International Journal of Mathematics and
Computer Science, **20**(2025), no. 1, 1–7
DOI: https://doi.org/10.69793/ijmcs/01.2025/mena

An Alternative Estimator for Path Sampling

Sudarat Nidsunkid, Mena Lao

Department of Statistics
Faculty of Science
Kasetsart University
Bangkok, Thailand

email: sudarat.n@ku.th, mena_patummasut@yahoo.com

(Received May 7, 2024, Accepted June 14, 2024, Published June 15, 2024)

Abstract

Path sampling is used in area population where paths are defined and sampled from the population. It is a cost saving sampling design because all units in the sample are observed when surveying. The purpose of this paper is to propose an alternative estimator for path sampling by utilizing modified ratio estimator for unequal probability sampling design. Mean square error of the proposed ratio estimator is obtained and compared with the classical ratio estimator. The situations when the proposed ratio estimator is better than the traditional ratio estimator are theoretically presented.

1 Introduction

Proposed in 2012, Path sampling is a cost-effective and convenient sampling design for a spatial population allowing the researcher to sample all units along a path traveled during the sampling. The population area is partitioned into quadrats or units. Then all paths are created with a certain connected starting and ending units. In path sampling, the paths are selected from all possible paths in the population by simple random sampling without replacement (SRSWOR). The unbiased estimator for the population mean

Keywords: Path sampling, Horvitz-Thompson estimator, unequal probability sampling, ratio estimator, auxiliary variable.
AMS (MOS) Subject Classifications: 62D05, 94A20.
The corresponding author is Mena Lao.
ISSN 1814-0432, 2025, http://ijmcs.future-in-tech.net

was proposed in path sampling. In 2021, a ratio estimator for path sampling was introduced by applying the generalized ratio estimator for unequal probability sampling designs [1].

A ratio estimator is widely created in many sampling designs because it provides more precision in estimation. There are many research articles proposing ratio estimators. For instance, Sisodia and Dwivedi [2], Upadhyaya and Singh [3] , Kadilar and Cingi [4] and Srisodaphol, Kingphai and Tanjai [5] developed the ratio estimators of a population mean for simple random sampling. In adaptive cluster sampling, ratio estimators were created by Chao [6], Dryver and Chao [7] and Chutiman and Chiangpradit [8]. In stratified adaptive cluster sampling, a ratio estimator was proposed by Chutiman [9]. Furthermore, Bacanli and Kadilar [10] proposed modified ratio estimators for unequal probability sampling design by using the Horvitz-Thompson ratio type estimators. Therefore, the objective of this research is to propose an alternative ratio estimator for path sampling in order to obtain more precision estimation.

2 Path sampling

The rectangular population region is divided into r rows and c columns forming rc quadrats or units, as shown in Figure 1. The symbol (i, j) refers to a unit in row i and column j for $i = 1, 2, 3, \ldots, r$ and $j = 1, 2, 3, \ldots, c$. A unit h for $h = 1, 2, 3, \ldots, N$ in the population can be represented by a symbol (i, j) where $i = (h \ mod \ r)+1$ and $j = h - (i - 1)c$.

Let y_h denote the value of the variable of interest of unit h, and the population mean is

$$\mu = \frac{1}{N} \sum_{h=1}^{N} y_h. \tag{2.1}$$

A path m is created, as shown in Figure 1, to begin traveling from a unit $(1, j^*)$ and to go straight on column j^*, then turn left at row m and go straight on this row to the end, then go up to row $m + 1$ and straight on this row to the end, then go down to row m and go straight on this row, then turn right at column $j^* + 1$, go straight on this column and stop at unit $(1, j^* + 1)$. By creating a path this way, there are $r - 1$ paths in the population [11], [1].

In path sampling, p paths are sampled using simple random sampling without replacement from $r - 1$ paths in the population. A path sample can be written as $p_s = (p_1, p_2, p_3, ..., p_p)$, where p_m denotes path m in the sample for $m = 1, 2, 3, \ldots, p$. Let $a = r - 1$. The number of all possible path

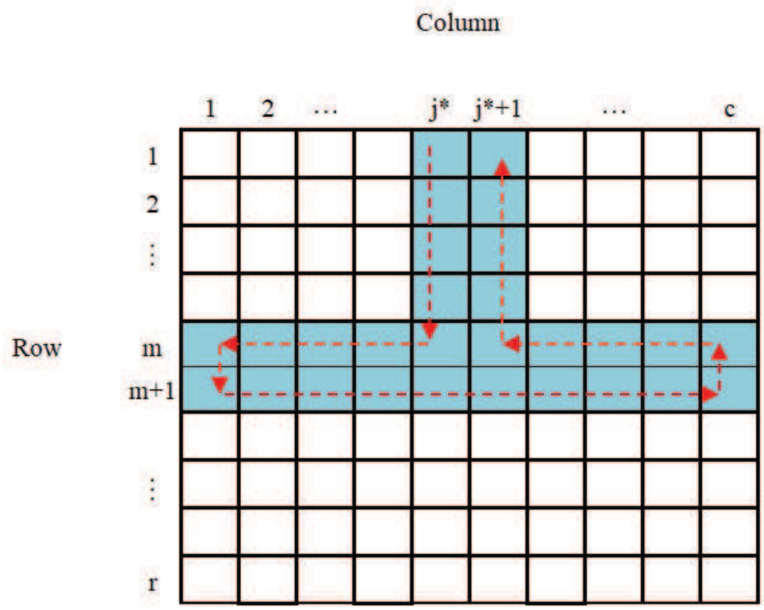

Figure 1: Path m traveling from the starting unit $(1, j^*)$ to the ending unit $(1, j^* + 1)$

samples of size p is the number of combinations of p paths from a paths in the population, denoted as ${}^aC_p$.

Let π_h be an inclusion probability of unit h, which is located at row i and column j and thus may also be denoted as $\pi_{(i,j)}$ as well, is calculated by the formula

$$\pi_h = \pi_{(i,j)} = \begin{cases} 1 - {}^{i-2}C_p/{}^aC_p & for\ (i,j) \in H_1 \\ 1 - {}^{a-2}C_p/{}^aC_p & for\ (i,j) \in H_2 \\ 1 - {}^{a-1}C_p/{}^aC_p & for\ (i,j) \in H_3 \end{cases} \tag{2.2}$$

where $i = (h\ mod\ r)+1$, $j = h - (i-1)c$, H_1, H_2 and H_3 are sets of units as follows. $H_1 = \{(i,j) \mid i = 1, 2, 3, \ldots, r \text{ and } j = j^* \text{ and } j^* + 1\}$ is a set of units in the starting column j^* and ending column $j^* + 1$. $H_2 = \{(i,j) \mid i = 2, 3, \ldots, r-1 \text{ and } j = 1, 2, 3, ..., j^* - 1, j^* + 2, j^* + 3, ..., c\}$ is a set of units not in column j^* and $j^* + 1$ and not in the first row or the last row. $H_3 = \{(i,j) \mid i = 1, r \text{ and } j = 1, 2, 3, ..., j^* - 1, j^* + 2, j^* + 3, ..., c\}$ is a set of units in the first and last row.

Let π_{hl} denote the probability that units h and l are included in the sample, or the joint inclusion probability, defined as

$$\pi_{hl} = \pi_h + \pi_l - (1 - {}^gC_p/{}^aC_p) \tag{2.3}$$

where g is the number of paths not containing either units h or l, and $g \geq 0$.

Let s be the set of distinct units in path sample p_s. The generalized ratio estimator [1] for path sampling is

$$\hat{\mu}_p = \frac{\bar{y}_p}{\bar{x}_p}\mu_x = \hat{R}_p\mu_x \tag{2.4}$$

where $\bar{y}_p = \frac{1}{N}\sum_{u\in s}\frac{y_u}{\pi_u}$ and $\bar{x}_p = \frac{1}{N}\sum_{u\in s}\frac{x_u}{\pi_u}$ are Horvitz-Thompson type estimators of the population means of the study and auxiliary variables, respectively, and $\hat{R}_p = \frac{\bar{y}_p}{\bar{x}_p}$. The approximation of MSE of the estimator $\hat{\mu}_p$ is

$$MSE(\hat{\mu}_p) = \frac{1}{N^2}\left[\sum_{u=1}^{N}\sum_{v=1}^{N}(\frac{\pi_{uv} - \pi_u\pi_v}{\pi_u\pi_v})y'_u y'_v\right] \tag{2.5}$$

where $y'_u = y_u - Rx_u$ and $R = \frac{\mu_y}{\mu_x}$.

3 A ratio estimator for unequal probability sampling

Bacanli and Kadilar [10] developed a ratio estimator for unequal probability sampling design:

$$\bar{y}_B = \frac{\bar{y}_{HT}(\mu_x + C_x)}{\bar{x}_{HT} + C_x} \tag{3.6}$$

where $\bar{y}_{HT} = \frac{1}{N}\sum_{i=1}^{n}\frac{y_i}{\pi_i}$ and $\bar{x}_{HT} = \frac{1}{N}\sum_{i=1}^{n}\frac{x_i}{\pi_i}$ are Horvitz-Thompson type estimators of the population means of the study and auxiliary variables. The MSE of this estimator can be given by

$$MSE(\bar{y}_B) = \frac{1}{N^2}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}(\frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j})y_i^* y_j^*\right] \tag{3.7}$$

where $y_i^* = y_i - R_B x_i$, $R_B = \frac{\mu_y}{\mu_x + C_x}$ and C_x is the population coefficient of variation of the auxiliary variable, and B_x is the population kurtosis of the auxiliary variable.

4 An alternative ratio estimator for path sampling

Combining the Bacanli and Kadilar ratio type estimator for unequal probability sampling and the generalized ratio estimator [1] for path sampling, we propose the alternative ratio estimator for path sampling:

$$\hat{\mu}_{p1} = \frac{\bar{y}_p(\mu_x + C_x)}{\bar{x}_p + C_x}$$

Using Eq. (3.7), the MSE of the proposed estimator can be given by

$$MSE(\hat{\mu}_{p1}) = \frac{1}{N^2}\left[\sum_{u=1}^{N}\sum_{v=1}^{N}(\frac{\pi_{uv} - \pi_u\pi_v}{\pi_u\pi_v})y_u^* y_v^*\right]$$

where $y_u^* = y_u - R_1 x_u$ and $R_1 = \frac{\mu_y}{\mu_x + C_x}$.

5 Efficiency comparison

The MSE of the proposed modified ratio estimator and MSE of classical ratio estimator are compared as follows:
Let $e_{uv} = \frac{\pi_{uv} - \pi_u\pi_v}{\pi_u\pi_v}$ and $\tau = \frac{\sum_{u=1}^{N}\sum_{v=1}^{N} e_{uv}(y_u x_v - y_v x_u)}{\sum_{u=1}^{N}\sum_{v=1}^{N} e_{uv} x_u x_v}$. Consider

$$MSE(\hat{\mu}_{p1}) < MSE(\hat{\mu}_p)$$

$$\sum_{u=1}^{N}\sum_{v=1}^{N} e_{uv}(y_u - R_1 x_u)(y_v - R_1 x_v) < \sum_{i=1}^{N}\sum_{j=1}^{N} e_{uv}(y_u - R x_u)(y_v - R x_v)$$

$$\sum_{u=1}^{N}\sum_{v=1}^{N} e_{uv}(y_u y_v - R_1 y_u x_v - R_1 y_v x_u + R_1^2 x_u x_v)$$

$$< \sum_{u=1}^{N}\sum_{v=1}^{N} e_{uv}(y_u y_v - R y_u x_v - R y_v x_u + R^2 x_u x_v)$$

$$\sum_{u=1}^{N}\sum_{v=1}^{N} e_{uv}\{(R_1 - R)(-y_u x_v - y_v x_u) + (R_1^2 - R^2)x_u x_v\} < 0,$$

$$R_1 + R < \tau; \qquad for \quad R_1 > R \tag{5.8}$$

or

$$R_1 + R > \tau; \qquad for \quad R_1 < R \tag{5.9}$$

Therefore, when the condition (5.8) or (5.9) is satisfied, the proposed ratio estimator is more efficient than the classical ratio estimator, $\hat{\mu}_p$.

6 Conclusion

In this paper, an alternative ratio estimator was proposed for path sampling and its efficiency was investigated. The mean square error was obtained and compared with the classical ratio estimator. When the proposed ratio estimator is better than the traditional ratio estimator, the conditions were presented theoretically. We discovered that when condition 1: $R_1 + R < \tau$ for $R_1 > R$ or condition 2: $R_1 + R > \tau$ for $R_1 < R$ was satisfied, the proposed ratio estimator was more efficient than the classical ratio estimator $\hat{\mu}_p$.

Acknowledgment. This work is supported by a research grant from the Department of Statistics, Faculty of Science, Kasetsart University.

References

[1] M. Patummasut, S. Nidsunkid, A ratio estimator in path sampling, Communications in Statistics - Simulation and Computation, **50**, no. 2, (2021), 376–387.

[2] B. V. S. Sisodia, V. K. Dwivedi, A modified ratio estimator using coefficient of variation of auxiliary variable, Journal of the Indian Society of Agricultural Statistics, **33**, (1981), 13–18.

[3] L. N. Upadhyaya, H. P. Singh, Use of transformed auxiliary variable in estimating the finite population mean, Biometrical Journal, **41**, no. 5, (1999), 627–636.

[4] C. Kadilar, H. Cingi, Ratio estimators in stratified random sampling, Biometrical Journal, **45**, no. 2, (2003), 218–225.

[5] W. Srisodaphol, K. Kingphai, N. Tanjai, New ratio estimators of a population mean using one auxiliary variable in simple random sampling, Chiang Mai Journal of Science **42**, (2015), 523–527.

[6] C. T. Chao, Ratio estimation on adaptive cluster sampling, Journal of Chinese Statistical Association **42**, no. 3, (2004), 307–327.

[7] A. L. Dryver, C. T. Chao, Ratio estimators in adaptive cluster sampling, Environmetric , **18**, no. 6, (2007), 607–620.

[8] N. Chutiman, M. Chiangpradit, Ratio estimator in adaptive cluster sampling without replacement of networks, Journal of Probability and Statistics, (2014), 1–6, Article ID 726398,.

[9] N. Chutiman, A New Ratio Estimator in Stratified Adaptive Cluster Sampling, Thailand Statistician, **8**, no. 2, (2010), 223–233.

[10] S. Bacanli, C. Kadilar, Ratio estimators with unequal probability designs, Pakistan Journal of Statistics **24**, no. 3, (2008), 167–172.

[11] M. Patummasut, A. L. Dryver, A New Sampling Design for a Spatial Population: Path Sampling, Journal of Applied Sciences, **12**, (2012), 1355–1363.