International Journal of Statistical Sciences *ISSN 1683-5603*
Vol.24(1), March, 2024, pp 103-114
DOI: https://doi.org/10.3329/ijss.v24i1.72028

Ratio Type Estimator for Balanced Sampling Plan excluding Adjacent Units

Neeraj Tiwari[1], Jharna Banerjie[2], Girish Chandra[3] and Shailja Bhari[1*]

[1]Department of Statistics, S. S. J. University, Almora-263601
[2]Department of Statistics, D. A. V. (P.G.) College, Dehradun-248001
[3]Department of Statistics, University of Delhi, Delhi-110007
Email: neerajtiwari.amo@gmail.com, jharna_banerjie@yahoo.co.in,
gchandra@stats.du.ac.in and shailjabhari36@gmail.com
*Correspondence should be addressed to Shailja Bhari
(Email: shailjabhari36@gmail.com)

[Received December 18, 2023; Accepted February 27, 2024]

Abstract

The study variables are usually correlated with the auxiliary variables and therefore their correlation could easily be computed. In this paper, the correlation coefficient is used for the estimation procedure, and therefore, we proposed a Horvitz-Thompson ratio-type estimator using correlation coefficient for Balanced Sampling plan excluding Adjacent units (BSA plan). It has been illustrated theoretically and empirically that the proposed Horvitz-Thompson ratio-type estimator is more precise than the Horvitz-Thompson estimator based on BSA plan. The proposed estimator provides an opportunity to utilise the auxiliary information for the estimation of population mean for BSA plan and useful for many real-life experiments.

Keywords and Phrases: BSA plan, Correlation coefficient; Horvitz-Thompson estimator; Ratio-type estimator; Relative precision.

AMS Classification: 62D05, 62G09, 94A20.

1. Introduction

More often, contiguous (non-preferred) and non-contiguous (preferred) units are available in the population in which one aims to select a smaller number of contiguous units and consequently more non-contiguous units. The contiguous units may have many forms. In some cases, the adjacent units might be contiguous since they share the similar information. The persons infected by the same contagious disease (say Covid-19) may be considered as the contiguous units. Balanced Sampling plan Excluding Contiguous units (BSEC) was introduced by Hedayat et al. (1988a, 1988b) with the aims to minimize the probability of contiguous units and simultaneously increase the probability of non-contiguous units. It is the sampling plan in which all non-contiguous units have constant second order inclusion probabilities and all pair of contiguous units has zero second order inclusion probabilities. Stufken (1993) generalised this idea to Balanced

Sampling plans excluding Adjacent units (BSA plans) where all pair of adjacent units have zero second order inclusion probabilities and all pair of non-adjacent units has constant second order inclusion probabilities. Here, two units are said to be adjacent if they are less than or equal to pre-defined distance m (such general BSA plans are denoted by BSA (m) plans).

Sometimes, an auxiliary information (say, variable X) corresponding to study variable (Y) is available, where X and Y are highly correlated to each other. Although, the standard ratio estimator (Cochran, 1977) for population mean is biased but it is preferred over unbiased estimators since it accumulates the additional information from X and have smaller mean square error (MSE) in comparison to estimators based on simple random sampling (SRS). Many authors gave modified version of ratio estimator and reduces the biasness (even upto zero) in it. Researchers developed an unbiased estimator with various sampling schemes which was explored by Hartley and Ross (1954), Robson (1957), Goodman and Hartley (1958), Mickey (1959), Robson and Vithayasai (1961) and others.

Interestingly, certain sampling schemes has been improved using ratio estimators. Lahiri (1951) explored the selection of small samples from many strata using the biased ratio estimator. Midzuno (1952), Horvitz and Thompson (1952), Mickey (1959), Nanjamma, Murthy, and Sethi (1960), Pathak (1964), Raj (1964, 1965) and Sukhatme (1962) investigated unbiased ratio-type estimators in two-stage and multi-stage samplings. Raj (1954), Williams (1961) and Pascual (1961) extended the use of these estimators in the stratified sampling techniques. Nanjamma, Murthy and Sethi (1960) discussed an application of unbiased ratio-type estimator in systematic sampling. Olkin (1958), Williams (1962) and Raj (1965) utilised the multi-auxiliary information for the unbiased ratio-type estimators. Dryver and Chao (2007) used ratio estimators in adaptive cluster sampling (ACS) and Chutiman and Chiangpradit (2014) further utilised the ratio estimators in ACS without replacement of networks. Kadilar, Unyazici, and Cingi (2009) improved the ranked set sampling (RSS) by the use of ratio estimators.

In this paper, we attempted the use of ratio estimator in BSA (m) plan of Stufken (1993) and improved the estimator of the population mean. We used the correlation coefficient along with the ratio estimator for the estimation procedure as there is a high correlation between study variables and auxiliary variables. Hence, Horvitz-Thompson ratio-type estimator based on correlation coefficient for BSA plan has been given. The generator sample procedure has been used for the sample selection in the proposed procedure. The structure of the present paper is given as under.
In Section 2, the BSA plan has been detailed. The proposed ratio-type estimator and its comparison with ratio-type estimator based on simple random sampling (SRS) as well as Horvitz-Thompson estimator for BSA plan is detailed in Section 3. Section 4 reveals the advantage of the proposed estimator with the help of some numerical illustrations. The detailed results of the study are discussed in Section 5.

2. Balanced Sampling plan excluding Adjacent units (BSA plan)

BSEC plan is a type of controlled sampling and a lot of developments has been taken place since its foundation laid by Hedayat et al. (1988a, 1988b). This is the plan in which every pair of non-contiguous units has constant second order inclusion probabilities and every pair of contiguous units has zero second order inclusion probabilities.

Stufken (1993) generalised this sampling plan as BSA plan in which every pair of adjacent units has zero second order inclusion probabilities and every pair of non-adjacent units has constant second order probabilities. The two units are said to be adjacent whenever they are within a distance of m units with m suitably selected by the investigator. However, BSA plan does not always exists for given population size N, sample size n and m. Stufken (1993), Stufken et al. (1999) and Wright (2008) further studied the existence conditions of BSA plans. Stufken et al. (1999) introduced polygonal designs to obtain BSA plans. A polygon design in v treatments and b blocks with each block of size k is an incomplete block design such that (i) no treatment appears more than once in a block, (ii) every treatment appears in r blocks in the design, and (ii) each pair of treatments which are at a distance of more than m units, appear together in λ blocks. The parameter v, b, r, k, λ and m satisfy the following necessary conditions (i) $vr = bk$ and (ii) $\lambda(v - 2m - 1) = r(k - 1)$.

Stufken and Wright (2001) developed an algorithm to obtain BSA plan for m = 1 and Stufken and Wright (2008) for $m \geq 1$. Mandal (2007) provided a catalogue for the construction of BSA plans of $N \leq 7, n \leq 7, m \leq 4$. Mandal et al. (2008) proposed a linear programming to obtain BSA plans using the idea of Rao and Nigam (1990, 1992) but this limits to large N and n, as it produces a large number of possible samples and the linear programming formulation becomes impractical to adopt. Recently, Tiwari et al. (2023) improved BSA (1) plan by the use of ACS.

Colbourn and Ling (1998, 1999) and Tahir et al. (2010, 2012) suggested some techniques to obtain BSA plans with the assumption that the population have a circular ordering. They constructed all the polygonal designs for $v \leq 100,\ k = 3$ for all permissible m using the approach. BSA plans with linear ordering and two-dimensional ordering was considered. Wright (2008) discussed about the two-dimensional BSA plans and Stufken and Wright (2008) showed the existence of linear BSA plans. Mandal et al. (2008) presented a linear programming approach to obtain linear BSA plans. Mandal et al. (2011) proposed an integer linear programming formulation to identify the generator blocks of cyclic polygonal designs. Wang et al. (2015) introduced two-dimensional BSA plan with different adjacency schemes, namely, Row and Column, Sharing a Border and Island. Kumar et al. (2016) obtained a small BSA plan using integer linear programming algorithm, which can produce designs that may or may not be cyclic. Gopinath, Prasad and Mandal (2018) gave an algorithm based on linear programming approach for the development of two-dimensional BSA plans under a border and island adjacency schemes for $m \leq 2$. Kumar et al. (2019) presented an algorithm based on linear integer programming for 66 new BSA plans with one dimensional population linear ordering of units in the parametric range $N \leq 60, n \leq 7, m < 7$.

2.1. Important theorems for the construction of BSA plan

Some theorems related to BSA plan are useful and hence states as follows:

Theorem 1 (Stufken and Wright, 2008): The necessary and sufficient conditions for the existence of a circular BSA$(N, n, m)'s, n \geq 3$ are:

(1) A circular BSA(N, n, m) exists if and only if $N \geq (2m + 1)n$ for the following combinations of (n, m)
$\{(n, 1); n = 3, 4\}, \{(n, 2); n = 3, 4, 5\}, \{(n, 3); n = 3, 4\}, \{(n, 4); n = 3, 4, 5\}$ and $\{(n, 5); n = 3, 4, 5\}$

(2) A circular BSA (N, n, m) exists if and only if $N \geq (2m + 1)n + 1$ for the following combinations of (n, m)
$\{(n, 1); 5 \leq n \leq 15\}, \{(n, 2); 6 \leq n \leq 9\}, \{(n, 3); n = 6, 7\}, \{(n, 4); n = 6, 7\}$ and $(6, 5)$.

where, N, n, m are respectively as population size, sample size and distance between two adjacent units.

Theorem 2 (Hedayat et al., 1988a): Let $A_1, A_2, \dots, A_t$ be subsets of $\{1, 2, \dots, v\}$ with cardinality k. For $A_i = \{a_{i1}, a_{i2}, \dots, a_{ik}\}$, compute the $tk(k-1)$ differences $\pm(a_{ij} - a_{il})$ modulo $v, j \neq l, i = 1, 2, \dots, t$. If the residues 0, 1 and $v-1$ do not appear among these differences, while all others appears equally often, a polygonal design based on number of treatments v block size k exists.

Theorem 3 (Mandal, 2008): Let $A_1, A_2, \dots, A_t$ be subsets of $\{1, 2, \dots, v\}$ with cardinality k. For $A_i = \{a_{i1}, a_{i2}, \dots, a_{ik}\}$, compute the $tk(k-1)$ differences $\pm(a_{ij} - a_{il})$ modulo $v, j \neq l, i = 1, 2, \dots, t$. If the residues 0, 1 to m and $v-m$ to $v-1$ do not appear among these differences, while all others appears equally often, then a polygonal design based on number of treatments v and block size k with $m > 1$ exists.

The first and second order inclusion probabilities of BSA plan of Stufken (1993) is given as
First order inclusion probability, $\pi_i = \frac{n}{N}$, $i \in \{1, \dots, N\}$ (1)

Second order inclusion probability,

$$\pi_{ij} = \begin{cases} 0, & if\ i \text{ and } j \text{ are adjacent} \\ \frac{n(n-1)}{N(N-2m-1)}, & i \neq j \in \{1, \dots, N\}, i \text{ and } j \text{ are not adjacent} \end{cases} \quad (2)$$

N, n and m which satisfy these two probability conditions can only make BSA plan.

2.2. Horvitz-Thompson estimator and its variance for BSA plan

Consider the population of N units with y-values $(y_1, y_2, \dots, y_N)$ and the sample of size n is drawn from the population with the distance m between two distinct units. The aim is to estimate the population mean μ of the y-values. The unbiased Horvitz-Thompson estimator of population mean for BSA plan is given by

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i \epsilon s} z_i \quad (3)$$

where, $z_i = \frac{y_i}{\pi_i}$ and π_i is the first order inclusion probability.

The variance of Horvitz-Thompson estimator $\bar{y}_{HT}$ is

$$V(\bar{y}_{HT}) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\pi_{ij} - \pi_i \pi_j) z_i z_j$$

Provided the second-order inclusion probabilities π_{ij} of all pairs of units in the population are positive, the unbiased estimate of variance is given by

$$\hat{V}(\bar{y}_{HT}) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \Delta_{ij} z_i z_j / \pi_{ij} \quad (4)$$

where $\frac{\Delta_{ij}}{\pi_{ij}} = \begin{cases} 1 - \frac{\pi_i \pi_j}{\pi_{ij}}, for\ i \neq j \\ 1 - \pi_i\ , for\ i = j \end{cases}$

The Horvitz-Thompson estimator $\bar{y}_{HT}$ of Y is an unbiased, however, an unbiased estimator of the variance of $\bar{y}_{HT}$ cannot be obtained.

After making appropriate substitutions and simplifying (4), Wright and Stufken (2011) obtains

$$V(\bar{y}_{HT}) = \frac{1}{N^2}\left[\left(\frac{N-n}{n}\right)\sum_{i=1}^{N} {y_i}^2 + \left(\frac{n(2m+1)-N}{n(N-(2m+1))}\right)\sum_i\sum_j y_{ij} \atop i-j \neq \pm 1, \dots, \pm m \ (mod\ N) - \sum_i\sum_j y_{ij} \atop i-j \neq \pm 1, \dots, \pm m \ (mod\ N)\right] \quad (5)$$

While the first two summations in (5) can be estimated unbiasedly, this is not possible for the last summation when using a circular BSA (N, n, m) since that summation is a sum of cross-products of adjacent terms.

Wright and Stufken (2011) gave the variance estimation using approximation techniques, one of the best approximations of the unbiased estimator of variance of the Horvitz-Thompson estimator is

$$\hat{V}(\bar{y}_{HT}) = \frac{N-(2m+1)n}{2Nn^2(n-1)}\sum_{i\epsilon s}\sum_{j\epsilon s}\left(y_i - y_j\right)^2 \quad (6)$$

This approximation gives the best result for variance estimation whereas the estimator underestimates the actual variance for any population (Wright and Stufken, 2011).
The estimated Coefficient of Variation for the estimator $\bar{y}_{HT}$ is given as

$$CV(\bar{y}_{HT}) = \frac{\sqrt{\hat{V}(\bar{y}_{HT})}}{\bar{y}_{HT}}$$

2.3. Generalized ratio-type estimator for unequal probability design

Kalidar and Bacanli (2008) mentioned the generalised ratio-type estimator of population mean for an unequal probability design in which i^{th} unit is selected with probability proportional to size π_i, and is given by

$$\bar{y}_{HT(R)} = \frac{\hat{\bar{y}}_{HT}}{\hat{\bar{x}}_{HT}}\,\bar{X} = \hat{\bar{R}}_{HT}\,\bar{X} \quad (7)$$

where $\hat{\bar{y}}_{HT} = \frac{1}{N}\sum_{i\epsilon s}\frac{y_i}{\pi_i}$ and $\hat{\bar{x}}_{HT} = \frac{1}{N}\sum_{i\epsilon s}\frac{x_i}{\pi_i}$ are Horvitz-Thompson estimators for the population means of the study and auxiliary variables respectively.
The approximate value of MSE of the estimator is given by

$$\text{MSE}\left(\bar{y}_{HT(R)}\right) \cong \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}{y_i}^*{y_j}^*$$

where ${y_i}^* = y_i - Rx_i$ with $R = \frac{\bar{y}}{\bar{x}}$
The estimate of this MSE is

$$\widehat{MSE}\left(\bar{y}_{HT(R)}\right) \cong \frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}\right)\frac{{y_i}^*{y_j}^*}{\pi_{ij}}$$

$$\widehat{MSE}\left(\bar{y}_{HT(R)}\right) \cong \frac{1}{N^2}\sum_i^n\sum_j^n\left(\frac{1}{\pi_i\pi_j} - \frac{1}{\pi_{ij}}\right){y_i}^*{y_j}^*$$

$$\cong \frac{1}{N^2}\sum_i^n\sum_j^n\left(\frac{1}{\pi_i\pi_j} - \frac{1}{\pi_{ij}}\right){y_i}^*{y_j}^* \quad (8)$$

Under SRS, due to Thompson (1990); we have
$\pi_i = \frac{n}{N}$ and $\pi_{ij} = \frac{n(n-1)}{N(N-1)}$
Therefore, for SRS, using Eq. (8), $\widehat{MSE}\ \left(\bar{y}_{HT(R)}\right)$ becomes

$$\widehat{MSE}\ \left(\bar{y}_{HT(R)}\right) \cong \frac{1}{N^2}\left(\frac{N^2}{n^2} - \frac{N(N-1)}{n(n-1)}\right)\sum_i^n\sum_j^n {y_i}^*{y_j}^*$$

Or, $\widehat{MSE}\left(\bar{y}_{HT(R)}\right) \cong \frac{n-N}{Nn^2(n-1)} \sum_i^n \sum_j^n y_i^* y_j^*$ (9)

with $y_i^* = y_i - \hat{\hat{R}} x_i$ with $\hat{\hat{R}} = \frac{\hat{\bar{y}}_{HT}}{\hat{\bar{x}}_{HT}}$

3. The Ratio-type Horvitz-Thompson Estimator in BSA plan

It is expected that the adjacent units (also call these as contiguous units) share the similar information, so one should avoid these contiguous units for selection in the sample. Although finding the correlation between study variable and auxiliary variable (which is assumed to be very high) is easy since either it is readily available for the population or might be quickly computed based upon the sampled data as the measurement of auxiliary variable is very cheap or easy, therefore the correlation between study variables and auxiliary variables has been used for the estimation. Therefore, the correlation could be used for the estimation of population parameters. Wright and Stufken (2011) used the Horvitz-Thompson estimator of population mean for BSA plan. Motivated by the estimator of Singh and Tailor (2003) in which gave ratio-type estimator for population mean, we proposed Horvitz-Thompson ratio type estimator using the correlation coefficient between the study variable Y and auxiliary variables X (i.e. ρ) for BSA plan as follows:

$$\bar{y}_B = \frac{\bar{y}_{HT}}{\bar{x}_{STHT}} \bar{X}_{ST} \tag{10}$$

where, $\bar{x}_{STHT} = \bar{x}_{HT} + \rho$ and $\bar{X}_{ST} = \bar{X} + \rho$

where, $\bar{X}_{ST}$ = Mean estimator of Single and Tailor (2003) which is the sum of average of auxiliary variables X and its correlation with population Y.

$\bar{x}_{HT}$ = Average of Horvitz Thompson mean estimator for sample combination of auxiliary variables

$\bar{X}$ = Mean of auxiliary variable

The MSE of $\bar{y}_B$ is given by

$$\text{MSE}\,(\bar{y}_B) \cong \frac{1}{N^2} \sum_i^N \sum_j^N \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right) y_i^{**} y_j^{**}$$

where $y_i^{**} = y_i - R x_i$ and $R = \frac{\bar{y}}{\bar{x}}$

The estimator of this MSE could be obtained by using $y_i^{**} = y_i - \hat{R} x_i$ in Horvitz-Thompson variance estimation formula

$$\widehat{MSE}\,(\bar{y}_B) \cong \frac{1}{N^2} \sum_i^n \sum_j^n \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right) \frac{y_i^{**} y_j^{**}}{\pi_{ij}}$$

where, $y_i^{**} = y_i - \hat{R} x_i$ and $\hat{R} = \frac{\bar{y}_{HT}}{\bar{x}_{STHT}}$

$$\widehat{MSE}\,(\bar{y}_B) \cong \frac{1}{N^2} \sum_i^n \sum_j^n \left(\frac{1}{\pi_i \pi_j} - \frac{1}{\pi_{ij}}\right) y_i^{**} y_j^{**}$$

substituting the values of $\pi_i = \frac{n}{N}$ and $\pi_{ij} = \frac{n(n-1)}{N(N-2m-1)}$, we get

$$\widehat{MSE}\,(\bar{y}_B) \cong \frac{1}{N^2} \left(\frac{N^2}{n^2} - \frac{N(N-2m-1)}{n(n-1)}\right) \sum_i^n \sum_j^n y_i^{**} y_j^{**}$$

or

$$\widehat{MSE}\,(\bar{y}_B) \cong \frac{n(2m+1) - N}{Nn^2(n-1)} \sum_i^n \sum_j^n y_i^{**} y_j^{**} \tag{11}$$

The estimated Coefficient of Variation of the estimator $\bar{y}_B$ is given as

$$CV(\bar{y}_B) = \frac{\sqrt{\widehat{MSE}\ (\bar{y}_B)}}{\bar{y}_B}$$

3.1. Relative Precisions

In this section, we compare the estimators $\bar{y}_{HT}$, $\bar{y}_{HT(R)}$ and $\bar{y}_B$ in terms of relative precision (RP).
RP of $\bar{y}_B$ with respect to $\bar{y}_{HT}$ using (6) and (11) is given by

$$\text{RP}_1 = \frac{\hat{V}(\bar{y}_{HT})}{\widehat{MSE}(\bar{y}_B)} = \frac{\frac{N-(2m+1)n}{2Nn^2(n-1)}\sum_{i\epsilon s}\sum_{j\epsilon s}(y_i - y_j)^2}{\frac{n(2m+1)-N}{Nn^2(n-1)}\sum_i^n \sum_j^n y_i^{**} y_j^{**}}$$

or

$$\text{RP}_1 = \frac{N-(2m+1)n}{2n(2m+1)-N}\frac{\sum_{i\epsilon s}\sum_{j\epsilon s}(y_i - y_j)^2}{\sum_i^n \sum_j^n y_i^{**} y_j^{**}} \tag{12}$$

From (12) we see that $N-(2m+1)n \geq 2n(2m+1)-N$ is equivalent to $N \geq \frac{3n}{2}(2m+1)$. Further, $\sum_{i\epsilon s}\sum_{j\epsilon s}(y_i - y_j)^2 > \sum_i^n \sum_j^n y_i^{**} y_j^{**}$.
Therefore, RP$_1$ > 1, when $N-(2m+1)n \geq 2n(2m+1)-N$, this makes the Horvitz-Thompson ratio-type estimator more efficient than the Horvitz-Thompson estimator for BSA plan.
The RP of $\bar{y}_B$ with respect to $\bar{y}_{HT(R)}$ using (9) and (11) is given by

$$\text{RP}_2 = \frac{\widehat{MSE}\ (\bar{y}_{HT(R)})}{\widehat{MSE}\ (\bar{y}_B)} = \frac{\delta \sum_i^n \sum_j^n y_i^* y_j^*}{\lambda \sum_i^n \sum_j^n y_i^{**} y_j^{**}}$$, where $\delta = \frac{n-N}{Nn^2(n-1)}$ and $\lambda = \frac{n(2m+1)-N}{Nn^2(n-1)}$

or

$$\text{RP}_2 = \frac{(n-N)\sum_i^n \sum_j^n y_i^* y_j^*}{(n(2m+1)-N)\sum_i^n \sum_j^n y_i^{**} y_j^{**}} \tag{13}$$

From (13), if N and m increases for fixed n, empirically $n(2m+1)-N$ decreases rapidly than $(n-N)$, so therefore $(n-N) > (n(2m+1)-N)$. Further,
$\sum_i^n \sum_j^n y_i^* y_j^* > \sum_i^n \sum_j^n y_i^{**} y_j^{**}$
Therefore, RP$_2$ >1 when $(n-N) > (n(2m+1)-N)$, this makes the Horvitz-Thompson ratio-type estimator for BSA plan more efficient than the Horvitz-Thompson ratio-type estimator for SRS.

4. Empirical examples

In this section, we elaborate three empirical examples in which population (Y) with its auxiliary variables (X). In the examples, the values of y-variate and x-variate are generated from the uniform population.

The sample selection has been done using the generator sample method of Hedayat et al. (1988b) and Mandal et al. (2008). The first and second order inclusion probabilities have been computed by equation (1) and (2) respectively. The population size N, sample size n, distance between two adjacent units m, y-values, x-values, possible number of sample combinations of size 3, the first order inclusion probability and the second order inclusion probability are given in the Table 1.

Table 1. Represents the population and sample in brief.

Population size N	Sample size n	Distance m	y-values	x-values	No of sample combinations of size 3	First order inclusion probability	Second order inclusion probability
17	3	2	11, 9, 15, 6, 1, 19, 10, 24, 30, 5, 13, 25, 27, 1, 4, 13, 23	52, 55, 61, 77, 69, 54, 63, 56, 80, 77, 66, 51, 77, 80, 64, 65, 74	34	3/17	1/34
25	3	3	28, 27, 7, 14, 16, 19, 6, 22, 10, 6, 39, 16, 25, 4, 19, 10, 20, 40, 12, 26, 38, 22, 26, 30, 14	99, 103, 78, 95, 105, 91, 79, 74, 100, 72, 79, 94, 86, 73, 91, 104, 72, 90, 86, 98, 74, 85, 75, 77, 84	75	3/25	1/75
33	3	4	5, 23, 30, 30, 20, 27, 11, 12, 8, 18, 28, 28, 22, 3, 17, 1, 10, 12, 14, 28, 27, 8, 9, 27, 25, 20, 18, 15, 22, 19, 23, 27, 24	72, 95, 70, 82, 97, 87, 91, 98, 86, 74, 82, 74, 86, 96, 81, 77, 84, 89, 98, 95, 100, 74, 71, 70, 79, 85, 85, 82, 72, 72, 73, 82, 72	132	1/11	1/132

After the computations, the results for the estimators $\bar{y}_{HT}$ and $\bar{y}_B$ are respectively given in the Table 2.

Table 2. Estimator, MSE/Variance and Coefficient of Variation of $\bar{y}_{HT}$ and $\bar{y}_B$ in BSA plan

Population size N	Sample size n	distance m	Population Mean	Population Variance	$\bar{y}_{HT}$		$\bar{y}_B$		Coefficient of Variation		RP_1
					Mean	Var	Mean	MSE	$\bar{y}_{HT}$	$\bar{y}_B$	
17	3	2	13.882	81.633	13.882	1.635	13.950	0.576	0.092	0.054	2.839
25	3	3	19.840	103.574	19.840	2.905	19.976	1.097	0.086	0.052	2.647
33	3	4	18.576	68.608	18.576	2.129	18.631	0.797	0.079	0.048	2.672

When population size $N = 17$, sample size $n = 3$ and the distance between two adjacent units $m = 2$, the number of combinations formed by SRS is 680, then the classical ratio-type estimator $\bar{y}_{HT(R)}$ is 13.964 and its MSE is 4.186. The RP_2 is computed as:

$RP_2 = \frac{\widehat{MSE}\ (\bar{y}_{HT(R)})}{\widehat{MSE}\ (\bar{y}_B)} = 7.268$

When the population size $N = 25$ and 33, and distance between two adjacent units $m = 4$ and 5, the number of combinations formed by SRS is 2,300 and 5,456 respectively for the fixed sample size n. The number of combinations is quite large when population size and distance increases, and therefore, it becomes difficult to calculate the variance of such a large number of combinations.

It has been observed that in all the cases, the RPs > 1. The value of Coefficient of Variation of $\bar{y}_B$ is also less than the value of Coefficient of Variation of $\bar{y}_{HT}$. The usefulness of auxiliary information remains valid as we see in these criterion indicates the suitability of using the proposed ratio-type Horvitz-Thompson estimator in place of Horvitz-Thompson estimator for BSA plan and the classical ratio-type estimator for SRS.

5. Conclusion

In this paper, the auxiliary information for improving the estimates of study variable under the Balanced Sampling plan excluding Adjacent units (BSA plan) has been used. Accordingly, the ratio-type Horvitz-Thompson estimator of population mean is proposed using the correlation coefficient between study and auxiliary variable for BSA plan. The correlation can either be positive or negative but might be of high value. This is the first attempt to establish the relationship between ratio estimator and BSA plan. The proposed estimator is superior to the existing Horvitz-Thompson estimator in terms of relative precision.

However, this suffers from some drawbacks that the BSA plan have limited combinations of N, n and m. Moreover, the sampling units should be in a circular arrangement for the sample selection. The number of combinations for BSA plan is much less than the number of combinations by SRS.

As the population size and distance between two adjacent units increases with fixed sample size, the relative precision (RP_1) of proposed estimator varies with no trend between the range of 2 and 3 whereas the value of Coefficient of Variation of estimator $\bar{y}_B$ decreases. When the population size $N = 17, 25$ and 33 and distance m = 2, 3 and 4 with the same sample size $n = 3$; the relative precision (RP_1) is 2.839, 2.647, 2.672 respectively, therefore these values show no trend. In all cases, RP_1 > 1 and $CV(\bar{y}_B)$ is less, this implies that the proposed ratio-type Horvitz-Thompson estimator using correlation coefficient is more efficient than Horvitz Thompson estimator for BSA plan. The numerical illustration given in Section 4 shows that the value of MSE of ratio-type Horvitz-Thompson estimator using correlation coefficient also varies with no trend between the range of 0 to 2, when population size N and distance m increases with the same sample size $n = 3$ but it is always less than the variance of Horvitz-Thompson estimator.

The proposed estimator can be used in many real life experiments such as (i) to find an average growth (say height or collar diameter) of a plant when the soil quality parameters such organic carbon, nitrogen, pH values are known. The adjacent units are considered as contiguous units as they might provide the similar information. The units could be selected at distance of 2, 3 or 4. In this way, an appropriate average growth of a plant can be estimated with slightly different soil quality parameters. (ii) the prediction of rainfall can be done with the help of given humidity in an air. The humidity can be recorded every two, three or four hours of a day. The prediction of rainfall would be interpreted. (iii) the mean production of tree lumber can be estimated when the diameter of trees (DBH, diameter at breast height) is known. It would be difficult to measure the diameter of all the trees, so trees could be selected at a distance of 2, 3 or 4. It is easy to measure the DBH of trees and can be used as auxiliary information for estimating the mean production of tree lumber. The proposed estimator has a potential to exhibit the relationship between study variable and auxiliary variable.

Acknowledgment: Authors are grateful to the two unknown reviewers and Editor for their constructive comments for the improvement of this paper.

References

[1] Chutiman, N., and Chiangpradit, M. (2014): Ratio estimator in adaptive cluster sampling without replacement of networks. Journal of Probability and Statistics, 2014.

[2] Cochran, W. G. (1977): Sampling Techniques, Third Edition, Wiley Eastern Limited.

[3] Colbourn, C. J., and Ling, A. C. H. (1998): A class of partial triple systems with applications in survey sampling. Communications in Statistics - Theory and Methods, 27(4), 1009-1018.

[4] Colbourn, C. J., and Ling, A. C. H. (1999): Balanced sampling plans with block size four excluding contiguous units. Australasian Journal of Combinatorics, 20, 37-46.

[5] de Paschal, J. N. (1961): Unbiased ratio estimators in stratified sampling. Journal of the American Statistical Association, 56(293), 70-87.

[6] Dryver, A. L., and Chao, C. T. (2007): Ratio estimators in adaptive cluster sampling. Environmetrics: The Official Journal of the International Environmetrics Society, 18(6), 607-620.

[7] Goodman, L. A., and Hartley, H. O. (1958): The precision of unbiased ratio-type estimators. Journal of the American Statistical Association, 53(282), 491-508.

[8] Gopinath, P. P., Parsad, R., & Mandal, B.N. (2018): Two-dimensional balanced sampling plans excluding adjacent units under sharing a border and island adjacency schemes, Communications in Statistics - Simulation and Computation, 47(3), 712-720.

[9] Hartley, H. O., and Ross, A. (1954): Unbiased ratio estimators. Nature, 174(4423), 270-271.

[10] Hedayat, A. S., Rao, C. R., and Stufken, J. (1988): 24 Designs in survey sampling avoiding contiguous units. Handbook of statistics, 6, 575-583.

[11] Hedayat, A. S., Rao, C. R., and Stufken, J. (1988): Sampling plans excluding contiguous units. Journal of Statistical Planning and Inference, 19(2), 159-170.

[12] Horvitz, D. G., and Thompson, D. J. (1952): A generalization of sampling without replacement from a finite universe. Journal of the American statistical Association, 47(260), 663-685.

[13] Kadilar, C., Unyazici, Y., and Cingi, H. (2009): Ratio estimator for the population mean using ranked set sampling. Statistical Papers, 50, 301-309.

[14] Kalidar, C., and Bacanli, S. (2008): Ratio estimators with unequal probability designs. Pakistan Journal of Statistics, 24(3), 167-172.

[16] Kumar, R., Parsad, R., and Mandal, B. N. (2016): Smaller Balanced Sampling Plans Excluding Adjacent Units for One Dimensional Population. International Journal of Computational and Theoretical Statistics, 3(2), 55-61.

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

[18] Mandal, B. N. (2007): Combinatorics and its applications with special reference to sample surveys. Unpublished Ph. D. Thesis, Indian Agricultural Research Institute, New Delhi.

[19] Mandal, B. N., Gupta, V. K., and Parsad, R. (2011): Construction of polygonal designs using linear integer programming. Communications in Statistics-Theory and Methods, 40(10), 1787-1794.

[20] Mandal, B. N., Parsad, R., and Gupta, V. K. (2008): Computer-aided construction of balanced sampling plans excluding contiguous units. Journal of Statistics and Applications, 3, 59-85.

[21] Mickey, M. R. (1959): Some finite population unbiased ratio and regression estimators. Journal of the American Statistical Association, 54(287), 594-612.

[22] Midzuno, H. (1952): On the sampling system with probability proportionate to sum of sizes. Annals of Institute of Statistical Mathematics, 3(2), 99-108.

[23] Nanjamma, N. S., Murthy, M. N., and Sethi, V. K. (1959): Some sampling systems providing unbiased ratio estimators. Sankhyā: The Indian Journal of Statistics, 299-314.

[24] Olkin, I. (1958): Multivariate ratio estimation for finite populations. Biometrika, 45, 154-165.

[25] Pathak, P. K. (1964): On sampling schemes providing unbiased ratio estimators. The Annals of Mathematical Statistics, 222-231.

[26] Raj, D. (1954): Ratio estimation in sampling with equal and unequal probabilities. Journal of the Indian Society of Agricultural Statistics, 6(2), 127-138.

[27] Raj, D. (1964): On double sampling for pps estimation. The Annals of Mathematical Statistics, 35(2), 900-902.

[28] Raj, D. (1965): On sampling over two occasions with probability proportionate to size. The Annals of Mathematical Statistics, 36(1), 327-330.

[29] Rajeev, K., Animesh, K., and Moury, P. K. (2019): Balanced sampling plans excluding adjacent units for linear ordering of population units. International Research Journal of Agricultural Economics and Statistics, 10(2), 272-277.

[30] Rao, J. N. K., and Nigam, A. K. (1990): Optimum controlled sampling designs. Biometrika, **77**, 807-814.

[31] Rao, J. N. K., and Nigam, A. K. (1992): Optimal controlled sampling: A unified approach. International Statistical Review, 60, 89-98.

[32] Robson, D. (1957): Applications of multivariate polykays to the theory of unbiased ratio-type estimation. Journal of the American Statistical Association, 52(280), 511-522.

[33] Robson, D. S., and Vithatasai, C. (1961): Unbiased componentwise ratio estimation. Journal of the American Statistical Association, 56(294), 350-358.

[34] Singh, H. P., and Tailor, R. (2003): Use of known correlation coefficient in estimating the finite population mean. Statistics in transition, 6(4), 555-560.

[35] Stufken, J. (1993): Combinatorial and statistical aspects of sampling plans to avoid the selection of adjacent units. Journal of Combinatorics Information and System Science, 18, 81-92.

[36] Stufken, J., and Wright, J. H. (2001): Polygonal designs with blocks of size $k \leq 10$. Metrika, 54, 179-184.

[37] Stufken, J., Song, S. Y., See, K., and Driessel, K. R. (1999): Polygonal designs: Some existence and non-existence results. Journal of statistical planning and inference, 77(1), 155-166.

[38] Sukhatme, B. V. (1962): Some ratio-type estimators in two-phase sampling. Journal of the American Statistical Association, 57(299), 628-632.

[39] Tahir, M. H., and Iqbal, I. (2012): Polygonal Designs with Block Size 3 and Single Distance. Hacettepe journal of Mathematics and Statistics, 41(4), 587-604.

[40] Tahir, M. H., Iqbal, I., Akhtar, M., and Shabbir, J. (2010): Cyclic polygonal designs with block size 3 and λ= 1 for joint distance α= 6 to 16. Journal of Statistical Theory and Practice, 4(2), 203-220.

[41] Tiwari, N., Banerjie, J., Chandra, G., and Bhari, S. (2023): Adaptive cluster sampling based on balanced sampling plan excluding contiguous units. Communications in Statistics-Theory and Methods, 1-14.

[42] Wang, X., Feng, T., Zhang, J., and Xu, Y. (2015): Two-dimensional balanced sampling plans avoiding adjacent units. Discrete Mathematics, 338(9), 1624-1642.

[43] Williams, W. H. (1961): Generating unbiased ratio and regression estimators. Biometrics, 17(2), 267-274.

[44] Williams, W. H. (1962): On two methods of unbiased estimation with auxiliary variates. Journal of the American Statistical Association, 57(297), 184-186.

[45] Wright, J. H., and Stufken, J. (2008): New balanced sampling plans excluding adjacent units. Journal of Statistical Planning and Inference, 138(11), 3326-3335.

[47] Wright, J. H., and Stufken, J. (2011): Variance approximation under balanced sampling plans excluding adjacent units. Journal of Statistical Theory and Practice, 5(1), 147-160.