İSTATİSTİK: JOURNAL OF THE TURKISH STATISTICAL ASSOCIATION
Vol. 13, No. 2, July 2021, pp. 39–51
ISSN 1300-4077 | 21 | 2 | 39 | 51

İSTATİSTİK

IMPROVED ESTIMATORS FOR ESTIMATING THE POPULATION MEAN IN TWO OCCASION SUCCESSIVE SAMPLING

Vishwantra Sharma *
Department of Statistics,
Jammu University,
J&K, India

Sunil Kumar
Department of Statistics,
Jammu University,
J&K, India

Abstract: This paper addresses the problem of estimating the population mean of the study variable in two occasions successive sampling. Based on the available information from the first and second occasions, class of estimators produced under two situations, i) when the information on a positively correlated auxiliary variable with the study variable is available on both the occasions and ii) when the information on the auxiliary variable which is negatively correlated with the study variable is available on both the occasions. Properties of the suggested class of estimators have been studied and compared with the sample mean estimator with no matching from the previous occasion and traditional successive sampling linear estimator. The study is supported by an optimal replacement policy. Empirical study also has been illustrated to show the performance of the recommended estimators theoretically.

Key words: Study variable, Auxiliary variable, Bias, Mean squared error, Successive sampling.

1. Introduction

In most surveys, the interest is on the current average despite looking at it from one occasion to the next occasion and all occasions. In successive (rotation) sampling, it is common to use the entire information gathered on the previous occasions to improve the precision of the estimator on the current occasion. The main objective of the sampling on two successive occasions is to estimate the population parameters viz. population total, mean, ratio, product, etc. for the most recent occasion as well as changes in the parameters from one occasion to the next occasion, see Okafor and Arnab [6]. Jessen [4] was the first who pioneered the procedure of utilizing the information obtained on the first occasion in improving the estimates of the current occasion. Patterson [7] extended the work of Jessen from two occasions to more. Further, Eckler [2], Rao and Graham [8], Singh et al. [10], Feng and Zou [3], Biradar and Singh [1], Singh and Vishwakarma [12], Singh and Vishwakarma [13], Singh and Pal [15] among others have suggested several estimators by using the auxiliary information for estimating the population mean on the current occasion successive (rotation) sampling.

In this paper, we extend a procedure of utilizing the information of the auxiliary variable readily available on both the equations under two different situations, by suggesting the estimator of the population mean $\bar{Y}$ of the study variable y:

Situation I: When the auxiliary variable z_1 is positively correlated with the study variable y.

Situation II: Readily available auxiliary variable z_2 is negatively correlated with the study variable y.

* Corresponding author. E-mail address:vishwantrasharma07@gmail.com

Keeping in view the situation I and II, we have suggested two estimators and studied the properties of the suggested estimators. The behaviour of the suggested estimator is explained through empirical study. We found that the proposed study is more efficient than the other considered estimators when there is close association between auxiliary and study variables.

2. Notations used and the proposed estimator

Consider a finite population $U=(U_1,U_2,...,U_N)$ of N distinct identifiable units. Let the variable under study on the first (second) occasion be denoted by $x(y)$ respectively. It is assumed that the information on the auxiliary variable z_1 and z_2 are known and have positive and negative correlation with x and y respectively readily available on both the occasions. A simple random sample (without replacement) of n units is taken on the first occasion from population U. A random sub-sample of $m(=n\lambda)$ units is retained (matched) for use on the second occasion. Now, at the current occasion, we again withdraw a simple random sample (without replacement) of size $u=n-m=n\mu$ units from the remaining $(N-n)$ units of the population so that the sample size on second occasion is also n. λ and μ are the fractions of matched and fresh samples respectively at the second (current) occasion such that $(\lambda+\mu=1)$. We shall use the following notations:

- $\bar{X},\bar{Y},\bar{Z}_1,\bar{Z}_2$: The population means of variables x,y,z_1 and z_2 respectively.
- $S_{xy},S_{yz_1},S_{yz_2},S_{xz_1},S_{xz_2}$: The population covariance between variables in suffixes.
- $\rho_{xy},\rho_{yz_1},\rho_{yz_2},\rho_{xz_1},\rho_{xz_2}$: The population correlation coefficients between variables in suffixes.
- $S_x^2,S_y^2,S_{z_1}^2,S_{z_2}^2$: The population variances of x,y, z_1 and z_2 respectively.

For obtaining the expression of bias and mean squared error of the proposed estimator, we assume that

$$y_u=\bar{Y}(1+e_{0u}),y_m=\bar{Y}(1+e_{0m}),x_m=\bar{X}(1+e_{1m}),x_n=\bar{X}(1+e_{1n}),z_{1u}=\bar{Z}_1(1+e_{1u}),$$

$$z_{2u}=\bar{Z}_2(1+e_{2u}),z_{1n}=\bar{Z}_1(1+e_{2n}),z_{2n}=\bar{Z}_2(1+e_{2n'}),b_{yz_1u}=\frac{s_{yz_1(u)}}{s^2_{z_1(u)}},b_{yx(m)}=\frac{s_{yx(m)}}{s^2_{x(m)}},$$

$$s_{yz_1(u)}=S_{yz_1(u)}(1+e_{3u}),s^2_{z_1(u)}=S^2_{z_1(u)}(1+e_{4u}),s_{yx(m)}=S_{yx(m)}(1+e_{3m}),s^2_{x(m)}=S^2_{x(m)}(1+e_{4m})$$

such that

$$E(e_{0u})=E(e_{0m})=E(e_{1m})=E(e_{1n})=E(e_{1u})=E(e_{2u})=E(e_{2n})=E(e_{2n'})=E(e_{3u})$$

$$=E(e_{4u})=E(e_{3m})=E(e_{4m})=0 \text{ and}$$

$$E(e_{0u}^2)=(\tfrac{1}{u}-\tfrac{1}{N})C_y^2,E(e_{0m}^2)=(\tfrac{1}{m}-\tfrac{1}{N})C_y^2,\ E(e_{1m}^2)=(\tfrac{1}{m}-\tfrac{1}{N})C_x^2,E(e_{1n}^2)=(\tfrac{1}{n}-\tfrac{1}{N})C_x^2,$$

$$E(e_{1u}^2)=(\tfrac{1}{u}-\tfrac{1}{N})C_{z_1}^2,E(e_{2u}^2)=(\tfrac{1}{u}-\tfrac{1}{N})C_{z_2}^2,E(e_{2n}^2)=(\tfrac{1}{n}-\tfrac{1}{N})C_{z_1}^2,E(e_{2n}^{\prime 2})=(\tfrac{1}{n}-\tfrac{1}{N})C_{z_2}^2,$$

$$E(e_{0u}e_{0m})=-(\tfrac{1}{N})C_y^2,E(e_{0u}e_{1m})=-(\tfrac{1}{N})\rho_{xy}C_yC_x,\ E(e_{0u}e_{1n})=-(\tfrac{1}{N})\rho_{xy}C_yC_x,$$

$$E(e_{0u}e_{1u})=(\tfrac{1}{u}-\tfrac{1}{N})\rho_{yz_1}C_yC_{z_1},\ E(e_{0u}e_{2u})=(\tfrac{1}{u}-\tfrac{1}{N})\rho_{yz_2}C_yC_{z_2},E(e_{0u}e_{2n})=-(\tfrac{1}{N})\rho_{xz_1}C_yC_{z_1},$$

$$E(e_{0u}e_{2n'})=-(\tfrac{1}{N})\rho_{xz_1}C_yC_{z_1},E(e_{0m}e_{1m})=(\tfrac{1}{m}-\tfrac{1}{N})\rho_{yx}C_yC_x,E(e_{0m}e_{1n})=(\tfrac{1}{n}-\tfrac{1}{N})\rho_{yx}C_yC_x,$$

$$E(e_{0m}e_{1u})=-(\tfrac{1}{N})\rho_{yz_1}C_yC_{z_1},E(e_{0m}e_{2u})=-(\tfrac{1}{N})\rho_{yz_2}C_yC_{z_2},E(e_{0m}e_{2n})=(\tfrac{1}{n}-\tfrac{1}{N})\rho_{yz_1}C_yC_{z_1},$$

$$E(e_{0m}e_{2n'})=(\tfrac{1}{n}-\tfrac{1}{N})\rho_{yz_2}C_yC_{z_2},E(e_{0m}e_{2n})=(\tfrac{1}{n}-\tfrac{1}{N})C_x^2,E(e_{1m}e_{1u})=-(\tfrac{1}{N})\rho_{xz_1}C_xC_{z_1},$$

$E(e_{1m}e_{2u}) = -(\frac{1}{N})\rho_{xz_2}C_xC_{z_2}, E(e_{1n}e_{1u}) = -(\frac{1}{N})\rho_{xz_1}C_xC_{z_1}, E(e_{1n}e_{2u}) = -(\frac{1}{N})\rho_{xz_2}C_xC_{z_2},$

$E(e_{1n}e_{2n}) = (\frac{1}{n} - \frac{1}{N})\rho_{xz_1}C_xC_{z_1}, E(e_{1n}e_{2n'}) = (\frac{1}{n} - \frac{1}{N})\rho_{xz_2}C_xC_{z_2}, E(e_{1u}e_{2u}) = (\frac{1}{u} - \frac{1}{N})\rho_{z_1z_2}C_{z_1}C_{z_2},$

$E(e_{1u}e_{2n}) = -(\frac{1}{N})C_{z_1}^2, E(e_{1u}e_{2n'}) = -(\frac{1}{N})C_{z_2}^2, E(e_{2n}e_{2n'}) = (\frac{1}{n} - \frac{1}{N})\rho_{z_1z_2}C_{z_1}C_{z_2},$

$E(e_{1u}e_{3u}) = \frac{(N-u)}{u(N-2)}\frac{\mu_{012}}{\bar{Z}\mu_{011}}, E(e_{1u}e_{4u}) = \frac{(N-u)}{u(N-2)}\frac{\mu_{003}}{\bar{Z}\mu_{002}}, E(e_{2u}e_{3u}) = \frac{(N-u)}{u(N-2)}\frac{\mu_{012}}{\bar{Z}\mu_{011}},$

$E(e_{2u}e_{4u}) = \frac{(N-u)}{u(N-2)}\frac{\mu_{003}}{\bar{Z}\mu_{002}}, E(e_{1m}e_{3m}) = \frac{(N-m)}{m(N-2)}\frac{\mu_{210}}{\bar{X}\mu_{110}}, E(e_{1n}e_{3m}) = \frac{(N-n)}{n(N-2)}\frac{\mu_{210}}{\bar{X}\mu_{110}},$

$E(e_{1m}e_{4m}) = \frac{(N-m)}{m(N-2)}\frac{\mu_{300}}{\bar{X}\mu_{200}}, E(e_{1n}e_{4m}) = \frac{(N-n)}{n(N-2)}\frac{\mu_{300}}{\bar{X}\mu_{200}},$

$\mu_{pqr} = E((x_i - \bar{X})^p(y_i - \bar{Y})^q(z_i - \bar{Z})^r), (p,q,r) = 0,1,2., C_x^2 = S_x^2/\bar{X}^2, C_y^2 = S_y^2/\bar{Y}^2,$

$C_{z_1}^2 = S_{z_1}^2/\bar{Z}_1^2, C_{z_2}^2 = S_{z_2}^2/\bar{Z}_2^2.$

We assume that information on the auxiliary character is readily available on both occasions under two different situations, one can define the estimator when

Situation I: Estimation of the population mean $\bar{Y}$ of the study variable y when the auxiliary variable 'z_1' is positively correlated with the study variable.
In a situation, when the regression of Y on X is a straight line that does not pass through the origin then regression estimators are used. Replacing regression estimator in place of a sample mean and using in exponential-type estimators of Singh and Pal [15]. We have suggested two independent estimators for estimating the population mean $\bar{Y}$ of the study variable y on the second occasion. One is based on the sample of size $u(= n\mu)$ drawn afresh on the second occasion defined by

$$T_u = t_{lregu} exp\left(\frac{\bar{Z}_1 - \bar{z}_{1u}}{\bar{Z}_1 + \bar{z}_{1u}}\right) \tag{2.1}$$

where $t_{lregu} = [\bar{y}_u + b_{yz_1(u)}(\bar{Z}_1 - \bar{z}_{1u})]$ and $b_{yz_1(u)}$ is the sample regression coefficient of y and z_1 based on the sample size u.
Second estimator is based on the sample of size $m(= n\lambda)$ common for both the occasions is defined by

$$T_m = t_{lregm} exp\left(\frac{\bar{x}_n - \bar{x}_m}{\bar{x}_n + \bar{x}_m}\right) exp\left(\frac{\bar{Z}_1 - \bar{z}_{1n}}{\bar{Z}_1 + \bar{z}_{1n}}\right) \tag{2.2}$$

where $t_{lregm} = [\bar{y}_m + b_{yx(m)}(\bar{x}_n - \bar{x}_m)]$ and $b_{yx(m)}$ is the sample regression coefficient of y and x based on the matched sample of size m.

The estimator T_u may be used to estimate the population mean on each occasion, while the estimator T_m is suitable to estimate the change over occasions. To device suitable estimation procedures for both the problems simultaneously, a convex linear combination of T_u and T_m is considered as a final estimator of the population mean $\bar{Y}$ and is given by

$$T = \phi T_u + (1 - \phi)T_m \tag{2.3}$$

where $\phi(0 \le \phi \le 1)$ is an unknown scalar to be defined such that the mean squared error (MSE) of T is minimum.

Remark 1 : For estimating the mean on each occasion the estimator T_u is suitable, which implies that for ϕ close to 1 while for estimating the change from one occasion to next occasion, the estimator T_m is more suitable so that the value of ϕ might be close to 0. For asserting both the problems simultaneously, the optimum (minimized) choice of ϕ is required.

3. Bias and Mean Square Error of T

Since T_u and T_m are biased estimators of $\bar{Y}$, therefore the estimator T is also a biased estimator of $\bar{Y}$. For bias, express the estimator T_u and T_m in terms of $\epsilon^{'}$s, we have

$$T_u = \{\bar{Y}(1+e_{0u}) - b_{yz_1}(u)\bar{Z}_1 e_{1u}\} exp\left[\frac{-e_{1u}}{2}\left(1+\frac{e_{1u}}{2}\right)^{-1}\right] \tag{3.1}$$

$$\begin{aligned} T_m = \bar{Y}\{(1+e_{0m}) + k_{yx}(e_{1n} - e_{1m} + e_{1n}e_{3m} - e_{1n}e_{4m} - e_{1m}e_{3m} + e_{1m}e_{4m})\} \\ exp\left[\frac{e_{1n}-e_{1m}}{2}\left(1+\frac{e_{1n}+e_{1m}}{2}\right)^{-1}\right] exp\left[\frac{-e_{2n}}{2}\left(1+\frac{e_{2n}}{2}\right)^{-1}\right] \end{aligned} \tag{3.2}$$

Expanding the right-hand side of equation (4) and (5) in terms of e's and neglecting the terms having power greater than two, we get

$$T_u \cong \bar{Y}\left[1 + e_{0u} - \left(\frac{1}{2}\right)e_{1u} - \left(\frac{1}{2}\right)e_{0u}e_{1u} + \left(\frac{3}{8}\right)e_{1u}^2 - k_{yz_1}(e_{1u} - \left(\frac{1}{2}\right)e_{1u}^2 - e_{1u}e_{4u} + e_{1u}e_{3u}\right]$$

$$(T_u - \bar{Y}) \cong \bar{Y}\left[e_{0u} - \left(\frac{1}{2}\right)e_{1u} - \left(\frac{1}{2}\right)e_{0u}e_{1u} + \left(\frac{3}{8}\right)e_{1u}^2 - k_{yz_1}(e_{1u} - \left(\frac{1}{2}\right)e_{1u}^2 - e_{1u}e_{4u} + e_{1u}e_{3u}\right] \tag{3.3}$$

$$\begin{aligned} T_m \cong \bar{Y}\Bigg[1 + e_{0m} - \left(\frac{1}{2}\right)(e_{1n} - e_{1m} - e_{2n}) + \left(\frac{3}{8}\right)e_{2n}^2 + \left(\frac{1}{2}\right)(e_{0m}e_{1n} - e_{0m}e_{1m} - \\ e_{0m}e_{2n}) - \left(\frac{1}{4}\right)(e_{1n}^2 - e_{1m}^2 + e_{1n}e_{2n} - e_{1m}e_{2n}) + \left(\frac{1}{8}\right)(e_{1n}^2 + e_{1m}^2 - 2e_{1n}e_{1m}) \\ + k_{yx}\{e_{1n} - e_{1m} + \left(\frac{1}{2}\right)(e_{1n}^2 + e_{1m}^2) - \left(\frac{1}{2}\right)(e_{1n}e_{2n} + e_{1n}e_{1m} - e_{1m}e_{2n} + \\ e_{1n}e_{1m}) + e_{1n}e_{3m} - e_{1n}e_{4m} - e_{1m}e_{3m} + e_{1m}e_{4m}\}\Bigg] \end{aligned}$$

$$\begin{aligned} (T_m - \bar{Y}) \cong \bar{Y}\Bigg[e_{0m} - \left(\frac{1}{2}\right)(e_{1n} - e_{1m} - e_{2n}) + \left(\frac{3}{8}\right)e_{2n}^2 + \left(\frac{1}{2}\right)(e_{0m}e_{1n} - e_{0m}e_{1m} - e_{0m}e_{2n}) - \\ \left(\frac{1}{4}\right)(e_{1n}^2 - e_{1m}^2 + e_{1n}e_{2n} - e_{1m}e_{2n}) + \left(\frac{1}{8}\right)(e_{1n}^2 + e_{1m}^2 - \\ 2e_{1n}e_{1m}) + k_{yx}(e_{1n} - e_{1m} + \left(\frac{1}{2}\right)(e_{1n}^2 + e_{1m}^2) - \left(\frac{1}{2}\right)(e_{1n}e_{2n} + e_{1n}e_{1m} \\ - e_{1m}e_{2n} + e_{1n}e_{1m}) + e_{1n}e_{3m} - e_{1n}e_{4m} - e_{1m}e_{3m} + e_{1m}e_{4m})\Bigg] \end{aligned} \tag{3.4}$$

where, $k_{yz_1} = \rho_{yz_1}\frac{C_y}{C_{z_1}}$ and $k_{yx} = \rho_{yx}\frac{C_y}{C_x}$.
Taking expectation on both sides of equation (6) and (7), one can obtain the bias of T_u and T_m to

the first degree of approximation as

$$B(T_u) = \bar{Y}\left[\left(\frac{1}{u}-\frac{1}{N}\right)\left[\left(\frac{3}{8}\right)+k_{yz_1}\right]C_{Z_1}^2-\left(\frac{1}{2}\right)\left(\frac{1}{u}-\frac{1}{N}\right)\rho_{yz_1}C_yC_{z_1}-k_{yz_1}\left(\frac{N-u}{u(N-2)\bar{Z}}\right)\left[\frac{\mu_{012}}{\mu_{011}}-\frac{\mu_{003}}{\mu_{002}}\right]\right] \tag{3.5}$$

$$B(T_m) = \bar{Y}\left[\left(\frac{3}{8}\right)\left(\frac{1}{n}-\frac{1}{N}\right)C_{z_1}^2+\left(\frac{1}{2}\right)\left[\left(\frac{1}{n}-\frac{1}{m}\right)\rho_{yx}C_yC_x-\left(\frac{1}{n}-\frac{1}{N}\right)\rho_{yz_1}C_yC_{z_1}\right]-\left(\frac{7}{8}\right)\left(\frac{1}{n}-\frac{1}{m}\right)C_x^2+\frac{(N-n)}{n(N-2)\bar{X}}\left[\left(\frac{1}{n}\right)\left(\frac{\mu_2 10}{\mu_1 10}-\frac{\mu_0 12}{\mu_0 11}\right)-\left(\frac{1}{m}\right)\left(\frac{\mu_{210}}{\mu_{110}}-\frac{\mu_{012}}{\mu_{011}}\right)\right]\right] \tag{3.6}$$

For MSE, squaring both side of equation (6) and (7), and neglecting terms of e's having power greater than two, we have

$$(T_u-\bar{Y})^2 \cong \bar{Y}^2\left[e_{0u}-\left(\frac{1}{2}\right)e_{1u}-k_{yz_1}e_{1u}\right]^2 \tag{3.7}$$

$$(T_m-\bar{Y})^2 \cong \bar{Y}^2\left[e_{0m}+\left(\frac{1}{2}\right)(e_{1n}-e_{1m}-e_{2n})-k_{yx}(e_{1n}-e_{1m})\right]^2 \tag{3.8}$$

Taking expectations to both sides of equation (10) and (11) we get the MSE of T_u and T_m respectively, as

$$MSE(T_u) = \bar{Y}^2\left(\frac{1}{u}-\frac{1}{N}\right)\left[C_{y^2}+C_{z_1}^2\left\{\left(\frac{1}{4}\right)+k_{yz_1}^2+k_{yz_1}\right\}-(1+2k_{yz_1})\rho_{yz_1}C_yC_{z_1}\right] \tag{3.9}$$

$$MSE(T_m) = \bar{Y}^2\left[\left(\frac{1}{m}-\frac{1}{N}\right)\left\{C_y^2+\left(\frac{1}{4}\right)C_x^2+k_{yx}^2C_x^2-\rho_{yx}C_yC_x+k_{yx}C_x^2-2k_{yx}\rho_{yx}C_yC_x\right\}+\left(\frac{1}{n}-\frac{1}{N}\right)\left\{\left(\frac{1}{4}\right)C_{z_1}^2-\left(\frac{1}{4}\right)C_x^2+k_{yx}^2C_x^2+\rho_{yx}C_yC_x-k_{yx}C_x^2-\rho_{yz_1}C_yC_{z_1}-2k_{yx}\rho_{yx}C_yC_x\right\}\right] \tag{3.10}$$

The covariance between the two estimators T_u and T_m to the first degree of approximation is obtained as follows:

$$\begin{aligned}Cov(T_u,T_m) &= E[(T_u-\bar{Y})(T_m-\bar{Y})]\\ &= \bar{Y}^2E\big[\big(e_{0u}-(1/2)e_{1u}-k_{yz_1}e_{1u}\big)\big(e_{0m}-(1/2)(e_{1n}-e_{1m}-e_{2n})+k_{yx}(e_{1n}-e_{1m})\big)\big]\\ &= \bar{Y}^2E\big[e_{0u}e_{0m}-(1/2)e_{0m}e_{1u}-k_{yz_1}e_{0m}e_{1u}+(1/2)\big(e_{0u}e_{1n}-(1/2)e_{1u}e_{1n}-k_{yz_1}e_{1u}e_{1n}-e_{0u}e_{1m}+(1/2)e_{1u}e_{1m}+k_{yz_1}e_{1u}e_{1m}-e_{0u}e_{2n}+(1/2)e_{1u}e_{2n}+k_{yz_1}e_{1u}e_{2n}\big)+k_{yx}\big(e_{0u}e_{1n}-(1/2)e_{1u}e_{1n}-k_{yz_1}e_{1n}e_{1u}-e_{0u}e_{1m}+(1/2)e_{1u}e_{1m}+k_{yz_1}e_{1u}e_{1m}\big)\big]\\ &= -(\bar{Y}^2/N)\big[C_y^2-\rho_{yz_1}C_yC_{z_1}+(1/4)C_{z_1}^2-k_{yz_1}\rho_{yz_1}C_yC_{z_1}+(1/2)k_{yz_1}C_{z_1}^2\big]\end{aligned} \tag{3.11}$$

Assumption 1. Considering the stability nature of the variables, the coefficient of variation of x, y, z_1, z_2 are assumed to be approximately equal ($C_y \cong C_x \cong C_{z_1} \cong C_{z_2}$), see Murthy [5], Reddy [9], Singh and Ruiz-Espejo [11]. Under Assumption 1, we state the following theorems without proof.

THEOREM 1. *The bias of the proposed estimator 'T' to the first degree of approximation is given by*

PROOF.

$$B(T) = \phi B(T_u) + (1-\phi)B(T_m) \tag{3.12}$$

where

$$B(T_u) = \bar{Y}\left[\left(\frac{1}{u}-\frac{1}{N}\right)\left(\frac{3}{8}+\frac{\rho_{yz_1}}{2}\right)C_y^2 - \rho_{yz_1}\left(\frac{N-u}{u(N-2)\bar{Z}_1}\right)\left(\frac{\mu_{012}}{\mu_{011}}-\frac{\mu_{003}}{\mu_{002}}\right)\right] \tag{3.13}$$

and

$$B(T_m) = \bar{Y}\Bigg[\left\{\left(\frac{3}{8}\right)\left(\frac{1}{n}-\frac{1}{N}\right)+\left(\frac{1}{2}\right)\left\{\left(\frac{1}{n}-\frac{1}{m}\right)\rho_{yx}-\left(\frac{1}{n}-\frac{1}{N}\right)\rho_{yz_1}\right\}C_y^2 - \left(\frac{7}{8}\right)\left(\frac{1}{n}-\frac{1}{m}\right)\right\} + \frac{N-n}{n(N-2)\bar{X}}\left\{\left(\frac{1}{n}\right)\left(\frac{\mu_{210}}{\mu_{110}}-\frac{\mu_{012}}{\mu_{011}}\right)-\left(\frac{1}{m}\right)\left(\frac{\mu_{210}}{\mu_{110}}-\frac{\mu_{300}}{\mu_{200}}\right)\right\}\Bigg] \tag{3.14}$$

THEOREM 2. *The MSE of 'T' to the first degree of approximation is obtained by*

PROOF.

$$MSE(T) = \phi^2 MSE(T_u) + (1-\phi)^2 MSE(T_m) + 2\phi(1-\phi)Cov(T_u,T_m), \tag{3.15}$$

where

$$MSE(T_u) = \left(\frac{1}{u}-\frac{1}{N}\right)\left[\frac{5}{4}-\rho_{yz_1}^2\right]S_y^2, \tag{3.16}$$

$$MSE(T_m) = \left[\frac{1}{m}\left(\frac{5}{4}+\rho_{yx}^2\right)+\frac{1}{n}\left(\rho_{yz_1}+\rho_{yx}^2\right)-\frac{1}{N}\left(\frac{5}{4}-\rho_{yz_1}\right)\right]S_y^2 \tag{3.17}$$

and

$$Cov(T_u,T_m) = -\frac{S_y^2}{N}\left[\frac{5}{4}-\frac{\rho_{yz_1}}{2}-\rho_{yz_1}^2\right] \tag{3.18}$$

4. Minimum mean squared error of the estimator 'T'

Since MSE(T) in equation (18) is a function of unknown constant ϕ, therefore, it can be minimized with respect to ϕ and equating it to zero, we get the optimum value of ϕ as

$$\phi_{opt} = \frac{[MSE(T_m) - Cov(T_u,T_m)]}{[MSE(T_u) + MSE(T_m) - 2Cov(T_u,T_m)]} \tag{4.1}$$

By substituting the value of optimum 'ϕ' from equation (22) in equation (18) we will have the minimum MSE of 'T' as

$$min.MSE(T) = \frac{[MSE(T_u)MSE(T_m) - (Cov(T_u,T_m))^2]}{[MSE(T_u) + MSE(T_m) - 2Cov(T_u,T_m)]} \tag{4.2}$$

Substituting the values of $MSE(T_u)$, $MSE(T_m)$ and $Cov(T_u,T_m)$ in equations (22) and (23), we will have the value of ϕ_{opt} and min.MSE(T), respectively.

For simplification, further we use the following notations,

$\delta_1 = N\alpha_2 - n\alpha_5,\ \delta_2 = N\alpha_1 - N\alpha_2 - n\alpha_5 - N\alpha_4,\ \delta_3 = n^2\alpha_8^2 - nN\alpha_2\alpha_4 - n^2\alpha_4\alpha_7,$

$\delta_4 = N^2\alpha_2\alpha_4 + n^2(\alpha_4\alpha_7 - \alpha_82), \delta_5 = N^2\alpha_1\alpha_4 - N^2\alpha_2\alpha_4 - nN\alpha_4\alpha_7,$

$\alpha_1 = (5/4) - \rho_{yx}^2, \alpha_2 = \rho_{yz_1} - \rho_{yx}^2, \alpha_3 = \rho_{yz_1}^2 - (\rho_{yz_1}/2),\ \alpha_4 = (5/4) - \rho_{yz_1}^2,$

$\alpha_5 = \rho_{yz_1}^2,\ \alpha_7 = (5/4) - \rho_{yz_1}, \alpha_8 = (5/4) - (\rho_{yz_1}/2) - \rho_{yz_1}^2.$

Now, we have the reduced form of ϕ_{opt} and min.MSE(T) from equation (22) and (23) as

$$\phi_{opt} = \frac{[\mu N\alpha_1 - \mu(1-\mu)(N\alpha_2 + n\alpha_3)]}{[\mu N\alpha_1 - (1-\mu)(\mu N\alpha_2 + n\mu\alpha_5 - N\alpha_4)]} \tag{4.3}$$

and

$$min.MSE(T) = \left(\frac{S_y^2}{nN}\right)\frac{\mu^2\delta_3 + \mu\delta_4 + \delta_5}{\mu^2\delta_1 + \mu\delta_2 + N\alpha_4} \tag{4.4}$$

5. Optimum replacement policy

For obtaining the optimum value of μ (fraction of a sample to be taken afresh at the second occasion) so that the population mean $\bar{Y}$ may be estimated with maximum precision, we minimize MSE of T in equation (25) by differentiating it with respect to 'μ' and hence we get the optimum value of 'μ' as

$$\mu^2\lambda_1 + \mu\lambda_2 + \lambda_3 = 0 \tag{5.1}$$

where $\lambda_1 = (\delta_2\delta_3 - \delta_1\delta_4);\ \lambda_2 = (2N\alpha_4\delta_3 - 2\delta_1\delta_5);\ \lambda_3 = (N\alpha_4\delta_4 - \delta_2\delta_5).$
Solving equation (26) for 'μ', we get

$$\hat{\mu} = \frac{-\lambda_2 \pm \sqrt{\lambda_2^2 - 4\lambda_1\lambda_3}}{2\lambda_1} \tag{5.2}$$

The value of $\hat{\mu}$ exists, if $\lambda_2^2 \geq 4\lambda_1\lambda_3$. For any combinations of correlations (ρ_{yx}, ρ_{yz_1}) that satisfy the condition of solution, two values of $\hat{\mu}$ are possible. If both the two values $\hat{\mu}$ are admissible, then the lowest one is best. Substituting the admissible values of $\hat{\mu}$, say μ_0, from equation (27) into (25), we get the optimum value of the mean squared error of 'T', which is given by

$$min.MSE(T)_{opt} = \left(\frac{S_y^2}{nN}\right)\left[\frac{\mu_0^2\delta_3 + \mu_0\delta_4 + \delta_5}{\mu_0^2\delta_1 + \mu_0\delta_2 + N\alpha_4}\right] \tag{5.3}$$

6. Efficiency comparison

The percent relative efficiencies of the estimators T with respect to (i) $\bar{y}_n$, when there is no matching, (ii) usual successive sampling estimator, $\hat{\hat{Y}} = \psi\bar{y}_u + (1-\psi)\bar{y}_{d'}$, when no auxiliary information is used at any occasion, where $[\bar{y}_{d'} = \bar{y}_m + b_{yx}^m(\bar{x}_n - \bar{x}_m)]$ have been obtained for different choices of ρ_{yx}, ρ_{yz_1} and ρ_{yz_2}. Since $\bar{y}_n$ and $\hat{\hat{Y}}$ are unbiased estimators of the population mean $\bar{Y}$, the variance of $\bar{y}_n$ and the minimum variance of $\hat{\hat{Y}}$ [as given in Sukhatme et al.[13]] are given by

$$V(\bar{y}_n) = \frac{1-f}{n} S_y^2 \tag{6.1}$$

$$V(\hat{\hat{Y}}) = \left[\left(\frac{1}{2}\right)\left\{1 + \sqrt{(1-\rho_{yx}^2)}\right\} - f\right]\frac{S_y^2}{n} \tag{6.2}$$

From (27), (28), (29), and (30) the percent relative efficiencies of the estimators 'T' with respect to $\bar{Y}_n$ are given by

$$E_1 = PRE(T, \bar{y}_n) = \frac{V(\bar{y}_n)}{min.MSE(T)_{opt}} \times 100$$

$$= \frac{N(1-f)(\mu_0^2\delta_1 + \mu_0\delta_2 + N\alpha_4)}{\mu_0^2\delta_3 + \mu_0\delta_4 + \delta_5} \times 100 \tag{6.3}$$

$$E_2 = PRE(T, \hat{\hat{Y}}) = \frac{V(\hat{\hat{Y}})}{min.MSE(T)_{opt}} \times 100$$

$$= \frac{N\left[\left\{1 + \sqrt{(1-\rho_{yx}^2)}\right\} - 2f\right](\mu_0^2\delta_1 + \mu_0\delta_2 + N\alpha_4)}{2(\mu_0^2\delta_3 + \mu_0\delta_4 + \delta_5)} \times 100 \tag{6.4}$$

For $N = 2000$, $n = 200$ and various choices of correlations (ρ_{yx}, ρ_{yz_1}) and using the formulae from equations (27), (31) and (32) we have computed the optimum values of μ_0 and percent relative efficiencies E_1 and E_2. The findings are displayed in Table 1.

TABLE 1. Optimum values μ_0 and percent relative efficiency of T with respect to $\bar{y}_n$ and $\hat{\hat{Y}}$.

ρ_{yx}	0.2			0.3			0.4			0.5		
ρ_{yz_1}	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.6	0.72	101.46	100.32	0.75	102.23	99.61	0.80	103.91	-	0.90	103.91	-
0.7	0.65	112.20	110.94	0.67	113.18	110.28	0.71	114.49	109.18	0.78	115.92	107.30
0.8	0.58	128.32	126.88	0.59	129.57	126.25	0.62	131.31	125.22	0.67	133.45	123.52
0.9	0.48	155.05	153.31	0.50	156.70	152.69	0.52	159.05	151.67	0.55	162.10	150.03
ρ_{yx}	0.6			0.7			0.8			0.9		
ρ_{yz_1}	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.6	1.13	102.77	91.35	2.16	-	-	-1.33	178.14	138.56	0.07	154.09	105.80
0.7	0.91	116.70	103.74	1.25	112.55	-	*	-	-	-0.27	184.80	126.89
0.8	0.75	135.61	120.54	0.9	135.96	114.37	1.59	119.04	-	-1.54	281.87	193.53
0.9	0.60	165.71	147.30	0.70	169.04	142.19	0.94	166.72	129.67	*	-	-

Note : * denotes μ_0 does not exist and − implies very low efficiency.

It is envisaged from Table 1 that the proposed estimator 'T' is more efficient than the estimators $\bar{y}_n$ and $\hat{\hat{Y}}$ for different levels of correlation between the variables (y and x) and (y and z_1). The following point have been noted from the Table 1 as

1. For moderate to high correlation between y and z_1, efficiency increases with respect to $\bar{y}_n$ and $\hat{\hat{Y}}$.

2. When the correlation between y and x is very high i.e, $\rho_{yx} = 0.9$ corresponding to the different levels of correlation between y and z_1 i.e, $(\rho_{yz_1} = 0.6 \text{ to } 0.9)$, the proposed estimator 'T' performs efficiently among $\bar{y}_n$ and $\hat{\hat{Y}}$ respectively.

3. With different levels of correlation between y and z_1 i.e, $(\rho_{yz_1} = 0.6 \text{ to } 0.9)$ and for different correlation between y and x i.e, $(\rho_{yx} = 0.2 \text{ to } 0.9)$, the PRE of the proposed estimator T increases except the case when $\rho_{yx} = 0.7\, and\, 0.8$ and $\rho_{yz_1} = 0.6\, to\, 0.9$ where the PRE of the proposed estimator first decreases then increases because the value of μ_0 first increases then decreases respectively.

Situation II : Estimation of the population mean $\bar{Y}$ of the study variable 'y' when the auxiliary variable z_2 is negatively correlated with the study variable 'y'.

This section deals with case II of our problem, where the correlation between study variable 'y' and the auxiliary variable z_2 is negative. In this case, for estimating the population mean $\bar{Y}$ at the current (second) occasion with negatively correlated auxiliary variable z_2 at the first (second) occasion, we suggest the following estimators as

$$T_u^* = \{\bar{y}_u + b_{yz_2(u)}(\bar{Z}_2 - \bar{z}_{2u})\} exp\left(\frac{\bar{z}_{2u} - \bar{Z}_2}{\bar{z}_{2u} + \bar{Z}_2}\right) \tag{6.5}$$

where $b_{yz_2}(u)$ is the sample regression coefficient of y and z_2 based on the sample size u.

$$T_m^* = \{\bar{y}_m + b_{yx(m)}(\bar{x}_n - \bar{x}_m)\} exp\left(\frac{\bar{x}_n - \bar{x}_m}{\bar{x}_n + \bar{x}_m}\right) exp\left(\frac{\bar{z}_{2n} - \bar{Z}_2}{\bar{z}_{2n} + \bar{Z}_2}\right) \tag{6.6}$$

where $b_{yx(m)}$ is the sample regression coefficient of y and x based on the matched sample of size m. Consider the linear combination of $T_u^*\, and\, T_m^*$, we define the following estimator as

$$T^* = \phi^* T_u^* + (1 - \phi^*) T_m^* \tag{6.7}$$

where ϕ^* is any suitably chosen scalar.
Using the result from section '2', one can obtain the bias and mean square error of $T_u^*\, and\, T_m^*$ respectively, results of which are mentioned in the form of theorems.

THEOREM 3. *The bias of the proposed estimator T^* to the first degree of approximation is*

PROOF.

$$B(T^*) = \phi^* B(T_u^*) + (1 - \phi^*) B(T_m^*) \tag{6.8}$$

where

$$B(T_u^*) = \bar{Y}\left[\left(\frac{1}{u} - \frac{1}{N}\right)\left[\frac{\rho_{yz_2} C_y C_{z_2}}{2} - \frac{C_{z_2}^2}{8}\right] - k_{yz_2}\left(\frac{1}{u} - \frac{1}{N}\right)\frac{C_{z_2}^2}{2} + k_{yz_2}\left(\frac{N-u}{u(N-2)\bar{Z}_2}\right)\left(\frac{\mu_{003}}{\mu_{002}} - \frac{\mu_{012}}{\mu_{011}}\right)\right] \tag{6.9}$$

and

$$B(T_m^*) = \bar{Y}\left[\left(\frac{1}{2}\right)\left(\frac{1}{n} - \frac{1}{N}\right)\rho_{yz_2} C_y C_{z_2} - \left(\frac{C_{z_2}^2}{8}\right)\left(\frac{1}{n} - \frac{1}{N}\right) - \left(\left(\frac{7}{8}\right)\frac{1}{n} - \frac{1}{m}\right) C_x^2 + \left(\frac{1}{2}\right)\left(\frac{1}{n} - \frac{1}{m}\right)\rho_{yx} C_y C_x + \left(\frac{1}{4}\right)\left(\frac{1}{n} - \frac{1}{N}\right)\rho_{xz_2} C_x C_{z_2} + k_{yx}\left\{\frac{1}{(N-2)\bar{X}}\left(\frac{\mu_{210}}{\mu_{110}} - \frac{\mu_{300}}{\mu_{200}}\right)\left(\frac{N-n}{n} - \frac{N-m}{m}\right)\right\}\right] \tag{6.10}$$

where $k_{yx} = \rho_{yx}\frac{C_y}{C_x}$, $k_{yz_2} = \rho_{yz_2}\frac{C_y}{C_{z_2}}$.

THEOREM 4. *To the first degree of approximation, the MSE of 'T*' is given by*

PROOF.

$$MSE(T^*) = \phi^{*2}MSE(T_u^*) + (1-\phi^*)^2 MSE(T_m^*) + 2\phi(1-\phi)Cov(T_u^*, T_m^*) \tag{6.11}$$

where,

$$MSE(T_u^*) = \bar{Y}^2\left(\frac{1}{u} - \frac{1}{N}\right)\left[C_y^2 + C_{z_2}^2((1/4) + k_{yz_2}^2 - \rho_{yz_2}) + (1-2k_{yz_2})\rho_{yz_2}C_yC_{z_2}\right] \tag{6.12}$$

$$MSE(T_m^*) = \bar{Y}^2\left[\left(\frac{1}{m} - \frac{1}{N}\right)\left\{C_y^2 + \left(\frac{1}{4}\right)C_x^2 + k_{yx}^2C_x^2 - \rho_{yx}C_yC_x + k_{yx}C_x^2 - 2k_{yx}\rho_{yx}C_yC_x\right\} + \left(\frac{1}{n} - \frac{1}{N}\right)\left\{\left(\frac{1}{4}\right)C_{z_2}^2 - \left(\frac{1}{4}\right)C_x^2 - k_{yx}^2C_x^2 + \rho_{yx}C_yC_x + \rho_{yz_2}C_yC_{z_2} - k_{yx}C_x^2 + 2k_{yx}\rho_{yx}C_yC_x\right\}\right] \tag{6.13}$$

and

$$Cov(T_u^*, T_m^*) = -(\bar{Y}^2/N)(C_y^2 + \rho_{yz_2}C_yC_{z_2} - k_{yz_2}\rho_{yz_2}C_yC_{z_2} + (1/4)C_{z_2}^2 - (1/2)\rho_{yz_2}C_{z_2}^2) \tag{6.14}$$

THEOREM 5. *Considering Assumption 1, the bias of the proposed estimator 'T*' reduces to*

PROOF.

$$B(T^*) = \phi^*B(T_u^*) + (1-\phi^*)B(T_m^*) \tag{6.15}$$

where

$$B(T_u^*) = \bar{Y}\left[\rho_{yz_2}\left\{\left(\frac{N-u}{u(N-2)\bar{Z}_2}\right)\left(\frac{\mu_{003}}{\mu_{002}} - \frac{\mu_{012}}{\mu_{011}}\right)\right\} - \left(\frac{1}{u} - \frac{1}{N}\right)\left(\frac{1}{8}\right)\right] \tag{6.16}$$

and

$$B(T_m^*) = \bar{Y}\left[\left\{\left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{\rho_{yz_2}}{2} - \frac{1}{8}\right) - \left(\frac{1}{n} - \frac{1}{m}\right)\left((\frac{7}{8}) - \frac{\rho_{yx}}{2} - \frac{\rho_{xz_2}}{4}\right)\right\}C_y^2 + \rho_{yx}\left\{\frac{1}{(N-2)\bar{X}}\left(\frac{\mu_{210}}{\mu_{110}} - \frac{\mu_{300}}{\mu_{200}}\right)\left(\frac{N-n}{n} - \frac{N-m}{m}\right)\right\}\right] \tag{6.17}$$

THEOREM 6. *Under Assumption 1, the MSE of T^* to the first degree of approximation reduces to*

PROOF.

$$MSE(T^*) = \phi^{*2}MSE(T_u^*) + (1-\phi^*)^2 MSE(T_m^*) + 2\phi(1-\phi)Cov(T_u^*, T_m^*) \tag{6.18}$$

where,

$$MSE(T_u^*) = (1/u - 1/N)[(5/4) - \rho_{yz_2}^2]S_y^2, \tag{6.19}$$

$$MSE(T_m^*) = [(1/m)(5/4 - \rho_{yx}^2) + (1/n)(\rho_{yx}^2 + \rho_{yz_2}) - (1/N)(5/4 + \rho_{yz_2})]S_y^2 \tag{6.20}$$

and

$$Cov(T_u^*, T_m^*) = -(S_y^2/N)((5/4) + (1/2)\rho_{yz_2} - \rho_{yz_2}^2) \tag{6.21}$$

7. Minimum mean squared error of the estimator T^*

For minimum MSE of T^*, we partially differentiate equation (46) with respect to the unknown constant ϕ^* and equating it to zero, we get the optimum value of ϕ^* as

$$\phi^*_{opt} = \frac{[MSE(T^*_m) - Cov(T^*_u, T^*_m)]}{[MSE(T^*_u) + MSE(T^*_m) - 2Cov(T^*_u, T^*_m)]}$$

$$= \frac{[\mu n N\alpha_1 - \mu(1-\mu)(N\alpha_2^{'} + n\alpha_3^{'})]}{[\mu N\alpha_1 - (1-\mu)(\mu N\alpha_2^{'} + n\mu\alpha_5^{'} - N\alpha_4^{'})]} \tag{7.1}$$

Putting the value of ϕ^*_{opt} from equation (50) in equation (46) we get the minimized MSE of T^* as

$$min.MSE(T^*) = \frac{[MSE(T^*_u)MSE(T^*_m) - Cov(T^*_u, T^*_m)^2]}{[MSE(T^*_u) + MSE(T^*_m) - 2Cov(T^*_u, T^*_m)]}$$

$$= \left(\frac{S_y^2}{nN}\right)\left[\frac{\mu^2\delta_3^{'} + \mu\delta_4^{'} + \delta_5^{'}}{\mu^2\delta_1^{'} + \mu\delta_2^{'} + N\alpha_4^{'}}\right] \tag{7.2}$$

where
$\delta_1^{'} = n\alpha_2^{'} - N\alpha_5^{'}$, $\delta_2^{'} = N\alpha_1 + N\alpha_2^{'} - n\alpha_5^{'} - N\alpha_4^{'}$, $\delta_3^{'} = n^2\alpha_8^{'2} - nN\alpha_2^{'}\alpha_4^{'} - n^2\alpha_4^{'}\alpha_7^{'}$,

$\delta_4^{'} = N^2\alpha_2^{'}\alpha_4^{'} + n^2(\alpha_4^{'}\alpha_7^{'} - \alpha_8^{'2}) + nN\alpha_4^{'}(\alpha_7^{'} + \alpha_2^{'} - \alpha_1)$,

$\delta_5^{'} = N^2\alpha_1\alpha_4^{'} - N^2\alpha_2^{'}\alpha_4^{'} - nN\alpha_4^{'}\alpha_7^{'}$,
$\alpha_1 = (5/4) - \rho_{yx}^2$, $\alpha_2^{'} = \rho_{yz_2} + \rho_{yx}^2, \alpha_3^{'} = \rho_{yz_2}^2 + \rho_{yz_2}/2$, $\alpha_4^{'} = (5/4) - \rho_{yz_2}^2$,

$\alpha_5^{'} = \rho_{yz_2}^2$, $\alpha_7^{'} = (5/4) + \rho_{yz_2}\alpha_8^{'} = (5/4) + (\rho_{yz_2}/2) - \rho_{yz_2}^2$

8. Optimum replacement policy in case of negative correlation between study and auxiliary variables.

In this section, we will obtain the optimum value of μ (fraction of sample to be drawn afresh at the second occasion) so that the population mean $\bar{Y}$ may be estimated with maximum precision. Differentiating the min.MSE(T^*) given by equation (52) with respect to μ and equating to zero we get

$$\mu^2(\delta_2^{'}\delta_3^{'} - \delta_1^{'}\delta_4^{'}) + \mu(2N\alpha_4^{'}\delta_3^{'} - 2\delta_1^{'}\delta_5^{'}) + (N\alpha_4^{'}\delta_4^{'} - \delta_2^{'}\delta_5^{'}) = 0$$

$$\mu^2\lambda_1^{'} + \mu\lambda_2^{'} + \lambda_3^{'} = 0 \tag{8.1}$$

where $\lambda_1^{'} = \delta_2^{'}\delta_3^{'} - \delta_1^{'}\delta_4^{'}$, $\lambda_2^{'} = 2N\alpha_4^{'}\delta_3^{'} - 2\delta_1^{'}\delta_5^{'}$, $\lambda_3^{'} = N\alpha_4^{'}\delta_4^{'} - \delta_2^{'}\delta_5^{'}$
Solving equation (52) for μ, we get

$$\hat{\mu} = \frac{-\lambda_2^{'} \pm \sqrt{(\lambda_2^{'2} - 4\lambda_1^{'}\lambda_3^{'})}}{2\lambda_1^{'}} \tag{8.2}$$

The value of $\hat{\mu}$ exists, if $(\lambda_2^{'2} - 4\lambda_1^{'}\lambda_3^{'}) \geq 0$. For any combinations of correlations (ρ_{yx}, ρ_{yz_2})) that satisfy the solution, two values of $\hat{\mu}$ are possible. Substituting the admissible values of $\hat{\mu}$, say μ_0, from equation (53) into (51), we get the optimum value of mean squared error of T^*, which is given by

$$min.MSE(T^*)_{opt} = \left(\frac{S_y^2}{nN}\right)\frac{(\mu_0^2\delta_3^{'} + \mu_0\delta_4^{'} + \delta_5^{'})}{(\mu_0^2\delta_1^{'} + \mu_0\delta_2^{'} + N\alpha_4^{'})} \tag{8.3}$$

TABLE 2. Optimum values μ_0 and percent relative efficiency of T^* with respect to $\bar{y}_n$ and $\hat{\hat{Y}}$.

ρ_{yx}	0.8			0.9		
ρ_{yz_2}	μ_0	E_1^*	E_2^*	μ_0	E_1^*	E_2^*
-0.70	0.086	131.14	102.00	*	-	-
-0.72	0.606	132.27	102.88	*	-	-
-0.74	0.431	131.20	102.04	2.161	97.40	60.77
-0.76	0.909	128.97	100.31	1.626	130.71	81.55
-0.78	0.221	126.15	98.12	3.690	68.37	46.94
-0.80	0.156	123.06	95.72	2.041	140.28	96.32
-0.82	0.109	119.89	93.25	1.281	170.62	117.15
-0.84	0.073	116.73	90.79	0.834	180.53	125.32
-0.86	0.047	113.63	88.38	0.551	184.95	126.98
-0.88	0.027	110.64	86.05	0.365	182.54	125.33
-0.90	0.012	107.77	83.82	0.241	177.82	122.09
-0.92	0.002	105.01	81.67	0.155	172.10	118.16
-0.94	-0.005	102.38	79.63	0.097	166.06	114.01
-0.96	-0.010	99.87	77.68	0.056	160.05	109.89
-0.98	-0.014	97.48	75.81	0.029	154.25	105.91

Note: $*$ denotes μ_0 does not exist and $-$ implies very low efficiency.

9. Efficiency comparison

The percent relative efficiencies of the estimators T with respect to (i) $\bar{y}_n$, when there is no matching, (ii) usual successive sampling estimator, $\hat{\hat{Y}} = \psi \bar{y}_u + (1-\psi)\bar{y}_{d'}$, when no auxiliary information is used at any occasion, where $[\bar{y}_{d'} = \bar{y}_m + b_{yx}^m(\bar{x}_n - \bar{x}_m)]$ have been obtained for different choices of ρ_{yx}, ρ_{yz_1} and ρ_{yz_2}. Since $\bar{y}_n$ and $\hat{\hat{Y}}$ are unbiased estimators of the population mean $\bar{Y}$, the variance of $\bar{y}_n$ and the minimum variance of $\hat{\hat{Y}}$ [as given in Sukhatme et al.[16]] are given by equation (29) and (30) in section 6.
From (29), (30) and (54), the percent relative efficiencies of the estimators T^* with respect to $\bar{y}_n$ and $\hat{\hat{Y}}$ are given by

$$E_1^* = PRE(T^*, \bar{y}_n) = \frac{V(\bar{y}_n)}{min.MSE(T^*)_{opt}} \times 100$$

$$= \frac{N(1-f)[\mu_0^2\delta_1' + \mu_0\delta_2' + N\alpha_4']}{\mu_0^2\delta_3' + \mu_0\delta_4' + \delta_5'} \times 100 \tag{9.1}$$

$$E_2^* = PRE(T^*, \hat{\hat{Y}}) = \frac{V(\hat{\hat{Y}})}{min.MSE(T^*)_{opt}} \times 100$$

$$= \frac{N\left[\{1+\sqrt{(1-\rho_{yx}^2)}\} - 2f\right](\mu_0^2\delta_1' + \mu_0\delta_2' + N\alpha_4')}{2(\mu_0^2\delta_3' + \mu_0\delta_4' + \delta_5')} \times 100 \tag{9.2}$$

For $N = 2000, n = 200$, and various choices of correlations(ρ_{yx}, ρ_{yz_2})and using the formulae from equations (53), (55) and (56) we have computed the optimum values of μ_0 and percent relative efficiencies E_1^* and E_2^*. The findings are displayed in Table 2.

It is noticed from Table 2 that for $\rho_{yx} = 0.8$ and $\rho_{yz_2} = -0.70\, to\, -0.94$, the performance of the proposed estimator T^* is efficient than $\bar{y}_n$ while T^* is efficient than $\hat{\hat{Y}}$ for different values of ρ_{yz_2} from $-0.70\, to\, -0.76$. For $\rho_{yx} = 0.8$ and $\rho_{yz_2} = -0.72$, the efficiency of the proposed estimator T^* over $\bar{y}_n$ is maximum, after that the efficiency decreases with increase in the value of ρ_{yz_2}.

Further, it is noticed that for $\rho_{yx} = 0.9$, the efficiency of T^* over $\bar{y}_n$ and $\hat{\hat{Y}}$ behaves in the following manner

- Efficiency increases with the increase in the value of ρ_{yz_2} i.e $\rho_{yz_2} = -0.80$ to -0.84,
- Efficiency is maximum when $\rho_{yz_2} = -0.86$, and
- Efficiency decreases with the increase in the value of ρ_{yz_2} i.e. $\rho_{yz_2} = -0.86$ to -0.98.

10. Conclusions

This article deals with the problem of estimating the population mean of the study variable on current (second) occasion in two-occasion successive sampling under two situations i) when the auxiliary variable is positively correlated with the study variable and ii) when the auxiliary variable is negatively correlated with the study variable. Properties of the suggested estimators have been discussed and the conditions where the suggested estimators are optimum are also obtained. It is found that the suggested estimator in both cases has shown efficient results when there is high correlation between study and auxiliary variables. From the empirical results, it can be concluded that the proposed estimator is more rewarding in the estimation of the population mean of the study variable at the current occasion in two occasion successive sampling. Finally, our recommendation is to use the proposed estimator by the survey practitioners in practice.

References

[1] Biradar R. S. and Singh, H. P. (2001). Successive sampling using auxiliary information on both occasions. Calcutta Statistical Association Bulletin, 51, 243-251.

[2] Eckler, A. R.(1955). Rotation sampling. Ann. Math. Stat. 26: 664-685.

[3] Feng, S, and Zou, G. (1997). Sampling rotation method with auxiliary variable. Communication in Statistics – Theory and Methods, 26(6), 1497-1509.

[4] Jessen, R. (1942). Statistical investigation of a sample survey for obtaining form facts. Iowa Agricultural Experiment Station Road Bulletin no. 304, Ames, USA.

[5] Murthy, M. (1967). Sampling Theory and Methods. Statistical Publishing Society, Kolkata, India.

[6] Okafor, F. C., Arnab, R.(1987). Some strategies of two-stage sampling for estimating population ratio over two occasions. Austrian Journal of Statistics 29(2): 128-142.

[7] Patterson, H. D. (1950). Sampling on successive occasions with partial replacement of units", Journal of the Royal Statistical Society B, 12, 241-255.

[8] Rao, J. N. K., and Graham, J.E. (1964). Rotation design for sampling on repeated occasions", Journal of the American Statistical Association, 59, 492-502.

[9] Reddy, V. (1978). A general class of estimators in successive sampling. Sankhya C, 40, pp. 29–37.

[10] Singh, Hari P., Singh, Hausila P., and Singh, V.P. (1992). A generalized efficient class of estimators of population mean in two-phase and successive sampling. International Journal of Management Systems, 8(2), 173-183.

[11] Singh, H. P., Ruiz-Espejo, M.R. (2003). On linear regression and ratio–product estimation of a finite population mean. Journal of the Royal Statistical Society: Series D (The Statistician), 52, no. 1, pp. 59–67.

[12] Singh, H. P., and Vishwakarma, G. K. (2007a). Modified exponential ratio estimators for finite population mean in double sampling. Austrian Journal of Statistics, 36(3), 217-225.

[13] Singh, H. P., and Vishwakarma, G. K. (2007b). A general class of estimators in successive sampling", Metron, 65(2), 201-227.

[14] Singh, H. P., and Vishwakarma, G. K. (2009). A general procedure for estimating the population mean in successive sampling. Communication in Statistics – Theory and Methods, 38, 293-308.

[15] Singh, H. P., and Pal, S. (2017). A Modified Procedure for Estimating the Population Mean in Two-occasion Successive Samplings. Afr. Stat. 12 (3), 1347-1365.

[16] Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S. and Ashok C.(1984). Sampling theory of University Press.