper supervision of field work are essential. The field investigator should correctly locate the selected units and record the necessary measurement as per instructions. The supervising staff should check part of their works in the field.
6. **Analysis of data:** On the basis of the sampling strategy and collected information, proper formula should be used to obtain estimates and precision of the estimates. Double check is always necessary for safeguard accuracy in the analysis.
7. **Pilot survey:** A statistically planned small scale “Pilot survey” should be conducted before undertaking a large scale survey. Such pilot survey will provide adequate knowledge about variability and afford opportunity to test and improve field procedures, train field workers and study the operational efficiency of a design. All data obtained through pilot survey may help in taking decision on the type of design and intensity of sampling. A data collection format can be pre-tested.

CHAPTER II

TERMINOLOGY AND CONCEPT

2.1 Population and sample

Population: A population is a group of units defined according to the aims of a survey.

Example: No. of individuals of a country, No. of trees of a forest, etc.

Population size: The number of units in a population is called population size and is denoted by N .

Sample: A sample is a part or section of a population under study.

Example: No. of trees of a section of a forest.

Sample size: The number of units in a sample is called sample size and is denoted by n . A sample should be selected in such a way so that a proper inference about a population can be drawn

2.2 Sampling unit and frame

Sampling unit: An element of a population is called a sampling unit and is denoted by 'i'.

Example: A tree of a forest area, a family, an individual, etc.

Frame: A frame is a list of all distinct sampling units in a population.

Example: No. of trees serially numbered in a forest area.

Characteristics of a sampling unit: A value or values of a variable defined on an unit is called characteristic of an unit and is denoted by " y_i ".

Example: Volume, diameter, height of a tree, etc.

2.3 Point estimation and interval estimation

Parameter: A parameter is an unknown constant of a population under study.

Example : Population mean, population total, population proportion, etc.

Statistical estimation: Statistical estimation is a mathematical procedure by which a parameter is estimated on the basis of sample observations.

Statistic: A statistic is a function of sample observations.

Example: Sample mean, sample standard deviation, sample median, sample mode, sample proportion, etc.

Estimator: A statistic used to estimate a parameter is called an estimator.

Estimate: The estimated value of an estimator is known as estimate.

Point estimation: Point estimation is a procedure for estimating a parameter of a population given by a single value. The single estimated value is called point estimate.

Interval estimation:

Interval estimation is a procedure for estimating a parameter of a population by two limits between which the estimate of the parameter lies. An interval estimate (or confidence interval (C.I)) consists of an upper confidence limit and a lower confidence limit which contains true population value.

Procedure for interval estimation:

Step-1: The particular estimate of a parameter is first determined.

Step-2: The confidence coefficient 95% or 99% is decided.

Step-3: The standard error of the estimate is calculated.

2.4 Village forest terminology

(1) **Household:**

A household means one or more persons normally living together in one house and eating together (khana). A household may have more than one houses. It is distinct from a family. Members of a family may live in different places but members of a household must share the same kitchen and live together.

(2) **Enumeration area:**

An enumeration area means a defined and specified area assigned to an enumerator. An enumeration area may consist of a few villages.

(3) **Homestead area:**

Homestead area is the land area of courtyard, pond and ditch excluding farm land of a household.

(4) **Compact tree garden area (or Compact plantation area or compact area):**

Compact tree garden area is an area of land covered with perennial fruit trees (including palm trees) and non-fruit trees (or mixture of both) grown together in a compact form of a household.

(5) **Bamboo plantation (or Bamboo area):**

Bamboo plantation is an area occupied by one or more bamboo clumps with compact growth of bamboos of a household.

(6) **Agriculture (farm) land:**

Agriculture land is an area where agriculture crops are grown. It is usually outside the homestead area, but sometimes may be close to homestead area.

(7) **Total land area:**

Total land area is an aggregate of homestead area, bamboo plantation, compact tree garden and agriculture land possessed by a household.

(8) **Compact trees:**

Compact trees are a compact growth of fruit trees (including palm trees) and non-fruit trees.

(9) **Scattered trees:**

Scattered trees are those trees grown or planted at different locations of a household area. The area occupied by each tree is not feasible for measurement. So, the area under scattered trees is not accounted.

(10) **Homestead woodlot:**

Homestead woodlot is considered to be a group of trees and bamboos grown or planted in and around homestead area. It is a combination of both compact and scattered trees.

(11) **Homestead vegetation:**

Homestead vegetation is a vegetation of fruit trees, non fruit trees, bamboos, shrubs, etc., in a homestead area.

(12) **Base diameter:**

Diameter at base of a tree is called base diameter.

(13) **Stump diameter:**

Diameter at stump height (normally 1.5 feet) of a tree is called stump diameter.

(14) **Breast height diameter:**

Breast height diameter of a tree means diameter measurement (cm) at the height of 4.5 feet.

(15) **Underbark volume:**

Volume of a tree when calculated excluding bark thickness is termed as underbark volume.

(16) **Stand table:**

A stand table is a bi-variate frequency table of a number of trees with respect to diameter and height classes.

(17) **Volume table:**

A volume table is a table of wood volume of trees estimated using different diameter and height classes. Each of the volume figures in the table indicates an average volume of a tree.

(18) **Local volume table:**

A local volume table is a volume table prepared on the basis of diameter at breast height. It is prepared for quick timber inventory.

(19) **Top diameter:**

The smallest diameter for commercial wood volume, usually 20 cm, is termed as top diameter.

(20) **Tip height:**

Tip height is a height of a tree from 20 cm top diameter to the last tip.

(21) **Volume equation:**

A regression function used to estimate wood volume from diameter, height and/or other tree variables.

(22) **Precision:**

Precision is defined here as percentage standard error (%SE). It is estimated by $(SE(Est.)/Est.) 100$.

(23) **Percentage accuracy:**

Percentage accuracy is defined here as percent deviation of the estimate from the true value after it is deducted from one hundred. It is estimated by $100 - \{(True\ value - Est.)/True\ value\} 100$ in case of unbiased estimates.

(24) **Percentage bias:**

Percentage bias is estimated by $(Bias\ (Est.)/Est.) 100$ in case of biased estimates.

(25) **Furnival index:**

Furnival index is an index estimated through multiplying square root of the residual mean square of a volume regression by geometric mean of the weight of the dependent variable. It was first developed by Furnival in 1961 for weighted models.

(26) Guiding variable (or Guiding factor or Influential characteristic):

A guiding variable is a variable which dictates all other variables under consideration. It is sometimes called guiding factor or influential characteristic.

2.5 The parameters and the variables

(Islam, 1996)

The parameters to be estimated are listed below:

- (a) Total land area (ha).
- (b) Total area (ha) utilized by compact trees.
- (c) Total homestead area (ha).
- (d) Total number of bamboo clumps.
- (e) Total number of trees (DBH>20cm) for selected species.
- (f) Total wood volume of trees (DBH > 20cm) for selected species.
- (g) Total volume of wood (m³) of compact trees.
- (h) Total number of bamboos (mature and immature) for selected species.
- (i) Total number of palm trees for selected species.
- (j) Average volume (compact trees) (DBH>20cm) for selected species per hectare,
- (k) Average number of bamboos (mature and immature) for selected species per clump.
- (l) Total number of households with bamboo plantation.

There is a provision in the questionnaire for defining many variables. But 9 variables are considered. These variables are chosen carefully such that complete enumeration and the chosen sampling methods can be directly applied for estimating the parameters. The main variables are defined as follows:

y_1 = Number of standing trees (DBH>20 cm) for selected species in a household.

y_2 = Volume (cubic metre) of standing trees (DBH>20 cm) for selected species in a household.

y_3 = Volume of standing compact trees (DBH>20 cm) for selected species in a household.

y_4 = Number of standing bamboos (mature and immature) for selected species in a household.

y_5 = Number of standing palm trees ($5\text{m} < \text{Height} < 20\text{m}$) for selected species in a household.

The variables are defined as variables having main characteristic. The variables y_2 and y_3 are defined through calculating under bark volume by Smalian formula (Jerram and Bourne, 1949).

2.6. Choice of sampling design

In forest and other surveys, the sampling units are different. Depending on the type and size of the sampling unit in such populations, the sampling designs are adopted. But in village woodlot survey, the sampling units are distinct and have various sizes. A household, a para, a village, a union, a thana and a district may be defined as a sampling unit. Since an up-to-date and reliable frame of these sampling units is always available, a number of typical probability sampling designs are used in village woodlot survey. A probability sampling can only assure precision and reliability of the estimates. The methods and applications of probability sampling designs are provided such that one may apply a suitable sampling design as per his need.

CHAPTER III

SIMPLE RANDOM SAMPLING (SRS)

3.1 Definition and discussion

Simple Random Sampling: Simple Random Sampling is a method of selecting a sample of units n from a population size N such that every one of the possible samples ${}^N C_n$ has an equal probability $1/{}^N C_n$ of being selected.

It is applied in each sub-population (stratum) of households which is internally homogeneous if heterogeneity exist in any area of homestead woodlot population. It is also applied in each of chosen sampling designs. It is not expensive to apply in the field. A sample of n households is drawn without replacement from a total of N households. The parameters are estimated and compared with the population values on the basis of this sample through computer. The following estimation formulas are used for the purpose:

3.2 Estimation of parameters

i) Estimate of total number of trees in the population:

$$\hat{Y}_1 = N\bar{y}_1 \text{ (unbiased), } \bar{y}_1 = 1/n \sum_{i=1}^n y_{1i}$$

$$SE(\hat{Y}_1) = S_{y_1} \sqrt{N(N-n)/n}$$

$$\%SE = \{SE(\hat{Y}_1)/\hat{Y}_1\} 100$$

$$\% \text{ Accuracy} = 100 - \{(Y_1 - \hat{Y}_1)/Y_1\} 100$$

$$95\% \text{ C.I. for } Y_1 : \hat{Y}_1 \pm Z_{.05} SE(\hat{Y}_1)$$

Similarly for $\hat{Y}_2, \hat{Y}_3, \hat{Y}_4$ & $\hat{Y}_5$.

Proportion estimation: An attribute is a qualitative characteristic which cannot be measured quantitatively, e.g. honesty, beauty, intellignency, etc. In the study of attribute, one may be interested in estimating the total number or proportion of (i) defective items in a large consignment of such items, (ii) smokers in a class, (iii) unemployed educated persons in a city.

(ii) Estimate of number of households with bamboo plantation:

$$\hat{NP} = Np \text{ (unbiased), } p = n/n$$

$$SE(\hat{NP}) = N\sqrt{\{(N-n)/N\}p(1-p)/(n-1)}$$

$$\% SE = \{SE(\hat{NP})/\hat{NP}\} 100$$

$$\% \text{ Accuracy} = 100 - \{(NP - \hat{NP})/\hat{NP}\} 100$$

$$95\% \text{ C.I. for NP : } \hat{NP} \pm Z_{.05} SE(\hat{NP})$$

Where,

N = Size of the population

n = Size of the sample

sx_i = Sample standard deviation for variable x_i .

sy_i = Sample standard deviation for variable y_i .

P = Population proportion for number of households with bamboo plantation.

p = Estimate of P .

n_1 = Number of households with bamboo plantation in the sample.

3.3 Advantages and disadvantages

Advantages:

SRS is the simplest form of sampling and is the basis for many other sampling methods. Since the sampling units are selected at random, the element of subjectivity and personal bias is completely eliminated. As such simple random sample is more representative as compared to purposive sampling. It is mostly applicable for an initial survey in an investigation and for studies which involve sampling from a small area where sample size is small. After acquiring knowledge regarding a population through initial survey, other methods which are likely to be efficient and convenient for organizing the survey may be adopted. It is an inexpensive sampling method.

Disadvantage:

Simple Random Sample is not suitable when the distribution of sampling units in a population is irregular.

3.4 Example:

A population of 50 households is considered in an area of Chittagong district in Bangladesh (appendix-2). A sample of 15 households is selected at random. Data of volume (m³) of compact trees (y_3) and no. of bamboo clumps (x_5) each of sample households are collected.

Estimate the following parameters with reliability of estimate in each case

- (a) Total volume of compact trees
- (b) No. of households with bamboo plantation

Solution:

Table-1. A simple random sample of 15 households

i	x_5	y_3
1	0	2.81
2	0	2.41
3	0	8.15
4	1	2.49
5	1	0.19
6	1	5.15
7	1	7.14
8	0	1.54
9	0	1.15
10	0	11.02
11	1	14.83
12	1	1.71
13	0	4.09
14	0	3.93
15	0	3.90
Total		

$$\bar{y}_3=4.70, s_3=4.03$$

$$(a) \quad \hat{Y}_3 = \bar{N}y_3 = 5 \times 4.70 = 235 \text{ m}^3, \quad \bar{y}_3 = 1/n \sum_{i=1}^n y_{3i} = 4.70$$

$$SE(N\bar{y}_3) = 50\sqrt{0.7 \times 1.0828} = 43.5287, \quad SE(N\bar{y}_3) = NSE(\bar{y}_3),$$

$$SE(N\bar{y}_3) = \sqrt{(N-n)/Nn} \times s^2_{y_3}$$

$$\% SE(N\bar{y}_3) = 43.5287/235 \times 100 = 18.5\%$$

$$\% \text{Accuracy} = 100 - \{192.04 - 235/192.04 \times 100\} = 78\%$$

95% Confidence Interval:

$$\text{Lower limit} : 235 - 1.96 \times 43.5287 = 149.68$$

$$\text{Upper limit} : 235 + 1.96 \times 43.5287 = 320.32$$

$$(b) \quad \hat{NP} = Np = 50 \times 6/15 = 50 \times .4 = 20$$

$$\begin{aligned} SE(Np) &= N\sqrt{(N-n)/N} \sqrt{P(1-P)/(n-1)} \\ &= 50\sqrt{.7} \times .0171 = .1095 \end{aligned}$$

$$\%SE(Np) = .1095/20 \times 10 = .55\%$$

$$NP = 50 \times 23/50 = 23$$

$$\% \text{Accuracy} = 100 - 23 - 20/23 \times 100 = 87\%$$

95% Confidence Interval :

$$\text{Lower limit} : 20 - 1.96 \times .1095 = 19.79$$

$$\text{Upper limit} : 20 + 1.96 \times .1095 = 20.22$$

CHAPTER IV

STRATIFIED RANDOM SAMPLING

4.1 Definition and discussion

Stratified random sampling: Stratified random sampling is a sampling procedure which consists of dividing a population in many sub-divisions and drawing a simple random sample from each sub-division. Each sub-division is called stratum.

The precision of an estimate depends not only on sample size, but also on variability of a population. Stratification should be done in such a way so that the heterogeneous population may be sub-divided into many strata each of which is internally homogeneous. There must be at least two units in each stratum for estimating variability.

Stratified Random Sampling is a common procedure in homestead woodlot survey. The basic idea in this sampling design is the stratification of the heterogeneous homestead woodlot population into the strata each of which is internally homogenous with respect to the characteristics under study. Stratification is always done in conducting sample survey in a heterogeneous population like homestead woodlot. It serves many useful purposes. It may be desired for administrative convenience. It, by natural characteristics, helps in improving the sampling design.

It makes possible to use different sampling designs in different strata. It ensures adequate representation of various strata. It also ensures selection of a better cross-section of the population than that under unstratified population. It brings a gain in precision in estimating a parameter.

4.2 Sampling scheme

Let a population size N be stratified in the following way:

L_i :	Stratum-1	Stratum-2	-----	Stratum-L
N_i :	N_1	N_2	-----	N_L
W_i :	$W_1=N_1/N$	$W_2=N_2/N$	-----	$W_L=N_L/N$
n_i :	nW_1	nW_2	-----	nW_L

(with proportional allocation)

Sample mean and variance of

$$y_i : \bar{y}_1, s_1^2, \bar{y}_2, s_2^2, \dots, \bar{y}_L, s_L^2$$

Allocation of Sample Size:

The allocation of sample size for different strata is done in the following ways:

- (a) Proportion allocation : $n_i = nW_i$,
- (b) Equal allocation : $n_i = n/L$,
- (c) Neyman allocation : $n_i = W_i s_i / \sum_{i=1}^L W_i s_i$,
- (d) Optimal allocation : $n_i = Co \cdot W_i s_i / \sqrt{C_i} / \sum_{i=1}^L W_i s_i \sqrt{C_i}$

Where $W_i = N_i/N$, L = no. of strata, s_i = standard deviation of i th stratum.

n = sample size, n_i = sample size of i th stratum.

C_i = Cost of collecting data and other cost per unit.

$$Co = \text{Total cost} = \sum_{i=1}^L C_i n_i$$

An efficient method of allocation for minimum cost is to take large samples from larger strata where sampling is cheap and variability is more. When this method is not possible, proportion allocation may be considered. Estimator corresponding to each allocation with smallest standard error gives best estimator with the allocation.

4.3 Estimation of parameters

(i) **Estimate of total number of trees in the population:**

$$\hat{Y}_1 = N\bar{y}_{1st} \text{ (Unbiased), } \bar{y}_{1st} = (1/N) \sum_{i=1}^L N_i \bar{y}_{1i}$$

$$SE(\hat{Y}_1) = N\sqrt{\{(1/n) \sum_{i=1}^L w_i s_{1i}^2 - (1/N) \sum_{i=1}^L w_i s_{1i}^2\}}$$

$$\%SE = \{SE(\hat{Y}_1)/\hat{Y}_1\}100$$

$$\% \text{ Accuracy} = 100 - \{Y_1 - \hat{Y}_1\}/Y_1\}100$$

$$95\% \text{ C I for } Y_1 : \hat{Y}_1 \pm Z_{.05} SE(\hat{Y}_1)$$

Similarly for $\hat{Y}_2, \hat{Y}_3, \hat{Y}_4$ & $\hat{Y}_5$

(ii) **Estimate of number of households with bamboo plantation:**

$$\hat{NP} = NP_{st} = \sum_{i=1}^L N_i p_i \text{ (unbiased), } p_i = n_{1i}/n_i$$

$$SE(\hat{NP}_{st}) = \sqrt{\sum N_i(N_i - n_i) p_i(1-p_i)/n_i}$$

$$\% SE = \{SE(NP_{st})/NP_{st}\}100$$

$$\% \text{ Accuracy} = \{(NP - \hat{NP}_{st})/NP\}100$$

$$95\% \text{ C.I. for } \hat{NP} : NP_{st} \pm Z_{.05} SE(\hat{NP}_{st})$$

Where,

- L = Number of strata
- N_i = Number of households for i th stratum in the population.
- W_i = N_i/N .
- y_{st} = Stratified sample mean for y -variable.
- $S^2 y_i$ = Sample variance for y -variable for i th stratum.
- P_{st} = Stratified sample proportion of households having bamboo plantation.
- N_{ji} = Number of households having bamboo plantation for i th stratum in the sample.
- n_i = Number of households for i th stratum in the sample.

4.4 Advantages

- (a) Stratified random sampling is more representative because of internally homogeneous strata,
- (b) Stratified random sampling is more convenient administratively,
- (c) Stratified random sampling solves the problem of surveying a human population living in hotels, hospitals, prisons, etc,
- (d) Stratified random sampling provides gain in precision in the estimates of population parameter.

4.5 Example

A population of 112 households in an area of Chittagong district in Bangladesh is stratified in to 3 strata in respect of regional effects (Appendix-3). Data of same variables as in example 3.4 are collected. A sample of 34 households (30%) is allocated to 3 strata by proportional allocation (Table-2). Estimate the following parameters:

- total volume (m^3) of compact trees
- total no. of households with bamboo plantation.

Solution:

Table-2. A sample of 34 stratified households

Stratum -1			Stratum -2			Stratum -3		
i	y_3	x_5	i	y_3	x_5	i	y_3	x_5
1	2.94	1	1	4.38	0	1	12.06	0
2	3.61	0	2	5.84	4	2	7.25	0
3	16.53	0	3	4.67	0	3	11.49	1
4	8.07	0	4	34.79	1	4	5.88	0
5	2.02	1	5	3.60	0	5	9.61	0
6	18.30	1	6	9.62	0	6	4.69	0
7	7.63	0	7	2.70	1	7	10.62	0
8	21.31	1	8	9.19	0	8	7.12	1
9	11.61	1	9	4.24	1	9	13.08	0
10	15.71	4	10	5.03	4	10	7.60	0
11	2.52	1	11			11	8.15	1
12	.55	0	12			12	9.67	0

Stratum-1

Stratum-2

Stratum-3

$$N_1 = 40$$

$$N_2 = 32$$

$$N_3 = 40$$

$$W_1 = 40/112 = .36$$

$$W_2 = 32/112 = .29$$

$$W_3 = .36$$

$$n_1 = 35 \times .35 = 12$$

$$n_2 = 35 \times .29 = 10$$

$$n_3 = .35 \times 36 = 12$$

$$\bar{y}_3 = 9.23 \quad s_3 = 7.24$$

$$\bar{y}_3 = 8.36 \quad s_3 = 9.56$$

$$\bar{y}_3 = 8.88 \quad s_3 = 2.55$$

$$p_1 = n_{10}/n_1$$

$$p_2 = n_{20}/n_2$$

$$p_3 = n_{30}/n_3$$

$$(a) \bar{Y}_3 = N\bar{y}_{3st} = 1001.79, \bar{y}_{3st} = (1/N \sum_{i=1}^L N_i \bar{y}_{3i} = 8.94,$$

$$SE(\hat{Y}_3) = N\sqrt{\{(1/n) \sum_{i=1}^L w_i s_{1i}^2 - (1/N) \sum_{i=1}^L w_i s_{1i}^2\}} = 15.86$$

$$\%SE = 15.86/1001.79 \times 100 = 1.6\%$$

$\%Accuracy = 91\%$, $\bar{Y}_3$ is taken from previous information.

95% Confidence Interval :

$$\text{Lower limit} = \hat{Y}_3 - Z_{.05} SE(\hat{Y}_3) = 970.70$$

$$\text{Upper limit} = \hat{Y}_3 + Z_{.05} SE(\hat{Y}_3) = 1032.09$$

$$(b) \hat{NP} = Np_{st} = 52.4, p_{st} = .47$$

$$SE(Np_{st}) = 7.5584, SE(Np_{st}) = \sqrt{\sum_{i=1}^L N_i(N_i - n_i) p_i(1-p_i)/n_i}$$

$$\%SE(Np_{st}) = 7.5584/52.4 \times 100 = 14\%$$

95% Confidence Interval :

$$\text{Lower limit} : Np_{st} - Z_{.05} SE(Np_{st}) = 38$$

$$\text{Upper limit} : Np_{st} + Z_{.05} SE(Np_{st}) = 67$$

CHAPTER V

RATIO METHOD OF ESTIMATION

5.1 Ratio method

In simple random sampling and stratified random sampling, estimation of parameters with reliabilities using main character y_i has been formulated and discussed. Definition of main character y_1, y_2, y_3, y_4 and y_5 in respect of village forest is introduced in section 2.5. But information on auxiliary character x_i in respect of village forest is also available and given below:

x_1 = Area (ha) of total land (agriculture land+compact tree garden + bamboo plantation+homestead area) in a household.

x_2 = Area (ha) of compact tree garden in a household

x_4 = Homestead area (ha) (courtyard area+pond area+ditch area) in a household.

x_5 = Number of bamboo clumps for selected species in a households.

There are two methods of estimation such as the ratio method and the regression method using auxiliary variable correlated with main variable. The two methods give more reliable estimates of parameters. Here only ratio method of estimation is discussed considering sampling with equal probability for selection of units. The method is also used in stratified random sampling applying separate ratio method. In ratio estimate, total of auxiliary character must be known from previous information or by estimation. The ratio method is extremely used in estimating main character per unit and total of main character. Estimation of parameters by the ratio method and corresponding reliabilities are given below.

5.2 Estimation of parameters

If the relationship between y_i and x_i is linear passing through origin and they are highly positively correlated ($r > .5$), ratio estimator provides more precise estimates. Assume X_i is known.

(i) Estimate of volume yield per hectare for compact trees:

$$r = (\sum_{i=1}^n y_3 x_2 - n \bar{y}_3 \bar{x}_2) / (s_{y_3} \cdot s_{x_2})$$

$$b = (\sum_{i=1}^n y_3 x_2 - n \bar{y}_3 \bar{x}_2) / s^2 x_2, a = \bar{y}_3 - b \bar{x}_2$$

$$\hat{R}_1 = \bar{y}_3 / \bar{x}_2 \text{ (biased)}$$

$$SE(\hat{R}_1) = s_{y_3, x_2} / (\bar{X}_2) \sqrt{\{(N-n)/Nn\}} \quad (\text{Islam, 1977})$$

$$\%SE = \{SE(\hat{R}_1) / \hat{R}_1\} 100$$

$$\text{Bias}(\hat{R}_1) = \sqrt{\{v(\bar{x}_2)v(\hat{R}_1)\}} / \bar{X}_2, \quad \sqrt{v(\bar{x}_2)} = s_{x_2} / \sqrt{n} \text{ and } \sqrt{v(\hat{R}_1)}$$

$$= s_{y_3, x_2} / \sqrt{n} \text{ for large } N.$$

$$\% \text{ Bias} = \{\hat{\text{Bias}}(\hat{R}_1) / \hat{R}_1\} 100$$

$$95\% \text{ C.I. for } R_1 : \hat{R}_1 \pm Z_{.05} SE(\hat{R}_1)$$

(ii) Estimate of total volume of compact trees:

$$\hat{Y}_{3R} = (\bar{y}_3 / \bar{x}_2) X_2 \text{ (biased)}$$

$$SE(\hat{Y}_{3R}) = s_{y_3, x_2} \sqrt{\{N(N-n)/n\}}$$

$$\% SE = \{SE(\hat{Y}_{3R}) / \hat{Y}_{3R}\} 100$$

$$\text{Bias}(\hat{Y}_{3R}) = (X_2) \text{Bias}(\hat{R}_1),$$

$$\% \text{ Bias} = \{\hat{\text{Bias}}(\hat{Y}_{3R}) / \hat{Y}_{3R}\} 100$$

$$95\% \text{ C.I. for } Y_{3R} : \hat{Y}_{3R} \pm Z_{.05} SE(\hat{Y}_{3R})$$

(iii) Estimate of number of bamboos per clump :

$$\hat{R}_2 = \bar{y}_4 / \bar{x}_5 \text{ (biased)}$$

$$SE(\hat{R}_2) = s_{y_4, x_5} / \bar{X}_5 \sqrt{\{(N-n)/Nn\}}$$

$$\% SE = \{SE(\hat{R}_2) / \hat{R}_2\} 100$$

$$Bias(\hat{R}_2) = \sqrt{\{v(\bar{x}_5)v(\hat{R}_2)\}} / \bar{X}_5, \sqrt{v(\bar{x}_5)} = s_{x_5} / \sqrt{n}$$

$$\text{and } \sqrt{v(\hat{R}_2)} = s_{y_4, x_5} / \sqrt{n} \text{ for large } N.$$

$$\% Bias = \{Bias(\hat{R}_2) / \hat{R}_2\} 100$$

$$95\% \text{ C.I. for } R_2 : \hat{R}_2 \pm Z_{.05} SE(\hat{R}_2)$$

(iv) Estimate of number of bamboos:

$$\hat{Y}_{4R} = (\bar{y}_4 / \bar{x}_5) X_5 \text{ (biased)}$$

$$SE(\hat{Y}_{4R}) = s_{y_4, x_5} \sqrt{\{N(N-n)/n\}}$$

$$\% SE = \{SE(\hat{Y}_{4R}) / \hat{Y}_{4R}\} 100$$

$$Bias(\hat{Y}_{4R}) = (X_5) Bias(\hat{R}_2).$$

$$\% Bias = \{Bias(\hat{Y}_{4R}) / \hat{Y}_{4R}\} 100$$

$$Bias(\bar{Y}_{4R}) = (\bar{X}_5) Bias(\hat{R}_2)$$

$$95\% \text{ C.I. for } Y_{4R} : \hat{Y}_{4R} \pm Z_{.05} SE(\hat{Y}_{4R})$$

(v) Estimate of volume per tree:

$$\hat{R}_3 = \bar{y}_2/\bar{x}_1 \text{ (biased)}$$

$$SE(\hat{R}_3) = s_{y_2.y_1}/Y_1 \sqrt{\{(N-n)/Nn\}}$$

$$\% SE = \{SE(\hat{R}_3)/\hat{R}_3\}100$$

$$\text{Bias}(\hat{R}_3) = \sqrt{\{v(\bar{y}_1)v(\hat{R}_3)\}}/\bar{Y}_1, \sqrt{v(\bar{y}_1)} = s_{y_1}/\sqrt{n}$$

$$\text{and } \sqrt{v(\hat{R}_3)} = s_{y_2.y_1}/\sqrt{n} \text{ for large } N.$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{R}_3)/\hat{R}_3\}100$$

$$95\% \text{ C.I. for } R_3 : \hat{R}_3 \pm Z_{.05} SE(\hat{R}_3)$$

(vi) Estimate of total volume:

$$\hat{Y}_{2R} = (\bar{y}_2/\bar{y}_1)Y_1 \text{ (biased)}$$

$$SE(\hat{Y}_{2R}) = s_{y_2.y_1} \sqrt{\{N(N-n)/n\}}$$

$$\% SE = \{SE(\hat{Y}_{2R})/\hat{Y}_{2R}\}100$$

$$\text{Bias}(\hat{Y}_{2R}) = (Y_1)\text{Bias}(\hat{R}_3),$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{Y}_{2R})/\hat{Y}_{2R}\}100$$

$$\text{Bias}(\bar{Y}_{2R}) = \bar{Y}_1 \text{Bias}(\hat{R}_3)$$

$$95\% \text{ C.I. for } Y_{2R} : \hat{Y}_{2R} \pm Z_{.05} SE(\hat{Y}_{2R})$$

Where,

N = Size of the population.

n = Size of the sample.

s_{x_i} = Sample standard deviation for variable x_i .

s_{y_i} = Sample standard deviation for variable y_i .

$s_{y_i \cdot x_i} = \sqrt{\{1/(n-1) \sum_{i=1}^n (y_i - \hat{R}x_i)^2\}}$

X_i = Population total for variable x_i .

Y_i = Population total for variable y_i .

R_i = Population ratio for variable y_i and x_i .

5.3 Example

A population of 50 households is considered in an area of Chittagong district in Bangladesh (appendix-2). A sample of 15 households is selected at random. Data of an area (ha) utilized by compact trees (x_2) and volume (m^3) of compact trees (y_3) for each of sample households are collected.

Estimate volume of compact trees per hectare

Solution:

(a) Table: 3. A sample of 15 households for ratio estimate.

i	x_2	y_3	$(y_3 - R x_2)^2$
1	.05	2.81	9.39
2	.04	2.41	5.24
3	.02	8.15	33.64
4	.01	2.49	1.73
5	.04	0.19	20.34
6	.05	5.15	0.53
7	.03	7.14	35.58
8	.01	1.54	0.13
9	.05	1.15	22.33
10	.08	11.02	2.62
11	.06	14.83	60.53
12	.01	1.71	0.29
13	.02	4.09	3.03
14	.02	3.93	2.49
15	.04	3.90	0.64
Total			198.51

$$\bar{x}_2 = .04, s_2 = .02, \bar{y}_3 = 4.70, s_3 = 4.03, \hat{R} = 117.5$$

$$\hat{R} = \bar{y}_3 / \bar{x}_2 = 4.70 / .04 = 117.5 m^3/ha.$$

$$S_{y_3 \cdot x_2} = \sqrt{1/(n-1) \sum (y_3 - \hat{R} x_2)^2} = 3.77$$

$$SE(\hat{R}) = S_{y_3 \cdot x_2} / \bar{x}_2 \sqrt{(N-n)/Nn} = 3.77 / .04 \sqrt{(50-15)/50} = .2953$$

$$\%SE(\hat{R}) = .2953 / 117.5 \times 100 = .0025 = 0.25\%$$

$$SE(\bar{x}_2) = \sqrt{(N-n)/N} s_2 = \sqrt{.0476} \times .0004 = .00044$$

$$\text{Bias}(\hat{R}) = SE(\hat{R}) \cdot SE(\bar{x}_2) / \bar{X}_2 = .2953 \times .00044 / 2.76 = .000047$$

$\bar{X}_2$ is given from previous information.

$$\% \text{ Bias}(\hat{R}) = .000047 / 117.5 \times 100 = .00004$$

95% Confidence Interval :

$$\text{Lower limit : } \hat{R} - 1.96 \times SE(\hat{R}) = 116.92$$

$$\text{Upper limit : } \hat{R} + 1.96 \times SE(\hat{R}) = 118.08$$

Similar examples maybe given for estimation of $R_2, R_3, Y_{3R}, Y_{4R}, Y_{2R}$.

5.4 Ratio method in stratified random sampling

Suppose $\bar{y}_i$ and $\bar{x}_i$ are sample means in i th stratum for y , the main character and for x , the auxiliary character in stratified random sampling done in Chapter – IV. The ratio of y and x and total of y may be estimated by separate ratio estimation and combined ratio estimation methods. Here separate ratio estimation method is preferred and formulated. It needs to obtain R_i in each stratum and average of y per unit and taking these estimates in all strata to obtain a single estimate for whole population. The notations and formulas are given in respect of village forest.

5.5 Separate ration estimate

If separate Ratio Estimate for R_1 : If $R_i = Y_{3i} / X_{2i}$ differs largely in strata and sample size in each stratum is large ($n_i > 20$).

(i) Estimate of volume of compact trees per hectare:

$$\hat{R}_{1(\text{sep})} = \hat{Y}_{3(\text{sep})}/X_2 \text{ (biased),}$$

$$\hat{Y}_{3(\text{sep})} = \sum_{i=1}^L (\bar{y}_{3i}/\bar{x}_{2i}) X_{2i} = \sum_{i=1}^L \hat{R}_{1i} X_{2i}$$

$$SE(\hat{R}_{1(\text{sep})}) = \sqrt{[(1/X_2^2) \sum_{i=1}^L \{ (N_i(N_i-n_i)/n_i) (s_{iy_3}^2 + \hat{R}_{1i}^2 s_{ix_2}^2 - 2\hat{R}_{1i} s_{iy_3x_2}) \}]}$$

$$\hat{R}_{1i} = (y_{3i}/x_{2i})$$

$$\%SE = \{SE(\hat{R}_1)/\hat{R}_1\} 100$$

$$\text{Bias}(\hat{R}_1) = \sum_{i=1}^L (1-f_i/n_i x_{2i}) \{ \hat{R}_i S_{ix_2}^2 - s_{ix_2iy_3} \}$$

$$\% \text{Bias} = \{ \text{Bias}(\hat{R}_1)/\hat{R}_1 \} 100$$

$$f_i = n_i/N_i$$

$$95\% \text{ C I for } R_1 : \hat{R}_{1(\text{sep})} \pm Z_{.05} SE(\hat{R}_{1(\text{sep})})$$

Separate Ratio Estimate for R_2 : If $R_i = Y_{4i}/X_{5i}$ differs largely in strata and sample size in each stratum is large ($n_i > 20$).

ii) Estimate of number of bamboos per clump:

$$\hat{R}_{2(\text{sep})} = \hat{Y}_{4(\text{sep})}/X_5 \text{ (biased),}$$

$$\hat{Y}_{4(\text{sep})} = \sum_{i=1}^L (\bar{y}_{4i}/\bar{x}_{5i}) X_{5i} = \sum_{i=1}^L \hat{R}_{2i} X_{5i}$$

$$SE(\hat{R}_{2(\text{sep})}) = \sqrt{[(1/X_5^2) \sum_{i=1}^L \{ (N_i(N_i-n_i)/n_i) (s_{iy_4}^2 + \hat{R}_{2i}^2 s_{ix_5}^2 - 2\hat{R}_{2i} s_{iy_4x_5}) \}]}$$

$$\hat{R}_{2i} = (y_{4i}/x_{5i})$$

$$\%SE = \{SE(\hat{R}_2)/\hat{R}_2\} 100$$

$$\text{Bias}(\hat{R}_2) = \sum_{i=1}^L (1-f_i/n_i \bar{x}_{5i}) \{ \hat{R}_{2i} S_{ix_5}^2 - s_{ix_5iy_4} \}$$

$$\% \text{Bias} = \{ \text{Bias}(\hat{R}_2) / \hat{R}_2 \} 100$$

$$f_i = n_i / N_i$$

$$95\% \text{ C I for } R_2 : \hat{R}_{2(\text{sep})} \pm Z_{.05} \text{SE}(\hat{R}_{2(\text{sep})})$$

Justification for (a) and (b)

The separate ratio estimate has higher precision than the combined ratio estimate. But the bias in the former estimate is larger than the latter. Besides, the sample size is large ($n_i > 20$) in each stratum and R_i differs considerably from stratum to stratum. The separate ratio estimate may, therefore, be accepted. It is almost as efficient as the combined ratio estimate.

Where,

L = Number of strata

N_i = Number of households for i th stratum in the population.

W_i = N_i / N .

$\bar{x}_{st}$ = Stratified sample mean for x -variable.

$\bar{y}_{st}$ = Stratified sample mean for y -variable.

s^2_{xi} = Sample variance for x -variable for i th stratum.

s^2_{yi} = Sample variance for y -variable for i th stratum.

$s_{yi.xi} = (1/n_i - 1) \{ \sum_{i=1}^{n_i} y_i x_i - n_i \bar{y}_i \bar{x}_i \}$.

= Sample covariance for y and x variable in i th stratum.

5.6 Advantages

Ratio method of estimation is recommended when there is high correlation between y and x and the regression between them is linear and passes through origin. The ratio estimate is biased although the bias is negligible in large samples. Thus if the sample size is fairly large, the ratio method of estimation is quite satisfactory.

5.7 Example

A population of 112 households in an area of Chittagong district in Bangladesh is stratified in to 3 strata in respect of regional effects. Data of same variables as in example 3.4 are collected. A sample of 34 households (30%) is allocated to 3 strata by proportional allocation (Table-4).

Estimate volume (m^3) of compact trees per hectare.

Solution:

Table-4. A sample of 34 stratified households for separate ratio estimate

Stratum -1			Stratum -2			Stratum -3		
i	x_2	y_3	i	x_2	y_3	i	x_2	y_3
1	.04	2.94	1	.06	4.38	1	.04	12.06
2	.02	3.61	2	.08	5.84	2	.04	7.25
3	.08	16.53	3	.08	4.67	3	.04	11.49
4	.04	8.07	4	.03	34.79	4	.05	5.88
5	.06	2.02	5	.02	3.60	5	.04	9.61
6	.32	18.30	6	.03	9.62	6	.02	4.69
7	.11	7.63	7	.01	2.70	7	.02	10.62
8	.08	21.31	8	.04	9.19	8	.03	7.12
9	.05	11.61	9	.05	4.24	9	.09	13.08
10	.13	15.71	10	.07	5.03	10	.02	7.60
11	.06	2.52	11			11	.10	8.15
12	.02	.55	12			12	.04	9.67

Stratum-1	Stratum-2	Stratum-3
$N_1 = 40$	$N_2 = 32$	$N_3 = 40$
$W_1 = 40/112 = .36$	$W_2 = 32/112 = .29$	$W_3 = .36$
$n_1 = 35 \times .35 = 12$	$n_2 = 35 \times 29 = 10$	$n_3 = .35 \times 36 = 12$
$\bar{x}_2 = .084 \quad s_2 = .081$	$\bar{x}_2 = .047 \quad s_2 = .025$	$\bar{x}_2 = 0.44 \quad s_2 = .026$
$\bar{y}_3 = 9.23 \quad s_3 = 7.24$	$\bar{y}_3 = 8.36 \quad s_3 = 9.56$	$\bar{y}_3 = 8.88 \quad s_3 = 2.55$
$p_1 = n_{10}/n_1$	$p_2 = n_{20}/n_2$	$p_3 = n_{30}/n_3$
$\hat{R} = 109.8809$	$\hat{R} = 177.9149$	$\hat{R} = 201.8182$
$X_2 = 3.36$	$X_2 = 1.51$	$X_2 = 1.76$
$s_{x_2, y_3} = .3489$	$s_{x_2, y_3} = -.0484$	$s_{x_2, y_3} = .0202$

$$\hat{R}_{1\text{sep}} = \sum (\bar{y}_{3i} / \bar{x}_{2i}) X_2 = 149.78$$

$$SE(\hat{R}_{1\text{sep}}) = \sqrt{9970.99/43.96} \\ = 15.0611$$

$$\% SE(\hat{R}_{1\text{sep}}) = SE(\hat{R}_{1\text{sep}}) / \hat{R}_{1\text{sep}} \times 100 \\ = 15.0511 / 149.78 \times 100 \\ = 10.06\%$$

X_2 is taken from previous information

$$\% \text{Accuracy} = 15.0611 / 149.78 \times 100 = 100\%$$

95% Confidence Interval for $\hat{R}_{\text{sep}}$:

$$\text{Lower limit : } R_{\text{sep}} - Z_{.05} SE(R_{\text{sep}}) = 120.26$$

$$\text{Upper limit : } R_{\text{sep}} + Z_{.05} SE(R_{\text{sep}}) = 179.30$$

Similar example maybe given for estimation of $R_{2\text{sep}}$

CHAPTER VI

CLUSTER SAMPLING

6.1 Definition and discussion

Cluster Sampling is applied in homestead woodlot survey mainly for two reasons namely; (i) when no complete and up-to-date list of sampling units (households, segment of land, etc.) are available and (ii) when cost of conducting the survey is limited. These two reasons are obvious in conducting woodlot survey. Sometimes it is difficult to get the up-to-date list of the households in a village and the fund placed for conducting the survey is not enough. The main function in cluster sampling is to form the clusters. A sample of small size cluster gives more precise results than a sample of large size cluster. Cluster should be as heterogeneous as possible and one cluster should look very much alike to one another. Cluster sampling is more convenient and less costly for a given number of sampling units it is chosen because clusters solve the problem of constructing the frame.

6.2 Estimation of parameters

(i) Estimate of total area of land:

$$\hat{X}_1 = M \bar{x}_1, \quad \bar{x}_1 = (1/n\bar{M}) \sum_{i=1}^n M_i \bar{x}_{1i} \text{ (unbiased),}$$

$$\bar{x}_{1i} = 1/M_i \sum_{j=1}^{M_i} x_{1ij}, \quad \bar{M} = 1/N \sum_{i=1}^N M_i = M/N$$

$$SE(\hat{X}_1) = M \sqrt{\{(N-n)/N\} \{1/n(n-1)\} \sum_{i=1}^n \{(M_i \bar{x}_{1i}/\bar{M}) - \bar{x}_1\}^2}$$

$$\%SE = \{SE(\hat{X}_1)/\hat{X}_1\} 100$$

$$\%Accuracy = 100 - \{X_1 - \hat{X}_1\}/X_1 \} 100$$

$$95\% \text{ C. I. for } X_1 : \hat{X}_1 \pm Z_{.05} SE(\hat{X}_1)$$

$$\text{Similarly for } \hat{X}_2, \hat{X}_4 \text{ \& } \hat{X}_5$$

(ii) **Estimate of total number of trees:**

$$\hat{Y}_1 = M\bar{y}_1, \quad \bar{y}_1 = (1/n\bar{M}) \sum_{i=1}^n M_i \bar{y}_{1i} \quad (\text{unbiased}),$$

$$\bar{y}_{1i} = 1/M_i \sum_{j=1}^{M_i} y_{1ij}, \quad \bar{M} = 1/N \sum_{i=1}^n M_i = M/N$$

$$SE(\hat{Y}_1) = M\sqrt{[(N-n)/N(1/n(n-1)\sum\{(M_i\bar{Y}_{1i})/\bar{M}-\bar{y}_1\}^2]}$$

$$\%SE = \{SE(\hat{Y}_1)/\hat{Y}_1\}100$$

$$\% \text{ Accuracy} = 100 - \{Y_1 - \hat{Y}_1\}/Y_1\}100$$

$$95\% \text{ C I for } Y_1 : \hat{Y}_1 \pm Z_{.05} SE(\hat{Y}_1)$$

$$\text{Similarly for } \hat{Y}_2, \hat{Y}_3, \hat{Y}_4 \text{ \& } \hat{Y}_5$$

(iii) (a) **Estimate of volume of compact trees per hectare:**

$$\hat{R}_1 = \sum_{i=1}^n y_{3i} / \sum_{i=1}^n x_{2i} \quad (\text{biased})$$

$$\begin{aligned} SE(\hat{R}_1) &= \sqrt{[(N-n)/Nn(n-1)\bar{X}_2^2] (\sum_{i=1}^n y_{3i}^2 + \hat{R}_1^2 \sum_{i=1}^n x_{2i}^2 - 2\hat{R}_1 \sum_{i=1}^n y_{3i} x_{2i})} \\ &= \sqrt{[(N-n)/\bar{N}n(n-1)\bar{X}_2^2] \sum_{i=1}^n (y_{3i} - \hat{R}_1 x_{2i})^2} \end{aligned}$$

$$\bar{x}_2 = \text{cluster mean of } x_2$$

$$\%SE = \{SE(\hat{R}_1)/\hat{R}_1\}100$$

$$\text{Bias } (\hat{R}_1) \leq \hat{V}(\bar{x}_2) \hat{V}(\hat{R}_1) / \bar{x}_2 = SE(\bar{x}_2) \cdot SE(\hat{R}_1) / \bar{x}_2$$

$$\bar{x}_2 = \text{unbiased estimate of } \bar{X}_2$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{R}_1)/\hat{R}_1\}100$$

$$95\% \text{ C I for } R_1 : \hat{R}_1 \pm Z_{.05} SE(\hat{R}_1)$$

iii (b) Estimate of number of bamboos per clump:

$$\hat{R}_2 = \sum_{i=1}^n y_{4i} / \sum_{i=1}^n x_{5i} \text{ (biased)}$$

$$\begin{aligned} SE(\hat{R}_2) &= \sqrt{\{(N-n)/Nn(n-1)\bar{X}_5^2\}(\sum_{i=1}^n y_{4i}^2 + \hat{R}_2^2 \sum_{i=1}^n x_{5i}^2 - 2\hat{R}_2 \sum_{i=1}^n (y_{4i}x_{5i}))} \\ &= \sqrt{\{(N-n)/Nn(n-1)\hat{X}_5^2\} \sum_{i=1}^n (y_{4i} - \hat{R}_2 x_{5i})^2} \end{aligned}$$

$$\bar{x}_5 = \text{cluster mean of } x_5$$

$$\% SE = \{SE(\hat{R}_2)/\hat{R}_2\} 100$$

$$\text{Bias}(\hat{R}_2) \leq \hat{V}(\bar{x}_5) \cdot \hat{V}(\hat{R}_2)/\bar{x}_5 = SE(\bar{x}_5) \cdot SE(\hat{R}_2)/\bar{x}_5,$$

$$\bar{x}_5 = \text{unbiased estimate of } \bar{X}_5$$

$$\% \text{Bias} = \{\text{Bias}(\hat{R}_2)/\hat{R}_2\} 100$$

$$95\% \text{ C I for } R_2 : \hat{R}_2 \pm Z_{0.05} SE(\hat{R}_2)$$

(iv) Estimate of total number of households with bamboo plantation:

$$\hat{NP} = M_p = M \sum_{i=1}^n a_i / \sum_{i=1}^n M_i \text{ Which is equivalent to Ratio Estimate.}$$

$$p_i = a_i/M_i = \text{Proportion of households with bamboo plantation in the } i\text{th cluster.}$$

$$\bar{M} = 1/N \sum_{i=1}^n M_i = M/N$$

$$\begin{aligned} SE(\hat{NP}) &= \sqrt{[N(N-n)/n(n-1)\bar{M}^2] (\sum_{i=1}^n a_i^2 + p^2 \sum_{i=1}^n M_i^2 - 2p \sum_{i=1}^n a_i M_i)} \\ &= \sqrt{N(N-n)/n(n-1)\bar{M}^2 \sum_{i=1}^n (a_i - pM_i)^2} \end{aligned}$$

$$\% SE = \{SE(\hat{NP})/\hat{NP}\} 100$$

$$\% \text{ Accuracy} = 100 - \{(\hat{NP} - NP)/NP\} 100$$

$$95\% \text{ C. I. for } NP : \hat{NP} \pm Z_{0.05} SE(\hat{NP})$$

Where,

N = Number of clusters in the population.

n = Number of clusters in the sample.

M_i = Number of households in i th cluster.

$\bar{M}$ = Average number of households per cluster in the population.

$\bar{x}_i$ = Average of x -variable for i th cluster.

$\bar{y}_i$ = Average of y -variable for i th cluster.

a_i = Number of households with bamboo plantation in i th sample cluster.

6.3 Example

A number of 6 villages with 251 households is formed into 6 clusters. A sample of 3 clusters with 94 households is selected at random. Data of same variables as in example 3.4 are collected (Appendix-4). Estimate the following parameters:

- (a) total volume (m^3) of compact trees
- (b) volume (m^3) of compact trees per hectare, and
- (c) total no. of households with bamboo plantations.

Solution:

No. of villages = N (Population Clusters) = 6

No. of villages = n (Sample Clusters) = 3

N = 6

Cl-1	Cl-2	Cl-3
(HH-43)	(HH-45)	(HH-64)
Cl-4	Cl-5	Cl-6
(HH-17)	(HH-48)	(HH-34)

n = 3

Cl-1	Cl-4	Cl-6
(HH-43)	(HH)17	HH-43
$M_1 = 43$	$M_2 = 17$	$M_3 = 34$
$\bar{x}_2 = 0.045$	0.038	0.04
$\sum x_2 x_2 = 0.159$	0.069	0.07
$\bar{y}_3 = 14.760$	8.08	9.93
$\sum y_3 = 16769.425$	2907.738	6566.62
$\sum x_2 y_3 = 47.012$	12.143	18.29
$\sum x_2 = 1.935$	0.646	1.36

$\sum y_3 = 634.68$	137.36	3737.62
$\hat{R}_1 = 328$	$\hat{R}_2 = 212.6$	$\hat{R}_3 = 248.3$
$\sum(y_3 - \hat{R}_1 x_2)^2 = 3035.409$	863.249	1799.508
$\text{Weighted } \sum(y_3 - \hat{R}_1 x_2)^2 = 2195.5435$		

(a) Estimate of total volume of compact trees:

$$\bar{y}_3 = 1/3 \times 42 \{ (43 \times 17.76) + (17 \times 8.08) + (34 \times 9.93) \}$$

$$= 8.766$$

y_3 = mean of y_3 for 3 sample clusters

$$(i) \hat{Y}_3 = M \times \bar{y}_3 = 251 \times 8.766 = 2200.266$$

$$(ii) SE(\hat{Y}_3) = 251 \sqrt{0.083 \times (40.26 + 30.13 + 0.52)}$$

$$= 608.92$$

$$\%SE(\hat{Y}_3) = 608.92 / 2200.266 \times 100$$

$$= 0.27 \times 100$$

$$= 27\%$$

$$(iii) \%Accuracy = 100 - 2609.98 - 2200.26 / 2609.98$$

$$= 100 - 15.79$$

$$= 85\%$$

(iv) 95% Confidence interval for Y_3

$$\text{Lower limit : } Y_3 - 1.96 \times 608.92$$

$$= 2200.266 - 1193.4832$$

$$= 1006.783$$

$$\text{Upper limit : } = 2200.266 + 1193.4832$$

$$= 3393.75$$

(b) Estimate of volume of compact tree per hectare:

$$\hat{R}_1 = \sum y_3 / \sum x_2 = 1109.66 / 3.94 = 281.5682$$

$$\begin{aligned} SE(\hat{R}) &= 1 / \sum x_2 \sqrt{(N-n)n / N(n-1) \sum (y_3 - \hat{R}x_2)^2} \\ &= 1 / 14.558 \sqrt{.75 \times 2195.543} = 2.7874 \end{aligned}$$

$$\%SE(\hat{R}_1) = .98\%$$

$$\text{Bias}(\hat{R}_1) < SE(\hat{R}), SE(\bar{x}_2) / \bar{x}_2 = .5575$$

$$\text{Bias}(\hat{R}_1) \% = 0.19\%$$

$$\bar{x}_2 = 1/3 \times 42 \{1.935 + 0.646 + 1.36\} = 0.031$$

$$SE(\bar{x}_2) = \sqrt{0.083 (.00022 + 0.00024 + .000019)} = 6.1753$$

$\bar{x}_2 = \text{mean of } x_2 \text{ for 3 sample clusters.}$

95% Confidence interval for R_1 :

Lower limit : 276.1046

Upper limit : 287.0313

Similarly R_2 maybe estimated

(c) Estimate of total no. of households with bamboo plantation:

$$\begin{aligned} \hat{NP} &= M \times p & p &= \sum a_i / \sum M_i \\ &= 251 \times 0.53 & (17 + 8 + 25) / 94 \\ &= 133 & = .53 \end{aligned}$$

$$(i) \quad SE(Mp) = 1/42 \sqrt{6(6-3)/(3 \times 2) \{978 + (.53)^2 \times 3294 - 2 \times .53 \times 1717\}} = .3763$$

$$\%SE(Mp) = .3763/133 \times 100 = .28\%$$

$$(ii) \quad \%Accuracy = 100 - (23 - 133)/123 \times 100 = 92\%$$

(iii) 95% Confidence interval for NP :

$$\text{Lower limit : } 133 - 1.96 \times .3763 = 132$$

$$\text{Upper limit : } 133 + 1.96 \times .3763 = 134$$

CHAPTER VII

TWO STAGE SAMPLING

7.1 Definition and discussion

Two Stage Sampling Scheme has been found to be very useful in practice and it is often used in large-scale surveys. It is the only choice in a number of practical situations where a frame of sampling units is not readily available and the cost of preparing such a frame is large. It also reduces the cost of operations of the survey and helps to minimize non-sampling errors. The selection at each stage may be done either with equal probability or with unequal probability and the method may be different at different stage depending upon the situations. This procedure may be used in woodlot survey where the first stage units are the villages and the second stage units are the households. The frame of the villages is available but the frame of the households is not readily available. Preparation of the list of plots, trees, bamboos, etc., of a household is expensive and is not necessary in Two Stage Sampling.

7.2 Estimation of parameters

(i) Estimate of total area of land:

$$\hat{X}_1 = M \bar{x}_1, \bar{x}_1 = 1/n\bar{M} \sum_{i=1}^n M_i \bar{x}_{1i} \text{ (unbiased), } \bar{x}_{1i} = 1/m_i \sum_{j=1}^n x_{1ij},$$

$$\bar{M} = 1/N \sum_{i=1}^N M_i = M/N$$

$$SE(\hat{X}_1) = M \sqrt{[(N-n)/Nn] s_b^2 + 1/nN\bar{M}^2 \sum_{i=1}^n \{M_i(M_i-m_i)/m_i\} s_{wi}^2}$$

$$s_b^2 = 1/(n-1) \sum_{i=1}^n \{(M_i/\bar{M})\bar{x}_{1i} - \bar{x}_1\}^2,$$

$$s_{wi}^2 = 1/(m_i-1) \sum_{j=1}^{m_i} (x_{1ij} - \bar{x}_{1i})^2$$

$$\%SE = \{SE(\hat{X}_1)/\hat{X}_1\} 100$$

$$\%Accuracy = 100 - \{X_1 - \hat{X}_1\}/X_1 \} 100$$

$$95\% \text{ C. I. for } X_1 : \hat{X}_1 \pm Z_{.05} SE(\hat{X}_1)$$

$$\text{Similarly for } \hat{X}_2, \hat{X}_4 \text{ \& } \hat{X}_5$$

(ii) **Estimate of total number of trees in the population:**

$$\hat{Y}_1 = M\bar{y}_1, \bar{y}_1 = 1/n\bar{M} \sum_{i=1}^n M_i \bar{y}_{1i} \text{ (unbiased), } \bar{y}_{1i} = 1/m_i \sum_{j=1}^{m_i} x_{1ij},$$

$$M = 1/N \sum_{i=1}^n M_i = M/N$$

$$SE(\hat{Y}_1) = M \sqrt{[(N-n)/Nn] s_b^2 + 1/nN\bar{M}^2 \sum_{i=1}^n \{M_i(M_i-m_i)/m_i\} s_{wi}^2}$$

$$s_b^2 = 1/(n-1) \sum_{i=1}^n \{(M_i/\bar{M})\bar{y}_{1i} - \bar{y}_1\}^2,$$

$$s_{wi}^2 = 1/(m_i-1) \sum_{j=1}^{m_i} (y_{1ij} - \bar{y}_{1i})^2$$

$$\%SE = \{SE(\hat{Y}_1)/\hat{Y}_1\} 100$$

$$\%Accuracy = 100 - \{Y_1 - \hat{Y}_1\}/Y_1 \} 100$$

$$95\% \text{ C. I. for } Y_1 : \hat{Y}_1 \pm Z_{.05} SE(\hat{Y}_1)$$

Similarly for $\hat{Y}_2, \hat{Y}_4$ & $\hat{Y}_5$

(iii) Estimate of volume of compact trees per hectare:

$$(a) \quad \hat{R}_1 = \bar{y}_3 / \bar{x}_2 \text{ (biased), } \bar{y}_3 = 1/n \bar{M} \sum_{i=1}^n M_i \bar{y}_{3i}, \bar{x}_2 = 1/n \bar{M} \sum_{i=1}^n M_i \bar{x}_{2i}$$

$$\bar{y}_{3i} = 1/m_i \sum_{j=1}^{m_i} y_{3ij}, \bar{x}_{2i} = 1/m_i \sum_{j=1}^{m_i} x_{2ij}$$

$$SE(\hat{R}_1) \doteq 1/\bar{x}_2 \sqrt{\{v(\bar{y}_3) + \hat{R}_1^2 v(\bar{x}_2) - 2\hat{R}_1 \text{COV}(\bar{y}_3, \bar{x}_2)\}}$$

$v(\bar{y}_3)$ and $v(\bar{x}_2)$ are given in (i) and (ii)

$$\text{COV}(\bar{y}_3, \bar{x}_2) = (1/n - 1/N) s_{bx_2y_3} + 1/Nn \bar{M}^2 \sum_{i=1}^n \{M_i(M_i - m_i)/m_i\} s_{ix_2y_3}$$

$$s_{bx_2y_3} = 1/(n-1) \sum_{i=1}^n \{(M_i/\bar{M})\bar{x}_{2i} - \bar{x}_2\} \{(M_i/\bar{M})\bar{y}_{3i} - \bar{y}_3\}$$

$$s_{ix_2y_3} = 1/(m_i-1) \sum_{j=1}^{m_i} (x_{2ij} - \bar{x}_{2i})(y_{3ij} - \bar{y}_{3i})$$

$$\%SE = \{SE(\hat{R}_1)/\hat{R}_1\} 100$$

$$\text{Bias}(\hat{R}_1) = 1/\bar{x}_2 \{v(\bar{x}_2)/\bar{x}_2 - \text{cov}(\bar{x}_2, \bar{y}_3)/\bar{y}_3, \bar{x}_2\}$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{R}_1)/\hat{R}_1\} 100$$

$$95\% \text{ C. I. for } R_1 : \hat{R}_1 \pm Z_{.05} SE(\hat{R}_1)$$

(iii) Estimate of number of bamboos per clump:

$$(b) \quad \hat{R}_2 = \bar{y}_4, \bar{x}_5 \text{ (biased), } \bar{y}_4 = 1/n \bar{M} \sum_{i=1}^n M_i \bar{y}_{4i}, \bar{x}_5 = 1/n \bar{M} \sum_{i=1}^n M_i \bar{x}_{5i}$$

$$\bar{y}_{4i} = 1/m_i \sum_{j=1}^{m_i} y_{4ij}, \bar{x}_{5i} = 1/m_i \sum_{j=1}^{m_i} x_{5ij}$$

$$SE(\hat{R}_2) \doteq 1/\bar{x}_5 \sqrt{\{v(\bar{y}_4) + \hat{R}_2^2 v(\bar{x}_5) - 2\hat{R}_2 \text{COV}(\bar{y}_4, \bar{x}_5)\}}$$

$v(\bar{y}_4)$ and $v(\bar{x}_5)$ are given in (i) and (ii)

$$\text{COV}(\bar{y}_4, \bar{x}_5) = (1/n - 1/N) s_{bx_5y_4} + 1/Nn \bar{M}^2 \sum_{i=1}^n \{M_i(M_i - m_i)/m_i\} s_{ix_5y_4}$$

$$s_{bx_5y_4} = 1/(n-1) \sum_{i=1}^n \{(M_i/\bar{M})\bar{x}_{5i} - \bar{x}_5\} \{(M_i/\bar{M})\bar{y}_{4i} - \bar{y}_4\}$$

$$s_{ix_5y_4} = 1/(m_i-1) \sum_{j=1}^{m_i} (x_{5ij} - \bar{x}_{5i})(y_{4ij} - \bar{y}_{4i})$$

$$\%SE = \{SE(\hat{R}_2)/\hat{R}_2\} 100$$

$$\text{Bias}(\hat{R}_2) = 1/\bar{x}_5 \{v(\bar{x}_5)/\bar{x}_5 - \text{cov}(\bar{x}_5, \bar{y}_4)/\bar{y}_4, \bar{x}_5\}$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{R}_2)/\hat{R}_2\} 100$$

$$95\% \text{ C. I. for } R_2 : \hat{R}_2 \pm Z_{.05} SE(\hat{R}_2)$$

(vi) Estimate of number of households with bamboo plantation:

$$\text{Proportion : } \hat{NP} = \hat{MP}, \hat{P} = \frac{\sum_{i=1}^n M_i p_i}{\sum_{i=1}^n M_i} \text{ (unbiased)}$$

$$SE(\hat{MP}) = M \sqrt{\{N-n\}/N(1/n\bar{M}^2)s^2 + 1/nN\bar{M}^2 \sum_{i=1}^n M_i(M_i - m_i) \hat{p}_i \hat{q}_i / (m_i - 1) \}}$$

$$\text{Where, } s^2 = 1/(n-1) \sum_{i=1}^n M_i^2 (\hat{p}_i - \hat{P})^2$$

$$SE = \{SE(\hat{MP})/\hat{MP}\} 100$$

$$\% \text{Accuracy} = 100 - \{(MP - \hat{MP})/\hat{MP}\} 100$$

$$95\% \text{ C. I. for MP : } \hat{MP} \pm Z_{.05} SE(\hat{MP})$$

Where,

N = Number of first stage units (villages) in the population.

n = Number of first stage units (villages) in the sample.

M_i = Number of second stage units (households) in i th first stage unit (village) in the population.

m_i = Number of second stage units (households) in the sample of i th first stage unit.

p_i = Proportion estimate of number of households with bamboo plantation in m_i .

7.3 Example

A number of 10 villages with 250 households is considered in an area of Chittagong district in Bangladesh. A number of 5 villages with 130 households are selected at random at first stage. A number of 67 households from 5 villages are selected at random at second stage. A structure of 2 stage sampling design is given below. Data of same variables as in example 3.4 are collected as given below. Estimate the following parameters:

- total volume (m³) of compact trees
- volume (m³) of compact trees per hectare, and
- total no. of households with bamboo plantations.

Population: Vill-1(20) Vill-2(10) Vill-3(25) Vill-4(18) Vill-5(29)

Vills(10) : Vill-6(33) Vill-7(25) Vill-8(23) Vill-9(38) Vill-10(29)
HH(250)

Stage 1 : Vill-06(33) Vill-08(23) Vill-07(25) Vill-10(29) Vill-01(20)
SRS: Vills(5)

Stage 2 : Vill-06(17) Vill-08(12) Vill-07(13) Vill-10(15) Vill-01(10)
SRS : HH(67)

Table 5. A sample of 67 households at second stage of two stage sampling

Sl. No.	Vill-06(17)			Vill-08(12)			Vill-07(13)			Vill-10(15)			Vill-01(10)		
	x ₂	y ₃	x ₅	x ₂	y ₃	x ₅	x ₂	y ₃	x ₅	x ₂	y ₃	x ₅	x ₂	y ₃	x ₅
1	.02	7.70	0	.06	5.01	1	.11	9.48	0	.02	7.31	0	.02	7.50	0
2	.06	11.04	1	.04	2.29	2	.01	4.99	0	.08	61.92	1	.05	6.72	0
3	.02	2.32	0	.05	5.00	1	.02	1.86	0	.02	4.49	0	.03	3.52	0
4	.04	4.15	4	.01	3.27	1	.16	24.19	0	.03	6.84	3	.02	7.98	1
5	.04	9.67	0	.04	2.29	2	.04	21.09	1	.02	6.14	0	.03	1.98	2
6	.04	10.35	0	.01	3.92	0	.02	6.10	0	.03	11.60	0	.03	9.73	3

Sl. No.	Vill-06(17)			Vill-08(12)			Vill-07(13)			Vill-10(15)			Vill-01(10)		
	x_2	y_3	x_5	x_2	y_3	x_5	x_2	y_3	x_5	x_2	y_3	x_5	x_2	y_3	x_5
7	.10	8.16	1	.04	2.29	0	.04	8.15	0	.03	7.70	1	.04	15.41	4
8	.04	3.28	2	.02	0.76	0	.02	3.91	0	.04	5.19	1	.16	52.10	12
9	.01	8.12	0	.02	7.76	1	.06	2.57	4	.06	3.74	1	.02	3.08	5
10	.02	12.26	13	.01	1.60	1	.02	8.66	0	.02	6.19	2	.02	7.98	1
11	.02	3.16	0	.05	5.00	1	.02	10.10	1	.04	9.09	0			
12	.04	4.90	0	.01	3.77	0	.06	15.37	1	.06	20.57	1			
13	.02	11.56	2				.03	6.14	1	.04	11.98	1			
14	.08	52.53	0							.19	21.27	1			
15	.04	12.06	0							.04	4.93	0			
16	.14	18.35	0												
17	.04	29.16	0												

$$M_6 = 33 \quad M_8 = 23 \quad M_7 = 25, \quad M_{10} = 29 \quad M_1 = 20$$

$$m_6 = 17 \quad m_8 = 12 \quad m_7 = 13 \quad m_{10} = 15 \quad m_1 = 10$$

$$M = 250 \quad \bar{M} = 25, \quad M_1 = 33 \quad M_2 = 23 \quad M_3 = 25 \quad M_4 = 29 \quad M_5 = 20$$

$$N = 10 \quad m_1 = 17 \quad m_2 = 12 \quad m_3 = 13 \quad m_4 = 15 \quad m_5 = 10$$

$$n = 5$$

Vill-6

Vill-8

Vill-7

$$\bar{x}_2 = 0.05, \bar{y}_3 = 12.28$$

$$\bar{x}_2 = 0.03, \bar{y}_3 = 3.58$$

$$\bar{x}_2 = 0.047, \bar{y}_3 = 9.43$$

$$S^2 x_2 = S^2 w = 0.00113$$

$$S^2 x_2 = S^2 w = 0.000345$$

$$S^2 x_2 = S^2 w = 0.00187$$

$$S^2 y_3 = 149.605$$

$$S^2 y_3 = 3.697$$

$$S^2 y_3 = 47.35$$

$$s_1 x_2 y_3 = 0.098$$

$$s_1 x_2 y_3 = 1.3714$$

$$s_1 x_2 y_3 = 8.006$$

$$\sum x_2 y_3 = 12.01$$

$$\sum x_2 y_3 = 0.0075$$

$$\sum x_2 y_3 = 0.187$$

Vill-10

$$\bar{x}_2=0.048, \bar{y}_3=12.60$$

$$S^2_{x_2} = S^2_w=0.00185$$

$$S^2_{y_3}=215.061$$

$$s_1x_2y_3=0.278$$

$$\sum x_2y_3=12.97$$

Vill-01

$$\bar{x}_2=0.042, \bar{y}_3=11.6$$

$$S^2_{x_2} = S^2_w=0.0081$$

$$S^2_{y_3}=217.28$$

$$s_1x_2y_3=0.6004$$

$$\sum x_2y_3=10.276$$

(a) Total volume (m³) of compact trees :

$$\begin{aligned}\hat{Y}_3 &= M. \bar{y}_3 \\ &= 250 \times 10.57 \\ &= 2642.5\end{aligned}$$

$$\begin{aligned}\text{Here, } \bar{y}_3 &= 1/nM \sum M_i. \bar{y}_{3i} \\ &= 1/125 [33 \times 12.28 + 23 \times 3.58 + 25 \times 9.43 + 29 \times 12.6 + 20 \times 11.6] \\ &= 10.57\end{aligned}$$

$$\begin{aligned}S^2_b &= 1/n-1 \sum (M_i/m \bar{y}_{3i} - \bar{y}_3)^2 \\ &= 1/5-1 [(33/25 \times 12.28 - 10.57)^2 + 23/25 \times 3.58 - 10.57)^2 + (25/25 \times 9.43 \\ &\quad - 10.57)^2 + (29/25 \times 12.6 - 10.57)^2 + (20/25 \times 11.6 - 10.57)^2] \\ &= 1/4 [31.81 + 52.95 + 1.30 + 16.37 + 1.66] \\ &= 26.023\end{aligned}$$

$$S.E(\hat{Y}_3) = M \sqrt{[N-n/Nn.s^2_b + 1/nNM^2 \sum \{ M_i(M_i-m_i)/m_i \} S^2_{x_{2wi}}]} \quad (\text{Talukdar, 1987})$$

$$\begin{aligned}
 \text{Here, } \sum \{M_i (M_i - m_i) / m_i\} S^2_{x_{2wi}} \\
 &= 33(33-17)/17 \times 149.605 + 23(23-12)/12 \times 3.697 + 25(25-13)/13 \times 47.35 \\
 &\quad + 29(29-15)/15 \times 215.061 + 20(20-10)/10 \times 217.28 \\
 &= 4646.56 + 77.95 + 1092.69 + 5820.98 + 4345.6 \\
 &= 15983.78
 \end{aligned}$$

$$\begin{aligned}
 SE(\hat{Y}_3) &= 250 \times \sqrt{[5/50 \times 26.023 + 1/31250 \times 15983.78]} \\
 &= 250 \times \sqrt{3.114} \\
 &= 441.16
 \end{aligned}$$

$$\& V(\bar{y}_3) = 3.114$$

$$\begin{aligned}
 \%SE &= SE(\hat{Y}_3) / \hat{Y}_3 \times 100 \\
 &= 441.16 / 2642.5 \times 100 \\
 &= 16.695\%
 \end{aligned}$$

$$\begin{aligned}
 \%Accuracy &= 100 \{ (Y_3 - \hat{Y}_3) / Y_3 \} \times 100 \\
 &= 100 - \{ 2141.04 - 2642.5 / 2141.04 \times 100 \} \\
 &= 100 - (+23.42) \\
 &= 77\%
 \end{aligned}$$

95% Confidence Interval for Y_3 :

$$\begin{aligned}
 &\hat{Y}_3 \pm Z_{0.05} SE(\hat{Y}_3) \\
 &= 2642.5 \pm 1.96 \times 441.16 \\
 &= 2642.5 \pm 864.67
 \end{aligned}$$

Lower limit : 1777.83

Upper limit : 3507.17

(b) Volume (m³) of compact trees per hectare

$$\hat{R}_1 = \bar{y}_3 / \bar{x}_2 = 10.57 / 0.046 = 229.78$$

$$\begin{aligned} \# \text{Sb} \bar{x}_2 \bar{y}_3 &= 1/n - 1 \sum \{ M_i / M \bar{x}_{2i} - \bar{x}_2 \} \{ M_i / M \bar{y}_{3i} - \bar{y}_3 \} \\ &= 1/4 [(23/25 \times 0.05 - 0.046) (33/25 \times 12.28 - 10.57) \\ &\quad + 23/25 \times 0.03 - 0.046) (23/25 \times 3.58 - 10.57) \\ &\quad + (25/25 \times 0.047 - 0.046) (25/25 \times 9.43 - 10.57) \\ &\quad + (29/25 \times 0.048 - 0.046) (29/25 \times 12.6 - 10.57) \\ &\quad + (20/25 \times 0.042 - 0.046) (20/25 \times 11.6 - 10.57) \\ &= 1/4 [(0.02 \times 5.46) + (0.0184 \times -7.2764) \\ &\quad + (0.001 \times -1.14) + (0.00968 \times 4.046) + (-0.0124 \times -1.29)] \\ &= 1/4 [0.1128 + 0.1339 - 0.00114 + 0.03392 + 0.016] \\ &= 0.07519 \end{aligned}$$

$$\text{Now, Six}_2 y_3 = 1/(m_i - 1) (\sum x_{2i} y_{3i} - m_i \bar{x}_{2i} \bar{y}_{3i})$$

$$\begin{aligned} \text{So, Cov}(y_3, x_2) &= (1/n - 1/N) \text{Sb} x_2 y_3 + 1/N n \bar{M}^2 \sum m_i (m_i - m_i) / m_i \} \text{Six}_2 y_3 \\ &= (1/5 - 1/10) \times 0.07519 + 1/31250 [33 \times 16/17 \times 0.098 \\ &\quad + 23 \times 11/12 \times 0.0075 + 25 \times 12/13 \times 0.185 + 29 \times 14/15 \\ &\quad \times 0.278 + 20 \times 10/10 \times 0.6004] \\ &= 0.007519 + 1/31250 [3.044 + 0.158 + 4.315 + 7.525 + 12.008] \\ &= 0.00838 \end{aligned}$$

$$\text{So, SE}(\hat{R}_1) = 20.304$$

$$\% \text{SE}(\hat{R}_1) = 8.84\%$$

$$\text{Bias}(\hat{R}_1) = -.3603$$

$$\% \text{Bias}(\hat{R}_1) = 0.157\%$$

95% Confidence Interval for R_1 :

Lower limit : 189.984

Upper limit : 269.576

(c) **Estimate of total number of households with bamboo plantations:**

$$\begin{aligned}\hat{NP} &= \hat{MP} \\ &= 250 \times 0.523 \\ &= 130.75\end{aligned}$$

$$\text{Here, } \hat{P} = \sum M_i \hat{P}_i / \sum M_i = 0.523$$

$$\begin{aligned}SE(\hat{MP}) &= M \sqrt{\{(N-n)/(NnM)^2 \times s^2 + 1/(NnM)^2 \sum m_i (M_i - m_i)/(m_i - 1)\}} \\ &= 15.5125 \text{ where } s^2 = 1/(n-1) \sum M_i^2 (\hat{P}_i - \hat{P})^2 \\ &= 17.975\end{aligned}$$

$$\begin{aligned}\%SE(MP) &= SE(\hat{MP})/\hat{MP} \times 100 \\ &= 15.5125/130.75 \times 100 \\ &= 11.864\%\end{aligned}$$

$$\begin{aligned}\%Accuracy &= 100\{MP - \hat{MP}/MP \times 100 \\ &= 100 - \{116.94 - 130.75/116.94 \times 100 \\ &= 100 - 11.809 = 88\%\end{aligned}$$

95% Confidence Interval for MP :

$$\begin{aligned}\hat{MP} \pm Z_{0.05} SE(\hat{MP}) \\ &= 130.75 \pm 1.96 \times 15.5125 \\ &= 130.75 \pm 30.4045 \\ \text{Lower limit : } &100.3455 \\ \text{Upper limit : } &161.1545\end{aligned}$$

CHAPTER VIII

STRATIFIED TWO STAGE SAMPLING (Self-Weighted)

8.1 Definition and discussion

The most common sampling design used in large-scale surveys is Stratified Multi Stage Sampling. Stratified Two Stage Sampling is the Two Stage Sampling when applied for stratified heterogeneous population. The village woodlot in the country is heterogeneous. Many woodlot surveys either partly or wholly are conducted previously in the country applying Stratified Two Stage Sampling procedure. This sampling scheme has the advantage of concentrating the sample within limited areas (1st stage units) rather than spreading it over the entire area to be surveyed. This reduces considerably the cost of operations of the survey and helps to reduce the non-sampling errors. It saves time and money in preparing list of elements in second stage units. For example, in woodlot survey the list of households is not readily available and it is expensive and time consuming to prepare the list of the household. But the list of villages, Unions, Thanas and Zilas are readily available.

8.2 Estimation of parameters

The following estimation formulas are used :

i) Estimation of total area of land:

$$\hat{X}_1 = N\bar{x}_1, \quad \bar{x}_1 = r = \sum_{h=1}^L x_{1h} / \sum_{h=1}^L m_h$$

$$SE(r) = \sqrt{\left\{ (1-f)/M^2 \left\{ \sum_{h=1}^L V(x_{1h}) + r^2 \sum_{h=1}^L v(m_h) - 2r \sum_{h=1}^L \text{cov}(x_{1h}, m_h) \right\} \right\}}$$

$$\% SE = (SE(\hat{X}_1) / \hat{X}_1) \times 100$$

$$\text{Bias}(r) = SE(\bar{x}_1) \cdot SE(\bar{m}) / \bar{M}, \quad SE(\bar{m}) = \sqrt{\sum_{h=1}^L w_h^2 v(m_h)}$$

$$95\% \text{ C. I. for } X_1 : \hat{X}_1 \pm Z_{0.5} SE(\hat{X}_1)$$

Similarly for $\hat{X}_2$, $\hat{X}_4$ & $\hat{X}_5$.

ii) Estimation of total number of trees in the population:

$$\hat{Y}_1 = N \bar{y}_1, \bar{y}_1 = r = \frac{\sum_{h=1}^L y_{1h}}{\sum_{h=1}^L m_h}$$

$$SE(r) = \sqrt{(1-f)/M^2 \{ \sum_{h=1}^L V(y_{1h}) + r^2 \sum_{h=1}^L v(m_h) - 2r \sum_{h=1}^L \text{cov}(y_{1h}, m_h) \}}$$

$$\% SE = (SE(\hat{Y}_1)/\hat{Y}_1) 100$$

$$\text{Bias}(r) = SE(\bar{y}_1).SE(\bar{m})/\bar{M}, \quad SE(\bar{m}) = \sqrt{\sum_{h=1}^L w_h^2 v(m_h)}$$

$$95\% \text{ C. I. for } Y_1 : \hat{Y}_1 \pm Z_{.05} SE(\hat{Y}_1)$$

Similarly for $\hat{Y}_2, \hat{Y}_3, \hat{Y}_4$ & $\hat{Y}_5$.

iii) Estimation of volume yield per hectar for compat trees:

$$(a) \quad \hat{R}_1 = y_3/x_2 = \frac{\sum_{h=1}^L y_{3h}}{\sum_{h=1}^L m_{2h}}$$

$$SE(\hat{R}_1) = \sqrt{[(1-f)/X_2^2] \{ \sum_{h=1}^L V(y_{3h}) + \hat{R}_1^2 \sum_{h=1}^L v(x_{2h}) - 2\hat{R}_1 \sum_{h=1}^L \text{cov}(y_{3h}, x_{2h}) \}}$$

$$\% SE = \{SE(\hat{R}_1)/\hat{R}_1\} 100$$

$$\text{Bias}(\hat{R}_1) = SE(\hat{R}_1).SE(\bar{x}_2)/\bar{x}_2$$

$$\% \text{ Bias} = \{ \text{Bias}(\hat{R}_1)/\hat{R}_1 \} 100$$

$$95\% \text{ C. I. for } R_1 : \hat{R}_1 \pm Z_{.05} SE(\hat{R}_1)$$

(b) Estimation of number of bamboos per clump:

$$\hat{R}_2 = y_4 / x_5 = \sum_{h=1}^L y_{4h} / \sum_{h=1}^L x_{5h}$$

$$SE(\hat{R}_2) = \sqrt{[(1-f)/X_5^2] \{ \sum_{h=1}^L V(y_{4h}) + \hat{R}_2^2 \sum_{h=1}^L V(x_{5h}) - 2\hat{R}_2 \sum_{h=1}^L \text{cov}(y_{4h}, x_{5h}) \}}$$

$$\% SE = \{SE(\hat{R}_2)/\hat{R}_2\} 100$$

$$\text{Bias}(\hat{R}_2) = SE(\hat{R}_2) \cdot SE(\bar{x}_5) / \bar{x}_5$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{R}_2)/\hat{R}_2\} 100$$

$$95\% \text{ C. I. for } \hat{R}_2 : \hat{R}_2 \pm Z_{.05} SE(\hat{R}_2)$$

iv) Estimation of total number of households with bamboo plantation:

$$\hat{NP} = \hat{NP}, \hat{P} = p = 1/N \sum_h N_h \hat{p}_h, \hat{P}_h = \sum_i^{nh} M_{hi} \hat{p}_{hi} / \sum_i^{nh} M_{hi}, P_{hi} = m_{hi}^l / m_{hi},$$

$$q_{hi} = 1 - p_{hi}$$

$$SE(\hat{R}) = \sqrt{1/N^2 \sum_h N_h^2 v(\hat{p}_h)}.$$

$$v(\hat{p}_h) = (N_h - n_h) / N_h n_h \bar{M}^2 s_h^2 + 1/n_h N_h \bar{M}^2 \sum_i^{nh} M_{hi} (M_{hi} - m_{hi}) (P_{hi} q_{hi}) / (m_{hi} - 1)$$

$$s_h^2 = 1/(n_h - 1) [\sum_i^{nh} (M_{hi} \hat{p}_{hi})^2 - 2\hat{p}_h \sum_i^{nh} M_{hi}^2 \hat{p}_{hi} + \hat{p}_h^2 \sum_i^{nh} M_{hi}^2]$$

$$\%SE = \{SE(\hat{P})/\hat{P}\} 100$$

$$\% \text{Accuracy} = 100 - \{(\hat{NP} - NP)/NP\} 100$$

$$95\% \text{ C. I. for } P : \hat{P} \pm Z_{.05} SE(\hat{P})$$

Where,

$\sum_{h=1}^L x_{1h} / f$ = Estimation of population total for x_1 .

$\sum_{h=1}^L n_h / f$ = Estimation of population size.

X_{1h} = Sample total of x_1 for hth stratum.

m_h = Sample total of households for hth stratum.

$\sum_{h=1}^L y_{1h}$ = Estimation of population total for y_1 .

y_{1h} = Sample total of y_1 for hth stratum.

$\sum_{h=1}^L x_{2h} / f$ = Estimation of population total for x_2 .

X_{2h} = Sample total of x_2 for hth stratum.

Similarly for X_{4h} , X_{5h} , y_{2h} , y_{3h} , y_{4h} and y_{5h} .

m_{hi}^1 = Number of households with bamboo plantation for ith unit (village) in hth stratum in the sample.

f = Overall constant sampling fraction in any stratum = $f_a \cdot f_b$,

f_a = n_h / N_h , f_b = m_h / M_h .

$V(x_{1h}) = 1/(a_h - 1) \{ a_h \sum_{i=1}^{a_h} x_{1hi}^2 - x_{1h}^2 \}$.

$V(m_h) = 1/(a_h - 1) \{ a_h \sum_{i=1}^{a_h} m_{hi}^2 - m_h^2 \}$

$COV(x_{1h}, m_h) = 1/(a_h - 1) \{ a_h \sum_{i=1}^{a_h} x_{1hi} m_{hi} - x_{1h} m_h \}$.

x_{1hi} = $\sum x_{1hij}$, $n_{hi} = \sum n_{hij}$, (ith FSU totals in hth stratum).

a_{hi} = Number of selected villages in hth stratum).

Similarly for $v(X_{2h})$, $v(X_{4h})$, $v(X_{5h})$, $v(y_{1h})$, $v(y_{2h})$, $v(y_{3h})$, $v(y_{4h})$ and $v(y_{5h})$.

8.3 Example

A total of 9 villages with 404 households is stratified into 3 strata in Chittagong district of Bangladesh. Sampling at stage-1 and stage-2 are shown below. Data of same variables as in example 3.4 are collected (Appendix – 5). Estimate the following parameters:

- total volume (m3) of compact trees,
- volume (m3) of compact trees per hectare, and
- total no. of households with bamboo plantations.

Strata :	I	II	III
Vills (9) :	1(36)	1(48)	1(42)
HH(404)	2(43)	2(39)	2(55)
	3(64)	3(30)	3(47)
Stage I :	1(36)	1(48)	1(42)
SRS of Vills(6):	2(43)	2(39)	2(55)
HH(263)			
Stage 2 :	1(14)	1(19)	1(17)
SRS of III(105) :	2(17)	2(16)	2(22)

$$\begin{aligned}
 \sum x_2 &= .88 + \sum x_2 = 1.06 & \sum x_2 &= 1 + 1.68 & \sum x_2 &= .61 + 1.27 \\
 \sum x_2^2 &= .103 + \sum x_2^2 = .1038 & \sum x_2^2 &= .0876 + .3804 & \sum x_2^2 &= .0291 + .1783 \\
 \sum y_3 &= 67.12 + \sum y_3 = 302.27 & \sum y_3 &= 144.59 + 183.36 & \sum y_3 &= 136.99 + 211.66 \\
 \sum y_3^2 &= 755.749 + \sum y_3^2 = 8975.781 & \sum y_3^2 &= 1295.178 + 3208.813 & \sum y_3^2 &= 152.838 + 3222.66 \\
 \sum x_2 y_3 &= 6.3315 + \sum x_2 y_3 = 28.5059 & \sum x_2 y_3 &= 9.4013 + 24.2123 & \sum x_2 y_3 &= 5.4782 + 20.558
 \end{aligned}$$

Table 6. Self-weighting nature

(Obaidullah, 1980)

Strata	N_h	n_h	$f_a m_h / M_h$	M_h	m_h	$fb = m_h / M_h$	$fa \cdot fb$
I	3	2	.667	79	31	.3924	.2616
II	3	2	.667	87	35	.4023	.2682
III	3	2	.607	97	39	.4021	.2681

(a) Estimate of total volume (m3) of compact trees:

$$\hat{Y}_3 = M\bar{y}_3, \quad \bar{y}_3 = \Sigma \bar{y}_3 h / \Sigma m h = r = 9.96$$

$$\therefore \hat{Y}_3 = 404 \times 9.96 = 4023.84$$

$$\Sigma v(y_3 h) = 62374.27$$

$$\Sigma v(mh) = 43$$

$$\Sigma \text{cov}(y_3 h, mh) = 962.49$$

$$\text{SE}(r) = M \text{SE}(r) = 187.42$$

$$\% \text{SE}(Mr) = 4.6\%$$

Similarly Bias@% may be calculated as in (a).

95% Confidence Interval for $\hat{Y}_3$:

$$\hat{Y}_3 \pm 1.96 \times \text{SE}(\hat{Y}_3)$$

$$\therefore 4023 \pm 1.96 \times 187.42$$

Lower limit : 3656.49

Upper limit : 4391.18

(b) Estimate of volume (m3) of compact trees per hectare:

$$\hat{R}_1 = \Sigma y_3 h / \Sigma x_3 h = 1045.99 / 6.5 = 160.92$$

$$\Sigma v(y_3 h) = 62374.27$$

$$\Sigma v(x_2 h) = .9304$$

$$\Sigma \text{cov}(y_3 h, x_2 h) = 117.9729$$

$$\text{SE}(\hat{R}_1) = 8.53$$

$$\% \text{SE}(\hat{R}_1) = 8.52/160.92 \times 100 = 5\%$$

(c) **Estimate of total number of households with bamboo plantations:**

$$P_{hi} = m'_{hi}/m_{hi}$$

St-I	St-II	St-III
$P_{h1} = 9/14$	$P_{h2} = 14/19$	$P_{h1} = 4/17$
$P_{h2} = 7/17$	$P_{h2} = 13/16$	$P_{h2} = 7/22$
$\Sigma M_{hp_h} = 40.85$	$\Sigma M_{hp_h} = 67.06$	$\Sigma M_{hp_h} = 27.38$
$\Sigma M_h = 79$	$\Sigma M_h = 87$	$\Sigma M_h = 97$
$\hat{P}_{h1} = .52$	$\hat{P}_{h2} = .77$	$\hat{P}_{h3} = .28$

$$\hat{P} = p = 1/N \Sigma N_h \hat{P}_h = .52$$

$$\hat{NP} = Mp = 404 \times .52 = 210$$

$$s^2_1 = 19.996$$

$$s^2_2 = .1003$$

$$s^2_3 = .0032$$

$$\begin{aligned} \text{Stratum I} &: (N_1 - n_1)/N_1 n_1 \bar{M}^2 s^2_1 \\ &= (3-2)/3 \times 2 \times 1/45^2 \times 19.996 = .0016 \end{aligned}$$

$$\text{Stratum II} : (N_2 - n_2)/N_2 n_2 \bar{M}^2 s^2_2 = .00083$$

$$\text{Stratum III} : (N_3 - n_3)/N_3 n_3 \bar{M}^2 s^2_3 = .00000026$$

$$1/N_h n_h \bar{M}^2 \Sigma M_{hi} (M_{hi} - m_{hi}) (P_{hi} Q_{hi}) / (m_h - 1)$$

$$1/12150 \{ (15.2064 + 17.4408) + (13.6957 + 25.3253) + (13.23 + 17.424) \}$$

$$\text{Stratum I} : (15.2064 + 17.408)/12150 = .0027$$

$$\text{Stratum II} : (13.6957 + 25.325)/12150 = .0032$$

$$\text{Stratum III} : (13.23 + 17.4241)/12150 = .0025$$

$$V(\hat{P}_1) = .016 + .0027 = .0043$$

$$V(\hat{P}_2) = .00083 + .0032 = .0040$$

$$V(\hat{P}_3) = .00000026 + .0025 = .0025$$

$$SE(Np) = \sqrt{\sum N h^2 V(\hat{P}_h)} = .3118$$

$$\% SE(Np) = (.3118/210) \times 100 = .15\%$$

$$\% Accuracy = 86\%$$

95% Confidence Interval for MP :

$$\text{Lower limit : } 210 - 1.96 \times .3118 = 209$$

$$\text{Upper limit : } 210 + 1.96 \times .3118 = 211$$

CHAPTER IX

DOUBLE SAMPLING

9.1 Definition and Discussion

When the population mean $\bar{X}$ is not known, the ratio and the regression methods of estimation cannot be used to estimate the population mean of Y. In this situation Double Sampling Method can be applied to estimate the parameters. Double Sampling can play an important role in surveys of village wood and bamboo resources. It is useful in situation where estimation of volume of wood or number of bamboos (main character) involves much cost or labour whereas an auxiliary character (area of compact tree garden or area of bamboo plantation) correlated to the main character can be easily measured. Thus, it may be convenient and economical to take a larger sample for the auxiliary variable at the first phase for obtaining precise estimates. In the second phase, a smaller sub sample for both the auxiliary and the main characters, can be easily selected and a precise estimate of the main character can be obtained. The information of the auxiliary variate x is required many times, either for estimation or for use in stratification in homestead woodlot survey in order to increase the efficiency of the estimator. Double Sampling reduces the cost of enumeration and increase the accuracy of the sample in village woodlot survey. It is advantageously used in resurveys of villages woodlot. It is used for predicting future growth by measuring the past growth in a sub-sample of trees and bamboos and use it as a guide in estimating the future growth. It has another use in stratification of homestead woodlot population. At first phase sample for an auxiliary character may be used to sub-divide the population under the study into strata which is absolutely necessary for heterogeneous homestead tree garden. It, with its first phase, helps to stratify the population and in the second phase the main character is precisely estimated with the help of both y and x variables.

At first phase, a sample of n^1 households is randomly selected for measuring the three auxiliary variables (area under compact tree plantation, number of bamboo clumps and number of trees). At the second

phase, a sample of n households is chosen at random for observing volume of compact trees, number of bamboos and volume of trees (the main characters y 's). Data on x_2, x_5 and y_1 at first phase and y_2, y_3, y_4 and x_2, x_5, y_1 at second phase are to be compiled and the parameters are to be estimated. The following formulas are applied for estimating the parameters of interest.

9.2 Estimation of parameters

i) Estimate of volume per tree:

$$(a) \bar{\bar{y}}_{2dR} = (\bar{y}_2 / \bar{y}_1) \bar{y}_1^1 \quad (\text{biased})$$

$$SE(\bar{y}_{2dR}) \doteq \sqrt{\{(n^1 - n)/(n^1 n) s^2_{y_2 \cdot y_1} + (N - n^1)/(N n^1) s^2_{y_2}\}}$$

$$\% SE = \{SE(\bar{y}_{2dR}) / \bar{y}_{2dR}\} 100$$

$$Bias(\bar{y}_{2dR}) = \bar{y}_{2dR} (1/n - 1/n^1)(c^2_{y_1} - r c_{y_1} c_{y_2}), c_{y_1} = s_{y_1} / \bar{y}_1$$

$$c_{y_2} = s_{y_2} / \bar{y}_2 \quad \text{and} \quad r = s_{y_1 y_2} / s_{y_1} s_{y_2}$$

$$Bias(\hat{\bar{Y}}_{2dR}) = Y^1_1 Bias(\hat{R}_{1d}), \hat{R}_{1d} = \bar{y}_2 / \bar{y}_1$$

$$\% Bias = \{Bias(\hat{\bar{Y}}_{2dR}) / \hat{\bar{Y}}_{2dR}\} 100$$

$$95 \% \text{ C.I. for } \bar{Y}_{2dR} : \bar{y}_{2dR} \pm z_{.05} SE(\bar{y}_{2dR})$$

$$b) \quad \bar{y}_{2d1} = \bar{y}_2 + \hat{B}(\bar{y}_1 - \bar{y}_1)(\text{biased}), \quad \hat{B} = s_{y_2 y_1} / s^2_{y_1}$$

$$SE(\bar{y}_{2d1}) \doteq \sqrt{[(n^1 - n^1/n) s^2_{y_2}(1 - r^2) + (N - n^1/Nn^1) s^2_{y_2}}$$

$$\text{and } r = s_{y_1 y_2} / s_{y_1} \cdot s_{y_2}$$

$$\% SE = \{SE(\bar{y}_{2d1}) / \bar{y}_{2d1}\} 100$$

$$\text{Bias}(\bar{y}_{2d1}) \doteq \hat{B}(1/n - 1/n^1)(\hat{M}_{21}/s_{y_2 y_1} - \hat{M}_{30}/s^2_{y_1}).$$

$$\hat{M}_{21} \doteq 1/(n-1) \sum_{i=1}^n (y_{i1} - \bar{y}_1)^2 (y_{i2} - \bar{y}_2) \text{ and } \hat{M}_{30} \doteq 1/(n-1) \sum_{i=1}^n (y_{i1} - \bar{y}_1)^3$$

$$\text{Bias}(\hat{Y}_{2d1}) = N \text{Bias}(\bar{y}_{2d1})$$

$$\% \text{ Bias} = \{\text{Bias}(\hat{Y}_{2d1}) / \hat{Y}_{2d1}\} 100$$

$$95 \% \text{ C.I. for } \bar{Y}_{2d1} : \bar{y}_{2d1} \pm z_{.05} SE(\bar{y}_{2d1})$$

ii) **Estimate of volume per hectare of compact trees:**

$$a) \bar{y}_{3dR} = (\bar{y}_3 / \bar{x}_2) \bar{x}_2^1 \text{ (biased)}$$

$$SE(\bar{y}_{3dR}) = \sqrt{\{(n^1 - n)/(n^1 n) s^2_{y_3 \cdot x_2} + (N - n^1)/(N n^1) s^2_{y_3}\}}$$

$$\% SE = \{SE(\bar{y}_{3dR}) / \bar{y}_{3dR}\} 100$$

$$Bias(\bar{y}_{3dR}) = \bar{y}_{3dR} (1/n - 1/n^1)(c^2_{x_2} - r c_{x_2} c_{y_3}), c_{x_2} = s_{x_2} / \bar{x}_2, c_{y_3} = s_{y_3} / \bar{y}_3$$

$$\text{and } r = s_{x_2 y_3} / s_{x_2} \cdot s_{y_3}$$

$$Bias(\hat{Y}_{3dR}) = X_2^1 Bias(\hat{R}_{1d}), \hat{R}_{1d} = \bar{y}_3 / \bar{x}_2$$

$$\% Bias = \{Bias(\hat{Y}_{3dR}) / \hat{Y}_{3dR}\} 100$$

$$95\% \text{ C. I. for } \bar{Y}_{3dR} : \bar{y}_{3dR} \pm z_{.05} SE(\bar{y}_{3dR})$$

$$b) \bar{y}_{3d1} = \bar{y}_3 + \hat{B}(\bar{x}_2^1 - \bar{x}_2) \text{ (biased)} \quad \hat{B} = s_{x_2 y_3} / s^2_{x_2}$$

$$SE(\bar{y}_{3d1}) = \sqrt{[(n^1 - n)/(n^1 n) s^2_{y_3} (1 - r^2) + (N - n^1)/(N n^1) s^2_{y_3}]}$$

$$r = s_{y_3 x_2} / s_{y_3} s_{x_2}$$

$$\% SE = \{SE(\bar{y}_{3d1}) / \bar{y}_{3d1}\} 100$$

$$Bias(\bar{y}_{3d1}) = \hat{B} (1/n - 1/n^1) (\hat{M}_{21} / s_{x_2 y_3} - \hat{M}_{30} / s^2_{x_2}),$$

$$\hat{M}_{21} = 1/(n-1) S(x_2 - \bar{x}_2)^2 (y_3 - \bar{y}_3) \text{ and } \hat{M}_{30} = 1/(n-1) \sum_{i=1}^n (x_2 - \bar{x}_2)^3$$

$$Bias(\hat{Y}_{3d1}) = N Bias(\bar{y}_{3d1})$$

$$\% Bias = \{Bias(\hat{Y}_{3d1}) / \hat{Y}_{3d1}\} 100$$

$$95\% \text{ C.I. for } \bar{Y}_{3d1} : \bar{y}_{3d1} \pm z_{.05} SE(\bar{y}_{3d1})$$

iii) Estimate of number of bamboos per clump:

$$a) \bar{y}_{4dr} \doteq (\bar{y}_4/\bar{x}_5) \bar{x}_5^1 \text{ (biased)}$$

$$SE(\bar{y}_{4dr}) \doteq \sqrt{(n^1-n)/(n^1 n) s^2 y_4 \cdot x_5 + (N-n^1)/(Nn^1) s^2 y_4}$$

$$\% SE = \{SE(\bar{y}_{4dr})/\bar{y}_{4dr}\} 100\}$$

$$Bias(\bar{y}_{4dr}) \doteq \bar{y}_{4dr} (1/n-1/n^1) (c^2 x_5 - rc x_5 c_{y4}), cx_5 = s_{x_5}/\bar{x}_5,$$

$$c_{y4} = s_{y4}/\bar{y}_4 \text{ and } r = s_{x_5 y_4} / s_{x_5} s_{y_4}$$

$$Bias(\hat{Y}_{4dr}) = X_5^1 Bias(\hat{R}_{2d}), \hat{R}_{2d} = \bar{y}_4/\bar{x}_5$$

$$\% Bias = \{Bias(\hat{Y}_{4dr})/\hat{Y}_{4dr}\} 100$$

$$95\% \text{ C.I. for } \bar{y}_{4dr} : \bar{y}_{4dr} \pm z_{0.5} SE(\bar{y}_{4dr})$$

$$b) \bar{y}_{4dl} = \bar{y}_4 + B(\bar{x}_5 - \bar{x}_5) \text{ (biased)}, \hat{B} = s_{x_5 y_4} / s^2 x_5$$

$$SE(\bar{y}_{4dl}) \doteq \sqrt{\{(n^1-n)/(n^1 n) s^2 y_4 (1-r^2) + (N-n^1)/(Nn^1) s^2 y_4\}},$$

$$r = s_{x_5 y_4} / s_{x_5} s_{y_4}$$

$$\% SE = \{SE(\bar{y}_{4dl})/\bar{y}_{4dl}\} 100$$

$$Bias(\bar{y}_{4dl}) \doteq \hat{B}(1/n-1/n^1) (\hat{M}_{21}/s_{x_5 y_4} - \hat{M}_{30}/s^2 x_5),$$

[Sukhatme and Sukhatme, 1970]

$$\hat{M}_{21} \doteq 1/(n-1) \sum_{i=1}^n (x_5 - \bar{x}_5)^2 (y_4 - \bar{y}_4) \text{ and } \hat{M}_{30} \doteq 1/(n-1) \sum (x_5 - \bar{x}_5)^3$$

$$Bias(\hat{Y}_{4dl}) = N Bias(\bar{y}_{4dl})$$

$$\% Bias = \{Bias(\hat{Y}_{4dl})/\hat{Y}_{4dl}\} 100$$

$$95\% \text{ C. I. for } \bar{y}_{4dl} : \bar{y}_{4dl} \pm z_{.05} SE(\bar{y}_{4dl})$$

9.3 Example

A total of 200 households is considered as a population in an area of Chittagong district in Bangladesh. At first phase a sample of 100 households from the population is selected at random. At second phase a sample of 40 households is selected at random from a sample of first phase. Data of area of compact tree garden and its volume of trees are collected and presented below. Estimate the following parameter by ratio and regression methods:

(a) Volume of compact trees per hectare.

Table 7. A sample of 40 households at second phase of double sampling

i	x_2	y_3	i	x_2	y_3
1	.02	7.76	21	.03	1.44
2	.02	12.4	22	.02	2.90
3	.02	34.50	23	.16	52.10
4	.02	0.38	24	.02	12.14
5	.02	3.93	25	.03	3.33
6	0.14	1.58	26	.01	1.60
7	0.15	16.52	27	.03	8.98
8	.05	1.03	28	.03	2.5
9	.03	5.16	29	.02	3.61
10	.02	.76	30	.01	5.57
11	.02	2.49	31	.08	9.30
12	.02	8.73	32	.02	6.36
13	.04	5.29	33	.06	4.72
14	.01	3.72	34	.06	5.64
15	.07	5.78	35	.07	21.56
16	.07	23.82	36	.06	2.24
17	.02	1.99	37	.02	2.64
18	.02	8.39	38	.02	3.80
19	.03	3.05	39	.04	6.01
20	.02	.68	40	.04	10.51

$$N = 200,$$

$$n_1 = 100,$$

$$n = 40$$

$$\bar{x}_2 = .038$$

$$\hat{B} = 123.309$$

$$\Sigma x_2^2 = .1123 \quad r = .4551$$

$$\bar{y}_3 = 7.704 \quad s^2 y_3 = 1/n-1 \{ \Sigma y_3^2 - n y_3^2 \} = 1/39 \{ 6321.9273 - 40 \times (7.704)^2 \}$$

$$= 1/39 \{ 3947.8627 \} = 101.2272, \quad s y_3 = 10.06$$

$$\Sigma y_3^2 = 6321.9273$$

$$\Sigma x_2 y_3 = 18.418$$

$$\hat{R} = \bar{y}_3 / \bar{x}_2 = 202.7368 \quad s^2 x_2 = .0014, \quad s x_2 = .037$$

$$(a) \quad \text{Est } \bar{Y} = \bar{y}_{3dR} = (\bar{y}_3 / \bar{x}_2) \bar{x}_2^1 = \hat{R} \bar{x}_2^1 = 202.7368 \times .038 = 7.7039$$

$$\text{SE}(\bar{y}_{3dR}) = \sqrt{(s^2 y_3 - 2 \hat{R} s_{x_2 y_3} + \hat{R}^2 s^2 x_2) + (2 \hat{R} s_{x_2 y_3} - R^2 s^2 x_2) / n^1 \text{ for large } N}$$

$$= \sqrt{(101.2272 - 2 \times 202.7368 \times .1719 + (202.7368)^2 \times .0014) / 40 + (69.7009 - 57.5403) / 100}$$

$$= \sqrt{(2.2267 + .1216)}$$

$$= 1.53$$

$$\% \text{ SE}(\bar{y}_{3dR}) = 1.53 / 7.7039 \times 100 = 19\%$$

$$s_{y_3 x_2} = 1/(n-1) \{ \Sigma \bar{x} \bar{y} - n x y \}$$

$$= 1/39 \{ 18.418 - 40 \times .038 \times 7.704 \}$$

$$= .1719$$

$$\text{Here, } C x_2 = s x_2 / \bar{x}_2 = 0.037 / 0.038 = 0.985$$

$$C y_3 = s y_3 / \bar{y}_3 = 10.06 / 7.704 = 1.306$$

$$\text{Bias}(\bar{y}_{3dR}) = \bar{y}_{3dR} (N-n/Nn) \{ (C^2 x_2 - r(C x_2 C y_3)) \}$$

$$= 7.7039 (100-40/100 \times 40) \{ (0.985)^2 - (0.4551)(0.985 \times 1.306) \}$$

$$= 0.1156 \times 0.5151 \times 1.2864$$

$$= 0.0766$$

$$\% \text{ Bias}(\bar{y}_{3dR}) = \text{Bias}(\bar{y}_{3dR}) / \bar{y}_{3dR} \times 100$$

$$= 0.0766 / 7.7039 \times 100$$

$$= 5\%$$

95% Confidence Interval for $\bar{Y}_{3dR}$:

$$\bar{y}_{3dR} \pm Z_{.05} SE(\bar{y}_{3dR})$$

$$\text{Lower limit : } 7.55 - 1.96 \times 1.53 = 4.5512$$

$$\text{Upper limit : } 7.55 + 1.96 \times 1.53 = 10.5488$$

$$\begin{aligned} \text{(b) } \bar{y}_{3dR} &= \bar{y}_3 + \hat{B}(\bar{x}_2 - \bar{x}_1) \\ &= 7.704 + 123.309(0.0432 - 0.038) \\ &= 8.3452 \end{aligned}$$

$$\begin{aligned} SE(y_{3dR}) &= \sqrt{\{(1/n_1 - 1/n) s^2 y_3 (1/r^2) + (N-n/Nn) s^2 y_3\}} \\ &= \sqrt{\{(100-40/100 \times 40) 101.2272\{1 - (0.4551)^2\} + 200-100/200 \times 100) 101.2272\}} \\ &= \sqrt{\{(6/400 \times 101.2272 \times 0.7929 + 1/200 \times 101.2272\}} \\ &= \sqrt{1.71} = 1.3077 \end{aligned}$$

$$\begin{aligned} \% SE(\bar{y}_{3dl}) &= 1.3077/8.3452 \times 100 \\ &= 15\% \end{aligned}$$

$$\begin{aligned} \text{Bias}(\bar{y}_{3dl}) &= \hat{B}(n_1 - n)/n_1 n_2 (\hat{M}_{21}/S_{x_2 y_3} - \hat{M}_{30}/s^2_{x_2}) \\ &= 123.309(100-40/100 \times 40) (0.014/3467.93 - 0.00023/0.0014) \\ &= 1.85(0.000004 - 0.16) \\ &= -0.3 \end{aligned}$$

$$\therefore I \text{ Bias}(\bar{y}_{3dl}) I = 0.3$$

$$\begin{aligned} \% \text{ Bias } (\bar{y}_{3dl}) &= .3/8.3452 \times 100 \\ &= 3\% \end{aligned}$$

95% Confidence Interval for $\bar{Y}_{3dl}$:

$$\text{Lower limit : } 8.342 - 1.96 \times 1.3077 = 5.78$$

$$\text{Upper limit : } 8.3452 + 1.96 \times 1.3077 = 10.91$$

CHAPTER X

SYSTEMATIC SAMPLING

10.1 Definition and discussion

Systematic sampling is a method in which the 1st unit is selected at random and the remaining units are selected automatically according to some pattern. Suppose a population size N is serially numbered from 1 to N . let $N = nk$, where k and n are two integers, n being sample size. Draw a random number less than or equal to k say i , and the unit with corresponding serial number and every k th unit in the population thereafter. The sample with n units will be thus $i, i+k, i+2k, \dots, i+(n-1)k$. Such a sample is called a systematic sample. This is also called one dimensional systematic sample. There is also two dimensional systematic samples available. The method of estimating the parameter and its standard error is similar to one dimensional case given later. The method is equivalent to select villages systematically and household systematically from each selected village in case of village forest surveys.

One independent systematic sample does not provide unbiased estimate of the parameter and its standard error. So, at least two independent systematic samples with random start are required to estimate the precision.

10.2 Estimation of parameters

The following parameters are estimated by systematic sampling:

i) Estimate of total of an auxiliary character:

$$\begin{aligned}\hat{X}_2 &= N\bar{x}_{2sy}, \quad \text{When } \bar{x}_{2sy} = (\bar{x}_{2sy1} + \bar{x}_{2sy2})/2 \\ &= SE(\hat{X}_2) = N SE(\bar{x}_{2sy}) = N\sqrt{1/2(2-1)} \{(\bar{x}_{2sy1} - \bar{x}_{2sy})^2 + (\bar{x}_{2sy2} - \bar{x}_{2sy})^2\}\end{aligned}$$

$$\text{When } \bar{x}_{2sy1} = 1/n \sum x_2$$

$$\% SE(\hat{X}_2) = SE(\hat{X}_2) / \hat{X}_2 \times 100$$

ii) Estimate of total of a main character:

$$\hat{Y}_3 = N\bar{y}_{3sy} \text{ when } \bar{y}_{3sy} = (\bar{y}_{3sy1} + \bar{y}_{3sy2})/2$$

$$\text{When } \bar{y}_{3sy1} = 1/n\Sigma y_3$$

$$SE(\hat{Y}_3) = N SE(\bar{y}_{3sy})$$

$$= N\sqrt{1/2(2-1)\{(\bar{y}_{3sy1}-\bar{y}_{3sy})^2 + (\bar{y}_{3sy2}-\bar{y}_{3sy})^2\}}$$

$$\% SE(\hat{Y}_3) = SE(\hat{Y}_3) / \hat{Y}_3 \times 100$$

10.3 Approximation to standard error using simple random sample formula (Chacko,1965)

If the units in the population can be assumed to be arranged in space at random, then a systematic sample with random start is clearly equivalent to a simple random sample of the units.

Treating the sample as a simple random sample the estimation of mean volume of compact trees per hectare and its standard error is given below:

$$\hat{R} = \bar{y}_{3sy} / \bar{x}_{2sy}$$

$$SE(\hat{R}) = \sqrt{N-n/Nn(n-1)\bar{x}_2^{-2} (\Sigma y_2^2 - 2\hat{R}\Sigma x_2y_3 + \hat{R}^2\Sigma x_2^2)}$$

$$\%SE(\hat{R}) = SE(\hat{R})/\hat{R} \times 100$$

95% Confidence Interval for R :

$$\text{Lower limit : } \hat{R} - 1.96 SE(\hat{R})$$

$$\text{Upper limit : } \hat{R} + 1.96 SE(\hat{R})$$

10.4 Approximation to standard error using stratified random sample formula

This method consists in considering data collected as divided into strata and sampling error is computed as if the units are selected at random from each stratum. The sampling in fact stratifies the population into n strata each of which consists of k sampling units. Treating the systematic sample of simple random sample as stratified random sample taking 2 units (households), the estimate of volume per hectare and its standard error is given below:

$$\hat{R} = \sum_{t=1}^k N_t \bar{y}_{3t} / \sum_{t=1}^k N_t \bar{x}_{2t}, \text{ where } N_t = \text{size of } t \text{ th stratum, } k = \text{no. of strata}$$

n = size of systematic sample, N = population size

$$SE(\hat{R}) = 1/N\bar{x}\sqrt{\sum_{t=1}^k N_t(N_t-n_t)/n_t \cdot 1/(nt-1) \sum_{t=1}^k (y_{3t}-\bar{y}_{3t})^2 + \hat{R}^2 \sum (x_{2t}-\bar{x}_{2t})^2 - 2\hat{R} \sum (y_{3t}-\bar{y}_{3t})(x_{2t}-\bar{x}_{2t})}$$

$$\%SE(\hat{R}) = SE(\hat{R})/\hat{R} \times 100$$

95% Confidence Interval for R:

$$\text{Lower limit : } \hat{R} - 1.96 SE(\hat{R})$$

$$\text{Upper limit : } \hat{R} + 1.96 SE(\hat{R})$$

10.5 Approximation to standard error using method of first difference

In case of systematic households for village forest surveys using one dimensional systematic sampling an approximation to standard error may be obtained from difference between pairs of successive units. If there are n units in the systematic sample, there will be (n-1) difference. The volume per hectare and its standard error may be estimated as given below:

$$\hat{R} = \Sigma y_3 / \Sigma x_2 = \bar{y}_3 / \bar{x}_2$$

$$SE(\hat{R}) = 1/\bar{x}_2 \sqrt{(N-n)/2Nn \{ \Sigma (d(y_3))^2 + \hat{R}^2 \Sigma (d(x_2))^2 - 2\hat{R} \Sigma d(x_2)d(y_3) \}}$$

$$\text{When } d(y_3) = y_{3i+1} - y_{3i} \text{ (i=1, 2,...n-1)}$$

$$\% SE(\hat{R}) = SE(\hat{R})/\hat{R} \times 100$$

95% Confidence Interval for R:

$$\text{Lower limit: } \hat{R} - 1.96 SE(\hat{R})$$

$$\text{Upper limit: } \hat{R} + 1.96 SE(\hat{R})$$

10.6 Approximation to standard error using method of second difference

Another method of approximation to standard error of a systematic sample is based on second differences. The estimate of volume per hectare and its standard error is given below:

$$\hat{R} = \bar{y}_3 / \bar{x}_2$$

$$SE(\hat{R}) = 1/\bar{x}_2 \sqrt{[(N-n)/6Nn(n-1)] \{ \Sigma (D(y_3))^2 + \hat{R}^2 \Sigma (D(x_2))^2 - 2\hat{R} \Sigma D(x_2) D(y_3) \}}$$

$$\text{Where } D(x_2) = d(x_2)_i - d(x_2)_{i-1} = D_{11}$$

$$D(y_3) = d(y_3)_i - d(y_3)_{i-1} = D_{22}$$

$$\% SE(\hat{R}) = SE(\hat{R})/\hat{R} \times 100$$

95% Confidence Interval for R :

$$\text{Lower limit : } \hat{R} - 1.96 SE(\hat{R})$$

$$\text{Upper limit : } \hat{R} + 1.96 SE(\hat{R})$$

10.7 Advantages and disadvantages

Systematic sampling is a common technique if the complete but up to date frame is available. It is easier and quicker to use. Systematic sampling is extensively used in various fields including village forest surreys on account of its low cost and simplicity in the selection of the sampling units. In case of village forest surreys a systematic sample of households or villages may be selected from a complete list of households or villages. It also offers great advantages in organizing control over field work.

Systematic sample will give more precise estimate sometimes than simple random sample. Systematic sampling is not efficient in comparison with stratified random sampling in presence of a linear trend in the population. If the population exhibits a periodic variation and if the sampling interval of the systematic sample coincides with the period of the cycle systematic sampling does not give precise estimates.

10.8 Example

A systematic sample of 10 households is selected from a population of 64 households in an area of Chittagong district in Bangladesh. Data of x_2 and y_3 of each selected household are collected. Estimate (a) total area of compact tree garden, (b) total volume of compact trees, (c) volume of wood per hectare by simple random sample method, stratified random sample method, method of first difference and method of second difference through systematic sampling.

$$N = 64, \quad n = 16, \quad k = 64 \div 16 = 4$$

Table 8. A population of 64 households of systematic sampling

i	x_2	y_3	i	x_2	y_3	i	x_2	y_3
1	.03	5.55	25	.05	3.41	49	.06	5.01
2	.04	4.06	26	.02	3.58	50	.03	4.07
3	.03	5.46	27	.02	5.45	51	.26	3.83
4	.04	3.90	28	.01	4.50	52	.03	3.52
5	.05	5.15	29	.02	4.4	53	.05	5.71
6	.02	4.15	30	.05	4.01	54	.04	4.54
7	.03	5.54	31	.02	5.01	55	.02	4.63
8	.04	4.75	32	.03	4.86	56	.14	5.17
9	.03	5.61	33	.07	3.53	57	.03	3.92
10	.02	5.20	34	.02	5.12	58	.02	4.96
11	.02	4.22	35	.04	5.38	59	.04	5.07
12	.02	4.09	36	.01	4.99	60	.02	3.93
13	.02	3.93	37	.02	3.60	61	.07	4.29
14	.04	5.19	38	.07	3.76	62	.05	4.23
15	.06	3.57	39	.05	4.47	63	.04	7.24
16	.04	4.93	40	.08	5.74	64	.03	3.05
17	.02	4.84	41	.07	4.79			
18	.06	3.74	42	.02	5.93			
19	.02	5.27	43	.02	3.73			
20	.03	5.12	44	.02	4.06			
21	.03	4.01	45	.01	3.86			
22	.01	5.88	46	.04	4.56			
23	.02	5.71	47	.02	4.63			
24	.10	4.11	48	.06	4.38			

Table 9. Two independent systematic samples each of size 16

Systematic sample-1	i:	1	5	9	13	17	21	25	29	33
	x ₂ :	.03	.05	.03	.02	.02	.03	.05	.02	.07
	y ₃ :	5.55	5.15	5.61	3.93	4.84	4.01	3.41	4.40	3.53
	i:	37	41	45	49	53	57	61		
	x ₂ :	.02	.07	.01	.06	.05	.03	.07		
Systematic sample-2	y ₃ :	3.60	4.79	3.86	5.01	5.71	3.92	4.29		
	i:	3	7	11	15	19	23	27	31	35
	x ₂ :	.03	.03	.02	.06	.02	.02	.02	.02	.04
	y ₃ :	5.46	5.54	4.22	3.57	5.27	5.71	5.45	5.01	5.38
	i:	39	43	47	51	55	59	63		
	x ₂ :	.05	.02	.02	.26	.02	.04	.04		
	y ₃ :	4.47	3.73	4.63	3.83	4.63	5.07	7.24		

(a) Estimate of total area of compact tree garden:

Systematic sample-1 : $\bar{x}_2 = .039$

Systematic sample-2 : $\bar{x}_2 = .044$

$$\begin{aligned}\hat{X}_2 &= N\bar{x}_{2sy} & \bar{x}_{2sy} &= (.039 + .044)/2 \\ &= 64 \times .0415 & &= .0415 \\ &= 2.656\end{aligned}$$

$$\begin{aligned}SE(\hat{X}_2) &= N\sqrt{1/m(m-1)[(\bar{x}_{2sy1} - \bar{x}_{2sy})^2 + (\bar{x}_{2sy2} - \bar{x}_{2sy})^2]} \\ &= 64\sqrt{1/2(2-1)[(.039 - .0415)^2 + (.044 - .0415)^2]} \\ &= 64\sqrt{1/2(.00000625 + .00000625)} \\ &= 0.16\end{aligned}$$

$$\%SE(\hat{X}_2) = .16/2.656 \times 100 = 6\%$$

$$\%Accuracy = 100 - (X_2 - \hat{X}_2)/X_2 \times 100 = 100 - (2.28 - 2.66)/2.28 \times 100 = 83\%$$

95% Confidence Interval for X_2 :

$$\text{Lower limit : } 2.656 - 1.96 \times .16 = 2.9696$$

$$\text{Upper limit : } 2.656 + 1.96 \times .16 = 2.3424$$

(b) Estimate of total volume of compact trees:

Systematic sample-1 : $\bar{y}_3 = 4.4756$

Systematic sample-2 : $\bar{y}_3 = 4.9506$

$$\begin{aligned}\hat{Y}_3 &= N \bar{y}_{3sy} & \bar{y}_{3sy} &= (4.4756 + 4.9506)/2 \\ &= 64 \times 4.7131 & &= 4.7131 \\ &= 301.6384\end{aligned}$$

$$SE(\hat{Y}_3) = N\sqrt{1/m(m-1)[(\bar{y}_{3sy} - \bar{y}_{3sy})^2 + (\bar{y}_{3sy} - \bar{y}_{3sy})^2]}$$

$$\begin{aligned}SE(\hat{Y}_3) &= 64\sqrt{1/2(2-1)[(4.4756 - 4.7131)^2 + (4.9506 - 4.7131)^2]} \\ &= 64\sqrt{1/2 (.0564 + .0564)} \\ &= 15.1992\end{aligned}$$

$$\% SE(\hat{Y}_3) = 15.1992/301.6384 \times 100 = 5\%$$

$$\begin{aligned}\% \text{Accuracy} &= 100 - (Y_3 - \hat{Y}_3)/Y_3 \times 100 = 100 - (219.48 - 301.64) \\ &\quad /219.48 \times 100 = 63\%\end{aligned}$$

95% Confidence Interval for Y_3 :

$$\text{Lower limit : } 301.6384 - 1.96 \times 15.1992 = 271.8479$$

$$\text{Upper limit : } 301.6384 + 1.96 \times 15.1992 = 331.4288$$

(c) Estimate of volume of wood per hectare
Simple random sample formula

Table 10. A systematic sample of size 16 considered as simple random sample

Selected HH No.	x_2	y_3
1	.03	5.55
5	.05	5.15
9	.03	5.61
13	.02	3.93
17	.02	4.84
21	.03	4.01
25	.05	3.41
29	.02	4.40
33	.07	3.53
37	.02	3.60
41	.07	4.79
45	.01	3.86
49	.06	5.01
53	.05	5.71
57	.03	3.92
61	.07	4.29

$$\hat{R} = \bar{y}_{3sy} / \bar{x}_{2sy} = 4.4756 / 0.039 = 114.7589$$

$$\Sigma x_2 y_3 = 2.8435 \quad \Sigma x_2^2 = .0311, \quad \Sigma y_3^2 = 329.4751$$

$$SE(\hat{R}) = \sqrt{(64-16) / (64 \times 16 \times 15 \times (.039)^2) (329.4751 - 2 \times 114.7589 \times 2.8435 + (114.7589)^2 \times .0311)}$$

$$= \sqrt{48 / 23.3626 (329.475 - 652.6339 + 409.5747)}$$

$$= \sqrt{2.0546 \times 86.4158} = 13.3248$$

$$\%SE(\hat{R}) = 13.3248 / 114.7589 \times 100 = 12\%$$

95% Confidence Interval for R :

$$\text{Lower limit : } 114.7589 - 1.96 \times 13.3248 = 80.64$$

$$\text{Upper limit : } 114.7589 + 1.96 \times 13.3248 = 140.86$$

Stratified random sample formula:

Table 11. A systematic sample of size 16 considered as stratified random sample

Stratum No	HH	x_2	Difference ($x_{2t1} - x_{2t2}$)	(Diff.) ²	y_3	Difference ($y_{3t1} - y_{3t2}$)	(Diff.) ²	Product Difference
I	1	.03	-.02	.0004	5.55	.4	.16	-.008
	5	.05			5.15			
II	9	.03	+.01	.0001	5.16	+1.68	2.8224	.0168
	13	.02			3.93			
III	17	.02	-.01	.0001	4.84	.83	.6889	-.0083
	21	.03			4.01			
IV	25	.05	+.03	.0009	3.41	-.99	.9801	-.0297
	29	.02			4.40			
V	33	.07	+.05	.0025	3.53	-.07	.0049	-.0035
	37	.02			3.60			
VI	41	.07	+.06	.0036	4.79	.91	.8281	.0546
	45	.01			3.86			
VII	49	.06	+.01	.0001	5.01	-.7	.49	-.007
	53	.05			5.71			
VIII	57	.03	-.04	.0016	3.92	-.37	.1369	.0148
	61	.07			4.29			
Total		.63		.0093	71.61	-.0565	6.1113	.0297

Here $N = 64$, $n = 16$, $k = \text{no. of strata} = 8$

$N_t = 8$, $n_t = 2$ in each stratum

$$\hat{R} = \sum N_t \bar{y}_{3t} / \sum N_t \bar{x}_{2t} = 71.61 / .63 = 113.6666$$

$$\begin{aligned}
 V(\hat{R}) &= 1/(N\bar{x}_2)^2 [\sum N_t - n_t / N_t n_t / n_t - 1 \{ \sum \bar{y}_{3t} - \bar{y}_{3t} \}^2 + \hat{R}^2 \sum (x_{2t} - \bar{x}_{2t})^2 - \\
 &\quad 2\hat{R} \sum (x_{2t} - \bar{x}_{2t})(y_{3t} - \bar{y}_{3t})] \\
 &= 1/.63 \{ 8 - 2/8 \times 2 \{ 6.113 + (113.6667)^2 \times .0093 - 2 \times 113.6667 \times .0297 \} \\
 &= 3/8 \times .63 \times 119.5166 \\
 &= .1488 \times 119.5166 = 17.7886
 \end{aligned}$$

$$\therefore SE(\hat{R}) = 4.2176$$

$$\% SE(\hat{R}) = 4.2176 / 113.6667 \times 100 = 4\%$$

∴ 95% Confidence Interval for R

Lower limit : $113.6667 - 1.96 \times 4.2176 = 105.40$

Upper limit : $113.6667 + 1.96 \times 4.2176 = 121.93$

First difference method:

Table 12. A systematic sample of size 16 analyzed by first difference method

Selected HH	x_2	dx_2^2	y_3	dy_3^2	$dx_2 \times dy_3$
1	.03	$(.02)^2$	5.55	$(-.4)^2$	-.008
5	.05	$(-.02)^2$	5.15	$(.46)^2$	-.0092
9	.03	$(-.01)^2$	5.61	$(-1.68)^2$	+.00168
13	.02	$(.00)^2$	3.93	$(.91)^2$	00
17	.02	$(.01)^2$	4.84	$(-.83)^2$	-.0083
21	.03	$(.02)^2$	4.01	$(-.6)^2$	-.012
25	.05	$(-.03)^2$	3.41	$(.99)^2$	-.0297
29	.02	$(.05)^2$	4.40	$(-.87)^2$	-.0435
33	.07	$(-.05)^2$	3.53	$(.07)^2$	-.0035
37	.02	$(.05)^2$	3.60	$(1.19)^2$	.0595
41	.07	$(-.06)^2$	4.79	$(-.93)^2$	.0558
45	.01	$(.05)^2$	3.86	$(1.15)^2$	.0575
49	.06	$(-.01)^2$	5.01	$(.7)^2$	-.007
53	.05	$(-.02)^2$	5.71	$(-1.79)^2$	+.0358
57	.03	$(.04)^2$	3.92	$(.37)^2$	.0148
61	.07	-	4.29	-	-
	.63	.015	71.61	13.4144	.0664

$N = 64,$ $n = 16,$ $x_2 = .0394$ $y_3 = 4.4756$

$$\hat{R} = \bar{y}_3 / \bar{x}_2 = 4.4756 / .0394 = 113.5939$$

$$\begin{aligned} V(\hat{R}) &= (N-n)/2Nn(n-1)\bar{x}^2 [13.4149 + (113.5939)^2 \times .015 - 2 \times \\ &\quad 113.5939 \times .0664] \\ &= (64 - 16)/2 \times 64 \times 16 \times 15 \times (.0394)^2 (191.88) \\ &= 191.88/9935 = 193.1354 \end{aligned}$$

$$\therefore SE(\hat{R}) = 13.8973$$

$$\% SE(\hat{R}) = 12\%$$

$\therefore$ 95% Confidence Interval for R :

$$\text{Lower limit : } 113.5939 - 1.96 \times 13.8973 = 86.36$$

$$\text{Upper limit : } 113.5939 + 1.96 \times 13.8973 = 140.83$$

Second difference method:

Table 13. A systematic sample of size 16 analyzed by second difference method

Selected	HH	D ₁	D ₁₁	D ₂	D ₂₂	D ₁₁ × D ₂₂
1		+.02	-.04	-.4	+.86	-.0344
5		-.02	+.01	+.46	-2.14	-.0214
9		-.01	+.01	-1.68	+2.59	.0259
13		00	+.01	+.91	-1.74	-.0174
17		+.01	+.01	-.83	+.23	+.0023
21		+.02	-.05	-.6	+1.59	-.0795
25		-.03	+.08	+.99	-1.86	-.1488
29		+.05	-.10	-.87	+.94	-.094
33		-.05	+.10	+.07	+1.12	+.112
37		+.05	-.11	+1.19	-2.12	+.2332
41		-.06	+.11	-.93	+2.08	+.2288
45		+.05	-.06	+1.15	-.45	+.027
49		-.01	-.01	+.7	-2.49	+.0249
53		-.02	+.06	-1.79	-2.16	.1296
57		+.04	-	+.37		
61		-	-	-	-	
Total			.02		.77	.3859

N=64,

n=16,

$\bar{x}_2 = .0394$

$\bar{y}_3 = 4.4756$

$$\hat{R} = \bar{y}_3 / \bar{x}_2 = 4.4756 / .0394 = 113.5939, \Sigma D_{11}^2 = .0587, \Sigma D_{22}^2 = 43.1225$$

$$V(\hat{R}) = (64-16) / (6 \times 64 \times 16 \times 15 \times (.0394)^2 [43.1225 + (113.5939)^2 \times .0587 - 2 \times 113.5939 \times .3859])$$

$$= 1/2.9805 [712.89.5] = 239.1849$$

$$\therefore SE(\hat{R}) = 15.4656$$

$$\%SE(\hat{R}) = 14\%$$

$\therefore$ 95% Confidence Interval. for R :

$$\text{Lower limit : } 113.5939 - 1.96 \times 15.4656 = 83.28$$

$$\text{Upper limit : } 113.5939 + 1.96 \times 15.4656 = 143.91$$

CHAPTER XI

PROBABILITY PROPORTIONAL TO SIZE (PPS) SAMPLING

Lihiri's method for selection of PPS sampling units:

A simple method developed by D.B. Lihiri (1951) for selecting sampling units with probability proportional to size (PPS) is given below:

- (a) Choose a pair of random numbers (i,j) where i in the range 1 to N (population size) and j in the range lowest value to highest value of x (size of unit).
- (b) Select the ith unit in the sample if x_i of that unit is larger than or equal to j. Otherwise reject the unit.
- (c) If rejected, repeat (a) and (b) until a selection of the sample is done.

Note : Sometimes the pair of random numbers (i,j) may be considered as remainder dividing selected digit random number i and j by highest value of i and j. The value of j even may be changed in order to compare with x-values.

PPS sampling technique:

Probability Proportional to Size (PPS) Sampling provides better estimates than Simple Random Sampling. In homestead woodlot survey, size (area of land, number of trees, etc.) of the sampling units (the households) vary largely. There are many households which have small number of trees having large area of homestead land. On the other hand, many households with small area of land have plenty of trees. So, PPS sampling would contribute better estimates in the homestead population containing unequal size sampling units. It gives unbiased estimates if the sampled units are properly weighted with the probabilities of selection of the units. The selection of the units is made with replacement so that the randomization leads to a unit already selected it should again be included in the sample as a different unit. This would ensure that the probability of drawing a specified unit at any given draw is the same (Sukhatme and Sukhatme, 1970). Lahiri (1951) suggested a procedure for selection of the

units with PPS in the sample which avoids the necessity of computing cumulative totals. This procedure has been followed in selecting the sample in choice of sampling designs under the study. The same sample size has been considered in each of the sampling designs under the study. A few selected parameters under the variables which involve the units with different sizes, was estimated through PPS sampling method. The following formulas are used for the purpose :

11.1 PPS Sampling (PPS of households with replacement)

A sample of n households out of N households was selected with PPS (area of compact tree garden, number of bamboo clumps and number of trees) by Lahiri's method.

i) Estimate of total volume of trees is $\hat{Y}_2 = R_{1(pps)} \bar{N} \bar{Y}_1$. (unbiased):

$\hat{R}_{1(pps)}$ (estimate of mean volume per tree)

$$= 1/n \sum v_i, \quad v_i = y_{2i} / y_{1i}$$

$N \bar{Y}_1$ = Total number of trees

$$SE(\hat{Y}_1) = N \bar{Y}_1 \sqrt{\{\sum (v_i - \bar{v})^2 / n(n-1)\}}$$

$$\% SE = \{SE(\hat{Y}_1) / \hat{Y}_1\} 100$$

$$\% \text{Accuracy} = 100 - \{(Y_1 - \hat{Y}_1) / Y_1\} 100$$

$$95\% \text{ C. I. for } Y_1 : \hat{Y}_1 \pm z_{.05} SE(\hat{Y}_1)$$

ii) **Estimate of total volume of compact trees is $Y_3 = \hat{R}_{2(pps)} \bar{NX}_2$. (unbiased):**

$\hat{R}_{2(pps)}$ (estimate of volume per hectare of compact trees)

$$= (1/n) \sum_{i=1}^n v_i, v_i = y_{3i} / x_{2i}$$

$\bar{NX}_2$ = Total areda of compact tree garden.

$$SE(\hat{Y}_3) = \bar{NX}_2 \sqrt{\{\sum (v_i - \bar{v})^2 / n(n-1)\}}$$

$$\% SE = \{SE(\hat{Y}_3) / \hat{Y}_3\} 100$$

$$\% \text{Accuracy} = 100 - \{(Y_3 - \hat{Y}_3) / Y_3\} 100$$

$$95\% \text{ C. I. for } Y_3 : \hat{Y}_3 \pm z_{.05} SE(\hat{Y}_3)$$

iii) **Estimate of total number of bamboos is $\hat{Y}_4 = \hat{R}_{3(pps)} \bar{NX}_5$. (unbiased):**

$\hat{R}_{3(pps)}$ (estimate of number of bamboos per clump)

$$= 1/n \sum v_i, v_i = y_{4i} / x_{5i}$$

$\bar{NX}_5$ = Total number of bamboos clumps.

$$SE(\hat{Y}_4) = \bar{NX}_5 \sqrt{\{\sum (v_i - \bar{v})^2 / n(n-1)\}}$$

$$\%SE(\hat{Y}_4) = \{(\hat{Y}_4) / \hat{Y}_4\} 100$$

$$\% \text{Accuracy} = 100 - \{(Y_4 - \hat{Y}_4) / Y_4\}$$

$$95\% \text{ C. I. for } Y_4 : \hat{Y}_4 \pm z_{.05} SE(\hat{Y}_4)$$

11.1.1 Example

A population of 250 households is considered in an area of Chittagong district of Bangladesh. A sample of 50 households is selected with probability proportion to its area of compact tree garden from the population by Lihiri's method. Data of compact tree garden area (x_2) and volume of compact trees (y_3) are collected and given below. Estimate total volume of compact trees with reliabilities.

Table 14. A PPS sample of 50 households with data of y_3 & x_2

S.L.	x_2	y_3	$ri= y_3/ x_2$
1	0.02	14.84	742
2	0.02	4.62	231
3	0.02	11.43	57.150
4	0.02	22.35	111.75
5	0.08	12.99	162.375
6	0.12	14.79	123.25
7	0.24	8.82	36.75
8	0.08	21.31	266.375
9	0.02	6.36	318
10	0.14	5.17	36.929
11	0.20	50.32	251.6
12	0.14	1.58	11.286
13	0.04	11.40	285
14	0.08	7.61	95.125
15	0.04	8.52	213
16	0.03	1.84	61.333
17	0.01	3.77	377
18	0.07	12.24	174.857
19	0.14	30.09	214.929

S.L.	x_2	y_3	$ri = y_3 / x_2$
20	0.08	8.73	109.125
21	0.13	46.32	356.308
22	0.07	3.76	57.714
23	0.16	1.04	6.5
24	0.45	17.89	39.756
25	0.65	15.15	23.308
26	0.02	12.20	610
27	0.05	4.19	838
28	0.00	3.15	63
29	0.12	14.79	123.25
30	0.03	4.68	156
31	0.16	53.15	332.188
32	0.24	8.82	36.75
33	0.05	12.59	251.8
34	0.06	14	233.333
35	0.09	13.08	145.333
36	0.02	4.96	248
37	0.08	2.40	30
38	0.26	12.84	49.385
39	0.04	13.50	337.5
40	0.04	9.03	225.75
41	0.12	8.73	72.75
42	0.12	50.82	423.5
43	0.04	13.21	330.25
44	0.02	0.68	32
45	0.04	1.09	27.25

S.L.	x_2	y_3	$ri = y_3 / x_2$
46	0.10	5.36	53.60
47	0.12	27.84	232
48	0.45	17.89	39.756
49	0.02	13.19	659.50
50	0.03	13.98	466

$$N = 250, \quad n = 50$$

$$\hat{Y}_3 = R_2 (PPS) N \bar{X}_2$$

$$\text{Here, } R_2 = 1/n \sum v_i$$

$$= 192.4938$$

$$\hat{Y}_3 = 192.4938 \times 40 \quad [\bar{X}_2 = .16, N = 250]$$

$$= 7699.75$$

$$SE(\hat{Y}) = 40 \times 172.4893 \div \sqrt{50}$$

$$= 975.747$$

$$\%SE = \{SE(\hat{Y}_3) / \hat{Y}_3\} 100$$

$$= 975.747 / 7699.75 \times 100$$

$$= 12\%$$

$$Y_3 = 9002 \text{ (Estimate from 250 PPS Sample)}$$

$$\% \text{Accuracy} = 100 - \{(9002 - 7699.75) / 9002\} 100$$

$$\approx 86\%$$

95% Confidence level Y_3 :

$$\hat{Y}_3 \pm Z_{.05} SE(\hat{Y}_3)$$

$$\text{Lower limit : } 7699.75 - 1.96 \times 975.747 = 5787$$

$$\text{Upper limit : } 7699.75 + 1.96 \times 975.747 = 9612$$

11.2 PPS Stratified Sampling (PPS of households with replacement

A sample of n households from N household is allocated to L trata proportionately and allocated households are selected with probability proportional to size (area of compact tree garden, nmber of bamboo clumps and number of trees) from each stratum by lahiri's method.

i) Estimate of total volume of trees is $\hat{Y}_2 = \hat{R}_{1(pps)} N\bar{Y}_1$. (unbiased):

$\hat{R}_{1(pps)}$ (estimate of volume per tree)

$$= 1/A \sum_h^L A_h (1/n_h) \sum_j^{n_h} v_{hj}, \quad v_{hj} = Y_{2hj} / Y_{1hj}$$

$$SE(\hat{Y}_2) = N\bar{Y}_1 \sqrt{\{1/A^2 \sum_h^L A_h^2 / n_h (n_h - 1) \sum_j^{n_h} (v_{hj} - \bar{v}_h)^2\}}$$

$$\%SE = \{SE(\hat{Y}_2) / \hat{Y}_2\} 100$$

$$\%Accuracy = 100 - \{(Y_2 - \hat{Y}_2) / Y_{12}\} 100$$

$$95\% \text{ C.I. for } Y_2 : \hat{Y}_2 \pm z_{.05} SE(\hat{Y}_2)$$

ii) Estimate of total volume of compact trees is $\hat{Y}_3 = \hat{R}_{2(pps)} N\bar{X}_2$ (unbiased):

$\hat{R}_{2(pps)}$ (estimate of volume of compact trees per hectare)

$$= 1/A \sum_h^L A_h (1/n_h) \sum_j^{n_h} v_{hj}$$

$$V_{hj} = Y_{3hj} / X_{2hj}$$

$$SE(\hat{Y}_3) = N\bar{X}_2 \sqrt{\{1/A^2 \sum_h^L A_h^2 / n_h (n_h - 1) \sum_j^{n_h} (v_{hj} - \bar{v}_h)^2\}}$$

$$\%SE = \{SE(\hat{Y}_3) / \hat{Y}_3\} 100$$

$$\%Accuracy = 100 - \{(Y_3 - \hat{Y}_3) / Y_3\} 100$$

$$95\% \text{ C. I. for } Y_3 : \hat{Y}_3 \pm z_{.05} SE(\hat{Y}_3)$$

iii) Estimate of total number of bamboos is $\hat{Y}_4 = \hat{R}_{3(PPS)} N\bar{X}_5$. (unbiased):

$\hat{R}_{3(PPS)}$ (estimate of number of bamboos per clump)

$$= 1/A \sum_{h=1}^L A_h (1/n_h) \sum_{j=1}^{n_h} v_{hj}, v_{hj} = Y_{2hj}/X_{1hj}$$

$$SE(\hat{Y}_4) = N\bar{X}_5 \sqrt{\{1/A^2 \sum_{h=1}^L A_h^2 h/n_h(n_h-1) \sum_{j=1}^{n_h} (V_{hj} - \bar{V}_h)^2\}}$$

$$\% SE = \{SE(\hat{Y}_4)/\hat{Y}_4\} 100$$

$$\% Accuracy = 100 - \{(Y_4 - \hat{Y}_4)/Y_4\} 100$$

$$95\% \text{ C.I. for } Y_4 : \hat{Y}_4 \pm z_{.05} SE(\hat{Y}_4)$$

Where,

A_h = Area under compact tree garden/number of bamboo clumps/ number of trees for hth stratum

$A = \sum_{h=1}^L A_h$ = Total area under compact tree garden/total number of bamboo clumps/total number of trees in the population.

n_h = Sample size for hth stratum.

L = Number of strata.

$Y_1, Y_2, Y_3, Y_4, X_2, X_5$ indicate as before.

11.2.1 Example

A Sample of 34 households from 112 households is allocated by proportional allocation to 3 strata and allocated households are selected with probability proportion to area of compact tree garden by Lihiri's method in Chittagong district of Bangladesh. Data of area of compact tree garden and volume of compact trees are collected from sample households. Estimate total volume of compact trees and its reliabilities.

Table 15. A PPS sample of 34 households allocated to 3 strata with data of y_3 & x_2

S_1	x_2	y_3	PPS HH $V_i=y_3/x_2$	S_2	x_2	y_3	PPS HH $V_i=y_3/x_2$	S_3	x_2	y_3	PPS HH $V_i=y_3/x_2$
1	1.25	177.06		1	0.24	8.82	36.75	1	0.08	5.74	71.75
2	1.25	177.06		2	0.24	8.82	36.75	2	0.02	10.10	
3	1.25	177.06		3	0.24	8.82	36.75	3	0.03	6.14	
4	1.25	177.06		4	0.16	8.31		4	0.04	8.15	
5	1.25	177.06		5	0.08	4.67		5	0.04	8.42	
6	0.05	12.04		6	0.26	13.26		6	0.02	8.63	
7	0.20	12.82		7	0.26	3.83	14.73077	7	0.08	24.91	
8	0.36	67.28		8	0.65	15.15	23.30769	8	0.06	15.37	
9	0.36	67.28		9	0.65	15.15	23.30769	9	0.16	24.19	
10	0.36	67.28		10	0.65	15.15	23.30769	10	0.16	24.19	
11	0.32	18.30	57.1875	11	0.03	3.45		11	0.16	24.19	
12	0.32	18.30	57.1875	12	0.10	5.36		12	0.01	20.11	
13	0.07	15.79		13	0.05	2.63		13	0.12	28.98	
14	0.04	4.75		14	0.08	7.61		14	0.12	28.98	
15	0.06	2.24	37.333333	15	0.02	2.74		15	0.02	11.00	
16	0.04	8.69		16	0.05	6.52		16	0.16	10.17	63.5625
17	0.02	1.26	63	17	0.14	10.88		17	0.10	9.69	96.9
18	0.02	0.88	44	18	0.14	10.88		18	0.10	15.81	
19	0.10	15.06		19	0.02	3.01		19	0.06	11.04	
20	0.07	6.50		20	0.04	10.96		20	0.04	4.56	114
21	0.14	15.71		21	0.26	16.53		21	0.32	29.67	92.71875
22	0.16	7.52	47	22	0.41	22.36		22	0.32	29.67	92.71875
23	0.16	6.45	40.3125	23	0.12	8.73		23	0.32	29.67	92.71875
24	0.12	23.57		24	0.07	17.09		24	0.24	18.51	77.125
25	0.16	2.66	16.625	25	0.13	8.64		25	0.24	18.51	77.125
26	0.49	22.01	44.918367	26	0.02	24.38		26	0.02	1.65	82.5
27	0.49	22.01	44.918367	27	0.26	11.26	43.30769	27	0.04	4.15	103.75
28	0.49	7.05	14.387755	28	0.26	11.26		28	0.03	1.84	61.33333
29	0.05	1.03	20.6	29	0.26	11.26		29	0.08	17.52	
30	0.03	5.06		30	0.20	6.04	30.2	30	0.06	8.66	
31	0.03	5.46		31	0.20	6.04	30.2	31	0.04	10.30	
32	0.07	13.99		32	0.08	12.08		32	0.06	8.77	

S_1-1	x_2	y_3	PPS HH $V_i=y_3/x_2$	S_2	x_2	y_3	PPS HH $V_i=y_3/x_2$	S_3	x_2	y_3	PPS HH $V_i=y_3/x_2$
33	0.02	12.04		33				33	0.06	8.77	
34	0.16	43.80		34				34	0.06	8.77	
35	0.16	43.80		35				35	0.16	85.05	
36	0.12	19.40		36				36	0.20	36.74	
37	0.16	45.54		37				37	0.09	13.08	
38	0.04	8.07		38				38	0.08	28.34	
39	0.10	39.84		39				39	0.02	12.59	
40	0.10	39.84		40				40	0.05	11.56	

$$\bar{V}_1 = 40.62$$

$$\bar{V}_2 = 29.86$$

$$\bar{V}_3 = 85.56$$

$$s_1 = 16.00$$

$$s_2 = 8.70$$

$$s_3 = 16.05$$

$$\hat{Y}_3 = R_2(\text{PPS}) \times NX_2$$

$$= 45.53 \times 22.21$$

$$= 1011.14$$

$$R_2(\text{PPS}) = 1/A \{A_1 \bar{V}_1 + A_2 \bar{V}_2 + A_3 \bar{V}_3\}$$

$$= 1/22.21 \{11.89 \times 40.62 + 6.37 \times 29.86 + 3.95 \times 85.56\}$$

$$= 45.53$$

$$SE(\hat{Y}_3) = NX_2 \sqrt{1/A^2 \sum A_h^2 / n_h \{ (1/n_h - 1) \sum (V_{hi} - \bar{V})^2 \}}$$

$$= \sqrt{\sum A_h^2 / n_h \{ (1/n_h - 1) \sum (V_{hi} - \bar{V})^2 \}}, \quad \text{as } A = NX_2$$

$$= \sqrt{(11.89^2/12 \times 16^2 + 6.37^2/10 \times 8.70^2 + 3.95^2/12 \times 10.05^2)}$$

$$= \sqrt{3658.002}$$

$$= 60.48$$

$$\%SE(\hat{Y}_3) = 60.48/1011.14 \times 100 = 6\%$$

$$\begin{aligned}\% \text{ Accuracy} &= 100 - (Y_3 - \hat{Y}_3) / Y_3 \times 100 \\ &= 100 - (1048.64 - 1011.14) / 1048.64 \times 100 \\ &= 92.99\%\end{aligned}$$

as $Y_3 = 1048.64$ from previous experience

95% Confidence level Y_3 :

$$\begin{aligned}\text{Lower limit : } Y_3 &= \hat{Y}_3 - 1.96 \times SE(\hat{Y}_3) \\ &= 1011.14 - 1.96 \times 60.28 \\ &= 892.99\end{aligned}$$

$$\begin{aligned}\text{Upper limit : } Y_3 &= \hat{Y}_3 + 1.96 \times SE(\hat{Y}_3) \\ &= 1011.14 + 1.96 \times 60.28 \\ &= 1129.29\end{aligned}$$

11.3 PPS Cluster sampling (PPS of clusters with replacement)

A sample of n clusters such that sample households out of N clusters is selected with PPS of cluster size M_i by Lahiri's method.

i) Estimate of total volume of trees:

$$\hat{Y}_1 = N\bar{y}_1 \quad (\text{unbiased})$$

$$\bar{y}_1 = \bar{z}_{1n} = 1/n \sum_{i=1}^n \bar{z}_{1i}, \quad \bar{z}_{1i} = (M_i/M_o) (\bar{y}_{1i}/p_i),$$

$$P_i = M_i/M_o, \quad M_o = \sum_{i=1}^N M_i$$

$$SE(\hat{Y}_1) = N \sqrt{[\sum_{i=1}^n (\bar{z}_{1i} - \bar{z}_{1n})^2 / n(n-1)]} = N \sqrt{\sum_{i=1}^n (\bar{y}_{1i} - \bar{y}_{1n})^2 / n(n-1)}$$

$$\% SE = \{SE(\hat{Y}_1)/\hat{Y}_1\} 100$$

$$\% \text{ Accuracy} = 100 - \{Y_1 - \hat{Y}_1\} / Y_1 \} 100$$

$$95\% \text{ C. I. for } Y_1 : \hat{Y}_1 \pm z_{.05} SE(\hat{Y}_1)$$

$$\text{Similarly for } \hat{Y}_2, \hat{Y}_3, \hat{Y}_4 \text{ \& } \hat{Y}_5.$$

Where,

N = Number of clusters in the population.

n = Number of clusters in the simple such that sample housholds = 250

$$z_{ij} = (M_i y_{ij}) / (M_o P_i).$$

$$\bar{z}_i = \bar{y}_i.$$

$$z_{ij} = y_{ij}.$$

$$(i = 1, 2, \dots, N,$$

$$j = 1, 2, \dots, M_i)$$

11.3.1 Example :

A sample of 3 clusters (villages of 100 households) is selected with probability proportion to its size with replacement from a population of 6 clusters (villages of 241 households). Data of volume of compact trees are collected. Estimate total volume of compact trees and its reliabilites.

Table 16. A PPS sample of 3 clusters of 241 households with data of y_3

Cl-1	y_3	Cl-2	y_3	Cl-3	y_3
1	1.44	1	4.64	1	5.2
2	10.31	2	5.03	2	11.3
3	1.68	3	4.28	3	19.66
4	4.29	4	2.63	4	5.54
5	6.17	5	3.16	5	3.83
6	0.96	6	12.92	6	12.68
7	1.98	7	8.77	7	4.99
8	16.52	8	3.75	8	2.77
9	12.3	9	1.79	9	9.34
10	1.58	10	3.15	10	4.05
11	8.62	11	7.24	11	2.07
12	5.17	12	1.37	12	6.08
13	6.77	13	3.9 4	13	4.19
14	0.82	14	7.61	14	14.52
15	6.36	15	4.01	15	8.52
16	8.73	16	14.94	16	1.53
17	0.76	17	1.5	17	11.19
18	12.86	18	2.52	18	11.33
19	2.15	19	2.74	19	12.99
20	0.93	20	29.23	20	8.73
21	4.56	21	8.81	21	11.28
22	3.01	22	7.38	22	9.59
23	8.39	23	6. 07	23	4.13
24	15.15	24	6.52	24	0.72
25	2.22	25	3.53	25	9.52
26	0.68	26	10.88	26	26.47

Cl-1	y_3	Cl-2	y_3	Cl-3	y_3
27	6.97	27	4.24	27	10.08
28	7.5	28	22.35	28	6.3
29	8.03	29	5.84	29	16.03
30	1.99	30	7.52	30	11.42
31	5.78			31	14.05
32	2.9			32	3.32
33	8.21			33	18.16
34	8.04			34	13.74
				35	9.51
				36	8.38
Mean	5.70088		6.945333		9.255833

Population size = 6 (Number of Clusters)

Cl-I (AH-47)

Cl-II (HH-34)

Cl-III (HH-30)

Cl-IV (HH-48)

Cl-V (HH-36)

Cl-VI (HH-46)

Sampling size = 3 (Number of Clusters)

Cl-I (HH-34)

Cl-II(HH-30)

Cl-III(HH-36)

$$\bar{y}_3 = 7.30$$

$$s\bar{y}_3 = 1.81$$

$$N = M_1 + M_2 + M_3 + M_4 + M_5 + M_6 = 241$$

$$\therefore \hat{Y}_3 = N\bar{y}_3 = 241 \times 7.30 = 1759$$

$$n = 34 + 30 + 34 = 100$$

$$SE(\hat{Y}_3) = N / \sqrt{n} \bar{s}y_3 = 241 / \sqrt{100} \times 1.81 = 43.62$$

$$\%SE(\hat{Y}_3) = 43.62 / 1759 \times 100 = 2\%$$

$$\%Acuracy = 100 - (Y_3 - \hat{Y}_3) / Y_3 \times 100$$

$$= 100 - (2070 - 1759) / 2070 \times 100 = 85\%$$

95% Confidence Interval for Y_3 :

Lower limit : $1759 - 1.96 \times 43.62 = 1674$

Upper limit : $1759 + 1.96 \times 43.62 = 1844$

11.4 PPS Two stage sampling (PPS of first stage units with replacement and SRS of second stage units without replacement)

A sample of n villages was drawn with PPS of households with replacement from a population of N villages by Lahiri's method and m_i is drawn with SRS without replacement from M_i in i th selected village.

i) Estimate of total volume of trees :

$$\bar{y}_{1i} = 1/m_i \sum_{j=1}^{m_i} y_{1ij}. \quad (\text{unbiased})$$

$$\hat{Y}_{1pps} = 1/n \sum (M_i \bar{y}_{1i}/p_i) = M_0(1/n \sum y_{1i}),$$

$$p_i = M_i/M_0, \quad M_0 = \sum M_i$$

$$v(\hat{Y}_{1pps}) = \sum (M_0 y_{1i} - \hat{Y})^2 / n(n-1) \quad (\text{unbiased})$$

$$SE = \sqrt{v(\hat{Y}_{1pps})}$$

$$\% SE = \{SE(\hat{Y}_1)/\hat{Y}_1\} 100$$

$$\% \text{ Accuracy} = 100 - \{Y_1 - \hat{Y}_1\}/Y_1 \} 100$$

$$95\% \text{ C.I. for } \hat{Y}_1 : Y_1 \pm z_{.05} SE(\hat{Y}_1)$$

Similarly for $\hat{Y}_2, \hat{Y}_3, \hat{Y}_4$ & $\hat{Y}_5$.

11.4.1 Example :

A population of 10 villages with 250 households is formed in an area of Chittagong district in Bangladesh. A sample of 5 villages with 96 households is selected at first stage with probability proportion to size (households) and a sample of 50 households is selected at random at second stage from selected villages at first stage. Data of volume of compact trees are collected from sample households which are given below:

Table 17 : A sample of 50 households at second stage with data of y_3

VILL -4	VILL -7	VILL -8	VILL -9	VILL -10
y_3	y_3	y_3	y_3	y_3
2.7	3.01	0.82	6.97	12.89
7.16	1.58	8.21	1.54	5.2
9.14	6.77	15.15	3.58	1.15
11.19	2.90	8.03	1.89	17.51
5.09	12.86	2.22	3.65	3.17
3.9	8.73	5.17		0.79
7.24	8.39	8.04		3.82
5.84	8.62	0.96		3.49
7.19	6.36	2.15		3.71
9.19				9.14
9.11				5.78
8.31				3.02
5.36				3.14
9.51				

$$N = 10, \quad n = 5$$

Population (N) = 10 : 1 VILL (HH-33), 2 VILL (HH-35), 3 VILL (HH-35)
 (VILLS (10) 4 VILL (HH-26), 5 VILL (HH-25), 6 VILL (HH-26)
 HH (250) 7 VILL (HH-18), 8 VILL (HH-18), 9 VILL (HH-9)
 10 VILL (HH-25)

Stage 1 :	4 VILL (HH-26),	7 VILL (HH-18),	8 VILL (HH-18)
PPS of VILLS (5) HH (96)	9 VILL (HH-9),	10 VILL (HH-25)	
Stage 2 :	4 VILL (HH-14),	7 VILL (HH-9),	8 VILL (HH-9)
SRS of HH-50	9 VILL (HH-5),	10 VILL (HH-13)	

$$\bar{y}_{31} = 100.93/14, \quad \bar{y}_{32} = 59.22/9, \quad \bar{y}_{33} = 50.75/9 \quad \bar{y}_{34} = 17.63/5$$

$$\bar{y}_{35} = 77.81/13 \quad \text{Here, } \Sigma \bar{y}_{3i} = \bar{y}_{31} + \bar{y}_{32} + \bar{y}_{33} + \bar{y}_{34} + \bar{y}_{35} = 28.9396$$

$$\text{Here, } M_0 = 250, M_1 = 14, m_2 = 9, m_3 = 9, m_4 = 5, m_5 = 13 \\ n = 5, N = 10$$

$$\begin{aligned} \hat{Y}_{3PPS} &= M_0 (1/n \Sigma y_{3i}) \\ &= 250 [1/5 \times (28.9396)] \\ &= 1446.98 \end{aligned}$$

$$\begin{aligned} V(\hat{Y}_{3PPS}) &= \Sigma (M_0 \bar{Y}_{3i} - \hat{Y})^2 / n(n-1) \\ &= 1/20 [(250 \times 7.21 - 1446.98)^2 + (250 \times 6.58 - 1446.98)^2 + (250 \times 5.64 - \\ &\quad 1446.98)^2 + 250 \times 3.53 - 1446.98)^2 + 250 \times 5.99 - 1446.98)^2] \\ &= 1/20 [126394.47 + 39211.92 + 1367.52 + 318637.67 + 2552.27] \\ &= 488163.85/20 \\ &= 24408.19 \end{aligned}$$

$$SE(\hat{Y}_{3PPS}) = 156.23$$

$$\%SE = SE(\hat{Y}_{3PPS}) / \hat{Y}_{3PPS} \times 100 = 156.23/1446.98 \times 100 = 10.797\%$$

$$\begin{aligned}\% \text{Accuracy} &= 100 - \{(Y_3 - \hat{Y}_3) / Y_3\} \times 100 \\ &= 100 - \{1549.78 - 1446.398 / 1549.78 \times 100 = 93\%\end{aligned}$$

95% Confidence Interval for Y_3 :

$$\text{Lower limit} : 1446.98 - 1.96 \times 156.23 = 1140.7692$$

$$\text{Upper limit} : 1446.98 + 1.96 \times 156.23 = 1753.1968$$

11.5 PPS Stratified Two Stage Sampling (PPS of first stage units with replacement and SRS of second stage units without replacement)

A sample of villages was drawn with PPS (number of households) in all strata by Lahiri's method at first stage. A sample of households is drawn with SRS without replacement. Since the volume of trees, number of bamboos and palm trees are related with the number of households in the villages, PPS of households will be reasonable.

i) Estimate of total volume of trees :

$$\hat{Y}_1 = \sum_{h=1}^L \hat{Y}_{1h}$$

$$\hat{Y}_{1h} = 1/n_h \sum_{i=1}^{n_h} \hat{Y}_{1hi} / P_{hi}$$

$$\hat{Y}_{1hi} = M_{hi} / m_{hi} \sum_{j=1}^{m_{hi}} y_{1hij} = M_{hi} \bar{y}_{1hi}$$

$$v(\hat{Y}_{1h}) = \sum_{i=1}^{n_h} \{(\hat{Y}_{1hi} / P_{hi}) - \hat{Y}_{1h}\}^2 / n_h (n_h - 1)$$

$$SE(\hat{Y}_1) = \sqrt{\sum_{h=1}^L v(\hat{Y}_{1h})}$$

$$\%SE = \{SE(\hat{Y}_1) / \hat{Y}_1\} 100$$

$$\% \text{Accuracy} = 100 - \{(Y_1 - \hat{Y}_1) / Y_1\} 100$$

$$95\% \text{ C.I. for } Y_1 = \hat{Y}_1 \pm Z_{0.05} SE(\hat{Y}_1)$$

Where,

y_{1hij} = Value of y_1 at j th household in i th PSU (village) in h th stratum.

$\bar{y}_{1hij}$ = Mean of y_1 for i th PSU in h th stratum.

M_{hi} = Number of households in i th PSU in h th stratum in the Population.

m_{hi} = Number of households in i th PSU in h th stratum in the sample.

n_h = Number of PSU (village) in h th stratum in the sample.

P_{hi} = Size of i th PSU in i th stratum selected households /
Number of households in h th stratum in the Population.

Similarly for $\hat{Y}_2, \hat{Y}_3, \hat{Y}_4$ & $\hat{Y}_5$.

11.5.1 Example

A population of 9 villages with 419 households is formed and stratified into 3 strata in an area of Chittagong district in Bangladesh. A sample of 6 villages with 268 households taking two villages in each stratum is selected with probability proportional to its size (households) at first stage. A sample of 50 households is selected at random at second stage from each selected first stage units (villages). Data of volume of compact trees of sample households are collected (given below). Estimate total volume of compact trees and its reliabilities.

Table 18. A sample of stratified 50 households at second stage with data of y_3

S-1		S-2		S-3	
1VILL (HH -12)	2VILL (HH -7)	1VILL (HH -9)	2VILL (HH -6)	1VILL (HH -8)	2VILL (HH -8)
SL y_3	y_3	y_3	y_3	y_3	y_3
1 4.13	3.66	4.54	5.84	8.51	3.42
2 7.81	2.74	5.84	4.01	5.74	6.50
3 8.23	2.70	4.67	3.53	4.99	7.42
4 7.98	2.13	5.36	6.52	9.11	6.43
5 4.58	1.55	4.96	4.64	3.47	7.59
6 4.75	3.90	5.09	4.28	9.48	4.07
7 5.27	3.29	5.97		5.12	4.57
8 9.38		6.57		7.17	5.37
9 3.80		6.64			
10 7.63					
11 9.93					
12 5.49					

- Strata : S-1 S-2 S-3
VILLS (9) : 1 VILL(HH-36). 1 VILL(HH-48), 1 VILL(HH-55)
HH (417) 2 VILL(HH-42). 2 VILL(HH-43), 2 VILL(HH-43)
3 VILL(HH-64) 3 VILL(HH-30) 3 VILL(HH-47)

2. Stage 1 : S-1 S-2 S-3
- PPS of 1 VILL(HH-64), 1 VILL(HH-48), 1 VILL(HH-43)
- 6 VILLS 2 VILL(HH-36), 2 VILL(HH-30), 2 VILL(HH-47)
- HH 268)
-
3. Stage 2 : S-1 S-2 S-3
- SRS of 1 VILL(HH-12), 1 VILL(HH-9), 1 VILL(HH-8)
- HH-50 2 VILL(HH-7), 2 VILL(HH-6), 2 VILL(HH-8)

$$\hat{Y}_{3hi} = M_{hi} / m_{hi} \sum y_{3hi}^{m_{hi}}$$

$$\hat{Y}_{311} = M_{11} / m_{11} \sum y_{311j}^{m_{11}} = 64/12 \times 78.92 = 420.91$$

$$\hat{Y}_{312} = M_{12} / m_{12} \sum y_{312j}^{m_{12}} = 36/7 \times 22.67 = 116.59$$

$$\hat{Y}_{321} = M_{21} / m_{21} \sum y_{321j}^{m_{21}} = 48/9 \times 49.64 = 264.75$$

$$\hat{Y}_{322} = M_{22} / m_{22} \sum y_{322j}^{m_{22}} = 30/28 \times 28.82 = 144.10$$

$$\hat{M}_{331} = M_{31} / m_{31} \sum y_{331j}^{m_{31}} = 43/8 \times 53.59 = 280.05$$

$$\hat{M}_{332} = M_{32} / m_{32} \sum y_{332j}^{m_{32}} = 47/8 \times 45.17 = 265.37$$

$$\hat{Y}_{3h} = 1/n_h \sum Y_{3hi} / P_{hi} [h = 1]$$

$$\begin{aligned} \therefore \hat{Y}_{31} &= 1/n_1 \sum Y_{31i} / P_{1i} \\ &= 1/2 [Y_{311}/P_{11} + Y_{312}/P_{12}] \\ &= 1/2 [420.91/12/64 + 116.59/7/36] \\ &= 1/2 [2244.85 + 599.61] \\ &= 1422.23 \end{aligned}$$

$$\begin{aligned}
\hat{Y}_{32} &= 1/n_2 \sum^{n_2} Y_{32i} / P_{2i} \\
&= 1/2 [Y_{321}/P_{21} + Y_{322}/P_{22}] \\
&= 1/2 [264.75/9/48 + 144.10/6/30] \\
&= 1/2 (1412 + 7205) \\
&= 1066.25
\end{aligned}$$

$$\begin{aligned}
\hat{Y}_{33} &= 1/n_3 \sum^{n_3} Y_{33i} / P_{3i} \\
&= 1/2 [Y_{331}/P_{31} + Y_{332}/P_{32}] \\
&= 1/2 [280.05/8/43 + 265.37/8/47] \\
&= 1/2 [1505.27 + 1559.05] \\
&= 1532.16
\end{aligned}$$

$$\therefore \hat{Y}_3 = \sum \hat{Y}_{3h} = 1422.23 + 1066.25 + 1532.16 = 4020.64$$

$$\begin{aligned}
V(\hat{Y}_{31}) &= \sum (\hat{Y}_{31i} / P_{1i} - Y_{31})^2 / n_h (n_h - 1) \\
&= 1/2 (2-1) [(Y_{311} / P_{11} - Y_{31})^2 + (Y_{312} / P_{12} - Y_{31})^2] \\
&= 1/2 [(2244.85 - 1422.23)^2 + 599.61 - 1422.23]^2 \\
&= 1/2 (67670.66 + 67670.66) \\
&= 33835.83
\end{aligned}$$

$$\begin{aligned} V(\hat{Y}_{32}) &= \frac{1}{2}[119543.06 + 37684.25] \\ &= 78613.66 \end{aligned}$$

$$\begin{aligned} V(\hat{Y}_{33}) &= \frac{1}{2}[723.07 + 723.07] \\ &= 723.07 \end{aligned}$$

$$\begin{aligned} SE(\hat{Y}_3) &= \sqrt{\sum V(\hat{V}_{3h})} \\ &= \sqrt{(338351.83 + 78613.66 + 723.07)} \\ &= \sqrt{417688.56} = 646.29 \end{aligned}$$

$$\%SE = 646.29/4020.64 \times 100 = 16\%$$

$$\begin{aligned} \%Accuracy &= 100 - \{(Y_3 - \hat{Y}_3)/Y_3\} \times 100 = 100 - (7345.25 - 4020.64) / \\ &\quad 7345.25 \times 100 = 55\% \end{aligned}$$

95% Confidence Interval for Y_3 :

$$\text{Lower limit : } 4020.64 - 1.96 \times 646.29 = 2753.91$$

$$\text{Upper limit : } 4020.64 + 1.96 \times 646.29 = 5287.37$$

11.6 PPS Double Sampling (PPS of households with replacement at second phase)

A preliminary sample of size (n^1) households is drawn from N households by SRS and a sub-sample of size (n) households is selected with PPS (area of compact tree garden, number of bamboo clumps and number of trees) by Lahiri's method.

i) Estimation of total number of trees:

$$\hat{Y}_2 = N \bar{y}_{2d(PPS)}$$

$$\bar{y}_{2d(PPS)} = y^1 / nn^1 \sum_{i=1}^n v_i = 1 / nn^1 \sum_{i=1}^n z_i, \quad v_i = y_{2i} / y_{1i}, \quad z_i = y^1 v_i$$

$$SE(\hat{Y}_2) = N \sqrt{\{(1/n^1 - 1/N) s_{y_2}^2 + (n^1 - 1)/nn^1 N s_z^2\}}$$

$$s_{y_2}^2 = 1/(n-1) \sum_{i=1}^n (y_{2i} - \bar{y}_2)^2$$

$$s_z^2 = 1/(n-1) \sum_{i=1}^n (z_i - \bar{z})^2$$

$$\% SE = \{SE(\hat{Y}_2)/\hat{Y}_2\} 100$$

$$\% \text{ Accuracy} = 100 - \{(\hat{Y}_2 - Y_2)/Y_2\} 100$$

$$95\% \text{ C. I. for } Y_2 : \hat{Y}_2 \pm z_{0.05} SE(\hat{Y}_2)$$

Where,

$$P_i = y_{1i} / y^1.$$

$$y_{2i} = \text{Volume of trees in the sub-sample.}$$

$$y_{1i} = \text{Number of trees in the sub-sample.}$$

$$y^1 = \sum_{i=1}^{n^1} y_{1i} = \text{total of } y_{1i} \text{ in the preliminary sample.}$$

ii) **Estimation of total volume of compact trees:**

$$\hat{Y}_3 = N \bar{y}_{3d(PPS)} \quad (\text{unbiased})$$

$$\bar{y}_{3d(PPS)} = x^1 / nn^1 \sum_{i=1}^n v_i = 1 / nn^1 \sum_{i=1}^n z_i, \quad v_i = y_{3i} / x_{2i}, \quad z_i = x^1 v_i$$

$$SE(\hat{Y}_3) = N \sqrt{\{(1/n^1 - 1/N)s_{y_3}^2 + (n^1 - 1)/nn^1 N s_z^2\}}$$

$$s_{y_2}^2 = 1/(n-1) \sum_{i=1}^n (y_{3i} - \bar{y}_2)^2$$

$$s_z^2 = 1/(n-1) \sum_{i=1}^n p_i (z_i - \bar{y}_3)^2$$

$$\% SE = \{SE(\hat{Y}_3) / \hat{Y}_3\} 100$$

$$\% \text{ Accuracy} = 100 - \{(Y_3 - \hat{Y}_3) / Y_3\} 100$$

$$95\% \text{ C.I. for } \hat{Y}_3 : Y_3 \pm z_{0.05} SE(\hat{Y}_3)$$

Where,

$$P_i = x_{2i} / x^1.$$

y_{3i} = Volume of compact trees in the sub-sample.

x_{2i} = Area of compact tree garden in the sub-sample.

$x^1 = \sum_{i=1}^{n^1} x_{2i}$ = total of x_{2i} in the preliminary sample.

iii) **Estimation of total number of bamboos:**

$$\hat{Y}_4 = N \bar{y}_{4d(PPS)} \quad (\text{unbiased})$$

$$\bar{y}_{4d(PPS)} = x^1 / nn^1 \sum_{i=1}^n v_i = 1 / nn^1 \sum_{i=1}^n z_i, \quad v_i = y_{4i} / x_{5i}, \quad z_i = x^1 v_i$$

$$SE(\hat{Y}_4) = N \sqrt{\{(1/n^1 - 1/N)s_{y_4}^2 + (n^1 - 1)/nn^1 N s_z^2\}}$$

$$s_{y_4}^2 = 1/(n^1 - 1) \sum_{i=1}^n (y_{4i} - \bar{y}_4)^2$$

$$s_z^2 = 1/(n^1 - 1) \sum_{i=1}^n p_i (z_i - \bar{y}_4)^2$$

$$\% SE = \{SE(\hat{Y}_4) / \hat{Y}_4\} 100$$

$$\% \text{ Accuracy} = 100 - \{(\hat{Y}_4 - Y_4) / Y_4\} 100$$

$$95\% \text{ C. I. for } Y_4 : \hat{Y}_4 \pm z_{.05} SE(\hat{Y}_4)$$

Where,

$$p_i = x_{5i} / x^1.$$

y_{4i} = Volume of bamboos in the sub-sample.

x_{5i} = Number of bamboo clumps in the sub-sample.

$x^1 = \sum_{i=1}^{n^1} x_{5i}$ = Total of x_{5i} in the preliminary sample.

11.6.1 Example :

A population of 200 households is formed in an area of Chittagong district in Bangladesh. A sample of 100 households is selected from the population at random at first phase. At second phase a sample of 40 households is selected from 100 households with probability proportional to its size (area of compact tree garden) without replacement. Data of area of compact tree garden and volume of compact trees of sample households are collected. Estimate total volume of compact trees with its reliabilities.

Table 19. A PPS sample of 40 households with data of y_3 and x_2

Sl no.	x_{2i}	y_{3i}	$v_i = y_{3i}/x_{2i}$	$z_i = x_1^1 v_i$	p_i	$p_i(z_i - y_3)^2$
1	0.02	8.00	400.00	5200.00	0.001538	41468.42
2	0.05	9.51	190.20	2472.60	0.003846	23358.15
3	0.05	9.51	190.20	2472.60	0.003846	23358.15
4	0.04	7.02	175.50	2281.50	0.003077	15900.79
5	0.02	8.38	419.00	5447.00	0.001538	45508.03
6	0.03	9.38	312.67	4064.67	0.002308	37972.33
7	0.03	9.86	328.67	4272.67	0.002308	41966.35
8	0.03	4.36	145.33	1889.33	0.002308	8165.88
9	0.08	8.73	109.13	1418.63	0.006154	12241.32
10	0.08	8.73	109.13	1418.63	0.006154	12241.32
11	0.05	4.19	83.80	1089.40	0.003846	4495.88
12	0.02	7.26	363.00	4719.00	0.001538	34140.54
13	0.03	11.12	370.67	4818.67	0.002308	53400.69
14	0.12	4.56	38.00	494.00	0.009231	2178.21
15	0.12	4.56	38.00	494.00	0.009231	2178.21
16	0.04	8.52	213.00	2769.00	0.003077	23451.85
17	0.20	6.04	30.20	392.60	0.015385	2272.93
18	0.06	10.48	174.67	2270.67	0.004615	23624.40
19	0.10	12.16	121.60	1580.80	0.007692	19022.90
20	0.10	12.16	121.60	1580.80	0.007692	19022.90
21	0.04	7.04	176.00	2288.00	0.003077	15991.85
22	0.02	10.08	504.00	6552.00	0.001538	65878.35
23	0.02	10.08	504.00	6552.00	0.001538	65878.35
24	0.05	4.01	80.20	1042.60	0.003846	4115.08
25	0.03	7.54	251.33	3267.33	0.002308	24511.74
26	0.03	5.73	191.00	2483.00	0.002308	14133.43
27	0.05	12.59	251.80	3273.40	0.003846	41005.14
28	0.02	5.09	254.50	3308.50	0.001538	16756.59
29	0.03	9.62	320.67	4168.67	0.002308	39944.38
30	0.32	9.93	31.03	403.41	0.024615	3844.04

Sl no.	x_{2i}	y_{3i}	$v_i = y_{3i}/x_{2i}$	$z_i = x^1 v_i$	p_i	$p_i(z_i - \bar{y}_3)^2$
31	0.32	9.93	31.03	403.41	0.024615	3844.04
32	0.32	9.93	31.03	403.41	0.024615	3844.04
33	0.32	9.93	31.03	403.41	0.024615	3844.04
34	0.03	6.24	208.00	2704.00	0.002308	16770.41
35	0.06	6.79	113.17	1471.17	0.004615	9877.77
36	0.05	6.77	135.40	1760.20	0.003846	11805.38
37	0.22	12.30	55.91	726.82	0.016923	8738.55
38	0.22	12.30	55.91	726.82	0.016923	8738.55
39	0.04	4.75	118.75	1543.75	0.003077	7254.84
40	0.02	4.06	203.00	2639.00	0.001538	10647.62
Total			7482.11	97267.43	0.27	823393.45

$$N = 200.$$

$$n1 = 100.$$

$$n = 40$$

$$a) \hat{Y}_3 = N\bar{y}_{3d(PPS)}$$

$$\bar{y}_{3d(PPS)} = x_2 / nn^1 \sum v_i, \quad x^1 = \sum x_2 \text{ of } n^1 = 100 = 13, \quad v_i = y_{3i} / x_{2i}$$

$$= 13/40 \times 100 \times 7482.11$$

$$= 24.3169$$

$$\therefore \hat{Y}_3 = 200 \times 24.3169$$

$$= 4863.3715$$

$$s_z^2 = 21112.65$$

$$\therefore SE(\hat{Y}_3) = N\sqrt{(N-n^1/Nn^1 s^2_{y_3} + (n^1-1)/nn^1 1/Ns^2_z)}$$

$$= 200\sqrt{(200-100)/200 \times 100 \times 6.776 + 99/40 \times 100 1/200 \times 21112.65}$$

$$= 200\sqrt{(.03388 + 99/800000 \times 21112.65)}$$

$$= 200\sqrt{(.03388 + 2.1627)}$$

$$= 325.3663$$

$$s^2_{y_3} = 1/39\{2973.58 - 40 \times (8.23)^2\} = 6.776$$

$$s^2_z = 1/n - 1 \sum P_i (z_i - y_3)^2 = 823393.45/39 = 21112.65$$

$$\%SE(\hat{Y}_3) = 325.3663/4863.3715 = 6\%$$

$$\%Accuracy = 100 - (4682.5 - 4863.3715)/4682.5 \times 100 = 96\%$$

95% Confidence Interval for Y_3 :

$$\begin{aligned} \text{Lower limit: } \hat{Y}_3 - 1.96 \times NSE(\hat{Y}_3) \\ = 4863.3715 - 1.96 \times 325.3663 \\ = 4225.6536 \end{aligned}$$

$$\begin{aligned} \text{Upper limit: } \hat{Y}_3 + 1.96 \times SE(\hat{Y}_3) \\ = 4863.3715 + 1.96 \times 325.3663 \\ = 5501.0894 \end{aligned}$$

CHAPTER XII

DETERMINATION OF SAMPLE SIZE

In planning a sample survey, determination of sample size is primarily important. The sampling intensity (percentage of a sample) required to obtain a desired precision of the estimates depends on the extent of a population to be surveyed. It also depends on the type, size of the sampling units, the sampling design and the variability of the character of the sampling units to be considered. Cost of survey should also be another most important factor to be considered. A optimum sampling design usually provides estimates with prescribed precision or standard error at minimum cost. The cost may be assessed in terms of money, time or any other similar units.

The decision of sample size depends entirely on desired precision and the cost of the survey. The desired precision of the estimate is usually specified as permissible margin of error. The standard error of an estimator depends on sample size, variability of character in a population and sampling design used. The population variability may be obtained through a preliminary survey (or pilot survey) if no earlier information is available. However, a useful preliminary information of the sample size can be obtained on the basis of computations assuming the following sampling designs.

12.1 Simple random sample

Let a sample of n units is selected at random from a population of N units through a preliminary survey (or pilot survey). Now the problem is to determine sample size. The following formulas are provided for a simple random sample in consideration with specified precession and prescribed cost of the survey:

- (a) **Sample Size (n) for Specified Precision** - Let us assume that it is desired to estimate population mean of character y and the desired precision of the estimate is p percent error at $1-\alpha$ probability level.

The value of n in terms of actual margin of error (δ) is given by

$$\begin{aligned} n &= (ts/\delta)^2 / \{ 1 + 1/N(ts/\delta)^2 \} && \text{(Cochran, 1977)} \\ &= (ts/\delta)^2 && \text{for large } N \end{aligned}$$

Where, $\delta = p\bar{y}$

p = percentage margin of error

$\bar{y}$ = sample mean obtained either from pilot survey or from any other sources

s = sample standard deviation obtained either from pilot survey or from any other sources.

t = value of student's t corresponding to 95% probability level at $n-1$ degrees of freedom.

- (b) Sample Size within Prescribed Cost of the Survey - Sometimes it is possible to limit the cost of the survey within sanctioned money and to be satisfied with the precision obtained within available resources.

The value of n in terms of cost is given by $n = (C - C_o)/C_1$

Where,

C = total cost of the survey

C_o = overhead cost including planning, organizing and analysis of data

C_1 = average cost in locating and enumerating a single unit in the population.

12.2 Stratified random sample

Let a population consisting of N units is divided into L -strata of $N_1, N_2, \dots, N_L$ units respectively. A sample of n units is to be selected from the population. Let n_i be sample size in i th stratum ($i=1, 2, \dots, L$) and it has to be decided.

- (a) **Cost Proportional to size of the sample-** In case the cost of enumeration is proportional to number of units enumerated, the optimum values of n_i which minimizes the variance is given as $n_i = n w_i S_i / \sum w_i S_i$

Where $W_i = N_i/N$ = weight in each stratum, S_i = Population standard deviation in i th stratum.

- (b) **Cost varying in strata** - Let us assume that the cost of enumeration is different for different strata. Let average cost of enumerating a single unit in i th stratum be C_i ($i = 1, 2, \dots, L$). If the total cost C is fixed, optimum allocation (Neyman allocation) is obtained by minimizing the variance and it is given by

$$n_i = \{(C - C_o)W_i S_i / \sqrt{C_i}\} / \sum (W_i S_i / \sqrt{C_i})$$

Where, C_o = overhead cost

W_i and S_i are same as in 12.1(b) and 12.2(a)

The sample size required for estimating population mean with fixed variance V_o at minimum cost is given by

$$n = (\sum w_i s_i)^2 / V_o + 1/N \sum w_i s_i, V_o = (\%SE \times \bar{y}_{st})^2$$

$\%SE = 1/2$ of $\%$ margin of error.

(c) **Specified precision using proportional allocation-** Unless the estimates of variability within each stratum is available from preliminary surveys or any other previous surveys, it may not be desirable to use n_i mentioned above under (b). When no such preliminary information is available one may use proportional allocation using stratification.

Proportional allocation of a sample of size n among strata is $n_i = n w_i$
 $= n N_i/N$

To obtain a desired value V_o for variance of the estimate for a stratified sample with proportional allocation, sample size is given by

$$n = (\sum w_i S_i^2 / (V_o + 1/N \sum w_i S_i^2))$$

If variation within each stratum is same i, e $S_i^2 = S^2$

$$n = S^2/(V_o + S^2/N), \text{ as } \sum w_i = 1$$

When W_i and S_i are same as in (a).

12.3 Two stage sample- Let a population consists of N first stage units each with M second stage units. For each of n randomly selected first stage units, m second stage units are chosen at random to form a sample of nm second stage units. The population mean is to be estimated by the sample mean of y with variance of the estimator.

$$V(\bar{y}_{nm}) = (1/n-1/N) S_b^2 + 1/n(1/m - 1/\bar{M})\bar{S}_w^2$$

When S_b^2 and $\bar{S}_w^2$ are given in Two Stage Sampling in Chapter-7

From the variance of sample mean, it is obvious that apart from $\bar{S}_b^2$ and $\bar{S}_w^2$, the precision of the estimate depends on n and m . The cost of the survey also depends on the values of n and m . The problem is to determine n and m so that the population mean is estimated with specified precision at minimum cost.

- (a) **Cost directly proportional to total sample-** If the cost of the survey is proportional to m in the sample, minimum variance for the estimate can be obtained with the value of m being unity. However if V_o is specified for the variance of the estimate the value of n will be determined from the estimate.

$$V_o = (1/n-1/N) S_b^2 + 1/n(1 - 1/M) \bar{S}_w^2$$

- (b) **Allocation for specified precision-** We may denote average cost of locating a first stage unit by C_1 and average cost of enumerating a second stage unit selected in first stage unit by C_2 , total cost is given by $C = C_1n + C_2nm$

The optimum value of n and m for obtaining desired value V_o and minimizing cost of the survey is given by

$$m = \sqrt{C_1/C_2} \bar{S}_w/S_b$$

$$n = (S_b^2 + \bar{S}_w^2/m)/V_o$$

If S_b^2 and $\bar{S}_w^2$ are given, optimum value of n and m may be obtained by taking successively $m = 1, 2, 3, \dots$ in above equation. Substituting the value of n and m in the cost function mentioned above, total cost can be determined and appropriate value of m and n to minimize the cost may be obtained.

12.1.1 Example :

Data of main character Y_2 (Volume of trees) of a simple random sample of size 10 (households) out of 50 households obtained after a pilot survey conducted in an area of Chittagong district in Bangladesh are given below. Determine sample size n for a specified margin of error of 10% and 15%.

Table 20. Data of volume of trees of a simple random sample of 10 households collected through pilot survey of 50 households

i	y_2
1	4.06
2	3.90
3	3.17
4	4.80
5	4.00
6	4.75
7	4.22
8	4.30
9	3.40
10	3.17

$$\bar{y}_2 = 3.986, s_2 = .5764$$

(a) For Specified margin of error of 10%:

$$\begin{aligned}\text{The desired margin of error } \delta &= \bar{p}y_2 \\ &= .10 \times 3.986 = .3986\end{aligned}$$

$$\text{The sample size } n = (ts/\delta)^2 / \{1 + 1/N(ts/\delta)^2\}$$

$$\begin{aligned}n &= \{(2.2622 \times .5764)/.3986\}^2 \cdot 1 + 1/50(2.2622 \times .5764/.3986)^2 \\ &= 10.7013/(1 + .2140) = 9\end{aligned}$$

$$\text{Intercity of sampling} = 9/50 \times 100 = 18\%$$

(b) For Specified Margin of Error 15%:

$$\text{The desired margin of error } \delta = \bar{p}y_2$$

$$\begin{aligned}\text{The sample size } n &= \{(2.2622 \times .5764)/.5979/1 + 1/50 (2.2622 \\ &\quad \times .5764)\} / .5979)^2 \\ &= 4.7561/(1 + .0951) \\ &= 5\end{aligned}$$

$$\text{The intercity of sampling} = 5/50 \times 100 = 10\%$$

12.1.2 Example

Same as 12.1.1 except $n = 10$
and $N = 250$ and prescribed margin of error 10%.

Table 21: Data of volume of trees of a simple random sample of 10 households collected through pilot survey of 250 households

i	y_2
1	1.88
2	1.15
3	1.94
4	1.99
5	1.04
6	1.83
7	1.44
8	1.61
9	1.88
10	1.58

$$\bar{y}_2 = 1.634 \quad s_2 = .3347$$

(a) **For Specified Margin of Error 10% :**

$$\begin{aligned} \text{The desired margin of error } \delta &= \bar{p}y_2 \\ &= .10 \times .634 \\ &= 0.1634 \end{aligned}$$

The sample size $n = (t_{s_2}/\delta)^2$

$$n = (2.0639 \times .3347 / .1634)^2 = 18$$

The intercity of sampling = 7%

12.2.1 Example

For same example of stratified random sampling, determine sample size n for specified precision of 10% and 15% margin of error using proportional allocation.

(a) For specified margin of error 10%:

We have $N = 112$, $N_1 = 40$, $N_2 = 32$, $N_3 = 40$

$$W_1 = 40/112, \quad W_2 = 32/112, \quad W_3 = 40/112$$

$$s_1^2 = (7.24)^2 \quad s_2^2 = (9.56)^2 \quad s_3^2 = (2.55)^2$$

$$V_o = (SE\% \bar{y}_{st})^2 = (5\% \times 8.9446)^2$$

$$= (05 \times 8.9446)^2$$

$$= (.4472)^2 = .2$$

$$n = \sum W_i s_i^2 / (V_o + 1/N \sum W_i s_i^2)$$

$$\sum W_i s_i^2 = 47.1515$$

$$\therefore n = 47.1515 / (.2 + 47.1515/112)$$

$$= 47.1515 / (.2 + .4208)$$

$$= 47.1515 / .6209 = 76$$

$$\therefore n_1 = 76 \times .3571 = 27$$

$$n_2 = 76 \times .2857 = 22$$

$$n_3 = 76 \times .3571 = 27$$

The intensity of sampling $= 76/112 \times 100 = 67\%$

(b) **For specified margin of error 15% :**

$$V_o = (7.5\% \times 8.9446)^2 = .45$$

$$\therefore n = 47.1515 / (.45 + 47.1515/122) = 47.1515 / .8709 = 54$$

$$\therefore \text{The intensity of sampling} = 54/112 \times 100 = 48\%$$

12.3.1 Example

Using example in Two Stage Sample in CHAPTER-VII

(a) **Keeping the margin of error 10% :**

$$n=5, N=10 \quad m=? \quad \bar{M} = 25$$

$$V(\bar{y}_3) = (1/n - 1/N) s_b^2 + 1/n(1/m - 1/\bar{M}) s_w^2$$

$$\begin{aligned} s_b^2 &= 1/n - 1 \sum (\bar{y}_i - \bar{y}_3)^2 \\ &= 3.7419 \end{aligned}$$

$$s_w^2 = 1/n \sum s^2_{wi} = 126.5984$$

$$\begin{aligned} V(\bar{y}_3) &= V_o = (SE\% \times \bar{y}_3)^2 \text{ as margin of error} = 10\% \\ &= (.05 \times 10.57)^2 \quad \therefore SE\% = 5\% \\ &= .2793 \end{aligned}$$

$$s_b^2 = 3.7419$$

$$\begin{aligned} \therefore .2793 &= (1/5 - 1/10 \times 3.7419 + 1/5(1/m - 1/25) \times 126.5984 \\ &= .3742 + (1/m - 1/25) \times 25.3197 \\ &= .3742 + 25.3197/m - 1.0128 \\ &= -.6386 + 25.3197/m \end{aligned}$$

$$\therefore .9179 = 25.3197/m$$

$$\therefore m = 27$$

$$\therefore \text{The intensity of sampling} = 5/10 \times 27/250 \times 100 = 5\%$$

(b) **Keeping the margin of error 15%:**

and $n = 5$

$$\begin{aligned} V_o &= (SE\% \times \bar{y}_3)^2 = (7.5\% \times 10.57)^2 \\ &= .6285 \end{aligned}$$

$$\therefore .6285 = -.0386 + 25.397/m$$

$$\begin{aligned} m &= 25.397/1.2671 \\ &= 20 \end{aligned}$$

$$\therefore \text{The intensity of sampling} = 5/10 \times 20/250 \times 100 = 4\%$$

BIBLIOGRAPHY

1. Research Paper

- Abedin, M.Z. and Quddus, M.A. 1988. Household fuel situation, home gardens and agro-forestry practices at six agro-ecological different locations of Bangladesh. Proceedings of a national workshop held in July 17-19, 1988, at Joydevpur, Bangladesh. Editors: M.Z. Abedin, C.K. Lai and M.O. Ali, 19-53 pp.
- Abedin, M. A., Akter, S., Haque, F. and Alam, S. 1988. Uses of multipurpose trees on the small farms of low rainfall Ganges flood plain, soils of Bangladesh. Proceedings of an international workshop on multipurpose tree for small farm use held in Nov. 2-5, 1987 in Pattya, Thailand. Winrock International for Agricultural Development and the International Development Research Centre of Canada. 31-47 p.
- Islam, S.S. 1977. Forest sample surveys of Bangladesh. A report submitted as a creative component for the master of science degree in statistics. Iowa State University, USA. 24 pp.
- Islam, S.S. 1996. Sampling design for assessing village wood and bamboo resources in Bangladesh – An exploratory study. A Ph.D. Thesis. University of Chittagong, Chittagong, Bangladesh 1996. 240 pp.
- Talukder, M.A.H. 1987. Sampling Methods of Estimation for 'Two Common Design of Multi-State Sampling with Unequal Clusters (unpublished).

2. Project Report

- Aleem, S.A. 1981. Project report on tree volume tariff and bamboo weight tariff for village forest inventory of Bangladesh. UNDP/FAO Project BGD/78/020. 55 pp.
- Byron, R.N. 1981. Future consumption of wood and wood products in Bangladesh. Field document No. 4. UNDP/FAO/BGD/78/010. "Supply and demand of forests and future development strategies. 57 pp.

- Byron, R.N. 1984. Status, trends and proposed directions of Bangladesh forestry. The paper is associated with the project UNDP/FAO/BGD/78/010. "Supply and demand of forest products and future development strategies. It was presented at seminar on forestry extension at Kaptai on 31 October, 1984. 20 pp.
- Davidson, J. 1984. Research in the forest management branch of the Bangladesh Forest Research Institute. Field document No. 4. Volume No. 1, UNDP/FAO/ BGD/79/017.
- Obaidullah, M. 1980. Consultant's report on sampling design and statistical analysis consideration for national village forest inventory of Bangladesh. UNDP/FAO/BGD/78/020. 53 pp.
- Ohlsson, B.O. 1984. Tree husbandry and rural forestry. Field document No. 8. UNDP/FAO/BGD/78/010. "Supply and demand for forest products and future development strategies. 43 pp.

3. Book

- Chacko, V.J. 1965. A manual on sampling techniques for forest surveys. Forest Research Institute, Dehra Dun, India. 167 pp.
- Cochran, W.G. 1977. Sampling Technique. Third edition. John Wiley & Sons, Inc. New York, U.S.A. 428 pp.
- Jerram, M.R.K. and Bourne, R. 1949. Elementary forest mensuration. Thomas Murby and Company, 40 Museum Street, W.C.I., London, P. 10.
- Lahiri, D.B. 1951. A method of sample selection providing unbiased ration estimates. Bull. Inter. Stat. Inst., 33(2), 133-140.
- Sukhatme, P.V. and Sukhatme B.V. 1970. Sampling Theory of surveys with applications. Second revised edition. Iowa State University, Ames, Iowa, U.S.A. 451 pp.

APPENDICES

Appendix – 1 : The field sheet

1. General Information :

- a) Name of the head of the household :
- b) Father's/ Husband's name :
- c) Village :
- d) Union :
- e) Thana :
- f) House Type : (Pucca/Semi-pucca/Tinned(wood)/Tinned(bamboo)/ Tinned(mud) / Kancha (mud)/ Kancha (Bera).

2. Economic status :

- a) Family member :

Male		Female	
0-15 yrs	15-above yrs	0-15 yrs	15-above yrs

- b) Annual income : Tk. Main Sorce :

3. Particulars of Area :

Sl. No	Item	Decimal
1.	Cultivated area. (permanent and temporary crops)	
2.	Homestead and other structural areas.	
3.	Area under ponds, ditches, etc. and, permanent fallow land.	

Signature of the
Investigator

4. a) Compact area :

b) Compact trees :

Sl. No.	Species	Diameter class (cm)							Planted in last 12 months (Nos)
		10-20	20-30	30-40	40-50	50-60	60-70	70-....	
01.	Am								
02.	Jam								
03.	Kathal								
04.	Mandar								
05.	Kala koroi								
06.	Shil koroi								
07.	Sissoo								
08.	Raintree								
09.	Others (specify)								

c) Compact plam trees :

Sl. No.	Species	Height class (m)			Planted in last 12 months
		5-10	10-15	15-....	
01.	Narikel				
02.	Tal				

Signature of the Investigator

5. a) Scattered trees :

Sl. No.	Species	Diameter class (cm)							Planted in last 12 months (Nos)
		10-20	20-30	30-40	40-50	50-60	60-70	70-....	
01.	Am								
02.	Jam								
03.	Kathal								
04.	Mandar								
05.	Kala koroj								
06.	Shil koroj								
07.	Sissoo								
08.	Raintree								
09.	Others (specify)								

b) Scattered plam trees :

Sl. No.	Species	Height class (m)			Planted in last 12 months
		5-10	10-15	15-....	
01.	Narikel				
02.	Tal				

Signature of the Investigator

6. a) Felled and cyclone damaged trees :

Sl. No.	Species	Base diameter (cm) class					
		Felled in last 12 months			Cyclone damaged (1991)		
		30-40	40-50	50-....	30-40	40-50	50-....
01.	Am						
02.	Jam						
03.	Kathal						
04.	Mandar						
05.	Kala koroi						
06.	Shil koroi						
07.	Sissoo						
08.	Raintree						
09.	Others (specify)						

b) Felled and cyclone damaged plam trees :

Sl. No.	Species	Height class (m)				
		Felled in last 12 months		Cyclone damaged (1991)		
		15-20	20-....	5-10	10-15	15-....
01.	Narikel					
02.	Tal					

Signature of the
Investigator

7. a) Bamboo area :

b) Bamboo clumps and bamboos :

Sl. No.	Species	Clumps serial	No. of bamboos	No. of bamboos cut in last 12 months	No. of rhizomes planted in last 12 months
01.	Baijjya				
02.	Mal				
03.	Mita/Nitai/Mingga				
04.	Others (Specify)				

Signature of the
Supervisor

Signature of the
Investigator

Appendix – 2 : The population of 50 households

H.H	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	1.79	0.04	0.12	1	8	1.88	0.79	1000	65
2.	0.44	0.03	0.09	1	14	5.55	5.55	5	3
3.	0.65	0.04	0.12	0	10	4.06	4.06	0	9
4.	0.93	0.03	0.08	2	12	6.62	5.46	140	2
5.	0.49	0.04	0.04	0	5	2.41	2.41	0	4
6.	0.34	0.04	0.06	0	11	3.90	3.90	0	2
7.	2.60	0.12	0.44	3	43	34.91	23.57	195	14
8.	0.97	0.05	0.11	1	15	6.10	5.15	12	15
9.	0.93	0.02	0.10	0	19	9.42	9.14	0	15
10.	0.93	0.02	0.10	2	9	3.17	3.17	33	8
11.	0.46	0.07	0.07	2	38	17.51	17.51	120	50
12.	0.17	0.05	0.05	0	4	1.15	1.15	0	10
13.	1.12	0.02	0.12	1	7	1.94	1.94	50	6
14.	0.05	0.01	0.04	1	9	2.49	2.49	32	0
15.	0.06	0.02	0.04	1	10	4.80	4.15	14	4
16.	0.59	0.03	0.07	0	7	5.54	5.54	0	0
17.	0.21	0.01	0.03	0	5	4.00	3.29	0	0
18.	2.18	0.04	0.10	2	10	4.75	4.75	1546	39
19.	0.19	0.03	0.02	0	15	6.08	5.67	0	9
20.	0.58	0.03	0.06	0	22	10.05	8.28	0	4
21.	0.11	0.02	0.04	0	9	5.20	5.20	0	1
22.	0.73	0.02	0.06	0	9	4.22	4.22	0	9
23.	1.19	0.08	0.12	2	16	7.11	3.21	436	37
24.	0.07	0.02	0.02	0	15	5.11	4.09	0	10
25.	0.98	0.10	0.06	2	62	34.94	31.95	85	30
26.	0.62	0.20	0.02	0	30	11.43	11.43	0	7
27.	5.68	0.45	0.05	0	29	17.89	17.89	0	12
28.	1.28	0.12	0.34	2	25	14.79	14.79	80	10
29.	0.47	0.01	0.06	0	15	14.34	1.54	0	14
30.	0.23	0.02	0.04	0	12	4.30	3.93	0	3
31.	0.58	0.03	0.07	0	20	11.24	11.24	0	2
32.	0.53	0.04	0.08	0	19	9.31	8.73	0	2
33.	0.06	0.03	0.02	0	9	3.09	3.09	0	5
34.	0.93	0.04	0.08	0	17	7.00	6.21	0	3
35.	0.11	0.07	0.04	2	19	7.18	6.05	54	1
36.	0.06	0.02	0.03	1	20	12.63	12.63	25	8
37.	1.16	0.32	0.19	2	16	10.70	9.93	160	14
38.	0.45	0.03	0.09	0	27	15.11	15.14	0	7
39.	0.06	0.05	0.01	0	8	2.81	2.81	0	1
40.	2.65	0.06	0.16	0	22	19.00	19.00	0	6
41.	0.35	0.02	0.01	0	9	3.49	2.64	0	1
42.	0.36	0.02	0.02	0	18	8.15	8.15	0	3
43.	0.49	0.03	0.07	1	24	7.74	7.74	8	20
44.	0.12	0.04	0.07	2	24	11.49	11.49	60	2
45.	0.91	0.04	0.06	1	17	5.91	5.19	18	2
46.	0.03	0.01	0.01	1	8	1.99	1.71	8	2
47.	0.50	0.04	0.13	1	28	12.70	11.98	8	16
48.	0.20	0.06	0.05	1	23	20.71	14.83	4	1
49.	0.95	0.04	0.10	0	19	7.59	6.60	0	9
50.	2.19	0.08	0.17	0	15	11.92	11.92	0	4

Note : HH = Household

Appendix - 3 : The population of 112 households stratified into 3 strata

S ₁ H.H	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	1.27	0.02	0.04	0	12	6.31	6.31	0	0
2.	1.59	0.03	0.10	0	14	6.05	6.05	0	7
3.	0.36	0.04	0.07	2	45	32.91	32.91	13	3
4.	1.48	0.14	0.13	0	54	30.09	30.09	0	3
5.	1.23	0.00	0.01	0	5	1.31	1.31	0	2
6.	0.17	0.05	0.11	2	16	12.50	12.50	53	0
7.	0.50	0.04	0.03	1	14	5.24	5.24	5	6
8.	0.34	0.04	0.06	0	13	8.07	8.07	0	2
9.	1.45	0.06	0.17	1	0	2.02	2.02	154	0
10.	1.61	0.14	0.24	4	25	16.83	15.71	172	3
11.	0.38	0.07	0.11	0	31	21.46	21.46	0	3
12.	0.49	0.11	0.05	0	19	9.25	7.63	0	1
13.	0.75	0.06	0.05	0	8	2.24	2.24	0	5
14.	0.34	0.08	0.14	0	30	22.67	16.53	0	10
15.	0.03	0.01	0.01	1	8	1.99	1.71	8	2
16.	0.04	0.02	0.02	0	8	2.09	2.09	0	3
17.	0.45	0.02	0.04	0	12	6.26	6.26	0	3
18.	1.16	0.32	0.18	1	34	25.54	18.30	60	23
19.	0.14	0.02	0.03	0	11	3.31	1.26	0	1
20.	0.09	0.03	0.05	1	8	3.87	3.87	100	1
21.	0.32	0.06	0.03	1	21	9.68	6.11	3	1
22.	2.56	0.05	0.03	0	36	12.04	12.04	0	12
23.	0.31	0.03	0.04	0	5	4.55	4.55	0	2
24.	0.28	0.05	0.07	1	23	17.25	11.61	140	9
25.	0.12	0.00	0.02	0	14	3.72	3.72	0	1
26.	0.04	0.01	0.03	0	6	2.86	1.36	0	1
27.	0.29	0.06	0.03	1	11	4.44	2.52	10	4
28.	0.28	0.02	0.03	0	3	1.26	0.75	0	1
29.	0.70	0.04	0.06	0	25	13.74	13.36	0	6
30.	0.34	0.02	0.03	0	9	3.61	3.61	0	1
31.	0.21	0.01	0.03	0	5	4.00	3.29	0	0
32.	1.50	0.16	0.53	1	14	6.45	6.45	60	5
33.	0.04	0.02	0.02	0	12	6.36	6.36	0	0
34.	0.32	0.01	0.02	0	3	0.76	0.76	0	1
35.	0.05	0.02	0.03	0	2	0.55	0.55	0	2
36.	0.11	0.04	0.07	1	11	3.49	2.94	25	7
37.	0.75	0.10	0.15	3	30	16.08	15.60	57	6
38.	0.55	0.08	0.15	1	53	21.31	21.31	25	19
39.	0.04	0.01	0.02	0	7	4.88	4.13	0	2
40.	0.61	0.03	0.07	1	15	10.90	10.35	12	4

Note : HH = Household

Appendix-3 Continued

S ₂ H.H	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	0.13	0.06	0.07	0	15	4.38	4.38	0	0
2.	1.32	0.03	0.07	1	13	7.29	7.29	60	10
3.	0.15	0.03	0.04	1	58	34.79	34.79	80	8
4.	1.06	0.12	0.11	1	24	10.34	8.73	102	4
5.	0.02	0.02	0.00	0	7	3.60	3.60	0	1
6.	0.14	0.07	0.02	1	34	17.09	17.09	70	1
7.	0.81	0.13	0.04	1	26	16.27	15.72	25	1
8.	1.47	0.08	0.07	4	18	7.28	5.84	755	8
9.	0.16	0.04	0.04	0	21	8.43	5.97	0	11
10.	0.45	0.13	0.21	3	21	9.97	9.97	390	3
11.	0.08	0.03	0.04	1	15	5.36	3.90	50	0
12.	0.04	0.01	0.02	1	8	2.70	2.70	80	2
13.	0.05	0.02	0.02	1	14	5.12	5.12	200	3
14.	0.22	0.02	0.03	1	15	7.19	7.19	150	7
15.	0.32	0.04	0.05	0	20	9.19	9.19	0	3
16.	0.45	0.16	0.11	3	27	10.11	8.31	700	11
17.	0.08	0.03	0.05	2	14	6.53	6.02	42	1
18.	0.24	0.03	0.09	0	15	9.86	9.86	0	0
19.	0.53	0.08	0.12	0	17	4.67	4.67	0	8
20.	0.09	0.02	0.02	2	11	3.01	3.01	135	3
21.	0.02	0.01	0.01	0	7	3.57	1.50	0	1
22.	0.04	0.02	0.02	0	12	5.25	4.64	0	2
23.	0.11	0.08	0.02	1	11	6.07	6.07	25	1
24.	0.41	0.06	0.02	0	16	7.90	7.38	0	0
25.	0.06	0.04	0.02	0	5	2.52	2.52	0	2
26.	0.42	0.05	0.05	1	15	4.24	4.24	15	6
27.	1.08	0.07	0.04	4	15	6.43	5.03	34	5
28.	0.31	0.03	0.05	0	23	9.62	9.62	0	4
29.	0.09	0.04	0.03	2	17	7.25	7.25	475	9
30.	0.14	0.03	0.06	1	17	7.54	7.54	250	9
31.	0.87	0.10	0.12	0	32	14.06	12.16	0	5
32.	0.33	0.05	0.04	0	18	7.48	7.24	0	2

Note : HH = Household

Appendix-3 Continued

S ₃ H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	0.08	0.05	0.03	0	13	5.80	5.80	0	0
2.	0.56	0.03	0.04	0	9	2.68	2.12	0	4
3.	1.64	0.07	0.11	0	31	29.66	29.66	0	5
4.	0.82	0.06	0.10	2	22	17.59	8.77	260	3
5.	0.71	0.02	0.04	1	33	24.58	15.46	65	4
6.	3.29	0.09	0.28	0	25	13.08	13.08	0	18
7.	1.78	0.08	0.41	0	19	13.19	5.74	0	7
8.	0.31	0.06	0.08	1	21	19.62	13.74	10	11
9.	0.29	0.01	0.04	0	21	20.11	20.11	0	1
10.	0.38	0.02	0.04	0	17	8.16	7.60	0	3
11.	0.20	0.02	0.02	0	26	11.00	11.00	0	3
12.	0.42	0.04	0.05	0	24	11.89	9.67	0	8
13.	0.21	0.02	0.03	0	16	10.62	10.62	0	3
14.	0.14	0.02	0.03	0	18	6.79	6.79	0	2
15.	3.45	0.12	0.09	0	63	31.88	28.98	0	7
16.	1.51	0.10	0.11	1	26	15.27	8.16	800	9
17.	0.05	0.02	0.04	0	12	4.96	4.69	0	11
18.	2.58	0.03	0.11	1	12	7.44	7.17	30	5
19.	0.29	0.02	0.03	0	21	13.21	12.20	0	3
20.	0.37	0.04	0.08	1	38	23.47	21.09	50	4
21.	0.09	0.02	0.07	1	10	5.12	5.12	15	8
22.	0.65	0.03	0.12	2	17	10.02	10.02	255	4
23.	0.69	0.10	0.11	0	13	9.69	9.69	0	3
24.	0.71	0.02	0.04	0	10	5.14	5.14	0	2
25.	0.50	0.04	0.06	0	47	18.89	12.06	0	7
26.	0.34	0.02	0.03	3	9	3.03	1.64	43	14
27.	0.26	0.06	0.04	3	24	17.07	14.00	51	40
28.	0.76	0.04	0.07	1	12	10.30	10.30	125	5
29.	1.40	0.04	0.06	0	10	7.25	7.25	0	3
30.	0.17	0.02	0.03	1	4	2.85	2.85	4	17
31.	0.06	0.02	0.04	1	6	2.58	2.10	58	13
32.	1.74	0.03	0.08	2	18	8.16	3.28	46	15
33.	0.12	0.04	0.07	2	24	11.49	11.49	60	2
34.	0.21	0.02	0.02	0	29	11.47	11.47	0	1
35.	0.73	0.02	0.06	0	9	4.22	4.22	0	9
36.	0.11	0.02	0.02	0	7	3.16	3.16	0	9
37.	0.17	0.02	0.07	1	6	1.65	1.65	4	7
38.	0.17	0.02	0.07	1	10	5.31	5.31	54	4
39.	0.38	0.03	0.03	0	49	30.83	29.16	0	3
40.	1.03	0.03	0.03	0	13	5.65	5.37	0	7

Note : HH = Household

Appendix - 4 : The sample of 3 clusters (villages) with 94 households

Cluster 1: P. Khiaychara

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	0.29	0.01	0.01	1	15	5.23	5.23	20	1
2.	0.34	0.02	0.03	0	9	3.61	3.61	0	1
3.	1.25	0.05	0.22	4	42	20.11	16.46	150	1
4.	0.02	0.01	0.01	0	3	1.35	1.35	0	0
5.	0.70	0.04	0.06	0	1	2.57	2.57	0	6
6.	0.70	0.04	0.06	0	25	13.74	13.36	0	6
7.	0.04	0.02	0.02	0	12	6.36	6.36	0	0
8.	0.04	0.02	0.02	0	13	4.55	2.40	0	0
9.	0.79	0.02	0.04	0	19	10.93	10.93	0	4
10.	0.02	0.00	0.01	0	13	4.74	4.74	0	0
11.	0.12	0.00	0.02	0	14	3.72	3.72	0	1
12.	0.25	0.03	0.10	0	39	26.27	8.98	0	5
13.	0.03	0.00	0.02	0	9	5.57	5.57	0	0
14.	0.24	0.07	0.05	0	40	18.17	13.99	0	0
15.	1.23	0.00	0.01	0	5	1.31	1.31	0	2
16.	0.63	0.02	0.04	0	23	12.04	12.04	0	3
17.	1.29	0.02	0.05	1	35	17.51	14.66	31	1
18.	0.45	0.02	0.02	0	20	10.54	7.38	0	1
19.	1.27	0.02	0.04	0	12	6.31	6.31	0	0
20.	0.07	0.04	0.04	0	53	20.51	13.50	0	0
21.	0.03	0.01	0.02	0	9	5.54	4.44	0	0
22.	2.80	0.16	0.20	5	94	43.80	43.80	350	22
23.	0.49	0.03	0.05	1	42	22.11	22.11	62	2
24.	0.50	0.04	0.03	1	14	5.24	5.24	5	6
25.	1.46	0.08	0.16	2	55	45.15	45.15	20	5
26.	0.36	0.04	0.07	2	45	32.91	32.91	13	3
27.	5.52	0.12	0.53	2	38	50.82	50.82	13	6
28.	1.45	0.06	0.17	1	6	2.02	2.02	154	0
29.	1.56	0.05	0.10	0	21	10.42	10.42	0	3
30.	1.48	0.14	0.13	0	54	30.09	30.09	0	3
31.	4.09	0.12	0.30	2	40	23.64	19.40	612	5
32.	1.57	0.03	0.11	2	25	16.49	15.36	77	3
33.	1.21	0.16	0.23	2	32	45.54	45.54	9	7
34.	0.17	0.05	0.11	2	16	12.50	12.50	53	0
35.	0.34	0.03	0.07	2	19	22.23	22.23	18	0
36.	1.76	0.03	0.11	1	16	10.94	9.87	24	2
37.	0.34	0.04	0.06	0	13	8.07	8.07	0	2
38.	0.63	0.06	0.07	2	18	10.38	10.38	34	1
39.	0.56	0.02	0.05	0	21	12.17	12.17	0	3
40.	0.39	0.06	0.09	0	31	21.47	21.47	0	1
41.	2.09	0.10	0.53	0	34	43.06	39.84	0	4
42.	1.59	0.03	0.10	0	14	6.05	6.05	0	7
43.	0.61	0.03	0.07	1	15	10.90	10.35	12	4

Note : HH = Household

Appendix-4 Continued

Cluster- 4 : Hazarbigha

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1	0.65	0.01	0.03	0	12	7.60	5.38	0	2
2	0.53	0.02	0.02	0	23	15.86	14.41	0	5
3	0.25	0.03	0.02	1	2	1.13	1.13	10	2
4	0.61	0.04	0.04	1	7	3.13	2.86	150	1
5	1.30	0.02	0.04	1	5	3.12	3.12	300	5
6	0.34	0.02	0.03	0	12	6.34	3.58	0	3
7	0.89	0.01	0.08	0	15	11.35	6.97	0	5
8	0.36	0.01	0.07	0	15	10.22	3.65	0	7
9	0.05	0.02	0.01	0	6	5.66	0.93	0	2
10	0.06	0.01	0.02	0	7	5.23	1.89	0	4
11	0.72	0.02	0.06	0	16	26.55	20.16	0	3
12	0.46	0.04	0.02	1	60	15.72	5.31	5	3
13	0.13	0.02	0.03	1	4	0.78	0.00	83	3
14	0.47	0.01	0.06	0	15	14.34	1.54	0	14
15	0.47	0.19	0.03	1	34	24.57	22.13	15	7
16	5.43	0.16	1.38	3	68	48.76	40.38	115	31
17	2.65	0.02	0.03	2	31	19.83	3.94	71	14

Cluster - 6 : Chatari

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1	0.26	0.05	0.05	0	16	6.15	5.26	0	5
2	0.04	0.02	0.02	0	17	8.83	7.66	0	1
3	0.13	0.02	0.06	2	20	11.75	10.08	50	3
4	0.10	0.05	0.05	0	22	8.46	6.72	0	0
5	0.86	0.02	0.03	1	22	9.95	7.98	18	2
6	1.15	0.05	0.12	5	25	12.08	9.83	308	6
7	0.25	0.05	0.07	2	28	18.27	16.94	24	7
8	1.10	0.03	0.09	3	18	12.49	12.49	65	8
9	0.07	0.04	0.03	1	21	12.33	10.15	30	1
10	1.10	0.02	0.10	2	20	10.51	8.53	108	17
11	0.29	0.02	0.03	0	22	8.64	7.50	0	2
12	0.49	0.01	0.06	1	22	7.87	6.47	152	4
13	0.45	0.02	0.02	0	9	2.23	2.23	0	2
14	0.62	0.03	0.09	4	10	3.16	3.16	456	3
15	0.69	0.02	0.02	5	13	4.04	3.08	236	2
16	0.25	0.04	0.04	4	26	15.41	15.41	56	2
17	0.25	0.04	0.04	2	12	5.88	5.88	32	5
18	0.25	0.04	0.04	2	14	6.17	6.17	67	2
19	0.03	0.01	0.01	2	6	2.24	2.24	33	3

Note : HH = Household

Appendix-4 Continued

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
20	0.22	0.03	0.02	1	10	3.27	3.27	5	5
21	0.21	0.03	0.02	0	11	6.24	6.24	0	3
22	0.14	0.03	0.03	0	9	3.52	3.52	0	6
23	1.03	0.02	0.03	2	47	34.50	34.50	55	5
24	2.01	0.16	0.22	12	72	52.10	52.10	320	16
25	1.09	0.03	0.08	3	15	9.75	9.75	39	4
26	1.21	0.03	0.19	5	21	10.52	10.52	120	6
27	0.56	0.02	0.05	2	22	10.39	10.39	85	3
28	0.39	0.04	0.02	0	18	11.40	11.40	0	6
29	0.31	0.05	0.09	2	15	7.70	7.70	30	4
30	0.60	0.07	0.05	0	28	23.82	23.82	0	3
31	0.40	0.04	0.03	2	17	6.92	6.40	32	11
32	0.68	0.04	0.14	3	10	5.07	5.07	158	6
33	0.06	0.03	0.02	1	12	5.05	3.05	5	3
34	0.21	0.03	0.02	2	1	1.98	1.98	30	2

Note : HH = Household

Appendix – 5 : The sample of 105 households at second stage in 3 strata

S 1 – Uttar Haitkandi

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	3.91	0.16	0.18	3	9	2.66	2.66	105	4
2.	0.10	0.06	0.03	4	7	3.66	3.66	117	1
3.	0.15	0.05	0.10	0	7	2.23	1.03	0	8
4.	0.10	0.02	0.05	2	6	1.98	0.38	25	0
5.	0.92	0.03	0.07	1	13	5.81	5.06	56	1
6.	1.30	0.16	0.41	1	12	7.52	7.52	72	3
7.	1.50	0.16	0.53	1	14	6.45	6.45	60	5
8.	0.05	0.01	0.04	0	7	2.13	2.13	0	0
9.	2.60	0.12	0.44	3	43	34.91	23.57	195	14
10.	0.29	0.06	0.18	2	9	4.31	4.31	174	3
11.	0.06	0.00	0.02	0	3	0.72	0.72	0	2
12.	0.05	0.01	0.04	0	8	3.50	3.50	0	0
13.	0.21	0.01	0.03	0	5	4.00	3.29	0	0
14.	0.42	0.02	0.01	1	8	2.84	2.84	20	1

Purba Khaiyachara

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	1.48	0.14	0.13	0	54	30.09	30.09	0	3
2.	0.39	0.06	0.09	0	31	21.47	21.47	0	1
3.	2.09	0.10	0.53	0	34	43.06	39.84	0	4
4.	0.04	0.02	0.02	0	12	6.36	6.36	0	0
5.	1.25	0.05	0.22	4	42	20.11	16.46	150	1
6.	1.29	0.02	0.05	1	35	17.51	14.66	31	1
7.	0.24	0.07	0.05	0	40	18.17	13.99	0	0
8.	0.12	0.00	0.02	0	14	3.72	3.72	0	1
9.	1.27	0.02	0.04	0	12	6.31	6.31	0	0
10.	0.03	0.01	0.02	0	9	5.54	4.44	0	0
11.	0.34	0.02	0.03	0	9	3.61	3.61	0	1
12.	1.57	0.03	0.11	2	25	16.49	15.36	77	3
13.	4.09	0.12	0.30	2	40	23.64	19.40	612	5
14.	5.52	0.12	0.53	2	38	50.82	50.82	13	6
15.	1.45	0.06	0.17	1	6	2.02	2.02	154	0
16.	1.56	0.05	0.10	0	21	10.42	10.42	0	3
17.	2.80	0.16	0.20	5	94	43.80	43.80	350	22

Note : HH = Household

Appendix-5 Continued

S₂- Musapur

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	1.46	0.06	0.10	0	28	13.85	11.19	0	9
2.	0.07	0.03	0.04	0	10	3.45	3.45	0	4
3.	0.87	0.10	0.12	0	32	14.06	12.16	0	5
4.	0.35	0.05	0.04	1	22	7.19	7.19	225	13
5.	0.36	0.05	0.05	2	25	9.48	9.48	300	13
6.	0.14	0.03	0.06	1	17	7.54	7.54	250	9
7.	0.09	0.04	0.03	2	24	9.13	9.13	355	7
8.	0.81	0.13	0.04	1	26	16.27	15.72	25	1
9.	0.09	0.04	0.04	1	17	7.04	7.04	85	8
10.	0.32	0.02	0.03	2	14	7.33	7.33	160	6
11.	0.06	0.02	0.04	0	12	4.96	4.96	0	2
12.	0.02	0.01	0.01	0	5	1.58	1.58	0	0
13.	0.45	0.16	0.11	3	27	10.11	8.31	700	11
14.	0.32	0.04	0.03	2	18	10.79	7.16	525	15
15.	0.05	0.02	0.02	1	14	5.12	5.12	200	3
16.	0.22	0.02	0.03	1	15	7.19	7.19	150	7
17.	0.08	0.03	0.04	1	15	5.36	3.90	50	0
18.	0.45	0.13	0.21	3	21	9.97	9.97	390	3
19.	0.06	0.02	0.03	2	12	6.17	6.17	155	3

Rahmatpur

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	0.99	0.06	0.12	0	31	15.44	0.00	0	0
2.	1.31	0.03	0.07	2	45	29.33	29.33	110	5
3.	1.42	0.20	0.24	1	21	0.72	6.04	60	0
4.	0.64	0.08	0.06	4	23	8.15	7.02	465	7
5.	0.11	0.04	0.05	2	13	6.83	6.83	40	2
6.	0.25	0.06	0.05	2	22	8.50	8.50	200	8
7.	0.16	0.13	0.03	1	41	20.51	20.51	48	4
8.	8.43	0.41	0.24	3	36	22.92	22.36	280	23
9.	0.09	0.02	0.02	2	11	3.01	3.01	135	3
10.	0.85	0.26	0.12	3	40	16.53	16.53	265	16
11.	0.10	0.02	0.04	6	53	24.38	24.38	395	4
12.	0.05	0.02	0.03	0	13	6.98	6.70	0	1
13.	0.02	0.02	0.00	0	7	3.60	3.60	0	1
14.	1.32	0.03	0.07	1	13	7.29	7.29	60	10
15.	0.16	0.04	0.04	1	23	10.42	8.41	125	8
16.	0.99	0.26	0.14	3	32	12.84	12.84	265	9

Note: HH = Household

Appendix-5 Continued

S₃-Pachim Forhadabad

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	1.78	0.08	0.41	0	19	13.19	5.74	0	7
2.	0.38	0.03	0.02	0	12	4.54	4.54	0	3
3.	0.05	0.02	0.03	1	4	1.11	0.55	22	2
4.	1.78	0.08	0.41	0	21	13.97	6.52	0	7
5.	0.65	0.03	0.12	2	17	10.02	10.02	255	4
6.	0.64	0.03	0.12	1	14	6.14	6.14	10	3
7.	0.48	0.04	0.08	0	21	12.57	12.57	0	8
8.	1.22	0.06	0.02	1	14	16.22	15.37	40	2
9.	0.10	0.05	0.05	0	13	8.51	8.51	0	3
10.	0.13	0.02	0.02	0	25	8.66	8.66	0	3
11.	0.10	0.05	0.05	0	34	22.36	20.98	0	2
12.	0.06	0.02	0.04	0	10	6.19	6.19	0	0
13.	0.11	0.03	0.08	0	16	9.48	9.48	0	4
14.	0.12	0.02	0.05	0	5	1.84	1.84	0	8
15.	0.04	0.01	0.03	0	10	4.99	4.99	0	2
16.	0.29	0.02	0.03	0	21	13.21	12.20	0	3
17.	0.05	0.02	0.02	0	6	2.69	2.69	0	2

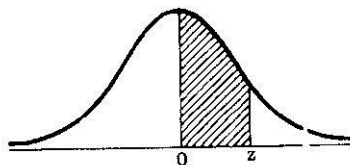
Enayetpur

H.H.	X ₁	X ₂	X ₄	X ₅	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅
1.	0.26	0.06	0.04	3	24	17.07	14.00	51	40
2.	0.73	0.02	0.06	0	9	4.22	4.22	0	9
3.	0.36	0.04	0.07	4	14	4.94	4.15	77	11
4.	0.34	0.02	0.03	3	9	3.03	1.64	43	14
5.	0.42	0.04	0.05	0	24	11.89	9.67	0	8
6.	0.06	0.02	0.04	1	6	2.58	2.10	58	13
7.	0.21	0.02	0.03	0	16	10.62	10.62	0	3
8.	0.07	0.01	0.02	0	11	3.86	3.86	0	2
9.	0.04	0.01	0.02	0	7	1.79	1.79	0	2
10.	0.72	0.12	0.12	0	45	28.72	27.84	0	12
11.	0.69	0.10	0.11	0	13	9.69	9.69	0	3
12.	2.12	0.32	0.65	4	45	38.16	29.67	200	13
13.	0.69	0.10	0.11	0	21	15.81	15.81	0	3
14.	0.44	0.03	0.09	1	14	5.55	5.55	5	3
15.	0.05	0.02	0.04	0	12	4.96	4.69	0	11
16.	0.11	0.02	0.02	0	7	3.16	3.16	0	9
17.	0.04	0.02	0.02	0	9	2.56	2.56	0	2
18.	0.21	0.02	0.02	0	29	11.47	11.47	0	1
19.	0.86	0.16	0.21	1	35	12.88	10.51	150	6
20.	0.34	0.04	0.02	0	49	22.25	22.25	0	1
21.	0.50	0.04	0.06	0	47	18.89	12.06	0	7
22.	0.06	0.04	0.02	0	24	10.35	10.35	0	17

Note : HH = Household

Appendix – 6 : Areas under the normal probability curve

A denotes the area between the line of symmetry (i.e. $z = 0$) and the given z -value.



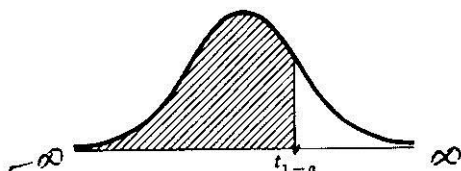
z	A	z	A	z	A	z	A
0.00	0.0000	0.47	0.1808	0.94	0.3264	1.41	0.4207
.01	.0040	.48	.1844	.95	.3289	1.42	.4222
.02	.0080	.49	.1879	.96	.3315	1.43	.4236
.03	.0120	.50	.1915	.97	.3340	1.44	.4251
.04	.0160	.51	.1950	.98	.3365	1.45	.4265
.05	.0199	.52	.1985	.99	.3389	1.46	.4279
.06	.0239	.53	.2019	1.00	.3413	1.47	.4292
.07	.0279	.54	.2054	1.01	.3438	1.48	.4306
.08	.0319	.55	.2088	1.02	.3461	1.49	.4319
.09	.0359	.56	.2123	1.03	.3485	1.50	.4332
.10	.0398	.57	.2157	1.04	.3508	1.51	.4345
.11	.0438	.58	.2190	1.05	.3531	1.52	.4357
.12	.0478	.59	.2224	1.06	.3554	1.53	.4370
.13	.0517	.60	.2258	1.07	.3577	1.54	.4382
.14	.0557	.61	.2291	1.08	.3599	1.55	.4394
.15	.0596	.62	.2324	1.09	.3621	1.56	.4406
.16	.0636	.63	.2357	1.10	.3643	1.57	.4418
.17	.0675	.64	.2389	1.11	.3665	1.58	.4430
.18	.0714	.65	.2422	1.12	.3686	1.59	.4441
.19	.0754	.66	.2454	1.13	.3708	1.60	.4452
.20	.0793	.67	.2486	1.14	.3729	1.61	.4463
.21	.0832	.68	.2518	1.15	.3749	1.62	.4474
.22	.0871	.69	.2549	1.16	.3770	1.63	.4485
.23	.0910	.70	.2580	1.17	.3790	1.64	.4495
.24	.0948	.71	.2612	1.18	.3810	1.65	.4505
.25	.0987	.72	.2642	1.19	.3830	1.66	.4515
.26	.1026	.73	.2673	1.20	.3849	1.67	.4525
.27	.1064	.74	.2704	1.21	.3869	1.68	.4535
.28	.1103	.75	.2734	1.22	.3888	1.69	.4545
.29	.1141	.76	.2764	1.23	.3907	1.70	.4554
.30	.1179	.77	.2794	1.24	.3925	1.71	.4564
.31	.1217	.78	.2823	1.25	.3944	1.72	.4573
.32	.1255	.79	.2852	1.26	.3962	1.73	.4582
.33	.1293	.80	.2881	1.27	.3980	1.74	.4591
.34	.1331	.81	.2910	1.28	.3997	1.75	.4599
.35	.1368	.82	.2939	1.29	.4015	1.76	.4608
.36	.1406	.83	.2967	1.30	.4032	1.77	.4616
.37	.1443	.84	.2996	1.31	.4049	1.78	.4625
.38	.1480	.85	.3023	1.32	.4066	1.79	.4633
.39	.1517	.86	.3051	1.33	.4082	1.80	.4641
.40	.1554	.87	.3079	1.34	.4099	1.81	.4649
.41	.1591	.88	.3106	1.35	.4115	1.82	.4656
.42	.1628	.89	.3133	1.36	.4131	1.83	.4664
.43	.1664	.90	.3159	1.37	.4147	1.84	.4671
.44	.1700	.91	.3186	1.38	.4162	1.85	.4678
.45	.1736	.92	.3212	1.39	.4177	1.86	.4686
.46	.1772	.93	.3238	1.40	.4192	1.87	.4693

Appendix-6 Continued

z	A	z	A	z	A	z	A
1.88	.04700	2.41	.04920	2.94	.04984	3.47	.04997
1.89	.4706	2.42	.4922	2.95	.4984	3.48	.4998
1.90	.4713	2.43	.4925	2.96	.4985	3.49	.4998
1.91	.4719	2.44	.4927	2.97	.4985	3.50	.4998
1.92	.4726	2.45	.4929	2.98	.4986	3.51	.4998
1.93	.4732	2.46	.4931	2.99	.4986	3.52	.4998
1.94	.4738	2.47	.4932	3.00	.4987	3.53	.4998
1.95	.4744	2.48	.4934	3.01	.4987	3.54	.4998
1.96	.4750	2.49	.4936	3.02	.4987	3.55	.4998
1.97	.4756	2.50	.4938	3.03	.4988	3.56	.4998
1.98	.4762	2.51	.4940	3.04	.4988	3.57	.4998
1.99	.4767	2.52	.4941	3.05	.4989	3.58	.4998
2.00	.4773	2.53	.4943	3.06	.4989	3.59	.4998
2.01	.4778	2.54	.4945	3.07	.4989	3.60	.4998
2.02	.4783	2.55	.4946	3.08	.4990	3.61	.4999
2.03	.4788	2.56	.4948	3.09	.4990	3.62	.4999
2.04	.4793	2.57	.4949	3.10	.4990	3.63	.4999
2.05	.4798	2.58	.4951	3.11	.4991	3.64	.4999
2.06	.4803	2.59	.4952	3.12	.4991	3.65	.4999
2.07	.4808	2.60	.4953	3.13	.4991	3.66	.4999
2.08	.4812	2.61	.4955	3.14	.4992	3.67	.4999
2.09	.4817	2.62	.4956	3.15	.4992	3.68	.4999
2.10	.4821	2.63	.4957	3.16	.4992	3.69	.4999
2.11	.4826	2.64	.4959	3.17	.4992	3.70	.4999
2.12	.4830	2.65	.4960	3.18	.4993	3.71	.4999
2.13	.4834	2.66	.4961	3.19	.4993	3.72	.4999
2.14	.4838	2.67	.4962	3.20	.4993	3.73	.4999
2.15	.4842	2.68	.4963	3.21	.4993	3.74	.4999
2.16	.4846	2.69	.4964	3.22	.4994	3.75	.4999
2.17	.4850	2.70	.4965	3.23	.4994	3.76	.4999
2.18	.4854	2.71	.4966	3.24	.4994	3.77	.4999
2.19	.4857	2.72	.4967	3.25	.4994	3.78	.4999
2.20	.4861	2.73	.4968	3.26	.4994	3.79	.4999
2.21	.4865	2.74	.4969	3.27	.4995	3.80	.4999
2.22	.4868	2.75	.4970	3.28	.4995	3.81	.4999
2.23	.4871	2.76	.4971	3.29	.4995	3.82	.4999
2.24	.4875	2.77	.4972	3.30	.4995	3.83	.4999
2.25	.4878	2.78	.4973	3.31	.4995	3.84	.4999
2.26	.4881	2.79	.4974	3.32	.4996	3.85	.4999
2.27	.4884	2.80	.4974	3.33	.4996	3.86	.4999
2.28	.4887	2.81	.4975	3.34	.4996	3.87	.5000
2.29	.4890	2.82	.4976	3.35	.4996	3.88	.5000
2.30	.4893	2.83	.4977	3.36	.4996	3.89	.5000
2.31	.4896	2.84	.4977	3.37	.4996		
2.32	.4898	2.85	.4978	3.38	.4996		
2.33	.4901	2.86	.4979	3.39	.4997		
2.34	.4904	2.87	.4980	3.40	.4997		
2.35	.4906	2.88	.4980	3.41	.4997		
2.36	.4909	2.89	.4981	3.42	.4997		
2.37	.4911	2.90	.4981	3.43	.4997		
2.38	.4910	2.91	.4982	3.44	.4997		
2.39	.4916	2.92	.4983	3.45	.4997		
2.40	.4918	2.93	.4983	3.46	.4997		

* Reproduced with permission from *Introduction to Probability and Statistics*, Fourth edition, by Henry L. Alder and Edward B. Roessler. W. H. Freeman and Company. Copyright © 1968.

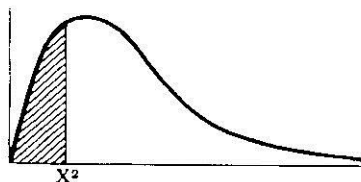
Appendix – 7 : Percentage points of the t – distribution



d. f.	$t_{.80}$	$t_{.70}$	$t_{.60}$	$t_{.50}$	$t_{.40}$	$t_{.30}$	$t_{.25}$	$t_{.20}$	$t_{.15}$	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$	$t_{.001}$
1	0.3250	0.7270	1.376	3.078	6.3138	12.706	31.821	63.657	636.619						
2	.2885	.6172	1.061	1.886	2.9200	4.3027	6.965	9.9248	31.598						
3	.2766	.5840	.978	1.638	2.3534	3.1825	4.541	5.8409	12.924						
4	.2707	.5692	.941	1.533	2.1318	2.7764	3.747	4.6041	8.610						
5	.2672	.5598	.920	1.476	2.0150	2.5706	3.365	4.0321	6.869						
6	.2648	.5536	.906	1.440	1.9432	2.4469	3.143	3.7074	5.959						
7	.2632	.5493	.896	1.415	1.8946	2.3646	2.998	3.4995	5.408						
8	.2619	.5461	.889	1.397	1.8595	2.3060	2.896	3.3554	5.041						
9	.2610	.5436	.883	1.383	1.8331	2.2622	2.821	3.2498	4.781						
10	.2602	.5416	.879	1.372	1.8125	2.2281	2.764	3.1693	4.587						
11	.2596	.5400	.876	1.363	1.7939	2.2010	2.718	3.1058	4.437						
12	.2590	.5387	.873	1.356	1.7823	2.1788	2.681	3.0545	4.318						
13	.2586	.5375	.870	1.350	1.7709	2.1604	2.650	3.0123	4.221						
14	.2582	.5366	.868	1.345	1.7613	2.1448	2.624	2.9768	4.140						
15	.2579	.5358	.866	1.341	1.7530	2.1315	2.602	2.9467	4.073						
16	.2576	.5351	.865	1.337	1.7459	2.1199	2.583	2.9208	4.015						
17	.2574	.5344	.863	1.333	1.7396	2.1098	2.567	2.8982	3.965						
18	.2571	.5338	.862	1.330	1.7341	2.1009	2.552	2.8784	3.922						
19	.2569	.5333	.861	1.328	1.7291	2.0930	2.539	2.8609	3.883						
20	.2567	.5329	.860	1.325	1.7247	2.0860	2.528	2.8453	3.850						
21	.2566	.5325	.859	1.323	1.7207	2.0796	2.518	2.8314	3.819						
22	.2564	.5321	.858	1.321	1.7171	2.0739	2.508	2.8188	3.792						
23	.2563	.5318	.858	1.319	1.7139	2.0687	2.500	2.9073	3.767						
24	.2562	.5315	.857	1.318	1.7109	2.0639	2.492	2.7969	3.745						
25	.2561	.5312	.856	1.316	1.7081	2.0595	2.485	2.7874	3.725						
26	.2560	.5309	.856	1.315	1.7056	2.0555	2.479	2.7787	3.707						
27	.2559	.5307	.855	1.314	1.7033	2.0518	2.473	2.7707	3.690						
28	.2558	.5304	.855	1.313	1.7011	2.0484	2.467	2.7633	3.674						
29	.2557	.5302	.854	1.311	1.6991	2.0452	2.462	2.7564	3.659						
30	.2556	.5300	.854	1.310	1.6973	2.0423	2.457	2.7500	3.616						
35	.2553	.5292	.8521	1.3062	1.6896	2.0301	2.438	2.7239	3.5919						
40	.2550	.5286	.8507	1.3031	1.6839	2.0211	2.423	2.7045	3.5511						
45	.2549	.5281	.8497	1.3007	1.6794	2.0141	2.412	2.6896	3.5207						
50	.2547	.5278	.8489	1.2987	1.6759	2.0086	2.403	2.6778	3.4965						
60	.2545	.5272	.8477	1.2959	1.6707	2.0003	2.390	2.6603	3.4606						
70	.2543	.5268	.8468	1.2938	1.6669	1.9945	2.381	2.6480	3.4355						
80	.2542	.5265	.8462	1.2922	1.6641	1.9901	2.374	2.6388	3.4169						
90	.2541	.5263	.8457	1.2910	1.6620	1.9867	2.368	2.6316	3.4022						
100	.2540	.5261	.8452	1.2901	1.6602	1.9840	2.364	2.6260	3.3909						
120	.2539	.5258	.8446	1.2887	1.6577	1.9799	2.358	2.6175	3.3736						
140	.2538	.5256	.8442	1.2876	1.6558	1.9771	2.353	2.6114	3.3615						
160	.2538	.5255	.8439	1.2869	1.6545	1.9749	2.350	2.6070	3.3527						
180	.2537	.5253	.8436	1.2863	1.6534	1.9733	2.347	2.6035	3.3456						
200	.2537	.5252	.8434	1.2858	1.6525	1.9719	2.345	2.6006	3.3400						
∞	.2533	.5244	.8416	1.2816	1.6449	1.9600	2.326	2.5758	3.2905						

* Reproduced from *Documenta Geigy Scientific Tables*, 7th edition, by permission of CIBA-GEIGY Limited, Basle, Switzerland.

Appendix - 8 : Percentage Points of the χ^2 -Distribution



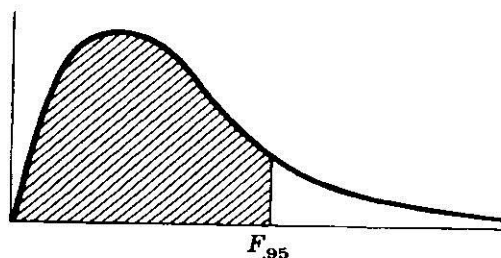
d.f.	$\chi^2_{.005}$	$\chi^2_{.010}$	$\chi^2_{.025}$	$\chi^2_{.050}$	$\chi^2_{.100}$	$\chi^2_{.250}$	$\chi^2_{.500}$
1	392704. 10 ⁻¹⁰	157088. 10 ⁻⁸	982069. 10 ⁻⁶	393214. 10 ⁻⁵	0.0157908	0.1015308	0.454936
2	0.0100251	0.0201007	0.0506356	0.102587	0.210721	0.575364	1.38629
3	0.0717218	0.114832	0.215795	0.351846	0.584374	1.212534	2.36597
4	0.206989	0.297109	0.484419	0.710723	1.063623	1.92256	3.35669
5	0.411742	0.554298	0.831212	1.145476	1.61031	2.67460	4.35146
6	0.675227	0.872090	1.23734	1.63538	2.20413	3.45460	5.34812
7	0.989256	1.239043	1.68987	2.16735	2.83311	4.25485	6.34581
8	1.34441	1.64650	2.17973	2.73264	3.48954	5.07064	7.34412
9	1.73493	2.08790	2.70039	3.32511	4.16816	5.89883	8.34283
10	2.15586	2.55821	3.24697	3.94030	4.86518	6.73720	9.34182
11	2.60322	3.05348	3.81575	4.57481	5.57778	7.58414	10.3410
12	3.07382	3.57057	4.40379	5.22603	6.30380	8.43942	11.3403
13	3.56593	4.10692	5.00875	5.89186	7.04150	9.29097	12.3398
14	4.07467	4.66043	5.62873	6.57963	7.78953	10.1653	13.3393
15	4.60092	5.22935	6.26214	7.26094	8.54676	11.0365	14.3389
16	5.14221	5.81221	6.90766	7.96165	9.31224	11.9122	15.3385
17	5.69722	6.40776	7.56419	8.67176	10.0852	12.7919	16.3382
18	6.26480	7.01491	8.23075	9.39046	10.8649	13.6753	17.3379
19	6.84397	7.63273	8.90652	10.1170	11.6509	14.5620	18.3377
20	7.43384	8.26040	9.59078	10.8508	12.4426	15.4518	19.3374
21	8.03365	8.89720	10.28293	11.5913	13.2396	16.3444	20.3372
22	8.64272	9.54240	10.9823	12.3380	14.0415	17.2396	21.3370
23	9.26043	10.19567	11.6886	13.0905	14.8480	18.1373	22.3369
24	9.88623	10.8564	12.4012	13.8484	15.6587	19.0373	23.3367
25	10.5197	11.5240	13.1197	14.6114	16.4734	19.9393	24.3366
26	11.1602	12.1981	13.8439	15.3792	17.2919	20.8434	25.3365
27	11.8076	12.8785	14.5734	16.1514	18.1139	21.7494	26.3363
28	12.4613	13.5647	15.3079	16.9279	18.9392	22.6572	27.3362
29	13.1211	14.2565	16.0471	17.7084	19.7677	23.5666	28.3361
30	13.7867	14.9535	16.7908	18.4927	20.5992	24.4776	29.3360
40	20.7065	22.1643	24.4330	26.5093	29.0505	33.6603	39.3353
50	27.9907	29.7067	32.3574	34.7643	37.6886	42.9421	49.3349
60	35.5345	37.4849	40.4817	43.1880	46.4589	52.2938	59.3347
70	43.2752	45.4417	48.7576	51.7393	55.3289	61.6983	69.3345
80	51.1719	53.5401	57.1532	60.3915	64.2778	71.1445	79.3343
90	59.1963	61.7541	65.6466	69.1260	73.2911	80.6247	89.3342
100	67.3276	70.0649	74.2219	77.9295	82.3581	90.1332	99.3341
X	-2.5758	-2.3263	-1.9600	-1.6449	-1.2816	-0.6745	0.0000

* Abridged with permission from *Biometrika Tables for Statisticians*, Vol. I, Edited by E. S. Pearson and H. O. Hartley, Cambridge University Press (1962).

Appendix-8 Continued

<i>d.f.</i>	$\chi^2_{.750}$	$\chi^2_{.900}$	$\chi^2_{.950}$	$\chi^2_{.975}$	$\chi^2_{.990}$	$\chi^2_{.995}$	$\chi^2_{.999}$
1	1.32330	2.70554	3.84146	5.02389	6.63490	7.87944	10.828
2	2.77259	4.60517	5.99146	7.37776	9.21034	10.5966	13.816
3	4.10834	6.25139	7.81473	9.34840	11.3449	12.8382	16.266
4	5.38527	7.77944	9.48773	11.1433	13.2767	14.8603	18.467
5	6.62568	9.23636	11.0705	12.8325	15.0863	16.7496	20.515
6	7.84080	10.6446	12.5916	14.4494	16.8119	18.5476	22.458
7	9.03715	12.0170	14.0671	16.0128	18.4753	20.2777	24.322
8	10.2189	13.3616	15.5073	17.5345	20.0902	21.9550	26.125
9	11.3888	14.6837	16.9190	19.0228	21.6660	23.5894	27.877
10	12.5489	15.9872	18.3070	20.4832	23.2093	25.1882	29.588
11	13.7007	17.2750	19.6751	21.9200	24.7250	26.7568	31.264
12	14.8454	18.5493	21.0261	23.3367	26.2170	28.2995	32.909
13	15.9839	19.8119	22.3620	24.7356	27.6882	29.8195	34.528
14	17.1169	21.0641	23.6848	26.1189	29.1412	31.3194	36.123
15	18.2451	22.3071	24.9958	27.4884	30.5779	32.8013	37.697
16	19.3689	23.5418	26.2962	28.8454	31.9999	34.2672	39.252
17	20.4887	24.7690	27.5871	30.1910	33.4087	35.7185	40.790
18	21.6049	25.9894	28.8693	31.5264	34.8053	37.1565	42.312
19	22.7178	27.2036	30.1435	32.8523	36.1909	38.5823	43.820
20	23.8277	28.4120	31.4104	34.1696	37.5662	39.9968	45.315
21	24.9348	29.6151	32.6706	35.4789	38.9322	41.4011	46.797
22	26.0393	30.8133	33.9244	36.7807	40.2894	42.7957	48.268
23	27.1413	32.0069	35.1725	38.0756	41.6384	44.1813	49.728
24	28.2412	33.1962	36.4150	39.3641	42.9798	45.5585	51.179
25	29.3389	34.3816	37.6525	40.6465	44.3141	46.9279	52.618
26	30.4346	35.5632	38.8851	41.9232	45.6417	48.2899	54.052
27	31.5284	36.7412	40.1133	43.1945	46.9629	49.6449	55.476
28	32.6205	37.9159	41.3371	44.4608	48.2782	50.9934	56.892
29	33.7109	39.0875	42.5570	45.7223	49.5879	52.3356	58.301
30	34.7997	40.2560	43.7730	46.9792	50.8922	53.6720	59.703
40	45.6160	51.8051	55.7585	59.3417	63.6907	66.7660	73.402
50	56.3336	63.1671	67.5048	71.4202	76.1539	79.4900	86.661
60	66.9815	74.3970	79.0819	83.2977	88.3794	91.9517	99.607
70	77.5767	85.5270	90.5312	95.0232	100.425	104.215	112.317
80	88.1303	96.5782	101.879	106.629	112.329	116.321	124.839
90	98.6499	107.565	113.145	118.136	124.116	128.299	137.208
100	109.141	118.498	124.342	129.561	135.807	140.169	149.449
X	+0.6745	+1.2816	+1.6449	+1.9600	+2.3263	+2.5758	+3.0902

Appendix - 9 : 95% Significance points of the F distribution*



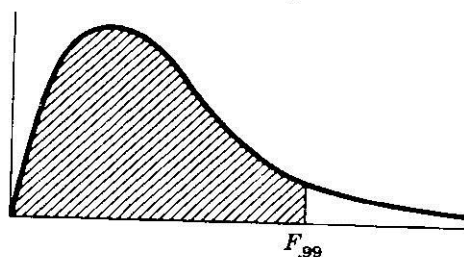
$\nu_1 \backslash \nu_2$	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.7	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.88	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

* Abridged with permission from *Biometrika Table for Statisticians*, Vol. 1. Edited by E. S. Person and H. O. Hartley. Cambridge University Press (1962).

Appendix-9 Continued

$\nu_1 \backslash \nu_2$	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.10	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	6.96	6.91	6.86	6.80	6.77	6.75	6.72	6.69	6.66	6.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.38
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Appendix – 10 : 99% Significance points of the F-Distribution*



$\nu_1 \backslash \nu_2$	1	2	3	4	5	6	7	8	9
1	4052	999.5	5403	5625	5764	5859	5928	5981	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.28	8.65	7.69	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.58	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.88	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

* Abridged with permission from *Biometrika Table for Statisticians*, Vol. 1. Edited by E. S. Person and H. O. Hartley, Cambridge University Press (1962).

Appendix-10 Continued

$\frac{V_1}{V_2}$	10	12	15	20	24	30	40	60	120	∞
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Appendix – 11 : Random numbers

93108	77033	68325	10160	38667	62441	87023	94372	06164	30700
28271	08589	83279	48838	60935	70541	53814	95588	05832	80235
21841	35545	11148	34775	17308	88034	97765	35959	52843	44895
22025	79554	19698	25255	50283	94037	57463	92925	12042	91414
09210	20779	02994	02258	86978	85092	54052	18354	20914	28460
90552	71129	03621	20517	16908	06668	29916	51537	93658	29525
01130	06995	20258	10351	99248	51660	38861	49668	74742	47181
22604	56719	21784	68788	38358	59827	19270	99287	81193	43366
06690	01800	34272	65497	94891	14537	91358	21587	95765	72605
59809	69982	71809	64984	48709	43991	24987	69246	86400	29559
56475	02726	58511	95405	70293	84971	06676	44075	32338	31980
02730	34870	83209	03138	07715	31557	55242	61308	26507	06186
74482	33990	13509	92588	10462	76546	46097	01825	20153	36271
19793	22487	94238	81054	95488	23617	15539	94335	73822	93481
19020	27856	60526	24144	98021	60564	46373	86928	52135	74919
69565	60635	65709	77887	42766	86698	14004	94577	27936	47220
69274	23208	61035	84263	15034	28717	76146	22021	23779	98562
83658	14204	09445	41081	49630	34215	89806	40930	97194	21747
78612	51102	66826	40430	54072	62164	68977	95583	11765	81072
14980	74158	78216	38985	60838	82836	42777	85321	90463	11813
63172	28010	29405	91554	75195	51183	65805	87525	35952	83204
71167	37984	52737	06869	38122	95322	41356	19391	96787	64410
78530	56410	19195	34434	83712	50397	80920	15464	81350	18673
98324	03774	07573	67864	06497	20758	83454	22756	83959	96347
55793	30055	08373	32652	02654	75980	02095	87545	88815	80086
05674	34471	61967	91266	38814	44728	32455	17057	08339	93997
15643	22245	07592	20278	73628	60902	41561	54608	41023	98345
66750	19609	70358	03622	64898	82220	69304	46235	97332	64539
42320	74314	50222	82339	51564	42885	50482	98501	02245	88990
73752	73818	15470	04914	24936	65514	56633	72030	30856	85183
97546	02188	46373	21486	28221	08155	23486	66134	88799	49496
32569	52162	38444	42004	78011	16909	94194	79732	47114	23919
36048	93973	82596	28739	86985	58144	65007	08786	14826	04896
40455	36702	38965	56042	80023	28169	04174	65533	52718	55255
33597	47071	55618	51796	71027	46690	08002	45066	02870	60012
22828	96380	35883	15910	17211	42358	14056	55438	98148	35384
00631	95925	19324	31497	88118	06283	84596	72091	53987	01477
75722	36478	07634	63114	27164	15467	03983	09141	60562	65725
80577	01771	61510	17099	28731	41426	18853	41523	14914	76661
10524	20900	65463	83680	05005	11611	64426	59065	06758	02892
93815	69446	75253	51915	97839	75427	90685	60352	96288	34248
81867	97119	93446	20862	46591	97677	42704	13718	44975	67145
64649	07689	16711	12169	15238	74106	60655	56289	74166	78561
55768	09210	52439	33355	57884	36791	00853	49969	74814	09270
38080	49460	48137	61589	42742	92035	21766	19435	92579	27683
22360	16332	05343	34613	24013	98831	17157	44089	07366	66196
40521	09057	00239	51284	71556	22605	41293	54854	39736	05113
19292	69862	59951	49644	53486	28244	20714	56030	39292	45166
79504	40078	06838	05509	68581	39400	85615	52314	83202	40313
64138	27983	84048	42631	58658	62243	82572	45211	37060	15017

* Abstracted with permission from *A Million Random Digits with 100,000 Normal Deviates* by The Rand Corporation, The Free Press of Glencoe, New York (1955).



Author's Brief Life

- Birth** : Born in a farm family of Village- Khurdo,
Post Office- Patkelghata, Police Station- Patkelghata,
District- Satkhira, Bangladesh.
- Born in August 01, 1947 as a second son of a family
of 7 brothers and 5 sisters.
- Father** : Late Sheikh Estab Ali
- Mother** : Late Maryam Khatun
- Wife** : Dr. Nazma Begum, MBBS,
Retaired Medical Officer
Bangladesh Forest Research Institute
- First Daughter** : Dr. Sadia Islam, MBBS
Canadian Citizen.
- Second Daughter** : Engineer Samia Islam
House Wife
- Third Son** : Sheikh Nazmul Islam
Businessman
- Present Position** : Dr. Sheikh Sirajul Islam
Professor of Statistics
Department of Business Administration
International Islamic University Chittagong
Bangladesh.

Previous Position : Teaching Faculty, BRAC University in 2005
Director, Bangladesh Forest Research Institute,
Chittagong in 1973-2004
Teaching Faculty, Al-Fateh University, Tripoli, Libya
in 1980-86
Teaching Faculty, Institute of Forestry and
Environmental Science in 1988-2002
Chittagong University, Bangladesh.

Research Publications : Published a total of 33 research articles

Overseas Visit : USA, COSTARICA, PANAMA, KSA, LIBYA,
KENYA, INDIA, PAKISTAN, SRI LANKA,
BHUTAN, NEPAL, THAILAND, MALAYSIA,
SINGAPORE, INDONESIA AND DUBAI.

