PATH SAMPLING

Mena Patummasut

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
2011

PATH SAMPLING

Mena Patummasut

School of Applied Statistics

Associate Professor ..... Major Advisor

(Arthur L. Dryver, Ph.D.)

The Examining Committee Approved This Disseration Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy (Statisitics).

Associate Professor .....Committee Chairperson

(Pachitjanut Siripanich, Ph.D.)

Professor .....Committee

(Prachoom Suwattee, Ph.D.)

Instructor .....Committee

(Ampai Thongteeraparp, Ph.D.)

Instructor .....Dean

(Lersan Bosuwan, Ph.D.)

April 2012

ABSTRACT

Title of Dissertation	Path Sampling
Author	Miss Mena Patummasut
Degree	Doctor of Philosophy (Statistics)
Year	2011

In sampling spatial populations, one part of the cost is due to the distance travelled to observe all of the units in a sample. Cluster sampling is one such sampling design which is often used specifically to address this issue. Even in cluster sampling, the researcher may have to travel great distances from the cluster to cluster selected. In an optimal setting, when sampling costs are mainly a function of distance travelled, researchers could sample all of the units in the path travelled during the sampling. For this reason, the authors are introducing a new sampling design, called "path sampling," which offers exactly the latter ability to sample all of the units in the researcher's path traversed during the sampling. Path sampling is a design in which the researcher selects a path or paths from start to finish, as opposed to selecting units. By applying the Horvitz-Thompson estimator, path sampling offers unbiased estimators for both mean and variance. This dissertation covers the pros and cons of path sampling in comparison to simple random sampling, cluster sampling, and random walk sampling.

The simulation results show that path sampling gives the smallest value of the expected number of units traveled for the same sample size among four sampling designs. Thus, path sampling has less traveling or less cost. However, path sampling is less efficient than cluster sampling, simple random sampling without replacement, and random walk sampling in the population with low variation of y-values among clusters. On the other hand, path sampling is more efficient than random walk sampling in a population with high variation of y-values among clusters. Moreover, path sampling is more efficient than cluster sampling and SRSWOR in a population with high variation of y-values among clusters with the path starting or ending point on high y-values.

ACKNOWLEDGEMENTS

I am greatly indebted to my advisor, Associate Professor Dr. Arthur L. Dryver, for his invaluable comments on the completion of this dissertation and for his patient guidance, supervision and encouragement throughout this study. Furthermore, he has been a good teacher for me and his wisdom and knowledge has been invaluable to the development of ideas in the study. I would also like to gratefully acknowledge all of the members of the committee for my dissertation, Professor Dr. Prachoom Suwattee, Associate Professor Dr. Pachitjanut Siripanich and Dr. Ampai Thongteeraparp for their helpful comments and suggestions.

I would like to express my gratitude to Associate Professor Dr. Virool Boonyasombat, Associate Professor Dr. Vichit Loirachoonkul, Associate Professor Dr. Jirawan Jitthavech and Associate Professor Dr. Samruam Chongcharoen for their invaluable lectures, imparting their knowledge and experience, which really helped to guide me throughout this study. I would like to extend my special thanks to the National Statistical Office for supplying me with useful data for the study.

I would like to thank Kasetsart University, and the Commission on Higher Education, Thailand, for financial support through a grant under the Strategic Scholarships Fellowships Frontier Research Networks. I am very grateful to Dr. Bruce Leeds for his kindness towards me by editing my English, which has made the manuscript more readable.

To my friends and colleagues, I could not have done this without their support. Thanks for your friendship and compassion. I would like to express my deepest thanks to all persons, whether their names have been mentioned above or not, for their energy and help, spirit, patience and co-operation. Finally, I want to express my gratitude to my parents and my sisters for their patience and encouragement throughout the period of this accomplishment.

Mena Patummasut

April 2012

TABLE OF CONTENTS

LIST OF TABLES

Tables **Page**

LIST OF FIGURES

Figures **Page**

CHAPTER 1

INTRODUCTION

1.1 Statement of the Problem

Many of the sample survey methods have been applied to natural populations for the purpose of estimating total numbers or population density (Seber, 1986: 267, 1992: 129; Thompson, 2002: 6, 289). The population study area is divided into spatial units (plots) generally of the same size, and the numbers of animals or organisms are counted on a selection of the units. There are many sampling designs that can be used, for example, simple random sampling, cluster sampling, systematic sampling, or adaptive sampling in the case of rare or clustered populations. The sampler may use simple random sampling in a spatial population because it is not a complicated design and the estimators are easy to calculate.

In simple random sampling, the sample consists of n units randomly selected from the N units in the spatial population. At each selection step, each unit has an equal chance of selection (Thompson, 2002: 11). Thus, the simple random sample may select units all over the study region, as shown in Figure 1.1 (a). Unfortunately, traveling from place to place to observe every unit selected can be costly, as the distance traveled can be quite long. For example, we have to visit 34 units to investigate a simple random sample of 10 units, as shown in Figure 1.1(b).

Cluster sampling is used in practice because it is usually much cheaper and more convenient to select clusters of units than randomly selected units in the population. In cluster sampling, a primary unit consists of a cluster of secondary units, usually in close proximity to each other. A sampling unit is a primary unit. For the spatial setting, cluster primary units include spatial arrangements as square collections of adjacent plots or long narrow strips of adjacent units (Thompson, 2002: 129). In one-stage cluster sampling, a simple random sample of n primary units is taken from

N primary units in the population, and the elements observed are all secondary units within the clusters (Lohr, 1999: 134).

(a) A Simple Random Sample (b) Sampling Travel Visiting 34 Units

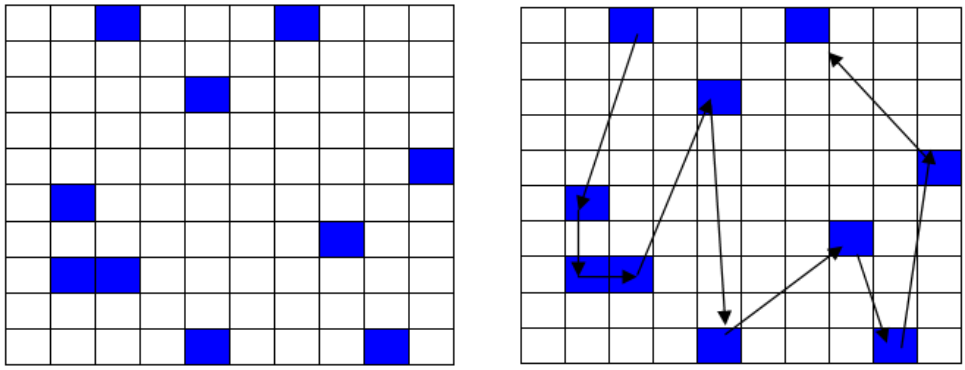


Figure 1.1 A Simple Random Sample of 10 Units from a Population of 100 Units

The advantage of cluster sampling is that it is often less costly to sample a collection of units in a cluster than to sample an equal number of secondary units selected at random from the population (Thompson, 2002: 139). However, cluster samples, a simple random sample of primary units, may cover all of the study region. It takes a lot of time and a high cost of sampling to travel from cluster to cluster.

It can be seen that the problem in simple random sampling and cluster sampling is that a sample may cover all of the region since each sampling unit has an equal chance of selection. Thus, traveling from place to place to sample every unit selected for sampling can be costly, as the distance traveled can be quite long. Therefore, path sampling is a new sampling design proposed in this dissertation to overcome this disadvantage.

Path sampling, as proposed in this study, is a sampling design in which *p* distinct paths are selected by simple random sampling without replacement (SRSWOR) from the *q* paths in the population, and the sample consists of all units in the selected paths. A path is defined as the course of sampling from the starting unit to the finishing unit.

Every unit in the sampled paths will be observed, as shown in Figure 1.2. This sampling scheme eliminates or at least reduces the distance travelled over units that are not to be sampled. Cost is considered as a function of distance traveled by counting units traveled to observe all units in the sample. In other words, we consider the number of units traveled. For example, to investigate a path sample of 26 units in Figure 1.2 (a), we visit only 26 units, as shown in Figure 1.2 (b).

Path sampling utilizes all of the observations of the units traveled. Thus, when the main cost of sampling a unit is the distance traveled, path sampling may be a very cost-effective design.

(a) A Path Sample (b) Sampling Travel

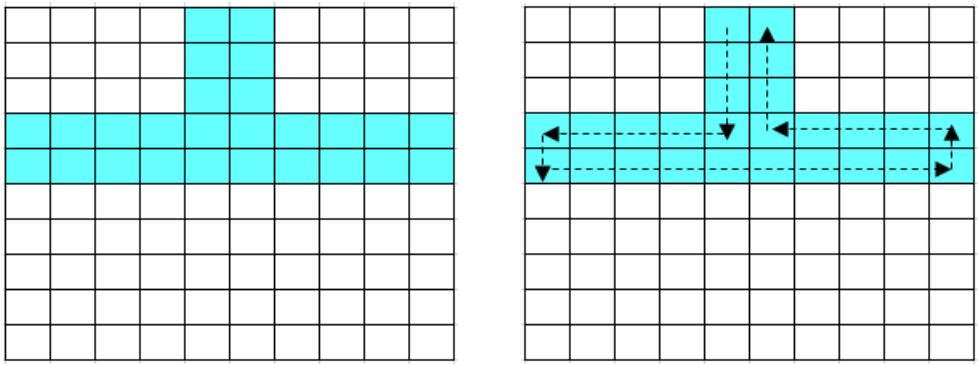


Figure 1.2 A Path Sample With Only One Path Selected

1.2 Objectives of the Study

This dissertation focuses on finding a new sampling design which is cost effective and convenient under certain circumstances. This new sampling is named path sampling. The objectives of the study are as follows:

1) To propose a new cost-effective and convenient sampling design, named path sampling, for the spatial setting population

2) To find an estimator of the population mean and its variance for path sampling

3) To investigate the cost and relative efficiencies of path sampling as opposed to other sampling designs

1.3 Scope of the Study

The scope of the study is as follows. This study considers sampling in a study area that can be divided into spatial units of equal size. The new sampling design, path sampling, will be studied in the spatial setting. A path sampling scheme is proposed. The parameter considered in this study is the population mean. An estimator of the mean for the path sampling is proposed, and the properties of the proposed estimator, such as unbiasedness and variance, are investigated. Simulation is used to investigate the relative efficiencies of path sampling in relation to other sampling design–simple random sampling, cluster sampling and random walks sampling. Cost is considered as a function of distance traveled or the number of units traveled to observe all of the units in a sample. The number of units travelled of path sampling, simple random sampling, cluster sampling and random walk sampling are compared through simulation.

CHAPTER 2

LITERATURE REVIEW

The main purpose of this dissertation is to propose a new cost-effective sampling design that can be applied to a spatial setting population. The previous sampling designs should be reviewed first however.

Sampling consists of selecting some part of a population to observe so that one may estimate something about the whole population. The objective in sampling is to estimate some characteristics of the population, such as the mean or the total.

Sampling design is the procedure by which a sample of units is selected from the population. The design is determined by assigning to each possible sample s the probability P(s) (Thompson, 2002: 2).

2.1 Review of Methods of Sampling and Estimation

The theory of independent random sampling was developed by Bernoulli more than 200 years ago. Poisson considered the theory of stratification. Later, Lexis provided the theory of cluster sampling. In the early 1900's, the theory of sampling a finite population with equal probabilities and without replacement was developed. The estimation of the mean for simple random sampling was proposed by Splawa-Neyman (1925: 472-479).

Neyman (1934: 558-625) introduced the concept of the optimum allocation of sampling units to different strata. Stratified sampling and purposive sampling were compared. In 1938, Neyman proposed double sampling, which provides a better estimator by using an auxiliary variable.

Hansen and Hurwitz's (1943: 333-362) paper was the first to introduce unequal probabilities to select the sampling unit in order to increase the precision of the estimators. They considered the sampling scheme under the population made up

of strata. Each stratum contained primary sampling units consisting of secondary units. They proposed the selection of only one primary unit per stratum with probabilities proportionate to some measure of their size. The secondary units were selected from the selected primary unit, with equal probabilities without replacement. Note that this method was confined to samples of only one primary unit per stratum.

Midzuno (1950: 149-156) generalized the Hansen and Hurwitz approach to sampling a combination of *n* units with a probability proportional to some measure of the size of the combination. Madow (1949: 333-354) made a contribution to the theory of the systematic selection of several clusters of sampling units with the idea of probability proportional to a measure of size.

Horvitz and Thompson (1952: 663-685) mentioned the limitations of the Hansen and Hurwitz scheme, that an unbiased estimate of the sampling variance of the estimator cannot be obtained from the sample elements. Horvitz and Thompson provided a general method for dealing with sampling without replacement from a finite population when unequal selection probabilities are used. They proposed an unbiased estimator of the total of a finite population, now called the Horvitz-Thompson estimator, and also estimated the variance of the estimator. The general nature of this approach to sampling a finite population without replacement was illustrated by considering Horvitz-Thompson's estimator and its variance for simple random, systematic, and stratified random sampling procedure.

The formula of the Horvitz-Thompson (1952) estimator is shown in the following. For any design, with or without replacement (Thompson 2002: 53), giving probability π_i that unit i is included in the sample, for $i = 1, 2, \ldots, N$, an unbiased estimator of the population total is

$$\hat{\tau}_{HT} = \sum_{k=1}^{\upsilon} \frac{y_i}{\pi_i} , \qquad (2.1)$$

where υ is the number of distinct units in the sample. This estimator does not depend on the number of times a unit may be selected. Each distinct unit of the sample is utilized only once.

Note that if $\pi_i = \upsilon y_i / \sum_{i=1}^{N} y_i$, then $\hat{\tau}_{HT}$ will have zero variance and the sampling will be optimum.

The variance of the estimator is

$$v(\hat{\tau}_{HT}) = \sum_{i=1}^{N} \left(\frac{1-\pi_i}{\pi_i} \right) y_k^2 + \sum_{i=1}^{N} \sum_{i' \neq i} \left(\frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_i \pi_{i'}} \right) y_i y_{i'} \, . \tag{2.2}$$

This formula is applied only when every element has a positive inclusion probability.

Let $\pi_{ii'}$ be the probability that both units *i* and unit i' are included in the sample. An unbiased estimator of this variance is

$$\hat{v}(\hat{\tau}_{HT}) = \sum_{i=1}^{\upsilon} \left(\frac{1}{\pi_i^2} - \frac{1}{\pi_i} \right) y_i^2 + \sum_{i=1}^{\upsilon} \sum_{i' \neq i} \left(\frac{1}{\pi_i \pi_{i'}} - \frac{1}{\pi_{ii'}} \right) y_i y_{i'} \, . \tag{2.3}$$

It is unbiased if all of the joint inclusion probabilities are greater than zero. This variance estimate may be negative in some designs. Rao and Singh (1973: 95-104) studied 34 natural populations, selecting samples of size n=2. They found that $\hat{v}(\hat{\tau}_{HT})$ frequently resulted in negative estimates.

From the fact that if $\pi_i = \upsilon y_i / \sum_{i=1}^{N} y_i$, then $\hat{\tau}_{HT}$ will have zero variance and the sampling will be optimum, as noted by Horvitz and Thompson (1952). Thus, if the inclusion probabilities π_i can be chosen approximately proportional to the value y_i, the variance of the Horvitz-Thompson estimator would be low. Since y_i's are unknown, if related auxiliary information on a characteristic x_i is available, then the suitable choice for a design would be one for which π_i is proportionate to x_i.

Sen (1953: 119-127); Yates and Grundy (1953: 253-261) have suggested, independently, that the Horvitz and Thompson variance estimator can be of negative values. Thus, they proposed an alternative variance estimator,

$$\hat{v}_{YGS}(\hat{\tau}_{HT}) = \frac{1}{2}\sum_{i=1}^{\upsilon}\sum_{i'\neq i}\left(\frac{\pi_i\pi_{i'} - \pi_{ii'}}{\pi_{ii'}}\right)\left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2 , \qquad (2.4)$$

which may be negative. $\hat{v}(\hat{\tau}_{HT})$ and $\hat{v}_{YGS}(\hat{\tau}_{HT})$ give different value, and it is believed that $\hat{v}_{YG}(\hat{\tau}_{HT})$ is less often negative than $\hat{v}(\hat{\tau}_{HT})$.

Since the disadvantage of $\hat{v}(\hat{\tau}_{HT})$ and $\hat{v}_{YGS}(\hat{\tau}_{HT})$ is that they can take a negative value and the inclusion probabilities are not easy to compute, Raj (1956: 269-284) proposed unbiased estimators of the population total whose estimated variance is always positive and utilizes the conditional probability of selection, given the units selected previously, instead of inclusion probabilities. The value of the estimator depends on the order in which the units in the sample are selected. Thus, Raj's estimator is an ordered estimator, that is, an estimator which takes into account the order in which the units are drawn. It is not claimed that this estimator is necessarily more efficient than the Horvitz-Thompson estimator, although this has been found to be the case in several examples. He also showed that the estimator of variance given by Yates and Grundy (1953) is positive in at least two important situations:

1) When the first unit is selected with probabilities proportional to some measure of size and the remaining units are selected with equal probability

2) When the first unit is selected with probabilities proportional to some measure of size and the second unit with probabilities proportional to the sizes of the remaining units, for the sample of size 2

Murthy (1957: 379-390) improved this estimator by removing the dependence on order, but the improved estimator is not easy to compute. Murthy (1957) showed that corresponding to any biased or unbiased ordered, there exists an unordered estimator, which is more efficient, in sampling with varying probabilities without replacement.

The technique of improving the ordered estimators by unordered ones is also explained in this paper. This method is applied to the set of estimators given by Raj's estimators, which provide unbiased estimates of the population total.

Brewer and Donadio (2003: 189-196) have shown that, under conditions of high entropy, the variance of Horvitz-Thompson estimator depends almost entirely on inclusion probabilities.

Sampling of a rare population can be tedious and expensive; many sampling designs have been proposed for this type of population. Such sampling designs include inverse sampling, capture-recapture sampling, line-intercept sampling, and the link-tracing design: network sampling, snowball sampling, random walk sampling, and adaptive sampling. Briefly the details of these sampling designs are as follows.

1) Inverse sampling

Haldane (1945: 222-225) introduced a sampling technique to handle sampling in rare populations. In sampling for a rare attribute, a sample of fixed size may result in having no individuals with the attribute presented in the sample. An inverse sample is selected unit by unit using simple random sampling until a specified number of m units possessing the rare attribute is selected. Haldane considered the infinite population case; the sample size is a random variable in this case and its probability distribution were derived. An unbiased estimator of the proportion and the variance of the estimate were presented.

2) Capture-recapture sampling

Capture-recapture sampling was developed by Seber (1973). In capture-recapture sampling, in order to estimate the total number of individuals in a population, an initial sample is obtained and the individuals in that sample are marked. A second sample is obtained independently and the marked individuals are counted. If the second sample is representative of the population as a whole, the sample proportion of marked individuals should be about the same as the population proportion of the marked individuals. From this relationship, the total number of individuals in the population can be estimated (Thompson, 2002: 233; Lindberg and Rexstad, 2002: 251-262).

3) Line-Intercept sampling

Line-intercept sampling (Lucas and Seber, 1977: 618-622) is appropriate for sampling in a rare population for the purpose of estimating the population total. It is a sampling design in which *n* transect lines are selected at random by selecting *n* positions along a baseline of length *b* that traverses the width of the study region and a transect is run across the study area perpendicular to the baseline at each of the selected points. Whenever an object of the population is intersected by one or more of the sample lines, a variable of interest associated with that object is recorded.

4) Link-tracing design

A link-tracing design is a design in which links or connections between units are used to obtain the sample. The link-tracing design explained here includes network sampling, snowball sampling, random walk sampling, and adaptive sampling (Thompson, 2002: 182; Felix-Medina and Monjardin, 2009:491).

1) Network sampling

Network sampling, or multiplicity sampling, can be a useful technique used to increase the efficiencies of sample surveys in a rare population. The network-based design was first introduced for the study of social networks by Coleman (1958: 28-36). In network sampling, a simple random sample or stratified random sample of units (selection units) is selected, and all of the observation units linked to any of the units selected are included or observed (Nafiu and Adewara, 2007: 5-9). A network is defined to be a set of observation units with a given linkage pattern. Birnbaum and Sirken (1965) proposed unbiased estimators for network sampling.

2) Snowball sampling

Snowball sampling was proposed by Goodman (1961: 148-170). Here, individuals in the random sample are asked to identify a fix number of other individuals, who in turn are asked to identify other individuals for a fixed number of stages for the purpose of estimating the number of mutual relationships in the population.

3) Random walk sampling

Random walk design is a sampling design for obtaining a probability sample of a large social network (Klovdahl et al., 1977: 169). The initial unit is selected by probability sampling. If unit i is selected at wave k, then one of the units linked from i is selected at random until n waves are reached. The random walk sample of size n is then obtained.

Assume that the initial unit is selected at random from the population of size *N*. Let d_{ij} take value 1 if unit *i* links to unit *j*; otherwise, it is 0. Let $d_{i.}$ be the number of links out from unit *i,* where $d_{i.} = \sum_{j=1}^{N} d_{ij}$. Thus, the probability that the initial unit is selected is $q_1 = \frac{1}{N}$. Suppose unit *i* is the current unit in wave *k*-1; the probability that unit *j* is selected in the next wave k is $q_{k_{ij}} = \frac{d_{ij}}{d_{i.}}$. The selection probability for the ordered sample *s* of size *n* is $P(s) = q_1 \prod_{k=2}^{n} q_{k_{ij}}$. An estimator of the mean is the sample mean

$$\bar{y}_1 = \frac{\sum_{i=1}^{n} y_i}{n} \tag{2.5}$$

using data from a random walk sample. This is not an unbiased estimator (Thompson, 2006b: 6). This is not a good estimator. However, one can obtain an approximately unbiased estimator based on the Hansen-Hurwitz estimator

$$\hat{u}_{rws} = \frac{\sum_{i=1}^{n} \frac{y_i}{d_i}}{\sum_{i=1}^{n} \frac{1}{d_i}} \tag{2.6}$$

(Salganik and Heckathorn, 2004: 217-218). This is the ratio estimator of two Hanson-Hurwitz estimators, and it is an asymptotically unbiased estimator with a bias on the order of *n*-1, where *n* is the sample size. Generally, this bias is considered negligible in samples of moderate size.

4) Adaptive sampling

Adaptive sampling design is a sampling design in which the procedure for selecting the units to be included in the sample may depend on values of the variable of interest observed during the survey. An adaptive procedure was proposed by Thompson and Ramsey in 1983. Thompson (1990: 1050-1059) proposed adaptive cluster sampling for rare, clustered populations.

Thompson (2006a: 1-24) proposed a new adaptive sampling design, called the adaptive web sampling (AWS) design, for sampling in network and spatial settings. In the designs, selections are made sequentially with a mixture distribution based on an active set that changes as the sampling progresses, using network or spatial relationships as well as sample values. This design has certain advantages compared with the previously-existing adaptive and link-tracing designs, including control over sample sizes.

Snowball type designs, for example, typically occur in waves, with a whole set of links selected from the previous wave of units or from all the units selected to that point. AWS designs have more flexibility than random walk designs by not being confined to only one unit at a time in the active set. They are more flexible than ordinary network, snowball, and adaptive cluster sampling designs by not requiring every link to be followed from a particular wave; nor do connected components intersected by the sample need to be sampled completely. This flexibility can be used to seek a balance between going deep into the population, following links for many waves, or going wide, with only one or a few waves (Thompson, 2006a: 1-24).

2.2 The Horvitz-Thompson Estimator and Its Application

The Horvitz-Thompson (1952) theorem provides a general theory and methodology for design-based inference from probability samples. The theorem prescribes an estimator to use with any probability sample, and its application to a variety of designs is a powerful heuristic in teaching the similarities and differences among these designs (Overton and Stehman, 1995: 261-268).

Godambe and Joshi (1965: 1707-1722) gave a class of estimators to have a uniformly smaller mean square error than that of the Horvitz-Thompson estimator in the case of (i) unbiasedness and (ii) when fixed sample size requirements are relaxed.

Deshpande (1985: 290-291) showed an estimator based on a non-fixed sample size design which had a smaller mean square error than that of the Horvitz-Thompson estimator.

Taga (1993: 163-173) generalized the Horvitz-Thompson estimator by redefining inclusion probabilities so that the generalized Horvitz-Thompson estimator and its variance formula could be represented in the same form in both cases, with replacement sampling and without replacement sampling. Then, he showed, in the case of a fixed sample size *n* design, that a given strategy with replacement sampling could be improved over a strategy without replacement sampling under suitable conditions.

Godambe (1955: 269-278) established that for any sampling design there does not exist a uniformly minimum variance unbiased estimator of the population total in the class of all linear unbiased estimators. He used the superpopulation concept introduced by Cochran (1946: 164-177) and established that under the class of distribution satisfying $E(Y_i \mid X_i) = aX_i$, $E(Y_i \mid X_i) = \sigma^2 X_i^g$ and $C(Y_i, Y_j \mid X_i, X_j) = 0$, an optimum strategy (with g =2) for which

(i) the inclusion probability of each unit is proportional to the value of the auxiliary information taken on that unit,

(ii) every sample has n distinct units, and

(iii) the estimator used is the corresponding Horvitz-Thompson estimator exists which has a minimum expected variance.

This result opened up the construction of πPS sampling designs, which insisted on non-negative variance estimation.

Hanurav (1962: 429-436) obtained a class of optimal sampling designs best suited for the use of the Horvitz-Thompson estimator and termed them as πPS (π_i's Proportional to Size) sampling designs; these estimate the population total when

auxiliary information is available for all of the units. Since there does not exist a design in which the variance is uniformly minimum, optimal designs are obtained by minimizing the expected variance under a realistic superpopulation set-up. He also showed that when g =2, the Horvitz-Thompson strategy is better than the symmetrised Des Raj strategy.

Vijayan (1966: 87-92) proved that in the usual superpopulation model, the symmetrised Des Raj strategy is superior to the Horvitz-Thompson strategy when g = 1 and inferior when g = 2, except when all p_i's are equal, in which case the two strategies coincide. Note that $p_i = \frac{X_i}{X}$.

Other π ps designs are proposed in much of the literature, for example, by Brewer (1963). The Horvitz-Thompson estimator is used in these π PS designs.

2.3 Sampling Design Using Horvitz-Thompson Estimator

The Horvitz-Thompson estimator is used in the following sampling designs.

2.3.1 Simple Latin Square Sampling ± k Designs

Borkowski (2003: 215-237) proposed simple Latin square sampling ± k designs, which is a new class of probability sampling designs that ensure that the sample is well-distributed over the study region when spatial correlation is present. This design improves the estimation of population abundance. Assume that the study region is partitioned into quadrats which represent the sampling units. A simple Latin square sample +k is composed of a simple Latin square sample (Munholland and Borkowski, 1996) and additional units selected in a systematic fashion. The inclusion probabilities are determined. The Horvitz-Thompson estimator is used in this design.

2.3.2 Adaptive Cluster Sampling

To estimate the population total, the population study area is divided into spatial units (plots) that are generally of the same size (Thompson and Seber, 1996: 8). Adaptive cluster sampling was motivated by the problem of sampling a rare,

clustered population. With adaptive cluster sampling, an initial sample of units is selected and, whenever the value of the variable of interest satisfies the condition, neighboring units are added to the sample.

The usual designed estimators for adaptive cluster sampling with an initial sample taken by with or without replacement are of a Hansen-Hurvitz and Horvitz-Thompson type (Thompson, 1990: 1050-1059). Dryver and Thompson (2005: 157-166) proposed an improved unbiased estimator in adaptive cluster sampling, which is derived by taking the expected value of the usual estimator conditional on a sufficient statistic which is not minimally sufficient. Moreover, Dryver and Chao (2006: 607-620) proposed new ratio estimators under adaptive cluster sampling.

Moreover, the estimator for systematic and strip adaptive cluster sampling and stratified adaptive cluster sampling is also of a Hansen-Hurvitz and Horvitz-Thompson type (Thompson, 1991a: 1103-1115; 1991b: 389-397).

2.3.3 Inverse Sampling

Inverse sampling design is generally considered to be an appropriate technique when the population is divided into two subpopulations, one of which contains only a few units. It is considered to be an efficient strategy to estimate the population total when only a few units represent the characteristic of interest. Mohammadi and Salehi (2011: 1-14) derived the Horvitz-Thompson estimator for the population mean under inverse sampling designs, where subpopulation sizes are known. The formula of inclusion probabilities and joint inclusion probabilities are obtained.

2.3.4 Network Sampling

In network sampling, a simple random sample or stratified random sample of units (selection units) is selected, and all observation units linked to any of the units selected are included or observed. A network is defined to be a set of observation units with a given linkage pattern. The inclusion probability for each network, which is in fact the inclusion probability for any of the observational units within such a network, can be obtained. The Horvitz-Thompson estimator is applied to estimate the population total.

2.3.5 Line-Intercept Sampling

Line-intercept sampling (Lucas and Seber, 1977: 618-622) is appropriate for sampling in a rare population for the purpose of estimating the population total. The inclusion probabilities and joint inclusion probabilities can be obtained by utilizing the width of the shadow cast by an object on the baseline. The Hansen-Hurwitz estimator and Horvitz-Thompson estimator are used in this sampling design.

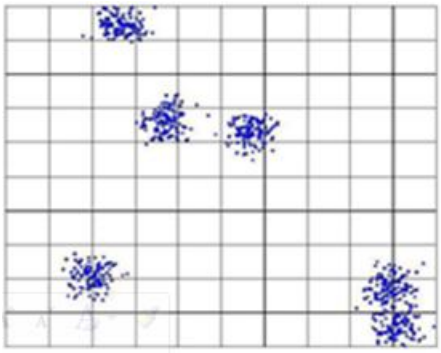
It can be seen that the Horvitz-Thompson estimator can be applied in a sampling design in which inclusion probabilities and joint inclusion probabilities can be obtained.

2.4 Sampling in Spatial Population

2.4.1 The Spatial Population

A spatial setting can be depicted as a geographical area partitioned into single units. For example, in the simulated spatial population presented in Figure 2.1, each unit is represented by a square, and the y_i variables take on the count of the number of point-objects in the square (Vincent, 2008: 4-5).

(a) Population



(b) Population y-values

0	2	80	12	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	19	4	0	0	0	0	0
0	0	0	90	8	68	23	0	0	0
0	0	0	1	0	32	3	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	42	73	0	0	0	0	0	21	23
0	50	54	0	0	0	0	0	57	53
0	0	0	0	0	0	0	0	55	64

Figure 2.1 A Simulated Spatial Population

Sampling in a spatial population, and there are many designs that can be used, for example, simple random sampling, stratified sampling, cluster sampling, and systematic sampling or adaptive sampling in case of rare or clustered population.

2.4.2 Simple Random Sampling in a Spatial Population

Simple random sampling is the most basic form of probability sampling and provides the theoretical basis for the more complicated forms (Lohr, 1999: 30).

Simple random sampling is a sampling design in which n units are randomly selected from the N units in the population. At each selection step, each unit has an equal chance of selection (Thompson, 2002: 11).

There are two ways for taking a simple random sampling: with replacement, in which the same unit may be included more than once in the sample, and without replacement, in which all units in the sample are distinct (Lohr, 1999: 30).

For a given sample of size n, simple random sampling with replacement, SRSWR, is inherently less efficient than simple random sampling without replacement-SRSWOR, (Thompson, 2002: 19). In addition, in a finite population sampling, sampling the same unit twice provides no additional information. We usually prefer to sample without replacement, so that the sample contains no duplicates (Lohr, 1999: 30). Hence, now we consider only SRSWOR.

Simple random sampling without replacement is a sampling design in which n distinct units are selected from the N units in the population in such a way that every possible combination of n units is equally likely to be the sample selected. The sample may be obtained through n selections in which at each step every unit of the population not already selected has an equal chance of selection. Equivalently, one may make a sequence of independent selections from the whole population, each unit having equal probability of selection at each step, discarding repeat selections and counting until n distinct units are obtained.

A simple random sample of $n = 10$ units from a population of $N = 100$ units is depicted in Figure 2.2. Another simple random sample is shown in Figure 2.3. Each such combination of 10 units has an equal probability of being the sample selected. With simple random sampling, the probability that the i^{th} unit of the population is

included in the sample is $\pi_i = n/N$, so that the inclusion probability is the same for each unit. Additionally, each possible sample of n units has the same probability. That is, the probability of selecting a sample s is $P(s) = \frac{1}{\binom{N}{n}}$. The unbiased estimator of the population mean is the sample mean $\bar{y}$, which is the average of the y-values in the sample of size n. That is,

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{2.7}$$

The variance of the estimator $\bar{y}$ is

$$v(\bar{y}) = \left(1 - \frac{n}{N}\right)\frac{\sigma^2}{n}, \tag{2.8}$$

where $\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu)^2$ is the population variance. An unbiased estimator of this variance is

$$\hat{v}(\bar{y}) = \left(1 - \frac{n}{N}\right)\frac{s^2}{n}, \tag{2.9}$$

where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ is the sample variance, which is an unbiased estimator of σ^2.

Notice that the simple random samples in Figure 2.2 and 2.3 cover all of the study region. To observe all sample units, the sampler must travel from unit to unit until every unit is observed. Unfortunately, it may be costly due to sampling travel if the population is quite a large region. Thus, this paper proposes a new sampling design, path sampling, to overcome this drawback.

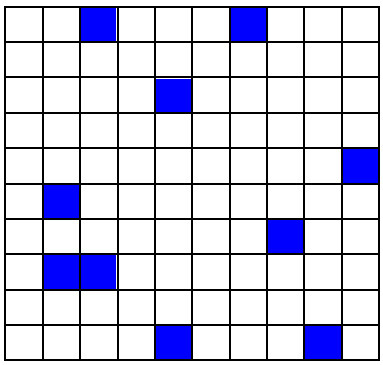


Figure 2.2 A Simple Random Sample of 10 Units from a Population of 100 Units

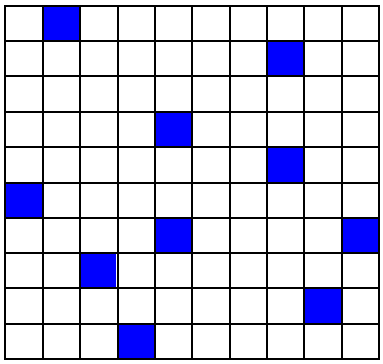


Figure 2.3 Another Simple Random Sample of 10 Units.

2.4.3 Stratified Sampling in a Spatial Population

In stratified sampling, the population is partitioned into regions or strata, and a sample is selected by some design within each stratum. Because the selections in different strata are made independently, the variance of estimators for individual strata can be added together to obtain variances of estimators for the whole population. Since only the within-stratum variances enter into the variances of the estimator, the principle of stratification is to partition the population in such a way that units within a stratum are as similar as possible. Then, even though one stratum may differ markedly from another, a stratified sample with the desired number of units from each stratum in the population will tend to be "representative" of the population as a whole.

A geographical region may be stratified into similar areas by means of some known variable, such as habitat type, elevation, or soil type. Even if a large geographic study area appears to be homogeneous, stratification into blocks can help ensure that the sample is spread out over the entire study area. Human populations may be stratified on the basis of geographic region, city size, sex, or socioeconomic factors (Thompson, 2002: 117-118).

In the following, it is assumed that a sample is selected by some probability design from each of the strata in the population, with selections in different strata independent of each other.

The design is called stratified random sampling if the design within each stratum is simple random sampling. Figure 2.4 shows a stratified random sample from a population of $N = 400$ units. The size of the $L = 4$ strata are $N_1 = 200$, $N_2 = 100$, $N_3 = N_4 = 50$. Within each stratum, a random sample without replacement has been selected independently. The total sample size of 40 has been allocated proportional to stratum size, so that $n_1 = 20$, $n_2 = 10$, and $n_3 = n_4 = 5$.

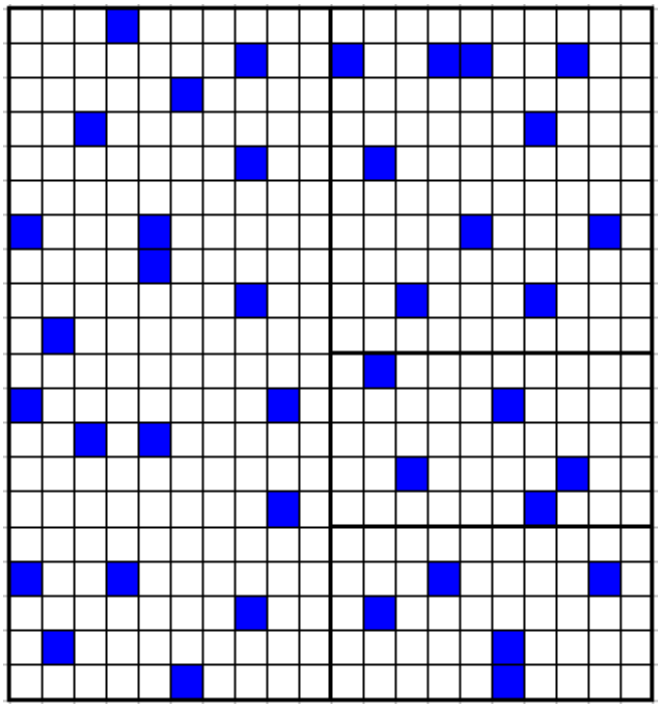


Figure 2.4 Stratified Random Sample in a Spatial Population

2.4.4 Cluster and Systematic Sampling in a Spatial Population

Although systematic sampling and cluster sampling seem on the surface to be opposites, the two designs share the same structure. The population is partitioned into a primary unit, each primary unit being composed of secondary units. Whenever a primary unit is included in the sample, the y-values of every secondary unit within it are observed (Thompson, 2002: 129-132).

In systematic sampling, a single primary unit consists of secondary units spaced in a systematic fashion throughout the population. In cluster sampling, a primary unit consists of a cluster of secondary units, usually in close proximity to each other. In the spatial setting, a systematic sample primary unit may be composed of a collection of plots in a grid pattern over the study area. Cluster primary units include such spatial arrangements as square collections of adjacent plots or long, narrow strips of adjacent units. A cluster sample consisting of a simple random sample of 40 primary units, each consisting of eight secondary units, is shown in Figure 2.5. A systematic sample with two randomly selected units is shown in Figure 2.6. The systematic sample consists of two primary units, each with 16 secondary units.

The key point in any of the systematic or clustered arrangements is that whenever any secondary unit of a primary unit is included in the sample, all of the secondary units of that primary unit are included. Even though the actual measurements may be made on secondary units, it is the primary units that are selected.

In systematic sampling, it is not uncommon to have a sample size of 1; that is, a single primary unit (Thompson, 2002: 129).

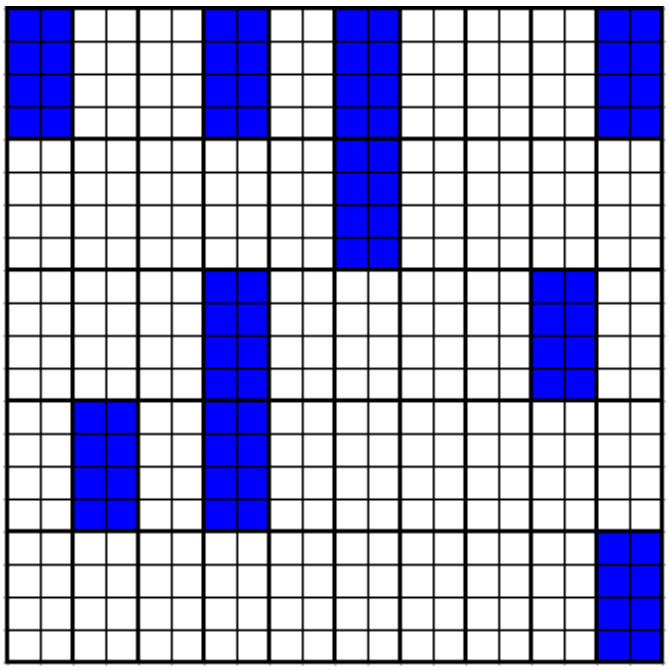


Figure 2.5 Cluster Sample

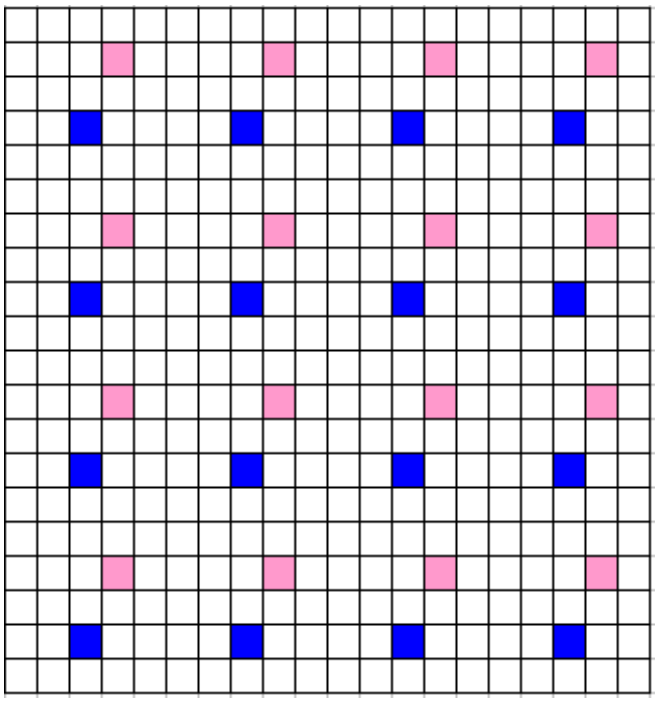


Figure 2.6 A Systematic Sample with Two Starting Points

Let N_p be the number of primary units in the population, n_p be the number of primary units in the sample, and M_i be the number of secondary unit in the i^{th} primary unit. The total number of secondary units in the population is $M = \sum_{i=1}^{N} M_i$. Let y_{ij} denote the value of the variable of interest of the j^{th} secondary unit in the i^{th} primary unit. The total of the y-values in the i^{th} primary unit is denoted by $y_i = \sum_{j=1}^{M_i} y_{ij}$. The population total is $\tau = \sum_{i=1}^{N}\sum_{j=1}^{M_i} y_{ij} = \sum_{i=1}^{N} y_i$. The population mean per primary unit is $\mu_1 = \frac{\tau}{N_p}$. The population mean per secondary unit is $\mu = \frac{\tau}{M}$.

The unbiased estimator of the mean per secondary unit is (Thompson, 2002: 132)

$$\hat{\mu}_{cls} = \frac{N_p \bar{y}}{M}, \tag{2.10}$$

where $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n_p}$ is the sample mean of the primary unit totals. The variance of $\hat{\mu}_{cls}$ is

$$v(\hat{\mu}_{cls}) = N_p(N_p - n_p)\frac{\sigma_u^2}{n_p M^2}, \tag{2.11}$$

where $\sigma_u^2 = \frac{1}{N_p - 1}\sum_{i=1}^{N_p}(y_i - \mu_1)^2$. The estimator of $v(\hat{\mu}_{cls})$ is

$$\hat{v}(\hat{\mu}_{cls}) = N_p(N_p - n_p)\frac{s_u^2}{n_p M^2}, \tag{2.12}$$

where $s_u^2 = \frac{1}{n_p - 1}\sum_{i=1}^{n_p}(y_i - \bar{y})^2$.

The advantage of cluster sampling is that it is often less costly to sample a collection of units in a cluster than to sample an equal number of secondary units selected at random from the population.

CHAPTER 3

PATH SAMPLING DESIGN AND ESTIMATION

This chapter details the definition of all possible paths in the spatial population, the path sampling scheme, and estimation.

Suppose the researcher's goal is to estimate the population mean of a study variable. Initially, it will be assumed that the study region can be partitioned into an $r \times c$ (r rows and c columns) grid of rc quadrats or secondary units. The population consists of rc spatial units. Each population unit is labeled with 2 coordinates, say (i, j), which are the row and column of the unit, respectively, for i = 1, 2, 3,…, r and j = 1, 2, 3, …, c. Associated with each unit (i, j) , the value of the population variable of interest is denoted as $y_{(i,j)}$. The parameter of interest in this paper is the population mean of y,

$$\mu = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c} y_{(i,j)} = \frac{1}{rc}\sum_{all\,(i,j)} y_{(i,j)} \,. \tag{3.1}$$

Path sampling design is a sampling design in which p distinct paths are selected by simple random sampling without replacement from q possible paths in the population, and the sample consists of all units in the selected paths. Thus, a path(s) is chosen instead of units. In this study, path sampling is used for the spatial population.

3.1 All Possible Paths in the Spatial Population

A path is basically the path or route taken from start to finish. Let q be the number of all possible paths. Let P_k denote the path k for k = 1, 2, 3, …, q. A path will be defined to start from row 1 and column j^*; that is, a unit labeled (1, j^*) is a starting unit, and ends at a unit (1, j^*+1). If a starting unit is inside a region, a

researcher has to travel from the edge of a region to such a starting unit, which is time consuming and of more distance without observing the units in the sample. We start sampling at an edge, at unit $(1, j^*)$, of a region because it is more convenient and less expensive than starting inside or in the middle of a region. The path k taken will be from such a starting unit and then goes to a particular row, say row k, to the end of that row on the left and then goes along row $k+1$ and comes back to the starting unit. That is, the path k taken will be from $(1, j^*)$ to $(2, j^*)$ then to (k, j^*) to (k, j^*-1) to (k, j^*-2) to $(k, 1)$ to $(k+1, 1)$ to $(k+1, 2)$ to $(k+1, c)$ to (k, c) to $(k, c-1)$ to $(k, c-2)$ to (k, j^*+1) to $(k,-1\ j^*+1)$ to $(k,-2\ j^*+1)$ and finally to $(1, j^*+1)$. Thus, for a spatial population of r rows, there are $q = r-1$ possible paths. In general, a path k in the spatial setting population of r rows and c columns can be written as

$$P_k = ((1, j^*), (2, j^*), (3, j^*), \ldots, (k, j^*), (k, j^*-1), (k, j^*-2), \ldots, (k,1), (k+1, 1),$$
$$(k+1,2), \ldots, (k+1,c), (k,c), (k,c-1), (k,c-2), \ldots, (k, j^*+1), (k-1, j^*+1),$$
$$(k-2, j^*+1), \ldots, (1, j*+1))$$

for $k = 1, 2, 3, \ldots, q = r-1$.

The number of units belonging to path P_k is $2c + 2(k-1)$. All possible paths are shown in Figure 3.1. Notice that the numbers of units in each path are not the same. We can see that the paths overlap in column j^* and j^*+1, which are in the going-out and coming-back column, respectively. Also, the paths next to each other overlap in the row between them. Thus, it can be said that path k-1 and path k overlap in row k for $k = 2, 3, \ldots, q = r-1$. It is assumed that the units are sampled in a logical manner such that all units will only be observed once. Finally, the researcher can define the rows and columns arbitrarily; thus, path sampling is not limited in its starting and ending position

In addition, there are four different edges in the rectangular region, so there may be four different starting points to be chosen. However, it can be rotated to set the starting unit at a starting unit $(1, j^*)$.

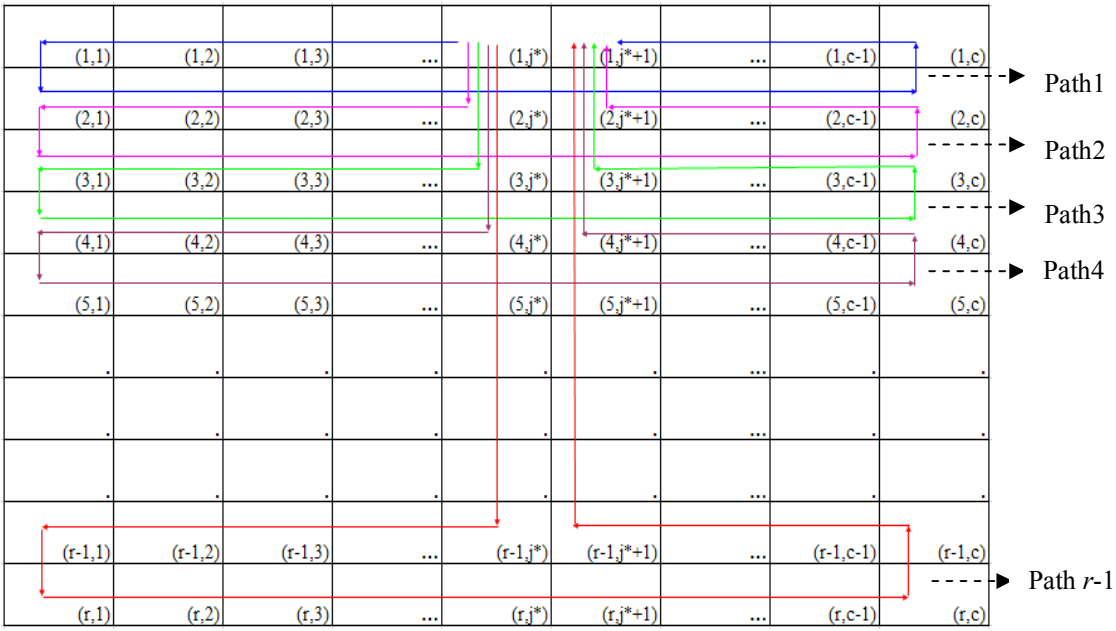


Figure 3.1 All Possible Paths with a Starting Unit $(1, j^*)$ and Each Unit Labeled with 2 Coordinates

Example: Define all possible paths when the starting unit $(1, j^*)$ is (1, 1)

In this example, a path will be defined as starting from row 1 and column 1, that is the unit labeled (1,1). That is $j^* = 1$. A row will be randomly selected from all rows, say row k. Then the path taken will be from (1, 1) to (2, 1) then to (k+1, 1) to (k+1, 2) to (k+1, c) to (k, c) to (k, c-1) to (k, 2) to (k-1, 2) to (2,2) and to (2, 1). From P_k on page 25, path k in the spatial population of r rows and c columns with starting unit (1, 1) or $j^* = 1$ can be written as

$$P_k = ((1,1), (2,1), (3,1),...,(k+1,1), (k+1,2), (k+1,3),..., (k+1,c), (k,c), (k,c-1), (k,c-2), ...,(k,2), (k-1,2), (k-2, 2), ..., (1,2))$$

for $k = 1, 2, 3, ..., q = r\text{-}1$.

The number of units belonging to P_k is $2c + 2(k\text{-}1)$. All possible paths are shown in Figure 3.2. There are $q = r\text{-}1$ possible paths. We can see that the paths

overlap in the first and second column. Also, the paths next to each other overlap in the row between them.

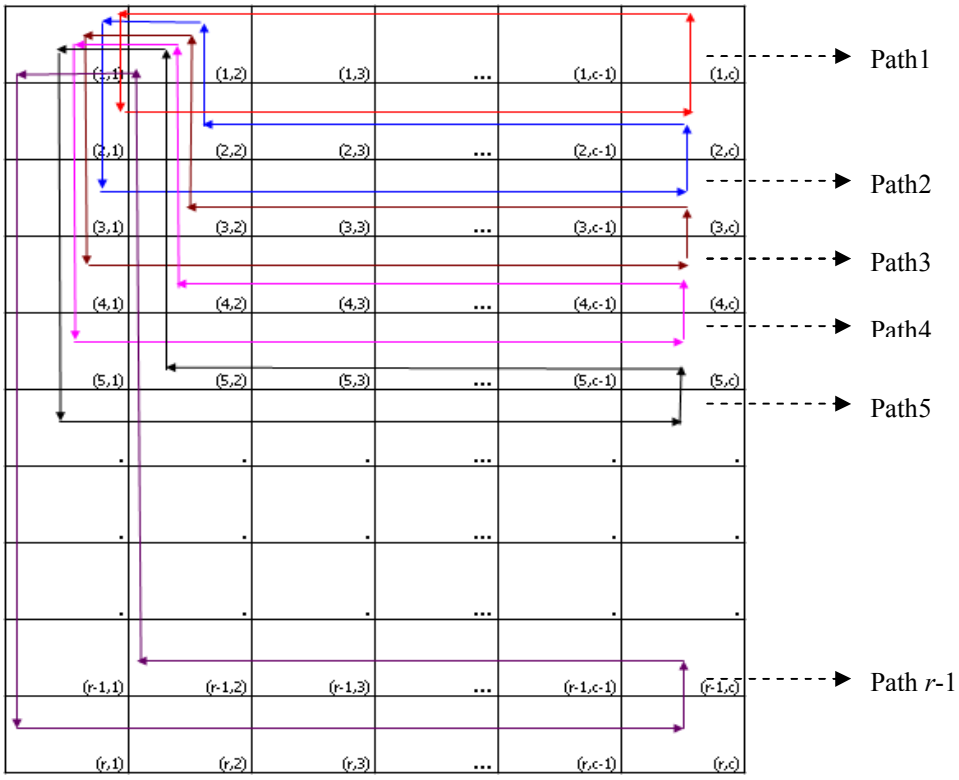


Figure 3.2 All Possible Paths with Starting Unit (1,1) and Each Unit Labeled with 2 Coordinates

To illustrate paths, the spatial setting population of 8 rows and 4 columns is considered, as shown in Figure 3.3. So, we have $r = 8$ and $c = 4$. Hence, there are $q = r\text{-}1 = 7$ possible paths, which have the same starting and ending unit. Notice that the seven paths overlap in the first and second columns. Path 1 and path 2 overlap in row 2; path 2 and path 3 overlap in row 3; path 3 and path 4 overlap in row 4 and so on.

The paths and labeled units belonging to them are shown in Table 3.1. Notice that the numbers of units in the seven paths are different. Now we consider the sample paths of size 2. Suppose that path 3 and path 4 are selected in the sample. They overlap in column 1 and 2 and in row 4. Thus, the overlap units, which are (1, 1), (1,

2), (2, 1), (2, 2), (3, 1), and (3, 2), are repeat observations. It is assumed that the repeat observation has the same value each time observed.

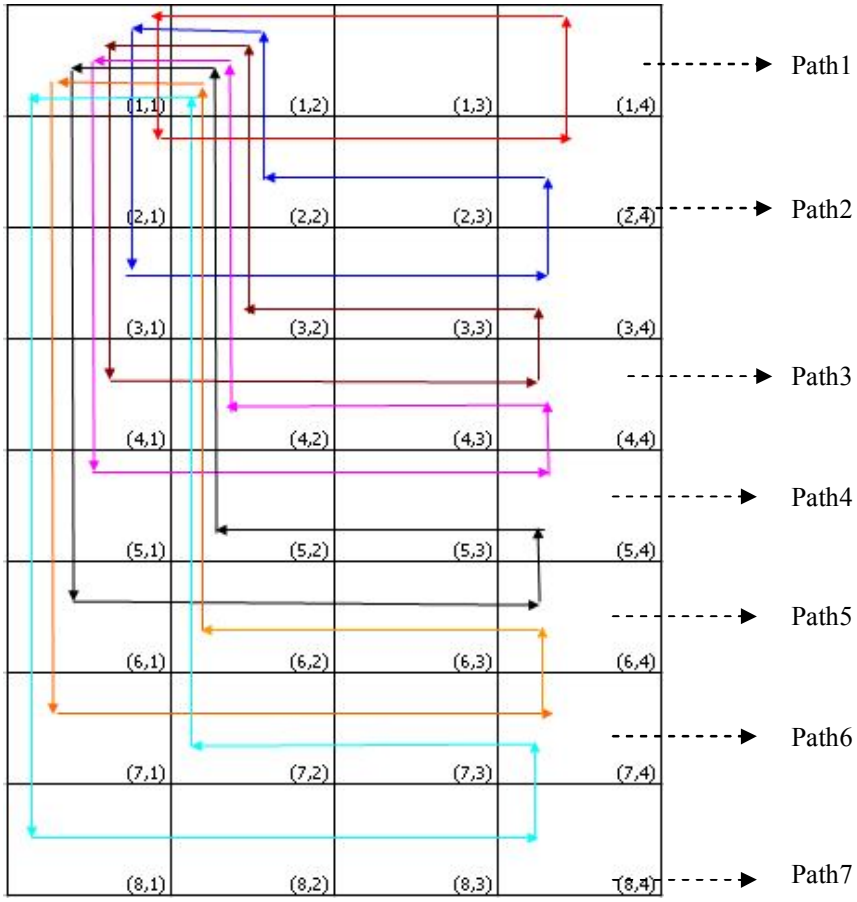


Figure 3.3 The Population Units Labeled with 2 Coordinates and All Possible Paths in a Population of 8 Rows and 4 Columns

Table 3.1 All Possible Paths and Their Units Labeled With Starting Unit (1,1)

All possible paths in the population						
P_1	P_2	P_3	P_4	P_5	P_6	P_7
(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)
(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)
(2,2)	(3,1)	(3,1)	(3,1)	(3,1)	(3,1)	(3,1)
(2,3)	(3,2)	(4,1)	(4,1)	(4,1)	(4,1)	(4,1)
(2,4)	(3,3)	(4,2)	(5,1)	(5,1)	(5,1)	(5,1)
(1,4)	(3,4)	(4,3)	(5,2)	(6,1)	(6,1)	(6,1)
(1,3)	(2,4)	(4,4)	(5,3)	(6,2)	(7,1)	(7,1)
(1,2)	(2,3)	(3,4)	(5,4)	(6,3)	(7,2)	(8,1)
	(2,2)	(3,3)	(4,4)	(6,4)	(7,3)	(8,2)
	(1,2)	(3,2)	(4,3)	(5,4)	(7,4)	(8,3)
		(2,2)	(4,2)	(5,3)	(6,4)	(8,4)
		(1,2)	(3,2)	(5,2)	(6,3)	(7,4)
			(2,2)	(4,2)	(6,2)	(7,3)
			(1,2)	(3,2)	(5,2)	(7,2)
				(2,2)	(4,2)	(6,2)
				(1,2)	(3,2)	(5,2)
					(2,2)	(4,2)
					(1,2)	(3,2)
						(2,2)
						(1,2)

Example: Define All Possible Paths When the Starting Unit $(1, j^*)$ is (1,3)

To illustrate paths, the spatial population of 8 rows and 6 columns is considered, as shown in Figure 3.4. So, we have $r = 8$ and $c = 6$. Hence, there are $q = r\text{-}1 = 7$ possible paths, which have the same starting and ending unit. Suppose that unit (1, 3) is the starting point. Notice that the seven paths overlap in column 3 and 4. Path 1 and path 2 overlap in row 2; path 2 and path 3 overlap in row 3; path 3 and path 4 overlap in row 4 and so on.

The paths and the labeled units belonging to them are shown in Table 3.2**.** Notice that the numbers of units in the seven paths are different.

Now we consider the sample paths of size 2. Suppose path 3 and path 4 are selected in the sample. They overlap in column 3 and 4 and in row 4. Thus, the overlapping units, which are (1, 3), (2, 3), (3, 3), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), (3, 4), (2, 4), and (1,4), are repeat observations. It is assumed that the repeat observation has the same value each time observed.

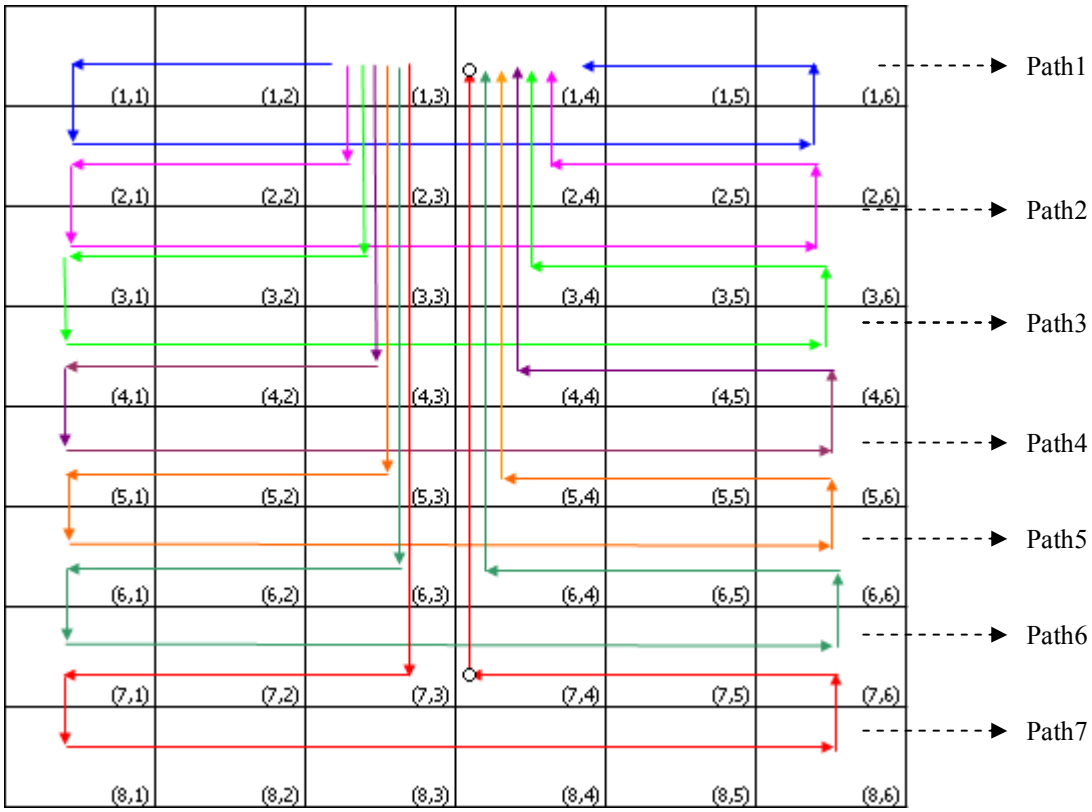


Figure 3.4 The Population Units Labeled with 2 Coordinates and All Possible Paths

Table 3.2 All Possible Paths and Their Units Labeled

All possible paths in the population						
P_1	P_2	P_3	P_4	P_5	P_6	P_7
(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)
(1,2)	(2,3)	(2,3)	(2,3)	(2,3)	(2,3)	(2,3)
(1,1)	(2,2)	(3,3)	(3,3)	(3,3)	(3,3)	(3,3)
(2,1)	(2,1)	(3,2)	(4,3)	(4,3)	(4,3)	(4,3)
(2,2)	(3,1)	(3,1)	(4,2)	(5,3)	(5,3)	(5,3)
(2,3)	(3,2)	(4,1)	(4,1)	(5,2)	(6,3)	(6,3)
(2,4)	(3,3)	(4,2)	(5,1)	(5,1)	(6,2)	(7,3)
(2,5)	(3,4)	(4,3)	(5,2)	(6,1)	(6,1)	(7,2)
(2, 6)	(3,5)	(4,4)	(5,3)	(6,2)	(7,1)	(7,1)
(1, 6)	(3,6)	(4,5)	(5,4)	(6,3)	(7,2)	(8,1)
(1, 5)	(2,6)	(4,6)	(5,5)	(6,4)	(7,3)	(8,2)
(1, 4)	(2,5)	(3,6)	(5,6)	(6,5)	(7,4)	(8,3)
	(2,4)	(3,5)	(4,6)	(6,6)	(7,5)	(8,4)
	(1,4)	(3,4)	(4,5)	(5,6)	(7,6)	(8,5)
		(2,4)	(4,4)	(5,5)	(6,6)	(8,6)
		(1,4)	(3,4)	(5,4)	(6,5)	(7,6)
			(2,4)	(4,4)	(6,4)	(7,5)
			(1,4)	(3,4)	(5,4)	(7,4)
				(2,4)	(4,4)	(6,4)
				(1,4)	(3,4)	(5,4)
					(2,4)	(4,4)
					(1,4)	(3,4)
						(2,4)
						(1,4)

3.2 Path Sampling Design

The spatial population of r rows and c columns consists of units labeled *(i,j)* for i=1, 2, 3, …, r and j = 1, 2, 3, …, c. There are q = r-1 possible paths in the population denoted by $P_1, P_2, P_3, ..., P_q$. By applying SRSWOR, p paths are selected from q possible paths in the population. Let p_k denote path k in the sample for k = 1, 2, 3, …, p. The sample consists of all units in the selected paths. The sample is represented as

$$p_s = (p_1, p_2, p_3, ..., p_p).$$

The probability of selecting a sample is $P(s) = \frac{1}{\binom{q}{p}} = \frac{1}{\binom{r-1}{p}}$ since paths are selected by SRSWOR and the inclusion probability of path k is $\pi_k = \frac{p}{q} = \frac{p}{r-1}$. That is, each path has an equal probability of selection. There are overlapping of paths, so there are repeat observations. Assume that the repeat observation has the same value each time it is observed.

3.2.1 Inclusion Probability

Suppose a unit $(1, j^*)$ is the starting unit. The inclusion probability of each unit is the probability that a unit is included in the sample. In path sampling, the inclusion probability of unit (i, j) is denoted as $\pi_{(i,j)}$.

Since paths overlap in rows and columns, the probabilities that units are included in the sample are not equal. That is, the inclusion probabilities of each unit in a path are not equal. All paths overlap in column j^* and j^*+1, and some paths overlap in row. Thus, the inclusion probabilities can be divided into three cases due to the overlapping of paths.

$$\pi_{(i,j)} = \begin{cases} 1-\dfrac{\binom{i-2}{p}}{\binom{q}{p}} & i = 1,2,3,\ldots,r \ \text{ and } \ j = j^{*} \text{ and } j^{*}+1 \\ 1-\dfrac{\binom{q-2}{p}}{\binom{q}{p}} & i = 2,3,\ldots,r-1 \ \text{ and } \ j = 1,2,3,\ldots,j^{*}\text{-}1, j^{*}+2, j^{*}+3,\ldots,c \\ 1-\dfrac{\binom{q-1}{p}}{\binom{q}{p}} & i = 1, r \text{ and } j = 1,2,3,\ldots,j^{*}\text{-}1, j^{*}+2, j^{*}+3,\ldots,c \end{cases} \tag{3.2}$$

Note that, for any constant $a < b$, it is defined that $\binom{a}{b} = 0$.

According to equation (3.2), each case of the formula is the inclusion probability for each type of unit. That is, the first case is the inclusion probability for the units in column $j^{*}$and j^{*}+1 (units of type 1). The second is for the units not in column $j^{*}$and j^{*}+1and not in the first or last row (units of type 2); and the third are for units in the first row and last row but not in column $j^{*}$and j^{*}+1 (units of type 3). Units of type 1, 2, and 3 in the population are shown in Figure 3.5. A proof of the inclusion probability in equation (3.2) can be found in the next section.

3.2.1.1 Proof of the Inclusion Probability

Case 1: For units in column $j^{*}$and j^{*}+1 (units of type 1), all paths overlap in column $j^{*}$and j^{*}+1, so the inclusion probabilities for the units in these columns are higher than in other columns. The inclusion probabilities for units of type 1can be written as

$$\begin{aligned} \pi_{(i,j)} &= P(unit\,(i,j)\ is\ in\ the\ sample) \\ &= 1 - P(unit\,(i,j)\ is\ not\ in\ the\ sample) \\ &= 1 - \frac{The\ number\ of\ sample\ not\ containing\ unit\,(i,j)}{The\ number\ of\ all\ possible\ sample} \end{aligned}$$

for i = 1, 2, 3, …, r and j= j^{*} and j^{*}+1.

The number of paths not containing units of type 1 (*i, j*) is *i*-2. A proof of this is shown in section 3.2.1.2. Thus, the number of samples not containing such units is $\binom{i-2}{p}$.

Then, we have

$$\pi_{(i,j)} = 1 - \frac{\binom{i-2}{p}}{\binom{q}{p}}. \tag{3.3}$$

Note that, for any constant a < b, it is defined that $\binom{a}{b} = 0$.

Case 2: For units not in column j^* or j^*+1 and not in the first row or the last row (units of type 2), the paths next to each other overlap in the row between them. That is, Path *k*-1 and path *k* overlap in row *k* for *k* = 2, 3,..., *q* = *r*-1. Thus, units of type 2 belong to two consecutive paths. Thus, the number of paths not containing such units is *q*-2, the proof of which is shown in section 3.1.1.2, and then the number of samples not containing such units is $\binom{q-2}{p}$. Hence, the inclusion probabilities for units of type 2 are

$$\begin{aligned}\pi_{(i,j)} &= 1 - \frac{The\ number\ of\ sample\ not\ containing\ unit\ (i,j)}{The\ number\ of\ all\ possible\ sample} \\ &= 1 - \frac{\binom{q-2}{p}}{\binom{q}{p}}\end{aligned} \tag{3.4}$$

for *i* = 2, 3, …, *r*-1 and *j* = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c.

Case3: For the units in the first row and last row but not in column j^* or j^*+1 (units of type 3), they belong to only one path. That is, unit (1, 1), (1, 2), (1, 3),

…, (1, j^*-1), (1, j^*+2), (1, j^*+3), …, (1, c-1), (1, c) belong to only one path, which is path 1. Also, units (*r*, 1), (*r*, 2), (*r*, 3), …, (*r*, j^*-1), (*r*, j^*+2), (*r*, j^*+3), …, (*r*, *c*-1), (*r*, *c*) belong to only one path, which is path *r*-1. Thus, the number of paths not containing such units is *q*-1. The number of samples not containing such units is $\binom{q-1}{p}$. Hence, for $i = 1$ and r and $j = 1, 2, 3, \ldots, j^*-1, j^*+2, j^*+3, \ldots, c$.

$$\pi_{(i,j)} = 1 - \frac{The\ number\ of\ sample\ not\ containing\ unit\ (i,j)}{The\ number\ of\ all\ possible\ sample}$$

$$= 1 - \frac{\binom{q-1}{p}}{\binom{q}{p}} \tag{3.5}$$

3.2.1.2 Proof of the Number of Paths not Containing Units of Type 1 is *i*-2

According to the principle of mathematical induction, it will be shown that the number of paths not containing a unit (*i, j*) type 1 is *i*-2 for $i \geq 1$.

Base case: $i = 1$. A unit (1, *j*) of type 1 belongs to $P_1, P_2, P_3, ..., P_{r-1}$. Note that the negative number of the number of paths not containing unit (1, j^*) is set equal to 0. Thus, the number of paths not containing unit (1, j^*) is *i*-2 =-1, which is equal to 0.

Induction step: suppose the number of paths not containing a unit (*i, j*) is *i*-2; we need to prove the number of paths not containing unit (*i+1, j*) is *i*+1-2= *i*–1. Unit (*i*+1, j^*) belongs to $P_i, P_{i+1}, P_{i+2}, ..., P_{r-1}$. It is not in $P_1, P_2, P_3, ..., P_{i-1}$. Thus, the number of paths not containing unit (*i*, j^*) is *i*-1.

3.2.1.3 Proof of the Number of Paths not Containing Units of Type 2 is *q*-2

For unit (i, *j*) where $i = 2, 3, \ldots, r$-1 and $j = 1, 2, 3, \ldots, j^*-1, j^*+2, j^*+3, \ldots$, c, unit (*i, j*) belongs to P_{i-1} and P_i. In general, the number of paths containing unit (*i, j*) is 2; thus the number of paths not containing such a unit is *q*-2.

3.2.1.4 Example: The Inclusion Probabilities in the Path with Starting Unit $(1, j^*)$=(1, 1)

With starting unit (1, 1), all paths overlap in column 1 and 2, and some paths overlap in the row. Thus, the inclusion probabilities for the three cases due to the overlapping of paths are as follows.

Case 1: The Inclusion Probabilities for Units of Type 1

All paths overlap in column 1 and 2, so the inclusion probabilities for the units in these columns are higher than in the other columns. The inclusion probabilities for the units in column 1 and 2 can be written as

$$\begin{aligned}\pi_{(i,j)} &= P(unit\,(i,j)\ is\ in\ the\ sample)\\ &= 1-P(unit\,(i,j)\ is\ not\ in\ the\ sample)\\ &= 1-\frac{The\ number\ of\ sample\ not\ containing\ unit\,(i,j)}{The\ number\ of\ all\ possible\ sample}\end{aligned}$$

for i = 1, 2, 3, …, r and j= 1 and 2.

The number of paths not containing units (i, j) in column 1 and 2 is i-2. For example, the number of paths not containing unit (2, 1) is i-2 = 2-2 = 0. Notice that unit (2, 1) belongs to every path. Unit (3, 1) belongs to all paths except path 1; thus the number of paths not containing such a unit is 1= i-2 = 3-2. Unit (5, 1) does not belong to path 1, 2 or 3; thus the number of paths not containing such a unit is 3= i-2 = 5-2. Consequently, the number of paths not containing any unit *(i, j)* not in column 1 and 2 is i-2. The number of samples not containing such units is $\binom{i-2}{p}$.

Hence,

$$\pi_{(i,j)} = 1-\frac{\binom{i-2}{p}}{\binom{q}{p}}$$

for i = 1, 2, 3, …, r and j= 1 and 2.

Note that, for any constant a < b, it is defined that $\binom{a}{b}=0$

(a) Units in Column j^* and j^*+1 (Units of type 1)

(1,1)	(1,2)	(1,3)	...	(1,j*-1)	(1,j*)	(1,j*+1)	(1,j*+2)	...	(1,c-1)	(1,c)
(2,1)	(2,2)	(2,3)	...	(2,j*-1)	(2,j*)	(2,j*+1)	(2,j*+2)	...	(2,c-1)	(2,c)
(3,1)	(3,2)	(3,3)	...	(3,j*-1)	(3,j*)	(3,j*+1)	(3,j*+2)	...	(3,c-1)	(3,c)
(4,1)	(4,2)	(4,3)	...	(4,j*-1)	(4,j*)	(4,j*+1)	(4,j*+2)	...	(4,c-1)	(4,c)
(5,1)	(5,2)	(5,3)	...	(5,j*-1)	(5,j*)	(5,j*+1)	(5,j*+2)	...	(5,c-1)	(5,c)
.	.	.	.	.	.	.	.	...	.	.
.	.	.	.	.	.	.	.	...	.	.
.	.	.	.	.	.	.	.	...	.	.
(r-1,1)	(r-1,2)	(r-1,3)	...	(r-1,j*-1)	(r-1,j*)	(r-1,j*+1)	(r-1,j*+2)	...	(r-1,c-1)	(r-1,c)
(r,1)	(r,2)	(r,3)	...	(r,j*-1)	(r,j*)	(r,j*+1)	(r,j*+2)	...	(r,c-1)	(r,c)

(b) Units not in Column j^* and $j^*+1$and Not in the First Row or Last Row (Units of type 2)

(1,1)	(1,2)	(1,3)	...	(1,j*-1)	(1,j*)	(1,j*+1)	(1,j*+2)	...	(1,c-1)	(1,c)
(2,1)	(2,2)	(2,3)	...	(2,j*-1)	(2,j*)	(2,j*+1)	(2,j*+2)	...	(2,c-1)	(2,c)
(3,1)	(3,2)	(3,3)	...	(3,j*-1)	(3,j*)	(3,j*+1)	(3,j*+2)	...	(3,c-1)	(3,c)
(4,1)	(4,2)	(4,3)	...	(4,j*-1)	(4,j*)	(4,j*+1)	(4,j*+2)	...	(4,c-1)	(4,c)
(5,1)	(5,2)	(5,3)	...	(5,j*-1)	(5,j*)	(5,j*+1)	(5,j*+2)	...	(5,c-1)	(5,c)
.	.	.	.	.	.	.	.	...	.	.
.	.	.	.	.	.	.	.	...	.	.
.	.	.	.	.	.	.	.	...	.	.
(r-1,1)	(r-1,2)	(r-1,3)	...	(r-1,j*-1)	(r-1,j*)	(r-1,j*+1)	(r-1,j*+2)	...	(r-1,c-1)	(r-1,c)
(r,1)	(r,2)	(r,3)	...	(r,j*-1)	(r,j*)	(r,j*+1)	(r,j*+2)	...	(r,c-1)	(r,c)

(c) Units in the First row and Last Row but Not in Column j^* and j^*+1 (Units of type 3)

(1,1)	(1,2)	(1,3)	...	(1,j*-1)	(1,j*)	(1,j*+1)	(1,j*+2)	...	(1,c-1)	(1,c)
(2,1)	(2,2)	(2,3)	...	(2,j*-1)	(2,j*)	(2,j*+1)	(2,j*+2)	...	(2,c-1)	(2,c)
(3,1)	(3,2)	(3,3)	...	(3,j*-1)	(3,j*)	(3,j*+1)	(3,j*+2)	...	(3,c-1)	(3,c)
(4,1)	(4,2)	(4,3)	...	(4,j*-1)	(4,j*)	(4,j*+1)	(4,j*+2)	...	(4,c-1)	(4,c)
(5,1)	(5,2)	(5,3)	...	(5,j*-1)	(5,j*)	(5,j*+1)	(5,j*+2)	...	(5,c-1)	(5,c)
.	.	.	.	.	.	.	.	...	.	.
.	.	.	.	.	.	.	.	...	.	.
.	.	.	.	.	.	.	.	...	.	.
(r-1,1)	(r-1,2)	(r-1,3)	...	(r-1,j*-1)	(r-1,j*)	(r-1,j*+1)	(r-1,j*+2)	...	(r-1,c-1)	(r-1,c)
(r,1)	(r,2)	(r,3)	...	(r,j*-1)	(r,j*)	(r,j*+1)	(r,j*+2)	...	(r,c-1)	(r,c)

Figure 3.5 Units of Type 1, 2, and 3

Case 2: The Inclusion Probabilities for Units of Type 2

The paths next to each other overlap in the row between them. That is, Path k-1 and path k overlap in row k for k = 2, 3,..., q = r-1. Thus, the units not in column 1 and 2 and not in the first row and the last row belong to two consecutive paths. As a result, the number of paths not containing such units is q-2, and the number of samples not containing such units is $\binom{q-2}{p}$. Hence, the inclusion probabilities for the units not in column 1 or 2 and not in the first or last row are

$$\begin{aligned}\pi_{(i,j)} &= P(unit\,(i,j)\ is\ in\ the\ sample)\\ &= 1 - P(unit\,(i,j)\ is\ not\ in\ the\ sample)\\ &= 1 - \frac{The\ number\ of\ sample\ not\ containing\ unit\,(i,j)}{The\ number\ of\ all\ possible\ sample}\\ &= 1 - \frac{\binom{q-2}{p}}{\binom{q}{p}}\end{aligned}$$

for i = 2, 3, …, r-1 and j = 3, 4, 5,…, c.

We can see that such units have the same inclusion probabilities.

Case 3: The Inclusion Probabilities for Units of Type 3

The units in the first row and the last row and not in column 1 and 2 belong to only one path. That is, units (1, 3), (1, 4), (1, 5),…, (1, c) belong to only one path, which is path 1. Also, units (r, 3), (r, 4), (r, 5),…, (r, c) belong to only one path, which is path r-1. Thus, the number of paths not containing such units is q-1. The number of samples not containing such units is $\binom{q-1}{p}$.

Hence, for i = 1 and r and j = 3, 4, 5,…, c

$$\begin{aligned}\pi_{(i,j)} &= P(unit\,(i,j)\ is\ in\ the\ sample)\\ &= 1-P(unit\,(i,j)\ is\ not\ in\ the\ sample)\\ &= 1-\frac{The\ number\ of\ sample\ not\ containing\ unit\,(i,j)}{The\ number\ of\ all\ possible\ sample}\\ &= 1-\frac{\binom{q-1}{p}}{\binom{q}{p}}\end{aligned}$$

We can see that such units have the same inclusion probabilities.

According to the three cases of inclusion probabilities due to overlapping, the inclusion probability of a unit (*i, j*) can be written, in generic formula, as

$$\pi_{(i,j)} = \begin{cases} 1-\dfrac{\binom{i-2}{p}}{\binom{q}{p}} & i = 1,2,3,...,r \ \ and \ \ j = 1,2 \\ 1-\dfrac{\binom{q-2}{p}}{\binom{q}{p}} & i = 2,3,...,r-1 \ \ and \ \ j = 3,4,5,...,c \\ 1-\dfrac{\binom{q-1}{p}}{\binom{q}{p}} & i = 1,\ r \ and \ j = 3,4,5,...,c. \end{cases}$$

which is equation (3.2) when j^* =1. Next, the inclusion probabilities are calculated when starting unit (1, j^*) = (1, 1) in a small population. The calculation of the inclusion probability of each unit in the spatial population of 8 rows and 4 columns, as shown in Figure 3.3, will be shown. The number of all possible paths is $q = r - 1 = 8 - 1 = 7$. Suppose that the number of sample paths is two; that is, $p = 2$.

First, the inclusion probabilities for the units in column 1 and 2 (case 1) will be calculated.

From $\pi_{(i,j)} = 1 - \dfrac{\binom{i-2}{p}}{\binom{q}{p}}$, we get

$$\pi_{(1,1)} = 1 - \frac{\binom{1-2}{2}}{\binom{7}{2}} = 1 - 0 = 1 = \pi_{(1,2)} \qquad \pi_{(2,1)} = 1 - \frac{\binom{2-2}{2}}{\binom{7}{2}} = 1 - 0 = 1 = \pi_{(2,2)}$$

$$\pi_{(3,1)} = 1 - \frac{\binom{3-2}{2}}{\binom{7}{2}} = 1 - 0 = 1 = \pi_{(3,2)} \qquad \pi_{(4,1)} = 1 - \frac{\binom{4-2}{2}}{\binom{7}{2}} = 1 - \frac{1}{21} = \frac{20}{21} = \pi_{(4,2)}$$

$$\pi_{(5,1)} = 1 - \frac{\binom{5-2}{2}}{\binom{7}{2}} = 1 - \frac{3}{21} = \frac{18}{21} = \pi_{(5,2)} \qquad \pi_{(6,1)} = 1 - \frac{\binom{6-2}{2}}{\binom{7}{2}} = 1 - \frac{6}{21} = \frac{15}{21} = \pi_{(6,2)}$$

$$\pi_{(7,1)} = 1 - \frac{\binom{7-2}{2}}{\binom{7}{2}} = 1 - \frac{10}{21} = \frac{11}{21} = \pi_{(7,2)} \qquad \pi_{(8,1)} = 1 - \frac{\binom{8-2}{2}}{\binom{7}{2}} = 1 - \frac{15}{21} = \frac{6}{21} = \pi_{(8,2)}$$

Next, we will calculate the inclusion probabilities for the units not in column 1 or 2 and not in the first or last row (case 2).

From $\pi_{(i,j)} = 1 - \dfrac{\binom{q-2}{p}}{\binom{q}{p}}$, we get

$$\pi_{(2,3)} = 1 - \frac{\binom{7-2}{2}}{\binom{7}{2}} = 1 - \frac{10}{21} = \frac{11}{21} = \pi_{(2,4)} = \pi_{(3,3)} = \pi_{(3,4)} = \pi_{(4,3)} = \pi_{(3,4)}$$

$$= \pi_{(5,3)} = \pi_{(5,4)} = \pi_{(6,3)} = \pi_{(6,4)} = \pi_{(7,3)} = \pi_{(7,4)}$$

Finally, the inclusion probabilities for the units in the first and last row will be calculated, but not in column 1 or 2 (case 3).

From $\pi_{(i,j)} = 1 - \frac{\binom{q-1}{p}}{\binom{q}{p}}$, we get

$$\pi_{(1,3)} = 1 - \frac{\binom{7-1}{2}}{\binom{7}{2}} = 1 - \frac{15}{21} = \frac{6}{21} = \pi_{(1,4)} = \pi_{(8,3)} = \pi_{(8,4)}$$

All inclusion probabilities are shown Figure 3.6.

3.2.1.5 Example: Calculating the Inclusion Probabilities When the Starting unit $(1, j^*) = (1, 3)$ in a Small Population

The calculation of the inclusion probability of each unit in a spatial population of 8 rows and 6 columns, as shown in Figure 3.4, will be shown. The number of all possible paths is $q = r - 1 = 8 - 1 = 7$. Suppose that the number of sample paths is 2; that is, $p = 2$ and the starting point is unit (1, 3), or $j^* = 3$.

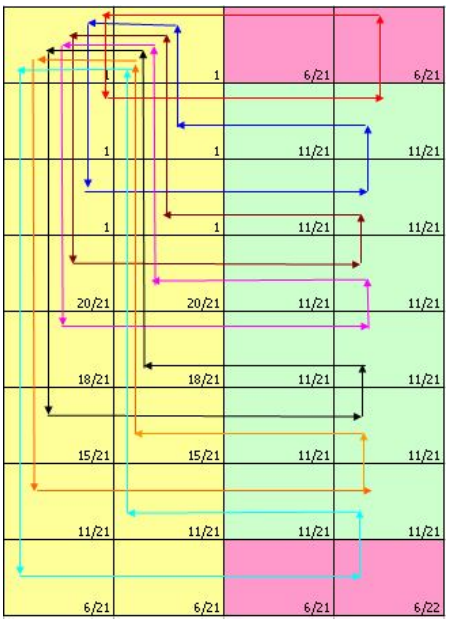


Figure 3.6 Inclusion Probabilities for the Population for 8 Rows and 4 Columns

Note: The unit in yellow is unit of type 1, green is unit of type 2, and pink is unit of type 3.

Using equation (3.2), first, the inclusion probabilities for the units in column 3 and 4 (case 1) will be calculated. For $i = 1, 2, 3, \ldots, 8$ and $j = 3$and 4, we have

$$\pi_{(i,j)} = 1 - \frac{\binom{i-2}{p}}{\binom{q}{p}}$$

Then, we get

$$\pi_{(1,3)} = 1 - \frac{\binom{1-2}{2}}{\binom{7}{2}} = 1 - 0 = 1 = \pi_{(1,4)} \qquad \pi_{(2,3)} = 1 - \frac{\binom{2-2}{2}}{\binom{7}{2}} = 1 - 0 = 1 = \pi_{(2,4)}$$

$$\pi_{(3,3)} = 1 - \frac{\binom{3-2}{2}}{\binom{7}{2}} = 1 - 0 = 1 = \pi_{(3,4)} \qquad \pi_{(4,3)} = 1 - \frac{\binom{4-2}{2}}{\binom{7}{2}} = 1 - \frac{1}{21} = \frac{20}{21} = \pi_{(4,4)}$$

$$\pi_{(5,3)} = 1 - \frac{\binom{5-2}{2}}{\binom{7}{2}} = 1 - \frac{3}{21} = \frac{18}{21} = \pi_{(5,4)} \qquad \pi_{(6,3)} = 1 - \frac{\binom{6-2}{2}}{\binom{7}{2}} = 1 - \frac{6}{21} = \frac{15}{21} = \pi_{(6,4)}$$

$$\pi_{(7,3)} = 1 - \frac{\binom{7-2}{2}}{\binom{7}{2}} = 1 - \frac{10}{21} = \frac{11}{21} = \pi_{(7,4)} \qquad \pi_{(8,3)} = 1 - \frac{\binom{8-2}{2}}{\binom{7}{2}} = 1 - \frac{15}{21} = \frac{6}{21} = \pi_{(8,4)}$$

Next, the inclusion probabilities for the units not in column 3 and 4 will be calculated, and not in the first or last row (case 2).

For $i = 2, 3, \ldots, 7$ and $j = 1, 2, 5, 6$

$$\pi_{(i,j)} = 1 - \frac{\binom{q-2}{p}}{\binom{q}{p}}$$

Then, we get

$$\pi_{(2,1)} = 1 - \frac{\binom{7-2}{2}}{\binom{7}{2}} = 1 - \frac{10}{21} = \frac{11}{21} = \pi_{(2,2)} = \pi_{(2,5)} = \pi_{(2,6)} = \pi_{(3,1)} = \pi_{(3,2)} = \pi_{(3,5)}$$

$$= \pi_{(3,6)} = \pi_{(4,1)} = \pi_{(4,2)} = \pi_{(4,5)} = \pi_{(4,6)} = \pi_{(5,1)} = \pi_{(5,2)} = \pi_{(5,5)} = \pi_{(5,6)}$$

$$= \pi_{(6,1)} = \pi_{(6,2)} = \pi_{(6,5)} = \pi_{(6,6)} = \pi_{(7,1)} = \pi_{(7,2)} = \pi_{(7,5)} = \pi_{(7,6)}$$

Finally, the inclusion probabilities for the units in the first row and last row will be calculated, but not in column 3 or 4 (case 3).

For $i = 1$ and 8 and $j = 1, 2, 5,6$

$$\pi_{(i,j)} = 1 - \frac{\binom{q-1}{p}}{\binom{q}{p}}$$

Then, we get

$$\pi_{(1,1)} = 1 - \frac{\binom{7-1}{2}}{\binom{7}{2}} = 1 - \frac{15}{21} = \frac{6}{21} = \pi_{(1,2)} = \pi_{(1,5)} = \pi_{(1,6)} = \pi_{(8,1)} = \pi_{(8,2)} = \pi_{(8,5)} = \pi_{(8,6)}$$

All inclusion probabilities are shown Figure 3.7.

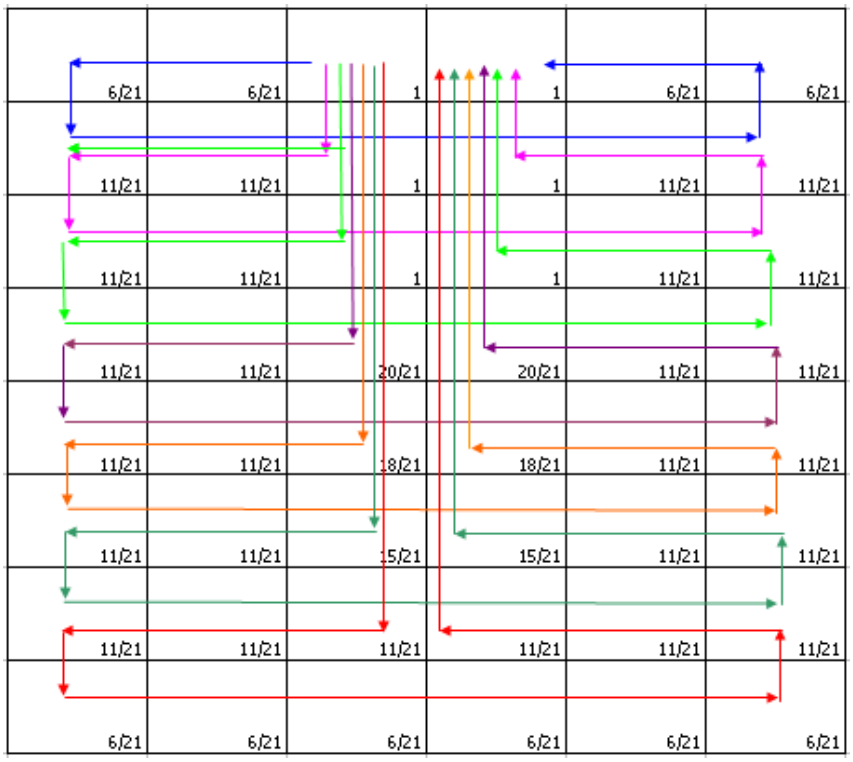


Figure 3.7 Inclusion Probabilities of the Population of 8 Rows and 6 Columns

3.2.2 Joint Inclusion Probability

Let the probability that both units (*i, j*) and (i', j') are included in the sample be denoted by $\pi_{(i,j),(i',j')}$, also called the joint inclusion probability. A formula for calculating under path sampling is

$$\pi_{(i,j),(i',j')} = \pi_{(i,j)} + \pi_{(i',j')} - (1 - \frac{\binom{f}{p}}{\binom{q}{p}}), \tag{3.6}$$

where f is the number of paths not containing either units *(i, j)* or (i', j'), and f is divided into 6 cases.

Case1: For Units of Type 1

For i, i' = 1, 2, 3, …, r and j, $j' = j^*$ and j^*+1

$$f = \min(i, i') - 2 \tag{3.7}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 2: For Units of Type 1 and 2

For i = 1, 2, 3, …, r and $j = j^*$ and j^*+1

i' = 2, 3, …, r-1 and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c

$$f = \begin{cases} i-2 & if\ i \le i' \\ i-3 & if\ i - i' = 1 \\ i-4 & if\ i - i' \ge 2 \end{cases} \tag{3.8}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 3: For Units of Type 1 and 3

For i = 1, 2, 3, …, r and $j = j^*$ and j^*+1

i' = 1 and r and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c

$$f = \begin{cases} i-2 & i \le i' \\ i-3 & i > i' \end{cases} \tag{3.9}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 4: For Units of Type 2

For $i, i' = 2, 3, \ldots, r\text{-}1$ and $j,\ j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots,$ c

$$f = \begin{cases} q-2 & if\ |i-i'| = 0 \\ q-3 & if\ |i-i'| = 1 \\ q-4 & if\ |i-i'| \ge 2 \end{cases} \tag{3.10}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 5: For Units of Type 2 and 3

For $i = 2, 3, \ldots, r\text{-}1$ and $j = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots,$ c

$i' = 1$ and r and $j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots, c$

$$f = \begin{cases} q-2 & |i-i'| = 1 \\ q-3 & |i-i'| \ge 2 \end{cases} \tag{3.11}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 6: For Units of Type 3

For $i, i' = 1$ and r and $j, j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots, c$

$$f = \begin{cases} q-1 & i = i' \\ q-2 & i \ne i' \end{cases} \tag{3.12}$$

Proof of the joint inclusion probability is shown in the next section.

3.2.2.1 Proof of Joint Inclusion Probability

Joint inclusion probability is defined as

$$\pi_{(i,j),(i',j')} = a\ probability\ that\ both\ unit\ (i,j)\ and\ (i',j')\ are\ included\ in\ the\ sample$$
$$= \frac{the\ number\ of\ samples\ containing\ both\ unit\ (i,j)\ and\ (i',j')}{the\ number\ of\ all\ possible\ samples}$$

Let A be an event that unit (*i, j*) is included in a sample. Thus P(A)= $\pi_{(i,j)}$. Let B be an event that unit (i', j') is included in a sample. Thus P(B)= $\pi_{(i',j')}$. $A \cap B$ is an event that both unit *(i, j)* and (i', j') are included in the sample. Thus, $P(A \cap B) = \pi_{(i,j),(i',j')}$. Thus,

$$\pi_{(i,j),(i',j')} = P(A \cap B)$$
$$= P(A)+P(B)-P(A \cup B)$$
$$= \pi_{(i,j)} + \pi_{(i',j')} - P(A \cup B) \quad \text{since } P(A)= \pi_{(i,j)} \text{ and } P(B)= \pi_{(i',j')}$$
$$= \pi_{(i,j)} + \pi_{(i',j')} - (1-P(A \cup B)^c)$$

Since $(A \cup B)^c = (A^c \cap B^c)$ = an event that unit *(i, j)* is not included in a sample and unit (i', j') is not included in the sample. Thus,

$$\pi_{(i,j),(i',j')} = \pi_{(i,j)} + \pi_{(i',j')} - (1-P(A^c \cap B^c)),$$

where

$P(A^c \cap B^c)$= *P(an event that unit (i, j) and* (i', j') *are not included in the same sample)*

$$= \frac{The\ number\ of\ sample\ not\ containing\ either\ units\ (i,j)\ \text{or}\ (i',j')}{The\ number\ of\ all\ possible\ sample}$$

$$= \frac{\binom{f}{p}}{\binom{q}{p}},$$

where f = the number of paths not containing either units *(i, j)* or (i', j'). Hence,

$$\pi_{(i,j),(i',j')} = \pi_{(i,j)} + \pi_{(i',j')} - (1 - \frac{\binom{f}{p}}{\binom{q}{p}})$$

Next is derivation of f for 6 cases.

Case 1: For Units of Type 1

For i, i' = 1, 2, 3, …, r and j, $j' = j^*$ and j^*+1, the number of paths not containing unit (i, j) is i-2, and the number of paths not containing unit (i', j') is i'-2.

Let C1 be a set of paths containing unit (i, j). Thus, C1 = $\{P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}\}$.

Let D1 be a set of paths containing unit (i', j'). Thus, D1 = $\{P_{i'-1}, P_{i'}, P_{i'+1}, \ldots, P_{r-1}\}$.

When $i < i'$, D1 is a subset of C1; that is D1 $\subset$ C1. A set of paths containing either unit (i, j) or (i', j') is C1$\cup$D1=C1. Thus, the number of paths not containing unit (i, j) and (i', j') is i-2.

When $i' < i$, C1 is a subset of D1; that is C1 $\subset$ D1. A set of paths containing either unit (i, j) or (i', j') is C1$\cup$D1=D1. Thus, the number of paths not containing unit *(i, j)* and (i', j') is $i' - 2$.

In conclusion, the number of paths not containing units *(i, j)* or (i', j') is $\min(i, i') - 2$

That is,

$$f = \min(i, i') - 2$$

Case 2: For Units of Type 1 and 2

For i = 1, 2, 3, …, r and $j = j^*$ and j^*+1, the number of paths not containing unit (i, j) is i-2. For i' = 2, 3, …, r-1 and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c**,** unit (i', j') belongs to $P_{i'-1}$ and $P_{i'}$**.** Let C2 be a set of paths containing unit (i, j).

Thus, C2 = $\{P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}\}$. Let D2 be a set of paths containing unit (i', j'). Thus, D2 = $\{P_{i'-1}, P_{i'}\}$.

When $i \le i'$, D2 is a subset of C2; that is D2 $\subset$ C2. A set of paths containing either unit (*i, j*) or (i', j') is C2 $\cup$D2=C2. Thus, the number of paths not containing unit *(i, j)* or (i', j') is *i*-2. Thus, f = *i*-2 when $i \le i'$.

When $i - i' = 1$, this means that unit (*i, j*) is in row *i* and unit (i', j') is in row $i' = i - 1$. Then, unit (*i, j*) belongs to $P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}$. Unit (i', j') belongs to P_{i-2} and P_{i-1}. The paths containing either unit (*i, j*) or (i', j') is $P_{i-2}, P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}$.

So, the paths not containing unit (*i, j*) or (i', j') is $P_1, P_2, P_3, \ldots, P_{i-3}$. Thus, the number of paths not containing unit *(i, j)* or (i', j') is *i*-3. Hence, f = *i*-3 when $i - i' = 1$.

When $i - i' \ge 2$, D2 is not a subset of C2, and D2$\cap$ C2=$\varnothing$. A set of paths containing either unit (*i, j*) or (i', j') is C2 $\cup$ D2 = $P_{i'-1}, P_{i'}, P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}$. Thus, the number of paths not containing unit *(i, j)* or (i', j') is (*i*-2)-2= *i*-4. Thus, f = *i*-4 when $i - i' \ge 2$. Thus,

$$f = \begin{cases} i-2 & \text{if } i \le i' \\ i-3 & \text{if } i - i' = 1 \\ i-4 & \text{if } i - i' \ge 2 \end{cases}$$

Case 3: For Units of Type 1 and 3

For *i* = 1, 2, 3, …, *r* and *j*= $j^*$and j^*+1, unit (*i, j*) belongs to $P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}$. For i' = 1 and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, *c*, unit (i', j') belongs to only one path, which is P_1. For i' = r and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, *c*, unit (i', j') belongs to only one path, which is P_{r-1}. Let C3 be a set of paths containing unit (*i, j*). Thus, C3 = $\{P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}\}$. Let D3 be a set of paths containing unit (i', j'). Thus, D3 = $\{P_1\}$ if i' = 1 . D3 = $\{P_{r-1}\}$ if i' = r.

For i' = 1 and i = 1, 2 we have C3 = $\{P_1, P_2, P_3, \ldots, P_{r-1}\}$ and D3 = $\{P_1\}$. We can see that D3 $\subset$ C3. A set of paths containing either unit (i, j) or (i', j') is C3 $\cup$ D3 = C3 = a set of all possible paths. Thus, the number of paths not containing unit *(i, j)* and (i', j') is zero.

For i' = r and i = 1,2,3,…r we have C3 = $\{P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}\}$ and D3 = $\{P_{r-1}\}$. We can see that D3 $\subset$ C3. A set of paths containing either unit (i, j) or (i', j') is C3 $\cup$ D3 =C3. Thus, the number of paths not containing unit *(i, j)* or (i', j') is i-2.

For i' = 1 and i = 3, 4, 5…r we have C3 = $\{P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}\}$ and D3 = $\{P_1\}$. We can see that D3 is not a subset of C3 and that C3 $\cap$ D3 =$\varnothing$. A set of paths containing either unit (i, j) or (i', j') is C3 $\cup$ D3 =$\{P_1, P_{i-1}, P_i, P_{i+1}, \ldots, P_{r-1}\}$. Thus, the number of paths not containing unit *(i, j)* or (i', j') is (i-2)-1= (i-3). Hence,

$$f = \begin{cases} 0 & \text{if } i' = 1 \text{ and } i = 1,2 \\ i-2 & \text{if } i' = r \text{ and } i = 1,23,..r \\ i-3 & \text{if } i' = 1 \text{ and } i = 3,4,5..r \end{cases}$$

Case 4: For Units of Type 2

For i, i' = 2, 3, …, r-1 and j, j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c**,** unit *(i, j)* belongs to P_{i-1} and P_i. Unit (i', j') belongs to $P_{i'-1}$ and $P_{i'}$. For two units (i, j) and (i', j') in the same row, that is $i = i'$, they are in the same paths P_{i-1} and P_i. Consequently, the number of paths containing either unit (i, j) or (i', j') is 2. Thus, the number of paths not containing unit (i, j) and (i', j') is q-2. Hence, for $i = i'$, or $|i - i'|$=0, we have f=q-2.

For two units *(i, j)* and (i', j') in different but consecutive rows, $|i - i'|$=1. When $i - i'$ =1, we have i=i' +1. Unit (i, j) is in path $P_{i'}$ and $P_{i'+1}$. Unit (i', j') is in path $P_{i'-1}$ and $P_{i'}$. A set of paths containing either unit (i, j) or (i', j')

is $\{P_{i'-1}, P_{i'}, P_{i'+1}\}$. So, the number of paths either containing unit (*i, j*) or (i', j') is 3. Thus, the number of paths not containing unit (*i, j*) or (i', j') is *q*-3.

When $i - i' = -1$, we have $i = i' - 1$. Unit (*i, j*) is in path $P_{i'-2}$ and $P_{i'-1}$. Unit (i', j') is in path $P_{i'-1}$ and $P_{i'}$. A set of paths containing either unit (*i, j*) or (i', j') is $\{P_{i'-2}, P_{i'-1}, P_{i'}\}$. So, the number of paths either containing unit (*i, j*) or (i', j') is 3. Thus, the number of paths not containing unit (*i, j*) or (i', j') is *q*-3. Hence, for $|i - i'| = 1$, we have $f = q-3$.

For the two units (*i, j*) and (i', j') in different and non-consecutive rows, $|i - i'| \geq 2$. Unit (*i, j*) is in path P_i and P_{i+1}. Unit (i', j') is in path $P_{i'-1}$ and $P_{i'}$. A set of paths containing either unit (*i, j*) or (i', j') is $\{P_{i'-1}, P_{i'}, P_{i-1}, P_i\}$. So, the number of paths either containing unit (*i, j*) or (i', j') is 4. Thus, the number of paths not containing unit (*i, j*) or (i', j') is *q*-4. Hence, for $|i - i'| \geq 2$, we have $f = q-4$. Hence,

$$f = \begin{cases} q-2 & if\ |i-i'| = 0 \\ q-3 & if\ |i-i'| = 1 \\ q-4 & if\ |i-i'| \geq 2 \end{cases}$$

Case 5: For Units of Type 2 and 3

For i = 2, 3, …, *r*-1 and j= 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c, unit *(i, j)* belongs to P_{i-1} and P_i. For $i' = 1$ and j'= 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, *c*, unit (i', j') belongs to only one path, which is P_1. For i' = r and j'= 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, *c*, a unit (i', j') belongs to only one path, which is P_{r-1}.

1) For the two units *(i, j)* and (i', j') in different row but consecutive rows, $|i - i'| = 1$.

For $i' = 1$ and i=2, that is $i - i' = 1$, we have unit $(i', j') = (1, j')$ in path P_1 and unit *(i, j)* =(2,*j*) in path P_1, P_2. A set of paths containing either unit (*i, j*) or (i', j') is $\{P_1, P_2\}$. So, the number of paths either containing unit (*i, j*) or (i', j') is 2. Thus, the number of paths not containing unit (*i, j*) or (i', j') is *q*-2.

For i' = r and i= r-1, that is $i - i'$ = -1, we have unit *(i, j)* =(r-*1*,j) in path $P_{r-2,}P_{r-1}$ and unit $(i', j') = (r, j')$ in path P_{r-1}. A set of paths containing either unit (*i, j*) or (i', j') is$\{P_{r-2,}P_{r-1}\}$. So, the number of paths either containing unit (*i, j*) or (i', j') is 2. Thus, the number of paths not containing unit (*i, j*) or (i', j') is q-2.

2) For two units (*i, j*) and (i', j') in different and non-consecutive rows, $|i - i'| \geq 2$.

For i' = 1 and i = 3, 4,…r, that is $i - i' \geq 2$**,** we have unit $(i', j') = (1, j')$ in path $\{P_1\}$=C5 and unit *(i, j)* in path $\{P_{i-1}, P_i\}$=D5. C5 is not a subset of D5 or C5 $\cap$ D5 =$\varnothing$. A set of paths containing either unit (*i, j*) or (i', j') is C5 $\cup$ D5 =$\{P_1, P_{i-1}, P_i\}$. So, the number of paths either containing unit (*i, j*) or (i', j') is 3. Thus, the number of paths not containing unit (*i, j*) and (i', j') is q-3.

For i' = r and i=1, 2, 3,…r-2, that is $i - i' \leq$ -2**,** we have unit $(i', j') = (r, j')$ in path $\{P_{r-1}\}$=C5 and unit *(i, j)* in path $\{P_{i-1}, P_i\}$=D5. C5 is not a subset of D5 and C5 $\cap$ D5 =$\varnothing$. A set of paths containing either unit (*i, j*) or (i', j') is C5 $\cup$ D5 =$\{P_{i-1}, P_i, P_{r-1}\}$ So, the number of paths either containing unit (*i, j*) or (i', j') is 3. Thus, the number of paths not containing unit (*i, j*) or (i', j') is q-3.

Hence,

$$f = \begin{cases} q-2 & |i - i'| = 1 \\ q-3 & |i - i'| \geq 2 \end{cases}$$

Case 6: For Units of Type 3

1) For the two units (*i, j*) and (i', j') in the same row, $i = i'$.

In row 1, *(i, j)*= *(*1, *j)* is in path P_1 and $(i', j') = (1, j')$ is also in path P_1. Both units *(i, j)* and (i', j') are in P_1, only one path. Thus, the number of paths not containing unit (*i, j*) or (i', j') is q-1. In row *r*, *(i, j)= (r, j)* in path P_{r-1} and $(i', j') = (r\ j')$ is also in path P_{r-1}. Both unit s*(i, j)* and (i', j') are in P_{r-1}, only one path.

Thus, the number of paths not containing unit (i, j) or (i', j') is q-1. Hence, f=q-1 when $i = i'$.

2) For the two units (i, j) and (i', j') in the different rows, $i \neq i'$.

For i=1, a unit (i, j)= $(1, j)$ is in path P_1. For i'=r and $(i', j') = (r\ j')$ they are also in path P_{r-1}. The paths containing either units (i, j) or (i', j') are P_1, P_{r-1}. Thus, the number of paths not containing unit (i, j) or (i', j') is q-2. On the other hand, for i=r, unit (i, j)= (r, j) is in path P_{r-1}. For i'=1, $(i', j') = (1\ j')$ is also in path P_1. The paths containing either unit (i, j) or (i', j') are P_1, P_{r-1}. Thus, the number of paths not containing unit (i, j) or (i', j') is q-2. Hence, f=q-2 when $i \neq i'$. Hence,

$$f = \begin{cases} q-1 & i = i' \\ q-2 & i \neq i' \end{cases}$$

3.3 Estimation of the Population Mean

Horvitz-Thompson (1952: 663-685) have proposed that with any design, with or without replacement, giving probability π_k that unit k is included in the sample, the unbiased estimator of the population total is

$$\hat{\tau}_{HT} = \sum_{k=1}^{\upsilon} \frac{y_k}{\pi_k}, \tag{3.13}$$

where υ is the number of distinct units in the sample.
The variance of the estimator is

$$v(\hat{\tau}_{HT}) = \sum_{k=1}^{N} \left(\frac{1-\pi_k}{\pi_k} \right) y_k^2 + \sum_{k=1}^{N} \sum_{k' \neq k} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'}} \right) y_k y_{k'} \tag{3.14}$$

The unbiased estimator of this variance is

$$\hat{v}(\hat{\tau}_{HT}) = \sum_{k=1}^{\upsilon}\left(\frac{1}{\pi_k^2} - \frac{1}{\pi_k}\right) y_k^2 + \sum_{k=1}^{\upsilon}\sum_{k'\neq k}\left(\frac{1}{\pi_k \pi_{k'}} - \frac{1}{\pi_{kk'}}\right) y_k y_{k'} \tag{3.15}$$

if all of the joint inclusion probabilities are greater than zero. Note that if there are zero joint inclusion probabilities, this estimator of variance may be not unbiased. This variance estimate may be negative in some designs.

The unbiased estimator of the population mean is

$$\hat{\mu}_{HT} = \frac{1}{N}\hat{\tau}_{HT} = \frac{1}{N}\sum_{k=1}^{\upsilon}\frac{y_k}{\pi_k} \tag{3.16}$$

having variance

$$v(\hat{\mu}_{HT}) = \frac{1}{N^2}\operatorname{var}(\hat{\tau}_{HT}) = \frac{1}{N^2}\left(\sum_{k=1}^{N}\left(\frac{1-\pi_k}{\pi_k}\right) y_k^2 + \sum_{k=1}^{N}\sum_{k'\neq k}\left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'}}\right) y_k y_{k'}\right) \tag{3.17}$$

and estimated variance proposed by Horvitz and Thompson (1952)

$$\hat{v}(\hat{\mu}_{HT}) = \frac{1}{N^2}\hat{\operatorname{var}}(\hat{\tau}_{HT}) = \frac{1}{N^2}\left(\sum_{k=1}^{\upsilon}\left(\frac{1}{\pi_k^2} - \frac{1}{\pi_k}\right) y_k^2 + \sum_{k=1}^{\upsilon}\sum_{k'\neq k}\left(\frac{1}{\pi_k \pi_{k'}} - \frac{1}{\pi_{kk'}}\right) y_k y_{k'}\right). \tag{3.18}$$

Now, we want to find the estimator of the mean for path sampling. Let $p_s = (p_1, p_2, p_3, ..., p_p)$ denote the sample of paths selected. Let s denote the set of distinct units in the sample. By using the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), the unbiased estimator of the population mean under path sampling is

$$\hat{\mu}_{ps} = \frac{1}{rc}\sum_{(i,j)\in s}\frac{y_{(i,j)}}{\pi_{(i,j)}}. \tag{3.19}$$

Let $I_{(i,j)}$ be the indicator function taking the value one if unit (i, j) is selected in the sample and 0 otherwise. It can be written as

$$I_{(i,j)} = \begin{cases} 1 & \textit{if unit}\,(i,j)\,\textit{is included in the sample} \\ 0 & \textit{otherwise} \end{cases}$$

Therefore, $\hat{\mu}_{ps}$ can be written in the alternative form

$$\hat{\mu}_{ps} = \frac{1}{rc} \sum_{all\,(i,j)} \frac{y_{(i,j)} I_{(i,j)}}{\pi_{(i,j)}}. \tag{3.20}$$

$\hat{\mu}_{ps}$ is the unbiased estimator for the population mean μ.

Proof.

$$E\left(\hat{\mu}_{ps}\right) = E\left(\frac{1}{rc} \sum_{all\,(i,j)} \frac{y_{(i,j)} I_{(i,j)}}{\pi_{(i,j)}} \right)$$

Each of the $I_{(i,j)}$ is a (Bernoulli) random variable, with expected value

$$E(I_{(i,j)}) = P(I_{(i,j)} = 1) = \pi_{(i,j)}.$$

Hence, the expected value of $\hat{\mu}_{ps}$ is

$$\begin{aligned} E\left(\hat{\mu}_{ps}\right) &= \frac{1}{rc} \sum_{all\,(i,j)} \frac{y_{(i,j)} E\left(I_{(i,j)}\right)}{\pi_{(i,j)}} \\ &= \frac{1}{rc} \sum_{all\,(i,j)} \frac{y_{(i,j)} \pi_{(i,j)}}{\pi_{(i,j)}} \\ &= \frac{1}{rc} \sum_{all\,(i,j)} y_{(i,j)} \\ &= \frac{1}{rc} \sum_{i=1}^{r} \sum_{j=1}^{c} y_{(i,j)} \\ &= \mu \end{aligned}$$

Another way to prove that $\hat{\mu}_{ps}$ is an unbiased estimator for μ is the following.

Proof. By definition

$$E(\hat{\mu}_{ps}) = \sum_{all\ s} \hat{\mu}_{ps} P(s)$$

$$= \sum_{all\ s} \hat{\mu}_{ps} \frac{1}{\binom{q}{p}} \qquad \text{since } P(s) = \frac{1}{\binom{q}{p}}.$$

$$= \sum_{all\ s} \left(\frac{1}{rc} \sum_{(i,j)\in s} \frac{y_{(i,j)}}{\pi_{(i,j)}} \right) \frac{1}{\binom{q}{p}} \quad \text{since } \hat{\mu}_{ps} = \frac{1}{rc} \sum_{(i,j)\in s} \frac{y_{(i,j)}}{\pi_{(i,j)}}$$

$$= \frac{1}{rc} \frac{1}{\binom{q}{p}} \sum_{all\ s} \left(\sum_{(i,j)\in s} \frac{y_{(i,j)}}{\pi_{(i,j)}} \right).$$

The sum extends over all $\binom{q}{p}$ samples. To evaluate this sum, we find out in how many samples any specific value $y_{(i,j)}$ appears. Now, let $a_{(i,j)}$ be the number of samples containing unit (i, j). Then, we have

$$E(\hat{\mu}_{ps}) = \frac{1}{rc} \frac{1}{\binom{q}{p}} \sum_{i=1}^{r} \sum_{j=1}^{c} a_{(i,j)} \frac{y_{(i,j)}}{\pi_{(i,j)}}.$$

From the inclusion probabilities $\pi_{(i,j)}$ for the 3 cases in equation (3.2), we have

$$E(\hat{\mu}_{ps}) = \frac{1}{rc} \frac{1}{\binom{q}{p}} \left(\sum_{i=1}^{r} \sum_{j=j^*}^{j^*+1} a_{(i,j)} \frac{y_{(i,j)}}{1 - \frac{\binom{i-2}{p}}{\binom{q}{p}}} + \sum_{i=2}^{r-1} \sum_{\substack{j=1 \\ j\neq j^*, j^*+1}}^{c} a_{(i,j)} \frac{y_{(i,j)}}{1 - \frac{\binom{q-2}{p}}{\binom{q}{p}}} \right.$$

$$+\sum_{i=1,r}\sum_{\substack{j=1\\ j\neq j^*,j^*+1}}^{c} a_{(i,j)}\frac{y_{(i,j)}}{1-\dfrac{\binom{q-1}{p}}{\binom{q}{p}}}\Bigg)$$

Now we want to find $a_{(i,j)}$. The number of all possible samples is $\binom{q}{p}$. For any unit (*i, j*) in case 1, the number of samples not containing unit (*i, j*) is $\binom{i-2}{p}$; thus the number of samples containing unit (*i, j*) is $\binom{q}{p}$-$\binom{i-2}{p}$. It can be written as $a_{(i,j)} = \binom{q}{p}$-$\binom{i-2}{p}$. For any unit (*i, j*) in case 2, $a_{(i,j)} = \binom{q}{p}$-$\binom{q-2}{p}$. For any unit (*i, j*) in case 3, $a_{(i,j)} = \binom{q}{p}$-$\binom{q-1}{p}$. Then we have

$$E(\hat{\mu}_{ps}) = \frac{1}{rc}\frac{1}{\binom{q}{p}}\Bigg(\sum_{i=1}^{r}\sum_{j=j^*}^{j^*+1}\left[\binom{q}{p}-\binom{i-2}{p}\right]\frac{y_{(i,j)}}{1-\dfrac{\binom{i-2}{p}}{\binom{q}{p}}} +$$

$$\sum_{i=2}^{r-1}\sum_{\substack{j=1\\ j\neq j^*,j^*+1}}^{c}\left[\binom{q}{p}-\binom{q-2}{p}\right]\frac{y_{(i,j)}}{1-\dfrac{\binom{q-2}{p}}{\binom{q}{p}}} + \sum_{i=1,r}\sum_{\substack{j=1\\ j\neq j^*,j^*+1}}^{c}\left[\binom{q}{p}-\binom{q-1}{p}\right]\frac{y_{(i,j)}}{1-\dfrac{\binom{q-1}{p}}{\binom{q}{p}}}\Bigg)$$

After each term is canceled out, then we have

$$E(\hat{\mu}_{ps}) = \frac{1}{rc}\frac{1}{\binom{q}{p}}\left(\sum_{i=1}^{r}\sum_{j=j^*}^{j^*+1}\binom{q}{p} y_{(i,j)} \;+\; \sum_{i=2}^{r-1}\sum_{\substack{j=1 \\ j\neq j^*, j^*+1}}^{c}\binom{q}{p} y_{(i,j)} \;+\; \sum_{i=1,r}\sum_{\substack{j=1 \\ j\neq j^*, j^*+1}}^{c}\binom{q}{p} y_{(i,j)} \right)$$

$$= \frac{1}{rc}\left(\sum_{i=1}^{r}\sum_{j=j^*}^{j^*+1} y_{(i,j)} \;+\; \sum_{i=2}^{r-1}\sum_{\substack{j=1 \\ j\neq j^*, j^*+1}}^{c} y_{(i,j)} \;+\; \sum_{i=1,r}\sum_{\substack{j=1 \\ j\neq j^*, j^*+1}}^{c} y_{(i,j)} \right)$$

$$= \frac{1}{rc}\left(\sum_{i=1}^{r}\sum_{j=1}^{c} y_{(i,j)} \right)$$

$$= \mu$$

The variance of $\hat{\mu}_{ps}$, applied equation (3.17), is

$$v(\hat{\mu}_{ps}) = \frac{1}{(rc)^2}\left(\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\frac{1-\pi_{(i,j)}}{\pi_{(i,j)}} \right) y_{(i,j)}^2 + \sum_{i=1}^{r}\sum_{i'\neq i}\sum_{j=1}^{c}\sum_{j'\neq j}\left(\frac{\pi_{(i,j),\,(i',j')} - \pi_{(i,j)}\pi_{(i',j')}}{\pi_{(i,j)}\pi_{(i',j')}} \right) y_{(i,j)} y_{(i',j')} \right)$$

(3.21)

and the estimated variance, applied equation (3.18), is

$$\hat{v}(\hat{\mu}_{ps}) = \frac{1}{(rc)^2}\left(\sum_{(i,j)\in s}\left(\frac{1}{\pi_{(i,j)}^2} - \frac{1}{\pi_{(i,j)}} \right) y_{(i,j)}^2 + \sum_{\substack{(i,j),\,(i',j')\in s \\ (i,j)\neq(i',j')}}\left(\frac{1}{\pi_{(i,j)}\pi_{(i',j')}} - \frac{1}{\pi_{(i,j),\,(i',j')}} \right) y_{(i,j)} y_{(i',j')} \right)$$

(3.22)

It is unbiased if all of the joint inclusion probabilities are greater than zero. This variance may be negative. A researcher may use an alternative variance estimator, such as that proposed by Sen (1953) and Yate and Grundy (1953); see equation (2.4). It is claimed that it is less often negative. Suppose X is an auxiliary variable. Based on the population coefficient of the variation of ratios Y/X, denoted as C. V. (Y/X), the

variance estimator proposed by Sen, Yate and Grundy is better when C. V. (Y/X) is very small. On the other hand, the variance estimator of Horvitz-Thompson is better when C. V. (Y/X) is larger (Stephan and Overton, 1987).

3.4 An Illustrative Example

Now the population of 4 rows and 6 columns will be considered, as shown in Figure 3.8. The population mean and variance are 8.208 and 549.6, respectively. The objective is to estimate the population mean by using path sampling. First, all possible paths are created. The number of rows in this population is r = 4, and the number of columns is c = 6. Thus, the number of all possible paths is $q = r\text{-}1=4\text{-}1=3$. In general, a path k in the spatial setting population of r rows and c columns is written as

$$P_k = ((1, j^*), (2, j^*), (3, j^*), \ldots, (k, j^*), (k, j^* - 1), (k, j^* - 2), \ldots, (k,1), (k+1, 1), (k+1, 2), \ldots, (k+1, c), (k, c), (k, c-1), (k, c-2), \ldots, (k, j^* + 1), (k-1, j^* + 1), (k-2, j^* + 1), \ldots, (1, j^* + 1))$$

for $k = 1, 2, 3, \ldots, q = r\text{-}1$.

Let the starting unit be (1, 3), so $j^* = 3$. Thus, we have all possible paths with their labeled units as follows:

P_1 = ((1, 3),(1, 2),(1, 1),(2, 1),(2, 2),(2, 3),(2, 4),(2, 5),(2, 6),(1, 6),(1, 5),(1, 4))

P_2 = ((1, 3),(2, 3),(2, 2),(2, 1),(3, 1),(3, 2),(3, 3),(3, 4),(3, 5),(3, 6),(2, 6),(2, 5), (2, 4),(1, 4))

P_3 = ((1, 3),(2, 3),(3, 3),(3, 2),(3, 1),(4, 1),(4, 2),(4, 3),(4, 4),(4, 5),(4, 6),(3, 6),(3, 5), (3, 4), (2, 4),(1, 4))

Since the number of units belonging to P_k is $2c + 2(k\text{-}1)$, the number of units belonging to P_1 is $2(6)+ 2(1\text{-}1) = 12$ units, the number of units belonging to P_2 is

2(6)+ 2(2-1) = 14 units, and the number of units belonging to P_3 is 2(6)+ 2(3-1) = 16 units.

Suppose the number of sampled paths is 2. According to the SRSWOR, p = 2 sample paths are selected. There are 3 possible samples, which are $p_{s1}=(P_1, P_2)$, $p_{s2}=(P_1, P_3)$ and $p_{s3}=(P_2, P_3)$. Since there is overlapping of paths, each sample is reduced to the set of distinct units in the sample for the purpose of applying the Horvitz-Thompson estimator as follows:

$p_{s1}=(P_1, P_2)$ reduces to s_1= {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2,1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)}

$p_{s2}=(P_1, P_3)$ reduces to s_2= {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2,1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6)}

$p_{s3}=(P_2, P_3)$ reduces to s_3= {(1, 3), (1, 4), (2,1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6)}

Next, the inclusion probabilities are calculated by the formula from equation (3.2). First, we will calculate the inclusion probabilities for the units in column 3 and 4 (units of type 1). For i = 1, 2, 3, 4 and j = 3 and 4, we have

$$\pi_{(i,j)} = 1 - \frac{\binom{i-2}{p}}{\binom{q}{p}}$$

Then, we get

$$\pi_{(1,3)} = 1 - \frac{\binom{1-2}{2}}{\binom{3}{2}} = 1 - 0 = 1 = \pi_{(1,4)} \qquad \pi_{(2,3)} = 1 - \frac{\binom{2-2}{2}}{\binom{3}{2}} = 1 - 0 = 1 = \pi_{(2,4)}$$

$$\pi_{(3,3)} = 1 - \frac{\binom{3-2}{2}}{\binom{3}{2}} = 1 - 0 = 1 = \pi_{(3,4)} \qquad \pi_{(4,3)} = 1 - \frac{\binom{4-2}{2}}{\binom{3}{2}} = 1 - \frac{1}{3} = \frac{2}{3} = \pi_{(4,4)}$$

Next, the inclusion probabilities for the units not in the column 3 and 4 will be calculated and not in the first row or the last row (units of type 2). For $i = 2, 3$ and $j = 1, 2, 5, 6$

$$\pi_{(i,j)} = 1 - \frac{\binom{q-2}{p}}{\binom{q}{p}}$$

Then, we get

$$\pi_{(2,1)} = 1 - \frac{\binom{3-2}{2}}{\binom{3}{2}} = 1 - 0 = 1 = \pi_{(2,2)} = \pi_{(2,5)} = \pi_{(2,6)} = \pi_{(3,1)} = \pi_{(3,2)} = \pi_{(3,5)} = \pi_{(3,6)}$$

Finally, the inclusion probabilities for the units in the first row and the last row will be calculated, but not in column 3 or 4 (units of type 3). For $i = 1$ and 4 and $j = 1, 2, 5, 6$

$$\pi_{(i,j)} = 1 - \frac{\binom{q-1}{p}}{\binom{q}{p}}$$

Then, we get

$$\pi_{(1,1)} = 1 - \frac{\binom{3-1}{2}}{\binom{3}{2}} = 1 - \frac{1}{3} = \frac{2}{3} = \pi_{(1,2)} = \pi_{(1,5)} = \pi_{(1,6)} = \pi_{(4,1)} = \pi_{(4,2)} = \pi_{(4,5)} = \pi_{(4,6)}$$

The inclusion probabilities are shown in Figure 3.9. The estimates of the mean for all possible samples are shown in Table 3.3. We can see that $\hat{\mu}_{ps}$ is an unbiased estimator since its bias is zero.

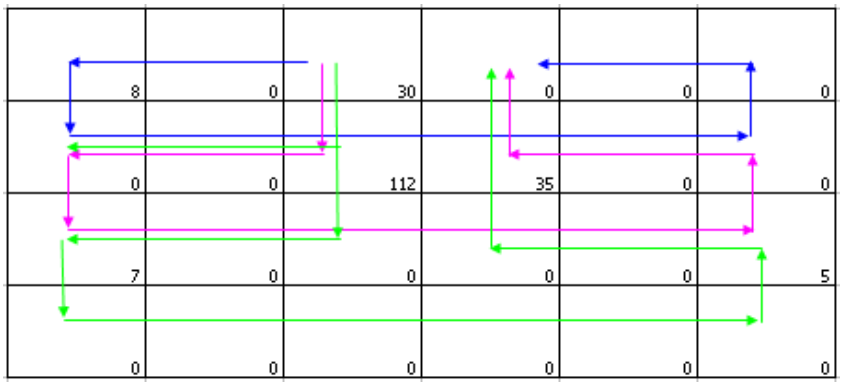


Figure 3.8 All Possible Paths of the Spatial Population of the 4 Rows and 6 Columns with the y-value of Each Unit

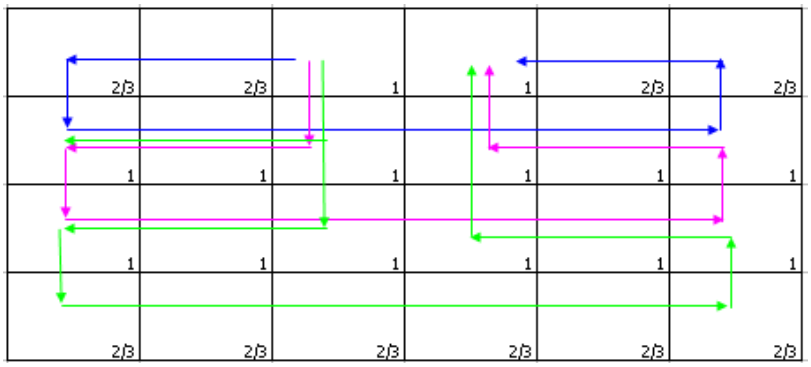


Figure 3.9 The Inclusion Probabilities of the Population of 4 Rows and 6 Columns

Table 3.3 Estimates of the Mean and Variance Estimator for all Possible Samples

Sample	$\hat{\mu}_{ps}$	Sample size	$\hat{v}(\hat{\mu}_{ps})$
$p_{s1}=(P_1, P_2)$	8.375	18	0.083
$p_{s2}=(P_1, P_3)$	8.375	24	0.083
$p_{s3}=(P_2, P_3)$	7.875	20	0.000
Mean	8.208	20.67	0.056
Bias	0		0
Variance	0.056		

The calculation of the estimate of the mean for sample $p_{s1}=(P_1, P_2)$ is shown in Table 3.4. $p_{s1}=(P_1, P_2)$ reduces to $s_1=$ {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2,1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)} corresponding to y = {8, 7, 30, 24, 6, 5, 0, 10, 112, 35, 5, 8, 7, 7, 32, 0, 0, 5}. According to Table 3.4, $\sum_{(i,j)\in s_1} \frac{y_{(i,j)}}{\pi_{(i,j)}} = 201$. By using equation (3.19),

$$\hat{\mu}_{ps} = \frac{1}{rc}\sum_{(i,j)\in s_1}\frac{y_{(i,j)}}{\pi_{(i,j)}} = \frac{1}{4(6)}\left[\frac{y_{(1,1)}}{\pi_{(1,1)}} + \frac{y_{(1,2)}}{\pi_{(1,2)}} + ... + \frac{y_{(3,6)}}{\pi_{(3,6)}}\right] = \frac{1}{24}\left[\frac{8}{2/3} + \frac{0}{2/3} + ... + \frac{5}{1}\right]$$

$$= \frac{201}{24} = 8.375$$

To calculate $\hat{v}(\hat{\mu}_{ps})$ of this sample and $v(\hat{\mu}_{ps})$, the joint inclusion is calculated using the formula in equation (3.6).

$$\pi_{(i,j),(i',j')} = \pi_{(i,j)} + \pi_{(i',j')} - (1 - \frac{\binom{f}{p}}{\binom{q}{p}})$$

$$= \pi_{(i,j)} + \pi_{(i',j')} - (1 - \frac{\binom{f}{2}}{\binom{3}{2}}) \quad \text{since } p = 2 \text{ and } q = 3$$

$$= \pi_{(i,j)} + \pi_{(i',j')} - (1 - \frac{\binom{f}{2}}{3}),$$

where f = the number of paths not containing unit (i, j) and (i', j'). Note that if $f < p$, then it is set that $\binom{f}{p} = 0$.

Case 1: For Units of Type 1

For i, i' = 1, 2, 3, …, r and j, $j' = j^*$ and j^*+1

$$f = \min(i, i') - 2$$

Note that if $f < 0$, then it is set that $f = 0$. Here, $j^* = 3$ and $j^*+1 = 4$. The units of type 1 are the units in column 3 and 4, which are (1, 3), (2, 3), (3, 3), (4, 3), (1, 4), (2, 4), (3, 4), (4, 4). The calculation of the joint inclusion probabilities for units of type 1 is shown in Table 3.5.

Table 3.4 The Calculation of the Estimate of the Mean for Sample s_1

i	j	$y_{(i,j)}$	$\pi_{(i,j)}$	$y_{(i,j)}/\pi_{(i,j)}$
1	1	8	0.67	12
1	2	0	0.67	0
1	3	30	1	30
1	4	0	1	0
1	5	0	0.67	0
1	6	0	0.67	0
2	1	0	1	0
2	2	0	1	0
2	3	112	1	112
2	4	35	1	35
2	5	0	1	0
2	6	0	1	0
3	1	7	1	7
3	2	0	1	0
3	3	0	1	0
3	4	0	1	0
3	5	0	1	0
3	6	5	1	5
			sum	201

Table 3.5 The Calculation of Joint Inclusion Probabilities for Units of type 1 (Case 1)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
1	3	1	3	0	0	1.00	1.00	1.00	1.00
1	3	2	3	0	0	1.00	1.00	1.00	1.00
1	3	3	3	0	0	1.00	1.00	1.00	1.00
1	3	4	3	0	0	1.00	1.00	0.67	0.67
1	3	1	4	0	0	1.00	1.00	1.00	1.00
1	3	2	4	0	0	1.00	1.00	1.00	1.00
1	3	3	4	0	0	1.00	1.00	1.00	1.00
1	3	4	4	0	0	1.00	1.00	0.67	0.67
2	3	1	3	0	0	1.00	1.00	1.00	1.00
2	3	2	3	0	0	1.00	1.00	1.00	1.00
2	3	3	3	0	0	1.00	1.00	1.00	1.00
2	3	4	3	0	0	1.00	1.00	0.67	0.67
2	3	1	4	0	0	1.00	1.00	1.00	1.00
2	3	2	4	0	0	1.00	1.00	1.00	1.00
2	3	3	4	0	0	1.00	1.00	1.00	1.00
2	3	4	4	0	0	1.00	1.00	0.67	0.67
3	3	1	3	0	0	1.00	1.00	1.00	1.00
3	3	2	3	0	0	1.00	1.00	1.00	1.00
3	3	3	3	1	0	1.00	1.00	1.00	1.00
3	3	4	3	1	0	1.00	1.00	0.67	0.67
3	3	1	4	0	0	1.00	1.00	1.00	1.00
3	3	2	4	0	0	1.00	1.00	1.00	1.00
3	3	3	4	1	0	1.00	1.00	1.00	1.00
3	3	4	4	1	0	1.00	1.00	0.67	0.67
4	3	1	3	0	0	1.00	0.67	1.00	0.67
4	3	2	3	0	0	1.00	0.67	1.00	0.67
4	3	3	3	1	0	1.00	0.67	1.00	0.67
4	3	4	3	2	1	0.67	0.67	0.67	0.67
4	3	1	4	0	0	1.00	0.67	1.00	0.67
4	3	2	4	0	0	1.00	0.67	1.00	0.67
4	3	3	4	1	0	1.00	0.67	1.00	0.67
4	3	4	4	2	1	0.67	0.67	0.67	0.67
1	4	1	3	0	0	1.00	1.00	1.00	1.00
1	4	2	3	0	0	1.00	1.00	1.00	1.00
1	4	3	3	0	0	1.00	1.00	1.00	1.00
1	4	4	3	0	0	1.00	1.00	0.67	0.67
1	4	1	4	0	0	1.00	1.00	1.00	1.00
1	4	2	4	0	0	1.00	1.00	1.00	1.00
1	4	3	4	0	0	1.00	1.00	1.00	1.00
1	4	4	4	0	0	1.00	1.00	0.67	0.67
2	4	1	3	0	0	1.00	1.00	1.00	1.00
2	4	2	3	0	0	1.00	1.00	1.00	1.00
2	4	3	3	0	0	1.00	1.00	1.00	1.00
2	4	4	3	0	0	1.00	1.00	0.67	0.67
2	4	1	4	0	0	1.00	1.00	1.00	1.00
2	4	2	4	0	0	1.00	1.00	1.00	1.00
2	4	3	4	0	0	1.00	1.00	1.00	1.00
2	4	4	4	0	0	1.00	1.00	0.67	0.67
3	4	1	3	0	0	1.00	1.00	1.00	1.00
3	4	2	3	0	0	1.00	1.00	1.00	1.00

Table3.5 (Continued)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
3	4	3	3	1	0	1.00	1.00	1.00	1.00
3	4	4	3	1	0	1.00	1.00	0.67	0.67
3	4	3	4	1	0	1.00	1.00	1.00	1.00
3	4	4	4	1	0	1.00	1.00	0.67	0.67
4	4	1	3	0	0	1.00	0.67	1.00	0.67
4	4	2	3	0	0	1.00	0.67	1.00	0.67
4	4	3	3	1	0	1.00	0.67	1.00	0.67
4	4	4	3	2	1	0.67	0.67	0.67	0.67
4	4	1	4	0	0	1.00	0.67	1.00	0.67
4	4	2	4	0	0	1.00	0.67	1.00	0.67
4	4	3	4	1	0	1.00	0.67	1.00	0.67
4	4	4	4	2	1	0.67	0.67	0.67	0.67

Case 2: For Units of Type 1 and 2

For i = 1, 2, 3, …, r and $j = j^*$and j^*+1 and i' = 2, 3, …, r-1 and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c

$$f = \begin{cases} i-2 & if\ i \le i' \\ i-3 & if\ i-i' = 1 \\ i-4 & if\ i-i' \ge 1 \end{cases}$$

Note that if $f < 0$, then it is set that $f = 0$. Units of type 1 are unit (i,j) = (1, 3), (2, 3), (3, 3), (4, 3), (1, 4), (2, 4),(3, 4), (4, 4). Units of type 2 are unit (i', j') = (2,1), (2,2), (2,5), (2,6), (3,1), (3,2), (3,5), (3,6). The calculation of the joint inclusion probabilities for units of type 1 and 2 is shown in Table 3.6.

Table 3.6 The Calculation of Joint Inclusion Probabilities for Units of type 1 and 2 (Case 2)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
1	3	2	1	0	0	1.00	1.00	1.00	1.00
1	3	2	2	0	0	1.00	1.00	1.00	1.00
1	3	2	5	0	0	1.00	1.00	1.00	1.00
1	3	2	6	0	0	1.00	1.00	1.00	1.00
1	3	3	1	0	0	1.00	1.00	1.00	1.00
1	3	3	2	0	0	1.00	1.00	1.00	1.00
1	3	3	5	0	0	1.00	1.00	1.00	1.00
1	3	3	6	0	0	1.00	1.00	1.00	1.00
2	3	2	1	0	0	1.00	1.00	1.00	1.00
2	3	2	2	0	0	1.00	1.00	1.00	1.00
2	3	2	5	0	0	1.00	1.00	1.00	1.00
2	3	2	6	0	0	1.00	1.00	1.00	1.00
2	3	3	1	0	0	1.00	1.00	1.00	1.00
2	3	3	2	0	0	1.00	1.00	1.00	1.00
2	3	3	5	0	0	1.00	1.00	1.00	1.00
2	3	3	6	0	0	1.00	1.00	1.00	1.00
3	3	2	1	0	0	1.00	1.00	1.00	1.00
3	3	2	2	0	0	1.00	1.00	1.00	1.00
3	3	2	5	0	0	1.00	1.00	1.00	1.00
3	3	2	6	0	0	1.00	1.00	1.00	1.00
3	3	3	1	1	0	1.00	1.00	1.00	1.00
3	3	3	2	1	0	1.00	1.00	1.00	1.00
3	3	3	5	1	0	1.00	1.00	1.00	1.00
3	3	3	6	1	0	1.00	1.00	1.00	1.00
4	3	2	1	0	0	1.00	0.67	1.00	0.67
4	3	2	2	0	0	1.00	0.67	1.00	0.67
4	3	2	5	0	0	1.00	0.67	1.00	0.67
4	3	2	6	0	0	1.00	0.67	1.00	0.67
4	3	3	1	1	0	1.00	0.67	1.00	0.67
4	3	3	2	1	0	1.00	0.67	1.00	0.67
4	3	3	5	1	0	1.00	0.67	1.00	0.67
4	3	3	6	1	0	1.00	0.67	1.00	0.67
1	4	2	1	0	0	1.00	1.00	1.00	1.00
1	4	2	2	0	0	1.00	1.00	1.00	1.00
1	4	2	5	0	0	1.00	1.00	1.00	1.00
1	4	2	6	0	0	1.00	1.00	1.00	1.00
1	4	3	1	0	0	1.00	1.00	1.00	1.00
1	4	3	2	0	0	1.00	1.00	1.00	1.00
1	4	3	5	0	0	1.00	1.00	1.00	1.00
1	4	3	6	0	0	1.00	1.00	1.00	1.00
2	4	2	1	0	0	1.00	1.00	1.00	1.00
2	4	2	2	0	0	1.00	1.00	1.00	1.00
2	4	2	5	0	0	1.00	1.00	1.00	1.00
2	4	2	6	0	0	1.00	1.00	1.00	1.00

Table3.6 (Continued)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
2	4	3	1	0	0	1.00	1.00	1.00	1.00
2	4	3	2	0	0	1.00	1.00	1.00	1.00
2	4	3	5	0	0	1.00	1.00	1.00	1.00
2	4	3	6	0	0	1.00	1.00	1.00	1.00
3	4	2	1	0	0	1.00	1.00	1.00	1.00
3	4	2	2	0	0	1.00	1.00	1.00	1.00
3	4	2	5	0	0	1.00	1.00	1.00	1.00
3	4	2	6	0	0	1.00	1.00	1.00	1.00
3	4	3	1	1	0	1.00	1.00	1.00	1.00
3	4	3	2	1	0	1.00	1.00	1.00	1.00
3	4	3	5	1	0	1.00	1.00	1.00	1.00
3	4	3	6	1	0	1.00	1.00	1.00	1.00
4	4	2	1	0	0	1.00	0.67	1.00	0.67
4	4	2	2	0	0	1.00	0.67	1.00	0.67
4	4	2	5	0	0	1.00	0.67	1.00	0.67
4	4	2	6	0	0	1.00	0.67	1.00	0.67
4	4	3	1	1	0	1.00	0.67	1.00	0.67
4	4	3	2	1	0	1.00	0.67	1.00	0.67
4	4	3	5	1	0	1.00	0.67	1.00	0.67
4	4	3	6	1	0	1.00	0.67	1.00	0.67

Case 3: For Units of Type 1 and 3

For $i = 1, 2, 3, \ldots, r$ and $j = j^*$ and j^*+1 and $i' = 1$ and r and $j' = 1, 2, 3, \ldots, j^*-1, j^*+2, j^*+3, \ldots, c$

$$f = \begin{cases} i-2 & i \le i' \\ i-3 & i > i' \end{cases}$$

Note that if $f < 0$, then it is set that $f = 0$. Units of type 1 are unit $(i,j) = (1, 3), (2, 3), (3, 3), (4, 3), (1, 4), (2, 4), (3, 4), (4, 4)$. Units of type 3 are unit $(i', j') = (1,1), (1,2), (1,5), (1,6), (4,1), (4,2), (4,5), (4,6)$. The calculation of the joint inclusion probabilities for units of type 1 and 3 is shown in Table 3.7.

Table 3.7 The Calculation of Joint Inclusion Probabilities for Units of type 1 and 3 (Case 3)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
1	3	1	1	0	0	1.00	1.00	0.67	0.67
1	3	1	2	0	0	1.00	1.00	0.67	0.67
1	3	1	5	0	0	1.00	1.00	0.67	0.67
1	3	1	6	0	0	1.00	1.00	0.67	0.67
1	3	4	1	0	0	1.00	1.00	0.67	0.67
1	3	4	2	0	0	1.00	1.00	0.67	0.67
1	3	4	5	0	0	1.00	1.00	0.67	0.67
1	3	4	6	0	0	1.00	1.00	0.67	0.67
2	3	1	1	0	0	1.00	1.00	0.67	0.67
2	3	1	2	0	0	1.00	1.00	0.67	0.67
2	3	1	5	0	0	1.00	1.00	0.67	0.67
2	3	1	6	0	0	1.00	1.00	0.67	0.67
2	3	4	1	0	0	1.00	1.00	0.67	0.67
2	3	4	2	0	0	1.00	1.00	0.67	0.67
2	3	4	5	0	0	1.00	1.00	0.67	0.67
2	3	4	6	0	0	1.00	1.00	0.67	0.67
3	3	1	1	0	0	1.00	1.00	0.67	0.67
3	3	1	2	0	0	1.00	1.00	0.67	0.67
3	3	1	5	0	0	1.00	1.00	0.67	0.67
3	3	1	6	0	0	1.00	1.00	0.67	0.67
3	3	4	1	1	0	1.00	1.00	0.67	0.67
3	3	4	2	1	0	1.00	1.00	0.67	0.67
3	3	4	5	1	0	1.00	1.00	0.67	0.67
3	3	4	6	1	0	1.00	1.00	0.67	0.67
4	3	1	1	1	0	1.00	0.67	0.67	0.33
4	3	1	2	1	0	1.00	0.67	0.67	0.33
4	3	1	5	1	0	1.00	0.67	0.67	0.33
4	3	1	6	1	0	1.00	0.67	0.67	0.33
4	3	4	1	2	1	0.67	0.67	0.67	0.67
4	3	4	2	2	1	0.67	0.67	0.67	0.67
4	3	4	5	2	1	0.67	0.67	0.67	0.67
4	3	4	6	2	1	0.67	0.67	0.67	0.67
1	4	1	1	0	0	1.00	1.00	0.67	0.67
1	4	1	2	0	0	1.00	1.00	0.67	0.67
1	4	1	5	0	0	1.00	1.00	0.67	0.67
1	4	1	6	0	0	1.00	1.00	0.67	0.67
1	4	4	1	0	0	1.00	1.00	0.67	0.67
1	4	4	2	0	0	1.00	1.00	0.67	0.67
1	4	4	5	0	0	1.00	1.00	0.67	0.67
1	4	4	6	0	0	1.00	1.00	0.67	0.67
2	4	1	1	0	0	1.00	1.00	0.67	0.67
2	4	1	2	0	0	1.00	1.00	0.67	0.67
2	4	1	5	0	0	1.00	1.00	0.67	0.67
2	4	1	6	0	0	1.00	1.00	0.67	0.67

Table3.7 (Continued)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
2	4	4	1	0	0	1.00	1.00	0.67	0.67
2	4	4	2	0	0	1.00	1.00	0.67	0.67
2	4	4	5	0	0	1.00	1.00	0.67	0.67
2	4	4	6	0	0	1.00	1.00	0.67	0.67
3	4	1	1	0	0	1.00	1.00	0.67	0.67
3	4	1	2	0	0	1.00	1.00	0.67	0.67
3	4	1	5	0	0	1.00	1.00	0.67	0.67
3	4	1	6	0	0	1.00	1.00	0.67	0.67
3	4	4	1	1	0	1.00	1.00	0.67	0.67
3	4	4	2	1	0	1.00	1.00	0.67	0.67
3	4	4	5	1	0	1.00	1.00	0.67	0.67
3	4	4	6	1	0	1.00	1.00	0.67	0.67
4	4	1	1	1	0	1.00	0.67	0.67	0.33
4	4	1	2	1	0	1.00	0.67	0.67	0.33
4	4	1	5	1	0	1.00	0.67	0.67	0.33
4	4	1	6	1	0	1.00	0.67	0.67	0.33
4	4	4	1	2	1	0.67	0.67	0.67	0.67
4	4	4	2	2	1	0.67	0.67	0.67	0.67
4	4	4	5	2	1	0.67	0.67	0.67	0.67
4	4	4	6	2	1	0.67	0.67	0.67	0.67

Case 4: For Units of Type 2

For $i, i' = 2, 3, \ldots, r\text{-}1$ and $j,\ j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots,$ c

$$f = \begin{cases} q-2 & if\ |i-i'| = 0 \\ q-3 & if\ |i-i'| = 1 \\ q-4 & if\ |i-i'| \geq 2 \end{cases}$$

Units of type 2 are (2,1), (2,2), (2,5), (2,6), (3,1), (3,2), (3,5), (3,6). The calculation of the joint inclusion probabilities for units of type 2 is shown in Table 3.8.

Table 3.8 The Calculation of Joint Inclusion Probabilities for Units of type 2 (case 4)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
2	1	2	1	1	0	1	1	1	1
2	1	2	2	1	0	1	1	1	1
2	1	2	5	1	0	1	1	1	1
2	1	2	6	1	0	1	1	1	1
2	1	3	1	0	0	1	1	1	1
2	1	3	2	0	0	1	1	1	1
2	1	3	5	0	0	1	1	1	1
2	1	3	6	0	0	1	1	1	1
2	2	2	1	1	0	1	1	1	1
2	2	2	2	1	0	1	1	1	1
2	2	2	5	1	0	1	1	1	1
2	2	2	6	1	0	1	1	1	1
2	2	3	1	0	0	1	1	1	1
2	2	3	2	0	0	1	1	1	1
2	2	3	5	0	0	1	1	1	1
2	2	3	6	0	0	1	1	1	1
2	5	2	1	1	0	1	1	1	1
2	5	2	2	1	0	1	1	1	1
2	5	2	5	1	0	1	1	1	1
2	5	2	6	1	0	1	1	1	1
2	5	3	1	0	0	1	1	1	1
2	5	3	2	0	0	1	1	1	1
2	5	3	5	0	0	1	1	1	1
2	5	3	6	0	0	1	1	1	1
2	6	2	1	1	0	1	1	1	1
2	6	2	2	1	0	1	1	1	1
2	6	2	5	1	0	1	1	1	1
2	6	2	6	1	0	1	1	1	1
2	6	3	1	0	0	1	1	1	1
2	6	3	2	0	0	1	1	1	1
2	6	3	5	0	0	1	1	1	1
2	6	3	6	0	0	1	1	1	1
3	1	2	1	0	0	1	1	1	1
3	1	2	2	0	0	1	1	1	1
3	1	2	5	0	0	1	1	1	1
3	1	2	6	0	0	1	1	1	1
3	1	3	1	1	0	1	1	1	1
3	1	3	2	1	0	1	1	1	1
3	1	3	5	1	0	1	1	1	1
3	1	3	6	1	0	1	1	1	1
3	2	2	1	0	0	1	1	1	1
3	2	2	2	0	0	1	1	1	1
3	2	2	5	0	0	1	1	1	1
3	2	2	6	0	0	1	1	1	1

Table3.8 (Continued)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
3	2	3	1	1	0	1	1	1	1
3	2	3	2	1	0	1	1	1	1
3	5	2	1	0	0	1	1	1	1
3	5	2	2	0	0	1	1	1	1
3	5	2	5	0	0	1	1	1	1
3	5	2	6	0	0	1	1	1	1
3	5	3	1	1	0	1	1	1	1
3	5	3	2	1	0	1	1	1	1
3	5	3	5	1	0	1	1	1	1
3	5	3	6	1	0	1	1	1	1
3	6	2	1	0	0	1	1	1	1
3	6	2	2	0	0	1	1	1	1
3	6	2	5	0	0	1	1	1	1
3	6	2	6	0	0	1	1	1	1
3	6	3	1	1	0	1	1	1	1
3	6	3	2	1	0	1	1	1	1
3	6	3	5	1	0	1	1	1	1
3	6	3	6	1	0	1	1	1	1

Case 5: For Units of Type 2 and 3

For $i = 2, 3, \ldots, r\text{-}1$ and $j = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots,$ c and $i' = 1$ and r and $j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*\text{+}2, j^*\text{+}3, \ldots, c$

$$f = \begin{cases} q-2 & |i-i'|=1 \\ q-3 & |i-i'|\geq 2 \end{cases}$$

Units of type 2 are unit (i,j) = (2,1), (2,2), (2,5), (2,6), (3,1), (3,2), (3,5), (3,6). Units of type 3 are unit (i',j') = (1,1), (1,2), (1,5), (1,6), (4,1), (4,2), (4,5), (4,6). The calculation of the joint inclusion probabilities for units of type 2 and 3 is shown in Table 3.9.

Table 3.9 The Calculation of Joint Inclusion Probabilities for Units of type 2 and 3 (Case 5)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
2	1	1	1	1	0	1	1.00	0.67	0.67
2	1	1	2	1	0	1	1.00	0.67	0.67
2	1	1	5	1	0	1	1.00	0.67	0.67
2	1	1	6	1	0	1	1.00	0.67	0.67
2	1	4	1	0	0	1	1.00	0.67	0.67
2	1	4	2	0	0	1	1.00	0.67	0.67
2	1	4	5	0	0	1	1.00	0.67	0.67
2	1	4	6	0	0	1	1.00	0.67	0.67
2	2	1	1	1	0	1	1.00	0.67	0.67
2	2	1	2	1	0	1	1.00	0.67	0.67
2	2	1	5	1	0	1	1.00	0.67	0.67
2	2	1	6	1	0	1	1.00	0.67	0.67
2	2	4	1	0	0	1	1.00	0.67	0.67
2	2	4	2	0	0	1	1.00	0.67	0.67
2	2	4	5	0	0	1	1.00	0.67	0.67
2	2	4	6	0	0	1	1.00	0.67	0.67
2	5	1	1	1	0	1	1.00	0.67	0.67
2	5	1	2	1	0	1	1.00	0.67	0.67
2	5	1	5	1	0	1	1.00	0.67	0.67
2	5	1	6	1	0	1	1.00	0.67	0.67
2	5	4	1	0	0	1	1.00	0.67	0.67
2	5	4	2	0	0	1	1.00	0.67	0.67
2	5	4	5	0	0	1	1.00	0.67	0.67
2	5	4	6	0	0	1	1.00	0.67	0.67
2	6	1	1	1	0	1	1.00	0.67	0.67
2	6	1	2	1	0	1	1.00	0.67	0.67
2	6	1	5	1	0	1	1.00	0.67	0.67
2	6	1	6	1	0	1	1.00	0.67	0.67
2	6	4	1	0	0	1	1.00	0.67	0.67
2	6	4	2	0	0	1	1.00	0.67	0.67
2	6	4	5	0	0	1	1.00	0.67	0.67
2	6	4	6	0	0	1	1.00	0.67	0.67
3	1	1	1	0	0	1	1.00	0.67	0.67
3	1	1	2	0	0	1	1.00	0.67	0.67
3	1	1	5	0	0	1	1.00	0.67	0.67
3	1	1	6	0	0	1	1.00	0.67	0.67
3	1	4	1	1	0	1	1.00	0.67	0.67
3	1	4	2	1	0	1	1.00	0.67	0.67
3	1	4	5	1	0	1	1.00	0.67	0.67
3	1	4	6	1	0	1	1.00	0.67	0.67
3	2	1	1	0	0	1	1.00	0.67	0.67
3	2	1	2	0	0	1	1.00	0.67	0.67
3	2	1	5	0	0	1	1.00	0.67	0.67
3	2	1	6	0	0	1	1.00	0.67	0.67

Table3.9 (Continued)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
3	2	4	1	1	0	1	1.00	0.67	0.67
3	2	4	2	1	0	1	1.00	0.67	0.67
3	5	1	1	0	0	1	1.00	0.67	0.67
3	5	1	2	0	0	1	1.00	0.67	0.67
3	5	1	5	0	0	1	1.00	0.67	0.67
3	5	1	6	0	0	1	1.00	0.67	0.67
3	5	4	1	1	0	1	1.00	0.67	0.67
3	5	4	2	1	0	1	1.00	0.67	0.67
3	5	4	5	1	0	1	1.00	0.67	0.67
3	5	4	6	1	0	1	1.00	0.67	0.67
3	6	1	1	0	0	1	1.00	0.67	0.67
3	6	1	2	0	0	1	1.00	0.67	0.67
3	6	1	5	0	0	1	1.00	0.67	0.67
3	6	1	6	0	0	1	1.00	0.67	0.67
3	6	4	1	1	0	1	1.00	0.67	0.67
3	6	4	2	1	0	1	1.00	0.67	0.67
3	6	4	5	1	0	1	1.00	0.67	0.67
3	6	4	6	1	0	1	1.00	0.67	0.67

Case 6: For Units of Type 3

For $i, i' = 1$ and r and $j, j' = 1, 2, 3, \ldots, j^*-1, j^*+2, j^*+3, \ldots, c$

$$f = \begin{cases} q-1 & i = i' \\ q-2 & i \neq i' \end{cases}$$

Note that if $f < 0$, then it is set that $f = 0$. Units of type 3 are (1,1), (1,2), (1,5), (1,6), (4,1), (4,2), (4,5), (4,6). The calculation of the joint inclusion probabilities for units of type 3 is shown in Table 3.10.

Table 3.10 The Calculation of Joint Inclusion Probabilities for Units of type 3 (Case 6)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
1	1	1	1	2	1	0.67	0.67	0.67	0.67
1	1	1	2	2	1	0.67	0.67	0.67	0.67
1	1	1	5	2	1	0.67	0.67	0.67	0.67
1	1	1	6	2	1	0.67	0.67	0.67	0.67
1	1	4	1	1	0	1.00	0.67	0.67	0.33
1	1	4	2	1	0	1.00	0.67	0.67	0.33
1	1	4	5	1	0	1.00	0.67	0.67	0.33
1	1	4	6	1	0	1.00	0.67	0.67	0.33
1	2	1	1	2	1	0.67	0.67	0.67	0.67
1	2	1	2	2	1	0.67	0.67	0.67	0.67
1	2	1	5	2	1	0.67	0.67	0.67	0.67
1	2	1	6	2	1	0.67	0.67	0.67	0.67
1	2	4	1	1	0	1.00	0.67	0.67	0.33
1	2	4	2	1	0	1.00	0.67	0.67	0.33
1	2	4	5	1	0	1.00	0.67	0.67	0.33
1	2	4	6	1	0	1.00	0.67	0.67	0.33
1	5	1	1	2	1	0.67	0.67	0.67	0.67
1	5	1	2	2	1	0.67	0.67	0.67	0.67
1	5	1	5	2	1	0.67	0.67	0.67	0.67
1	5	1	6	2	1	0.67	0.67	0.67	0.67
1	5	4	1	1	0	1.00	0.67	0.67	0.33
1	5	4	2	1	0	1.00	0.67	0.67	0.33
1	5	4	5	1	0	1.00	0.67	0.67	0.33
1	5	4	6	1	0	1.00	0.67	0.67	0.33
1	6	1	1	2	1	0.67	0.67	0.67	0.67
1	6	1	2	2	1	0.67	0.67	0.67	0.67
1	6	1	5	2	1	0.67	0.67	0.67	0.67
1	6	1	6	2	1	0.67	0.67	0.67	0.67
1	6	4	1	1	0	1.00	0.67	0.67	0.33
1	6	4	2	1	0	1.00	0.67	0.67	0.33
1	6	4	5	1	0	1.00	0.67	0.67	0.33
1	6	4	6	1	0	1.00	0.67	0.67	0.33
4	1	1	1	1	0	1.00	0.67	0.67	0.33
4	1	1	2	1	0	1.00	0.67	0.67	0.33
4	1	1	5	1	0	1.00	0.67	0.67	0.33
4	1	1	6	1	0	1.00	0.67	0.67	0.33
4	1	4	1	2	1	0.67	0.67	0.67	0.67
4	1	4	2	2	1	0.67	0.67	0.67	0.67
4	1	4	5	2	1	0.67	0.67	0.67	0.67
4	1	4	6	2	1	0.67	0.67	0.67	0.67
4	2	1	1	1	0	1.00	0.67	0.67	0.33
4	2	1	2	1	0	1.00	0.67	0.67	0.33
4	2	1	5	1	0	1.00	0.67	0.67	0.33
4	2	1	6	1	0	1.00	0.67	0.67	0.33

Table3.10 (Continued)

i	j	i'	j'	f	$\binom{f}{2}$	$1-\frac{\binom{f}{2}}{3}$	$\pi_{(i,j)}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$
4	2	4	1	2	1	0.67	0.67	0.67	0.67
4	2	4	2	2	1	0.67	0.67	0.67	0.67
4	5	1	1	1	0	1.00	0.67	0.67	0.33
4	5	1	2	1	0	1.00	0.67	0.67	0.33
4	5	1	5	1	0	1.00	0.67	0.67	0.33
4	5	1	6	1	0	1.00	0.67	0.67	0.33
4	5	4	1	2	1	0.67	0.67	0.67	0.67
4	5	4	2	2	1	0.67	0.67	0.67	0.67
4	5	4	5	2	1	0.67	0.67	0.67	0.67
4	5	4	6	2	1	0.67	0.67	0.67	0.67
4	6	1	1	1	0	1.00	0.67	0.67	0.33
4	6	1	2	1	0	1.00	0.67	0.67	0.33
4	6	1	5	1	0	1.00	0.67	0.67	0.33
4	6	1	6	1	0	1.00	0.67	0.67	0.33
4	6	4	1	2	1	0.67	0.67	0.67	0.67
4	6	4	2	2	1	0.67	0.67	0.67	0.67
4	6	4	5	2	1	0.67	0.67	0.67	0.67
4	6	4	6	2	1	0.67	0.67	0.67	0.67

The estimated variance, using equation (3.22), for sample s_1, is

$$\hat{v}(\hat{\mu}_{ps}) = \frac{1}{(rc)^2}\left(\sum_{(i,j)\in s_1} \left(\frac{1}{\pi^2_{(i,j)}} - \frac{1}{\pi_{(i,j)}} \right) y^2_{(i,j)} + \sum_{\substack{(i,j)\,,\ (i',j')\in s_1 \\ (i,j)\neq(i',j')}} \left(\frac{1}{\pi_{(i,j)}\pi_{(i',j')}} - \frac{1}{\pi_{(i,j),\,(i',j')}} \right) y_{(i,j)} y_{(i',j')} \right)$$

$$= \frac{1}{(rc)^2}\left(\sum_{(i,j)\,,\ (i',j')\in s_1} \left(\frac{1}{\pi_{(i,j)}\pi_{(i',j')}} - \frac{1}{\pi_{(i,j),\,(i',j')}} \right) y_{(i,j)} y_{(i',j')} \right)$$

where s_1= {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2,1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)}.

To calculate $\hat{v}(\hat{\mu}_{ps})$, $y_{(i,j)}$, $\pi_{(i,j)}$ and $\pi_{(i,j),(i',j')}$ in sample s_1 are used in that formula. Only units with a y-value greater than zero are utilized in the formula; otherwise each term in the formula yields zero. The calculation of $\hat{v}(\hat{\mu}_{ps})$ is shown in Table 3.11. We get

$$\sum_{(i,j)\,,\,(i',j')\,\in s_1}\left(\frac{1}{\pi_{(i,j)}\pi_{(i',j')}}-\frac{1}{\pi_{(i,j),\,(i',j')}}\right)y_{(i,j)}y_{(i',j')}=48.$$

Thus,

$$\hat{v}(\hat{\mu}_{ps}) \;=\frac{1}{(24)^2}(48)=0.083$$

To calculate $v(\hat{\mu}_{ps})$, $y_{(i,j)}$, $\pi_{(i,j)}$ and $\pi_{(i,j),(i',j')}$ in the population are used in the formula from equation (3.21).

$$v(\hat{\mu}_{ps})=\frac{1}{(rc)^2}\left(\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\frac{1-\pi_{(i,j)}}{\pi_{(i,j)}}\right)y_{(i,j)}^2+\sum_{i=1}^{r}\sum_{i'\neq i}\;\sum_{j=1}^{c}\sum_{j'\neq j}\left(\frac{\pi_{(i,j),\,(i',j')}-\pi_{(i,j)}\pi_{(i',j')}}{\pi_{(i,j)}\pi_{(i',j')}}\right)y_{(i,j)}\,y_{(i',j')}\right)$$

$$=\frac{1}{(rc)^2}\left(\sum_{all\;(i,j)\,and\,(i',j')}\left(\frac{\pi_{(i,j),\,(i',j')}-\pi_{(i,j)}\pi_{(i',j')}}{\pi_{(i,j)}\pi_{(i',j')}}\right)y_{(i,j)}\,y_{(i',j')}\right)$$

Only units with a y-value greater than zero are utilized in the formula; otherwise each term in the formula yields zero. The calculation is shown in Table 3.12. We get

$$\sum_{all\;(i,j)\,and\,(i',j')}\left(\frac{\pi_{(i,j),\,(i',j')}-\pi_{(i,j)}\pi_{(i',j')}}{\pi_{(i,j)}\pi_{(i',j')}}\right)y_{(i,j)}\,y_{(i',j')}=32$$

Then,

$$v(\hat{\mu}_{ps})=\frac{1}{(rc)^2}\left(\sum_{all\;(i,j)\,and\,(i',j')}\left(\frac{\pi_{(i,j),\,(i',j')}-\pi_{(i,j)}\pi_{(i',j')}}{\pi_{(i,j)}\pi_{(i',j')}}\right)y_{(i,j)}\,y_{(i',j')}\right)=\frac{1}{(24)^2}(32)\;=0.056$$

Table 3.11 The Calculation of $\hat{v}(\hat{\mu}_{ps})$

i	j	$y_{(i,j)}$	$\pi_{(i,j)}$	i'	j'	$y_{(i',j')}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$	$\left(\frac{1}{\pi_{(i,j)}\pi_{(i',j')}}-\frac{1}{\pi_{(i,j),\,(i',j')}}\right)y_{(i,j)}y_{(i',j')}$
1	1	8	0.67	1	1	8	0.67	0.67	48
1	1	8	0.67	1	3	30	1	0.67	0
1	1	8	0.67	2	3	112	1	0.67	0
1	1	8	0.67	2	4	35	1	0.67	0
1	1	8	0.67	3	1	7	1	0.67	0
1	1	8	0.67	3	6	5	1	0.67	0
1	3	30	1	1	1	8	0.67	0.67	0
1	3	30	1	1	3	30	1	1	0
1	3	30	1	2	3	112	1	1	0
1	3	30	1	2	4	35	1	1	0
1	3	30	1	3	1	7	1	1	0
1	3	30	1	3	6	5	1	1	0
2	3	112	1	1	1	8	0.67	0.67	0
2	3	112	1	1	3	30	1	1	0
2	3	112	1	2	3	112	1	1	0
2	3	112	1	2	4	35	1	1	0
2	3	112	1	3	1	7	1	1	0
2	3	112	1	3	6	5	1	1	0
2	4	35	1	1	1	8	0.67	0.67	0
2	4	35	1	1	3	30	1	1	0
2	4	35	1	2	3	112	1	1	0
2	4	35	1	2	4	35	1	1	0
2	4	35	1	3	1	7	1	1	0
2	4	35	1	3	6	5	1	1	0
3	1	7	1	1	1	8	0.67	0.67	0
3	1	7	1	1	3	30	1	1	0
3	1	7	1	2	3	112	1	1	0
3	1	7	1	2	4	35	1	1	0
3	1	7	1	3	1	7	1	1	0
3	1	7	1	3	6	5	1	1	0
3	6	5	1	1	1	8	0.67	0.67	0
3	6	5	1	1	3	30	1	1	0
3	6	5	1	2	3	112	1	1	0
3	6	5	1	2	4	35	1	1	0
3	6	5	1	3	1	7	1	1	0
3	6	5	1	3	6	5	1	1	0
								sum	48

Table 3.12 The Calculation of $v(\hat{\mu}_{ps})$

i	j	$y_{(i,j)}$	$\pi_{(i,j)}$	i'	j'	$y_{(i',j')}$	$\pi_{(i',j')}$	$\pi_{(i,j),(i',j')}$	$\left(\frac{\pi_{(i,j),(i',j')}-\pi_{(i,j)}\pi_{(i',j')}}{\pi_{(i,j)}\pi_{(i',j')}}\right)y_{(i,j)}y_{(i',j')}$
1	1	8	0.67	1	1	8	0.67	0.67	32
1	1	8	0.67	1	3	30	1	0.67	0
1	1	8	0.67	2	3	112	1	0.67	0
1	1	8	0.67	2	4	35	1	0.67	0
1	1	8	0.67	3	1	7	1	0.67	0
1	1	8	0.67	3	6	5	1	0.67	0
1	3	30	1	1	1	8	0.67	0.67	0
1	3	30	1	1	3	30	1	1	0
1	3	30	1	2	3	112	1	1	0
1	3	30	1	2	4	35	1	1	0
1	3	30	1	3	1	7	1	1	0
1	3	30	1	3	6	5	1	1	0
2	3	112	1	1	1	8	0.67	0.67	0
2	3	112	1	1	3	30	1	1	0
2	3	112	1	2	3	112	1	1	0
2	3	112	1	2	4	35	1	1	0
2	3	112	1	3	1	7	1	1	0
2	3	112	1	3	6	5	1	1	0
2	4	35	1	1	1	8	0.67	0.67	0
2	4	35	1	1	3	30	1	1	0
2	4	35	1	2	3	112	1	1	0
2	4	35	1	2	4	35	1	1	0
2	4	35	1	3	1	7	1	1	0
2	4	35	1	3	6	5	1	1	0
3	1	7	1	1	1	8	0.67	0.67	0
3	1	7	1	1	3	30	1	1	0
3	1	7	1	2	3	112	1	1	0
3	1	7	1	2	4	35	1	1	0
3	1	7	1	3	1	7	1	1	0
3	1	7	1	3	6	5	1	1	0
3	6	5	1	1	1	8	0.67	0.67	0
3	6	5	1	1	3	30	1	1	0
3	6	5	1	2	3	112	1	1	0
3	6	5	1	2	4	35	1	1	0
3	6	5	1	3	1	7	1	1	0
3	6	5	1	3	6	5	1	1	0
								sum	32

3.5 Path Sampling in a Non-rectangular Region

Previously, path sampling was used for the rectangular study region. Now, the non-rectangular study region is considered. An example of such a region is shown in Figure 3.10.

To apply path sampling in a non-rectangular study region, the first step is to try to create a rectangular region around the non-rectangular region, as shown in Figure 3.11. Then, ordinary path sampling can be used.

Next, the new rectangular region is partitioned into an $r \times c$ (r rows and c columns) grid of rc quadrats or secondary units, as shown in Figure 3.12. Units with no area in the non-rectangular region, which are unit (1, 6), (3, 1), (5, 1), (5, 2), and (5, 6), are called "artificial units" from the rectangular region.

Next step is creating all possible paths as shown in Figure 3.14. The created paths look like ordinary paths, as shown Figure 3.13; however, they are different in that the artificial units will not be visited. This means that the artificial units will not be in any path and will also not be observed.

To estimate the population mean and total, the inclusion probabilities can be calculated by using the formula from (3.2), as with ordinary path sampling.

$$\pi_{(i,j)} = \begin{cases} 1 - \dfrac{\binom{i-2}{p}}{\binom{q}{p}} & i = 1,2,3,4,5 \ \text{and} \ j = 3,4 \\ 1 - \dfrac{\binom{q-2}{p}}{\binom{q}{p}} = 1 - \dfrac{\binom{4-2}{2}}{\binom{4}{2}} = \dfrac{5}{6} = 0.83 & i = 2,3,4 \ \text{and} \ j = 1,2,5 \\ 1 - \dfrac{\binom{q-1}{p}}{\binom{q}{p}} = 1 - \dfrac{\binom{4-1}{2}}{\binom{4}{2}} = \dfrac{1}{2} = 0.5 & i = 1,5 \ \text{and} \ j = 1,2,5 \end{cases}$$

They are not affected by an artificial unit because the inclusion probability of unit (*i, j*) is a function of *i*, *p,* and *q*. Then, the Horvitz-Thompson estimator can be applied.

To illustrate, the population y-value with a total of 264 and a mean of 10.56 is considered, as shown in Figure 3.15, and let the number of sample paths be $p = 2$. Then the inclusion probability of each unit is shown in Figure 3.16. The estimates for each sample are shown in Table 3.13.

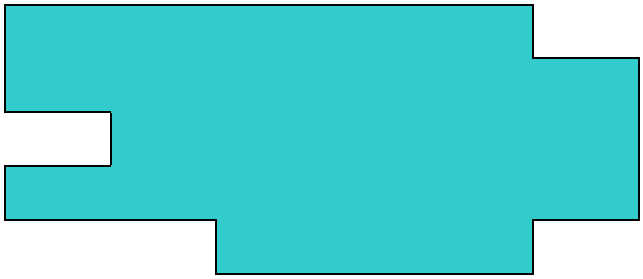


Figure 3.10 A Non-rectangular Region

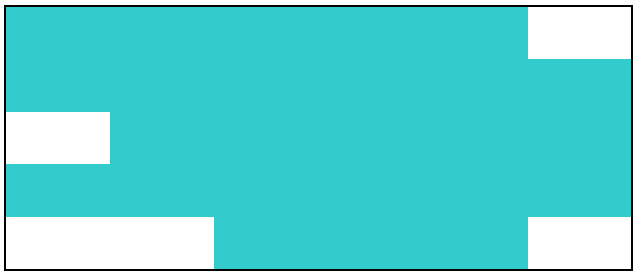


Figure 3.11 New Rectangular Region

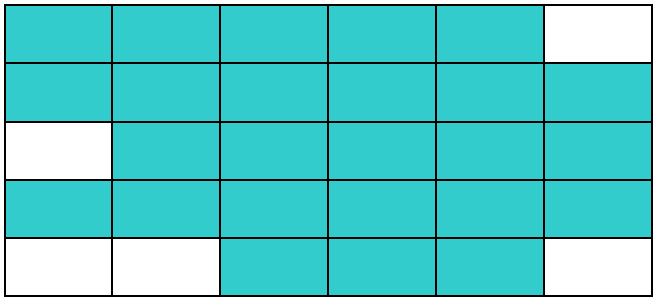


Figure 3.12 New Rectangular Region Partitioned into 5x6

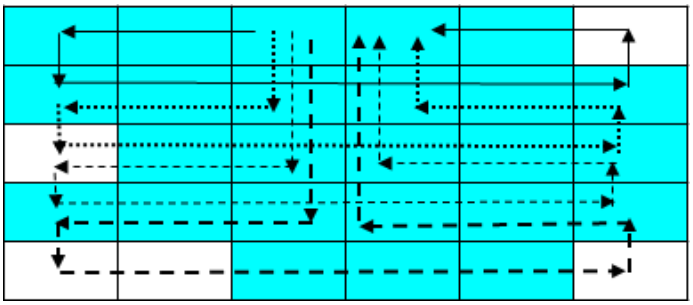


Figure 3.13 All Ordinary Possible Paths for Rectangular Region

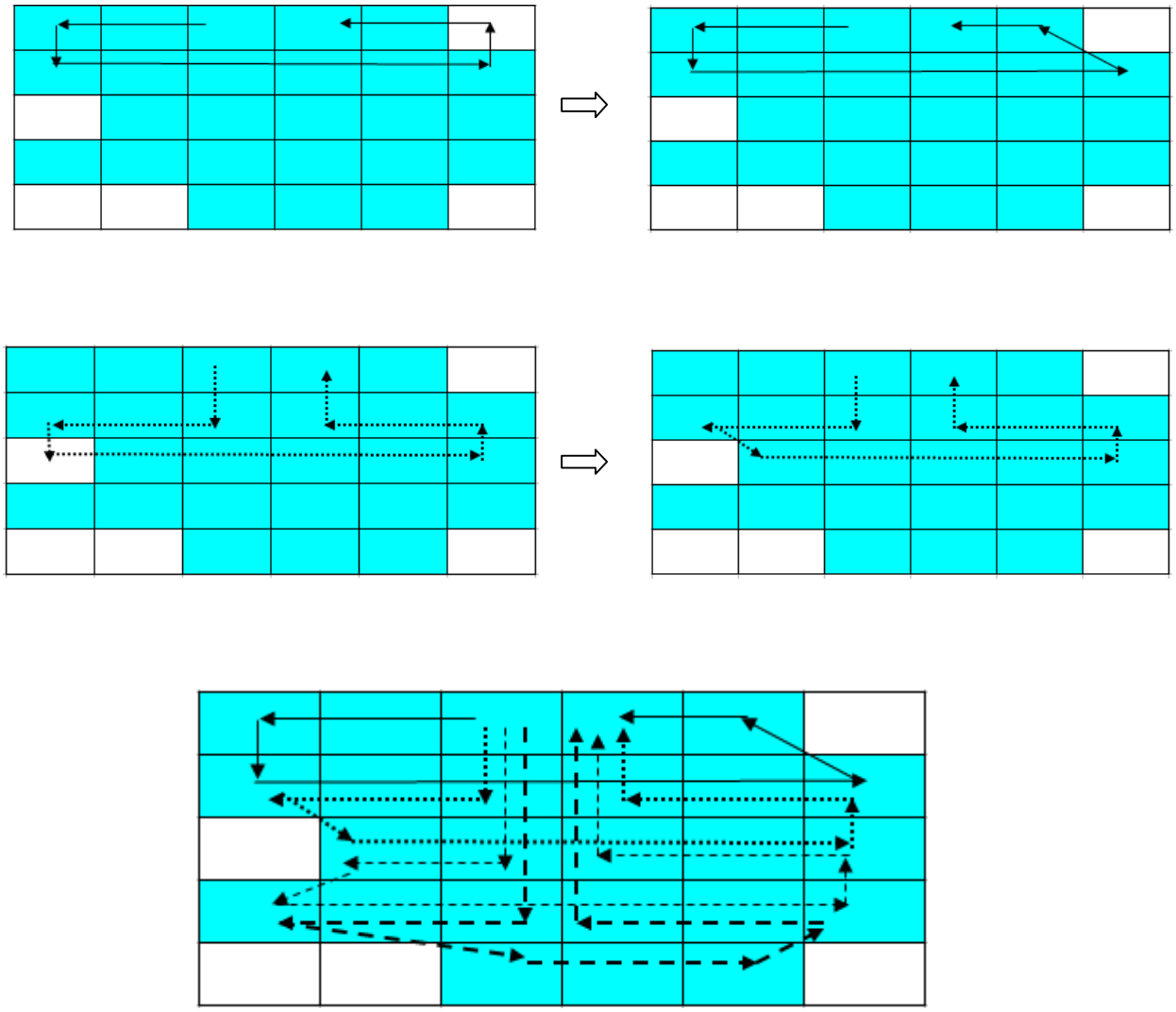


Figure 3.14 All New Possible Paths for Non-rectangular Region

8	7	30	24	6	
7	10	12	35	5	8
	7	32	7	5	5
5	7	7	9	3	5
		10	4	6	

Figure 3.15 Population y-values with a Total of 264 and a Mean of 10.56

0.50	0.50	1.00	1.00	0.50	
0.83	0.83	1.00	1.00	0.83	0.83
	0.83	1.00	1.00	0.83	0.83
0.83	0.83	0.83	0.83	0.83	0.83
		0.50	0.50	0.50	

Figure 3.16 Inclusion Probabilities of Non-rectangular Region

Table 3.13 Estimates of the Mean and Variance Estimator for All Possible Samples for Non-rectangular Region

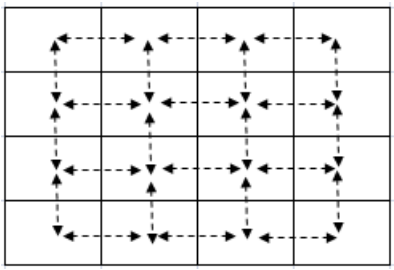
Sample	$\hat{\mu}_{ps}$	$\hat{\tau}_{ps}$
$s_1=(P_1, P_2)$	6.34	304.18
$s_2=(P_1, P_3)$	7.48	359.16
$s_3=(P_1, P_4)$	8.10	388.86
$s_4=(P_2, P_3)$	5.59	268.16
$s_5=(P_2, P_4)$	6.96	334.13
$s_6=(P_3, P_4)$	6.05	290.22
Mean	10.56	264
Bias	0	0

CHAPTER 4

COMPARSION OF THE SAMPLING DESIGNS

In this chapter, path sampling will be compared to cluster sampling, SRSWOR, and random walk sampling. Details of these sampling designs were discussed in chapter 2. Cluster sampling is used in practice because it is usually much cheaper and more convenient to sample in a cluster than randomly in a population; also, it is cost saving. Path sampling is a new sampling design proposed with the objective of being cost effective and convenient for sampling travel-it is more convenient and cost effective to sample and travel along the paths than randomly in the population. Thus, to investigate the efficiency of path sampling, it is compared to cluster sampling and SRSWOR in the present study. For the random walk design in a spatial setting, each unit links to the adjacent units, as shown in Figure 4.1 (a). Since sampling travel in random walk sampling for a spatial setting is from the initial unit to another adjacent unit until the last unit in the sample, which is like a route, as shown figure 4.1 (b), path sampling is compared to random walk sampling.

(a) Population and Links



(b) Random Walk Sample

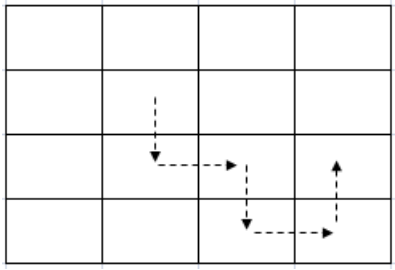


Figure 4.1 Random Walk Sampling

4.1 Simulation Study

Rare and non-rare population data are used in a simulation to examine the performance of path sampling compared to comparable sampling design, which in this research are SRSWOR, cluster sampling, and random walk sampling**.** For path sampling, SRSWOR, and cluster sampling, the estimator of the mean is unbiased, but for random walk sampling it is biased. It is known that $MSE(\hat{\mu}) = V(\hat{\mu}) + (bias(\hat{\mu}))^2$. If $\hat{\mu}$ is unbiased, then $MSE(\hat{\mu}) = V(\hat{\mu})$ since $bias(\hat{\mu}) = 0$. Thus, for a simulation of 1000 iterations, the formula used to estimate the MSE for four sampling designs is

$$\hat{MSE}(\hat{\mu}) = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\mu}_i - \mu)^2 \qquad (4.1)$$

4.1.1 Simulation Study for Rare Population

Blue-winged teal data are used (Smith et al, 1995: 777-778) in Figure 4.2. for part of the simulation study, as it is a rare population; that is, most of the units have zero y-value.

In cluster sampling, let a cluster be an entire column, consisting of 10 secondary units, as shown in Figure 4.3. It is a cluster of size 10. The expected sample size of path sampling will be denoted as $E(\upsilon)$, and the sample size used in the other designs is set equal the ceiling of the $E(\upsilon)$ for path sampling. For cluster sampling, the number of clusters samples is set equal to the ceiling of $\frac{E(\upsilon)}{10}$ for comparison purposes. Let m_c be the number of secondary units in a cluster sample. Let the C.V. among clusters be the coefficient of variation of cluster totals. Let the total of cluster *i* be $y_{i.}$. The formula of C.V. among clusters is $\frac{\sqrt{\mathrm{v}(y_{i.})}}{mean(y_{i.})}$. In this population data, the C.V. among clusters is 4.26.

In SRSWOR, the sample size is set equal to $E(\upsilon)$ in order to compare it to path sampling. The sampling unit is each unit (*i, j*). For random walk sampling, the

link is defined as a link of adjacent units, and the number of waves for random walk sampling is set equal to $E(\upsilon)$ for comparison purposes.

The results from the simulations are shown in Table 4.1. According to these results, for starting units (1, 1) and (1, 10), path sampling is more efficient than cluster sampling since the relative efficiency is greater than 1. Noticeably, the y-values in columns 17, 18 and 19 are much higher than the others, so there is high variation among clusters in this population. This makes cluster sampling less efficient. However, path sampling is less efficient than SRSWOR since the relative efficiency is less than 1. Notice that when the starting unit is in a high–value column, as in unit (1,17), path sampling is more efficient than SRSWOR since the relative efficiency is greater than 1, and much more efficient than cluster sampling since the relative efficiency is greater than 4. Moreover, for any p and any starting unit, path sampling is more efficient than random walk sampling since the relative efficiency is greater than 1.

4.1.2 Simulation Study for a Non-Rare Population

Two simulated data are considered. First, the simulated data in Figure 4.4 are used. Each unit is Poisson distributed with a mean of 50. To compare path sampling to cluster sampling, let a cluster be an entire column. In this population, the C.V. among clusters is 0.04. The simulation results are shown in Table 4.2.

According to the simulation results in Table 4.2, path sampling is less efficient than cluster sampling, SRSWOR, and random walk sampling because the relative efficiency is less than 1. Noticeably, there is small variation of y-values, so there is a low variation among clusters (C.V. among clusters is 0.04) in this population. This makes cluster sampling more efficient.

Next, simulated data are used, as shown in Figure 4.5. All units are the same as the population data in Figure 4.4, except column 6, 10, and 15. The y-values in these 3 columns are replaced with higher values. To compare path sampling to cluster sampling, let a cluster be an entire column. This population data have high variation among clusters, with a C.V among clusters at 1.46. The simulation results are shown in Table 4.3.

0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	20	4	2	12	0	0	0	0	0	10	103	0	0	0
0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	150	7144	1	0
0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	6	6339	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	122	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	114	60
0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	3	0

Figure 4.2 Blue-winged Teal Data With C.V. among Clusters 4.26

cls1	cls2	cls3	cls4	cls5	cls6	cls7	cls8	cls9	cls10	cls11	cls12	cls13	cls14	cls15	cls16	cls17	cls18	cls19	cls20
0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	20	4	2	12	0	0	0	0	0	10	103	0	0	0
0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	150	7144	1	0
0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	6	6339	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	122	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	114	60
0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	3	0

Figure 4.3 Clusters in Blue-winged Teal Data

43	51	40	40	55	56	49	43	61	61	49	39	38	42	53	61	57	50	48	47
47	51	51	53	51	43	60	55	50	60	61	49	56	50	57	55	59	49	40	47
60	52	46	49	54	51	58	45	48	44	43	54	61	50	63	50	57	45	47	50
55	56	51	51	47	38	55	50	54	51	61	42	48	35	50	41	67	48	47	48
49	55	55	44	39	61	47	54	60	55	67	43	41	50	52	55	44	45	54	54
60	50	49	46	57	49	57	49	44	56	37	44	47	47	46	48	46	42	29	52
53	57	35	52	43	51	49	65	54	51	55	52	55	68	39	44	39	48	68	56
55	51	56	34	50	57	49	58	52	64	41	49	47	61	52	50	55	57	41	47
51	62	41	45	41	55	43	51	46	33	49	54	56	41	51	46	61	55	43	35
47	47	41	46	47	56	47	43	61	44	43	59	39	52	46	37	48	59	49	61

Figure 4.4 Simulated Data: Each Unit is Poisson Distributed with a Mean of 50 with C.V. among Clusters of 0.04

43	51	40	40	55	556	49	43	61	1610	49	39	38	42	553	61	57	50	48	47
47	51	51	53	51	643	60	55	50	55	61	49	56	50	657	55	59	49	40	47
60	52	46	49	54	651	58	45	48	1404	43	54	61	50	563	50	57	45	47	50
55	56	51	51	47	638	55	50	54	55	61	42	48	35	689	41	67	48	47	48
49	55	55	44	39	561	47	54	60	67	67	43	41	50	552	55	44	45	54	54
60	50	49	46	57	665	57	49	44	155	37	44	47	47	546	48	46	42	29	52
53	57	35	52	43	551	49	65	54	1501	55	52	55	68	639	44	39	48	68	56
55	51	56	34	50	457	49	58	52	64	41	49	47	61	457	50	55	57	41	47
51	62	41	45	41	555	43	51	46	33	49	54	56	41	551	46	61	55	43	35
47	47	41	46	47	356	47	43	61	133	43	59	39	52	446	37	48	59	49	61

Figure 4.5 Simulated Data: Poisson Distributed with a Mean of 50 and Change 3 Columns with High Value with C.V. among Clusters of 1.46

Table 4.1 Results from the Simulations on Blue-winged Teal Data

p	$E(v)$	m_c	$\hat{MSE}(\hat{\mu}_{ps})$			$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	(1,1)			(1, 10)		
			(1,1)	(1, 10)	(1,17)*				R.E.cls	R.E.srs	R.E. rws	R.E.cls	R.E.srs	R.E. rws
1	48	50 (5)	10389.35	11235.81	2728.62	13719.06	7445.01	36491.29	1.32	0.72	3.51	1.22	0.66	3.25
2	83.33	90 (9)	4147.10	4385.75	452.91	5520.17	3256.30	25471.09	1.33	0.79	6.14	1.26	0.74	5.81
3	113	120 (12)	2124.90	2092.34	106.04	3038.16	1774.36	22842.35	1.43	0.84	10.75	1.45	0.85	10.92
4	138	140(14)	1043.16	1017.37	10.45	2010.08	1024.95	17174.65	1.93	0.98	16.46	1.98	1.01	16.88
5	158.6	160 (16)	496.88	557.02	0.27	1124.91	628.68	17189.50	2.26	1.27	34.59	2.02	1.13	30.86

Table 4.1 (Continued)

p	$E(v)$	m_c	$\hat{MSE}(\hat{\mu}_{ps})$			$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	(1,17)*		
			(1,1)	(1, 10)	(1,17)*				R.E.cls	R.E.srs	R.E. rws
1	48	50 (5)	10389.35	11235.81	2728.62	13719.06	7445.01	36491.29	5.03	2.73	13.37
2	83.33	90 (9)	4147.10	4385.75	452.91	5520.17	3256.30	25471.09	12.19	7.19	56.24
3	113	120 (12)	2124.90	2092.34	106.04	3038.16	1774.36	22842.35	28.65	16.73	215.42
4	138	140(14)	1043.16	1017.37	10.45	2010.08	1024.95	17174.65	192.29	98.05	1642.99
5	158.6	160 (16)	496.88	557.02	0.27	1124.91	628.68	17189.50	4233.67	2366.08	64693.61

Note: The number in parentheses is the number of clusters selected in cluster sampling.

* means that such a starting unit is on a high y-value column j* or has high y-value column j*+1

R.E.cls = $\hat{MSE}(\hat{\mu}_{cls}) / \hat{MSE}(\hat{\mu}_{ps})$

R.E.srs = $\hat{MSE}(\hat{\mu}_{srs}) / \hat{MSE}(\hat{\mu}_{ps})$

R.E.rws = $\hat{MSE}(\hat{\mu}_{rws}) / \hat{MSE}(\hat{\mu}_{ps})$

Table 4.2 Results from the Simulation on Non-rare Population with Low C.V. among Clusters

p	$E(\nu)$	m_c	$\hat{MSE}(\hat{\mu}_{ps})$		$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	(1,10)			(1, 17)		
			(1,10)	(1, 17)				R.E. cls	R.E. srs	R.E. rws	R.E. cls	R.E. srs	R.E. rws
1	48	50 (5)	92.46	94.72	0.80	0.86	2.71	0.0086	0.0093	0.0293	0.0084	0.0091	0.0286
2	83.33	90 (9)	71.43	65.38	0.31	0.39	1.71	0.0043	0.0055	0.0239	0.0047	0.0060	0.0262
3	113	120 (12)	51.82	55.73	0.15	0.21	1.43	0.0030	0.0041	0.0276	0.0028	0.0038	0.0257
4	138	140(14)	40.21	39.49	0.11	0.12	1.13	0.0027	0.0031	0.0281	0.0027	0.0032	0.0286
5	158.6	160 (16)	30.55	30.63	0.06	0.07	0.99	0.0021	0.0023	0.0324	0.0021	0.0023	0.0323

According to the simulation results in Table 4.3, for starting units (1, 2) and (1, 17), path sampling is more efficient than cluster sampling because the relative efficiency is greater than 1. Noticeably, the y-values in column 6, 10 and 15 are much higher than others, so there is high variation among clusters (C.V. of 1.46) in this population. This makes cluster sampling less efficient. Notice that when the starting unit is in a high-value column, as in units (1, 5), (1, 10), and (1, 15), path sampling is much more efficient than cluster sampling since the relative efficiency is greater than 2.

For starting units (1, 2) and (1, 17), path sampling is less efficient than SRSWOR since the relative efficiency is less than 1 for any p. However, for the starting unit in a high-value column, as with units (1, 5), (1, 10), and (1, 15), path sampling is more efficient than SRSWOR for $p = 1$ since the relative efficiency is greater than 1; however, it is less efficient than SRSWOR for $p > 2$ because the relative efficiency is less than 1. Moreover, path sampling is more efficient than random walk sampling since the relative efficiency is greater than 1 for any p and all starting units.

4.1.3 Simulation Results Summary

In this simulation, for a rare population and a non-rare population with high variation of y-values among clusters, path sampling is more efficient than random walk sampling. With the starting or ending point on high y-value column for this kind of population, path sampling is more efficient than cluster sampling but less efficient than SRSWOR. However, for a non-rare population with low variation of y-values among clusters, path sampling is less efficient than cluster sampling, SRSWOR, or random walk sampling.

Now we want to investigate the value C.V. among clusters that makes path sampling more efficient than cluster sampling, SRSWOR, and random walk sampling. According to the simulation on the population data with C.V. among clusters from 0.3 to 2.4 and the path starting or ending point on a high y-value column, the results are shown in Figure 4.6. and appendix A.

Table 4.3 Results from Simulation on Non-rare Population with High C.V. among Clusters

p	$E(\upsilon)$	m_c	$\hat{MSE}(\hat{\mu}_{ps})$					$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	(1,2)			(1,5)*			(1, 10)*		
			(1,2)	(1,5)*	(1, 10)*	(1, 15)*	(1,17)				R.E.cls	R.E.srs	R.E.rws	R.E.cls	R.E.srs	R.E.rws	R.E.cls	R.E.srs	R.E.rws
1	48	50 (5)	1025.19	803.89	275.34	795.22	1084.99	4708.79	880.05	4564.87	4.59	0.86	4.45	5.86	1.09	5.68	17.10	3.20	16.58
2	83.33	90 (9)	657.02	586.39	297.47	567.97	671.26	1960.08	387.34	2786.88	2.98	0.59	4.24	3.34	0.66	4.75	6.59	1.30	9.37
3	113	120 (12)	478.45	354.00	239.42	361.43	460.14	1114.89	206.29	2297.49	2.33	0.43	4.80	3.15	0.58	6.49	4.66	0.86	9.60
4	138	140(14)	368.14	264.31	160.38	262.70	320.87	700.68	115.73	1853.33	1.90	0.31	5.03	2.65	0.44	7.01	4.37	0.72	11.56
5	158.6	160 (16)	266.58	198.86	130.28	207.81	271.03	413.80	70.07	1710.16	1.55	0.26	6.42	2.08	0.35	8.60	3.18	0.54	13.13

Table 4.3 (Continued)

p	$E(\upsilon)$	m_c	$\hat{MSE}(\hat{\mu}_{ps})$					$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	(1,15) *			(1,17)		
			(1,2)	(1,5)*	(1, 10)*	(1, 15)*	(1,17)				R.E.cls	R.E.srs	R.E.rws	R.E.cls	R.E.srs	R.E.rws
1	48	50 (5)	1025.19	803.89	275.34	795.22	1084.99	4708.79	880.05	4564.87	5.92	1.11	5.74	4.34	0.81	4.21
2	83.33	90 (9)	657.02	586.39	297.47	567.97	671.26	1960.08	387.34	2786.88	3.45	0.68	4.91	2.92	0.58	4.15
3	113	120 (12)	478.45	354.00	239.42	361.43	460.14	1114.89	206.29	2297.49	3.08	0.57	6.36	2.42	0.45	4.99
4	138	140(14)	368.14	264.31	160.38	262.70	320.87	700.68	115.73	1853.33	2.67	0.44	7.05	2.18	0.36	5.78
5	158.6	160 (16)	266.58	198.86	130.28	207.81	271.03	413.80	70.07	1710.16	1.99	0.34	8.23	1.53	0.26	6.31

From this simulation, with the path starting or ending point on high a y-value column, when the C.V. among clusters is greater than 1, path sampling is more efficient than cluster sampling. When the C.V. among clusters is greater than 0.5, path sampling is more efficient than random walk sampling. When the C.V. among clusters is greater than or equal to 2.0, path sampling is more efficient than SRSWOR.

4.1.4 Efficiency of Path Sampling

4.1.4.1 The number of sample paths (p)

From the simulation, it can be seen that the greater the number of sample paths the more efficient is path sampling.

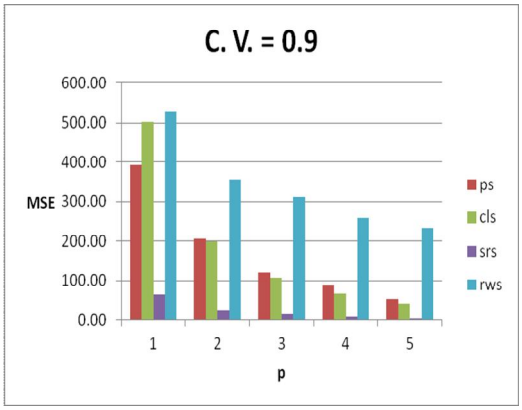
4.1.4.2 The starting and ending point

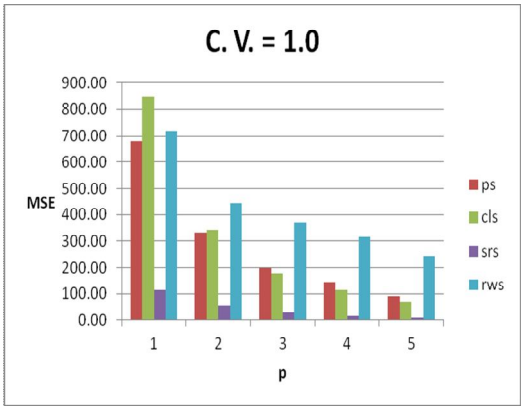
From the simulation, we notice that if the starting or ending point is on high y-value column, then path sampling is more efficient. This can be explained by the following. The Horvitz-Thompson estimator, used in path sampling, is more efficient when inclusion probabilities are proportional to the y-value (Horvitz and Thompson, 1952: 663-685). Under path sampling, all units in the column with the starting unit, say starting column j^*, have higher inclusion probabilities than those of units in other columns because in every path, column j^* is the way out from the starting unit to observe data, while column j^*+1 is the way back to the starting unit. Thus, the units in column j^* and j^*+1 have high probabilities to be included in the sample. Therefore, if column j^* or j^*+1 have high y-value, then path sampling is more efficient. If a researcher could set a starting point in a high y-value column, path sampling would be more efficient. Since y-value is unknown, an auxiliary variable can be used to identify the high-value column.

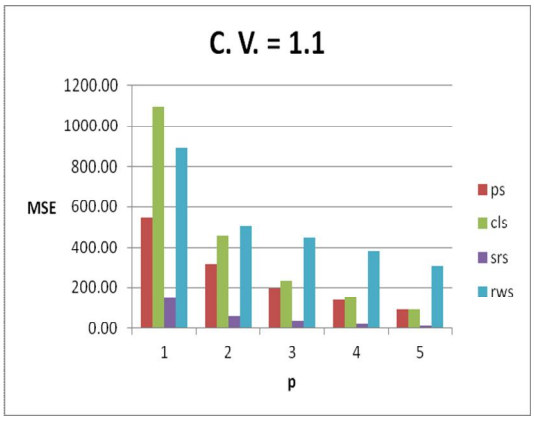
4.1.4.3 The population data

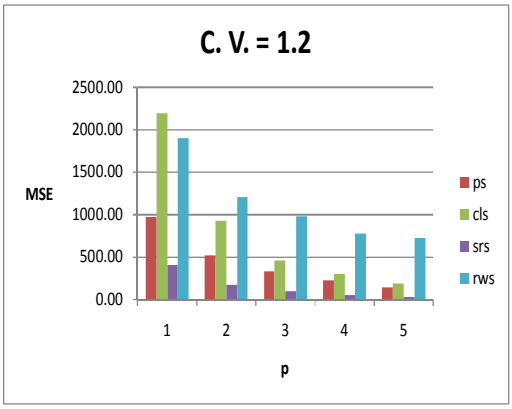
According to the simulation of the non-rare population with low variation of y-value among clusters, path sampling is not efficient, while in the rare and non-rare population data with high variation of y-value among cluster with the starting or ending point on high y-values, path sampling is efficient.

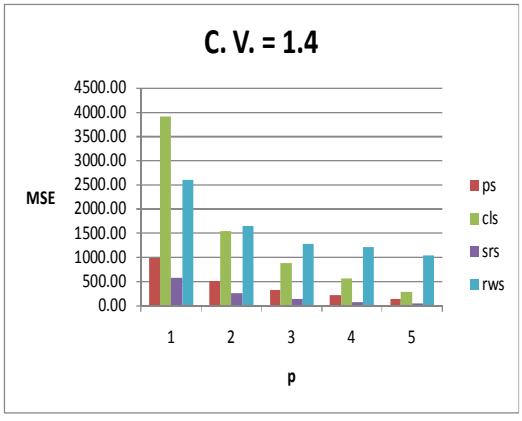
Figure 4.6 The Estimated MSE of Path Sampling, Cluster Sampling, SRSWOR, and Random Walk Sampling for Different C.V. among Clusters

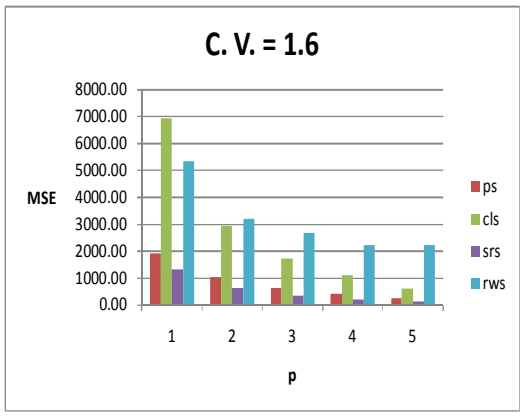

C. V. = 0.9
600.00
500.00
400.00
300.00
200.00
100.00
0.00
MSE
1
2
3
4
5
p
ps
cls
srs
rws

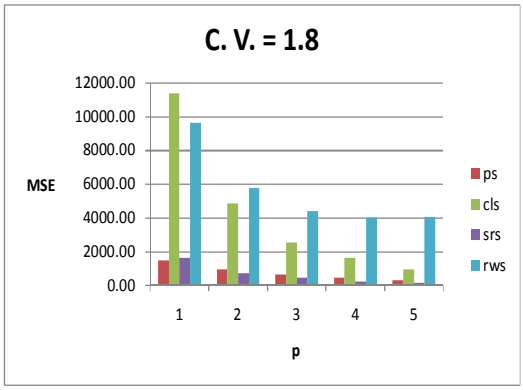

C. V. = 1.0
900.00
800.00
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
MSE
1
2
3
4
5
p
ps
cls
srs
rws

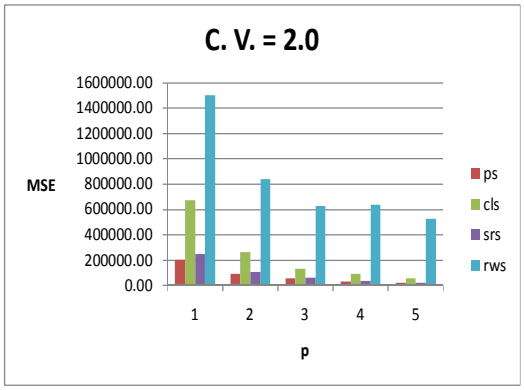

C. V. = 1.1
1200.00
1000.00
800.00
600.00
400.00
200.00
0.00
MSE
1
2
3
4
5
p
ps
cls
srs
rws

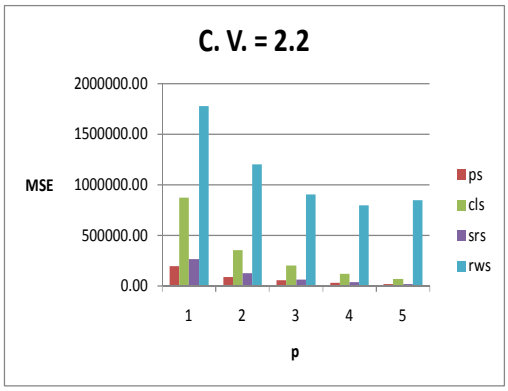

C. V. = 1.2
2500.00
2000.00
1500.00
1000.00
500.00
0.00
MSE
1
2
3
4
5
p
ps
cls
srs
rws

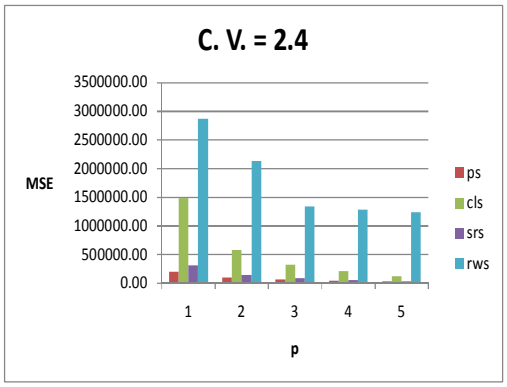

C. V. = 1.4
4500.00
4000.00
3500.00
3000.00
2500.00
2000.00
1500.00
1000.00
500.00
0.00
MSE
1
2
3
4
5
p
ps
cls
srs
rws


C. V. = 1.6
8000.00
7000.00
6000.00
5000.00
4000.00
3000.00
2000.00
1000.00
0.00
MSE
1
2
3
4
5
p
ps
cls
srs
rws









4.2 Comment on the Case $p = 1$

An unbiased estimator of variance $v(\hat{\tau}_{HT})$ is $\hat{v}(\hat{\tau}_{HT})$ if all of the joint inclusion probabilities are greater than zero (Horvitz and Thompson, 1952: 670). In path sampling, there exist zero joint inclusion probabilities when the number of sampled paths is set equal to 1; that is $p = 1$. To investigate the biasedness of the estimator of variance in this case, 10 simulated population data were studied. The results are shown in Table 4.4.

From the simulation results in Table 4.4, when $p = 1$, it can be seen that there exist zero joint inclusion probabilities, except for a population of two columns. Thus, $\hat{v}(\hat{\mu}_{ps})$ is a biased estimator for variance when p=1, except for a population of two columns.

Table 4.4 Unbiasedness of $\hat{v}(\hat{\mu}_{ps})$ When p =1

population	r	c	q	p	All $\pi_{(i,j)}$>0	Unbiasedness of $\hat{v}(\hat{\mu}_{ps})$
1	3	2	2	1	yes	ue
2	4	2	3	1	yes	ue
				2	yes	ue
3	5	2	4	1	yes	ue
				2	yes	ue
				3	yes	ue
4	3	3	2	1	no	be
5	4	3	3	1	no	be
				2	yes	ue
6	5	3	4	1	no	be
				2	yes	ue
				3	yes	ue
7	6	3	5	1	no	be
				2	yes	ue
				3	yes	ue
				4	yes	ue
8	7	3	6	1	no	be
				2	yes	ue
				3	yes	ue
				4	yes	ue
9	8	6	7	1	no	be
				2	yes	ue
				3	yes	ue
				4	yes	ue
10	10	20	9	1	no	be
				2	yes	ue
				3	yes	ue
				4	yes	ue

Note: r is the number of population columns.
c is the number of population rows.
q is the number of population paths
p is the number of paths selected.
ue means that $\hat{v}(\hat{\mu}_{ps})$ is unbiased.
be means that $\hat{v}(\hat{\mu}_{ps})$ is biased.

4.3 Comparison of Cost of Path Sampling, SRSWOR, Cluster Sampling, and Random Walk Sampling

We consider cost as a function of the distance traveled by counting the units traveled to observe all units in a sample. In other words, we consider the number of units traveled.

Let d_p, d_s, d_c and d_r be the number of units traveled to observe all units in a sample under path sampling, SRSWOR, cluster sampling, and random walk sampling, respectively.

4.3.1 Comparison of the Cost of Path Sampling to SRSWOR

When the samples from the two sampling designs for the sample same size are obtained, we want to know which one is more cost effective. Since a sample size of path sampling varies from sample to sample, the sample size under SRSWOR will be set equal to the expected sample size under path sampling for comparison purposes.

Let $I_{(i,j)}$ take the value 1 when a unit (*i, j*) is included in the sample, and 0 otherwise. The expectation of $I_{(i,j)}$ is $E(I_{(i,j)}) = P(I_{(i,j)} = 1)$ .The number of distinct units, υ, is a random variable; namely,

$$\upsilon = \sum_{i=1}^{r}\sum_{j=1}^{c} I_{(i,j)} \tag{4.3}$$

If $\pi_{(i,j)}$ is the probability that unit (*i, j*) is included in the sample, then $\pi_{(i,j)} = P(I_{(i,j)} = 1) = E(I_{(i,j)})$. The expected sample size under path sampling is

$$E(\upsilon) = \sum_{i=1}^{r}\sum_{j=1}^{c} \pi_{(i,j)} \tag{4.4}$$

(Thompson and Seber, 1996).

The number of units traveled to observe all units in a sample under path sampling is the expected sample size. It can be written as

$$E(d_p) = E(\upsilon). \tag{4.5}$$

On the other hand, a simple random sample without replacement of size $E(\upsilon)$ is taken. To observe all units in a simple random sample, other units not in the sample must be traveled.

Let c_s be the number of additional units, not included in the sample, traveled to observe all $E(\upsilon)$ units in the simple random sample. c_s is always greater than or equal to zero, so $E(c_s) \geq 0$. Thus, the number of units traveled to observe all units in a sample under SRSWOR is

$$d_s = E(\upsilon) + c_s \tag{4.6}$$

Then, its expectation is

$$\begin{aligned} E(d_s) &= E(\upsilon) + E(c_s) \\ &= E(d_p) + E(c_s) \text{ since } E(d_p) = E(\upsilon) \\ &\geq E(d_p) \text{ since } E(c_s) \geq 0 \end{aligned}$$

$E(d_s) = E(d_p)$ if $c_s = 0$. Moreover, $c_s = 0$ when the sample size is equal to the population size. $E(d_s) > E(d_p)$. If $c_s > 0$. Moreover, $c_s > 0$ when the sample size is less than the population size.

When the sample size is less than the population size, for the same sample size, path sampling yields the smaller number of units traveled than SRSWOR. Thus, path sampling is more cost effective.

4.3.2 Comparison Cost of Path Sampling to Cluster Sampling

Suppose that a population of r rows and c columns consists of N_c clusters. Each cluster consists of M units. That is, $MN_c = rc$. A cluster sample is a simple random sample of n_c clusters.

To compare the cost of path sampling to cluster sampling, a final sample size under cluster sampling will be set equal to the expected sample size under path sampling for comparison purposes. That is, $n_c M = E(\upsilon)$.

For the same sample size, which sampling design is more cost effective? To observe all of the units in a cluster sample, the units outside a sample must be traveled. Let c_c be the number of additional units, not included in the sample, traveled to observe all units in a cluster sample. c_c is always greater than or equal to zero. So, $E(c_c) \geq 0$. Thus, the number of units traveled to observe all units in a cluster sample is

$$d_c = E(\upsilon) + c_c \tag{4.7}$$

and the expectation number is

$$\begin{aligned} E(d_c) &= E(\upsilon) + E(c_c) \\ &= E(d_p) + E(c_c) \text{ since } E(d_p) = E(\upsilon) \\ &\geq E(d_p) \text{ since } E(c_c) \geq 0 \end{aligned}$$

$E(d_c) = E(d_p)$ if $c_c = 0$. Moreover, $c_c = 0$ when the sample size is equal to the population size. $E(d_c) > E(d_p)$. If $c_c > 0$. Moreover, $c_c > 0$ when the sample size is less than the population size.

When the sample size is less than the population size, for the same sample size, path sampling yields the smaller number of units traveled than cluster sampling. Thus, path sampling is more cost effective.

4.3.3 Comparison of the Cost of Path Sampling to Random Walk Sampling

The sample size under random walk sampling will be set equal to the expected sample size under path sampling, $E(\upsilon)$, for comparison purposes. The initial unit is selected by simple random sampling. To continue a random walk in each wave, the next unit is randomly selected from the adjacent units of the current unit until the $E(\upsilon)^{\text{th}}$ wave is reached. Finally, the random walk sample of size $E(\upsilon)$ is obtained.

To observe all units in the random walk sample, other units not included in the sample must be traveled. The starting point for traveling is the unit on edge of the rectangular region nearest the initial unit of random walk sample. Ending point is the unit on the edge of the region nearest the unit in the last wave. Let c_r be the number of additional units, not included in the sample, traveled to observe all $E(\upsilon)$ units in the random walk sample. c_r is always greater than or equal to zero, so $E(c_r) \geq 0$. Thus, the number of units traveled to observe all units in a sample under the random walk sample is

$$d_r = E(\upsilon) + c_r \tag{4.8}$$

Then, its expectation is

$$\begin{aligned} E(d_r) &= E(\upsilon) + E(c_r) \\ &= E(d_p) + E(c_r) \text{ since } E(d_p) = E(\upsilon) \\ &\geq E(d_p) \text{ since } E(c_r) \geq 0 \end{aligned}$$

$E(c_r) = E(d_p)$ if $c_r = 0$. Moreover, $c_r = 0$ when the sample size is equal to the population size. $E(c_r) > E(d_p)$. If $c_r > 0$. Moreover, $c_r > 0$ when the sample size is less than the population size.

When the sample size is less than the population size, for the same sample size, path sampling yields a smaller number of units traveled than random walk sampling. Thus, path sampling is more cost effective.

4.4 Cost Simulation

The simulation consists of 1000 iterations by using R program to examine the efficiency of path sampling, SRSWOR, cluster sampling, and random walk sampling. Visual Basic program is used to examine the number of units traveled for the cluster and simple random sample.

To find the number of units traveled in a simple random sample and cluster sample, the starting unit is set to be the sampled unit nearest the edge of the region. Then, all possible traveling routes of all units in a sample are created. For each traveling route, the number of units traveled is counted. The route with minimum number of units traveled will be used for comparison to path sampling.

To find the number of units traveled in a random walk sample, the starting point for traveling is the unit on the edge of the rectangular region nearest the initial unit of the random walk sample. The ending point is the unit on the edge of the region nearest the unit in the last wave.

The Longleaf Pin data in Figure 4.7, with a total abundance of 584, are used in the simulation. They consists of 400 secondary units and 100 primary units (clusters) of size 4. The simulation results are shown in Table 4.5 For $p = 1$ in path sampling, the expected number of units traveled for cluster sampling is 1.762 times the expected number of units traveled for path sampling. That is, the expected number of units traveled is 76% higher for cluster sampling. Also, the expected number of units traveled for SRSWOR is 4.373 times the expected number of units traveled for path sampling. The expected number of units traveled for random walk sampling is 1.091 times the expected number of units traveled for path sampling. From the table 4.5, we can see that among the four sampling designs, path sampling has the smallest value of the expected number of units traveled for any p.

However, for any p, the estimated variances of path sampling are greater than those of cluster sampling and SRSWOR, so path sampling is less efficient than cluster sampling and SRSWOR. Notice that in this population, there is a low variation of y-values among clusters (C.V. among clusters of 0.77), so cluster sampling is more

efficient. For this population, path sampling is more efficient than random walk sampling since path sampling gives smaller estimated variances (MSE) than random walk sampling.

Consider the estimated variance multiplied by the expected number of units traveled. If this value is small, this means that the sampling design is more efficient and less distance is travelled. According to table 4.5, path sampling gives a slightly larger value of multiplication than cluster sampling and SRSWOR but smaller than random walk sampling.

Note that the Horvitz-Thompson estimator is efficient when the y-value is proportional to an inclusion probability. In path sampling, the units in starting and ending column have high inclusion probabilities due to the overlapping of paths. Therefore, path sampling is more efficient when starting and ending columns have high y-value, as can be seen in the following simulation for the modified Longleaf Pin data.

Next, the modified Longleaf Pin data shown in Figure 4.8 are used for simulation. This population has a higher y-value in column 10 than the original Longleaf Pin data. Moreover, these population data have high variation of y-values among clusters (C.V. among clusters of 2.02). The results of the simulation with starting units (1, 10), which is in the high y-value column, are shown in Table 4.6.

According to the results in Table 4.6, the starting and ending columns have high y-value and high inclusion probabilities. Thus, for any p, the estimated variance for path sampling is smaller than cluster sampling and SRSWOR, and the relative efficiencies are greater than 1. Hence, path sampling is more efficient than cluster sampling and SRSWOR. Random walk sampling has a larger MSE than path sampling, so path sampling is more efficient than random walk sampling for this population.

For $p = 1$ in path sampling, the expected number of units traveled for cluster sampling is 1.779 times the expected number of units traveled for path sampling. Also, the expected number of units traveled for SRSWOR is 4.479 times the expected number of units traveled for path sampling. The expected number of units traveled for random walk sampling is 1.081 times the expected number of units traveled for path

sampling. From the table we can see that among the four sampling designs, path sampling has the smallest value of expected number of units traveled for any p.

Consider the estimated variance multiplied by the expected number of units traveled. If this value is small, this means that the sampling design is more efficient and less distance is travelled. From table 4.6, path sampling gives the smallest value among the four sampling designs.

Summary Results of Cost Simulation is following. From the simulation results, it can be seen that among the four sampling designs, path sampling has the smallest value of the expected number of units traveled for the same sample size. Thus, path sampling yields less distance travelled. When the main cost of sampling is the number of units traveled, path sampling saves cost. However, path sampling is less efficient than cluster sampling and SRSWOR in a population data with low variation of y-values among clusters (the Longleaf Pin data with C.V. among clusters of 0.77). On the other hand, in the case of population data with high variation of y-values among clusters (the modified Longleaf Pin data with C.V. among clusters of 2.02) with the path starting or ending on a high y-value column, path sampling is more efficient than cluster sampling and SRSWOR.

Compared to random walk sampling, path sampling is more efficient for both original and modified Longleaf Pin data. The expected number of units traveled under random walk sampling is slightly greater than that of path sampling.

1	1	1	1	1	2	1	0	0	0	4	5	0	1	0	1	2	1	0	1
3	2	1	0	1	0	0	0	1	2	2	2	0	2	2	2	0	2	0	1
7	4	1	1	1	1	0	0	0	2	2	0	4	3	2	4	2	1	2	2
0	1	2	0	0	0	0	0	4	6	5	1	5	0	0	0	2	1	2	0
1	1	0	2	3	2	0	0	2	1	3	1	4	1	1	1	2	2	1	1
2	0	0	0	4	3	3	0	1	16	5	0	1	3	8	0	0	1	3	3
0	0	1	14	3	3	1	2	0	8	0	2	0	3	9	0	4	2	1	0
0	0	5	1	8	7	6	6	6	1	0	4	0	0	1	2	2	0	1	2
0	0	2	2	3	2	2	3	1	1	1	3	0	0	2	2	0	3	4	0
0	0	0	0	1	0	3	1	1	1	2	0	2	0	2	0	2	1	1	0
1	8	7	7	8	0	5	0	1	0	1	2	0	0	2	4	2	2	2	4
0	9	1	0	0	1	1	1	0	0	0	1	2	4	0	2	1	3	3	1
0	0	0	1	0	2	4	3	1	2	2	0	0	1	1	2	2	0	2	4
0	1	0	0	1	2	0	2	3	5	2	0	0	2	1	1	2	0	1	3
1	0	0	1	1	0	0	0	2	2	2	1	1	1	0	0	2	0	0	0
0	2	0	2	2	0	1	1	0	2	0	0	1	0	0	1	1	1	5	3
0	0	0	3	2	1	0	0	0	0	0	2	1	0	1	1	1	3	1	2
1	0	0	1	0	3	0	1	0	0	2	1	2	0	0	0	1	1	1	0
0	0	0	0	0	0	0	1	1	1	0	1	0	3	0	2	0	1	1	0
2	0	0	0	0	0	0	0	1	2	0	1	3	0	0	1	0	1	2	4

Figure 4.7 The Longleaf Pin Data with 100 Clusters of Size 4 and C.V. among Clusters of 0.77

1	1	1	1	1	2	1	0	0	0	4	5	0	1	0	1	2	1	0	1
3	2	1	0	1	0	0	0	1	2	2	2	0	2	2	2	0	2	0	1
7	4	1	1	1	1	0	0	0	2	2	0	4	3	2	4	2	1	2	2
0	1	2	0	0	0	0	0	4	6	5	1	5	0	0	0	2	1	2	0
1	1	0	2	3	2	0	0	2	1	3	1	4	1	1	1	2	2	1	1
2	0	0	0	4	3	3	0	1	106	5	0	1	3	8	0	0	1	3	3
0	0	1	14	3	3	1	2	0	8	0	2	0	3	9	0	4	2	1	0
0	0	5	1	8	7	6	6	6	1	0	4	0	0	1	2	2	0	1	2
0	0	2	2	3	2	2	3	1	1	1	3	0	0	2	2	0	3	4	0
0	0	0	0	1	0	3	1	1	1	2	0	2	0	2	0	2	1	1	0
1	8	7	7	8	0	5	0	1	0	1	2	0	0	2	4	2	2	2	4
0	9	1	0	0	1	1	1	0	0	90	1	2	4	0	2	1	3	3	1
0	0	0	1	0	2	4	3	1	2	2	0	0	1	1	2	2	0	2	4
0	1	0	0	1	2	0	2	3	5	2	0	0	2	1	1	2	0	1	3
1	0	0	1	1	0	0	0	2	2	2	1	1	1	0	0	2	0	0	0
0	2	0	2	2	0	1	1	0	82	0	0	1	0	0	1	1	1	5	3
0	0	0	3	2	1	0	0	0	0	0	2	1	0	1	1	1	3	1	2
1	0	0	1	0	3	0	1	0	0	2	1	2	0	0	0	1	1	1	0
0	0	0	0	0	0	0	1	1	1	0	1	0	3	0	2	0	1	1	0
2	0	0	0	0	0	0	0	1	2	0	1	3	0	0	1	0	1	2	4

Figure 4.8 The Modified Longleaf Pin Data with 100 Clusters of Size 4 and C.V. among Clusters of 2.02

Table 4.5 Estimated Variance and MSE and Expected Number of Units Traveled Under Path Sampling, SRSWOR, Cluster Sampling (Cluster size = 4) and Random Walk Sampling for Longleaf Pin Data with a Simulation of 1000 Iterations (with Starting Unit (1, 10))

p	Expected sample size			$\hat{MSE}(\hat{\mu}_{ps})$	$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	R.E.cls	R.E.srs	R.E.rws
	PS	Cls	SRS							
1	58	56 (14)	58	0.188	0.074	0.057	0.456	0.394	0.300	2.421
2	98.77	100(25)	99	0.100	0.037	0.029	0.310	0.372	0.293	3.115
3	134.31	136(34)	135	0.066	0.027	0.019	0.299	0.402	0.294	4.519
4	166.63	168(42)	167	0.048	0.018	0.013	0.238	0.368	0.271	4.917
5	196.38	196(49)	196	0.039	0.013	0.010	0.200	0.333	0.249	5.102
6	223.86	224(56)	224	0.029	0.010	0.007	0.189	0.338	0.241	6.408
7	249.21	248(62)	249	0.024	0.008	0.006	0.177	0.334	0.236	7.273

Table 4.5 (Continued)

p	Expected sample size			Expected number of units traveled				R.D.cls	R.D.srs	R.D.rws	$\hat{MSE}$ * Expected number of units traveled			
	PS	Cls	SRS	PS	Cls	SRS	RWS				PS	Cls	SRS	RWS
1	58	56 (14)	58	58.4	102.9	255.4	63.7	1.762	4.373	1.091	10.998	7.638	14.431	29.047
2	98.77	100(25)	99	97.7	156.4	327.9	104.9	1.601	3.356	1.074	9.724	5.790	9.551	32.519
3	134.31	136(34)	135	135.3	191.3	350.2	140.8	1.414	2.588	1.041	8.952	5.094	6.819	42.099
4	166.63	168(42)	167	166.8	216.5	361.5	173.9	1.298	2.167	1.043	8.073	3.858	4.739	41.388
5	196.38	196(49)	196	196.4	241.6	372.4	201.8	1.230	1.896	1.027	7.699	3.158	3.635	40.360
6	223.86	224(56)	224	225.1	261.3	378.3	229.8	1.161	1.681	1.021	6.639	2.608	2.694	43.432
7	249.21	248(62)	249	247.8	281.5	381.7	253.8	1.136	1.540	1.024	6.031	2.287	2.193	44.923

Note: R.D. denotes the relative distance.

Table 4.6 Estimated Variance and Expected Number of Units Traveled Under Path Sampling, Cluster Sampling (Cluster size = 4) and Random Walk Sampling for the Modified Longleaf Pin Data with a Simulation of 1000 Iterations (with Starting Unit (1,10))

p	Expected sample size			$\hat{MSE}(\hat{\mu}_{ps})$	$\hat{MSE}(\hat{\mu}_{cls})$	$\hat{MSE}(\hat{\mu}_{srs})$	$\hat{MSE}(\hat{\mu}_{rws})$	R.E.cls	R.E.srs	R.E.rws
	PS	Cls	SRS							
1	58	56 (14)	58	0.334	1.012	0.995	2.596	3.033	2.982	7.782
2	98.77	100(25)	99	0.133	0.550	0.500	1.701	4.133	3.758	12.794
3	134.31	136(34)	135	0.079	0.348	0.335	1.375	4.398	4.244	17.403
4	166.63	168(42)	167	0.056	0.258	0.221	1.337	4.582	3.929	23.789
5	196.38	196(49)	196	0.043	0.169	0.158	1.144	3.910	3.646	26.456
6	223.86	224(56)	224	0.035	0.143	0.125	1.082	4.073	3.555	30.854
7	249.21	248(62)	249	0.030	0.117	0.101	0.989	3.851	3.351	32.694

Table 4.6 (Continued)

p	Expected sample size			Expected number of units traveled				R.D.cls	R.D.srs	R.D.rws	$\hat{MSE}$ * Expected number of units traveled			
	PS	Cls	SRS	PS	Cls	SRS	RWS				PS	Cls	SRS	RWS
1	58	56 (14)	58	58.9	104.8	263.8	63.7	1.779	4.479	1.081	19.648	106.039	262.417	165.365
2	98.77	100(25)	99	98.7	156.7	325.2	104.8	1.588	3.295	1.062	13.123	86.113	162.477	178.265
3	134.31	136(34)	135	134.3	191.2	352.1	140.8	1.424	2.622	1.048	10.611	66.444	118.059	193.600
4	166.63	168(42)	167	167.1	216.3	362.8	173.0	1.294	2.171	1.035	9.392	55.704	80.124	231.301
5	196.38	196(49)	196	194.7	241.8	373.2	201.7	1.242	1.917	1.036	8.419	40.886	58.835	230.745
6	223.86	224(56)	224	224.9	262.3	379.8	229.8	1.166	1.689	1.022	7.887	37.465	47.345	248.644
7	249.21	248(62)	249	247.9	282.2	383.4	255.0	1.138	1.547	1.029	7.499	32.878	38.859	252.195

CHAPTER 5

SUMMARY DISCUSSION AND FUTURE RESEARCH

5.1 Summary and conclusions

Consider the population region partitioned into an $r \times c$ (r rows and c columns) grid of rc quadrats or secondary units. The population consists of rc spatial units. Each population unit is labeled with 2 coordinates, say (i, j), which are the row and column of the unit, respectively, for $i = 1, 2, 3, \ldots, r$ and $j = 1, 2, 3, \ldots, c$. Associated with each unit (i, j) , the value of the population variable of interest is denoted as $y_{(i,j)}$. The parameter of interest in this paper is the population mean

$$\mu = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c} y_{(i,j)} = \frac{1}{rc}\sum_{all\,(i,j)} y_{(i,j)} \,. \tag{5.1}$$

For a spatial population of r rows, there are $q = r-1$ possible paths. In general, a path k in the spatial setting population of r rows and c columns with the starting unit $(1, j^*)$ can be written as

$$\begin{aligned} P_k = ((1, j^*), (2, j^*), (3, j^*), \ldots, (k, j^*), (k, j^*-1), (k, j^*-2), \ldots, (k,1), (k+1, 1), \\ (k+1,2), \ldots, (k+1,c), (k,c), (k,c-1), (k,c-2), \ldots, (k, j^*+1), (k-1, j^*+1), \\ (k-2, j^*+1), \ldots, (1, j^*+1)) \end{aligned}$$

for $k = 1, 2, 3, \ldots, q = r-1$.

p paths are selected by SRSWOR from q all possible paths in the population. Let p_k denote a path k in the sample for $k = 1, 2, 3, \ldots, p$. The sample consists of all units in the selected paths. The sample is represented as $p_s = (p_1, p_2, p_3, \ldots, p_p)$.

The probability of selecting a sample is $P(s)=\frac{1}{\binom{q}{p}}=\frac{1}{\binom{r-1}{p}}$ since paths are selected by SRSWOR, and the inclusion probability of path k is $\pi_k=\frac{p}{q}=\frac{p}{r-1}$. Although each path has an equal probability of selection, the units do not have an equal probability of selection, as the same unit may be in one or more paths. There is an overlapping of paths. It is assumed that the units are sampled in a logical manner such that all units will only be observed once. Finally, the researcher can define the rows and columns arbitrarily; thus, path sampling is not limited in its starting or ending position.

The inclusion probability of a unit *(i, j)* can be written, in generic formula, as

$$\pi_{(i,j)}=\begin{cases} 1-\frac{\binom{i-2}{p}}{\binom{q}{p}} & i=1,2,3,\ldots,r \text{ and } j=j^* \text{and } j^*+1 \\ 1-\frac{\binom{q-2}{p}}{\binom{q}{p}} & i=2,3,\ldots,r-1 \text{ and } j=1,2,3,\ldots,j^*\text{-}1,j^*+2,j^*+3,\ldots,c \\ 1-\frac{\binom{q-1}{p}}{\binom{q}{p}} & i=1,r \text{ and } j=1,2,3,\ldots,j^*\text{-}1,j^*+2,j^*+3,\ldots,c \end{cases} \tag{5.2}$$

Note that, for any constant a < b, it is defined that $\binom{a}{b}=0$.

The joint inclusion probabilities is

$$\pi_{(i,j),(i',j')}=\pi_{(i,j)}+\pi_{(i',j')}\text{-}(1\text{-}\frac{\binom{f}{p}}{\binom{q}{p}}). \tag{5.3}$$

f can be found as follows.

Case1: For Units of Type 1

For i, i' = 1, 2, 3, …, r and j, $j' = j^*$ and j^*+1

$$f = \min(i, i') - 2 \tag{5.4}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 2: For Units of Type 1 and 2

For i = 1, 2, 3, …, r and $j = j^*$ and j^*+1 and i' = 2, 3, …, r-1 and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c

$$f = \begin{cases} i-2 & if\ i \le i' \\ i-3 & if\ i - i' = 1 \\ i-4 & if\ i - i' \ge 2 \end{cases} \tag{5.5}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 3: For Units of Type 1 and 3

For i = 1, 2, 3, …, r and $j = j^*$ and j^*+1 and i' = 1 and r and j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c

$$f = \begin{cases} i-2 & i \le i' \\ i-3 & i > i' \end{cases} \tag{5.6}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 4: For Units of Type 2

For i, i' = 2, 3, …, r-1 and j, j' = 1, 2, 3,…, j^*-1, j^*+2, j^*+3,…, c

$$f = \begin{cases} q-2 & if\ |i-i'|=0 \\ q-3 & if\ |i-i'|=1 \\ q-4 & if\ |i-i'|\geq 2 \end{cases} \tag{5.7}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 5: For Units of Type 2 and 3

For $i = 2, 3, \ldots, r\text{-}1$ and $j = 1, 2, 3, \ldots, j^*\text{-}1, j^*+2, j^*+3, \ldots,$ c and $i' = 1$ and r and $j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*+2, j^*+3, \ldots, c$

$$f = \begin{cases} q-2 & |i-i'|=1 \\ q-3 & |i-i'|\geq 2 \end{cases} \tag{5.8}$$

Note that if $f < 0$, then it is set that $f = 0$.

Case 6: For Units of Type 3

For $i, i' = 1$ and r and $j, j' = 1, 2, 3, \ldots, j^*\text{-}1, j^*+2, j^*+3, \ldots, c$

$$f = \begin{cases} q-1 & i = i' \\ q-2 & i \neq i' \end{cases} \tag{5.9}$$

Let $p_s = (p_1, p_2, p_3, ..., p_p)$ denote the sampled paths. Let s denote the set of distinct units in the sample. By using the Horvitz-Thompson estimator (Horvitz-Thompson, 1952), the unbiased estimator of the population mean under path sampling is

$$\hat{\mu}_{ps} = \frac{1}{rc} \sum_{(i,j)\in s} \frac{y_{(i,j)}}{\pi_{(i,j)}}. \tag{5.10}$$

Let $I_{(i,j)}$ be the indicator function taking the value of one if unit (i, j) is selected in the sample and 0 otherwise. It can be written as

$$I_{(i,j)} = \begin{cases} 1 & \text{if unit } (i,j) \text{ is included in the sample} \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

Therefore, $\hat{\mu}_{ps}$ can be written in the alternative form

$$\hat{\mu}_{ps} = \frac{1}{rc} \sum_{all(i,j)} \frac{y_{(i,j)} I_{(i,j)}}{\pi_{(i,j)}}. \tag{5.12}$$

$\hat{\mu}_{ps}$ is the unbiased estimator for population mean μ. The variance of $\hat{\mu}_{ps}$ is

$$v(\hat{\mu}_{ps}) = \frac{1}{(rc)^2} \left(\sum_{i=1}^{r} \sum_{j=1}^{c} \left(\frac{1-\pi_{(i,j)}}{\pi_{(i,j)}} \right) y_{(i,j)}^2 + \sum_{i=1}^{r} \sum_{i' \neq i} \sum_{j=1}^{c} \sum_{j' \neq j} \left(\frac{\pi_{(i,j),(i',j')} - \pi_{(i,j)} \pi_{(i',j')}}{\pi_{(i,j)} \pi_{(i',j')}} \right) y_{(i,j)} y_{(i',j')} \right) \tag{5.13}$$

and the estimated variance is

$$\hat{v}(\hat{\mu}_{ps}) = \frac{1}{(rc)^2} \left(\sum_{(i,j) \in s} \left(\frac{1}{\pi_{(i,j)}^2} - \frac{1}{\pi_{(i,j)}} \right) y_{(i,j)}^2 + \sum_{\substack{(i,j),(i',j') \in s \\ (i,j) \neq (i',j')}} \left(\frac{1}{\pi_{(i,j)} \pi_{(i',j')}} - \frac{1}{\pi_{(i,j),(i',j')}} \right) y_{(i,j)} y_{(i',j')} \right) \tag{5.14}$$

It is unbiased. This variance estimate may be negative. A researcher may use an alternative variance estimator, such as that proposed by Sen (1953) and Yate and Grundy (1953), which is claimed to be less often negative.

According to the study, when p = 1, there exist zero joint inclusion probabilities, except for a population of two columns. Thus, $\hat{v}(\hat{\mu}_{ps})$ is a biased estimator for variance $v(\hat{\mu}_{ps})$ when p = 1, except for a population of two columns.

To apply path sampling in a non-rectangular region, the first step is to try to create a rectangular region around non-rectangular region. Then ordinary path sampling is applied.

5.2 Discussion

Path sampling utilizes all of the observations over the units traveled, while other sampling methods, such as SRSWOR and cluster sampling, travel from place to place to observe all of the units in the sample. Thus, for the same sample size, path sampling yields a smaller number of units traveled to observe all units in the sample than cluster sampling, SRSWOR, and random walk sampling.

Path sampling can be very cost effective for sampling many units. This is true when cost is mainly a function of distance travelled, as the number of units sampled equals the number of units travelled. In the simulations in the present study the number of units sampled for all sampling designs were comparable; however, in situations with budget constraints it is possible that a researcher could sample more units with path sampling, thus giving it an added advantage in this respect. Unfortunately, for path sampling the number of units in the final sample is random and can vary a lot as a result of the number of units in each path. Therefore the expense of sampling when cost is a function of distance travelled would also be random, possibly creating budget problems. In this study, path sampling was used in the case of objects that could not move, such as trees and rocks.

To investigate efficiency, path sampling was compared to cluster sampling, SRSWOR, and random walk sampling. From the results of the simulation, it was seen that for rare and non-rare populations with high variation of y-values among clusters, path sampling was more efficient than random walk sampling. With the path starting or ending point on a high y-value column for this kind of population, path sampling is more efficient than cluster sampling and SRSWOR. However, for a non-rare population with low variation of y-value among clusters, path sampling is less efficient than cluster sampling, SRSWOR, and random walk sampling.

The simulation results with the path starting or ending point on the high y-value column show that when the C.V. among clusters is greater than 1.0, path sampling is more efficient than cluster sampling. When the C.V. among clusters is greater than 0.5, path sampling is more efficient than random walk sampling. When

the C.V. among clusters is greater than or equal to 2.0, path sampling is more efficient than SRSWOR.

In this study, it was found that, if a starting point could be set in a high-value column, path sampling would be more efficient. Since the y-value is unknown, the auxiliary variable can be used to identify the high-value column.

The cost of path sampling is compared to cluster sampling, SRSWOR, and random walk sampling. The cost is considered as the number of units traveled to observe all of the units in a sample. According to the simulation results, path sampling has a smaller value of the expected number of units traveled than cluster sampling and SRSWOR for any p, but a little smaller than random walk sampling.

Path sampling is good in the situation where traveling from place to place is difficult and of high cost. Assume, for example, that one wants to estimate the number of plants in a pond. With SRSWOR or cluster sampling, travel by boat from place to place for sampling is not convenient and is time consuming. Using path sampling, one can travel in the sample paths by boat, which is more convenient and saves time. Moreover, one may use random walk sampling since one can travel in the route or path to observe all of the sampled units, which is convenient, as with path sampling. However, travelling from the edge of the region to the starting unit and travelling back from the last unit in a wave results in more units traveled, and are more time consuming than with path sampling. Path sampling should be implemented when two conditions are met–when the cost of the sampling is mainly a function of the distance travelled, and when it is believed that the y-values are positively correlated with the probability of selection.

5.3 Recommendations for Future Research

In this study, all possible paths are created in a certain way, so that inclusion probabilities and joint inclusion probabilities can be obtained and the Horvitz-Thompson estimator can be applied. Another form of path could be created that is more convenient and less expensive. Moreover, other estimators could be created to improve the precision. In a rare and clustered population, adaptive path sampling could be of interest.

BIBLIOGRAPHY

Birnbaum, Z. W. and Sirken, M. G. 1965. Design of Sampling Surveys to Estimate the Prevalence of Rare Diseases: Three Unbiased Estimates. **Vital and Health Statistics.** Washington, DC: U.S. Government Printing Office.

Brewer, K. R. W. 1963. Ratio Estimation and Finite Populations: Some Results Deducible from the Assumption of an Underlying Stochastic Process. **Australian Journal of Statistics.** 6: 93-105.

Brewer, K. and Donadio, M. 2003. The High Entropy Variance of Horvitz-Thompson Estimator**. Survey Methodology.** 29: 189-196.

Brewer, K. R. W. and Hanif, M. 1983. **Sampling with Unequal Probabilities.** NewYork: Springer-Verlag.

Borkowski, J. J. 2003. Simple Latin Square Sampling -K Designs. **Communications in Statistics Theory and Methods.** 32: 215-237.

Chaudhuri, A. and Vos, J. W. E. 1988. **Unified Theory and Strategies of Survey Sampling.** Amsterdam: North-Holland.

Cochran, W. G. 1946. Relative Accuracy of Systematic and Stratified Random Samples for a Certain Class of Population. **Annual of Mathematical Statistics.** 17: 164-177.

Cochran, W. G. 1977. **Sampling Techniques.** 2nd ed. New York: John Wiley & Sons.

Coleman, J. S. 1958. Relational Analysis: the Study of Social Organizations with Survey Methods. **Human Organization**. 17: 28-36.

Deshpande, M. N. 1985. Improving on Horvitz-Thompson's Estimator. **The Indian Journal of Statistics.** 47: 290-291.

Dryver, A. L. and Chao, C. T. 2006. Ratio Estimators in Adaptive Cluster Sampling. **Environmetrics.** 18 (6): 607-620.

Dryver, A. L. and Thompson, S. K. 2005. Improving Unbiased Estimators in Adaptive Cluster Sampling. **Journal of Royal Statistical Society.** 67 (1): 157-166.

Felix-Medina, M. H. and Monjardin, P. E. 2009. Link-Tracing Sampling with an Initial Sequential Sample of Sites: Estimating the Size of a Hidden Human Population. **Statistical Methodology.** 6: 490-502.

Godambe, V. P. 1955. A unified Theory of Sampling from Finite Populations. **Journal of Royal Statistical Society.** 17: 269-278.

Godambe, V. P. and Joshi, V. M. 1965. Admissibility and Bayes Estimation in Sampling finite Populations I. **Annuals of Mathematical Statistics.** 36: 1707-1722.

Goodman, L. A. 1961. Snowball Sampling. **Annuals of Mathematical Statistics.** 32: 148-170.

Haldane, J. B. S. 1945. On a Method of Estimating Frequencies. **Biometrika.** 33 (3): 222-225.

Hansen, M. M. and Hurwitz, W. N. 1943. On the Theory of Sampling from Finite Populations. **Annuals of Mathematical Statistics.** 14: 333-362.

Hanurav, T. V. 1962. On 'Horvitz and Thompson Estimator'. **The Indian Journal of Statistics.** 24 (4): 429-436.

Horvitz, D. G. and Thompson, M. E. 1952. A Generalization of Sampling without Replacement from a Finite Universe. **Journal of the American Statistical Association.** 47 (260): 663-685.

Klovdahl, A. S. et al. 1977. Social Networks in an Urban area: First Canberra Study. **Australian and New Zealand Journal of Sociology.** 3 (2): 169-172.

Lindberg, M. and Rexstad, E. 2002. Capture-Recapture Sampling Designs. **Encyclopedia of Environmetrics.** 1: 251-262.

Lohr, S. L. 1999. **Sampling: Design and Analysis.** Pacific Grove, CA: Brooks/ Cole Pub

Lucas, H. A. and Seber, G. A. F. 1977. Estimating Coverage and Particle Density Using the Line Intercept Method. **Biometrika.** 64: 618-622.

Madow, W. G. 1949. On the Theory of Systematic Sampling, II. **Annals of Mathematical Statistics.** 20: 333-354.

Midzuno, H. 1950. An Outline of the Theory of Sampling Systems. **Annals of the Institute of Statistical Mathematics.** 1 **(**1): 149-156.

Mohammadi, M. and Salehi, M. M. 2011. Horvitz Thompson Estimator of Population Mean under Inverse Sampling Designs. **Iranian Mathematical Society.** 1-14.

Munholand, P. L. and Borkowski, J. J. 1996. Latin Square Sampling +1 Designs. **Biometrics.** 52: 125-136.

Murthy, M. N. 1957. Ordered and Unordered Estimators in Sampling Without Replacement. **The Indian Journal of Statistics.** 18, 3/4 (September): 379-390.

Nafiu, L. A. and Adewara, A. A. 2007. On the Use of Network Sampling in Diabetic Surveys. **Journal of Research in National Development.** 5 (2): 5-9.

Neyman, J. 1934. On the Two Different Aspects of the Representative Methods: the Method of Stratified Sampling and the Method of Purposive Selection. **Journal of Royal Statistical Society.** 97: 558-625.

Neyman, J. 1938. Contribution to the Theory of Smpling Human Populations. **Journal of American Association.** 35: 101-116.

Overton, W. S. and Stehman, S. V. 1995. The Horvitz-Thompson Theorem as a Unifying Perspective for Probability Sampling: With Examples from Natural Resource Sampling. **The American Statistician.** 49 (3): 261-268.

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. and Singh, M. P. 1973. On the Choice of Estimator in Survey Sampling. **Austral. Journal of Statist.** 15 (2): 95-104.

Salganik, M. J. and Heckathorn, D. D. 2004. Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. **Sociological Methodology.** 34: 193-239.

Seber, G. A. F. 1973. **The Estimation of Animal Abundance.** London: Griffin.

Seber, G. A. F. 1982. **The Estimation of Animal Abundance.** 2nd ed. London: Griffin.

Seber, G. A. F. 1986. A Review of Estimating Animal Abundance. **Biometrics.** 42 (2): 267-292.

Seber, G. A. F. 1992. A Review of Estimating Animal Abundance II. **International Statistical Review/Revue International de Statistique.** 60 (2): 129.

Sen, A. R. 1953. On the Estimate of the Variance in Sampling with Varying Probabilities. **Journal of the Indian Society of Agricultural Statistics.** 5: 119-127.

Smith, D. R.; Conroy, M. J. and Brakhage, D. H. 1995. Efficiency of Adaptive Cluster Sampling for Estimating Density of Wintering Waterfowl. **Biometrics**. 51: 777-788.

Splawa-Neyman, J. 1925. Contributions to the Theory of Small Samples Drawn from a Finite Population**. Biometrika.** 17: 472-479.

Stehman, S. V. and Overton, W. S. 1987. **Estimating the Variance of the Horvitz-Thompson Emator in Variable Probability, Systematic Samples. Proceeding of the Section on Survey Research Methods.** American Statistical Association Annual Meeting. Pp. 743-748.

Taga, Y. 1993. Generalization of HT Strategy and Its Application. **Yokohama Math. Journal.** 40: 163-173.

Thompson, S. K. 1990. Adaptive Cluster Sampling. **Journal of the American Statistical Association.** 85: 1050-1059.

Thompson, S. K. 1991a. Adaptive Cluster Sampling: Designs with Primary and Secondary Units. **Biometrics.** 47: 1103-1115.

Thompson, S. K. 1991b. Stratified Adaptive Cluster Sampling. **Biometrika**. 78: 389-397.

Thompson, S. K. 2002. **Sampling.** 2nd ed. New York: Wiley.

Thompson, S. K. 2006a. Adaptive web sampling. **Biometrics.** 62: 1-24.

Thompson, S. K. 2006b. Targeted Random Walk Designs. **Survey Methodology.** 32: 1-36.

Thompson, S. K. and Ramsey, F. L. 1983. **Adaptive Sampling of Animal Populations.** Technical Report 82. Department of Statistics, Corvallis: Oregon State University.

Thompson, S. K. and Seber, G. A. F. 1996. **Adaptive Sampling.** New York: Wiley.

Vijayan, K. 1966. On Horvitz-Thompson and Des Raj Estimator. **The Indian Journal of Statistics.** 28: 87-92.

Vincent, K. S. 2008. **Design Variations in Adaptive Web Sampling.** Master's thesis, Simon Fraser University.

Yates, F. and Grundy, P. M. 1953. Selection without Replacement from Within Strata with Probability Proportional to Size. **Journal of the Royal Statistical Society.** 15 (1): 253-261.

APPENDICES

Appendix A

Estimated Variance, Mean Squared Error Estimated, and Relative Efficiency of Path Sampling (ps), Cluster Sampling (cls), SRSWOR (srs), and Random Walk Sampling (rws) with starting unit in high y-value column (1, 15)

C.V.	p	E(υ)	m_c	$\hat{MSE}(\mu_{ps})$	$\hat{MSE}(\mu_{cls})$	$\hat{MSE}(\mu_{srs})$	$\hat{MSE}(\mu_{rws})$	R.E.cls	R.E.srs	R.E.rws
0.3	1	48	50 (5)	222.09	43.26	13.52	64.57	0.19	0.06	0.29
	2	83.33	90 (9)	128.57	16.72	6.18	56.53	0.13	0.05	0.44
	3	113	120 (12)	87.67	9.17	3.46	43.22	0.10	0.04	0.49
	4	138	140(14)	61.78	5.65	2.09	40.22	0.09	0.03	0.65
	5	158.6	160 (16)	44.66	3.56	1.21	39.27	0.08	0.03	0.88
0.4	1	48	50 (5)	193.91	74.51	27.21	113.12	0.38	0.14	0.58
	2	83.33	90 (9)	132.13	29.54	11.97	75.25	0.22	0.09	0.57
	3	113	120 (12)	84.75	16.24	6.29	71.16	0.19	0.07	0.84
	4	138	140(14)	62.52	10.50	3.51	59.23	0.17	0.06	0.95
	5	158.6	160 (16)	41.67	6.64	2.20	53.33	0.16	0.05	1.28

(Continued)

C.V.	p	E(υ)	m_c	$\hat{MSE}(\mu_{ps})$	$\hat{MSE}(\mu_{cls})$	$\hat{MSE}(\mu_{srs})$	$\hat{MSE}(\mu_{rws})$	R.E.cls	R.E.srs	R.E.rws
0.5	1	48	50 (5)	285.17	123.94	32.25	235.1	0.43	0.11	0.82
	2	83.33	90 (9)	159.74	50.09	15.57	144.77	0.31	0.10	0.91
	3	113	120 (12)	103.75	27.49	9.14	103.11	0.26	0.09	0.99
	4	138	140(14)	70.46	16.99	4.78	98.34	0.24	0.07	1.40
	5	158.6	160 (16)	50.79	10.53	2.93	92.54	0.21	0.06	1.82
0.6	1	48	50 (5)	239.85	190.96	37.83	242	0.80	0.16	1.01
	2	83.33	90 (9)	135.45	73.76	17.12	164.21	0.54	0.13	1.21
	3	113	120 (12)	91.71	40.36	9.58	102.22	0.44	0.10	1.11
	4	138	140(14)	60.63	25.60	5.95	111.5	0.42	0.10	1.84
	5	158.6	160 (16)	43.77	15.84	3.23	86.96	0.36	0.07	1.99
0.7	1	48	50 (5)	238.26	278.66	42.14	302.83	1.17	0.18	1.27
	2	83.33	90 (9)	134.31	107.17	19.69	207.61	0.80	0.15	1.55
	3	113	120 (12)	89.59	58.52	9.92	167.53	0.65	0.11	1.87
	4	138	140(14)	61.63	36.39	61.63	137.36	0.59	1.00	2.23
	5	158.6	160 (16)	43.23	22.78	3.69	121.8	0.53	0.09	2.82

(Continued)

C.V.	p	E(υ)	m_c	$\hat{MSE}(\mu_{ps})$	$\hat{MSE}(\mu_{cls})$	$\hat{MSE}(\mu_{srs})$	$\hat{MSE}(\mu_{rws})$	R.E.cls	R.E.srs	R.E.rws
0.8	1	48	50 (5)	325.96	332.77	56.16	445.05	1.02	0.17	1.37
	2	83.33	90 (9)	195.16	131.49	27.07	292.59	0.67	0.14	1.50
	3	113	120 (12)	116.99	72.68	116.99	230.51	0.62	1.00	1.97
	4	138	140(14)	82.03	8.04	8.04	187.12	0.10	0.10	2.28
	5	158.6	160 (16)	52.26	4.79	4.79	179.3	0.09	0.09	3.43
0.9	1	48	50 (5)	392.98	503.16	64.44	527.1	1.28	0.16	1.34
	2	83.33	90 (9)	207.04	199.09	25.94	357	0.96	0.13	1.72
	3	113	120 (12)	120.20	108.06	15.25	311.98	0.90	0.13	2.60
	4	138	140(14)	90.31	67.64	8.49	257.83	0.75	0.09	2.86
	5	158.6	160 (16)	54.33	42.74	5.33	233.19	0.79	0.10	4.29
1.0	1	48	50 (5)	680.67	847.11	115.05	721.55	1.24	0.17	1.06
	2	83.33	90 (9)	333.35	342.93	56.56	442.86	1.03	0.17	1.33
	3	113	120 (12)	199.19	177.43	30.84	369.55	0.89	0.15	1.86
	4	138	140(14)	145.71	114.44	16.08	317.74	0.79	0.11	2.18
	5	158.6	160 (16)	91.23	69.69	9.48	245.58	0.76	0.10	2.69

(Continued)

C.V.	p	E(υ)	m_c	$\hat{MSE}(\mu_{ps})$	$\hat{MSE}(\mu_{cls})$	$\hat{MSE}(\mu_{srs})$	$\hat{MSE}(\mu_{rws})$	R.E.cls	R.E.srs	R.E.rws
1.1	1	48	50 (5)	549.40	1096.92	146.02	892.14	2.00	0.27	1.62
	2	83.33	90 (9)	318.00	456.90	62.65	504.27	1.44	0.20	1.59
	3	113	120 (12)	198.21	232.21	37.98	450.42	1.17	0.19	2.27
	4	138	140(14)	139.31	152.31	20.83	379.57	1.09	0.15	2.72
	5	158.6	160 (16)	92.61	92.90	13.10	309.16	1.00	0.14	3.34
1.2	1	48	50 (5)	971.60	2197.71	411.30	1897.4	2.26	0.42	1.95
	2	83.33	90 (9)	524.25	931.62	171.78	1209.16	1.78	0.33	2.31
	3	113	120 (12)	330.02	463.00	102.33	981.04	1.40	0.31	2.97
	4	138	140(14)	225.01	304.73	56.41	778.94	1.35	0.25	3.46
	5	158.6	160 (16)	145.21	191.57	31.65	722.51	1.32	0.22	4.98
1.4	1	48	50 (5)	992.62	3921.42	583.56	2602.58	3.95	0.59	2.62
	2	83.33	90 (9)	517.33	1544.92	260.50	1649.18	2.99	0.50	3.19
	3	113	120 (12)	326.87	885.29	150.89	1280.47	2.71	0.46	3.92
	4	138	140(14)	216.99	565.70	80.38	1211.06	2.61	0.37	5.58
	5	158.6	160 (16)	139.16	285.02	53.12	1044.11	2.05	0.38	7.50

(Continued)

C.V.	p	E(υ)	m_c	$\hat{MSE}(\mu_{ps})$	$\hat{MSE}(\mu_{cls})$	$\hat{MSE}(\mu_{srs})$	$\hat{MSE}(\mu_{rws})$	R.E.cls	R.E.srs	R.E.rws
1.6	1	48	50 (5)	1924.32	6951.18	1313.09	5355.34	3.61	0.68	2.78
	2	83.33	90 (9)	1042.90	2955.36	644.38	3197.78	2.83	0.62	3.07
	3	113	120 (12)	640.77	1715.54	342.57	2668.85	2.68	0.53	4.17
	4	138	140(14)	419.66	1097.71	193.45	2240.85	2.62	0.46	5.34
	5	158.6	160 (16)	268.09	599.92	122.94	2238.1	2.24	0.46	8.35
1.8	1	48	50 (5)	1494.76	11399.61	1657.03	9636.85	7.63	1.11	6.45
	2	83.33	90 (9)	959.53	4850.70	745.82	5760.54	5.06	0.78	6.00
	3	113	120 (12)	647.79	2535.40	462.57	4396.89	3.91	0.71	6.79
	4	138	140(14)	460.33	1620.48	251.24	4039.48	3.52	0.55	8.78
	5	158.6	160 (16)	301.15	955.87	149.13	4086.27	3.17	0.50	13.57
2.0	1	48	50 (5)	204926.00	675254.70	246886.90	1505684	3.30	1.20	7.35
	2	83.33	90 (9)	90773.99	266261.20	109102.30	842355.3	2.93	1.20	9.28
	3	113	120 (12)	53918.19	133291.50	59779.90	631202.34	2.47	1.11	11.71
	4	138	140(14)	31208.34	92419.92	35123.77	639559.8	2.96	1.13	20.49
	5	158.6	160 (16)	18682.83	57969.27	19985.13	524275.7	3.10	1.07	28.06

(Continued)

C.V.	p	E(υ)	m_c	$\hat{MSE}(\mu_{ps})$	$\hat{MSE}(\mu_{cls})$	$\hat{MSE}(\mu_{srs})$	$\hat{MSE}(\mu_{rws})$	R.E.cls	R.E.srs	R.E.rws
2.2	1	48	50 (5)	198023.00	873207.10	269798.40	1779027	4.41	1.36	8.98
	2	83.33	90 (9)	91425.72	352873.80	125139.70	1200783.71	3.86	1.37	13.13
	3	113	120 (12)	58314.68	205389.10	65553.19	904518.9	3.52	1.12	15.51
	4	138	140(14)	34407.61	120223.80	37576.56	796036.5	3.49	1.09	23.14
	5	158.6	160 (16)	20390.92	67493.85	21693.10	848517.9	3.31	1.06	41.61
2.4	1	48	50 (5)	202116.20	1483137.00	315018.80	2871031.21	7.34	1.56	14.20
	2	83.33	90 (9)	101762.70	582591.50	137557.50	2135696.54	5.73	1.35	20.99
	3	113	120 (12)	61595.53	325623.30	84768.76	1337381.46	5.29	1.38	21.71
	4	138	140(14)	37563.94	209380.90	48411.94	1285569.7	5.57	1.29	34.22
	5	158.6	160 (16)	26651.64	118284.20	28041.72	1239168.9	4.44	1.05	46.50

BIOGRAPHY

NAME	Miss Mena Patummasut
ACADEMIC BACKGROUND	B.Sc. (Mathematics), Mahidol University Bangkok, Thailand, 2001 M.S. (Applied Statistics), National Institute of Development Administration, Bangkok, Thailand, 2004
PRESENT POSITION	Lecturer, Department of Statistics, Faculty of Science, Kasetsart University, Bangkok, Thailand
EXPERIENCE	Lecturer in Statistics, Department of Statistics, Faculty of Science, Kasetsart University, Bangkok, Thailand
PUBLICATIONS	Patummasut, M. and Dryver, A. L. 2011 Path Sampling**.** In **Proceedings of Applied Statistics 2011 International Conference.** Ribno, September 25-28, 2011. Bled, Slovenia. Pp.25.