b107494

A STUDY OF CONTROLLED SAMPLING WITH VARYING PROBABILITIES WITHOUT REPLACEMENT

By

Wimon Sonchaem

A Dissertation Submitted in Partial
Fulfillment of The Requirements for The Degree of
Doctor of Philosophy (Statistics)
School of Applied Statistics
National Institute of Development Administration

ISBN 974-231-180-3
2001

A STUDY OF CONTROLLED SAMPLING WITH VARYING PROBABILITIES WITHOUT REPLACEMENT

Wimon Sonchaem

School of Applied Statistics

The Examining Committee Approved This Dissertation Submitted in Partial Fulfillment of The Requirements for The Degree of Doctor of Philosophy (Statistics)

Professor Prachoom Suwattee .....Committee Chairman
(Prachoom Suwattee, Ph.D.)

Assistant Professor Pachitjanut Siripanich .....Committee
(Pachitjanut Siripanich, Ph.D.)

Assistant Professor .....Committee
(Samruam Chongcharoen, Ph.D.)

Associate Professor .....Committee
(Anumongkol Sirivedhin, Ph.D.)

Associate Professor Jirawan Jitthavech .....Acting Dean
(Jirawan Jitthavech, Ph.D.)

Date Feb 21 ....., 2001.

B 11122473
O 10149168

ABSTRACT

Title of Dissertation	A Study of Controlled Sampling with Varying Probabilities without Replacement
Author	Miss Wimon Sonchaem
Degree	Doctor of Philosophy in Statistics
Year	2000

The objectives of this study are to develop a new controlled sampling scheme with probability proportional to size without replacement, and to present an estimator of the population total and an estimator of the variance of the estimator under the sampling scheme. The computational expressions of the first and second orders inclusion probabilities were derived. Unbiasedness and consistency properties of the proposed estimator of the population total were considered. The estimator of the variance of the estimator is found unbiased.

The proposed sampling scheme is in fact a rejective method under which a sample is drawn until it is accepted. A real-life population was employed to illustrate the sampling scheme. The probability of obtaining a nonpreferred sample under the controlled plan and the uncontrolled plan has been compared for seven natural populations.

The results of this study show that the inclusion probabilities of the first and second orders are easily obtained for this sampling scheme.

The study of this sampling plan can be extended to other sampling designs. In addition, nonnegativeness of variance estimator should also be further investigated.

ACKNOWLEDGEMENTS

My dissertation was completed and finally got approved through my great effort and kind assistance from the following Professors and colleagues to whom I would like to express my thanks and appreciation.

My deep gratitude and appreciation are extended specially to Professor Dr. Prachoom Suwattee, my major thesis advisor for his supervision, guidance and encouragement he kindly and closely gave me throughout my thesis conducting. Without his support, this thesis might have been impossible. I am equally grateful to the committee members, Associate Professor Dr. Anumongkol Sirivedhin, Assistant Professor Dr. Pachitjanut Siripanich and Assistant Professor Dr. Sumruam Chongcharoen, for their constructive comments and helpful advice.

I also feel grateful to Assistant Professor Chamnian Pribwai for his help in correcting my English language. From his correction at my presence, I have gained a wealth of knowledge of the English language.

My heartfelt thanks and appreciation also go to Ms. Pachaporn and Khun Anan Suwannarat, my sister and brother-in-law, and Khun Luepol Punnakanta, my colleague at Mahidol University, for supplying me with all conveniences and making me feel confident.

I also would like to profoundly thank Khun Supawadee Yimsricharoenkit from the Royal Irrigation Department and the officials from the Business Economics Department, for supplying me with some useful data for analysis.

Last but not least, my special thanks are due to Khun Phera Srisongmuang and Khun Vichuda Saenpakdee of the School of Applied Statistics who provided me with the useful computer software and their helps to facilitate my work.

Wimon Sonchaem

February 2001

b107494

TABLE OF CONTENTS

Page

ABSTRACT iii

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS v

LIST OF TABLES vii

CHAPTER

1. INTRODUCTION

1.1 Background 1
1.2 Statement of the Problem 2
1.3 Scope of the Research 4
1.4 Notations 5

2. LITERATURE REVIEW

2.1 Introductory Statement 7
2.2 Methods of Sampling with Varying Probabilities without Replacement 7
2.3 Variance Estimation and Its Comparison 17
2.4 Previous Works on Controlled Sampling 19

3. A PROPOSED SAMPLING SCHEME

3.1 Introduction 21
3.2 The Sampling Scheme 22

4. ESTIMATION OF PARAMETER

4.1 The Horvitz-Thompson Estimator 41
4.2 An Alternative Estimator 42
4.3 Estimation of the Variance of the Estimator 46

Table of Contents (Continued)

CHAPTER		Page
	4.4 An Example	48
5. MORE ILLUSTRATIONS		
	5.1 An Example	51
	5.2 Some More Natural Populations	57
	5.3 Numerical Results	60
6. SUMMARY		
	6.1 Summary and Conclusions	68
	6.2 Research Recommendations	69
BIBLIOGRAPHY		71
APPENDICES		75
Appendix A		76
Appendix B		78
Appendix C	Investment funds of 112 steel production factories	80
Appendix D	The adjusted value of the investment fund of 112 steel production factories	81
Appendix E	Population Number 1 and 2	84
Appendix F	Population Number 3 and 4	90
Appendix G	Population Number 5	98
Appendix H	Population Number 6	102
Appendix I	Population Number 7	107
VITA		111

LIST OF TABLES

Tables		Page
3.1	The Probability of Getting a Nonpreferred Sample Under the Uncontrolled Plan, P(s), and the Controlled Plan, P*(s)	34
4.1	Production of the Factories, Y_{ij}, the Number of Workers, X_{ij} and the Adjusted Value of the Number of the Workers in the Population	48
5.1	Stages in the Sample Selection by the Proposed Sampling Scheme in the Population	52
5.2	Production Total, the 1st Order Inclusion Probability and the Adjusted Value of the Investment Fund for each Element in an Accepted Sample of 20 Factories	54
5.3	The Probability of Getting a Nonpreferred Sample for Possible c's Under the Controlled Plan and the Uncontrolled Plan	56
5.4	Descriptions of Natural Populations Used in Studying the Proposed Sampling Scheme	58
5.5	The Estimates of the Population Total and their Variance Estimates for the Natural Populations	64
5.6	Probability of Obtaining a Nonpreferred Sample for Possible c's Under the Controlled Plan, P*, and the Uncontrolled Plan, P, for the Seven Natural Populations	65

CHAPTER 1

INTRODUCTION

1.1 Background

The basic problems of a sample survey are the getting of nondesirable units and the precisions of estimators of the population parameters. The problem of such nondesirable units acquired as some elements in a sample located in inconvenient locations can be solved by a technique of controlled sampling. This technique is a procedure that minimizes the probabilities of nondesirable units. Goodman (1950) was the first to introduce the technique of controlled sampling. A variety of procedures and theoretical orientations have been used in these controlled sampling studies. More studies concerning controlled sampling have tended to focus on the theoretical frameworks of simple random sampling such as those of Avadhani and Sukhatme in 1965,1968 and 1973. The only advantage of these studies is their simplicity. However, previous investigations suggest that simple random sampling produces low precision of estimators. Therefore, an attempt was made to increase the precision for controlled sampling. Several studies of controlled sampling are rooted in the theoretical frameworks of sampling with varying probability. In sampling with varying probability, the probabilities of selecting the units in the population are unequal. Avadhani and Sukhatme (1967) proposed a controlled sampling scheme but there is some evidence that its estimation suffers a slight loss of precision (Cochran, 1977). Other suggestions are made by Rao and Nigam (1990; 1992). However, their

studies are complicated for large sample size. Although controlled sampling researches have been conducted for several years, their results have not been widely applied to sampling with varying probability. This dissertation will develop a new method of sampling, one of the alternatives to be considered in the planning process of controlled sampling.

1.2 Statement of the Problem

This study is intended to present a method of sampling to acquire a desirable sample for large sample size as applied to varying probabilities without replacement for controlled sampling. In sampling with probability proportional to size without replacement (PPSWOR), it is usual to select the first unit with probability proportional to size, and the second with probability proportional to the size of the remaining units, and similarly for successive units. The usual estimator of the population total for this method of sampling is the Horvitz-Thompson estimator, $\hat{Y}_{HT}$,

$$\hat{Y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\pi_i}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1.1)$$

with its variance estimator

$$v_{HT}\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{n} \frac{(1-\pi_i)}{\pi_i^2} y_i^2 + 2\sum_{i=1}^{n}\sum_{j<i}^{n} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j \pi_{ij}} \right) y_i y_j, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1.2)$$

where π_i is the probability of drawing unit i in the sample, the first order inclusion probability and π_{ij} is the probability of drawing units i and j in the same sample, the second order inclusion probability.

Another its variance estimator expression introduced by Sen (1953); Yates and Grundy (1953) and denoted by $v_{SYG}\left(\hat{Y}_{HT}\right)$, is as follows:

$$v_{SYG}\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{n}\sum_{j<i}^{n}\left(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}}\right)\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots (1.3)$$

The complications of unit selection and the computational problem of π_i, π_{ij} are the limitations of this estimator since the size of the units will be computed in each time of unit selection.

To solve the problems of sample selection and estimation of parameters, some studied the case of small sample size, namely n = 2. Some proposed new sampling schemes and an estimator as in the random grouping (Rao, Hartley and Cochran, 1962) and some proposed alternative estimators (Hansen and Hurwitz, 1943; Das, 1951; Raj, 1956; Murthy, 1957).

The inclusion probability proportional to size (IPPS) design, sometimes called πPS design, is another approach for solving this problem of estimation (Cochran, 1977). The IPPS sampling designs make the probability of inclusion of unit i in the sample, π_i, proportional to its size, X_i, i.e., $\pi_i = n\frac{X_i}{X}$ where $X = \sum_{i=1}^{N} X_i$.

To explicitly introduce this strategy, we will use the following definitions:

Definition 1.1 A preferred unit is a unit that has some required and desirable properties. Otherwise, it is referred to as **a nonpreferred unit.**

Definition 1.2 A preferred sample is a sample containing less than c nonpreferred units, where c is a pre-chosen positively small integer much less than the sample size.

Otherwise, it is **a nonpreferred sample**.

Definition 1.3 Controlled sampling is a sampling scheme that reduces the probability of obtaining a nonpreferred sample.

The objectives of this dissertation are as follows:

1. To develop a controlled sampling scheme with probability proportional to size without replacement so that the inclusion probabilities of the 1st and 2nd orders may be easily obtained.

2. To present an unbiased estimator of the population total as an alternative to the ordinary Horvitz-Thompson estimator.

3. To present an unbiased estimator of the variance of the estimator as compared to the existing estimator of the variance of the Horvitz-Thompson estimator.

1.3 Scope of the Research

In this dissertation a new sampling scheme will be presented so that the probability of obtaining a nonpreferred sample for the controlled plan will not be larger than that for the uncontrolled one. It will be shown also that, for the new sampling scheme, the computations of the inclusion probabilities of first and second orders are easy.

An alternative estimator of the population total as applied to this sampling scheme will be introduced. Unbiasedness and consistency properties of the proposed

estimator of the population total will be investigated. Furthermore, an unbiased estimator of the variance of such estimator will be presented.

In Chapter III, we will suggest the sampling scheme with probability proportional to size. The computational expressions of π_i, π_{ij} will be shown. The reduction of the probability of getting a nonpreferred sample for this sampling scheme will also be illustrated.

In Chapter IV, an unbiased total estimator for this sampling scheme will be introduced. The variance of this estimator and an unbiased estimator of the variance of the estimator of population total will be found.

In Chapter V, as an illustration of the sampling scheme, a sample will be drawn from a natural population. The probability of obtaining a nonpreferred sample under the controlled plan will be compared with the uncontrolled one.

1.4 Notations

Let a population of size N under study be divided into two groups, the first of which is the group of N_1 preferred units and the second is the group of N_2 nonpreferred units, $N_1+N_2 = N$. In application, assume that N_2 is smaller than N_1. Let y be the study variable and x an auxiliary variable highly correlated to y. For the unit j in group i as denoted by $U_{j(i)}$, i = 1,2, j = 1,2,...,N_i, let Y_{ij} and X_{ij} denote the values of y and x, respectively. Assume that $Y_{ij} > 0$ and $X_{ij} > 0$, for all i and j. Furthermore, assume that the values of the auxiliary variable, X_{ij}, are all known.

Let S be the collection of all possible samples s of size n drawn from the population of size N. The values X_{ij} and Y_{ij} for elements in a sample are respectively denoted by lower case letters x_{ij} and y_{ij} ; i = 1,2, j = 1,2,..,n_i and $n_1+n_2 =$ n. For each unit in a sample, the values x_{ij} and y_{ij} are observed values from one of the units of the population, $x_{ij} \in \{X_{i1}, X_{i2}, ..., X_{iN_i}\}$ and $y_{ij} \in \{Y_{i1}, Y_{i2}, ..., Y_{iN_i}\}$, i = 1,2, j = 1,2,...,$n_i$. A sample of size n will be denoted by $s = \{u_{1(1)}, u_{2(1)}, ..., u_{n_1(1)}, u_{1(2)}, ..., u_{n_2(2)}\}$ where $u_{j(i)} \in \{U_{1(i)}, U_{2(i)}, ..., U_{N_i(i)}\}$, i = 1,2, j = 1,2,...,$n_i$.

CHAPTER 2

LITERATURE REVIEW

2.1 Introductory Statement

The objective of this dissertation as earlier stated is to study an estimator of the population total and its variance in a new sampling scheme in controlling sampling with varying probabilities without replacement. To attain the aims of this dissertation it is necessary to know the suitable sampling scheme that does not disturb the fundamental principle of sampling with varying probabilities design, and the technique to reduce selection probabilities of nonpreferred units that induces the minimizing of getting a nonpreferred sample. The solutions of these topics have been studied in many papers that will be reviewed. Then, in this survey the review is divided into three parts: methods of sampling with varying probabilities without replacement, variance estimation and its comparison and previous works on controlled sampling.

2.2 Methods of Sampling with Varying Probabilities without Replacement

The theory of fixed sized sampling design with varying probabilities was developed to solve the problem where sampling units vary considerably in size. The usual method of sampling with varying probabilities without replacement is the probability proportional to size without replacement (PPSWOR) sampling design.

In the PPSWOR sampling scheme, the first unit is selected with probability proportional to size and the next unit is selected with probability proportional to size of the remaining units and so on until the sample size can be reached. It seems impractical since the sizes of the remaining units are recalculated after each unit is selected. Lahiri (1951) developed a simple method for solving the problem of the unit selection. Another alternative convenient method for implementing the PPS sampling scheme is the following. Select the first unit using probability proportional to size by Lahiri's method. Select the next unit from the entire population using probability proportional to size. This procedure is repeated until n distinct units are obtained (Hedayat and Singha, 1991: 134-135). The Horvitz-Thompson estimator and its variance are suitable for the sampling design. The expression for the first order inclusion probability, π_i and the second order inclusion probability, π_{ij} for small sample size, n = 2, are

$$\pi_i = P_i + \sum_{j \neq i} P_j \frac{P_i}{(1 - P_j)}$$

$$\pi_{ij} = P_i \frac{P_j}{(1 - P_i)} + P_j \frac{P_i}{(1 - P_j)}$$

and it is satisfactory for the nonnegative variance estimate property, $0 < \pi_{ij} < \pi_i \pi_j$, $1 \leq i \neq j \leq N$. However, these expressions are quite complicated for $n > 2$.

Grundy (1954) considered a method of drawing sampling unit with probabilities exactly proportional to size. This method is modified from systematic sampling with unequal probabilities and is analogous to the selection method of Lahiri (1951). The units were selected by probabilities proportional to sizes of those not already selected to avoid the problem of periodicity in the population.

There are some papers which use estimators other than the Horvitz-Thompson estimator.

Raj (1956) presented an unequal probability without replacement sampling scheme where the conditional probability of drawing unit i in the population given that the subsample has been drawn was used. The order of the selected units in a sample plays an important role in the expression of the estimator. So it is called an ordered estimator. The variance estimator is also both unbiased and nonnegative.

Murthy (1957) modified the Des Raj estimator in the symmetrized form. All equivalent samples are weighted by the conditional probability of the sample given that it contains the unit i. Since the Des Raj estimator is unbiased, the Murthy's is also unbiased.

Rao, Hartley and Cochran (1962) presented a PPS sampling scheme with the concept of random grouping. A population of size N is divided randomly into n groups where n is a sample size. For each group, one unit is selected with PPS sampling. The estimator of the population total is unbiased and the estimator of variance of the estimator is a nonnegative unbiased variance estimator.

On the other hand, the inclusion probability proportional to size (πPS) sampling scheme is a sampling scheme that the probability of selecting a unit is proportional to its size, $\pi_i = n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$. The reasonable requirements of πPS sampling design are $\pi_{ij} \leq \pi_i \pi_j$ and $\pi_{ij} > 0$ for all $i \neq j$ in order that variance estimator

of the Horvitz-Thompson estimator will be nonnegative (Hedayat and Singha, 1991: 105)

Madow (1949) extended the theory of systematic sampling to the probability proportional to size (PPS) sampling design. The units in a population are numbered in any order and divided into n interval. Then elements in a sample are drawn from these segments of the population. This method is useful for single stage sampling with probability proportional to size without replacement. However the limitations which are the second order inclusion probability, π_{ij}, may be equal to zero and impracticable in case of large populations.

Narain (1951) considered a sampling scheme where all possible samples of 2 elements were assigned probabilities P(S). Then, one sample was drawn from these samples with probability P(S). The assigned probabilities P(S) which are equal to π_{ij}; $i \neq j$ will be assigned in order that $\pi_i = n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$ for each i.

Horvitz and Thompson (1952) presented a general method for the sampling without replacement from a finite population with unequal selection probabilities. They provided an unbiased estimator of the population total as

$$\hat{Y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\pi_i}$$

where y_i is the measurement for the i^{th} unit and π_i is a probability of the i^{th} element which will be included in a sample of size n. The variance of the estimator and the unbiased estimator of its variance are given respectively as

$$V\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{N} \frac{(1-\pi_i)}{\pi_i} Y_i^2 + \sum_{i \neq}^{N} \sum_{j}^{N} \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j} Y_i Y_j$$

and $$v\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{n} \frac{(1-\pi_i)}{\pi_i^2} y_i^2 + \sum_{i \neq}^{n} \sum_{j}^{n} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j \pi_{ij}} \right) y_i y_j$$

Yates and Grundy (1953) used a method of successive approximation that two units of a sample were drawn with probability proportional to revised relative size, Z_i, and without replacement. In such a way, the probabilities proportional to revised relative sizes will keep the inclusion probabilities proportional to the size, $\pi_i = n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$. However, when n > 2, calculating the probability of the i^{th} element included in a sample of size n is complicated. They also gave another unbiased variance estimator of the population total

$$v\left(\hat{Y}_{HT}\right) = \sum_{i}^{n} \sum_{j>i}^{n} \frac{(\pi_i \pi_j - \pi_{ij})}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2 .$$

Hanurav gave a method that each unit is continuously selected until n different elements will be received. It is called the sequential method that keeps the inclusion probability proportional to sizes. The first element is selected with probability proportional to its size P_i then the subsequent units are selected with replacement. The probability proportional to revised size for the k^{th} draw, $Z_i^{(k)}$, will be determined in order that the inclusion probability of unit i within the first k draw, $\pi_i^{(k)}$, is proportional to its size. Another method of Hanurav in 1967 is the grouping method. In this method, a population is divided into n groups with aggregate size and one element is selected with probability proportional to size from each group, then $\pi_i = nP_i$. (Hanurav, 1967: 421-428 quoted in Chaudhuri and Vos, 1988: 194-195)

Fuller (1970) presented a method that is similar to Hanurav's grouping method except that the sample size is fixed and junctional units cannot be drawn more than once. The elements are drawn with conditional probabilities depending on whether the preceding junctional unit was drawn in the previous group or not.

Stevens (1958) developed a regrouping method that is an extension of Durbin's grouping method. The N units in a population are subdivided into k groups on the basis of nondecreasing size of the ordered units. A sample of size n is drawn with probability proportional to the total size of the elements in each group and with replacement. If elements in any group are drawn t times, t units are selected from that group with equal probabilities and without replacement.

Brewer presented a simple method in 1963 that keeps inclusion probabilities proportional to sizes. The first element is drawn with revised probabilities proportional to $P_i(1-P_i)/(1-2P_i)$ and the second unit with probabilities $P_i(1-P_j)$, where j is the unit drawn first. (Brewer, 1963a: 5-13 quoted in Cochran, 1977: 261)

Brewer (1975) extended the method of Brewer for n = 2 in the general sample size. Conditional probabilities proportional to size are used to select the elements, after the first one. For this method, the second order inclusion probabilities are very difficult to calculate.

Rao (1963) proposed a method of a general sample size that the probabilities proportional to revised relative sizes are used to draw elements one by one and without replacement. To determine the probabilities of units selection and the

probability of getting units i and j included in a sample, Rao applied the approximate solution.

Fellegi (1963) suggested a sampling procedure that elements are drawn successively and without replacement with selection probability proportional to revised relative size. The selection probabilities are repeatedly calculated until they become stationary. Fellegi's method is proper for rotating samples since overall inclusion probability π_i does not change. However, this method is not suitable for implementing because of many calculations involved for $n > 2$.

Ha'jek (1964) presented a rejective sampling procedure with varying probabilities in such a way that each element in the population is selected with a Bernoulli trial. The results obtained are the desired property, $\pi_i = n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$. This method is called Poisson sampling.

Brewer, Early and Joyce (1972) gave a method that is much related to Poisson sampling, i.e., collocated sampling. The different point is that elements are drawn in a systematic way. In such a systematic sampling, it keeps the actual sample size close to its target sample size.

Durbin (1967) presented a sampling scheme of the sample size 2 where the first unit is drawn with probability proportional to size. And the second element is obtained by conditional probability given the unit i is selected.

Sampford (1967) suggested three procedures of drawing a sample that unit i

obtains probability $n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$ included in a sample. One is the whole sample method where the selection probabilities for all possible samples are evaluated. A sample is selected from these probabilities of samples. The second is the drawn by drawn method. The units are selected with probabilities that must be recalculated after each draw and without replacement. The last one is a rejective method. The first element is drawn with probabilities proportional to sizes and the remaining subsequent elements are drawn with probabilities proportional to $P_i/(1-nP_i)$, $P_i = \frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$, and with replacement. If all elements in a sample are not different, the entire sample is rejected. Among these methods, the rejective sampling method is the most convenient in practice and an unbiased variance can be obtained.

Vijayan (1968) extended the method of scheme A by Hanurav (1967) which discussed the problem of using auxiliary information in estimating the population total for the case of samples of size two. This method which is a general method of sample sizes n obtained $\pi_i = n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$, and $\pi_{ij} \leq \pi_i \pi_j$.

Dodds and Fryer (1971) generalized the Brewer's method for sampling two units from a population. The first element is drawn with probability proportional to its size and without replacement. The other one is selected with conditional probabilities given that the first one is drawn. Horvitz-Thompson estimator was used and the plan was more efficient than the Brewer scheme.

Sunter (1977) presented a method that is convenient for the unit selection by a computer. It is called a list sequential sampling where the first unit in the list population elements is drawn with probability nP_1. Some units with sizes less than N-n+1 are chosen with prescribed probabilities. After that the remaining elements with sizes less than n are drawn with simple random sampling. This method is not strictly πPS design and a fixed sample size.

Singh (1978) suggested a sampling scheme with probabilities proportional to sizes for a sample of size n = 2. It is similar to Durbin's method. However, the two elements are drawn with replacement at the first step. If they are the same, conditional probabilities are used to select one more element from the remaining. In this method, the πPS property, $\pi_i = n\frac{X_i}{X}$, $X=\sum_{i=1}^{N} X_i$, a nonnegative and stable variance estimator will be obtained.

In 1979, Choudhry and Singh considered a method that used both inclusion probability proportional to size sampling and a simple random sampling. First, a sample of size m< n is selected with assigned probabilities. Then the remaining elements are drawn with a simple random sampling. The result is that $\pi_i = n\frac{X_i}{X}$, $X = \sum_{i=1}^{N} X_i$, and the second order inclusion probability is shown in terms of the assigned probabilities for the initial sample. (Choudhry and Singh, 1979: 162-177 quoted in Chaudhuri and Vos, 1988: 249-250)

Srivastava and Singh (1981) presented a simple technique for inclusion probability proportional to size sampling. However, the sample size may be smaller than n, and through this method, the probability including unit i in the sample may not be strictly proportional to size.

Samiuddin and Asad (1981) suggested an exact grouping method with fixed sample size. Elements of a population are divided into n + 1 groups. One group is drawn with assigned probabilities. Then, each unit is drawn from the remaining n groups with probabilities proportional to size. They obtained $\pi_i = n\frac{X_i}{X}$, $X=\sum_{i=1}^{N} X_i$, and $\pi_{ij} = 0$ which also satisfy $\pi_{ij} \leq \pi_i \pi_j$. However, they do not give any solution to the problem of $\pi_{ij} = 0$.

Deshpande (1982) presented a procedure with unequal probabilities of selection and fixed sample size. The first element is drawn with revised relative probabilities and then rejected. The random sample of size n is chosen without replacement from the remaining of that population.

Gupta, Nigam and Kumar (1982) used the properties of balanced incomplete block designs which are available for any sample size. One block from this design is chosen with unequal specified probabilities. This method is simple in selection and an unbiased variance estimator can be obtained. However, for the larger number of N and n it becomes impracticable.

Nigam, Kumar and Gupta (1984) suggested a sampling scheme that ensures inclusion probability proportional to size selection. The method consists of

constructions of a sample space S in such a way that units as well as every pair of units are controlled to occur in proportion to their sizes. Then, one of the samples is selected with equal probabilities of selection.

Saxena, Singh and Srivastava (1986) proposed a new sampling scheme for a sample of size n. A subsample is drawn with prescribed probabilities. Then, the remaining units are selected with simple random sampling from the rest of a population.

2.3 Variance Estimation and Its Comparison

An estimator with least variance and nonnegativity is always chosen to be employed in a fixed size design. However, variance estimators from samples drawn with varying probabilities may turn out negative. Many researchers have tried to find out uniformly nonnegative estimators or conditions under which they are nonnegative (Chaudhuri and Vos, 1988). Horvitz-Thompson estimator is unbiased and does not always have nonnegative variance estimators for this method. The Yates-Grundy variance estimator, v_{SYG} for the Horvitz-Thompson estimator is more stable and always nonnegative than the Horvitz-Thompson variance estimator, v_{HT} (Cochran, 1977). The ratio estimator, as an alternative estimator, such as Lahiri (1951) estimator in sampling with probability proportional to aggregate size was investigated for nonnegativity.

Rao and Bayless (1969) performed a numerical study in comparison with unbiased variance estimators. The efficiencies of the estimators and the coefficient of variation of variance estimators were compared in 20 natural populations.

Rao and Vijayan (1977) proposed two unbiased variance estimators of the ratio estimator concerned with probability proportional to aggregate size for $n > 2$. They are nonnegative under sufficient conditions. One is easier than the other for numerical computation. For $n = 2$, the form of nonnegative unbiased variance estimators is at most one. The performance of variance estimators was investigated as regards their stability and probability of getting negative values.

Cochran (1977) introduced three situations in comparison with the performances of strategies. The first one is available on artificial populations. The second one is under the linear regression model. And the last one is on natural populations. The choice of any situation depends on the appropriation in comparisons. The superpopulation models were used in comparison with the accuracy of three techniques for estimating the population total of unequal sizes. And the expressions of their variances can be compared easily. If it seems that the expressions involved are quite complicated and difficult to handle artificial populations or natural populations are chosen.

To compare the methods of sampling with probability proportional to size, Chaudhuri and Vos (1988) used six artificial populations. All populations have the same set of x values. The values of y are constructed according to the specific model. The criteria used for comparisons are the expected variance and the variability of the variance estimator.

2.4 Previous Works on Controlled Sampling

Goodman and Kish (1950) introduced the method of controlled selection as a technique of decreasing the selection probability of nonpreferred combination. This paper presents a major usefulness in the selection of the first-stage sampling units in multi-stage sampling. The result is that the variances of estimates are reduced as compared with the corresponding variances for stratified random sampling.

Avadhani and Sukhatme (1965 and 1968) suggested a controlled simple random sampling design and simplified procedures, respectively. In 1965, they presented the controlled sampling probabilities that reduced nonpreferred samples without affecting the precision of an estimate of a parameter under consideration. However, this design will be complicated when the population size and the sample size are considerably large. Then in 1968, they presented a selection mechanism which eliminated the cumbersome equation system.

Avadhani and Sukhatme (1973) consolidated their previous results to be the new method of controlled simple random sampling without replacement. The balanced incomplete block design with parameters is used to construct controlled simple random sampling designs for reducing the risk of getting one of the nonpreferred samples. The ratio and regression estimators were considered efficient of the corresponding estimators under simple random sampling. However, the selection probabilities of nonpreferred samples may not remain the same if the balanced incomplete block design is chosen in different solutions. Then, the choice of this experimental design would require a trial and error in practice.

Rao and Nigam (1990) presented an optimal controlled simple random sampling design. This design used the linear programming technique to minimize the objective function with respect to the probability of selecting nonpreferred samples subject to the linear constraints. They obtained nonnegative unbiased variance estimators. They also pointed out that the optimal plan is in fact repeated blocks of the balanced incomplete block design.

Rao and Nigam (1992) extended the optimal controlled sampling to the Horvitz-Thompson estimator, Murthy's estimator and ratio estimator. The controlled plans will retain the fundamental principle of their designs and at the same time minimize probabilities of selecting nondesirable combinations. Hence, the plans obtained match the variance of a linear unbiased estimator associated with a specified uncontrolled plan.

The above literature is about controlled sampling plans in which the combinations of units are drawn. Avadhani and Sukhatme (1967) suggested a technique of controlled sampling. A positive single value function is applied to assign the probability to units in a population. They presented a necessary and sufficient condition that makes the controlled scheme more informative than the uncontrolled scheme.

CHAPTER 3

A PROPOSED SAMPLING SCHEME

3.1 Introduction

In this chapter, we suggest a sampling scheme with probability proportional to size of the sample elements. Under this sampling scheme, the probability of getting a sample and the technique of controlled sampling will be presented.

To explicitly introduce the idea, we use the following symbols:

X_{ij} = the actual measure of the j^{th} unit size of group i in the population, $i = 1,2, j = 1,2,...,N_i$,

n = the sample size,

S = the collection of all possible simple random samples each of size n, drawn without replacement from the population of size N,

$\pi_{j(i)}$ = the inclusion probability of the j^{th} unit in group i in samples of size n, $j = 1,2,...,N_i$, $i = 1,2$,

$\pi_{j(i),j'(i')}$ = the inclusion probability of different units j and j′ in groups i and i′ respectively, in samples of size n, $j(i) \neq j'$ (i')[1] .

---

[1] $j(i) \neq j'(i')$ if and only if $i \neq i'$ or $j \neq j'$ where $j = 1,2,...,N_i$, $j' = 1,2,...,N_{i'}$, i, i′ = 1,2.

3.2 The Sampling Scheme

A population is divided into two groups, the first consisting of N_1 preferred units and the second of N_2 nonpreferred units where $N_1+N_2 = N$. A sample of size n is to be drawn from this population by the controlled sampling with probability proportional to size without replacement.

$$\text{Let } X_{i.} = \sum_{j=1}^{N_i} X_{ij} \text{ , } i = 1,2 \text{ and } X_{..} = \sum_{i=1}^{2}\sum_{j=1}^{N_i} X_{ij} .$$

Now let

$$X'_{ij} = \begin{cases} X_{ij} + B\dfrac{X_{ij}}{X_{1.}} & \text{, for } i = 1,\ j = 1,2,\ldots N_1, \\ X_{min} & \text{, for } i = 2,\ j = 1,2,\ldots, N_2, \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

where $X_{min} = \text{Min } X_{ij}$, $j = 1,2,..,N_i$, $i = 1,2$, and $B = X_{2.} - N_2 X_{min}$. Note that $X'_{..} = \sum_{i=1}^{2}\sum_{j=1}^{N_i} X'_{ij} = X_{..}$.[2]

We propose a controlled sampling scheme consisting of the following steps:

Step(1): Draw a simple random sample s of size n without replacement from the population of size N.

Step(2): Calculate $\left(x'_{..}\right)_s = \left(\sum_{i=1}^{2}\sum_{j=1}^{n_i} x'_{ij}\right)_s$ where x'_{ij} is the value as given in (3.1). Here $\left(x'_{..}\right)_s$ is the sum of all x'_{ij}'s in the sample s.

Step(3): Draw a random number r between constants m and M, where m is a largest

---

217769

[2] $\sum_{i=1}^{2}\sum_{j=1}^{N_i} X'_{ij} = N_2X_{min}+X_{1.}+B=X_{1.}+X_{2.}=X_{..}$

integer smaller than the sum of n smallest X'_{ij}'s and M is a smallest integer larger than or equal to the sum of n largest X'_{ij}'s in the population.

Step(4): If $r \leq (x'_{..})_s$, then the sample s is accepted. Otherwise, reject the sample and repeat the process from Steps (1) to (4) until a sample is obtained.

Note: Under this sampling scheme, the larger the interval [m,M] the more the sample from Step(1) will be rejected.

Definition 3.1 This method of sampling scheme is actually a rejective method and will be called **a controlled plan**. The sampling scheme using the original measure of size X_{ij} will be called **an uncontrolled plan**. A sample obtained by the above process is called **an accepted sample**.

Definition 3.2 An effective round is any round of drawing from Steps (1) to (4) which yields an accepted sample. Otherwise, it is **an ineffective round**.

Example 3.1 Suppose that a population consists of 12 factories of ceramic dolls divided into 2 groups, preferred group with the first 9 factories and nonpreferred group with the last 3 factories. We also suppose again that a nonpreferred factory is a factory that was in operation for less than 3 years. The size of each factory is its number of workers as follows:

Group(i)	1									2		
Unit No.(j)	1	2	3	4	5	6	7	8	9	1	2	3
Factory No.	1	2	3	4	5	6	7	8	9	10	11	12
No. of workers (X_{ij})	97	100	87	103	92	83	110	69	75	96	89	116

Suppose that a sample of size n = 3 is to be drawn from this population. From (3.1), we have B = 94 where $X_{1.} = 816$, $X_{2.} = 301$ and X'_{1j} = 108.1740, 111.5196, 97.0220, 114.8652, 102.5980, 92.5613, 122.6716, 76.9485, 83.6397 for j = 1,2,...,9, respectively. And $X'_{2j} = 69$ for all j. For example,

$$X'_{11} = X_{11} + B\frac{X_{11}}{X_{1.}} = 97 + 94(\frac{97}{816}) = 108.1740, \text{ for } i = 1, j = 1$$

and $$X'_{21} = X_{min} = 69 \text{ for } i = 2, j = 1.$$

In the first round of drawing, suppose that the sample drawn in Step (1) contains units numbered 2,1,10. Then for Step (2), $(x'_{..})_s$ = 288.6936. In Step (3), suppose that a random number r = 300 is drawn between m = 206 and M = 350. Then by Step (4) the sample s = $\{U_{2(1)}, U_{1(1)}, U_{1(2)}\}$ is rejected.

The process is now returned to Steps (1) to (4) in the 2nd round. Suppose that a sample consisting of factories numbered 12,4,3 is chosen in Step (1). Then $(x'_{..})_s$ = 280.8872. If a random number, r = 219 is drawn in Step (3), then the sample $\{U_{3(2)}, U_{4(1)}, U_{3(1)}\}$ is accepted. Here, the 2nd round of drawing is an effective round.

Theorem 3.1 Under this sampling scheme,

$$\frac{1}{\binom{N}{n}} \sum_{s \in S} (x'_{..})_s = \frac{n}{N} X_{..}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.2)$$

where $\sum_{s \in S}$ is the summation over all possible samples $s \in S$ and $(x'_{..})_s$ is the sum of all x'_{ij}'s in the sample s.

Proof Since the sample is selected by simple random sampling without replacement, each unit of the population can occur in $\binom{N-1}{n-1}$ possible samples. Hence,

$$\frac{1}{\binom{N}{n}} \sum_{s \in S} (x'_{..})_s = \frac{1}{\binom{N}{n}} \binom{N-1}{n-1} \sum_{i=1}^{2} \sum_{j=1}^{N_i} X'_{ij} = \frac{n}{N} X_{..}.$$

Theorem 3.2 The probability of getting a sample s in the first effective round by this sampling scheme is given by

$$P(s) = \frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.3)$$

Proof The probability for getting a sample in Step (1) is $\frac{1}{\binom{N}{n}}$. For Step (3) and Step (4), a random number r is drawn between m and M and the draw is effective if $r \le (x'_{..})_s$. Hence,

$$P[r \le (x'_{..})_s] = \frac{(x'_{..})_s - m}{M - m}$$

and

$$P[r > (x'_{..})_s] = 1 - \frac{(x'_{..})_s - m}{M - m}.$$

This implies that the probability of getting a sample in the 1st round of drawing, i.e., the first round of drawing is effective, is equal to $\frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m}$.

The sample s drawn in Step (1) is rejected in the 1st round with probability

$$\frac{1}{\binom{N}{n}}\left[1-\frac{(x'_{..})_s - m}{M - m}\right].$$

However, an ineffective round may happen in the 1st round with any sample so long as r is greater than $(x'_{..})_s$.

So, P[getting a sample in the 1st round] $= \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m}$ and

$$P[\text{the 1st round is ineffective}] = \sum_{s\in S} \frac{1}{\binom{N}{n}}\left[1-\frac{(x'_{..})_s - m}{M - m}\right]$$

$$= 1-\sum_{s\in S} \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m}$$

$$= 1-\left(\frac{nX_{..} - Nm}{NM - Nm}\right), \text{ by Theorem 3.1.}$$

The probability of getting a sample in the second round given that the first round is ineffective is equal to

$$\left(1-\frac{nX_{..} - Nm}{NM - Nm}\right)\frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m}.$$

In general, the probability that the sample is chosen in the k^{th} round given that the previous (k-1) rounds of drawing are ineffective is

$$\left(1-\frac{nX_{..} - Nm}{NM - Nm}\right)^{k-1} \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m} \text{ for } k = 1,2,3,\ldots .$$

Therefore, the probability of getting a particular sample s in the first effective round is

$$P(s) = \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m} + \left(1 - \frac{nX_{..} - Nm}{NM - Nm}\right) \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m}$$

$$+ \left(1 - \frac{nX_{..} - Nm}{NM - Nm}\right)^2 \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m} + \ldots$$

$$= \frac{1}{\binom{N}{n}} \frac{(x'_{..})_s - m}{M - m} \frac{1}{\left[1 - \left(1 - \frac{nX_{..} - Nm}{NM - Nm}\right)\right]}$$

$$= \frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm}.$$

Under the proposed sampling scheme, the expressions for the first and second orders inclusion probabilities are available and easy to compute.

Theorem 3.3 For each i = 1,2 and j =1,2,...,N_i;

$$\sum_{s \ni j(i)} \frac{(x'_{..})_s - m}{nX_{..} - Nm} = \frac{\binom{N-2}{n-2}}{nX_{..} - Nm} \left(X_{..} + \frac{N-n}{n-1} X'_{ij} - m \frac{N-1}{n-1}\right), \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3.4)$$

where $\sum_{s \ni j(i)}$ is the summation over all samples s containing the unit j of group i.

Proof Let $\sum_{s \not\ni j(i)}$ be the summation over all samples s not containing the unit j of group i. Then

$$\sum_{s \ni j(i)} \frac{(x'_{..})_s - m}{nX_{..} - Nm} = \sum_{s \in S} \frac{(x'_{..})_s - m}{nX_{..} - Nm} - \sum_{s \not\ni j(i)} \frac{(x'_{..})_s - m}{nX_{..} - Nm}$$

$$= \frac{1}{nX_{..} - Nm}\left[X'_{11}\binom{N-1}{n-1} + X'_{12}\binom{N-1}{n-1} + \ldots + X'_{2N_2}\binom{N-1}{n-1} - m\binom{N}{n}\right]$$

$$- \frac{1}{nX_{..} - Nm}\left[X'_{11}\binom{N-2}{n-1} + \ldots + X'_{i,j-1}\binom{N-2}{n-1} + X'_{i,j+1}\binom{N-2}{n-1}\right.$$

$$\left. + \ldots + X'_{2N_2}\binom{N-2}{n-1} - m\binom{N-1}{n}\right],^{[3]}$$

$$= \frac{1}{nX_{..} - Nm}\left[X_{..}\binom{N-1}{n-1} - X_{..}\binom{N-2}{n-1} + X'_{ij}\binom{N-2}{n-1} - m\binom{N}{n}\right.$$

$$\left. + m\binom{N-1}{n}\right]$$

$$= \frac{1}{nX_{..} - Nm}\left[X_{..}\binom{N-2}{n-2} + X'_{ij}\binom{N-2}{n-2}\frac{N-n}{n-1} - m\binom{N-2}{n-2}\frac{N-1}{n-1}\right]$$

$$= \frac{\binom{N-2}{n-2}}{nX_{..} - Nm}\left(X_{..} + X'_{ij}\frac{N-n}{n-1} - m\frac{N-1}{n-1}\right).$$

Theorem 3.4 Under our proposed sampling scheme, we obtain

$$\pi_{j(i)} = \begin{cases} \dfrac{X_{..}(n-1) + (N-n)\left[X_{ij} + B\dfrac{X_{ij}}{X_{i.}}\right] - m(N-1)}{(N-1)(X_{..} - Nm/n)}, & i = 1,\ j = 1,\ldots, N_1, \\ \dfrac{X_{..}(n-1) + (N-n)X_{min} - m(N-1)}{(N-1)(X_{..} - Nm/n)}, & i = 2,\ j = 1,\ldots, N_2, \end{cases} \quad \ldots\ldots\ (3.5)$$

where $\pi_{j(i)}$ is the first order inclusion probability of unit j of group i.

---

[3] Each unit of the population other than unit j in group i can be drawn to be an element in $\binom{N-2}{n-1}$ possible samples (those samples do not contain unit j in group i) so $\sum_{s \neq j(i)} (x'_{..})_s = \sum_{k=1}^{2} \sum_{\substack{l=1 \\ l(k) \neq j(i)}}^{N_k} X'_{kl}\binom{N-2}{n-1}$

Proof $\pi_{j(i)} = \sum_{s \ni j(i)} P(s)$

$$= \sum_{s \ni j(i)} \frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm}, \text{ by Theorem 3.2}$$

$$= \frac{N}{\binom{N}{n}} \frac{\binom{N-2}{n-2}}{nX_{..} - Nm} \left(X_{..} + X'_{ij} \frac{N-n}{n-1} - m \frac{N-1}{n-1} \right), \text{ by Theorem 3.3,}$$

$$= \frac{n-1}{N-1} \frac{1}{X_{..} - Nm/n} \left(X_{..} + X'_{ij} \frac{N-n}{n-1} - m \frac{N-1}{n-1} \right)$$

$$= \frac{1}{N-1} \frac{1}{X_{..} - Nm/n} \left(X_{..}(n-1) + (N-n)X'_{ij} - m(N-1) \right), \text{ replace } X'_{ij} \text{ from (3.1),}$$

$$= \begin{cases} \dfrac{X_{..}(n-1) + (N-n)\left[X_{ij} + B \dfrac{X_{ij}}{X_{1.}} \right] - m(N-1)}{(N-1)(X_{..} - Nm/n)}, & i = 1, j = 1, \ldots, N_1, \\ \dfrac{X_{..}(n-1) + (N-n)X_{min} - m(N-1)}{(N-1)(X_{..} - Nm/n)}, & i = 2, j = 1, \ldots, N_2. \end{cases}$$

Theorem 3.5 Under our proposed sampling scheme, for each pair $j(i) \neq j'(i')$; $i, i' = 1,2$, $j = 1,2,\ldots,N_i$, and $j' = 1,2,\ldots,N_{i'}$ it is true that

$$\sum_{\substack{s \ni j(i), j'(i') \\ j(i) \neq j'(i')}} \frac{(x'_{..})_s - m}{nX_{..} - Nm} = \frac{1}{nX_{..} - Nm} \left[\left\{ \binom{N-2}{n-1} - \binom{N-3}{n-1} \right\} (X'_{ij} + X'_{i'j'}) \right.$$

$$\left. + \left\{ \binom{N-1}{n-1} - 2\binom{N-2}{n-1} + \binom{N-3}{n-1} \right\} X'_{..} - m \binom{N-2}{n-2} \right], \quad (3.6)$$

where $\sum_{\substack{s \ni j(i),j'(i') \\ j(i)\neq j'(i')}}$ is the summation over all possible samples $s \in S$ containing both the units j and j′ in groups i and i′ respectively, for $j(i) \neq j'(i')$.

Proof Let $\sum_{s \not\ni j(i),j'(i') \atop j(i)\neq j'(i')}$ be the summation over all possible samples $s \in S$ not containing both the units j and j′ in groups i and i′ respectively and $j(i) \neq j'(i')$.

Consider the left hand side in (3.6)

$$= \frac{1}{nX_{..} - Nm}\left[\sum_{\substack{s \ni j(i),j'(i') \\ j(i)\neq j'(i')}} (x'_{..})_s - m\binom{N-2}{n-2}\right]$$

$$= \frac{1}{nX_{..} - Nm}\left[\sum_{s\in S}(x'_{..})_s - \sum_{s \not\ni j(i)}(x'_{..})_s - \sum_{s \not\ni j'(i')}(x'_{..})_s + \sum_{s \not\ni j(i),j'(i') \atop j(i)\neq j'(i')} (x'_{..})_s - m\binom{N-2}{n-2}\right]$$

$$= \frac{1}{nX_{..} - Nm}\left[\binom{N-1}{n-1}X_{..} - \binom{N-2}{n-1}(X_{..} - X'_{ij}) - \binom{N-2}{n-1}(X_{..} - X'_{i'j'})\right.$$

$$\left. + \binom{N-3}{n-1}(X_{..} - X'_{ij} - X'_{i'j'}) - m\binom{N-2}{n-2}\right]$$

$$= \frac{1}{nX_{..} - Nm}\left[\left\{\binom{N-2}{n-1} - \binom{N-3}{n-1}\right\}(X'_{ij} + X'_{i'j'}) + \left\{\binom{N-1}{n-1} - 2\binom{N-2}{n-1} + \binom{N-3}{n-1}\right\}X_{..}\right.$$

$$\left. - m\binom{N-2}{n-2}\right].$$

Theorem 3.6 The second order inclusion probability for this sampling method is

$$\pi_{j(i),j'(i')} = \frac{n-1}{N-2}(\pi_{j(i)}+\pi_{j'(i')}) - \frac{n(n-1)X_{..}}{(N-1)(N-2)(X_{..}-Nm/n)} + \frac{mN(n-1)}{(N-1)(N-2)(X_{..}-Nm/n)}, \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.7)$$

$j(i) \neq j'(i')$; $i, i' = 1, 2$, $j = 1,\ldots, N_i$, $j' = 1,\ldots,N_{i'}$ where $\pi_{j(i)}$, $\pi_{j'(i')}$ are in Theorem 3.4.

Proof $\pi_{j(i),j'(i')} = \sum_{\substack{s \ni j(i),j'(i') \\ j(i)\neq j'(i')}} P(s)$

$$= \sum_{\substack{s \ni j(i),j'(i') \\ j(i)\neq j'(i')}} \frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm}, \text{ by Theorem 3.2}$$

$$= \frac{N}{\binom{N}{n}} \frac{1}{nX_{..} - Nm} \left[\left\{ \binom{N-2}{n-1} - \binom{N-3}{n-1} \right\} (X'_{ij} + X'_{i'j'}) \right.$$

$$\left. + \left\{ \binom{N-1}{n-1} - 2\binom{N-2}{n-1} + \binom{N-3}{n-1} \right\} X_{..} - m\binom{N-2}{n-2} \right], \text{ by Theorem 3.5}$$

$$= \frac{1}{nX_{..} - Nm} \frac{1}{(N-1)(N-2)} [\{(N-2)(N-n)-(N-n)(N-n-1)\}(X'_{ij} + X'_{i'j'})$$

$$+\{(N-1)(N-2)-2(N-2)(N-n)+(N-n)(N-n-1)\} X_{..} - m(N-2)(n-1)]$$

$$= \frac{1}{nX_{..} - Nm} \frac{1}{(N-1)(N-2)} [(N-n)(n-1)(X'_{ij} + X'_{i'j'}) + (n-1)(n-2) X_{..}$$

$$-m(N-2)(n-1)]$$

$$= \frac{n-1}{N-2}(\pi_{j(i)} + \pi_{j'(i')}) - \frac{n(n-1)X_{..}}{(N-1)(N-2)(X_{..}-Nm/n)}$$

$$+ \frac{mN(n-1)}{(N-1)(N-2)(X_{..}-Nm/n)}.$$

Definition 3.3 **An admissible scheme** is a sampling scheme which yields the probability of obtaining a nonpreferred sample not larger than that under the uncontrolled ones, i.e., X_{ij}'s are not adjusted.

Theorem 3.7 If $c \geq n\left(\frac{N_2}{N}\right)$, where c is as in definition 1.2, then the proposed sampling scheme is admissible.

Proof Let L be the set of all nonpreferred samples, i.e., samples containing at least c nonpreferred units. Then the number of all nonprefered samples, $\sum_{s\in L}$, is equal to $\sum_{k=c}^{Min(n,N_2)} \binom{N_1}{n-k}\binom{N_2}{k}$. Each unit in that preferred group occurs $\binom{N_1-1}{n-k-1}\binom{N_2}{k}$ times and each unit in that nonpreferred group occurs $\binom{N_1}{n-k}\binom{N_2-1}{k-1}$ times for k = c,c+1,..,Min(n, N_2).

The probability of getting a nonpreferred sample is given by

$$\sum_{s\in L} \frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm} = \frac{N}{\binom{N}{n}} \frac{1}{nX_{..} - Nm} \sum_{k=c}^{Min(n,N_2)} \left[\binom{N_1-1}{n-k-1}\binom{N_2}{k}\left(\sum_{j=1}^{N_1} X'_{1j}\right) \right.$$

$$\left. -\binom{N_1-1}{n-k-1}\binom{N_2}{k}m + \binom{N_1}{n-k}\binom{N_2-1}{k-1}\sum_{j=1}^{N_2} X'_{2j} - \binom{N_1}{n-k}\binom{N_2-1}{k-1}m \right].$$

Therefore, the probability of getting a nonpreferred sample where X'_{ij}'s are used given in (3.1) is

$$\frac{N}{\binom{N}{n}}\frac{1}{nX_{..}-Nm}\sum_{k=c}^{Min(n,N_2)}\left[\binom{N_1-1}{n-k-1}\binom{N_2}{k}\sum_{j=1}^{N_1}\left(X_{1j}+B\frac{X_{1j}}{X_{1.}}\right)-\binom{N_1-1}{n-k-1}\binom{N_2}{k}m\right.$$

$$\left.+\binom{N_1}{n-k}\binom{N_2-1}{k-1}N_2X_{min}-\binom{N_1}{n-k}\binom{N_2-1}{k-1}m\right]$$

$$=\frac{N}{\binom{N}{n}}\frac{1}{nX_{..}-Nm}\sum_{k=c}^{Min(n,N_2)}\left[\binom{N_1-1}{n-k-1}\binom{N_2}{k}X_{1.}+\binom{N_1-1}{n-k-1}\binom{N_2}{k}B\right.$$

$$\left.+\binom{N_1}{n-k}\binom{N_2-1}{k-1}N_2X_{min}-\binom{N_1-1}{n-k-1}\binom{N_2}{k}m-\binom{N_1}{n-k}\binom{N_2-1}{k-1}m\right]. \quad \ldots (3.8)$$

The probability of getting a sample in L where X_{ij}'s are used is

$$\frac{N}{\binom{N}{n}}\frac{1}{nX_{..}-Nm}\sum_{k=c}^{Min(n,N_2)}\left[\binom{N_1-1}{n-k-1}\binom{N_2}{k}X_{1.}-\binom{N_1-1}{n-k-1}\binom{N_2}{k}m+\binom{N_1}{n-k}\binom{N_2-1}{k-1}X_{2.}\right.$$

$$\left.-\binom{N_1}{n-k}\binom{N_2-1}{k-1}m\right]. \quad \ldots (3.9)$$

By subtracting (3.9) from (3.8) we have the difference

$$\frac{N}{\binom{N}{n}}\frac{1}{nX_{..}-Nm}\sum_{k=c}^{Min(n,N_2)}\left[\binom{N_1-1}{n-k-1}\binom{N_2}{k}B+\binom{N_1}{n-k}\binom{N_2-1}{k-1}N_2X_{min}-\binom{N_1}{n-k}\binom{N_2-1}{k-1}X_{2.}\right]$$

$$=\frac{N}{\binom{N}{n}}\frac{1}{nX_{..}-Nm}\sum_{k=c}^{Min(n,N_2)}\left[\binom{N_1-1}{n-k-1}\binom{N_2}{k}(X_{2.}-N_2X_{min})\right.$$

$$\left.-\binom{N_1}{n-k}\binom{N_2-1}{k-1}(X_{2.}-N_2X_{min})\right]$$

$$=\frac{N}{\binom{N}{n}}\frac{X_{2.}-N_2X_{min}}{nX_{..}-Nm}\sum_{k=c}^{Min(n,N_2)}\left[\binom{N_1}{n-k}\binom{N_2}{k}\frac{n-k}{N_1}-\binom{N_1}{n-k}\binom{N_2}{k}\frac{k}{N_2}\right]. \quad \ldots\ldots (3.10)$$

For each c that is greater than or equal to $n\left(\frac{N_2}{N}\right)$, we have

$$\left[\binom{N_1}{n-c}\binom{N_2}{c}\left(\frac{n}{N_1}\cdot c\frac{1}{N_1}\cdot c\frac{1}{N_2}\right)\right] = \left[\binom{N_1}{n-c}\binom{N_2}{c}\left(\frac{n}{N_1}\cdot c\left(\frac{N}{N_1N_2}\right)\right)\right] \le 0.$$

Therefore, the probability of getting a nonpreferrd sample under X'_{ij} is not larger than that under X_{ij}. Hence, the proposed sampling scheme is admissible.

Example 3.2 From example 3.1, a sample of size n = 3 is selected from a population of size N = 12. The population is divided into two groups: preferred units with $N_1 = 9$ and nonpreferred units with $N_2 = 3$. Let c be an integer greater than or equal to $n\left(\frac{N_2}{N}\right) = \frac{3}{4}$, i.e., c = 1,2 or 3. The probability of getting each of the possible nonpreferred samples using X_{ij}'s and X'_{ij}'s may be calculated with the results as presented in Table 3.1.

Table 3.1 The Probability of Getting a Nonpreferred Sample Under the Uncontrolled Plan, P(s) and the Controlled Plan, $P^*(s)$

The Number of Nonpreferred Units in a Sample (c)	Nonpreferred Sample (s)	P(s)	$P^*(s)$
1	$\{U_{1(1)},U_{2(1)},U_{1(2)}\}$	0.0054	0.0051
	$\{U_{1(1)},U_{2(1)},U_{2(2)}\}$	0.0050	0.0051
	$\{U_{1(1)},U_{2(1)},U_{3(2)}\}$	0.0066	0.0051
	$\{U_{1(1)},U_{3(1)},U_{1(2)}\}$	0.0046	0.0042
	$\{U_{1(1)},U_{3(1)},U_{2(2)}\}$	0.0042	0.0042
	$\{U_{1(1)},U_{3(1)},U_{3(2)}\}$	0.0058	0.0042
	$\{U_{1(1)},U_{4(1)},U_{1(2)}\}$	0.0056	0.0053

Table 3.1 (Continued)

The Number of Nonpreferred Units in a Sample (c)	Nonpreferred Sample (s)	P(s)	$P^*(s)$
1	$\{U_{1(1)},U_{4(1)},U_{2(2)}\}$	0.0052	0.0053
	$\{U_{1(1)},U_{4(1)},U_{3(2)}\}$	0.0068	0.0053
	$\{U_{1(1)},U_{5(1)},U_{1(2)}\}$	0.0049	0.0046
	$\{U_{1(1)},U_{5(1)},U_{2(2)}\}$	0.0045	0.0046
	$\{U_{1(1)},U_{5(1)},U_{3(2)}\}$	0.0061	0.0046
	$\{U_{1(1)},U_{6(1)},U_{1(2)}\}$	0.0043	0.0039
	$\{U_{1(1)},U_{6(1)},U_{2(2)}\}$	0.0039	0.0039
	$\{U_{1(1)},U_{6(1)},U_{3(2)}\}$	0.0056	0.0039
	$\{U_{1(1)},U_{7(1)},U_{1(2)}\}$	0.0060	0.0058
	$\{U_{1(1)},U_{7(1)},U_{2(2)}\}$	0.0056	0.0058
	$\{U_{1(1)},U_{7(1)},U_{3(2)}\}$	0.0073	0.0058
	$\{U_{1(1)},U_{8(1)},U_{1(2)}\}$	0.0035	0.0029
	$\{U_{1(1)},U_{8(1)},U_{2(2)}\}$	0.0030	0.0029
	$\{U_{1(1)},U_{8(1)},U_{3(2)}\}$	0.0047	0.0029
	$\{U_{1(1)},U_{9(1)},U_{1(2)}\}$	0.0038	0.0034
	$\{U_{1(1)},U_{9(1)},U_{2(2)}\}$	0.0034	0.0034
	$\{U_{1(1)},U_{9(1)},U_{3(2)}\}$	0.0051	0.0034
	$\{U_{2(1)},U_{3(1)},U_{1(2)}\}$	0.0048	0.0044
	$\{U_{2(1)},U_{3(1)},U_{2(2)}\}$	0.0043	0.0044
	$\{U_{2(1)},U_{3(1)},U_{3(2)}\}$	0.0060	0.0044
	$\{U_{2(1)},U_{4(1)},U_{1(2)}\}$	0.0058	0.0055
	$\{U_{2(1)},U_{4(1)},U_{2(2)}\}$	0.0053	0.0055
	$\{U_{2(1)},U_{4(1)},U_{3(2)}\}$	0.0070	0.0055
	$\{U_{2(1)},U_{5(1)},U_{1(2)}\}$	0.0051	0.0047
	$\{U_{2(1)},U_{5(1)},U_{2(2)}\}$	0.0046	0.0047
	$\{U_{2(1)},U_{5(1)},U_{3(2)}\}$	0.0063	0.0047
	$\{U_{2(1)},U_{6(1)},U_{1(2)}\}$	0.0045	0.0041
	$\{U_{2(1)},U_{6(1)},U_{2(2)}\}$	0.0041	0.0041

Table 3.1 (Continued)

The Number of Nonpreferred Units in a Sample (c)	Nonpreferred Sample (s)	P(s)	$P^*(s)$
1	$\{U_{2(1)},U_{6(1)},U_{3(2)}\}$	0.0058	0.0041
	$\{U_{2(1)},U_{7(1)},U_{1(2)}\}$	0.0062	0.0060
	$\{U_{2(1)},U_{7(1)},U_{2(2)}\}$	0.0058	0.0060
	$\{U_{2(1)},U_{7(1)},U_{3(2)}\}$	0.0074	0.0060
	$\{U_{2(1)},U_{8(1)},U_{1(2)}\}$	0.0036	0.0032
	$\{U_{2(1)},U_{8(1)},U_{2(2)}\}$	0.0032	0.0032
	$\{U_{2(1)},U_{8(1)},U_{3(2)}\}$	0.0049	0.0032
	$\{U_{2(1)},U_{9(1)},U_{1(2)}\}$	0.0040	0.0036
	$\{U_{2(1)},U_{9(1)},U_{2(2)}\}$	0.0036	0.0036
	$\{U_{2(1)},U_{9(1)},U_{3(2)}\}$	0.0053	0.0036
	$\{U_{3(1)},U_{4(1)},U_{1(2)}\}$	0.0049	0.0046
	$\{U_{3(1)},U_{4(1)},U_{2(2)}\}$	0.0045	0.0046
	$\{U_{3(1)},U_{4(1)},U_{3(2)}\}$	0.0062	0.0046
	$\{U_{3(1)},U_{5(1)},U_{1(2)}\}$	0.0043	0.0039
	$\{U_{3(1)},U_{5(1)},U_{2(2)}\}$	0.0038	0.0039
	$\{U_{3(1)},U_{5(1)},U_{2(2)}\}$	0.0038	0.0039
	$\{U_{3(1)},U_{5(1)},U_{3(2)}\}$	0.0055	0.0039
	$\{U_{3(1)},U_{6(1)},U_{1(2)}\}$	0.0037	0.0033
	$\{U_{3(1)},U_{6(1)},U_{2(2)}\}$	0.0033	0.0033
	$\{U_{3(1)},U_{6(1)},U_{3(2)}\}$	0.0050	0.0033
	$\{U_{3(1)},U_{7(1)},U_{1(2)}\}$	0.0054	0.0051
	$\{U_{3(1)},U_{7(1)},U_{2(2)}\}$	0.0050	0.0051
	$\{U_{3(1)},U_{7(1)},U_{3(2)}\}$	0.0066	0.0051
	$\{U_{3(1)},U_{8(1)},U_{1(2)}\}$	0.0029	0.0023
	$\{U_{3(1)},U_{8(1)},U_{2(2)}\}$	0.0024	0.0023
	$\{U_{3(1)},U_{8(1)},U_{3(2)}\}$	0.0041	0.0023
	$\{U_{3(1)},U_{9(1)},U_{1(2)}\}$	0.0032	0.0027
	$\{U_{3(1)},U_{9(1)},U_{2(2)}\}$	0.0028	0.0027

Table 3.1 (Continued)

The Number of Nonpreferred Units in a Sample (c)	Nonpreferred Sample (s)	P(s)	$P^*(s)$
1	$\{U_{3(1)},U_{9(1)},U_{3(2)}\}$	0.0045	0.0027
	$\{U_{4(1)},U_{5(1)},U_{1(2)}\}$	0.0053	0.0049
	$\{U_{4(1)},U_{5(1)},U_{2(2)}\}$	0.0048	0.0049
	$\{U_{4(1)},U_{5(1)},U_{3(2)}\}$	0.0065	0.0049
	$\{U_{4(1)},U_{6(1)},U_{1(2)}\}$	0.0047	0.0044
	$\{U_{4(1)},U_{6(1)},U_{2(2)}\}$	0.0043	0.0044
	$\{U_{4(1)},U_{6(1)},U_{3(2)}\}$	0.0060	0.0044
	$\{U_{4(1)},U_{7(1)},U_{1(2)}\}$	0.0064	0.0062
	$\{U_{4(1)},U_{7(1)},U_{2(2)}\}$	0.0060	0.0062
	$\{U_{4(1)},U_{7(1)},U_{3(2)}\}$	0.0076	0.0062
	$\{U_{4(1)},U_{8(1)},U_{1(2)}\}$	0.0038	0.0034
	$\{U_{4(1)},U_{8(1)},U_{2(2)}\}$	0.0034	0.0034
	$\{U_{4(1)},U_{8(1)},U_{3(2)}\}$	0.0051	0.0034
	$\{U_{4(1)},U_{9(1)},U_{1(2)}\}$	0.0042	0.0038
	$\{U_{4(1)},U_{9(1)},U_{2(2)}\}$	0.0038	0.0038
	$\{U_{4(1)},U_{9(1)},U_{3(2)}\}$	0.0055	0.0038
	$\{U_{5(1)},U_{6(1)},U_{1(2)}\}$	0.0040	0.0036
	$\{U_{5(1)},U_{6(1)},U_{2(2)}\}$	0.0036	0.0036
	$\{U_{5(1)},U_{6(1)},U_{3(2)}\}$	0.0053	0.0036
	$\{U_{5(1)},U_{7(1)},U_{1(2)}\}$	0.0057	0.0055
	$\{U_{5(1)},U_{7(1)},U_{2(2)}\}$	0.0053	0.0055
	$\{U_{5(1)},U_{7(1)},U_{3(2)}\}$	0.0070	0.0055
	$\{U_{5(1)},U_{8(1)},U_{1(2)}\}$	0.0032	0.0026
	$\{U_{5(1)},U_{8(1)},U_{2(2)}\}$	0.0027	0.0026
	$\{U_{5(1)},U_{8(1)},U_{3(2)}\}$	0.0044	0.0026
	$\{U_{5(1)},U_{9(1)},U_{1(2)}\}$	0.0035	0.0030
	$\{U_{5(1)},U_{9(1)},U_{2(2)}\}$	0.0031	0.0030
	$\{U_{5(1)},U_{9(1)},U_{3(2)}\}$	0.0048	0.0030

Table 3.1 (Continued)

The Number of Nonpreferred Units in a Sample (c)	Nonpreferred Sample (s)	P(s)	$P^*(s)$
	$\{U_{6(1)},U_{7(1)},U_{1(2)}\}$	0.0051	0.0048
	$\{U_{6(1)},U_{7(1)},U_{2(2)}\}$	0.0047	0.0048
	$\{U_{6(1)},U_{7(1)},U_{3(2)}\}$	0.0064	0.0048
	$\{U_{6(1)},U_{8(1)},U_{1(2)}\}$	0.0026	0.0020
	$\{U_{6(1)},U_{8(1)},U_{2(2)}\}$	0.0022	0.0020
	$\{U_{6(1)},U_{8(1)},U_{3(2)}\}$	0.0038	0.0020
	$\{U_{6(1)},U_{9(1)},U_{1(2)}\}$	0.0030	0.0024
	$\{U_{6(1)},U_{9(1)},U_{2(2)}\}$	0.0025	0.0024
	$\{U_{6(1)},U_{9(1)},U_{3(2)}\}$	0.0042	0.0024
	$\{U_{7(1)},U_{8(1)},U_{1(2)}\}$	0.0043	0.0029
	$\{U_{7(1)},U_{8(1)},U_{2(2)}\}$	0.0038	0.0029
	$\{U_{7(1)},U_{8(1)},U_{3(2)}\}$	0.0055	0.0029
	$\{U_{7(1)},U_{9(1)},U_{1(2)}\}$	0.0047	0.0043
	$\{U_{7(1)},U_{9(1)},U_{2(2)}\}$	0.0042	0.0043
	$\{U_{7(1)},U_{9(1)},U_{3(2)}\}$	0.0059	0.0043
	$\{U_{8(1)},U_{9(1)},U_{1(2)}\}$	0.0021	0.0015
	$\{U_{8(1)},U_{9(1)},U_{2(2)}\}$	0.0017	0.0015
	$\{U_{8(1)},U_{9(1)},U_{3(2)}\}$	0.0034	0.0015
2	$\{U_{1(1)},U_{1(2)},U_{2(2)}\}$	0.0047	0.0025
	$\{U_{2(1)},U_{1(2)},U_{2(2)}\}$	0.0049	0.0027
	$\{U_{3(1)},U_{1(2)},U_{2(2)}\}$	0.0041	0.0018
	$\{U_{4(1)},U_{1(2)},U_{2(2)}\}$	0.0051	0.0029
	$\{U_{5(1)},U_{1(2)},U_{2(2)}\}$	0.0044	0.0021
	$\{U_{6(1)},U_{1(2)},U_{2(2)}\}$	0.0038	0.0015
	$\{U_{7(1)},U_{1(2)},U_{2(2)}\}$	0.0055	0.0033
	$\{U_{8(1)},U_{1(2)},U_{2(2)}\}$	0.0030	0.0005
	$\{U_{9(1)},U_{1(2)},U_{2(2)}\}$	0.0034	0.0009
	$\{U_{1(1)},U_{2(2)},U_{3(2)}\}$	0.0060	0.0025

Table 3.1 (Continued)

The Number of Nonpreferred Units in a Sample (c)	Nonpreferred Sample (s)	P(s)	$P^*(s)$
2	$\{U_{2(1)},U_{2(2)},U_{3(2)}\}$	0.0061	0.0027
	$\{U_{3(1)},U_{2(2)},U_{3(2)}\}$	0.0053	0.0018
	$\{U_{4(1)},U_{2(2)},U_{3(2)}\}$	0.0063	0.0029
	$\{U_{5(1)},U_{2(2)},U_{3(2)}\}$	0.0056	0.0021
	$\{U_{6(1)},U_{2(2)},U_{3(2)}\}$	0.0051	0.0015
	$\{U_{7(1)},U_{2(2)},U_{3(2)}\}$	0.0068	0.0033
	$\{U_{8(1)},U_{2(2)},U_{3(2)}\}$	0.0042	0.0005
	$\{U_{9(1)},U_{2(2)},U_{3(2)}\}$	0.0046	0.0009
	$\{U_{1(1)},U_{1(2)},U_{3(2)}\}$	0.0064	0.0025
	$\{U_{2(1)},U_{1(2)},U_{3(2)}\}$	0.0066	0.0027
	$\{U_{3(1)},U_{1(2)},U_{3(2)}\}$	0.0058	0.0018
	$\{U_{4(1)},U_{1(2)},U_{3(2)}\}$	0.0068	0.0029
	$\{U_{5(1)},U_{1(2)},U_{3(2)}\}$	0.0061	0.0021
	$\{U_{6(1)},U_{1(2)},U_{3(2)}\}$	0.0055	0.0015
	$\{U_{7(1)},U_{1(2)},U_{3(2)}\}$	0.0072	0.0033
	$\{U_{8(1)},U_{1(2)},U_{3(2)}\}$	0.0047	0.0005
	$\{U_{9(1)},U_{1(2)},U_{3(2)}\}$	0.0050	0.0009
3	$\{U_{1(2)},U_{2(2)},U_{3(2)}\}$	0.0059	0.00006

P(s): The probability of a nonpreferred sample under the uncontrolled plan
$P^*(s)$: The probability of a nonpreferred sample under the controlled plan

From Table 3.1, for c = 1, the probability for obtaining a nonpreferred sample, under the controlled plan is 0.489 and under the uncontrolled plan is 0.656. For c = 2, this probability under the controlled plan is 0.054 and under the uncontrolled plan is 0.143. For the last case, where c = 3, the probability for obtaining

a nonpreferred sample under the controlled plan is 0.00006 and under the uncontrolled plan is 0.059.

Therefore, the proposed sampling scheme is a controlled one according to definition 1.3. In other words, this sampling scheme is admissible.

CHAPTER 4

ESTIMATION OF PARAMETER

4.1 The Horvitz-Thompson Estimator

In sampling with varying probability without replacement, an unbiased estimator of the population total, $Y_{..} = \sum_{i=1}^{2}\sum_{j=1}^{N_i} Y_{ij}$ as given by Horvitz-Thompson (1952) and applied to our sampling scheme may be written as

$$\hat{Y}_{HT} = \sum_{i=1}^{2}\sum_{j=1}^{n_i} \frac{y_{ij}}{\pi_{j(i)}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4.1)$$

where y_{ij} is the observed value from unit j of group i and $\pi_{j(i)}$ is the first order inclusion probability of unit j of group i. The variance of this estimator given by Horvitz-Thompson is

$$V\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{2}\sum_{j=1}^{N_i}\left(\frac{1-\pi_{j(i)}}{\pi_{j(i)}}\right)Y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}}\left(\frac{\pi_{j(i),j'(i')} - \pi_{j(i)}\pi_{j'(i')}}{\pi_{j(i)}\pi_{j'(i')}}\right)Y_{ij}Y_{i'j'}, \quad \ldots\ldots\ldots \quad (4.2)$$

where $\pi_{j(i)}$ and $\pi_{j(i),j'(i')}$ are the first and second order inclusion probabilities, respectively. This variance may be expressed in another form as follows:

$$^{[1]}V\left(\hat{Y}_{HT}\right) = \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}}\left(\pi_{j(i)}\pi_{j'(i')} - \pi_{j(i)j'(i')}\right)\left(\frac{Y_{ij}}{\pi_{j(i)}} - \frac{Y_{i'j'}}{\pi_{j'(i')}}\right)^2 \quad \ldots\ldots\ldots\ldots\ldots \quad (4.3)$$

---

[1] See Cochran, 1977: 260 for a discussion of another form of the variance expression.

An unbiased estimator of the variance of $\hat{Y}_{HT}$, as given by Horvitz-Thompson is

$$v_{HT}\left(\hat{Y}_{HT}\right)=\sum_{i=1}^{2}\sum_{j=1}^{n_i}\frac{\left(1-\pi_{j(i)}\right)}{\pi_{j(i)}^{2}}y_{ij}^{2}+\sum_{\substack{i,i'=1\\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\left(\frac{\pi_{j(i),j'(i)}-\pi_{j(i)}\pi_{j'(i')}}{\pi_{j(i)}\pi_{j'(i')}\pi_{j(i),j'(i')}}\right)y_{ij}y_{i'j'} \quad \ldots\ldots\ldots (4.4)$$

Another unbiased variance estimator of $\hat{Y}_{HT}$ modified by Yates and Grundy (1953) and by Sen (1953) as applied to our present situation is

$$v_{SYG}\left(\hat{Y}_{HT}\right)=\sum_{\substack{i,i'=1\\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{j'}}\left(\frac{\pi_{j(i)}\pi_{j'(i')}-\pi_{j(i),j'(i')}}{\pi_{j(i),j'(i')}}\right)\left(\frac{y_{ij}}{\pi_{j(i)}}-\frac{y_{i'j'}}{\pi_{j'(i')}}\right)^{2} \quad \ldots\ldots\ldots\ldots (4.5)$$

It was found that the estimator $v_{SYG}\left(\hat{Y}_{HT}\right)$ would take negative values less frequently than $v_{HT}\left(\hat{Y}_{HT}\right)$.(Cochran, 1977)

4.2 An Alternative Estimator

We will now present an alternative unbiased estimator of the population total under the proposed sampling scheme.

Theorem 4.1 Under the new sampling scheme, an unbiased estimator of the population total, $Y_{..}=\sum_{i=1}^{2}\sum_{j=1}^{N_i}Y_{ij}$ is given by

$$\hat{Y}_{..}=\frac{nX_{..}-Nm}{x'_{..}-m}\frac{y_{..}}{n}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.6)$$

where $y_{..}=\sum_{i=1}^{2}\sum_{j=1}^{n_i}y_{ij}$ is the sum of all y_{ij}'s in the sample. All other symbols are as given in Chapter III.

Proof Let S be the set of all possible samples s each of size n under the sampling scheme. Then

$$E\left(\hat{Y}_{..}\right) = \sum_{s\in S} \frac{nX_{..} - Nm}{\left(x'_{..}\right)_s - m} \frac{\left(y_{..}\right)_s}{n} P(s)$$

$$= \sum_{s\in S} \frac{nX_{..} - Nm}{\left(x'_{..}\right)_s - m} \frac{\left(y_{..}\right)_s}{n} \frac{N}{\binom{N}{n}} \frac{\left(x'_{..}\right)_s - m}{nX_{..} - Nm}, \text{ by Theorem 3.2}$$

$$= \sum_{s\in S} \frac{1}{\binom{N-1}{n-1}} \left(y_{..}\right)_s$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{N_i} Y_{ij}, \text{ since } \sum_{s\in S}\left(\sum_{i=1}^{2}\sum_{j=1}^{n_i} y_{ij}\right)_s = \binom{N-1}{n-1}\sum_{i=1}^{2}\sum_{j=1}^{N_i} Y_{ij}.^{[2]}$$

Definition 4.1 An estimator is **consistent** if the estimator converges to the population value as the sample size n approaches the population size N (Cochran, 1977: 21).

Theorem 4.2 The proposed estimator $\hat{Y}_{..} = \frac{nX_{..} - Nm}{x'_{..} - m} \frac{y_{..}}{n}$ of the population total, $Y_{..}$, is consistent.

Proof
$$\lim_{n\to N} \hat{Y}_{..} = \lim_{n\to N}\left(\frac{nX_{..} - Nm}{x'_{..} - m} \frac{y_{..}}{n}\right)$$

$$= \frac{NX_{..} - Nm}{X_{..} - m} \frac{Y_{..}}{N}$$

$$= Y_{..}.$$

---

[2] Since a sample is drawn by a method of simple random sampling without replacement then each unit will occur in $\binom{N-1}{n-1}$ possible samples.

Theorem 4.3 Under the proposed sampling scheme, the variance of the estimator $\hat{Y}_{..}$ is given by

$$V\left(\hat{Y}_{..}\right) = \sum_{i=1}^{2}\sum_{j=1}^{N_i} a_{j(i)} Y_{ij}^2 + \sum_{i,i'=1}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} Y_{ij} Y_{i'j'}, \quad \ldots\ldots\ldots\ldots \quad (4.7)$$
$$j(i)\neq j'(i')$$

where
$$a_{j(i)} = \left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{s \ni j(i)} \frac{1}{(x'_{..})_s - m} - 1\right]$$

and
$$a_{j(i)j'(i')} = \left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{\substack{s \ni j(i), j'(i') \\ j(i)\neq j'(i')}} \frac{1}{(x'_{..})_s - m} - 1\right], \text{ for } i, i' = 1,2,$$

$j = 1,2,\ldots,N_i$, $j' = 1,2,\ldots, N_{i'}$ and $j(i) \neq j'(i')$.

Proof The variance of $\hat{Y}_{..}$ is

$$V\left(\hat{Y}_{..}\right) = E\left(\hat{Y}_{..}^2\right) - Y_{..}^2$$

$$= E\left(\frac{nX_{..} - Nm}{x'_{..} - m}\frac{y_{..}}{n}\right)^2 - Y_{..}^2$$

$$= \sum_{s\in S}\left(\frac{nX_{..} - Nm}{(x'_{..})_s - m}\frac{(y_{..})_s}{n}\right)^2 \frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm} - Y_{..}^2, \text{ by Theorem 3.2}$$

$$= \sum_{s\in S} \frac{nX_{..} - Nm}{n^2} \frac{N}{\binom{N}{n}} \frac{\left(\sum_{i=1}^{2}\sum_{j=1}^{n_i} y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}} y_{ij}y_{i'j'} \right)_s}{\left(x'_{..}\right)_s - m} - Y_{..}^2, \quad [3]$$

$$= \frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \left(\sum_{\substack{s\ni j(i),j'(i') \\ j(i)\neq j'(i')}} \frac{1}{\left(x'_{..}\right)_s - m} \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} Y_{ij}Y_{i'j'} \right.$$

$$\left. + \sum_{s\ni j(i)} \frac{1}{\left(x'_{..}\right)_s - m} \sum_{i=1}^{2}\sum_{j=1}^{N_i} Y_{ij}^2 \right) - Y_{..}^2$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{N_i} Y_{ij}^2 \left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{s\ni j(i)} \frac{1}{\left(x'_{..}\right)_s - m} - 1 \right]$$

$$+ \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} Y_{ij}Y_{i'j'} \left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{\substack{s\ni j(i),j'(i') \\ j(i)\neq j'(i')}} \frac{1}{\left(x'_{..}\right)_s - m} - 1 \right], \quad [4]$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{N_i} a_{j(i)} Y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} Y_{ij}Y_{i'j'}.$$

It should be noted that the expression for the new estimator is actually a ratio estimator which is very easy to calculate and its variance, $V\left(\hat{Y}_{..}\right)$ is a quadratic form.

---

[3] By expansion of $\left(y_{..}\right)_s^2 = \sum_{i=1}^{2}\sum_{j=1}^{n_i} y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}} y_{ij}y_{i'j'}$

[4] $Y_{..}^2 = \sum_{i=1}^{2}\sum_{j=1}^{N_i} Y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} Y_{ij}Y_{i'j'}$

4.3 Estimation of the Variance of the Estimator

In this section, we propose an unbiased estimator of the variance of the estimator. The method of estimating the variance is modified from the work of Rao and Vijayan (1977).

Definition 4.2 A function of Y_{ij} say

$$Q(Y_{11},\ldots,Y_{1N_1},Y_{21},\ldots,Y_{2N_2}) = \sum_{i=1}^{2}\sum_{j=1}^{N_i} q_{j(i)} Y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} q_{j(i)j'(i')} Y_{ij} Y_{i'j'}$$

is said to be in the form of **homogeneous quadratic function** in Y_{ij}.

The quadratic expression of $V(\hat{Y}_{..})$ in (4.7) can be rewritten in an alternate quadratic form

$$^{[5]}V(\hat{Y}_{..}) = \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_{i'}} b_{j(i)j'(i')}\left(X'_{ij} - \frac{m}{n}\right)\left(X'_{i'j'} - \frac{m}{n}\right)\left(\frac{Y_{ij}}{X'_{ij} - m/n} - \frac{Y_{i'j'}}{X'_{i'j'} - m/n}\right)^2, \quad (4.8)$$

where $b_{j(i)j'(i')} = -a_{j(i)j'(i')}$.

Theorem 4.4 An unbiased variance estimator of $V(\hat{Y}_{..})$ is

$$v(\hat{Y}_{..}) = \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}} b(s)\left(x'_{ij} - \frac{m}{n}\right)\left(x'_{i'j'} - \frac{m}{n}\right)\left(\frac{y_{ij}}{x'_{ij} - m/n} - \frac{y_{i'j'}}{x'_{i'j'} - m/n}\right)^2, \quad \ldots\ldots\ (4.9)$$

where $b(s) = \dfrac{X_{..} - Nm/n}{(x'_{..})_s - m}\left(\dfrac{N-1}{n-1} - \dfrac{X_{..} - Nm/n}{(x'_{..})_s - m}\right)$.

---

[5] See Appendix B.

Proof Taking the expectation on both sides of (4.9), we have

$$E(v(\hat{Y}_{..})) = \sum_{s\in S} P(s) \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{n_i} \sum_{j'=1}^{n_{i'}} b(s) \left(x'_{ij} - \frac{m}{n}\right)\left(x'_{i'j'} - \frac{m}{n}\right)\left(\frac{y_{ij}}{x'_{ij} - m/n} - \frac{y_{i'j'}}{x'_{i'j'} - m/n}\right)^2$$

$$= \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} \left[\sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} b(s)\ P(s)\right] \left(X'_{ij} - \frac{m}{n}\right)\left(X'_{i'j'} - \frac{m}{n}\right)\left(\frac{Y_{ij}}{X'_{ij} - m/n} - \frac{Y_{i'j'}}{X'_{i'j'} - m/n}\right)^2,$$

where $\sum_{s\ni S}$ is the summation for all possible samples and $\sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')}$ is the summation for over all samples containing units j and j′ in groups i and i′, $j(i) \neq j'(i')$ respectively.

Consider

$$\sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} b(s)\ P(s) = \sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} \frac{X_{..} - Nm/n}{(x'_{..})_s - m}\left(\frac{N-1}{n-1} - \frac{X_{..} - Nm/n}{(x'_{..})_s - m}\right)_s P(s)$$

$$= \sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} \left(\frac{N-1}{n-1}\frac{1}{\binom{N-1}{n-1}} - \frac{1}{\binom{N-1}{n-1}}\frac{X_{..} - Nm/n}{(x'_{..})_s - m}\right), \text{ by Theorem 3.2}$$

$$= \frac{N-1}{n-1}\frac{\binom{N-2}{n-2}}{\binom{N-1}{n-1}} - \frac{1}{\binom{N-1}{n-1}} \sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} \frac{X_{..} - Nm/n}{(x'_{..})_s - m}, \text{ since } \sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} 1 = \binom{N-2}{n-2}$$

$$= 1 - \frac{1}{\binom{N-1}{n-1}} \sum_{s\ni j(i),j'(i') \atop j(i)\neq j'(i')} \frac{X_{..} - Nm/n}{(x'_{..})_s - m}.$$

$$\text{Hence, } E\left(v\left(\hat{Y}_{..}\right)\right) = \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} b_{j(i)j'(i')} \left(X'_{ij} - \frac{m}{n}\right)\left(X'_{i'j'} - \frac{m}{n}\right)\left(\frac{Y_{ij}}{X'_{ij} - m/n} - \frac{Y_{i'j'}}{X'_{i'j'} - m/n}\right)^2$$

$$= V\left(\hat{Y}_{..}\right).$$

4.4 An Example

From example 3.1, suppose that the sample $\{U_{3(2)}, U_{4(1)}, U_{3(1)}\}$ was chosen from a population of size N = 12 by the proposed sampling scheme.

Under the sampling plan, we have X'_{ij}, j = 1, 2, ...,N_i, i = 1,2 and we assume that the production of the factories in the population, Y_{ij}'s in Example 3.1 is as follows:

Table 4.1 Production of the Factories, Y_{ij}, the Number of Workers, X_{ij} and the Adjusted Value of the Number of the Workers in the Population

Group (i)	Factory No. (j(i))	Y_{ij} (pieces)	X_{ij} (persons)	X'_{ij}
1	1(1)	365	97	108.1740
	2(1)	430	100	111.5196
	3(1)	225	87	97.0220
	4(1)	410	103	114.8652
	5(1)	270	92	102.5980
	6(1)	200	83	92.5613
	7(1)	460	110	122.6716
	8(1)	190	69	76.9485
	9(1)	225	75	83.6397
2	1(2)	350	96	69.0000
	2(2)	250	89	69.0000
	3(2)	520	116	69.0000

X'_{ij} : The adjusted value of the number of workers in the j^{th} unit of the group i

The total of X'_{ij} is $\sum_{i=1}^{2}\sum_{j=1}^{N_i} X'_{ij} = X'_{..} = X_{..} = 1117$.

Using Theorem 3.2, the probability of obtaining the sample in Example 3.1 is

$$\frac{N}{\binom{N}{n}} \frac{(x'_{..})_s - m}{nX_{..} - Nm} = 4.0847/879 = 0.0046$$

where $(x'_{..})_s = 280.8872$ and $m = 206$. According to (3.5), the first order inclusion probabilities of units in the sample are $\pi_{3(1)} = 0.2609$, $\pi_{4(1)} = 0.3108$ and $\pi_{3(2)} = 0.1827$. And their second order inclusion probabilities according to (3.7), are $\pi_{3(1),4(1)} = 0.0598$, $\pi_{3(1),3(2)} = 0.0342$ and $\pi_{4(1),3(2)} = 0.0441$.

By using (4.6), the estimate of the total production is

$$\hat{Y}_{..} = \frac{nX_{..} - Nm}{(x'_{13} + x'_{14} + x'_{23} - m)} \frac{y_{13} + y_{14} + y_{23}}{n}$$

$$= 4518.9760 \text{ pieces.}$$

The Horvitz-Thompson estimate of the population total as shown in (4.1) is

$$\hat{Y}_{HT} = \left(\frac{y_{13}}{\pi_{3(1)}} + \frac{y_{14}}{\pi_{4(1)}} + \frac{y_{23}}{\pi_{3(2)}} \right)$$

$$= 5027.7716 \text{ pieces.}$$

Then, the unbiased variance estimates of the estimators as shown in (4.8), (4.4) and (4.5) are

$$v(\hat{Y}_{..}) = \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{n_i} \sum_{j'=1}^{n_{i'}} b(s) \left(x'_{ij} - \frac{m}{n} \right) \left(x'_{i'j'} - \frac{m}{n} \right) \left(\frac{y_{ij}}{x'_{ij} - m/n} - \frac{y_{i'j'}}{x'_{i'j'} - m/n} \right)^2$$

$$= 3.7154 \times 10^8$$

$$v_{HT}\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{2}\sum_{j=1}^{n_i}\frac{\left(1-\pi_{j(i)}\right)}{\pi_{j(i)}^{2}}y_{ij}^{2} + \sum_{\substack{i,i'=1\\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\left(\frac{\pi_{j(i),j'(i)}-\pi_{j(i)}\pi_{j'(i')}}{\pi_{j(i)}\pi_{j'(i')}\pi_{j(i),j'(i')}}\right)y_{ij}y_{i'j'}$$

$$= 3.4738\times10^{6}$$

$$v_{SYG}\left(\hat{Y}_{HT}\right) = \sum_{\substack{i,i'=1\\ j(i)\neq j'(i')}}^{2}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\left(\frac{\pi_{j(i)}\pi_{j'(i')}-\pi_{j(i),j'(i')}}{\pi_{j(i),j'(i')}}\right)\left(\frac{y_{ij}}{\pi_{j(i)}}-\frac{y_{i'j'}}{\pi_{j'(i')}}\right)^{2}$$

$$= 2.2949\times10^{6}$$

with relative standard error of 426%, 37% and 30%, respectively. In this case, a nonpreferred sample will contain nonpreferred units more than or equal to $n\left(\frac{N_2}{N}\right) = \frac{3}{4}$, i.e., c = 1,2,or 3. Therefore, the accepted sample is a nonpreferred one since it consists of 1 nonpreferred unit.

CHAPTER 5

MORE ILLUSTRATIONS

5.1 An Example

A real-life population of steel production factories, is chosen for illustration. The studied variable, y, in this case is the production quantity of a factory in metric-tons which correlates to the investment fund, x, of the factory. The total number of factories, N, is 112. Each factory is a sampling unit. The size of each factory is its investment as shown in Appendix C. These sampling units are divided into two groups, preferred group with the first 106 factories and nonpreferred group with the last 6 factories. The classification of these sampling units is based on the location of the factories. A sampling unit will be considered a nonpreferred unit if it is located in Kanchana Buri, Chon Buri, Rayong, Chachoengsao and Sara Buri provinces.

A sample of 20 factories is to be drawn by the proposed controlled sampling scheme from the population. From Definition 1.2, a sample in this case is defined to be a nonpreferred sample if it contains nonpreferred units more than or equal to

$n \frac{N_2}{N} = \frac{15}{14}$ i.e., $c = 2,3,...,6$.

Results of the sampling process are illustrated in Table 5.1. By using (3.1), the values of X'_{ij}, for $i = 1,2$, $j = 1,2,...,N_i$ are calculated as shown in Appendix D.

Table 5.1 Stages in the Sample Selection by the Proposed Sampling Scheme in the Population

Round No.	A Sample (s)	$(x'_{..})_s$	r
1	$U_{72(1)}, U_{42(1)}, U_{40(1)}, U_{97(1)}, U_{81(1)}, U_{33(1)}, U_{35(1)}, U_{64(1)}, U_{23(1)}, U_{107(2)}, U_{2(1)}, U_{78(1)}, U_{74(1)}, U_{94(1)}, U_{63(1)}, U_{31(1)}, U_{28(1)}, U_{91(1)}, U_{108(2)}, U_{99(1)}$	105450.8953×10^4	177000×10^4
2	$U_{41(1)}, U_{26(1)}, U_{104(1)}, U_{89(1)}, U_{40(1)}, U_{68(1)}, U_{15(1)}, U_{44(1)}, U_{88(1)}, U_{62(1)}, U_{38(1)}, U_{7(1)}, U_{109(2)}, U_{112(2)}, U_{60(1)}, U_{107(2)}, U_{10(1)}, U_{29(1)}, U_{25(1)}, U_{67(1)}$	106459.2689×10^4	294000×10^4
3	$U_{2(1)}, U_{9(1)}, U_{95(1)}, U_{107(2)}, U_{88(1)}, U_{85(1)}, U_{69(1)}, U_{91(1)}, U_{89(1)}, U_{78(1)}, U_{93(1)}, U_{38(1)}, U_{49(1)}, U_{81(1)}, U_{64(1)}, U_{72(1)}, U_{65(1)}, U_{8(1)}, U_{3(1)}, U_{56(1)}$	107044.0500×10^4	49276×10^4

$(x'_{..})_s$: The sum of x'_{ij}'s in the sample
r: A random number

From Table 5.1 in the 1st round, we draw a simple random sampling without replacement sample of size of 20 factories in the first step. We obtain a sample containing the factories numbered 72, 42, 40, 97, 81, 33, 35, 64, 23, 107, 2, 78, 74,

94, 63, 31, 28, 91, 108, and 99. In the second step, $(x'_{..})_s = 105450.8953\times10^4$. In Step (3), suppose that a random number, r = 177000×10^4 is drawn between m = 1675.6162×10^4 and M = 406862.7057×10^4.[1] In Step (4), the sample is rejected because $r \geq (x'_{..})_s$. Hence, the first round is an ineffective one.

Then the process is returned to Steps (1) to (4) in the second round. For Step (1), suppose that a sample consists of the factories numbered 41, 26, 104, 89, 40, 68, 15, 44, 88, 62, 38, 7, 109, 112, 60, 107, 10, 29, 25 and 67. Then by Step (2), $(x'_{..})_s$ = 106459.2689×10^4. In Step (3), if a random number, r = 294000×10^4 is drawn between 1675.6162×10^4 and 406862.7057×10^4, then in Step (4), the sample is also rejected because $r \geq (x'_{..})_s$.

Similarly, the process is repeated from Steps (1) to (4) in the third round. If in the first step, a sample consisting of the factories numbered 2, 9, 95, 107, 88, 85, 69, 91, 89, 78, 93, 38, 49, 81, 64, 72, 65, 8, 3 and 56 is chosen, then the second step, $(x'_{..})_s$ = 107044.0500×10^4. In the third step, if a random number, r = 49276×10^4 is drawn between 1675.6162×10^4 and 406862.7057×10^4, then the fourth step, the sample is accepted because $r \leq (x'_{..})_s$.

Here, the accepted sample was attained in the third round. Hence, this round of sampling is an effective one. The sampling units in the accepted sample are shown

---

[1] m is a largest integer smaller than the sum of n smallest x'_{ij}'s and M is a smallest integer larger than or equal to the sum of n largest x'_{ij}'s in the population

in Table 5.2. From the accepted sample, an unbiased estimate of the total production quantity and their unbiased variance estimates are calculated. The probability for obtaining a nonpreferred sample under the controlled plan and under the uncontrolled plan will be compared.

Table 5.2 Production Total, the 1st Order Inclusion Probability and the Adjusted Value of the Investment Fund for each Element in an Accepted Sample of 20 Factories

Factory Number	$U_{j(i)}$	Production Total(y_{ij})	$\pi_{j(i)}$	The Adjusted Value of the Investment Fund (x'_{ij})
2	$U_{2(1)}$	600	0.1725	9396844.1570
9	$U_{9(1)}$	36	0.1714	2582338.8520
95	$U_{95(1)}$	5000	0.1823	71731634.7900
107	$U_{107(2)}$	49500	0.1711	70000.0000
88	$U_{88(1)}$	12000	0.1891	114770615.7000
85	$U_{85(1)}$	48	0.1715	3155861.9650
69	$U_{69(1)}$	15	0.1712	1291169.4260
91	$U_{91(1)}$	1000	0.1729	12085345.8300
89	$U_{89(1)}$	374	0.1725	9325112.5220
78	$U_{78(1)}$	30000	0.2079	233845129.4000
93	$U_{93(1)}$	1000	0.1730	12409572.8200
38	$U_{38(1)}$	36000	0.2195	307728713.2000
49	$U_{49(1)}$	96	0.1718	5021214.43500
81	$U_{81(1)}$	3200	0.1780	44473616.5700
64	$U_{64(1)}$	150	0.1719	5738530.7830
72	$U_{72(1)}$	30000	0.1974	167134709.1000
65	$U_{65(1)}$	5000	0.1801	57385307.8300
8	$U_{8(1)}$	90	0.1718	4877751.1650
3	$U_{3(1)}$	57.60	0.1716	3600928.0660
56	$U_{56(1)}$	60	0.1716	3816122.9710

Also notice that the accepted sample is a preferred sample since there is only one nonpreferred unit.

From Appendix D, the total sum of X'_{ij} i.e., $X'_{..} = X_{..}$ equals 5352.0061×10^{6}.

Using (4.6), the proposed unbiased estimator of the population total is

$$\hat{Y}_{..} = \frac{nX_{..} - Nm/n}{(x'_{12} + x'_{19} + \ldots + x'_{1,56} - m)}(y_{12} + y_{19} + \ldots + y_{1,56})$$

$$= 869{,}438.1339 \text{ metric-tons.}$$

The Horvitz-Thompson estimate of the population total according to (4.1) is,

$$\hat{Y}_{HT} = \frac{y_{12}}{\pi_{2(1)}} + \frac{y_{19}}{\pi_{9(1)}} + \ldots + \frac{y_{1,56}}{\pi_{56(1)}}$$

$$= 906{,}684.7796 \text{ metric-tons.}$$

Using (4.8), the estimate of the variance of $\hat{Y}_{..}$ is, $v\left(\hat{Y}_{..}\right) = 5.7589\times10^{9}$, where its coefficient of variation of $\hat{Y}_{..}$ is, $c.v.\left(\hat{Y}_{..}\right) = 0.087$.

According to (4.4), the estimate of the variance of $\hat{Y}_{HT}$ is $v_{HT}\left(\hat{Y}_{HT}\right)$ $= 1.0089\times10^{11}$ where its coefficient of variation of $\hat{Y}_{HT}$ is, $c.v.\left(\hat{Y}_{HT}\right) = 0.350$.

And the estimate of the variance of $\hat{Y}_{HT}$ in (4.5) is, $v_{SYG}\left(\hat{Y}_{HT}\right) = 9.9586\times10^{10}$ where its coefficient of variation of $\hat{Y}_{HT}$ is, $c.v.\left(\hat{Y}_{HT}\right) = 0.348$.

The comparison of the probability of getting a nonpreferred sample under the controlled plan and under the uncontrolled plan is considered. By using (3.8) and (3.9)

these probabilities are calculated for each c = 2,3,4,5 and 6 where c is the number of nonpreferred units in a nonpreferred sample as shown in Table 5.3.

Table 5.3 The Probability of Getting a Nonpreferred Sample for Possible c's Under the Controlled Plan and the Uncontrolled Plan

c	P^*	P
2	0.2887	0.3587
3	0.0759	0.1050
4	0.0076	0.0159
5	0.00048	0.0012
6	0.000012	0.000036

c: The number of nonpreferred units contains in a nonpreferred sample
P^*: The probability of getting a nonpreferred sample for each c under the controlled plan
P: The probability of getting a nonpreferred sample for each c under the uncontrolled plan

From Table 5.3, under the controlled plan with c = 2, the probability of obtaining a nonpreferred sample according to (3.8), is 0.2887.

Similarly, the probabilities with c = 3, 4, 5 and 6 under the controlled plan are 0.0759, 0.0076, 0.00048 and 0.000012, respectively.

Under the uncontrolled plan with c = 2, 3, 4, 5 and 6, the probabilities of obtaining a nonpreferred sample according to (3.9), are 0.3587, 0.1050, 0.0159, 0.0012 and 0.000036, respectively.

The probability for obtaining a nonpreferred sample under the controlled plan is less than that of the uncontrolled one for any value of c under consideration. Therefore, the proposed sampling scheme is admissible according to definition 3.3.

5.2 Some More Natural Populations

From Table 5.4, the seven different natural populations can be grouped into two main types: daily suspended sediment (pop. nos. 1, 2, 3, 4) and factories (pop. nos. 5, 6, 7). Population types 1 and 2 are from the Ping River during the years 1993-1995 while types 3 and 4 are from the Huai Phak River during the years 1986-1997. Of the three factory populations, type 5 is crop factories numbered 88 while type 6 is steel factories numbered 112 and type 7 is noodle factories numbered 84.

The study variable, y, for population type 1 and 3 is the weight of suspended sediment (tons) which is correlated to the river discharge rate (cubic-metre per second, or c.m.s.), an auxiliary variable, x, and for population type 2 and 4 is the height of water level (metre) which is correlated to the river discharge rate (c.m.s.), an auxiliary variable, x. For population type 5 and 6, crop factories and steel factories, the study variable, y, is the production quantity (metric-tons) which is correlated to the investment fund, an auxiliary variable, x, while for population type 7, the production quantity is measured in kilograms.

For population types 1,2,3,4, each day of the data collecting constitutes one sampling unit and the size of each sampling unit is its river discharge rate. The number, N, of data collecting days is 136 days for types 1 and 2 and 196 days for types 3 and 4. For population types 5, 6, 7, each factory is a sampling unit and the size

of each sampling is its investment. For details, see Appendix E for types 1 and 2, Appendix F for types 3 and 4, Appendix G for type 5, Appendix H for type 6 and Appendix I for type 7.

Table 5.4 Descriptions of Natural Populations Used in Studying the Proposed Sampling Scheme

Pop No.	Types of Population	Y	X	N	N_1	N_2 [2]	Corr. [3]	c [4]
1	Daily Suspended Sediment from Ping River	Suspended Sediment (ton)	River Discharge (c.m.s)	136 Days	131	5	0.93	1
2	Daily Suspended Sediment from Ping River	Gage Height (metre)	River Discharge (c.m.s)	136 Days	131	5	0.59	1
3	Daily Suspended Sediment from Huai Phak River	Suspended Sediment (ton)	River Discharge (c.m.s)	196 Days	188	8	0.80	1
4	Daily Suspended Sediment from Huai Phak River	Gage Height (metre)	River Discharge (c.m.s)	196 Days	188	8	0.87	1
5	Crop Factory	Production (metric-ton)	Investment (baht)	88 Factories	78	10	0.86	3
6	Steel Factory (Central Part)	Production (metric-ton)	Investment (baht)	112 Factories	106	6	0.87	2
7	Noodle Factory	Production (kilogram)	Investment (baht)	84 Factories	79	5	0.41	2

In the population of type 1, of the total sampling units of 136 days, it was found that the first 131 were preferred units and the remaining 5 nonpreferred units,

[2] N represents a population size N_1 represents a number of preferred units and N_2 represents a number of nonpreferred units

[3] corr. stands for a correlation coefficient between y and x

[4] c represents a minimum number of nonpreferred units allowed in a nonpreferred sample

and a correlation coefficient between variable y and variable x is 0.93. A nonpreferred unit for this type is a unit that takes an irregular river discharge rate. And a minimum number of nonpreferred units allowed in this nonpreferred sample is 1.

The finding in the population of type 2 is different from that in type 1 only in a correlation coefficient between variable y and variable x, which is 0.59.

In the population of type 3, of the total sampling units of 196 days, it was found that the first 188 were preferred units and the remaining 8 nonpreferred units, and a correlation coefficient between variable y and variable x is 1.80. A nonpreferred unit for this is a unit that takes an irregular river discharge rate. And a minimum number of nonpreferred units allowed in this nonpreferred sample is 1.

The finding in the population of type 4 is different from that in 3 only in a correlation coefficient between variable y and variable x, which is 0.87.

In the population of type 5, crop factory, of the total sampling units of 88, it was found that the first 188 were preferred units and the remaining 10 nonpreferred units, and a correlation coefficient between variable y and variable x is 0.86. A nonpreferred unit for this type is a unit that does its trade on crops of fewer than three kinds. And a minimum number of nonpreferred units allowed in this nonpreferred sample is 3.

In the population of type 6, steel factory, of the total sampling units of 112, it was found that the first 106 were preferred units and the remaining 6 nonpreferred units, and a correlation coefficient between variable y and variable x is 0.87. A nonpreferred unit for this type is a unit that is located far away from the group of

factories studied. A minimum number of nonpreferred units allowed in this nonpreferred sample is 2.

In the last type of population, noodle factory, of the total sampling units of 84, it was found that the first 79 were preferred units and the remaining 5 were nonpreferred units, and a correlation coefficient between variable y and variable x is 0.41. A nonpreferred unit for this type is aunit that does not use raw material from that lacal area. A minimum number of nonpreferred units allowed in this nonpreferred sample is 2.

5.3 Numerical Results

A sample of size 20 from each of the populations described in Table 5.4 was drawn by the proposed sampling scheme. From each selected sample, the estimate of the population total, $\hat{Y}_{..}$, and the Horvitz-Thompson estimate, $\hat{Y}_{HT}$, are calculated. The estimates of variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ as given by different authors are also illustrated.

Results of the sample selection and the estimates and their variance estimates are also given.

In the first population as shown in Appendix E, the accepted sample is

$s = \{U_{70(1)}, U_{91(1)}, U_{99(1)}, U_{76(1)}, U_{125(1)}, U_{38(1)}, U_{6(1)}, U_{97(1)}, U_{63(1)}, U_{7(1)}, U_{116(1)}, U_{33(1)}, U_{69(1)}, U_{123(1)}, U_{42(1)}, U_{84(1)}, U_{53(1)}, U_{20(1)}, U_{104(1)}, U_{32(1)}\}$ where $\hat{Y}_{..} = 258{,}673.1606$ tons and $\hat{Y}_{HT} = 286{,}441.7067$ tons . The estimates of the variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ are

$v\left(\hat{Y}_{..}\right) = 7.4765\times10^{8}$ with c.v.$\left(\hat{Y}_{..}\right) = 0.105$ for the proposed variance estimator and $v_{HT}\left(\hat{Y}_{HT}\right) = -1.2088\times10^{10}$ for the Horvitz-Thompson variance estimator. The variance estimate of Yates-Grundy and Sen is $v_{SYG}\left(\hat{Y}_{HT}\right) = 3.6084\times10^{11}$ with c.v.$\left(\hat{Y}_{HT}\right) =$ 2.097.

From Appendix E again, the accepted sample for the second population is $s = \{U_{80(1)}, U_{14(1)}, U_{109(1)}, U_{61(1)}, U_{133(2)}, U_{135(2)}, U_{24(1)}, U_{132(2)}, U_{21(1)}, U_{85(1)}, U_{36(1)}, U_{117(1)}, U_{62(1)}, U_{59(1)}, U_{8(1)}, U_{79(1)}, U_{1(1)}, U_{28(1)}\}$ where $\hat{Y}_{..} = 30{,}945.1676$ metres and $\hat{Y}_{HT} = 40{,}709.3457$ metres. The estimate of the variance of $\hat{Y}_{..}$ is $v\left(\hat{Y}_{..}\right) = 2.4960\times10^{9}$ with c.v.$\left(\hat{Y}_{..}\right) = 1.61$. The estimate of the variance of the Horvitz-Thompson estimator is $v_{HT}\left(\hat{Y}_{HT}\right) = -1.1245\times10^{9}$. The variance estimate of Yates-Grundy and Sen is $v_{SYG}\left(\hat{Y}_{HT}\right) = 3.2361\times10^{6}$ with c.v.$\left(\hat{Y}_{HT}\right) = 0.044$.

For the third population as in Appendix F, the accepted sample is $s = \{U_{92(1)}, U_{171(1)}, U_{82(1)}, U_{72(1)}, U_{101(1)}, U_{76(1)}, U_{131(1)}, U_{54(1)}, U_{89(1)}, U_{58(1)}, U_{175(1)}, U_{11(1)}, U_{132(1)}, U_{167(1)}, U_{146(1)}, U_{115(1)}, U_{51(1)}, U_{187(1)}, U_{177(1)}, U_{178(1)}\}$ where $\hat{Y}_{..} = 16{,}671.0709$ tons and $\hat{Y}_{HT} = 14{,}859.1921$ tons. The estimates of the variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ are $v\left(\hat{Y}_{..}\right) = 2.1444\times10^{6}$ with c.v.$\left(\hat{Y}_{..}\right) = 0.087$ for the proposed variance estimator and $v_{HT}\left(\hat{Y}_{HT}\right) = 3.3666\times10^{7}$ with c.v.$\left(\hat{Y}_{HT}\right) = 0.397$ for the Horvitz-Thompson variance estimator. For the Yates-Grundy and Sen variance estimate, $v_{SYG}\left(\hat{Y}_{HT}\right) = 3.2744\times10^{7}$ with c.v.$\left(\hat{Y}_{HT}\right) = 0.385$.

The fourth population in Appendix F again, the accepted sample is

$s = \{U_{179(1)}, U_{3(1)}, U_{96(1)}, U_{145(1)}, U_{156(1)}, U_{84(1)}, U_{135(1)}, U_{97(1)}, U_{175(1)}, U_{90(1)}, U_{110(1)}, U_{136(1)}, U_{10(1)}, U_{154(1)}, U_{139(1)}, U_{1(1)}, U_{116(1)}, U_{166(1)}, U_{12(1)}, U_{125(1)}\}$. The estimator of the population total are $\hat{Y}_{..} = 511.8418$ metres for the proposed estimator and $\hat{Y}_{HT} = 550.7101$ metres for the Horvitz-Thompson estimator. The estimates of the variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ are $v\left(\hat{Y}_{..}\right) = 2.0365\times10^{5}$ with $c.v.\left(\hat{Y}_{..}\right) = 0.881$ and $v_{HT}\left(\hat{Y}_{HT}\right) = 882.7589$ with $c.v.\left(\hat{Y}_{HT}\right) = 0.053$ for the Horvitz-Thompson variance estimator. The variance estimate of Yates-Grundy and Sen, $v_{SYG}\left(\hat{Y}_{HT}\right) = 826.2571$ with $c.v.\left(\hat{Y}_{HT}\right) = 0.052$.

For the fifth population as in Appendix G, the accepted sample is

$s = \{U_{57(1)}, U_{3(1)}, U_{15(1)}, U_{73(1)}, U_{16(1)}, U_{35(1)}, U_{8(1)}, U_{84(1)}, U_{45(1)}, U_{72(1)}, U_{24(1)}, U_{51(1)}, U_{55(1)}, U_{32(1)}, U_{29(1)}, U_{82(2)}, U_{42(1)}, U_{6(1)}, U_{80(2)}, U_{61(1)}\}$ where $\hat{Y}_{..} = 133,193.9206$ metric-tons and $\hat{Y}_{HT} = 208,585.4655$ metric-tons. The estimates of the variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ are $v\left(\hat{Y}_{..}\right) = 4.8540\times10^{7}$ with $c.v.\left(\hat{Y}_{..}\right) = 0.052$ for the proposed variance estimator and $v_{HT}\left(\hat{Y}_{HT}\right) = 9.1831\times10^{9}$, $c.v.\left(\hat{Y}_{HT}\right) = 0.459$ for the Horvitz-Thompson variance estimator. The estimate of the variance, $v_{SYG}\left(\hat{Y}_{HT}\right) = 8.9938\times10^{9}$ with $c.v.\left(\hat{Y}_{HT}\right) = 0.454$ for the Yates-Grundy and Sen variance estimator.

The sixth population as shown in Appendix H, the accepted sample is

$s = \{U_{39(1)}, U_{34(1)}, U_{85(1)}, U_{100(1)}, U_{79(1)}, U_{38(1)}, U_{71(1)}, U_{13(1)}, U_{36(1)}, U_{93(1)}, U_{91(1)}, U_{47(1)}, U_{92(1)}, U_{110(2)}, U_{16(1)}, U_{4(1)}, U_{77(1)}, U_{28(1)}, U_{33(1)}, U_{60(1)}\}$. The proposed estimator of the population total is $\hat{Y}_{..} = 471,952.2671$ metric-tons and the Horvitz-Thompson

estimator is $\hat{Y}_{HT}$ = 479,460.4996 metric-tons. The estimates of the variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ are $v(\hat{Y}_{..}) = 3.9337\times10^{9}$ with $c.v(\hat{Y}_{..}) = 0.132$ for the proposed variance estimator, $v_{HT}(\hat{Y}_{HT}) = 3.1348\times10^{10}$ with $c.v.(\hat{Y}_{HT}) = 0.369$ for the Horvitz-Thompson variance estimator and $v_{SYG}(\hat{Y}_{HT}) = 2.9600\times10^{10}$ with $c.v.(\hat{Y}_{HT}) = 0.358$ for the Yates-Grundy and Sen variance estimator.

For the last population as shown in Appendix I, the accepted sample is $s = \{U_{68(1)}, U_{13(1)}, U_{11(1)}, U_{41(1)}, U_{49(1)}, U_{78(1)}, U_{34(1)}, U_{43(1)}, U_{52(1)}, U_{15(1)}, U_{65(1)}, U_{61(1)}, U_{60(1)}, U_{69(1)}, U_{27(1)}, U_{62(1)}, U_{75(1)}, U_{64(1)}, U_{74(1)}, U_{23(1)}\}$. The proposed estimator of the population total is $\hat{Y}_{..}$ = 7,463,646.2582 kilograms. The Horvitz-Thompson estimator of the population total is $\hat{Y}_{HT}$ = 11,086,387.6750 kilograms. The estimates of the variance of $\hat{Y}_{..}$ and $\hat{Y}_{HT}$ are $v(\hat{Y}_{..}) = 1.2653\times10^{14}$ with $c.v.(\hat{Y}_{..}) = 1.507$ for the proposed variance estimator, $v_{HT}(\hat{Y}_{HT}) = 8.309610^{12}$ with $c.v.(\hat{Y}_{HT}) = 0.260$ for the Horvitz-Thompson variance estimator and $v_{SYG}(\hat{Y}_{HT}) = 8.459410^{12}$ with $c.v.(\hat{Y}_{HT}) =$ 0.262 for the Yates-Grundy and Sen variance estimator.

Table 5.5 below presents the results of the estimates of the population total and their variance estimates for the natural populations. From these results, the Horvitz-Thompson variance estimate may take negative values on the population numbered 1 and 2.

Table 5.5 The Estimates of the Population Total and their Variance Estimates for the Natural Populations

Pop No.	$\hat{Y}_{..}$	$\hat{Y}_{HT}$	$v\left(\hat{Y}_{..}\right)$ (c.v.$\left(\hat{Y}_{..}\right)$)	$v_{HT}\left(\hat{Y}_{HT}\right)$ (c.v.$\left(\hat{Y}_{HT}\right)$)	$v_{SYG}\left(\hat{Y}_{HT}\right)$ (c.v.$\left(\hat{Y}_{HT}\right)$)
1	258673.1606	286441.7067	7.4765×10^{8} (0.105)	-1.2088×10^{10}	3.6084×10^{11} (2.097)
2	30945.1676	40709.3457	2.4960×10^{9} (1.610)	-1.1245×10^{9}	3.2361×10^{6} (0.044)
3	16671.0707	14859.1921	2.1444×10^{6} (0.087)	3.3666×10^{7} (0.397)	3.2744×10^{7} (0.385)
4	511.8418	550.7101	2.0365×10^{5} (0.881)	882.7589 (0.053)	826.2571 (0.052)
5	133193.9206	208585.4655	4.8540×10^{7} (0.052)	9.1831×10^{9} (0.459)	8.9938×10^{9} (0.454)
6	471952.2971	479460.4996	3.9337×10^{9} (0.132)	3.1348×10^{10} (0.369)	2.9600×10^{10} (0.358)
7	7463646.2582	11086387.6750	1.2653×10^{14} (1.507)	8.3096×10^{12} (0.260)	8.4594×10^{12} (0.262)

For the populations numbered 2 and 7 which are low in values of the correlation coefficient, the relative standard error of the proposed estimator is 161% for the population numbered 2 and 150.7% for the population numbered 7. For both populations, the relative standard error of the Horvitz-Thompson estimator is 4.4% for the population numbered 2 and 26.2% for the population numbered 7. The populations that are high in values of the correlation coefficient consist of the populations numbered 1,3,4,5 and 6. The values of the relative standard error of the proposed estimator are 10.5%, 8.7%, 5.2% and 13.2% for the populations numbered 1,3,5 and 6 respectively. In contrast, for the population numbered 4, the relative

standard error is 88.1%. The values of the relative standard error for the Horvitz-Thompson estimator are 209.7%, 38.5% 45.4% and 35.8% for the populations numbered 1,3,5 and 6 respectively and 5.2% for the population numbered 4.

The probability of obtaining a nonpreferred sample for possible c's under the controlled plan and under the uncontrolled plan is presented in Table 5.6.

Table 5.6 Probability of Obtaining a Nonpreferred Sample for Possible c's Under the Controlled Plan, P^*, and the Uncontrolled Plan, P, for the Seven Natural Populations

Pop.	c	P^*	P
1	1	0.5819	0.6364
	2	0.1534	0.1904
	3	0.0205	0.0293
	4	0.0013	0.00220
	5	0.000032	0.000063
2	1	0.5819	0.6364
	2	0.1534	0.1904
	3	0.0205	0.0293
	4	0.0013	0.00220
	5	0.000032	0.000063
3	1	0.5755	0.6224
	2	0.1791	0.2160
	3	0.0326	0.0443
	4	0.0036	0.0055
	5	0.00023	0.00041
	6	2.0737×10^{-7}	1.8935×10^{-5}
	7	1.9807×10^{-7}	4.5105×10^{-7}
	8	1.6633×10^{-9}	4.3479×10^{-9}

Table 5.6 (Continued)

Pop.	c	P^*	P
	4	1	0.5755	0.6224
		2	0.1791	0.2160
		3	0.0326	0.0443
		4	0.0036	0.0055
		5	0.00023	0.00041
		6	2.0737×10^{-7}	1.8935×10^{-5}
		7	1.9807×10^{-7}	4.5105×10^{-7}
		8	1.6633×10^{-9}	4.3479×10^{-9}
	5	3	0.3821	0.5011
		4	0.1438	0.2212
		5	0.0366	0.0666
		6	0.0061	0.0133
		7	0.00067	0.0017
		8	4.4000×10^{-5}	1.5300×10^{-4}
		9	1.5640×10^{-6}	5.6157×10^{-6}
		10	2.2453×10^{-8}	9.6366×10^{-8}
	6	2	0.2887	0.3587
		3	0.0759	0.1050
		4	0.0219	0.0159
		5	0.00048	0.0012
		6	1.2000×10^{-5}	3.6000×10^{-5}
	7	2	0.3279	0.4451
		3	0.0766	0.1285
		4	0.0088	0.0183
		5	0.00039	0.00099

c: The number of nonpreferred units contained in a nonpreferred sample
P^*: The probability of getting a nonpreferred sample for each c under the controlled plan
P: The probability of getting a nonpreferred sample for each c under the uncontrolled plan

From Table 5.6, it is clear that the probability for obtaining a nonpreferred sample under the controlled plan is less than that of the uncontrolled plan for all populations and all restricted values of c's. Therefore, the proposed sampling scheme is admissible according to definition 3.3 for any population.

CHAPTER 6

SUMMARY

6.1 Summary and Conclusions

In this design, the method of controlled sampling with varying probability without replacement has been proposed. The parameter of interest is $Y_{..}$, the population total. The estimator of $Y_{..}$, $\hat{Y}_{..}$, and an unbiased estimator of the variance of the estimator, $v\left(\hat{Y}_{..}\right)$, are presented.

It is noted that our proposed method of sampling is practical and effective in deriving computational form of the first order inclusion probabilities of unit j in group i, $\pi_{j(i)}$, and the second order inclusion probabilities, $\pi_{j(i)j'(i')}$, of units j and j′ in groups i and i′, respectively, for $j(i) \neq j'(i')$.[1] It is also interesting to note that, through this method, the probability for obtaining all nonpreferred samples is reduced.

In addition, the estimator of $Y_{..}$, $\hat{Y}_{..}$, derived from the study is a ratio estimator, which indicates that this ratio estimator of the population total is both unbiased and consistent. The computational form of the ratio estimator is not complicated. The variance of the estimator can be expressed in two forms. Both of them are in the form

---

[1] $j(i) \neq j'(i')$ if and only if $i \neq i'$ or $j \neq j'$ where $j = 1,2,\ldots,N_i$, $j' = 1,2,\ldots,N_{i'}$, $i, i'=1,2$

of homogeneous quadratic function. The estimator of variance of the estimator is an unbiased estimator.

Finally, the study of the proposed sampling scheme using seven populations concluded that this sampling scheme can employ to several kinds of populations. Although relative efficiency of these variance estimators is unable to draw any conclusions, the value of the proposed variance estimator may take negative value less frequent than the variance of Horvitz-Thompson estimator, $v\left(\hat{Y}_{HT}\right)$. However, to solve the problem of nonnegativeness of variance of the estimator, Lanke (1974) and Rao and Vijayan (1977) suggested some nonnegative variance estimators appropriate to the ratio estimator.

6.2 Research Recommendations

Findings in this thesis have led to several recommendations for further research. The results should be used further for a stratified sampling and a multi-stage sampling. A large population, for example, may be stratified and the method of controlled sampling should be used in each stratum.

The nonnegativeness of variance estimator in this thesis should be further investigated. However, there are some existing concepts about the nonnegative variance estimator earlier studied. For example, Lanke (1974) used the idea of the nonnegative variance estimator based on the conditional probabilities of subsamples. Rao and Vijayan (1977) also used the superpopulation model to find out the nonnegative variance estimates, both of which are appropriate to be used as a ratio

estimator. Other writers have tried to set a negative estimate to be at zero value, which is unsatisfactory for a large true variance (Rao and Vijayan, 1977).

Further investigation should center around a study of properties of the estimator proposed in this dissertation. However, this investigation is to be effectively carried out if limitations on the use of personal computer should be minimized or eliminated.

Lastly, for further research on this matter some other alternative techniques of controlled sampling of the same type of the sampling scheme should be developed.

APPENDICES

APPENDIX A

Under the sampling scheme introduced in Chapter III

$$-a_{j(i)}(X'_{ij} - m/n) = \sum_{\substack{i'=1 \\ j'(i')\neq j(i)}}^{2} \sum^{N_{i'}}_{j'=1} a_{j(i)j'(i')}(X'_{i'j'} - m/n)$$

where $$a_{j(i)} = \left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{s \ni j(i)} \frac{1}{(x'_{..})_s - m} - 1 \right]$$ and

$$a_{j(i)j'(i')} = \left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{\substack{s \ni j(i), j'(i') \\ j(i)\neq j'(i')}} \frac{1}{(x'_{..})_s - m} - 1 \right] \text{ for } i, i' = 1,2 \text{ and } j = 1,2,\ldots, N_i ,$$

$j'=1,2,\ldots, N_{i'}$ and $j(i) \neq j'(i')$.

Proof $-a_{j(i)}(X'_{ij} - m/n)$

$$= -(X'_{ij} - m/n)\left[\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \sum_{s \ni j(i)} \frac{1}{(x'_{..})_s - m} - 1 \right]$$

$$= -\left(\frac{X_{..} - Nm/n}{\binom{N-1}{n-1}} \right) \sum_{s \ni j(i)} \frac{X'_{ij} - m/n}{(x'_{..})_s - m} + (X'_{ij} - m/n) + (X_{..} - Nm/n) - (X_{..} - Nm/n), [1]$$

---

[1] By expansion then add and subtract by $(X_{..} - Nm/n)$.

$$=\left(\frac{X_{..}-Nm/n}{\binom{N-1}{n-1}}\sum_{s\ni j(i)}1\right)-\frac{X_{..}-Nm/n}{\binom{N-1}{n-1}}\sum_{s\ni j(i)}\frac{X'_{ij}-m/n}{(x'_{..})_s-m}-(X_{..}-Nm/n)+(X'_{ij}-m/n),^{[2]}$$

$$=\frac{X_{..}-Nm/n}{\binom{N-1}{n-1}}\sum_{s\ni j(i)}\frac{1}{(x'_{..})_s-m}\left\{\left((x'_{..})_s-m\right)-\left(X'_{ij}-m/n\right)\right\}$$

$$-\left(\left(X_{..}-Nm/n\right)-\left(X'_{ij}-m/n\right)\right)$$

$$=\frac{X_{..}-Nm/n}{\binom{N-1}{n-1}}\sum_{s\ni j(i)}\frac{1}{(x'_{..})_s-m}\left(\sum_{\substack{i'=1\,j'=1\\ j'(i')\neq j(i)\in s}}^{2\;N_{i'}}(X'_{i'j'}-m/n)\right)-\sum_{\substack{i'=1\,j'=1\\ j'(i')\neq j(i)}}^{2\;N_{i'}}(X'_{i'j'}-m/n),^{[3]}$$

$$=\sum_{\substack{i'=1\,j'=1\\ j'(i')\neq j(i)}}^{2\;N_{i'}}\left[\frac{X_{..}-Nm/n}{\binom{N-1}{n-1}}\sum_{\substack{s\ni j(i),j'(i')\\ j(i)\neq j'(i')}}\frac{1}{(x'_{..})_s-m}-1\right](X'_{i'j'}-m/n)$$

$$=\sum_{\substack{i'=1\,j'=1\\ j'(i')\neq j(i)}}^{2\;N_{i'}}a_{j(i)j'(i')}(X'_{i'j'}-m/n).$$

---

[2] since $\sum_{s\ni j(i)}1=\binom{N-1}{n-1}$.

[3] $\left((x'_{..})_s-m\right)-\left(X'_{ij}-m/n\right)=(x'_{11}+\ldots+x'_{2n_2}-m)-\left(X'_{ij}-m/n\right)=\left(\sum_{\substack{i'=1\,j'=1\\ j'(i')\neq j(i)\in s}}^{2\;N_{i'}}(X'_{i'j'}-m/n)\right)$

$$\left(X_{..}-Nm/n\right)-\left(X'_{ij}-m/n\right)=(X'_{11}+\ldots+X'_{2n_2}-Nm/n)-\left(X'_{ij}-m/n\right)=\sum_{\substack{i'=1\,j'=1\\ j'(i')\neq j(i)}}^{2\;N_{i'}}(X'_{i'j'}-m/n).$$

APPENDIX B

Another expression of the variance of the estimator of the population total is

$$V\left(\hat{Y}_{..}\right) = \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} b_{j(i)j'(i')} \left(X'_{ij} - \frac{m}{n}\right)\left(X'_{i'j'} - \frac{m}{n}\right)\left(\frac{Y_{ij}}{X'_{ij} - m/n} - \frac{Y_{i'j'}}{X'_{i'j'} - m/n}\right)^2$$

where $b_{j(i)j'(i')} = - a_{j(i)j'(i')}$.

Proof From the right hand side of (4.7), we have

$$\sum_{i=1}^{2} \sum_{j=1}^{N_i} a_{j(i)} Y_{ij}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} Y_{ij} Y_{i'j'}$$

$$= \sum_{j=1}^{N_1} a_{j(1)} Y_{1j}^2 + \sum_{j'=1}^{N_2} a_{j'(2)} Y_{2\,j'}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} Y_{ij} Y_{i'j'}$$

$$= - \sum_{j=1}^{N_1} \left(- a_{j(1)}\right) Y_{1j}^2 - \sum_{j'=1}^{N_2} \left(- a_{j'(2)}\right) Y_{2\,j'}^2 + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} Y_{ij} Y_{i'j'}$$

$$= \sum_{j=1}^{N_1} \frac{Y_{1j}^2}{X'_{1j} - m/n} \left(- \sum_{\substack{i'=1 \\ j'(i')\neq j(1)}}^{2} \sum_{j'=1}^{N_{i'}} a_{j(1)j'(i')} (X'_{i'j'} - m/n) \right)$$

$$+ \sum_{j'=1}^{N_2} \frac{Y_{2\,j'}^2}{X'_{2\,j'} - m/n} \left(- \sum_{\substack{i=1 \\ j(i)\neq j'(2)}}^{2} \sum_{j=1}^{N_i} a_{j'(2)j(i)} (X'_{ij} - m/n) \right) + \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} Y_{ij} Y_{i'j'}, \quad [1]$$

---

[1] See Appendix A. $- a_{j(i)} (X'_{ij} - m/n) = \sum_{\substack{i'=1 \\ j'(i')\neq j(i)}}^{2} \sum_{j'=1}^{N_{i'}} a_{j(i)j'(i')} (X'_{i'j'} - m/n)$.

$$= \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} \left(\frac{Y_{ij}}{X'_{ij} - m/n} \right)^2 \left(- a_{j(i)j'(i')} (X'_{i'j'} - m/n)(X'_{ij} - m/n) \right)$$

$$+ \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} \left(\frac{Y_{i'j'}}{X'_{i'j'} - m/n} \right)^2 \left(- a_{j(i)j'(i')} (X'_{ij} - m/n)(X'_{i'j'} - m/n) \right)$$

$$+ \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} \left(\frac{Y_{ij}}{X'_{ij} - m/n} \frac{Y_{i'j'}}{X'_{i'j'} - m/n} \right) \left(a_{j(i)j'(i')} (X'_{ij} - m/n)(X'_{i'j'} - m/n) \right)$$

$$= \sum_{\substack{i,i'=1 \\ j(i)\neq j'(i')}}^{2} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_{i'}} b_{j(i)j'(i')} \left(X'_{ij} - \frac{m}{n} \right) \left(X'_{i'j'} - \frac{m}{n} \right) \left(\frac{Y_{ij}}{X'_{ij} - m/n} - \frac{Y_{i'j'}}{X'_{i'j'} - m/n} \right)^2 .$$

APPENDIX C

Investment funds of 112 steel production factories

Factory Number	X_{ij}	Factory Number	X_{ij}	Factory Number	X_{ij}
1	4000000	41	28800000	81	31000000
2	6550000	42	1900000	82	14600000
3	2510000	43	60000000	83	65000000
4	2100000	44	15492244	84	30000000
5	2250000	45	75000000	85	2199770
6	1000000	46	82000000	86	4050000
7	3600000	47	4000000	87	10500000
8	3400000	48	6000000	88	80000000
9	1800000	49	3500000	89	6500000
10	2700000	50	3500000	90	4300000
11	7500000	51	3750000	91	8424000
12	60000000	52	3000000	92	20024000
13	4950000	53	6700000	93	8650000
14	1166000	54	16300000	94	22015000
15	4000000	55	600000	95	50000000
16	2390000	56	17000000	96	24000000
17	2525000	57	2660000	97	30000000
18	10500000	58	62000000	98	10500000
19	500000	59	5700000	99	20000000
20	430000	60	39984400	100	3700000
21	10850000	61	300000	101	1000000
22	70000	62	4800000	102	98000000
23	6500000	63	220000	103	7450000
24	4000000	64	86000000	104	30000000
25	1000000	65	4000000	105	20700000
26	4000000	66	40000000	106	20000000
27	51700000	67	5000000	107	38000000
28	40000000	68	7500000	108	400000
29	13600000	69	2500000	109	90006000
30	84062087	70	8000000	110	53000
31	3250000	71	75200000	111	510000
32	14000000	72	11650000	112	34000000
33	6000000	73	3400000		
34	19500000	74	8000000		
35	301000	75	73900000		
36	1200000	76	13000000		
37	1200000	77	34000000		
38	21450000	78	16300000		
39	38000000	79	19000000		
40	150000	80	7400000		

X_{ij} : investment funds of the j^{th} unit in group i

APPENDIX D

The adjusted value of the investment fund of 112 steel production factories

Group No.(i)	Unit No.(j)	Factory No.	the Adjusted Value of the Investment Funds (X'_{ij})
1	1	1	5738530.780
	2	2	9396844.150
	3	3	3600928.070
	4	4	3012728.660
	5	5	3227923.560
	6	6	1434632.690
	7	7	5164677.700
	8	8	4877751.160
	9	9	2582338.850
	10	10	3873508.280
	11	11	10759745.220
	12	12	86077961.740
	13	13	7101431.840
	14	14	1672781.720
	15	15	5738530.780
	16	16	3428772.140
	17	17	3622447.560
	18	18	15063643.300
	19	19	717316.350
	20	20	616892.060
	21	21	15565764.750
	22	22	100424.290
	23	23	9325112.520
	24	24	5738530.780
	25	25	1434632.690
	26	26	5738530.780
	27	27	74170510.370
	28	28	57385307.830
	29	29	19511004.660
	30	30	120598218.480
	31	31	4662556.260
	32	32	20084857.740
	33	33	8607796.170
	34	34	27975337.570
	35	35	431824.440
	36	36	1721559.230
	37	37	94685757.920
	38	38	307728713.230

Appendix D (Continued)

Group No.(i)	Unit No.(j)	Factory No.	the Adjusted Value of the Investment Funds (X'_{ij})
1	39	39	54516042.440
	40	40	2151949.040
	41	41	413174216.370
	42	42	2725802.120
	43	43	86077961.740
	44	44	22225679.770
	45	45	107597452.180
	46	46	117639881.050
	47	47	5738530.780
	48	48	8607796.170
	49	49	5021214.430
	50	50	5379872.610
	51	51	4303898.090
	52	52	9612039.060
	53	53	23384512.940
	54	54	8607796.170
	55	55	24388755.830
	56	56	3816122.970
	57	57	88947227.130
	58	58	8177406.360
	59	59	573629275.590
	60	60	430389.810
	61	61	6886236.940
	62	62	315619.190
	63	63	123378411.830
	64	64	5738530.780
	65	65	57385307.830
	66	66	7173163.480
	67	67	10759745.220
	68	68	3586581.740
	69	69	1291169.430
	70	70	11477061.560
	71	71	107884378.720
	72	72	167134709.050
	73	73	4877751.160
	74	74	11477061.560
	75	75	106019356.210
	76	76	18650225.040
	77	77	487775116.550
	78	78	233845129.400
	79	79	27258021.220
	80	80	10616281.950
	81	81	44473613.570

Appendix D (Continued)

Group No.(i)	Unit No.(j)	Factory No.	the Adjusted Value of the Investment Funds (X'_{ij})
1	82	82	20945637.360
	83	83	93251125.220
	84	84	43038980.870
	85	85	3155861.960
	86	86	5810262.420
	87	87	15063643.300
	88	88	114770615.660
	89	89	9325112.520
	90	90	6168920.590
	91	91	12085345.830
	92	92	28727085.100
	93	93	12409572.820
	94	94	31583438.800
	95	95	71731634.790
	96	96	34431184.700
	97	97	43038980.870
	98	98	15063643.300
	99	99	286926539.140
	100	100	5308140.970
	101	101	1434632.700
	102	102	140594004.180
	103	103	10688013.580
	104	104	43038980.870
	105	105	296968968.010
	106	106	28692653.910
2	1	107	70000.000
	2	108	70000.000
	3	109	70000.000
	4	110	70000.000
	5	111	70000.000
	6	112	70000.000

APPENDIX E

Population Number 1 and 2

The river discharge rate, X_{ij}, the adjusted value of the river discharge rate, X'_{ij}, of 136 daily suspended sediment from the Ping River, (c.m.s. : cubic-metre per second)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	1	$U_{1(1)}$	6.954	7.757751
	2	$U_{2(1)}$	27.953	31.18384
	3	$U_{3(1)}$	9.923	11.06991
	4	$U_{4(1)}$	56.852	63.42302
	5	$U_{5(1)}$	73.393	81.87585
	6	$U_{6(1)}$	32.571	36.33559
	7	$U_{7(1)}$	23.992	26.76502
	8	$U_{8(1)}$	25.137	28.04236
	9	$U_{9(1)}$	17.877	19.94324
	10	$U_{10(1)}$	16.117	17.97982
	11	$U_{11(1)}$	12.512	13.95815
	12	$U_{12(1)}$	19.172	21.38792
	13	$U_{13(1)}$	13.953	15.5657
	14	$U_{14(1)}$	12.969	14.46797
	15	$U_{15(1)}$	6.519	7.272474
	16	$U_{16(1)}$	4.388	4.89517
	17	$U_{17(1)}$	10.867	12.12302
	18	$U_{18(1)}$	6.243	6.964573
	19	$U_{19(1)}$	7.719	8.611171
	20	$U_{20(1)}$	8.094	9.029514
	21	$U_{21(1)}$	17.601	19.63534

Appendix E (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	22	$U_{22(1)}$	58.14	64.85989
	23	$U_{23(1)}$	56.386	62.90316
	24	$U_{24(1)}$	48.194	53.76432
	25	$U_{25(1)}$	30.305	33.80769
	26	$U_{26(1)}$	34.39	38.36484
	27	$U_{27(1)}$	29.944	33.40496
	28	$U_{28(1)}$	28.292	31.56202
	29	$U_{29(1)}$	270.524	301.7915
	30	$U_{30(1)}$	132.352	147.6494
	31	$U_{31(1)}$	277.262	309.3083
	32	$U_{32(1)}$	282.573	315.2331
	33	$U_{33(1)}$	183.92	205.1777
	34	$U_{34(1)}$	126.551	141.1779
	35	$U_{35(1)}$	121.216	135.2263
	36	$U_{36(1)}$	78.076	87.10011
	37	$U_{37(1)}$	71.299	79.53982
	38	$U_{38(1)}$	64.274	71.70286
	39	$U_{39(1)}$	79.242	88.40088
	40	$U_{40(1)}$	74.59	83.2112
	41	$U_{41(1)}$	57.566	64.21955
	42	$U_{42(1)}$	41.45	46.24084
	43	$U_{43(1)}$	38.608	43.07036
	44	$U_{44(1)}$	32.604	36.37241
	45	$U_{45(1)}$	15.963	17.80802
	46	$U_{46(1)}$	13.161	14.68216
	47	$U_{47(1)}$	8.913	9.943175
	48	$U_{48(1)}$	11.242	12.54136

Appendix E (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	49	$U_{49(1)}$	43.151	48.13844
	50	$U_{50(1)}$	21.063	23.49749
	51	$U_{51(1)}$	22.069	24.61976
	52	$U_{52(1)}$	100.648	112.281
	53	$U_{53(1)}$	48.278	53.85803
	54	$U_{54(1)}$	35.742	39.8731
	55	$U_{55(1)}$	125.044	139.4967
	56	$U_{56(1)}$	155.132	173.0623
	57	$U_{57(1)}$	165.94	185.1195
	58	$U_{58(1)}$	142.257	158.6992
	59	$U_{59(1)}$	171.121	190.8994
	60	$U_{60(1)}$	157.219	175.3906
	61	$U_{61(1)}$	95.631	106.6841
	62	$U_{62(1)}$	56.156	62.64658
	63	$U_{63(1)}$	24.48	27.30943
	64	$U_{64(1)}$	23.225	25.90937
	65	$U_{65(1)}$	23.226	25.91049
	66	$U_{66(1)}$	12.929	14.42335
	67	$U_{67(1)}$	12.724	14.19465
	68	$U_{68(1)}$	13.847	15.44745
	69	$U_{69(1)}$	40.91	45.63843
	70	$U_{70(1)}$	31.011	34.59529
	71	$U_{71(1)}$	27.353	30.51449
	72	$U_{72(1)}$	21.791	24.30963
	73	$U_{73(1)}$	35.71	39.8374
	74	$U_{74(1)}$	34	37.92976
	75	$U_{75(1)}$	50.095	55.88504

Appendix E (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	76	$U_{76(1)}$	81.176	90.55842
	77	$U_{77(1)}$	49.87	55.63403
	78	$U_{78(1)}$	47.049	52.48698
	79	$U_{79(1)}$	53.641	59.84089
	80	$U_{80(1)}$	60.498	67.49043
	81	$U_{81(1)}$	54.512	60.81256
	82	$U_{82(1)}$	99.936	111.4867
	83	$U_{83(1)}$	123.308	137.5601
	84	$U_{84(1)}$	319.5	356.4282
	85	$U_{85(1)}$	94.973	105.9501
	86	$U_{86(1)}$	66.304	73.96749
	87	$U_{87(1)}$	69.02	76.99741
	88	$U_{88(1)}$	66.842	74.56768
	89	$U_{89(1)}$	35	39.04534
	90	$U_{90(1)}$	93.654	104.4786
	91	$U_{91(1)}$	40.052	44.68126
	92	$U_{92(1)}$	32.869	36.66804
	93	$U_{93(1)}$	27.5	30.67848
	94	$U_{94(1)}$	22.8	25.43525
	95	$U_{95(1)}$	12.14	13.54316
	96	$U_{96(1)}$	12.14	13.54316
	97	$U_{97(1)}$	9.66	10.77651
	98	$U_{98(1)}$	9.66	10.77651
	99	$U_{99(1)}$	9.04	10.08485
	100	$U_{100(1)}$	8.42	9.393193
	101	$U_{101(1)}$	7.47	8.333391
	102	$U_{102(1)}$	7.47	8.333391

Appendix E (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	103	$U_{103(1)}$	8.42	9.393193
	104	$U_{104(1)}$	9.04	10.08485
	105	$U_{105(1)}$	17.096	19.07198
	106	$U_{106(1)}$	11.563	12.89947
	107	$U_{107(1)}$	12.243	13.65806
	108	$U_{108(1)}$	10.766	12.01035
	109	$U_{109(1)}$	9.37	10.453
	110	$U_{110(1)}$	33.774	37.67764
	111	$U_{111(1)}$	15.695	17.50905
	112	$U_{112(1)}$	20.202	22.53697
	113	$U_{113(1)}$	18.917	21.10345
	114	$U_{114(1)}$	12.238	13.65248
	115	$U_{115(1)}$	11.769	13.12927
	116	$U_{116(1)}$	160.983	179.5896
	117	$U_{117(1)}$	27.018	30.14077
	118	$U_{118(1)}$	32.888	36.68923
	119	$U_{119(1)}$	110.929	123.7503
	120	$U_{120(1)}$	66.362	74.0322
	121	$U_{121(1)}$	31.965	35.65955
	122	$U_{122(1)}$	98.199	109.549
	123	$U_{123(1)}$	86.371	96.35386
	124	$U_{124(1)}$	47.353	52.82611
	125	$U_{125(1)}$	57.125	63.72757
	126	$U_{126(1)}$	45.635	50.90955
	127	$U_{127(1)}$	33.557	37.43556
	128	$U_{128(1)}$	26.846	29.94889
	129	$U_{129(1)}$	8.393	9.363073

Appendix E (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	130	$U_{130(1)}$	7.145	7.970827
	131	$U_{131(1)}$	9.913	11.05876
2	1	$U_{1(2)}$	4.008	2.649
	2	$U_{2(2)}$	3.732	2.649
	3	$U_{3(2)}$	403.737	2.649
	4	$U_{4(2)}$	399.504	2.649
	5	$U_{5(2)}$	2.649	2.649

APPENDIX F

Population Number 3 and 4

The river discharge rate, X_{ij}, the adjusted value of the river discharge rate, X'_{ij}, of 196 daily suspended sediment from the Huai Phak River, (c.m.s. : cubic-metre per second)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	1	$U_{1(1)}$	41.829	46.80229
	2	$U_{2(1)}$	16.374	18.3208
	3	$U_{3(1)}$	22.721	25.42243
	4	$U_{4(1)}$	10.266	11.48658
	5	$U_{5(1)}$	4.891	5.472519
	6	$U_{6(1)}$	4.551	5.092095
	7	$U_{7(1)}$	0.275	0.307696
	8	$U_{8(1)}$	0.201	0.224898
	9	$U_{9(1)}$	0.798	0.892879
	10	$U_{10(1)}$	0.91	1.018195
	11	$U_{11(1)}$	0.821	0.918613
	12	$U_{12(1)}$	12	13.42675
	13	$U_{13(1)}$	57	63.77706
	14	$U_{14(1)}$	2.305	2.579055
	15	$U_{15(1)}$	16.792	18.7885
	16	$U_{16(1)}$	25.427	28.45016
	17	$U_{17(1)}$	1.737	1.943522
	18	$U_{18(1)}$	2.3	2.57346
	19	$U_{19(1)}$	34.399	38.4889
	20	$U_{20(1)}$	41.274	46.18131
	21	$U_{21(1)}$	16.976	18.99438

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	22	$U_{22(1)}$	12.954	14.49418
	23	$U_{23(1)}$	13.969	15.62986
	24	$U_{24(1)}$	36.974	41.37005
	25	$U_{25(1)}$	16.62	18.59605
	26	$U_{26(1)}$	10.372	11.60519
	27	$U_{27(1)}$	7.107	7.951993
	28	$U_{28(1)}$	5.755	6.439245
	29	$U_{29(1)}$	39.727	44.45037
	30	$U_{30(1)}$	49.169	55.01499
	31	$U_{31(1)}$	34.922	39.07408
	32	$U_{32(1)}$	1.096	1.22631
	33	$U_{33(1)}$	0.667	0.746304
	34	$U_{34(1)}$	0.305	0.341263
	35	$U_{35(1)}$	5.808	6.498547
	36	$U_{36(1)}$	3.453	3.863547
	37	$U_{37(1)}$	2.334	2.611503
	38	$U_{38(1)}$	2.888	3.231371
	39	$U_{39(1)}$	1.187	1.328129
	40	$U_{40(1)}$	0.727	0.813437
	41	$U_{41(1)}$	0.414	0.463223
	42	$U_{42(1)}$	0.694	0.776514
	43	$U_{43(1)}$	2.903	3.248155
	44	$U_{44(1)}$	0.533	0.596371
	45	$U_{45(1)}$	3.637	4.069424
	46	$U_{46(1)}$	4.927	5.5128
	47	$U_{47(1)}$	3.87	4.330127
	48	$U_{48(1)}$	8.259	9.240961

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	49	$U_{49(1)}$	13.005	14.55124
	50	$U_{50(1)}$	13.32	14.90369
	51	$U_{51(1)}$	10.832	12.11988
	52	$U_{52(1)}$	10.37	11.60295
	53	$U_{53(1)}$	5.717	6.396727
	54	$U_{54(1)}$	4.074	4.558382
	55	$U_{55(1)}$	5.921	6.624982
	56	$U_{56(1)}$	44.766	50.08849
	57	$U_{57(1)}$	40.998	45.87249
	58	$U_{58(1)}$	24.732	27.67253
	59	$U_{59(1)}$	12.277	13.73668
	60	$U_{60(1)}$	1.673	1.871913
	61	$U_{61(1)}$	2.244	2.510802
	62	$U_{62(1)}$	1.351	1.511628
	63	$U_{63(1)}$	2.283	2.554439
	64	$U_{64(1)}$	0.668	0.747422
	65	$U_{65(1)}$	0.28	0.313291
	66	$U_{66(1)}$	1.152	1.288968
	67	$U_{67(1)}$	0.548	0.613155
	68	$U_{68(1)}$	0.296	0.331193
	69	$U_{69(1)}$	3.184	3.562564
	70	$U_{70(1)}$	3.061	3.42494
	71	$U_{71(1)}$	0.951	1.06407
	72	$U_{72(1)}$	1.481	1.657085
	73	$U_{73(1)}$	5.916	6.619388
	74	$U_{74(1)}$	5.702	6.379944
	75	$U_{75(1)}$	2.149	2.404507

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	76	$U_{76(1)}$	1.377	1.54072
	77	$U_{77(1)}$	21.639	24.21179
	78	$U_{78(1)}$	21.215	23.73737
	79	$U_{79(1)}$	16.168	18.09031
	80	$U_{80(1)}$	4.371	4.890694
	81	$U_{81(1)}$	2.144	2.398913
	82	$U_{82(1)}$	28.112	31.4544
	83	$U_{83(1)}$	31.695	35.4634
	84	$U_{84(1)}$	18.614	20.82713
	85	$U_{85(1)}$	34.919	39.07072
	86	$U_{86(1)}$	11.133	12.45667
	87	$U_{87(1)}$	7.159	8.010175
	88	$U_{88(1)}$	5.676	6.350853
	89	$U_{89(1)}$	2.522	2.821855
	90	$U_{90(1)}$	0.52	0.581826
	91	$U_{91(1)}$	2.846	3.184378
	92	$U_{92(1)}$	1.529	1.710792
	93	$U_{93(1)}$	6.151	6.882328
	94	$U_{94(1)}$	1.39	1.555265
	95	$U_{95(1)}$	0.965	1.079734
	96	$U_{96(1)}$	0.901	1.008125
	97	$U_{97(1)}$	0.265	0.296507
	98	$U_{98(1)}$	0.336	0.375949
	99	$U_{99(1)}$	0.561	0.627701
	100	$U_{100(1)}$	9.589	10.72909
	101	$U_{101(1)}$	3.978	4.450968
	102	$U_{102(1)}$	1.673	1.871913

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	103	$U_{103(1)}$	0.86	0.96225
	104	$U_{104(1)}$	0.592	0.662386
	105	$U_{105(1)}$	0.532	0.595253
	106	$U_{106(1)}$	0.2	0.223779
	107	$U_{107(1)}$	0.607	0.67917
	108	$U_{108(1)}$	0.476	0.532594
	109	$U_{109(1)}$	0.217	0.2428
	110	$U_{110(1)}$	0.687	0.768681
	111	$U_{111(1)}$	0.36	0.402802
	112	$U_{112(1)}$	0.502	0.561686
	113	$U_{113(1)}$	3.788	4.238377
	114	$U_{114(1)}$	4.172	4.668033
	115	$U_{115(1)}$	9.607	10.74923
	116	$U_{116(1)}$	6.777	7.582757
	117	$U_{117(1)}$	36.376	40.70095
	118	$U_{118(1)}$	54.741	61.24948
	119	$U_{119(1)}$	15.607	17.46261
	120	$U_{120(1)}$	5.058	5.659375
	121	$U_{121(1)}$	1.378	1.541838
	122	$U_{122(1)}$	4.21	4.710551
	123	$U_{123(1)}$	0.668	0.747422
	124	$U_{124(1)}$	4.691	5.24874
	125	$U_{125(1)}$	3.398	3.802008
	126	$U_{126(1)}$	3.18	3.558089
	127	$U_{127(1)}$	2.73	3.054586
	128	$U_{128(1)}$	2.662	2.978501
	129	$U_{129(1)}$	2.352	2.631643

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	130	$U_{130(1)}$	2.268	2.537656
	131	$U_{131(1)}$	2.086	2.334017
	132	$U_{132(1)}$	2.237	2.50297
	133	$U_{133(1)}$	1.396	1.561979
	134	$U_{134(1)}$	32.792	36.69083
	135	$U_{135(1)}$	29.13	32.59344
	136	$U_{136(1)}$	14.193	15.88049
	137	$U_{137(1)}$	10.871	12.16352
	138	$U_{138(1)}$	9.828	10.99651
	139	$U_{139(1)}$	6.504	7.277298
	140	$U_{140(1)}$	0.579	0.647841
	141	$U_{141(1)}$	0.945	1.057357
	142	$U_{142(1)}$	0.462	0.51693
	143	$U_{143(1)}$	8.135	9.102218
	144	$U_{144(1)}$	2.625	2.937102
	145	$U_{145(1)}$	1.629	1.822681
	146	$U_{146(1)}$	1.008	1.127847
	147	$U_{147(1)}$	1.485	1.66156
	148	$U_{148(1)}$	0.739	0.826864
	149	$U_{149(1)}$	1	1.118896
	150	$U_{150(1)}$	3.103	3.471934
	151	$U_{151(1)}$	1.529	1.710792
	152	$U_{152(1)}$	0.396	0.443083
	153	$U_{153(1)}$	1.636	1.830514
	154	$U_{154(1)}$	2.073	2.319471
	155	$U_{155(1)}$	3.664	4.099634
	156	$U_{156(1)}$	4.326	4.840343

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	157	$U_{157(1)}$	4.958	5.547485
	158	$U_{158(1)}$	9.911	11.08938
	159	$U_{159(1)}$	6.275	7.021071
	160	$U_{160(1)}$	26.271	29.39451
	161	$U_{161(1)}$	8.701	9.735513
	162	$U_{162(1)}$	27.279	30.52236
	163	$U_{163(1)}$	1.703	1.90548
	164	$U_{164(1)}$	1.185	1.325892
	165	$U_{165(1)}$	0.851	0.95218
	166	$U_{166(1)}$	7.087	7.929615
	167	$U_{167(1)}$	8.167	9.138022
	168	$U_{168(1)}$	5.93	6.635052
	169	$U_{169(1)}$	4.535	5.074193
	170	$U_{170(1)}$	13.844	15.48999
	171	$U_{171(1)}$	11.355	12.70506
	172	$U_{172(1)}$	3.85	4.307749
	173	$U_{173(1)}$	17.986	20.12446
	174	$U_{174(1)}$	10.417	11.65554
	175	$U_{175(1)}$	4.757	5.322587
	176	$U_{176(1)}$	6.766	7.570449
	177	$U_{177(1)}$	3.962	4.433065
	178	$U_{178(1)}$	3.247	3.633055
	179	$U_{179(1)}$	3.787	4.237258
	180	$U_{180(1)}$	13.231	14.80411
	181	$U_{181(1)}$	10.225	11.44071
	182	$U_{182(1)}$	12.86	14.389
	183	$U_{183(1)}$	2.279	2.549964

Appendix F (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (c.m.s.)	X'_{ij} (c.m.s.)
1	184	$U_{184(1)}$	1.313	1.46911
	185	$U_{185(1)}$	3.126	3.497668
	186	$U_{186(1)}$	6.297	7.045687
	187	$U_{187(1)}$	36.491	40.82963
	188	$U_{188(1)}$	4.767	5.333776
2	1	$U_{1(2)}$	0.17	0.029
	2	$U_{2(2)}$	71	0.029
	3	$U_{3(2)}$	57.779	0.029
	4	$U_{4(2)}$	0.08	0.029
	5	$U_{5(2)}$	0.037	0.029
	6	$U_{6(2)}$	0.029	0.029
	7	$U_{7(2)}$	0.164	0.029
	8	$U_{8(2)}$	65.724	0.029

APPENDIX G

Population Number 5

The investment fund, X_{ij}, and the adjusted value of the investment fund, X'_{ij}, of 88 crop factories

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	1	$U_{1(1)}$	210000	386338.1
	2	$U_{2(1)}$	800000	1471764
	3	$U_{3(1)}$	390000	717485
	4	$U_{4(1)}$	500000	919852.5
	5	$U_{5(1)}$	1530000	2814749
	6	$U_{6(1)}$	23000000	42313217
	7	$U_{7(1)}$	290000	533514.5
	8	$U_{8(1)}$	1000000	1839705
	9	$U_{9(1)}$	365000	671492.4
	10	$U_{10(1)}$	150000	275955.8
	11	$U_{11(1)}$	1050000	1931690
	12	$U_{12(1)}$	2120000	3900175
	13	$U_{13(1)}$	2000000	3679410
	14	$U_{14(1)}$	500000	919852.5
	15	$U_{15(1)}$	130000	239161.7
	16	$U_{16(1)}$	1110000	2042073
	17	$U_{17(1)}$	420000	772676.1
	18	$U_{18(1)}$	1600000	2943528
	19	$U_{19(1)}$	650000	1195808
	20	$U_{20(1)}$	640500	1178331
	21	$U_{21(1)}$	3600000	6622938

Appendix G (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	22	$U_{22(1)}$	1030000	1894896
	23	$U_{23(1)}$	1750000	3219484
	24	$U_{24(1)}$	2750000	5059189
	25	$U_{25(1)}$	163000	299871.9
	26	$U_{26(1)}$	250000	459926.3
	27	$U_{27(1)}$	182000	334826.3
	28	$U_{28(1)}$	1920000	3532234
	29	$U_{29(1)}$	400000	735882
	30	$U_{30(1)}$	6900000	12693965
	31	$U_{31(1)}$	2400000	4415292
	32	$U_{32(1)}$	714000	1313549
	33	$U_{33(1)}$	786000	1446008
	34	$U_{34(1)}$	193000	355063.1
	35	$U_{35(1)}$	14500000	26675724
	36	$U_{36(1)}$	650000	1195808
	37	$U_{37(1)}$	4150000	7634776
	38	$U_{38(1)}$	240000	441529.2
	39	$U_{39(1)}$	640000	1177411
	40	$U_{40(1)}$	14600000	26859694
	41	$U_{41(1)}$	158000	290673.4
	42	$U_{42(1)}$	602000	1107502
	43	$U_{43(1)}$	610000	1122220
	44	$U_{44(1)}$	2512500	4622259
	45	$U_{45(1)}$	500000	919852.5
	46	$U_{46(1)}$	220000	404735.1
	47	$U_{47(1)}$	250000	459926.3

Appendix G (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	48	$U_{48(1)}$	200000	367941
	49	$U_{49(1)}$	86000	158214.6
	50	$U_{50(1)}$	505000	929051.1
	51	$U_{51(1)}$	8350000	15361537
	52	$U_{52(1)}$	2150000	3955366
	53	$U_{53(1)}$	410000	754279.1
	54	$U_{54(1)}$	100000	183970.5
	55	$U_{55(1)}$	100000	183970.5
	56	$U_{56(1)}$	2200000	4047351
	57	$U_{57(1)}$	760000	1398176
	58	$U_{58(1)}$	900000	1655735
	59	$U_{59(1)}$	125000	229963.1
	60	$U_{60(1)}$	1000000	1839705
	61	$U_{61(1)}$	1200000	2207646
	62	$U_{62(1)}$	375000	689889.4
	63	$U_{63(1)}$	6300000	11590142
	64	$U_{64(1)}$	20000000	36794101
	65	$U_{65(1)}$	8790000	16171008
	66	$U_{66(1)}$	780000	1434970
	67	$U_{67(1)}$	3150000	5795071
	68	$U_{68(1)}$	4800000	8830584
	69	$U_{69(1)}$	1400000	2575587
	70	$U_{70(1)}$	3000000	5519115
	71	$U_{71(1)}$	2400000	4415292
	72	$U_{72(1)}$	21000000	38633807
	73	$U_{73(1)}$	780000	1434970
	74	$U_{74(1)}$	275000	505918.9

Appendix G (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	75	$U_{75(1)}$	14789297	27207945
	76	$U_{76(1)}$	1400000	2575587
	77	$U_{77(1)}$	2300000	4231322
	78	$U_{78(1)}$	5200000	9566466
2	1	$U_{1(2)}$	47000000	25000
	2	$U_{2(2)}$	25000	25000
	3	$U_{3(2)}$	80000	25000
	4	$U_{4(2)}$	30436000	25000
	5	$U_{5(2)}$	44400000	25000
	6	$U_{6(2)}$	35000000	25000
	7	$U_{7(2)}$	80000	25000
	8	$U_{8(2)}$	39400	25000
	9	$U_{9(2)}$	25000	25000
	10	$U_{10(2)}$	24500000	25000

APPENDIX H

Population Number 6

The investment fund, X_{ij}, and the adjusted value of the investment fund, X'_{ij}, of 112 steel factories

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	1	$U_{1(1)}$	4000000	5738530.7829
	2	$U_{2(1)}$	6550000	9396844.1570
	3	$U_{3(1)}$	2510000	3600928.0663
	4	$U_{4(1)}$	2100000	3012728.6610
	5	$U_{5(1)}$	2250000	3227923.5654
	6	$U_{6(1)}$	1000000	1434632.6957
	7	$U_{7(1)}$	3600000	5164677.7046
	8	$U_{8(1)}$	3400000	4877751.1655
	9	$U_{9(1)}$	1800000	2582338.8523
	10	$U_{10(1)}$	2700000	3873508.2785
	11	$U_{11(1)}$	7500000	10759745.2179
	12	$U_{12(1)}$	60000000	86077961.7434
	13	$U_{13(1)}$	4950000	7101431.8438
	14	$U_{14(1)}$	1166000	1672781.7232
	15	$U_{15(1)}$	4000000	5738530.7829
	16	$U_{16(1)}$	2390000	3428772.1428
	17	$U_{17(1)}$	2525000	3622447.5567
	18	$U_{18(1)}$	10500000	15063643.3051
	19	$U_{19(1)}$	500000	717316.3479
	20	$U_{20(1)}$	430000	616892.0592
	21	$U_{21(1)}$	10850000	15565764.7486

Appendix H (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	22	$U_{22(1)}$	70000	100424.2887
	23	$U_{23(1)}$	6500000	9325112.5222
	24	$U_{24(1)}$	4000000	5738530.7829
	25	$U_{25(1)}$	1000000	1434632.6957
	26	$U_{26(1)}$	4000000	5738530.7829
	27	$U_{27(1)}$	51700000	74170510.3689
	28	$U_{28(1)}$	40000000	57385307.8289
	29	$U_{29(1)}$	13600000	19511004.6618
	30	$U_{30(1)}$	84062087	120598218.4810
	31	$U_{31(1)}$	3250000	4662556.2611
	32	$U_{32(1)}$	14000000	20084857.7401
	33	$U_{33(1)}$	6000000	8607796.1743
	34	$U_{34(1)}$	19500000	27975337.5666
	35	$U_{35(1)}$	301000	431824.4414
	36	$U_{36(1)}$	1200000	1721559.2349
	37	$U_{37(1)}$	66000000	94685757.9178
	38	$U_{38(1)}$	215000000	307728713.2327
	39	$U_{39(1)}$	38000000	54516042.4375
	40	$U_{40(1)}$	1500000	2151949.0436
	41	$U_{41(1)}$	288000000	413174216.3684
	42	$U_{42(1)}$	1900000	2725802.1219
	43	$U_{43(1)}$	60000000	86077961.7434
	44	$U_{44(1)}$	15492244	22225679.7725
	45	$U_{45(1)}$	75000000	107597452.1793
	46	$U_{46(1)}$	82000000	117639881.0493
	47	$U_{47(1)}$	4000000	5738530.7829
	48	$U_{48(1)}$	6000000	8607796.1743

Appendix H (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	49	$U_{49(1)}$	3500000	5021214.4350
	50	$U_{50(1)}$	3750000	5379872.6090
	51	$U_{51(1)}$	3000000	4303898.0872
	52	$U_{52(1)}$	6700000	9612039.0613
	53	$U_{53(1)}$	16300000	23384512.9403
	54	$U_{54(1)}$	6000000	8607796.1743
	55	$U_{55(1)}$	17000000	24388755.8273
	56	$U_{56(1)}$	2660000	3816122.9706
	57	$U_{57(1)}$	62000000	88947227.1349
	58	$U_{58(1)}$	5700000	8177406.3656
	59	$U_{59(1)}$	400000000	573629275.5889
	60	$U_{60(1)}$	300000	430389.8087
	61	$U_{61(1)}$	4800000	6886236.9395
	62	$U_{62(1)}$	220000	315619.1931
	63	$U_{63(1)}$	86000000	123378411.8322
	64	$U_{64(1)}$	4000000	5738530.7829
	65	$U_{65(1)}$	40000000	57385307.8289
	66	$U_{66(1)}$	5000000	7173163.4786
	67	$U_{67(1)}$	7500000	10759745.2179
	68	$U_{68(1)}$	2500000	3586581.7393
	69	$U_{69(1)}$	900000	1291169.4262
	70	$U_{70(1)}$	8000000	11477061.5658
	71	$U_{71(1)}$	75200000	107884378.7184
	72	$U_{72(1)}$	117000000	167134709.0518
	73	$U_{73(1)}$	3400000	4877751.1655
	74	$U_{74(1)}$	8000000	11477061.5658
	75	$U_{75(1)}$	73900000	106019356.2140

Appendix H (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	76	$U_{76(1)}$	13000000	18650225.0444
	77	$U_{77(1)}$	340000000	487775116.5460
	78	$U_{78(1)}$	163000000	233845129.4030
	79	$U_{79(1)}$	19000000	27258021.2187
	80	$U_{80(1)}$	7400000	10616281.9484
	81	$U_{81(1)}$	31000000	44473613.5674
	82	$U_{82(1)}$	14600000	20945637.3576
	83	$U_{83(1)}$	65000000	93251125.2220
	84	$U_{84(1)}$	30000000	43038980.8717
	85	$U_{85(1)}$	2199770	3155861.9651
	86	$U_{86(1)}$	4050000	5810262.4177
	87	$U_{87(1)}$	10500000	15063643.3051
	88	$U_{88(1)}$	80000000	114770615.6579
	89	$U_{89(1)}$	6500000	9325112.5222
	90	$U_{90(1)}$	4300000	6168920.5916
	91	$U_{91(1)}$	8424000	12085345.8288
	92	$U_{92(1)}$	20024000	28727085.0992
	93	$U_{93(1)}$	8650000	12409572.8180
	94	$U_{94(1)}$	22015000	31583438.7964
	95	$U_{95(1)}$	50000000	71731634.7862
	96	$U_{96(1)}$	24000000	34431184.6974
	97	$U_{97(1)}$	30000000	43038980.8717
	98	$U_{98(1)}$	10500000	15063643.3051
	99	$U_{99(1)}$	200000000	286926539.1447
	100	$U_{100(1)}$	3700000	5308140.9742
	101	$U_{101(1)}$	1000000	1434632.6957
	102	$U_{102(1)}$	98000000	140594004.1809

Appendix H (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	103	$U_{103(1)}$	7450000	10688013.5831
	104	$U_{104(1)}$	30000000	43038980.8717
	105	$U_{105(1)}$	207000000	296968968.0148
	106	$U_{106(1)}$	20000000	28692653.9145
2	1	$U_{1(2)}$	380000000	70000.0000
	2	$U_{2(2)}$	400000	70000.0000
	3	$U_{3(2)}$	900000000	70000.0000
	4	$U_{4(2)}$	510000	70000.0000
	5	$U_{5(2)}$	753000	70000.0000
	6	$U_{6(2)}$	340000000	70000.0000

APPENDIX I

Population Number 7

The investment fund, X_{ij}, and the adjusted value of the investment fund, X'_{ij}, of 84 noodle factories

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	1	$U_{1(1)}$	195000	291088.8000
	2	$U_{2(1)}$	210000	313480.2461
	3	$U_{3(1)}$	410000	612032.8615
	4	$U_{4(1)}$	115000	171667.7538
	5	$U_{5(1)}$	265000	395582.2153
	6	$U_{6(1)}$	340000	507539.4461
	7	$U_{7(1)}$	20000	29855.2615
	8	$U_{8(1)}$	1000000	1492763.0767
	9	$U_{9(1)}$	120000	179131.5692
	10	$U_{10(1)}$	265000	395582.2153
	11	$U_{11(1)}$	360000	537394.7076
	12	$U_{12(1)}$	504000	752352.5907
	13	$U_{13(1)}$	196200	292880.1157
	14	$U_{14(1)}$	20000	29855.2615
	15	$U_{15(1)}$	510000	761309.1691
	16	$U_{16(1)}$	300000	447828.9230
	17	$U_{17(1)}$	55000	82101.9692
	18	$U_{18(1)}$	310000	462756.5538
	19	$U_{19(1)}$	290000	432901.2923
	20	$U_{20(1)}$	168000	250784.1969
	21	$U_{21(1)}$	360000	537394.7076
	22	$U_{22(1)}$	1100000	1642039.3844

Appendix I (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	23	$U_{23(1)}$	300000	447828.9230
	24	$U_{24(1)}$	320000	477684.1846
	25	$U_{25(1)}$	4600000	6866710.1529
	26	$U_{26(1)}$	350000	522467.0769
	27	$U_{27(1)}$	510000	761309.1691
	28	$U_{28(1)}$	1270000	1895809.1074
	29	$U_{29(1)}$	190000	283624.9846
	30	$U_{30(1)}$	1000000	1492763.0767
	31	$U_{31(1)}$	700000	1044934.1537
	32	$U_{32(1)}$	4880000	7284683.8144
	33	$U_{33(1)}$	700000	1044934.1537
	34	$U_{34(1)}$	2700000	4030460.3072
	35	$U_{35(1)}$	520000	776236.7999
	36	$U_{36(1)}$	155000	231378.2769
	37	$U_{37(1)}$	830000	1238993.3537
	38	$U_{38(1)}$	3100000	4627565.5379
	39	$U_{39(1)}$	300000	447828.9230
	40	$U_{40(1)}$	70000	104493.4154
	41	$U_{41(1)}$	4800000	7165262.7683
	42	$U_{42(1)}$	444000	662786.8061
	43	$U_{43(1)}$	77000	114942.7569
	44	$U_{44(1)}$	400000	597105.2307
	45	$U_{45(1)}$	70000	104493.4154
	46	$U_{46(1)}$	30000	44782.8923
	47	$U_{47(1)}$	100000	149276.3077
	48	$U_{48(1)}$	2800000	4179736.6148
	49	$U_{49(1)}$	1100000	1642039.3844

Appendix I (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	50	$U_{50(1)}$	500000	746381.5384
	51	$U_{51(1)}$	230000	343335.5076
	52	$U_{52(1)}$	245000	365726.9538
	53	$U_{53(1)}$	360000	537394.7076
	54	$U_{54(1)}$	740000	1104644.6768
	55	$U_{55(1)}$	1800000	2686973.5381
	56	$U_{56(1)}$	601000	897150.6091
	57	$U_{57(1)}$	65000	97029.6000
	58	$U_{58(1)}$	1370000	2045085.4151
	59	$U_{59(1)}$	5000000	7463815.3836
	60	$U_{60(1)}$	1850000	2761611.6919
	61	$U_{61(1)}$	130000	194059.2000
	62	$U_{62(1)}$	4100000	6120328.6146
	63	$U_{63(1)}$	111000	165696.7015
	64	$U_{64(1)}$	955000	1425588.7383
	65	$U_{65(1)}$	100000	149276.3077
	66	$U_{66(1)}$	510000	761309.1691
	67	$U_{67(1)}$	125000	186595.3846
	68	$U_{68(1)}$	100000	149276.3077
	69	$U_{69(1)}$	2400000	3582631.3841
	70	$U_{70(1)}$	1350000	2015230.1536
	71	$U_{71(1)}$	1120000	1671894.6459
	72	$U_{72(1)}$	1800000	2686973.5381
	73	$U_{73(1)}$	3549000	5297816.1593
	74	$U_{74(1)}$	3050000	4552927.3840
	75	$U_{75(1)}$	1100000	1642039.3844
	76	$U_{76(1)}$	505000	753845.3537

Appendix I (Continued)

Group No.(i)	Unit No.(j)	$U_{j(i)}$	X_{ij} (Baht)	X'_{ij} (Baht)
1	76	$U_{76(1)}$	505000	753845.3537
	77	$U_{77(1)}$	269700	402598.2018
	78	$U_{78(1)}$	5100000	7613091.6913
	79	$U_{79(1)}$	3300000	4926118.1532
2	1	$U_{1(2)}$	200000	20000.0000
	2	$U_{2(2)}$	740000	20000.0000
	3	$U_{3(2)}$	6400000	20000.0000
	4	$U_{4(2)}$	300000	20000.0000
	5	$U_{5(2)}$	32800000	20000.0000

BIBLIOGRAPHY

Avadhani, M.S., and Sukhatme, B.V. 1965. Controlled simple random sampling. **Journal of the Indian Society of Agricultural Statistics.** 17: 34-42.

_____. 1967. Controlled sampling with varying probabilities with and without replacement. **The Australian Journal Statistics.** 9(1): 8-15.

_____. 1968. Simplified procedures for designing controlled simple random sampling. **The Australian Journal Statistics.** 10(1): 1-7.

_____. 1973. Controlled sampling with equal probabilities and without replacement. **International Statistical Review.** 41(2): 175-182.

Brewer, K.R.W. 1963a. A model of systematic sampling with unequal probabilities. **Australian Journal of Statistics.** 5: 5-13. Quoted in Cochran, W.G. 1977. **Sampling Techniques.** 3rd ed. New York: John Wiley and Sons, Inc.

_____. 1975. A simple procedure for sampling πpswor. **Australian Journal of the Statistics.** 17(3): 166-172.

Brewer, K.R.W., Early, L.J., and Joyce, S.F. 1972. Selecting several samples from a single population. **Australian Journal of the Statistics.** 14(3): 231-239.

Brien, L.G.O. 1987. User control versus randomisation in geographical probability sampling: a compromise solution using controlled sampling. **Environmental and Planning A.** 19: 949-958.

Chaudhuri, Arijit, and Vos, J.W.E. 1988. **Unified Theory and Strategies of Survey Sampling**. 4 Vols. Netherlands: Elsevier Science Publishing Company, Inc.

Choudhry, G.H., and Singh, M.P. 1979. Quoted in Chaudhuri, Arijit, and J.W.E. Vos. 1988. **Unified Theory and Strategies of Survey Sampling**. 4 Vols. Netherlands: Elsevier Science Publishing Company, Inc.

Cochran, W.G. 1977. **Sampling Techniques**. 3rd ed. New York: John Wiley and Sons, Inc.

Das, A.C. 1951. On two-phase sampling and sampling with varying probabilities. **Bulletin of the International Statistical Institute.** 33(2): 105-112. Quoted in Chaudhuri, Arijit, and J.W.E. Vos. 1988. **Unified Theory and Strategies of Survey Sampling**. 4 Vols. Netherlands: Elsevier Science Publishing Company, Inc.

Deshpande, M.N. 1982. A new fixed size sampling procedure with unequal probabilities of selection. **Journal of the Indian Society of Agricultural Statistics.** 34(2): 71-75.

Dodds, D.J., and Fryer, J.G. 1971. Some families of selection probabilities for sampling with probability proportional to size. **Journal of the Royal Statistical Society, Series B.** 33(2): 263-274.

Durbin, J. 1967. Design of multi-stage surveys for the estimation of sampling errors. **Applied Statistics.** 16: 152-164.

Fellegi, I.P. 1963. Sampling with varying probabilities without replacement: rotating and non-rotating samples. **Journal of the American Statistical Association.** 58(March): 183-201.

Fuller, W.A. 1970. Sampling with random stratum boundaries. **Journal of the Royal Statistical Society, Series B.** 32(2): 209-226.

Goodman, R., and Kish, L. 1950. Controlled selection-a technique in probability sampling. **Journal of the American Statistical Association.** 45(September): 350-372.

Grundy, P.M. 1954. A method of sampling with probability exactly proportional to size. **Journal of the Royal Statistical Society, Series B.** 16(2): 236-238.

Gupta, V.K., Nigam, A.K., and Kumar, P. 1982. On a family of sampling schemes with inclusion probability proportional to size. **Biometrika.** 69(1): 191-196.

Ha'jek, J. 1964. Asymptotic theory of rejective sampling with varying probabilities from a finite population. **The Annals of Mathematical Statistics.** 35: 1491-1523.

Hansen, M.H., and Hurwitz, W.N. 1943. On the theory of sampling from finite populations. **The Annals of Mathematical Statistics.** 14: 333-362. Quoted in Chaudhuri, Arijit, and J.W.E. Vos. 1988. **Unified Theory and Strategies of Survey Sampling**. 4 Vols. Netherlands: Elsevier Science Publishing Company, Inc.

Hanurav, T.V. 1967. Optimum utilization of auxiliary information: πPS sampling of two units from a stratum. **Journal of the Royal Statistical Society, Series B.** 29: 374-391. Quoted in Chaudhuri, Arijit, and J.W.E. Vos. 1988. **Unified Theory and Strategies of Survey Sampling**. 4 Vols. Netherlands: Elsevier Science Publishing Company, Inc.

Hedayat, A.S., and Singha, Bikas K. 1991. **Design and Inference in Finite Population Sampling**. New York: John Wiley and Sons, Inc.

Horvitz, D.G., and Thompson, D.J. 1952. A generalization of sampling without replacement from a finite universe. **Journal of the American Statistical Association.** 47(December): 663-685.

Lahiri, D.B. 1951. A method of sample selection providing unbiased ratio estimates. **Bulletin of the International Statistical Institute.** 33(2): 133-140.

Lanke, J. 1974. On non-negative estimators in survey Sampling. **Sankhya, Series C.** 36: 33-42.

Madow, W.G. 1949. On the theory of systematic sampling II. **The Annals of Mathematical Statistics.** 20: 333-354.

Murthy, M.N. 1957. Ordered and unordered estimators in sampling without replacement. **Sankhya.** 18: 379-390.

Narain, R.D. 1951. On sampling without replacement with varying probabilities. **Journal of the Indian Society of Agricultural Statistics.** 3(2): 169-175.

Nigam, A.K., Kumar, P., and Gupta, V.K. 1984. Some methods of inclusion probability proportional to size sampling. **Journal of the Royal Statistical Society, Series B.** 46(3): 564-571.

Raj, D. 1956. Some estimators in sampling with varying probabilities without replacement. **Journal of the American Statistical Association.** 51: 269-284.

Rao, J.N.K. 1963. On three procedures of unequal probability sampling without replacement. **Journal of the American Statistical Association.** 58(March): 202-215.

Rao, J.N.K., and Bayless, D.L. 1969. An empirical study of the stabilities of estimators and variance estimators in unequal probability sampling of two units per-statum. **Journal of the American Statistical Association.** 64: 540-559.

Rao, J.N.K., and Nigam, A.K. 1990. Optimal controlled sampling designs. **Biometrika.** 77(4): 807-814.

_____. 1992. Optimal controlled sampling: a unified approach. **International Statistical Review.** 60(1): 89-98.

Rao, J.N.K., and Vijayan, K. 1977. On estimating the variance in sampling with probability proportional to aggregate size. **Journal of the American Statistical Association.** 72(359): 579-584.

Rao, J.N.K., Hartley, H.O., and Cochran, W.G. 1962. On a simple procedure of unequal probability sampling without replacement. **Journal of the Royal Statistical Society, Series B.** 24(2): 482-491.

Samiuddin, M., and Asad, H. 1981. A simple procedure of unequal probability sampling. **Biometrika.** 68(3): 728-731.

Sampford, M.R. 1967. On sampling without replacement with unequal probabilities of selection. **Biometrika.** 54(3-4): 499-513.

Saxena, R.R., Singh, P., and Srivastava, A.K. 1986. An unequal probability sampling scheme. **Biometrika.** 73(3): 761-763.

Sen, A.R. 1953. On the estimator of variance in sampling with varying probabilities. **Journal of the Indian Society of Agricultural Statistics.** 5(2): 119-127.

Singh, P. 1978. The selection of samples of two units with inclusion probabilities proportional to size. **Biometrika.** 65(2): 450-454.

Srivastava, A.K., and Singh, D. 1981. A sampling procedure with inclusion probabilities proportional to size. **Biometrika.** 68(3): 732-734.

Stevens, W.L. 1958. Sampling without replacement with probability proportional to size. **Journal of the Royal Statistical Society, Series B.** 20(2): 393-397.

Sunter, A.B. 1977. List sequential sampling with equal or unequal probabilities without replacement. **Applied Statistics.** 26(3): 261-268.

Vijayan, K. 1968. An exact πps sampling scheme-generalization of a method of Hanurav. **Journal of the Royal Statistical Society, Series B.** 30(3): 556-566.

Yates, F., and Grundy, P.M. 1953. Selection without replacement from within strata with probability proportional to size. **Journal of the Royal Statistical Society, Series B.** 15(2): 253-261.

VITA

Miss Wimon Sonchaem was born on October 16, 1963, in Bangkok, Thailand. Having received her bachelor's degree in Medical Technology at Mahidol University in 1985, she was offered a job at Diethelm Company Co., Ltd. In 1991, she earned a Master's degree in Applied Statistics at the National Institute of Development Administration (NIDA). Upon graduation from NIDA, she taught at the Prince of Songkhla University, Haad Yai Campus, for one semester before she moved to Bangkok to teach at Mahidol University, Salaya Campus. While teaching at Mahidol University's Faculty of Environment and Resource Studies, she enrolled in NIDA's Graduate School of Applied Statistics' Doctoral Program in Statistics. She is expected to complete her Ph.D. degree by the end of the second semester of the academic year 2000.