STATISTICS IN TRANSITION-new series, Spring 2013
Vol. 14, No. 1, pp. 57–74

BEST LINEAR UNBIASED ESTIMATORS OF POPULATION MEAN ON CURRENT OCCASION IN TWO-OCCASION ROTATION PATTERNS

G. N. Singh, S. Prasad[1]

ABSTRACT

Best linear unbiased estimators have been proposed to estimate the population mean on current occasion in two-occasion successive (rotation) sampling. Behavior of the proposed estimators have been studied and their respective optimum replacement policies are discussed. Empirical studies are carried out to examine the performance of the proposed estimators and consequently the suitable recommendations are made.

Key words: successive sampling, auxiliary information, unbiased, variance, optimum replacement policy.

1. Introduction

Often in sample surveys on successive occasions for the same population, the current or most recent estimates are of the primary interest if the characteristics of the population are likely to change rapidly over time. For example, monthly surveys are carried out to collect data on prices of goods to determine the consumer price index, labor force surveys are conducted on monthly basis to estimate the numbers of people in employment and industries, collect information at regular intervals to know popularity of their products, etc. In such studies, successive (rotation) sampling may be an impressive statistical tool to generate reliable and cost effective estimates of different population parameters on successive points of time (occasions) in chronological order. It also provides effective estimates of changing patterns over a period of time.

The problem of successive (rotation) sampling with a partial replacement of sampling units was initiated by Jessen (1942) in the analysis of agricultural survey data. He pioneered using the entire information collected during the previous investigations. The theory of successive (rotation) sampling was further

[1] Department of Applied Mathematics, Indian School of Mines, Dhanbad-826004, India. E-mail: gnsingh_ism@yahoo.com.

extended by Patterson (1950), Rao and Graham (1964), Gupta (1979), Das (1982) and Chaturvedi and Tripathi (1983), among others. Sen (1971) applied this theory with success in designing the strategies for estimating the population mean on the current occasion using information on two auxiliary variables readily available on the previous occasion. Sen (1972, 1973) extended his work for several auxiliary variables. Singh *et al.* (1991) and Singh and Singh (2001) used the auxiliary information available on the current occasion and proposed estimators for the current population mean in two-occasion successive (rotation) sampling. Singh (2003) generalized his work for h-occasion successive sampling.

In many situations, information on an auxiliary variable may be readily available on the first as well as on the second occasion; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation, number of beds in different hospitals may be known in hospital surveys, number of polluting industries and vehicles is known in environmental surveys, nature of employment status, educational status, food availability and medical aids of a locality is well known in advance for estimating various demographic parameters in demographic surveys. Utilizing auxiliary information on both the occasions, Feng and Zou (1997), Biradar and Singh (2001), Singh (2005), Singh and Priyanka (2006, 2007, 2008), Singh and Karna (2009a, b) have proposed several estimators for estimating the population mean on current (second) occasion in two-occasion successive (rotation) sampling. Recently Singh and Vishwakarma (2009) have suggested a general estimation procedure for population mean in successive (rotation) sampling. Motivated with the above works and utilizing the information on an auxiliary variable, readily available on both the occasions, we have proposed best linear unbiased estimators for estimating the current population mean in two-occasion successive (rotation) sampling. Behaviors of the proposed estimators are examined through empirical means of comparison and subsequently the suitable recommendations are made.

2. Sample structures and notations on two occasions

Let $U = (U_1, U_2, - - -, U_N)$ be the finite population of N (large) units which is assumed to remain unchanged over two occasions. Let x (y) be the character under study on the first (second) occasion respectively. It is assumed that the information on an auxiliary variable z (stable over occasion), is readily available for both the occasions, whose population mean is known and it is highly positively correlated to x and y on the first and second occasions respectively. A simple random sample (without replacement) of size n units is drawn on the first occasion and a random sub-sample of size $m = n\lambda$ units from the sample on the first occasion is retained (matched) for its use on the current (second) occasion. A fresh (un-matched) sample of size $u = (n\text{-}m) = n\mu$ units is drawn on the current occasion from the entire population by simple random sampling (without replacement) method so that the sample size on the current occasion is also n. λ and μ ($\lambda+\mu=1$) are the fractions of the matched and fresh samples, respectively, on the current occasion. We consider the following notations for further use:

$\bar{X}$, $\bar{Y}$, $\bar{Z}$: Population means of the variables x, y and z respectively.

$\bar{x}_m$, $\bar{x}_n$, $\bar{y}_u$, $\bar{y}_m$, $\bar{z}_u$, $\bar{z}_m$, $\bar{z}_n$: Sample means of the respective variables based on the sample sizes shown in suffices.

ρ_{yx}, ρ_{yz}, ρ_{xz}: Correlation coefficients between the variables shown in suffices.

$S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}(x_i-\bar{X})^2$: Population mean square of x.

S_y^2, S_Z^2 : Population mean squares of y and z respectively.

3. Formulation of the estimator

To estimate the population mean $\bar{Y}$ on the current (second) occasion, we consider the following minimum variance linear unbiased estimator of $\bar{Y}$, which is as follows:

$$T_1 = \left\{a_1\bar{y}_u + a_2\bar{y}_m\right\} + \left\{a_3\bar{x}_m + a_4\bar{x}_n\right\} + \left\{a_5\bar{z}_u + a_6\bar{z}_m + a_7\bar{z}_n + a_8\bar{Z}\right\} \tag{1}$$

where a_1, a_2, a_3, a_4, a_5, a_6, a_7 and a_8 are constants to be determined so that

(i) T_1 becomes an unbiased estimator of $\bar{Y}$ and

(ii) the variance of T_1 attains a minimum value.

For unbiasedness condition, we must have

$$\left(a_1+a_2\right) = 1,\ \left(a_3+a_4\right) = 0 \text{ and} \left(a_5+a_6+a_7+a_8\right) = 0\,.$$

Substituting $a_1 = \varphi_1$, $a_3 = \beta_1$ and $a_8 = -\left(a_5 + a_6 + a_7\right)$, the estimator T_1 defined in equation (1) reduces to the following form

$$T_1 = \left\{\varphi_1\bar{y}_u + \left(1-\varphi_1\right)\bar{y}_m\right\} + \beta_1\left\{\bar{x}_m-\bar{x}_n\right\} + \left\{a_5\left(\bar{z}_u-\bar{Z}\right) + a_6\left(\bar{z}_m-\bar{Z}\right) + a_7\left(\bar{z}_n-\bar{Z}\right)\right\}$$

$$= \varphi_1\left[\bar{y}_u + k_1\left(\bar{z}_u-\bar{Z}\right)\right] + \left(1-\varphi_1\right)\left[\bar{y}_m + k_2\left\{\bar{x}_m-\bar{x}_n\right\} + k_3\left(\bar{z}_m-\bar{Z}\right) + k_4\left(\bar{z}_n-\bar{Z}\right)\right]$$

$$= \varphi_1 T_{1u} + \left(1-\varphi_1\right)T_{1m} \tag{2}$$

where $T_{1u} = \bar{y}_u + k_1\left(\bar{z}_u-\bar{Z}\right)$; an estimator based on the fresh sample of size u and $T_{1m} = \bar{y}_m + k_2\left(\bar{x}_m - \bar{x}_n\right) + k_3\left(\bar{z}_m - \bar{Z}\right) + k_4\left(\bar{z}_n - \bar{Z}\right)$; an estimator based on the matched sample of size m, $k_1 = \frac{a_5}{\varphi_1}$, $k_2 = \frac{\beta_1}{1-\varphi_1}$, $k_3 = \frac{a_6}{1-\varphi_1}$, $k_4 = \frac{a_7}{1-\varphi_1}$ and φ_1 are the unknown constants to be determined under certain criterions.

Remark 3.1. For estimating the population mean on each occasion the estimator T_{1u} is suitable, which implies that more belief on T_{1u} could be shown by choosing φ_1 as 1 (or close to 1), while for estimating the change over the occasions, the estimator T_{1m} could be more useful and hence φ_1 might be chosen as 0 (or close to 0). For asserting both the problems simultaneously, the suitable (optimum) choice of φ_1 is desired.

4. Properties of the estimator T_1

T_1 is an unbiased estimator of $\bar{Y}$ whose variance, ignoring finite population corrections, is derived in the following theorem.

Theorem 4.1. Variance of the estimator T_1 is obtained as

$$V(T_1) = \varphi_1^2\, V\left(T_{1u}\right) + \left(1-\varphi_1\right)^2 V\left(T_{1m}\right) \tag{3}$$

where

$$V(T_{1u}) = \frac{1}{u}\eta_1 S_y^2 \tag{4}$$

$$V(T_{1m}) = \left[\frac{1}{m}\eta_2 + \left(\frac{1}{m} - \frac{1}{n}\right)\eta_3 + \frac{1}{n}\eta_4\right]S_y^2 \tag{5}$$

$$\eta_1 = \left(1 + k_1^2 + 2k_1\rho_{yz}\right),\ \eta_2 = \left(1 + k_3^2 + 2k_3\rho_{yz}\right),$$

$$\eta_3 = \left(k_2^2 + 2k_2\rho_{yx} + 2k_2k_3\rho_{xz}\right) \text{ and } \eta_4 = \left(k_4^2 + 2k_4\rho_{yz} + 2k_3k_4\right).$$

Proof: It is obvious that the variance of the estimator T_1 is given by

$$V(T_1) = E\left[T_1 - \bar{Y}\right]^2 = E\left[\varphi_1\left(T_{1u} - \bar{Y}\right) + \left(1-\varphi_1\right)\left(T_{1m} - \bar{Y}\right)\right]^2$$

$$= \varphi_1^2\, V(T_{1u}) + \left(1-\varphi_1\right)^2 V\left(T_{1m}\right) + 2\varphi_1\left(1-\varphi_1\right)C_{11} \tag{6}$$

where

$$V\left(T_{1u}\right) = E\left[T_{1u} - \bar{Y}\right]^2,\ V(T_{1m}) = E\left[T_{1m} - \bar{Y}\right]^2 \text{ and } C_{11} = E\left[\left(T_{1u} - \bar{Y}\right)\left(T_{1m} - \bar{Y}\right)\right]$$

Substituting the expressions of T_{1u} and T_{1m} in equation (6), taking expectations and ignoring finite population corrections, we have the expression of the variance of the estimator T_1 as given in equation (3).

Remark 4.1. Results in equation (3) are derived under the assumption that the population mean squares of the variables x, y and z are almost equal.

Remark 4.2. T_{1u} and T_{1m} are based on two independent samples of sizes u and m respectively and they are unbiased estimators of $\bar{Y}$, hence the covariance term C_{11} between T_{1u} and T_{1m} vanishes.

5. Minimum variance of the estimator T_1

Since the variance of the estimator T_1 in equation (3) is the function of the unknown constants k_1, k_2, k_3, k_4 and φ_1, therefore it is minimized with respect to these constants, and subsequently the optimum values of k_1, k_2, k_3, k_4 and φ_1 are obtained as

$$k_1^* = -\rho_{yz} \tag{7}$$

$$k_2^* = \frac{\rho_{yz}\rho_{xz} - \rho_{yx}}{1-\rho_{xz}^2} \tag{8}$$

$$k_3^* = \frac{\rho_{yx}\rho_{xz} - \rho_{yz}}{1-\rho_{xz}^2} \tag{9}$$

$$k_4^* = \frac{\rho_{xz}\left(\rho_{yz}\rho_{xz} - \rho_{yx}\right)}{1-\rho_{xz}^2} \tag{10}$$

$$\varphi_{1opt} = \frac{V(T_{1m})}{V(T_{1u}) + V(T_{1m})} \tag{11}$$

Substituting the values of k_1^*, k_2^*, k_3^* and k_4^* in equations (4) and (5), we get the optimum variances of T_{1u} and T_{1m} as

$$V(T_{1u})_{opt} = \frac{1}{u} A_1 S_y^2 \tag{12}$$

$$V(T_{1m})_{opt} = \left[\frac{1}{m}A_2 + \left(\frac{1}{m} - \frac{1}{n}\right)A_3 + \frac{1}{n}A_4\right]S_y^2 \tag{13}$$

where $A_1 = 1-\rho_{yz}^2$, $A_2 = \dfrac{1+\rho_{xz}^2\left(\rho_{xz}^2 - 2 + \rho_{yx}^2 + 2\rho_{yz}^2 - 2\rho_{yz}\rho_{yx}\rho_{xz}\right) - \rho_{yz}^2}{\left(1-\rho_{xz}^2\right)^2}$, $A_3 = -k_2^{*2}$

and $A_4 = -k_4^{*2}$.

Further, substituting the values of $V(T_{1u})_{opt}$ and $V(T_{1m})_{opt}$ from equations (12) and (13) in equation (11), we get the optimum value $\varphi_{1_{opt}}$ with respect to k_1^*, k_2^*, k_3^* and k_4^* as

$$\varphi_{1opt}^* = \frac{V(T_{1m})_{opt}}{V(T_{1u})_{opt}+V(T_{1m})_{opt}} \tag{14}$$

Again from equation (14) substituting the value of φ_{1opt}^* in equation (3), we get the optimum variance of T_1 as

$$V\left(T_1\right)_{opt} = \frac{V\left(T_{1m}\right)_{opt} . V\left(T_{1u}\right)_{opt}}{V\left(T_{1u}\right)_{opt}+V\left(T_{1m}\right)_{opt}} \tag{15}$$

Further, substituting the values from equations (12) and (13) in equations (14) and (15), the simplified values of φ_{1opt}^* and $V\left(T_1\right)_{opt}$ are obtained as

$$\varphi_{1opt}^* = \left[\frac{\mu_1\left(A_5+\mu_1 A_6\right)}{A_1+\mu_1 A_7+\mu_1^2 A_6}\right] \tag{16}$$

$$V(T_1)_{opt} = \frac{1}{n}\left[\frac{A_1\left(A_5+\mu_1 A_6\right)}{A_1+\mu_1 A_7+\mu_1^2 A_6}\right] S_y^2 \tag{17}$$

where $A_5 = A_2+A_4$, $A_6 = A_3-A_4$, $A_7 = A_5-A_1$ and μ_1 is the fraction of fresh sample for the estimator T_1.

6. Optimum replacement policy

To determine the optimum value of μ_1 (fraction of a sample to be drawn afresh on the current occasion) so that the population mean $\bar{Y}$ may be estimated with the maximum precision, we minimize the $V\left(T_1\right)_{opt}$ given in equation (17) with respect to μ_1, which result in a quadratic equation in μ_1 and respective solutions of μ_1 say μ_1^0 is given below:

$$Q_1\mu_1^2+2Q_2\mu_1+Q_3=0 \tag{18}$$

$$\mu_1^0=\frac{-Q_2\pm\sqrt{Q_2^2-Q_1Q_3}}{Q_1} \tag{19}$$

where $Q_1 = A_6^2$, $Q_2 = A_5A_6$ and $Q_3 = A_7A_5-A_1A_6$.

From equation (19), it is obvious that the real values of μ_1^0 exist if the quantity under square root is greater than or equal to zero. Two real values of μ_1^0 are possible. Hence, while choosing the value of μ_1^0, it should be remembered that $0 \le \mu_1^0 \le 1$. All other values of μ_1^0 are inadmissible. Substituting the admissible value of μ_1^0 say $\hat{\mu}_1$ from equation (19) into equation (17), we have the optimum value of $V(T_1)_{opt}$ as

$$V(T_1^0)_{opt} = \frac{1}{n}\left[\frac{A_1(A_5+\hat{\mu}_1 A_6)}{A_1+\hat{\mu}_1 A_7+\hat{\mu}_1^2 A_6}\right]S_y^2 \tag{20}$$

7. Efficiency comparison

To study the performance of the estimator T_1 the percent relative efficiencies of the estimator T_1 with respect to (i) $\bar{y}_n$, when there is no matching, and (ii) the estimator T_2, when no auxiliary information is used at any occasion, have been computed for different choices of correlations. The estimator T_2 is defined under the same circumstances as the estimator T_1, but in the absence of the auxiliary variable z on both the occasions and proposed as

$$T_2 = \{b_1\bar{y}_u + b_2\bar{y}_m\} + \{b_3\bar{x}_m + b_4\bar{x}_n\} \tag{21}$$

where b_1, b_2, b_3 and b_4 are constants to be determined so that

(i) T_2 becomes an unbiased estimator of $\bar{Y}$ and

(ii) The variance of T_2 attains a minimum value.

For unbiasedness condition, we must have $(b_1+b_2) = 1$ and $(b_3+b_4) = 0$.

Substituting $b_1 = \varphi_2$ and $b_3 = \beta_2$, the estimator T_2 defined in equation (21) reduces to the following form

$$\begin{aligned} T_2 &= \{\varphi_2\bar{y}_u+(1-\varphi_2)\bar{y}_m\}+\beta_2\{\bar{x}_m-\bar{x}_n\} \\ &= \varphi_2\bar{y}_u+(1-\varphi_2)\left[\bar{y}_m+k_5\{\bar{x}_m-\bar{x}_n\}\right] \\ &= \varphi_2 T_{2u}+(1-\varphi_2)T_{2m} \end{aligned} \tag{22}$$

where $T_{2u} = \bar{y}_u$; an estimator based on the fresh sample of size u and $T_{2m} = \bar{y}_m + k_5(\bar{x}_m - \bar{x}_n)$; an estimator based on the matched sample of size m, $k_5 = \frac{\beta_2}{1-\varphi_2}$ and φ_2 are the unknown constants to be determined in such a way that they minimize the variance of the estimator T_2. Following the methods discussed in Sections 4, 5 and 6, the optimum values of k_5, μ_2 (fraction of fresh sample for the estimator T_2), variance of $\bar{y}_n$ and optimum variance of T_2 for large N are given by

$$k_5^* = -\rho_{yx} \tag{23}$$

$$\hat{\mu}_2 = \frac{1 \pm \sqrt{1-\rho_{yx}^2}}{\rho_{yx}^2} \tag{24}$$

$$V(\bar{y}_n) = \frac{S_y^2}{n} \tag{25}$$

$$V(T_2^0)_{opt} = \frac{1}{n}\left[\frac{1 - \hat{\mu}_2\rho_{yx}^2}{1 - \hat{\mu}_2^2\rho_{yx}^2}\right]S_y^2 \tag{26}$$

For different choices of ρ_{yx}, ρ_{xz} and ρ_{yz}, Table 1 shows the optimum values of μ_1 and percent relative efficiencies E_1 and E_2 of the estimator T_1 with respect to the estimators $\bar{y}_n$ and T_2 respectively, where

$$E_1 = \frac{V(\bar{y}_n)}{V(T_1^0)_{opt}} \times 100 \quad \text{and} \quad E_2 = \frac{V(T_2^0)_{opt}}{V(T_1^0)_{opt}} \times 100.$$

8. Analysis of results for estimator T_1

The following conclusions can be read out from Table 1.

(a) For fixed values of ρ_{xz} and ρ_{yz}, the values of μ_1 and E_1 are increasing with the increasing values of ρ_{yx}. The values of E_2 are decreasing for the lower values of ρ_{yx} while increasing pattern may be seen for the higher values of ρ_{yx}.

(b) For fixed values of ρ_{xz} and ρ_{yx}, the values of μ_1 are decreasing with the increasing values of ρ_{yz}.. Values of E_1 and E_2 are increasing with the increasing values of ρ_{yz}..This behavior is highly desirable, since it concludes that if highly correlated auxiliary variable is available, it pays in terms of enhance precision of the estimates as well as it reduces the cost of the survey.

(c) For fixed values of ρ_{yz} and $\rho_{yx,}$ the values of μ_1 are decreasing with the increasing values of ρ_{xz} . Similar patterns are visible for the efficiencies E_1 and E_2.

(d) Minimum value of μ_1 is 0.4329, which indicates that only 43 percent of the total sample size is to be replaced on the current occasion for the corresponding choices of the correlations.

9. Use of auxiliary variable only at the current occasion

In section 3 we have formulated the estimator T_1 on the assumption that information on a stable auxiliary variable z was readily available on both the occasions. If the duration between two successive occasions is small then one may expect the stability of the auxiliary variable but the stability character of the auxiliary variable may not sustain if the duration between two successive occasions is appreciably large. In such situation it may not be wise to use the auxiliary information from the previous occasion. Motivated with the above argument, we formulate the estimator T_3 when the information on an auxiliary variable z is available only on the current (second) occasion. The estimator T_3 is formulated as

$$T_3 = \left\{c_1\bar{y}_u + c_2\bar{y}_m\right\} + \left\{c_3\bar{x}_m + c_4\bar{x}_n\right\} + \left\{c_5\bar{z}_u + c_6\bar{z}_m + c_7\bar{Z}\right\} \qquad (27)$$

where c_1, c_2, c_3, c_4, c_5, c_6 and c_7 are constants to be determined so that

(i) T_3 becomes an unbiased estimator of $\bar{Y}$ and

(ii) The variance of T_3 attains a minimum value.

For unbiasedness condition, we must have

$$\left(c_1+c_2\right) = 1,\ \left(c_3+c_4\right) = 0 \text{ and } \left(c_5+c_6+c_7\right) = 0.$$

Substituting $c_1 = \varphi_3$, $c_3 = \beta_3$ and $c_7 = -\left(c_5 + c_6\right)$, the estimator T_3 defined in equation (27) reduces to the following form

$$T_3 = \left\{\varphi_3\bar{y}_u + \left(1-\varphi_3\right)\bar{y}_m\right\} + \beta_3\left\{\bar{x}_m - \bar{x}_n\right\} + \left\{c_5\left(\bar{z}_u - \bar{Z}\right) + c_6\left(\bar{z}_m - \bar{Z}\right)\right\}$$

$$= \varphi_3\left[\bar{y}_u + l_1\left(\bar{z}_u - \bar{Z}\right)\right] + \left(1-\varphi_3\right)\left[\bar{y}_m + l_2\left\{\bar{x}_m - \bar{x}_n\right\} + l_3\left(\bar{z}_m - \bar{Z}\right)\right]$$

$$= \varphi_3 T_{3u} + \left(1-\varphi_3\right)T_{3m} \qquad (28)$$

where $T_{3u} = \bar{y}_u + l_1(\bar{z}_u - \bar{Z})$; an estimator based on the fresh sample of size u and $T_{3m} = \bar{y}_m + l_2(\bar{x}_m - \bar{x}_n) + l_3(\bar{z}_m - \bar{Z})$; an estimator based on the matched sample of size m, $l_1 = \frac{c_5}{\varphi_3}$, $l_2 = \frac{\beta_3}{1-\varphi_3}$, $l_3 = \frac{c_6}{1-\varphi_3}$ and φ_3 are the unknown constants to be determined under certain criterions.

9.1. Properties of the estimator T_3

T_3 is an unbiased estimator of $\bar{Y}$ whose variance is given in the following theorem.

Theorem 9.1. Variance of the estimator T_3 is obtained as

$$V(T_3) = \varphi_3^2\, V(T_{3u}) + (1-\varphi_3)^2\, V(T_{3m}) \tag{29}$$

where
$$V(T_{3u}) = \frac{1}{u}\left(1 + l_1^2 + 2l_1\rho_{yz}\right)S_y^2 \tag{30}$$

$$V(T_{3m}) = \left[\frac{1}{m}\left(1 + l_3^2 + 2l_3\rho_{yz}\right) + \left(\frac{1}{m} - \frac{1}{n}\right)\left(l_2^2 + 2l_2\rho_{yx} + 2l_2l_3\rho_{xz}\right)\right]S_y^2 \tag{31}$$

Proof: It is obvious that the variance of the estimator T_3 is given by

$$V(T_3) = E\left[T_3 - \bar{Y}\right]^2 = E\left[\varphi_3\left(T_{3u} - \bar{Y}\right) + (1-\varphi_3)\left(T_{3m} - \bar{Y}\right)\right]^2$$
$$= \varphi_3^2\, V(T_{3u}) + (1-\varphi_3)^2\, V(T_{3m}) + 2\varphi_3(1-\varphi_3)R_{11} \tag{32}$$

where $V(T_{3u}) = E\left[T_{3u} - \bar{Y}\right]^2$, $V(T_{3m}) = E\left[T_{3m} - \bar{Y}\right]^2$ and

$R_{11} = E\left[\left(T_{3u} - \bar{Y}\right)\left(T_{3m} - \bar{Y}\right)\right]$

Substituting the expressions of T_{3u} and T_{3m} in equation (32), taking expectations and ignoring finite population corrections, we have the expression of the variance of T_3 as given in equation (29).

Remark 9.1. Results in theorem 9.1 is derived similar to the results obtained in theorem 4.1.

9.2. Minimum variance of the estimator T_3

Since the variance of the estimator T_3 in equation (29) is the function of the unknown constants l_1, l_2, l_3 and φ_3, therefore it is minimized with respect to these constants and subsequently the optimum values of l_1, l_2, l_3 and φ_3 are obtained as

$$l_1^* = -\rho_{yz} \tag{33}$$

$$l_2^* = \frac{\rho_{yz}\ \rho_{xz} - \rho_{yx}}{1-\mu_3\rho_{xz}^2} \tag{34}$$

$$l_3^* = \frac{\mu_3\rho_{yx}\rho_{xz} - \rho_{yz}}{1-\mu_3\rho_{xz}^2} \tag{35}$$

$$\varphi_{3_{opt}} = \frac{V\left(T_{3m}\right)}{V\left(T_{3u}\right)+V\left(T_{3m}\right)} \tag{36}$$

Now, substituting the values of l_1^*, l_2^* and l_3^* in equations (30) and (31), we get the optimum variances of T_{3u}, T_{3m} as

$$V\left(T_{3u}\right)_{opt} = \frac{1}{u}B_1 S_y^2 \tag{37}$$

$$V\left(T_{3m}\right)_{opt} = \frac{1}{m}\left[\frac{B_1+\mu_3 B_5+\mu_3^2 B_2}{\left(1-\mu_3\rho_{xz}^2\right)^2}\right]S_y^2 \tag{38}$$

where $B_1 = 1-\rho_{yz}^2$, $B_2 = \rho_{xz}^2\left(\rho_{xz}^2+\rho_{yx}^2-2\rho_{yz}\rho_{yx}\rho_{xz}\right)$, $B_3 = -2\rho_{xz}^2 B_1$, $B_4 = -\left(\rho_{yz}\rho_{xz}-\rho_{yx}\right)^2$ and $B_5 = B_3+B_4$.

Further, substituting the values of $V(T_{3u})_{opt}$ and $V(T_{3m})_{opt}$ from equations (37) and (38) in equation (36), we get the optimum value $\varphi_{1_{opt}}$ with respect to l_1^*, l_2^* and l_3^* as

$$\varphi_{3opt}^* = \frac{V\left(T_{3m}\right)_{opt}}{V\left(T_{3u}\right)_{opt}+V\left(T_{3m}\right)_{opt}} \tag{39}$$

Again, from equation (39) substituting the value of φ_{3opt}^* in equation (29), we get the optimum variance of T_3 as

$$V(T_3)_{opt} = \frac{V(T_{3m})_{opt}\ V(T_{3u})_{opt}}{V(T_{3m})_{opt} + V(T_{3u})_{opt}} \tag{40}$$

Further, substituting the values from equations (37) and (38) in equations (39) and (40), the simplified values of φ^*_{3opt} and $V(T_3)_{opt}$ are obtained as

$$\varphi^*_{3opt} = \left[\frac{\mu_3(B_1 + \mu_3 B_5 + \mu_3^2 B_2)}{\mu_3^3 B_6 + \mu_3^2 B_7 + \mu_3 B_3 + B_1}\right] \tag{41}$$

$$V(T_3)_{opt} = \frac{1}{n}\left[\frac{B_8 + \mu_3 B_9 + \mu_3^2 B_{10}}{\mu_3^3 B_6 + \mu_3^2 B_7 + \mu_3 B_3 + B_1}\right] S_y^2 \tag{42}$$

where $B_6 = B_2 - B_1\rho_{xz}^4$, $B_7 = B_1\rho_{xz}^4 + 2B_1\rho_{xz}^2 + B_5$, $B_8 = B_1^2$, $B_9 = B_1B_5$, $B_{10} = B_1B_2$ and μ_3 is the fraction of fresh sample for the estimator T_3.

9.3. Optimum replacement policy

To determine the optimum value of μ_3 (fraction of a sample to be drawn afresh on the current occasion) so that population mean $\bar{Y}$ may be estimated with the maximum precision, we minimize the $V(T_3)_{opt}$ given in equation (42) with respect to μ_3, which result in fourth degree equation in μ_3 and respective solutions of μ_3 is discussed below:

$$P_1\mu_3^4 + P_2\mu_3^3 + P_3\mu_3^2 + P_4\mu_3 + P_5 = 0 \tag{43}$$

where $P_1 = -B_6B_{10}$, $P_2 = -2B_6B_9$,

$P_3 = B_3B_{10} - B_7B_9 - 3B_6B_8$, $P_4 = 2(B_1B_{10} - B_7B_8)$,

$P_5 = B_1B_9 - B_3B_8$

From equations (43) it is obvious that the four real values of μ_3 are possible. Hence, while choosing the values of μ_3, it should be remembered that $0 \le \mu_3 \le 1$. All the other values of μ_3 are inadmissible. If more than one admissible values are obtained, the lowest admissible value is the best choice as it reduces the cost of the survey. From equation (43), substituting the admissible value of μ_3 say $\hat{\mu}_3$ into equation (42), we have the optimum value of $V(T_3)_{opt}$ as

$$V\left(T_3^0\right)_{opt} = \frac{1}{n}\left[\frac{B_1\left(B_1 - \hat{\mu}_3 B_4 + \hat{\mu}_3^2 B_5\right)}{\hat{\mu}_3^3 B_6 + \hat{\mu}_3^2 B_7 + \hat{\mu}_3 B_8 + B_1}\right] S_y^2 \tag{44}$$

9.4. Efficiency comparison

To study the performance of the estimator T_3, the percent relative efficiencies of the estimator T_3 with respect to (i) $\bar{y}_n$, when there is no matching, and (ii) the estimator T_2 , when no auxiliary information is used at any occasion, have been obtained for different choices of correlations. For different choices of ρ_{yx}, ρ_{xz} and ρ_{yz}, Table 2 shows the optimum values of μ_3 and percent relative efficiencies E_3 and E_4 of the estimator T_3 with respect to the estimators $\bar{y}_n$ and T_2 respectively, where

$$E_3 = \frac{V\left(\bar{y}_n\right)}{V\left(T_3^0\right)_{opt}} \times 100 \quad \text{and} \quad E_4 = \frac{V\left(T_2^0\right)_{opt}}{V\left(T_3^0\right)_{opt}} \times 100.$$

9.5. Analysis of results for estimator T_3

The following conclusions can be read out from Table 2:

(a) For fixed values of ρ_{xz} and $\rho_{yz,}$ the values of μ_3 and E_3 are increasing with the increasing values of $\rho_{yx.}$ Efficiencies E_4 are decreasing for the increasing values of ρ_{yx}.

(b) For fixed values of ρ_{xz} and $\rho_{yx,}$ the values of μ_3 increase for the lower values of ρ_{yz} and decrease for the higher values of ρ_{yz}. Efficiencies E_3 and E_4 are increasing with the increasing values of ρ_{yz} .

(c) For fixed values of ρ_{yz} and ρ_{yx}, the values of μ_3 are increasing with the increasing values of ρ_{xz}. Efficiencies E_3 and E_4 increase for the lower values of ρ_{xz} while decreasing patteren may also be seen for the higher values of ρ_{xz} .

(d) Minimum value of μ_3 is 0.5365, which indicates that only 54 percent of the total sample size is to be replaced at the current occasion for the corresponding choices of the correlations.

Table 1. Optimum values of μ_1 and percent relative efficiencies of T_1 with respect to $\bar{y}_n$ and T_2

$\rho_{xz}\downarrow$	$\rho_{yz}\downarrow$	$\rho_{yx}\rightarrow$	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	0.5	$\hat{\mu}_1$	0.5006	0.5051	0.5147	0.5307	0.5556	0.5953	0.6672
		E_1	133.48	134.69	137.25	141.51	148.14	158.74	177.91
		E_2	130.41	129.07	128.06	127.35	126.97	126.99	127.72
	0.7	$\hat{\mu}_1$	0.4956	0.4965	0.5039	0.5190	0.5450	0.5899	0.6818
		E_1	194.36	194.70	197.60	203.53	213.74	231.35	267.36
		E_2	189.88	186.58	184.36	183.17	183.18	185.08	191.95
	0.9	$\hat{\mu}_1$	0.4352	0.4329	0.4404	0.4597	0.4982	0.5823	*
		E_1	458.11	455.69	463.55	483.92	524.45	612.93	-
		E_2	447.55	436.67	432.49	435.53	449.49	490.35	-
0.7	0.5	$\hat{\mu}_1$	0.4987	0.4996	0.5070	0.5221	0.5481	0.5929	0.6844
		E_1	132.97	133.22	135.19	139.22	146.16	158.11	182.51
		E_2	129.91	127.65	126.13	125.30	125.27	126.49	131.03
	0.7	$\hat{\mu}_1$	0.5187	0.5040	0.5000	0.5060	0.5232	0.5574	0.6271
		E_1	203.39	197.62	196.09	198.41	205.17	218.58	245.91
		E_2	198.71	189.37	182.96	178.57	175.85	174.86	176.55
	0.9	$\hat{\mu}_1$	0.6616	0.5334	0.4915	0.4796	0.4897	0.5286	0.6436
		E_1	696.46	561.47	517.36	504.86	515.52	556.42	677.46
		E_2	680.42	538.03	482.70	454.37	441.84	445.13	486.37
0.9	0.5	$\hat{\mu}_1$	0.4820	0.4883	0.5003	0.5201	0.5515	0.6050	0.7230
		E_1	128.54	130.20	133.43	138.68	147.07	161.33	192.78
		E_2	125.58	124.77	124.48	124.81	126.05	129.06	138.41
	0.7	$\hat{\mu}_1$	0.6548	0.5435	0.5049	0.4943	0.5051	0.5440	0.6564
		E_1	256.79	213.15	198.01	193.85	198.07	213.33	257.41
		E_2	250.87	204.25	184.74	174.47	169.76	170.66	184.81
	0.9	$\hat{\mu}_1$	*	*	*	*	0.5509	0.5003	0.5317
		E_1	-	-	-	-	579.85	526.68	559.70
		E_2	-	-	-	-	496.96	421.35	401.83

Note: "*" indicates $\hat{\mu}_1$ do not exist.

Table 2. Optimum values of μ_3 and percent relative efficiencies of T_3 with respect to $\overline{y}_n$ and T_2

$\rho_{xz}\downarrow$	$\rho_{yz}\downarrow$	$\rho_{yx}\rightarrow$	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	0.5	$\hat{\mu}_3$	0.5365	0.5410	0.5505	0.5663	0.5907	0.6294	0.6983
		E_3	133.46	134.50	136.70	140.32	145.89	154.64	169.95
		E_4	130.38	128.88	127.54	126.28	125.04	123.71	122.02
	0.7	$\hat{\mu}_3$	0.5367	0.5367	0.5434	0.5580	0.5834	0.6273	0.7163
		E_3	196.35	196.35	198.63	203.56	212.11	226.71	255.75
		E_4	191.83	188.15	185.32	183.20	181.79	181.37	183.61
	0.9	$\hat{\mu}_3$	0.5572	0.5381	0.5381	0.5572	0.6065	0.7550	*
		E_3	545.67	528.32	528.32	545.66	589.99	719.76	-
		E_4	533.09	506.26	492.93	491.09	505.66	575.81	-
0.7	0.5	$\hat{\mu}_3$	0.5842	0.5842	0.5907	0.6049	0.6294	0.6712	0.7538
		E_3	133.48	133.48	134.72	137.40	141.96	149.55	163.89
		E_4	130.41	127.91	125.70	123.66	121.67	119.64	117.66
	0.7	$\hat{\mu}_3$	0.6014	0.5872	0.5834	0.5892	0.6058	0.6381	0.7019
		E_3	201.09	197.15	196.09	197.69	202.30	211.13	227.92
		E_4	196.46	188.92	182.95	177.92	173.38	168.89	163.63
	0.9	$\hat{\mu}_3$	*	0.6751	0.6065	0.5845	0.5897	0.6257	0.7378
		E_3	-	593.11	543.53	527.17	531.03	557.65	636.23
		E_4	-	568.34	507.12	474.45	455.13	446.12	456.78
0.9	0.5	$\hat{\mu}_3$	0.7143	0.6983	0.6983	0.7143	0.7538	0.8596	*
		E_3	135.37	133.55	133.55	135.37	139.64	149.42	-
		E_4	132.25	127.97	124.60	121.83	119.68	119.54	-
	0.7	$\hat{\mu}_3$	*	0.7730	0.7163	0.6974	0.7019	0.7325	0.8217
		E_3	-	208.23	199.41	196.24	197.001	202.04	215.00
		E_4	-	199.53	186.05	176.62	68.84	161.63	154.36
	0.9	$\hat{\mu}_3$	*	*	*	*	0.7378	0.6967	0.7226
		E_3	-	-	-	-	544.564	526.45	538.03
		E_4	-	-	-	-	66.72	421.16	386.27

Note: "*" indicates $\hat{\mu}_3$ do not exist.

10. General conclusions

The estimators T_1 and T_3 proposed in this work are proved to be the best linear unbiased estimators of population mean $\bar{Y}$ with their respective minimum variance. These estimators may be seen as new innovative ideas in survey literature as they nicely utilized the information on an auxiliary variable in order to improve the precision of the estimates. From the analysis of the results shown in Tables 1-2, the propositions of the estimators T_1 and T_3 are vindicated because it enhances the precision of estimates as well as reduces the cost of the survey. Therefore, the proposed estimators may be recommended to survey practitioners for use in real life problems.

Acknowledgements

Authors are thankful to the referee for his valuable suggestions. Authors are also thankful to the UGC, New Delhi and Indian School of Mines, Dhanbad for providing financial assistance and necessary infrastructures to carry out the present research work.

REFERENCES

BIRADAR, R. S. and SINGH, H. P., (2001). Successive sampling using auxiliary information on both occasions. *Cal. Statist. Assoc. Bull*. 51, 243-251.

CHATURVEDI, D. K. and TRIPATHI, T. P., (1983). Estimation of population ratio on two occasions using multivariate auxiliary information. *Jour. Ind. Statist. Assoc*., 21, 113-120.

DAS, A. K., (1982). Estimation of population ratio on two occasions, *Jour Ind. Soc. Agr. Statist.* 34, 1-9.

FENG, S. and ZOU, G., (1997). Sample rotation method with auxiliary variable. *Communications in Statistics-Theory and Methods*, 26, 6, 1497-1509.

GUPTA, P. C., (1979). Sampling on two successive occasions. *Jour. Statist. Res*. 13, 7-16.

JESSEN, R. J., (1942). Statistical Investigation of a Sample Survey for obtaining farm facts, *Iowa Agricultural Experiment Station Research Bulletin No. 304, Ames, Iowa, U. S. A.*, 1-104.

PATTERSON, H. D., (1950). Sampling on successive occasions with partial replacement of units, *Journal of the Royal Statistical Society*, 12, 241-255.

RAO, J. N. K. and Graham, J. E., (1964). Rotation design for sampling on repeated occasions. *Jour. Amer. Statist. Assoc.* 59, 492-509.

SEN, A. R., (1971). Successive sampling with two auxiliary variables, *Sankhya*, 33, Series B, 371-378.

SEN, A. R., (1972). Successive sampling with p $(p \geq 1)$ auxiliary variables, *Ann. Math. Statist.*, 43, 2031-2034.

SEN, A. R., (1973). Theory and application of sampling on repeated occasions with several auxiliary variables, *Biometrics* 29, 381-385.

SINGH, V. K., SINGH, G. N. and SHUKLA, D., (1991). An efficient family of ratio-cum-difference type estimators in successive sampling over two occasions, *Jour. Sci. Res*. 41 C, 149-159.

SINGH, G. N., (2003). Estimation of population mean using auxiliary information on recent occasion in h-occasion successive sampling, *Statistics in Transition*, 6, 523-532.

SINGH, G. N., (2005). On the use of chain-type ratio estimator in successive sampling, *Statistics in Transition*, 7, 21-26.

SINGH, G. N. and SINGH, V. K., (2001). On the use of auxiliary information in successive sampling, *J. Indian Soc. Agric. Statist*., 54 (1), 1-12.

SINGH, G. N. and PRIYANKA, K., (2006). On the use of chain-type ratio to difference estimator in successive sampling, *IJAMAS,* 5 (S06), 41-49.

SINGH, G. N. and PRIYANKA, K.**,** (2007). On the use of auxiliary information in search of good rotation patterns on successive occasions, *Bulletin of Statistics and Economics,* 1 (A07), 42-60.

SINGH, G. N. and PRIYANKA, K., (2008). Search of good rotation patterns to improve the precision of estimates at current occasion, *Communications in Statistics- Theory and Methods,* 37(3), 337-348.

SINGH, G. N. and KARNA, J. P., (2009, a). Estimation of population mean on current occasion in two-occasion successive sampling, *METRON*, 67(1), 69-85.

SINGH, G. N. and KARNA, J. P., (2009, b). Search of effective rotation patterns in presence of auxiliary information in successive sample over two-occasions, *Statistics in Transition, new series* 10(1), 59-73.

SINGH, H. P. and VISHWAKARMA, G. K., (2009). A general procedure for estimating population mean in successive sampling, *Communications in Statistics - Theory and Methods*, 38(2), 293-308.