STATISTICS IN TRANSITION-new series, July 2010
Vol. 11, No. 1, pp. 105—126

ESTIMATION OF POPULATION MEAN AT CURRENT OCCASION IN PRESENCE OF SEVERAL VARYING AUXILIARY VARIATES IN TWO-OCCASION SUCCESSIVE SAMPLING

G. N. Singh, Kumari Priyanka[1]

ABSTRACT

The present work intended to emphasize the role of several varying auxiliary variates at both the occasions to improve the precision of estimates at current occasion in two-occasion successive sampling. Two different efficient estimators are proposed and their theoretical properties are examined. Relative comparison of efficiencies of the proposed estimators with the sample mean estimator when there is no matching from previous occasion, and the optimum successive sampling estimator when no auxiliary information is used have been incorporated. Empirical studies are significantly justifying the composition of proposed estimators.

Key words: Successive sampling, varying auxiliary variates, optimum replacement policy, bias, mean square error.

1. Introduction

There are many practical instances, where the survey often needs to be repeated many times. The aim of a repeated survey is to allow one or more items to be monitored over time. For survey design purposes this aim has often been simplified to two objectives: good estimates of the item for each occasion, and good estimates of change from occasion to occasion. The follow-up of these objectives is achieved by mean of sampling on successive occasions according to a specified rule, with partial replacement of units, called successive sampling. A key issue is the extent to which elements sampled at a previous occasion should be retained in the sample selected at the current occasion; this is termed as optimum replacement policy.

[1] Department of Applied Mathematics, Indian School of Mines, Dhanbad-826 004, India, E-mail: gnsingh_ism@yahoo.com.

The problem of sampling on two successive occasions with a partial replacement of units was first considered by Jessen (1942). Estimation on more than two occasions is due to Patterson (1950). Patterson's theory was examined and extended by Eckler (1955), Rao and Graham (1964), Cochran (1977), Gupta (1979), Das (1982), Chaturvedi and Tripathi (1983), among others. Sen (1971) developed estimators for the population mean on the current occasion using information on two auxiliary variates available on previous occasion. Sen (1972, 73) extended his work for more than two auxiliary variates. In addition to the information from previous occasion, Singh et al. (1991) and Singh and Singh (2001) used an additional auxiliary information on current occasion for estimating the current population mean in two-occasion successive sampling. Singh (2003) extended this methodology for h-occasions successive sampling.

In many situations, information on an auxiliary variate 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 polluting industries is known in environmental survey. Many other situations in biological (life) sciences could be explored to show the benefits of the present study. Utilizing the auxiliary information available for both the occasions Feng and Zou (1997), Biradar and Singh (2001), Singh (2005), Singh and Priyanka (2006, 2007, 2008 a, b) have proposed several chain-type ratio, difference and regression estimators for estimating the population mean at current (second) occasion in two-occasion successive sampling.

Following the works of Singh and Priyanka (2008 a, b), the objective of the present work is to propose estimators for estimating the population mean at current occasion using several varying auxiliary variates available at both the occasions. Chain-type difference and regression estimators are proposed, utilizing p-dynamic auxiliary variates. Practicability of the proposed estimator has been discussed. A relative comparison of efficiencies of the proposed estimators with the sample mean estimator when there is no matching from previous occasion, and the optimum successive sampling estimator when no auxiliary information has been used is discussed. Empirical studies show highly significant gains for the proposed estimators.

2. Formulation of Estimator

2.1. Notations

Let the character under study on the first (second) occasion be denoted by x (y). It is assumed that information on p (non negative integer constant) auxiliary variates z_{1j} (z_{2j}), j = 1, 2, - - -, p whose population means are known, are available on the first (second) occasion respectively. The auxiliary variates z_{1j} (z_{2j}), are correlated to x and y on the first and second occasions respectively. For convenience, it is assumed that the population under consideration is considerably large enough. A simple random sample (without replacement) of n units is taken on the first occasion. A random sub-sample of m = n λ units is retained (matched) for its use on the second occasion, while a fresh simple random sample (without replacement) of u = (n - m) = nμ units is drawn on the second occasion from the non-sampled units of the population so that the sample size on the second occasion is also n. λ and μ are the non-negative fractions of matched and fresh samples respectively at the current (second) occasion such that $\lambda + \mu = 1$. Hence onwards the following notations are employed:

$\overline{X}, \overline{Y}$: Population means of the study variate x and y respectively.

$\overline{Z}_{1j}, \overline{Z}_{2j}$: Population means of the jth (j = 1, 2, - - -, *p)* auxiliary variates at first and second occasion respectively.

$\overline{x}_n, \overline{x}_m, \overline{y}_u, \overline{y}_m, \overline{z}_{1nj}, \overline{z}_{1mj}, \overline{z}_{2mj}, \overline{z}_{2uj}$: Sample means of the respective variates of the sample sizes shown in suffices, where j = 1, 2, - - -, *p*.

$\rho_{yx}, \rho_{xz1j}, \rho_{yz1j}, \rho_{z1jz2j}, \rho_{xz2j}, \rho_{yz2j}$: Correlation coefficients between the variates shown in suffices, where j = 1, 2, - - -, *p*.

$S_x^2, S_y^2, S_{z1j}^2, S_{z2j}^2$: Population mean squares of x, y, z_{1j} and z_{2j} (j = 1, 2, - - -, *p*) respectively.

fpc: Finite population correction

2.2.Formulation of the Estimator T

To estimate the population mean $\overline{Y}$ on the second occasion, utilizing information on p varying auxiliary variates, two different estimators are suggested. One is a difference estimator based on sample of size u = (n μ) drawn afresh on the second occasion and given by

$$T_1 = \overline{y}_u + \sum_{j=1}^{p} \beta_{yz2j}\left(\overline{Z}_{2j} - \overline{z}_{2uj}\right) \tag{1}$$

Second estimator is a chain-type difference to difference estimator based on the sample of size m (= n λ) common with both the occasions and defined as

$$T_2 = \overline{y}_m^* + \beta_{yx}\left(\overline{x}_n^* - \overline{x}_m^*\right) \tag{2}$$

where $\overline{y}_m^* = \overline{y}_m + \sum_{j=1}^{p} \beta_{yz2j}\left(\overline{Z}_{2j} - \overline{z}_{2mj}\right)$

$$\overline{x}_n^* = \overline{x}_n + \sum_{j=1}^{p} \beta_{xz1j}\left(\overline{Z}_{1j} - \overline{z}_{1nj}\right)$$

$$\bar{x}_m^* = \bar{x}_m + \sum_{j=1}^{p} \beta_{xz1j}\left(\bar{Z}_{1j} - \bar{z}_{1mj}\right)$$

and β_{yx}, β_{xz1j} and β_{yz2j} (j = 1, 2, - - -, *p*) are known population regression coefficients between the variates shown in suffices. We assume that these population regression coefficients are known from past surveys or may be known through a pilot survey.

Combining the estimators T_1 and T_2, we have the final estimator of $\bar{Y}$ as

$$T = \varphi T_1 + \left(1 - \varphi\right) T_2 \tag{3}$$

where φ is an unknown constant to be determined so as to minimize the variance of the estimator T.

Remark 2.1: For estimating the mean on each occasion the estimator T_1 is suitable, which implies that more belief on T_1 could be shown by choosing φ as 1 (or close to 1), while for estimating the change from one occasion to the next, the estimator T_2 could be more useful so φ might be chosen as 0 (or close to 0). For asserting both the problems simultaneously, the suitable (optimum) choice of φ is required.

2.3. Estimator T in Practice

The main difficulty in using the estimator T, is the non-availability of population regression coefficients β_{yz2j}, β_{yx} and β_{xz1j}. Under such situations, the following working version of estimator T may be considered.

$$\Delta = \psi\, \Delta_1 + \left(1 - \psi\right) \Delta_2 \tag{4}$$

where $\Delta_1 = \bar{y}_u + \sum_{j=1}^{p} b_{yz2j}(u)\left(\bar{Z}_{2j} - \bar{z}_{2uj}\right)$ and $\Delta_2 = \bar{y}_m^{**} + b_{yx}\left(m\right)\left(\bar{x}_n^{**} - \bar{x}_m^{**}\right)$

where $\bar{y}_m^{**} = \bar{y}_m + \sum_{j=1}^{p} b_{yz2j}\left(m\right)\left(\bar{Z}_{2j} - \bar{z}_{2mj}\right)$

$$\bar{x}_n^{**} = \bar{x}_n + \sum_{j=1}^{p} b_{xz1j}\left(n\right)\left(\bar{Z}_{1j} - \bar{z}_{1nj}\right)$$

$$\bar{x}_m^{**} = \bar{x}_m + \sum_{j=1}^{p} b_{xz1j}\left(m\right)\left(\bar{Z}_{1j} - \bar{z}_{1mj}\right)$$

ψ is an unknown constant to be determined so that it minimizes the mean square error of the estimator Δ, $b_{yz2j}(u)$, $b_{yx}(m)$, $b_{yz2j}(m)$, $b_{xz1j}(n)$ and $b_{xz1j}(m)$

are the sample regression coefficients between the variates shown in suffices and based on the sample sizes indicated in the braces, where j = 1, 2, - - - , *p*.

2.4. Particular Cases

Case 1: If p = 1, estimators T and Δ defined in equations (3) and (4) respectively reduces to the estimators proposed by Priyanka (2008).

Case 2: If the auxiliary variates considered over the two-occasion are not changing much rapidly over time, i.e., $z_{1j} \approx z_{2j} \approx z_j$ (say), then the estimator T defined in equation (3) reduces to Singh and Priyanka (2008 b) estimator.

3. Properties of the Estimator T

Theorem 3.1: T is an unbiased estimator of $\bar{Y}$.

Proof: Since T_1 and T_2 are difference type estimators, they are unbiased for $\bar{Y}$. The final estimator T is a convex linear combination of T_1 and T_2, therefore, T is also an unbiased estimator of $\bar{Y}$.

Theorem 3.2: The variance of the estimator T (ignoring fpc) is given by

$$V(T) = \varphi^2 V(T_1) + (1-\varphi)^2 V(T_2) \tag{5}$$

$$V(T_1) = \frac{1}{u}\left(1 - \sum_{j=1}^{p}\rho_{yz2j}^2 + \sum_{j \neq k=1}^{p}\rho_{yz2j}\,\rho_{yz2k}\,\rho_{z2jz2k}\right)S_y^2 \tag{6}$$

$$\begin{aligned} V(T_2) = \Bigg[&\frac{1}{m}\left(1 - \sum_{j=1}^{p}\rho_{yz2j}^2 + \sum_{j \neq k=1}^{p}\rho_{yz2j}\,\rho_{yz2k}\,\rho_{z2jz2k}\right) \\ &+\left(\frac{1}{m}-\frac{1}{n}\right)\Bigg(-\rho_{yx}^2\left(1+\sum_{j=1}^{p}\rho_{xz1j}^2\right) + 2\,\rho_{yx}\sum_{j=1}^{p}\left\{\rho_{xz1j}\rho_{yz1j} + \rho_{yz2j}\rho_{xz2j} - \rho_{yz2j}\rho_{xz1j}\rho_{z2jz1j}\right\} \\ &+ \sum_{j \neq k=1}^{p}\left\{\rho_{yx}^2\rho_{xz1j}\,\rho_{xz1k}\,\rho_{z1jz1k} - 2\rho_{yx}\rho_{yz2j}\,\rho_{xz1k}\,\rho_{z2jz1k}\right\}\Bigg)\Bigg]S_y^2 \end{aligned} \tag{7}$$

Proof: It is clear that the variance of the estimator T is given by

$$V(T) = \varphi^2 V(T_1) + (1-\varphi)^2 V(T_2) + 2\,\varphi(1-\varphi)\,Cov(T_1, T_2) \tag{8}$$

Variance and covariance terms of equation (8) can be derived as

$$V(T_1) = E\left[T_1 - \bar{Y}\right]^2 = \left(\frac{1}{u} - \frac{1}{N}\right)\left(1 - \sum_{j=1}^{p} \rho_{yz2j}^2 + \sum_{j \neq k = 1}^{p} \rho_{yz2j}\,\rho_{yz2k}\,\rho_{z2jz2k}\right) S_y^2 \quad (9)$$

$$V(T_2) = E\left[T_2 - \bar{Y}\right]^2 = E\left[\left(\bar{y}_m^* - \bar{Y}\right) + \beta_{yx}\left\{\left(\bar{x}_n^* - \bar{X}\right) - \left(\bar{x}_m^* - \bar{X}\right)\right\}\right]^2$$

Taking expectation we get the variance of T_2 as

$$V(T_2) = \left[\left(\frac{1}{m} - \frac{1}{N}\right)\left(1 - \sum_{j=1}^{p} \rho_{yz2j}^2 + \sum_{j \neq k = 1}^{p} \rho_{yz2j}\,\rho_{yz2k}\,\rho_{z2jz2k}\right)\right.$$

$$+ \left(\frac{1}{m} - \frac{1}{n}\right)\left(-\rho_{yx}^2\left(1 + \sum_{j=1}^{p} \rho_{xz1j}^2\right) + 2\rho_{yx}\sum_{j=1}^{p}\left\{\rho_{xz1j}\rho_{yz1j} + \rho_{yz2j}\rho_{xz2j} - \rho_{yz2j}\rho_{xz1j}\rho_{z2jz1j}\right\}\right.$$

$$\left.\left. + \sum_{j \neq k = 1}^{p}\left\{\rho_{yx}^2\rho_{xz1j}\,\rho_{xz1k}\,\rho_{z1jz1k} - 2\rho_{yx}\rho_{yz2j}\,\rho_{xz1k}\,\rho_{z2jz1k}\right\}\right)\right] S_y^2 \quad (10)$$

Since the estimators T_1 and T_2 are unbiased estimators based on two independent samples of sizes u and m respectively,

$$Cov\,(T_1, T_2) = 0 \quad (11)$$

Further, we consider population size is sufficiently large, therefore, finite population corrections (fpc) are ignored.

Applying the above assumption in the equations (9) and (10), we get the expressions for the variances of the estimators T_1 and T_2 as given in equations (6) and (7) respectively.

Now, substituting the values of V (T_1), V (T_2) and Cov (T_1, T_2) in equation (8) we get the V (T) as in equation (5).

Since variance of the estimator T in equation (5) is a function of the unknown constant φ, it is minimized with respect to φ and subsequently the optimum value of φ is obtained as

$$\varphi_{opt.} = \frac{V(T_2)}{V(T_1) + V(T_2)} \quad (12)$$

Substituting this optimum value φ in equation (5) we obtain the minimum variance of T with respect to φ as

$$V(T)_{opt.} = \frac{V(T_1)\,V(T_2)}{V(T_1) + V(T_2)} \quad (13)$$

Further, substituting the values from equations (6) and (7) in equation (13), the simplified value of $V(T)_{opt.}$ is obtained as shown below in theorem 3.3.

Theorem 3.3: The $V(T)_{opt.}$ is derived as

$$V(T)_{opt.} = \frac{A[A+\mu B]}{n[A+\mu^2 B]} S_y^2 \tag{14}$$

Where $A = \left(1 - \sum_{j=1}^{p} \rho_{yz2j}^2 + \sum_{j \neq k=1}^{p} \rho_{yz2j}\, \rho_{yz2k}\, \rho_{z2jz2k}\right)$,

$$B = \left[-\rho_{yx}^2 \left(1 + \sum_{j=1}^{p} \rho_{xz1j}^2\right) + 2\,\rho_{yx} \sum_{j=1}^{p} \left\{\rho_{xz1j}\rho_{yz1j} + \rho_{yz2j}\rho_{xz2j} - \rho_{yz2j}\rho_{xz1j}\rho_{z2jz1j}\right\} \right.$$

$$\left. + \sum_{j \neq k=1}^{p} \left\{\rho_{yx}^2 \rho_{xz1j}\, \rho_{xz1k}\, \rho_{z1jz1k} - 2\rho_{yx}\rho_{yz2j}\, \rho_{xz1k}\, \rho_{z2jz1k}\right\} \right]$$, and $\mu\left(=\frac{u}{n}\right)$.

Corollary 3.1: If there is complete matching, i.e., $\mu = 0$ then

$$V(T)_{opt.} = \frac{A}{n} S_y^2 \tag{15}$$

Corollary 3.2: If there is no matching, i.e., $\mu = 1$ then

$$V(T)_{opt} = \frac{A}{n} S_y^2 \tag{16}$$

In both the cases $V(T)_{opt.}$ has the same value, this gives an implication that there must be an optimum choice of ¯ other than extreme values so that $V(T)_{opt.}$ will be smaller than the quantity given in equations (15) or (16). Thus, for making current estimate (neither the case of "complete matching" nor the case of "no matching") better, it is always preferable to replace the sample partially. For such situation a sensible strategy is to look for an optimum choice of μ, which is an important factor in reducing the cost of the survey.

Hence, $V(T)_{opt.}$ defined in equation (14) is minimized with respect to μ, resulting in a quadratic equation in μ, which shown as

$$B\,\mu^2 + 2\,A\,\mu - A = 0 \tag{17}$$

Solving equation (17) for μ, the solutions are given as

$$\hat{\mu} = \frac{-A \pm \sqrt{A(A+B)}}{B} \tag{18}$$

The real values of $\hat{\mu}$ exists if $A\,(A+B) \geq 0$. For any combination of the correlations that satisfies the above condition two real values of $\hat{\mu}$ are possible, however only the value of $\hat{\mu}$ that falls in the unit interval, $0 \leq \hat{\mu} \leq 1$ is admissible. Substituting this value of $\hat{\mu}$ (say μ_0) from equation (18) in equation (14), we have

$$V(T)_{opt.^*} = \frac{A[A+\mu_0 B]}{n[A+\mu_0^2\ B]} S_y^2 \tag{19}$$

4. Properties of the Estimator Δ

Since Δ_1 and Δ_2 are the simple linear regression and chain type regression estimators respectively, they are biased for $\bar{Y}$. Therefore, the resulting estimator Δ defined in equation (4) is also a biased estimator of $\bar{Y}$. The bias B (.) and mean square error MSE (.) up to the first order of approximations and for large population size (ignoring fpc) are derived under large sample approximations, considering the following transformations:

$$\bar{y}_u = \bar{Y}(1+e_1),\ \bar{y}_m = \bar{Y}(1+e_2),\ \bar{x}_n = \bar{X}(1+e_3),\ \bar{x}_m = \bar{X}(1+e_4),$$

$$s_{yx}(m) = S_{yx}(1+e_5),\ s_x^2(m) = S_x^2(1+e_6),\ s_{yz2j}(u) = S_{yz2j}(1+e_{7j}),$$

$$s_{yz2j}(m) = S_{yz2j}(1+e_{7j}^*),\ s_{z2j}^2(u) = S_{z2j}^2(1+e_{8j}),\ s_{z2j}^2(m) = S_{z2j}^2(1+e_{8j}^*),$$

$$s_{xz1j}(n) = S_{xz1j}(1+e_{9j}),\ s_{xz1j}(m) = S_{xz1j}(1+e_{9j}^*),\ s_{z1j}^2(m) = S_{z1j}^2(1+e_{10j}),$$

$$s_{z1j}^2(n) = S_{z1j}^2(1+e_{10j}^*),\ \bar{z}_{2uj} = \bar{Z}_{2j}(1+e_{11j}),\ \bar{z}_{2mj} = \bar{Z}_{2j}(1+e_{12j}),$$

$\bar{z}_{1mj} = \bar{Z}_{1j}(1+e_{13j})$, $\bar{z}_{1nj} = \bar{Z}_{1j}(1+e_{14j})$ such that $|e_i| < 1$, $|e_{lj}| < 1$ and $|e_{kj}^*| < 1$ $\forall$ $i = 1, 2, ---, 6$; $l = 7, 8, 9, 10, 11, 12, 13, 14$; $k = 7, 8, 9, 10$ and $j = 1, 2, ---, p$.

Under the above transformations Δ_1 and Δ_2 take the following forms:

$$\Delta_1 = \left[\bar{Y}(1+e_1) - \sum_{j=1}^{p} \beta_{yz2j}\ \bar{Z}_{2j}\ e_{11j}(1+e_{7j})(1+e_{8j})^{-1}\right] \tag{20}$$

$$\Delta_2 = \left[\bar{Y}(1+e_2) + \beta_{yx}\bar{X}(e_3 - e_4 + e_3e_5 - e_4e_5 - e_3e_6 + e_4e_6)\right.$$

$$+\sum_{j=1}^{p}\Big\{-\beta_{yz2j}\bar{Z}_{2j}\left(e_{12j}-e_{12j}e_{8j}^{*}+e_{7j}^{*}e_{12j}\right)$$

$$+\beta_{xz1j}\,\bar{Z}_{1j}\,\beta_{yx}\Big(e_{13j}-e_{10j}e_{13j}+e_{9j}e_{13j}-e_{14j}-e_{9j}^{*}e_{14j}+e_{14j}e_{10j}^{*}$$

$$+e_{5}e_{13j}-e_{5}e_{14j}-e_{6}e_{13j}+e_{6}e_{14j}\Big)\Big\}\Big] \tag{21}$$

Thus, we have the following theorems.

Theorem 4.1: $B(\Delta)=\psi\,B(\Delta_1)+(1-\psi)B(\Delta_2)$ (22)

where $B(\Delta_1)=-\left(\frac{1}{u}\right)\sum_{j=1}^{p}\beta_{yz2j}\left[\frac{C_{0102}}{S_{yz2j}}-\frac{C_{0003}}{S_{z2j}^{2}}\right]$ (23)

and $B(\Delta_2)=\frac{1}{m}\sum_{j=1}^{p}\beta_{yz2j}\left[\frac{C_{00003}}{S_{z2j}^{2}}-\frac{C_{0102}}{S_{yz2j}}\right]+\left(\frac{1}{m}-\frac{1}{n}\right)\left[\beta_{yx}\left\{\frac{C_{3000}}{S_{x}^{2}}-\frac{C_{2100}}{S_{yx}}\right\}+\right.$

$$\left.\beta_{yx}\sum_{j=1}^{p}\beta_{xz1j}\left\{\frac{C_{1110}}{S_{yx}}-\frac{C_{2010}}{S_{x}^{2}}-\frac{C_{0030}}{S_{z1j}^{2}}+\frac{C_{1020}}{S_{xz1j}}\right\}\right] \tag{24}$$

where $C_{abcd}=E\left[\left(x_i-\bar{X}\right)^{a}\left(y_i-\bar{Y}\right)^{b}\left(z_{1ij}-\bar{Z}_{1j}\right)^{c}\left(z_{2ij}-\bar{Z}_{2j}\right)^{d}\right]$; ((a, b, c, d) ≥ 0 are integers), j = 1, 2, - - -, *p*.

Proof: $B(\Delta)=E\left[\Delta-\bar{Y}\right]=\psi B(\Delta_1)+(1-\psi)B(\Delta_2)$

where

$$B(\Delta_1)=E\left[\Delta_1-\bar{Y}\right]$$

$$=E\left[\bar{Y}(1+e_1)-\sum_{j=1}^{p}\beta_{yz2j}\,\bar{Z}_{2j}\,e_{11j}\left(1+e_{7j}\right)\left(1+e_{8j}\right)^{-1}-\bar{Y}\right]$$

Expanding the right-hand side of the above expression binomially, taking expectations and collecting the terms up to the order $o(n^{-1})$, we have

$$B(\Delta_1)=-\left(\frac{1}{u}-\frac{1}{N}\right)\sum_{j=1}^{p}\beta_{yz2j}\left[\frac{C_{0102}}{S_{yz2j}}-\frac{C_{0003}}{S_{z2j}^{2}}\right] \tag{25}$$

Similarly,

$$B(\Delta_2)=E\left[\Delta_2-\bar{Y}\right]$$

$$= E\Big[\bar{Y}(1+e_2) + \beta_{yx}\bar{X}(e_3 - e_4 + e_3e_5 - e_4e_5 - e_3e_6 + e_4e_6)$$

$$+ \sum_{j=1}^{p}\Big\{ - \beta_{yz2j}\bar{Z}_{2j}\left(e_{12j} - e_{12j}e_{8j}^{*} + e_{7j}^{*}e_{12j}\right)$$

$$+ \beta_{xz1j}\ \bar{Z}_{1j}\ \beta_{yx}\Big(e_{13j} - e_{10j}e_{13j} + e_{9j}e_{13j} - e_{14j} - e_{9j}^{*}e_{14j}$$

$$+ e_{14j}e_{10j}^{*} + e_5e_{13j} - e_5e_{14j} - e_6e_{13j} + e_6e_{14j}\Big)\Big\} - \bar{Y}\Big]$$

Again, expanding the right-hand side of the above expression binomially, taking expectations and retaining the terms up to the first order of approximations, we have

$$B(\Delta_2) = \left(\frac{1}{m} - \frac{1}{N}\right)\sum_{j=1}^{p}\beta_{yz2j}\left[\frac{C_{00003}}{S_{z2j}^2} - \frac{C_{0102}}{S_{yz2j}}\right] + \left(\frac{1}{m} - \frac{1}{n}\right)\Bigg[\beta_{yx}\left\{\frac{C_{3000}}{S_x^2} - \frac{C_{2100}}{S_{yx}}\right\}$$

$$+ \beta_{yx}\sum_{j=1}^{p}\beta_{xz1j}\left\{\frac{C_{1110}}{S_{yx}} - \frac{C_{2010}}{S_x^2} - \frac{C_{0030}}{S_{z1j}^2} + \frac{C_{1020}}{S_{xz1j}}\right\}\Bigg] \tag{26}$$

Ignoring fpc, for sufficiently large population size in equations (25) and (26) we get the expressions for the bias of the estimators Δ, Δ_1 and Δ_2 up to the first order of approximations as shown in equations (22), (23) and (24) respectively.

Theorem 4.2: The mean square error of the estimator Δ of the population mean $\bar{Y}$, to the first order of approximations and ignoring fpc, is given by

$$M(\Delta) = \psi^2 M(\Delta_1) + (1-\psi)^2 M(\Delta_2) \tag{27}$$

where $M(\Delta_1) = \dfrac{1}{u} A\, S_y^2$ (28)

and $M(\Delta_2) = \left[\dfrac{1}{m}A + \left(\dfrac{1}{m} - \dfrac{1}{n}\right)B\right]S_y^2$ (29)

Proof: By the definition of mean square error we have

$$M(\Delta) = E\left[\Delta - \bar{Y}\right]^2 = E\left[\psi(\Delta_1 - \bar{Y}) + (1-\psi)(\Delta_2 - \bar{Y})\right]^2$$

$$= \psi^2 M(\Delta_1) + (1-\psi)^2 M(\Delta_2) + 2\psi(1-\psi)E\left[(\Delta_1 - \bar{Y})(\Delta_2 - \bar{Y})\right] \tag{30}$$

where $M(\Delta_1) = E\left[\Delta_1 - \bar{Y}\right]^2$ and $M(\Delta_2) = E\left[\Delta_2 - \bar{Y}\right]^2$

Now, using the expressions given in equations (20) and (21), expanding binomially, taking expectations, retaining the terms up to the first order of approximations, we have the following results

$$\mathrm{MSE}\left(\Delta_1\right) = \left(\frac{1}{u} - \frac{1}{N}\right) A\, S_y^2 \tag{31}$$

$$\mathrm{MSE}\left(\Delta_2\right) = \left[\left(\frac{1}{m} - \frac{1}{N}\right) A + \left(\frac{1}{m} - \frac{1}{n}\right) B\right] S_y^2 \tag{32}$$

Now, ignoring fpc, we have

$$\mathrm{MSE}\left(\Delta_1\right) = \frac{1}{u} A\, S_y^2 \tag{33}$$

$$\mathrm{MSE}\left(\Delta_2\right) = \left[\frac{1}{m} A + \left(\frac{1}{m} - \frac{1}{n}\right) B\right] S_y^2 \tag{34}$$

Since the estimators Δ_1 and Δ_2 are biased estimators based on two independent samples of sizes u and m respectively, for large population size we have

$$E\left[\left(\Delta_1 - \bar{Y}\right)\left(\Delta_2 - \bar{Y}\right)\right] = 0 \tag{35}$$

Substituting the values of $\mathrm{MSE}\left(\Delta_1\right)$, $\mathrm{MSE}\left(\Delta_2\right)$ and $E\left[\left(\Delta_1 - \bar{Y}\right)\left(\Delta_2 - \bar{Y}\right)\right]$ from equations (33), (34) and (35) into equation (30) we get the value of $\mathrm{MSE}\left(\Delta\right)$ as shown in equation (27).

Remark 4.1 From equation (27), it is visible that the mean square error up to the first order of approximations of the estimator Δ is exactly similar to that of the variance of the estimator T in equations (5). This is one of the positive aspects of the estimator Δ. The estimator Δ is based on sample estimates and up to the first order of approximations; it is equally precise to that of estimator T.

5. Special Case

When the p-auxiliary variates are mutually uncorrelated, i.e., $\rho_{zijzik} = 0 \ \forall\ i = 1, 2; j \neq k = 1, 2, ---, p$ then the expression for the optimum value of μ and $V\left(T\right)_{opt.^*}$ reduces to

$$\hat{\mu} = \frac{-A^* \pm \sqrt{A^*(A^* + B^*)}}{B^*} \tag{36}$$

and

$$V(T)_{opt.^*} = \frac{A^*\left[A^* + \mu_0 B^*\right]}{n\,[A^* + \mu_0^2 B^*]} S_y^2 \tag{37}$$

where $A^* = 1 - \sum_{j=1}^{p} \rho_{yz2j}^2$

and $B^* = -\rho_{yx}^2\left(1 + \sum_{j=1}^{p} \rho_{xz1j}^2\right) + 2\rho_{yx} \sum_{j=1}^{p}\left(\rho_{xz1j}\,\rho_{yz1j} + \rho_{yz2j}\,\rho_{xz2j} - \rho_{yz2j}\,\rho_{xz1j}\,\rho_{z2jz1j}\right)$.

6. Efficiency Comparison

The percent relative efficiencies of T with respect to (i) $\bar{y}_n$, when there is no matching and (ii) $\hat{\bar{Y}} = \varphi^* \bar{y}_u + \left(1 - \varphi^*\right)\bar{y}'_m$, when no additional auxiliary information was used at any occasion, where $\bar{y}'_m = \bar{y}_m + \beta_{yx}\left(\bar{x}_n - \bar{x}_m\right)$, have been obtained for different choices of the correlations involved. Since, $\bar{y}_n$ and $\hat{\bar{Y}}$ are unbiased estimators of $\bar{Y}$, following on the line of Sukhatme et al. (1984) the variance of $\bar{y}_n$ and the optimum variance of $\hat{\bar{Y}}$ for large N are respectively given by

$$V\left(\bar{y}_n\right) = \frac{S_y^2}{n} \tag{38}$$

$$V\left(\hat{\bar{Y}}\right)_{opt.^*} = \left[1 + \sqrt{\left(1 - \rho_{yx}^2\right)}\right]\frac{S_y^2}{2n} \tag{39}$$

The percent relative efficiencies E_1 and E_2 of T (under optimal condition) with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$ respectively are given by

$$E_1 = \frac{V\left(\bar{y}_n\right)}{V(T)_{opt.^*}} \times 100 \quad \text{and} \quad E_2 = \frac{V\left(\hat{\bar{Y}}\right)_{opt.^*}}{V(T)_{opt.^*}} \times 100.$$

6.1. Empirical Study

The expressions of the optimum μ (i.e. μ_0) and the percent relative efficiencies E_1 and E_2 are in terms of population correlation coefficients. Therefore, the values of μ_0, E_1 and E_2 have been computed for different choices of positive correlations. For empirical studies, cases of p = 1 and 2 auxiliary variates have been considered.

Case 1: For p = 1, the values of A and B takes the form $A = 1 - \rho^2_{yz21}$ and $B = -\rho^2_{yx}\left(1+\rho^2_{xz11}\right) + 2\,\rho_{yx}\left(\rho_{xz11}\rho_{yz11} + \rho_{xz21}\rho_{yz21} - \rho_{yz21}\rho_{xz11}\rho_{z21z11}\right)$. For convenience we assume $\rho_{xz11} = \rho_{xz21} = \rho_{yz11} = \rho_{yz21} = \rho_1$. Hence, the values of A and B becomes $A = 1 - \rho_1^2$ and $B = -\rho^2_{yx}\left(1+\rho_1^2\right) + 2\rho_{yx}\rho_1^2\left(2 - \rho_{z21z11}\right)$, which is the work of Priyanka (2008).

Case 2: For p = 2 and assuming that the two auxiliary variates are mutually correlated i.e., $\rho_{zijzik} \neq 0$. The values of A and B are given by

$$A = 1 - \rho^2_{yz21} - \rho^2_{yz22} + 2\rho_{yz21}\rho_{yz22}\rho_{z22z21}$$

and

$$\begin{aligned} B = & -\rho^2_{yx}\left(1+\rho^2_{xz11}+\rho^2_{xz12}\right) + 2\rho_{yx}\left(\rho_{xz11}\rho_{yz11} + \rho_{yz21}\rho_{xz21} - \rho_{yz21}\rho_{xz11}\rho_{z21z11}\right. \\ & \left. + \rho_{xz12}\rho_{yz12} + \rho_{yz22}\rho_{xz22} - \rho_{yz22}\rho_{xz12}\rho_{z22z12}\right) \\ & + 2\rho^2_{yx}\rho_{xz11}\rho_{xz12}\rho_{z11z12} - 4\rho_{yx}\rho_{yz21}\rho_{xz12}\rho_{z21z12} \end{aligned}$$

For convenience we assume

$$\rho_{xz11} = \rho_{xz21} = \rho_{yz11} = \rho_{yz21} = \rho_2,$$

$$\rho_{xz12} = \rho_{xz22} = \rho_{yz12} = \rho_{yz22} = \rho_3$$

and

$$\rho_{z11z12} = \rho_{z21z22} = \rho_{z11z22} = \rho_{z12z21} = \rho_4.$$

In the light of these assumptions, the values of A and B are given as

$$A = 1 - \rho_2^2 - \rho_3^2 + 2\rho_2\rho_3\rho_4$$

and

$$B = -\rho^2_{yx}\left(1+\rho_2^2+\rho_3^2 - 2\rho_2\rho_3\rho_4\right) + 2\rho_{yx}\left(2\rho_2^2 - \rho_2^2\rho_{z21z11} + 2\rho_3^2 - \rho_3^2\rho_{z22z12} - 2\rho_2\rho_3\rho_4\right).$$

Substituting these values of A and B in equations (18) and (19), we have the values of optimum μ and $V(T)_{opt.^*}$ respectively. For different choices of correlations, Tables 1-3 show the optimum values of i.e., μ_0 and percent relative efficiencies E_1 and E_2 of T (under optimal condition) with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$ respectively.

Table 1. Optimum values of μ and Percent Relative Efficiencies of T with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$

$\rho_{z21z11} = 0.3$, $\rho_{z22z12} = 0.4$											
ρ_4			0.1			0.3			0.5		
ρ_3	ρ_2	ρ_{yx} ↓	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.4	0.5	0.3	0.46	144.9	141.7	0.47	131.4	128.3	0.47	120.1	117.4
		0.6	0.45	144.4	29.92	0.47	131.8	118.6	0.48	121.4	109.2
		0.9	0.49	156.1	112.0	0.51	142.6	102.4	0.52	131.5	**
	0.7	0.3	0.40	199.1	194.5	0.43	164.6	160.8	0.44	140.8	137.5
		0.6	0.38	187.7	168.9	0.41	157.6	141.8	0.41	136.5	122.8
		0.9	0.39	193.3	138.7	0.42	162.8	116.9	0.45	141.4	101.5
	0.9	0.3	0.25	500.6	489.1	0.34	275.3	269.0	0.38	196.6	192.1
		0.6	0.22	430.1	387.1	0.31	245.5	220.9	0.35	179.8	161.8
		0.9	0.21	419.8	301.4	0.30	241.6	173.5	0.35	178.1	127.9
0.6	0.5	0.3	0.42	185.6	181.3	0.44	153.9	150.3	0.45	131.8	128.7
		0.6	0.40	177.3	159.6	0.43	149.2	134.3	0.45	129.4	116.5
		0.9	0.42	184.8	132.7	0.44	156.0	112.0	0.47	135.7	**
	0.7	0.3	0.34	291.9	285.2	0.39	195.9	191.4	0.43	149.5	146.0
		0.6	0.30	262.7	236.4	0.36	181.7	163.5	0.40	141.7	127.5
		0.9	0.31	262.7	188.6	0.37	182.86	131.3	0.41	143.4	103.0
	0.9	0.3	*	–	–	0.28	370.7	362.1	0.37	200.1	195.5
		0.6	*	–	–	0.25	319.7	287.7	0.33	180.2	162.2
		0.9	*	–	–	0.24	308.2	221.3	0.33	175.8	126.2
0.8	0.5	0.3	0.32	338.1	330.3	0.38	217.5	212.5	0.42	163.1	159.4
		0.6	0.28	301.3	271.2	0.35	200.0	180.0	0.39	153.5	138.2
		0.9	0.29	300.7	215.9	0.35	200.9	144.3	0.39	155.1	111.4
	0.7	0.3	*	–	–	0.31	307.7	300.6	0.39	180.6	176.4
		0.6	*	–	–	0.28	270.2	243.2	0.35	165.1	148.5
		0.9	*	–	–	0.27	263.1	188.9	0.35	162.4	116.6
	0.9	0.3	*	–	–	*	–	–	0.33	243.7	238.1
		0.6	*	–	–	*	–	–	0.29	212.7	191.4
		0.9	*	–	–	*	–	–	0.27	203.0	145.7

Note: ** denotes μ_0 does not exist and ** indicate no gain.*

Table 2. Optimum values of μ and Percent Relative Efficiencies of T with respect to $\overline{y}_n$ and $\hat{\hat{Y}}$

$\rho_{z21z11} = 0.7$, $\rho_{z22z12} = 0.5$											
ρ_4			0.1			0.3			0.5		
ρ_3	ρ_2	$\rho_{yx}\downarrow$	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.4	0.5	0.3	0.47	148.2	144.8	0.48	134.1	131.0	0.48	122.5	119.6
		0.6	0.48	150.9	135.8	0.49	137.6	123.8	0.50	126.5	113.8
		0.9	0.54	170.7	122.6	0.55	155.6	111.7	0.57	143.1	102.8
	0.7	0.3	0.42	208.4	203.6	0.44	171.6	167.6	0.46	146.2	142.8
		0.6	0.41	203.7	183.3	0.44	170.1	153.1	0.46	146.7	132.0
		0.9	0.45	223.1	160.2	0.48	186.5	133.9	0.51	161.0	115.6
	0.9	0.3	0.29	552.2	539.5	0.37	298.1	291.2	0.41	210.5	205.6
		0.6	0.25	497.2	447.5	0.34	278.4	250.6	0.39	201.5	181.4
		0.9	0.27	525.1	377.0	0.36	294.3	211.3	0.42	213.6	153.4
0.6	0.5	0.3	0.43	190.9	186.5	0.45	157.9	154.2	0.46	134.9	131.8
		0.6	0.42	186.7	168.0	0.45	156.6	140.9	0.47	135.5	121.9
		0.9	0.46	202.4	145.3	0.48	170.0	122.1	0.51	147.3	105.8
	0.7	0.3	0.36	308.9	301.7	0.41	205.4	200.7	0.44	155.7	152.1
		0.6	0.34	287.6	258.8	0.40	196.9	177.2	0.43	152.5	137.3
		0.9	0.35	303.6	218.0	0.42	208.5	149.7	0.46	162.0	116.3
	0.9	0.3	*	–	–	0.31	405.1	395.8	0.40	214.5	209.5
		0.6	*	–	–	0.28	364.4	328.0	0.37	201.5	181.4
		0.9	*	–	–	0.29	373.8	268.4	0.38	208.0	149.4
0.8	0.5	0.3	0.33	352.4	344.3	0.39	225.1	219.9	0.43	168.1	164.2
		0.6	0.31	321.3	289.2	0.37	211.7	190.6	0.41	161.7	145.5
		0.9	0.32	332.3	238.6	0.38	219.9	157.9	0.43	168.6	121.1
	0.7	0.3	*	–	–	0.34	326.7	319.1	0.41	189.4	185.1
		0.6	*	–	–	0.30	295.8	266.2	0.38	178.5	160.6
		0.9	*	–	–	0.31	300.8	215.9	0.39	182.8	131.3
	0.9	0.3	*	–	–	*	–	–	0.36	263.1	257.0
		0.6	*	–	–	*	–	–	0.32	238.4	214.6
		0.9	*	–	–	*	–	–	0.32	238.6	171.3

Note: ** denotes* μ_0 *does not exist.*

Table 3. Optimum values of μ and Percent Relative Efficiencies of T with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$

$\rho_{z21z11} = 0.9$, $\rho_{z22z12} = 0.8$											
ρ_4			0.1			0.3			0.5		
ρ_3	ρ_2	ρ_{yx} ↓	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.4	0.5	0.3	0.48	151.1	147.7	0.48	136.6	133.4	0.49	124.6	121.7
		0.6	0.49	157.5	141.8	0.51	143.3	129.0	0.52	131.6	118.4
		0.9	0.59	189.3	135.9	0.61	171.9	123.4	0.62	157.6	113.1
	0.7	0.3	0.44	215.8	210.9	0.46	177.0	173.0	0.47	150.4	146.9
		0.6	0.44	218.2	196.4	0.47	181.4	163.2	0.49	155.8	140.2
		0.9	0.53	259.1	186.0	0.56	214.6	154.1	0.58	183.9	132.0
	0.9	0.3	0.30	594.9	581.2	0.39	316.0	308.8	0.43	221.1	216.0
		0.6	0.29	562.0	505.8	0.38	308.9	278.0	0.43	221.2	199.1
		0.9	0.34	666.6	478.6	0.44	361.0	259.1	0.50	256.9	184.4
0.6	0.5	0.3	0.44	197.9	193.3	0.46	163.0	159.2	0.48	138.8	135.6
		0.6	0.46	200.3	180.3	0.48	167.2	150.5	0.50	144.1	129.7
		0.9	0.53	234.7	168.5	0.56	195.5	140.3	0.58	168.1	120.7
	0.7	0.3	0.38	327.3	319.8	0.43	215.4	210.4	0.46	162.2	158.4
		0.6	0.37	318.6	286.7	0.43	215.4	193.9	0.47	165.4	148.9
		0.9	0.43	370.4	265.9	0.50	248.8	178.6	0.54	190.6	136.8
	0.9	0.3	*	–	–	0.34	439.8	429.7	0.42	228.2	222.9
		0.6	*	–	–	0.32	417.0	375.3	0.42	225.4	202.9
		0.9	*	–	–	0.37	477.0	342.5	0.47	255.7	183.6
0.8	0.5	0.3	0.36	378.1	369.3	0.42	238.4	232.9	0.45	176.5	172.4
		0.6	0.34	361.6	325.4	0.41	234.8	211.3	0.45	177.5	159.8
		0.9	0.39	412.6	296.2	0.47	266.4	191.3	0.51	201.1	144.4
	0.7	0.3	*	–	–	0.36	353.7	345.6	0.43	201.4	196.8
		0.6	*	–	–	0.35	337.5	303.8	0.43	199.6	179.6
		0.9	*	–	–	0.39	379.4	272.4	0.48	223.5	160.4
	0.9	0.3	*	–	–	*	–	–	0.39	286.6	280.0
		0.6	*	–	–	*	–	–	0.37	275.0	247.5
		0.9	*	–	–	*	–	–	0.41	303.7	218.1

***Note:** * denotes μ_0 does not exist.*

Case3: For p = 2 and assuming that the two auxiliary variates are mutually uncorrelated i.e., $\rho_{zijzik} = 0$. The values of A^* and B^* are given by

$$A^* = 1 - \rho_2^2 - \rho_3^2$$

and $B^* = -\rho_{yx}^2\left(1+\rho_2^2+\rho_3^2\right)+2\rho_{yx}\left(2\rho_2^2-\rho_2^2\rho_{z21z11}+2\rho_3^2-\rho_3^2\rho_{z22z12}\right)$

Substituting A^* and B^* in equations (36) and (37), Tables 4 – 6, show the optimum values of μ, percent relative efficiencies E_1 and E_2 for different choices of correlations under the present assumption.

Table 4. Optimum values of μ and Percent Relative Efficiencies of T with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$ when auxiliary variates are uncorrelated

$\rho_{z22z12} = 0.3$											
ρ_{z21z11}			0.2			0.5			0.8		
ρ_3	ρ_2	$\rho_{yx}\downarrow$	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.4	0.5	0.3	0.45	151.8	148.3	0.45	154.0	150.5	0.46	156.4	152.7
		0.6	0.44	149.5	134.5	0.45	153.6	138.3	0.47	158.3	142.5
		0.9	0.47	159.5	114.5	0.50	167.9	120.5	0.53	178.4	128.1
	0.7	0.3	0.38	219.9	214.9	0.40	227.0	221.8	0.41	235.0	229.6
		0.6	0.36	203.7	183.3	0.38	214.8	193.3	0.40	228.5	205.7
		0.9	0.36	206.6	148.4	0.39	225.3	161.8	0.44	252.1	181.0
	0.9	0.3	0.16	1033.8	1010.0	0.17	1114.3	1088.6	0.18	1219.0	1191.0
		0.6	0.13	852.1	766.9	0.14	943.0	848.7	0.16	1073.9	966.5
		0.9	0.12	814.6	584.8	0.14	940.9	675.5	0.17	1161.8	834.1
0.6	0.5	0.3	0.40	204.8	200.1	0.41	208.0	203.2	0.41	211.4	206.6
		0.6	0.37	192.2	173.0	0.39	197.5	177.7	0.40	203.2	182.9
		0.9	0.38	197.3	197.3	0.40	206.2	148.0	0.42	216.8	155.6
	0.7	0.3	0.29	389.6	380.7	0.30	404.4	395.0	0.32	421.2	411.5
		0.6	0.26	340.4	306.3	0.27	359.5	323.6	0.29	382.9	344.6
		0.9	0.25	334.2	239.9	0.27	362.5	260.2	0.30	400.9	287.8
	0.9	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–
0.8	0.5	0.3	0.26	478.4	467.4	0.27	487.9	476.7	0.27	498.1	486.6
		0.6	0.23	412.8	371.5	0.23	424.6	382.1	0.24	437.5	393.8
		0.9	0.22	404.0	290.1	0.23	420.9	302.2	0.24	440.3	316.1
	0.7	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–
	0.9	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–

Note: ** denotes* μ_0 *does not exist.*

Table 5. Optimum values of μ and Percent Relative Efficiencies of T with respect to $\overline{y}_n$ and $\hat{\overline{Y}}$ when auxiliary variates are uncorrelated

$\rho_{z22z12} = 0.5$											
ρ_{z21z11}			0.2			0.5			0.8		
ρ_3	ρ_2	$\rho_{yx}\downarrow$	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.4	0.5	0.3	0.45	152.7	149.2	0.46	154.9	151.4	0.46	157.3	153.7
		0.6	0.45	151.2	136.1	0.46	155.6	140.0	0.47	160.5	144.4
		0.9	0.48	162.9	116.9	0.51	172.1	123.6	0.54	183.9	132.0
	0.7	0.3	0.39	221.4	216.3	0.40	228.7	223.4	0.41	236.9	231.4
		0.6	0.36	205.9	185.3	0.38	217.6	195.8	0.41	232.0	208.8
		0.9	0.37	210.2	150.9	0.40	230.2	165.3	0.45	259.8	186.5
	0.9	0.3	0.16	1043.3	1019.3	0.17	1126.4	1100.5	0.19	1235.3	1206.9
		0.6	0.13	862.5	776.3	0.14	957.4	861.7	0.16	1095.9	986.3
		0.9	0.12	828.2	594.6	0.14	962.6	691.1	0.18	1205.3	865.4
0.6	0.5	0.3	0.41	207.9	203.1	0.41	211.3	206.4	0.42	214.9	210.0
		0.6	0.38	197.2	177.5	0.40	203.0	182.7	0.41	209.4	188.5
		0.9	0.40	205.8	147.8	0.42	216.3	155.3	0.45	229.1	164.5
	0.7	0.3	0.30	396.6	387.5	0.31	412.3	402.8	0.32	430.4	420.5
		0.6	0.26	349.3	314.4	0.28	370.4	333.3	0.30	396.5	356.8
		0.9	0.26	347.1	249.2	0.28	379.7	272.6	0.32	425.7	305.6
	0.9	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–
0.8	0.5	0.3	0.27	495.0	483.6	0.28	505.7	494.1	0.28	517.2	505.3
		0.6	0.24	433.6	390.2	0.25	447.6	402.8	0.25	463.1	416.8
		0.9	0.24	434.3	311.8	0.25	455.9	327.3	0.26	481.6	345.8
	0.7	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–
	0.9	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–

Note: ** denotes* μ_0 *does not exist.*

Table 6. Optimum values of μ and Percent Relative Efficiencies of T with respect to $\overline{y}_n$ and $\hat{\hat{Y}}$ when auxiliary variates are uncorrelated

$\rho_{z22z12} = 0.9$											
ρ_{z21z11}			0.2			0.5			0.8		
ρ_3	ρ_2	$\rho_{yx}\downarrow$	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
0.4	0.5	0.3	0.46	154.6	151.1	0.46	156.9	153.3	0.47	159.4	155.7
		0.6	0.46	154.9	139.4	0.47	159.7	143.7	0.49	165.1	148.6
		0.9	0.50	170.6	122.5	0.54	181.9	130.6	0.58	197.2	141.6
	0.7	0.3	0.39	224.5	219.3	0.41	232.1	226.8	0.42	240.8	235.2
		0.6	0.37	210.7	189.6	0.39	223.4	201.1	0.42	239.4	215.5
		0.9	0.38	218.1	156.6	0.42	241.5	173.4	0.49	278.2	199.8
	0.9	0.3	0.16	1063.3	1038.8	0.17	1152.1	1125.6	0.19	1270.2	1240.9
		0.6	0.13	884.6	796.2	0.15	988.5	889.6	0.17	1144.5	1030.1
		0.9	0.13	857.9	615.9	0.15	1011.3	726.1	0.20	1311.0	941.2
0.6	0.5	0.3	0.42	214.6	209.7	0.43	218.5	213.5	0.43	222.6	217.5
		0.6	0.41	208.9	188.0	0.42	216.1	194.5	0.44	224.3	201.8
		0.9	0.44	228.0	163.7	0.48	244.0	175.1	0.52	265.4	190.6
	0.7	0.3	0.31	412.0	402.5	0.32	430.0	420.1	0.34	451.1	440.7
		0.6	0.28	369.9	333.0	0.30	395.9	356.3	0.32	429.4	386.4
		0.9	0.28	378.9	272.0	0.32	424.6	304.8	0.37	497.8	357.4
	0.9	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–
0.8	0.5	0.3	0.29	535.1	522.8	0.30	549.2	536.5	0.31	564.6	551.6
		0.6	0.27	488.6	439.8	0.28	509.8	458.9	0.29	534.5	481.1
		0.9	0.27	528.1	379.2	0.31	572.3	410.9	0.35	632.6	454.2
	0.7	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–
	0.9	0.3	*	–	–	*	–	–	*	–	–
		0.6	*	–	–	*	–	–	*	–	–
		0.9	*	–	–	*	–	–	*	–	–

Note: ** denotes μ_0 does not exist.*

7. Discussion and Conclusion

Following conclusions can be made from the Tables 1–6:

1. For fixed values of ρ_{yx}, ρ_{z21z11}, ρ_{z22z12}, ρ_3 and ρ_4, the values of μ_0 decrease with the increase in the values of ρ_2 and the values of E_1 and E_2 increase with the increase in the values of ρ_2. Similar trends are visible for increasing values of ρ_3 and ρ_4 provided other correlations are fixed. This pattern indicates that smaller fresh sample is required at current occasion, if highly correlated auxiliary variates are available. The increase in precision indicates the efficient utilization of available auxiliary variates.

2. For fixed values of ρ_2, ρ_{z21z11}, ρ_{z22z12}, ρ_3 and ρ_4, the values of μ_0 first decrease and then increase with the increase in the values of ρ_{yx}, however the values of E_1 and E_2 decrease uniformly with the increase in the values of ρ_{yx}.

3. For fixed values of ρ_2, ρ_3, ρ_4 and ρ_{yx}, it is observed that the values of μ_0, E_1 and E_2 increase with the increase in the values of ρ_{z21z11} and ρ_{z22z12}. This indicates that higher mutual correlation between the auxiliary variates results in increase in precision of estimates at current occasion.

4. A close perusal of Tables 1–3, revels that μ_0 attains the minimum value of 0.21 if two dynamic auxiliary variates are used. However, if only one dynamic auxiliary variate is considered, the minimum value acquired by μ_0 is 0.2729 (Priyanka (2008)).This indicates that use of more than one dynamic auxiliary variate is fruitful in sampling on two occasions.

5. From Tables 4–6, i.e., when the auxiliary variates are mutually uncorrelated, it is observed that the precision of estimates increases much more as compared to the situation when the auxiliary variates are mutually correlated. The minimum value attained by μ_0 for this case is 0.12, which is very low as compared to the previous situation. Hence, the cost of survey will reduce to a greater extent.

Maximum gain and minimum μ_0 is observed for higher values of ρ_2. Hence, the use of varying (dynamic) p-auxiliary variates is highly advantageous in terms of the proposed estimators. They not only provide estimates with increased accuracy but also tremendously reduce the cost of the survey. Thus, the proposed estimators may be recommended for its further practical use.

Acknowledgement

Authors are thankful to the honourable referee for his inspiring comments. Authors are also thankful to the Indian School of Mines, Dhanbad for providing the financial assistance to carry out the present work.

REFERENCES

BIRADAR, R. S. and SINGH, H. P. (2001). Successive sampling using auxiliary information on both occasions. Cal. Stat. 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.

COCHRAN, W. G. (1977): Sampling Techniques. New York: John Wiley & Sons.

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

ECKLER, A. R. (1955). Rotation sampling. Ann. Math. Stat., 26, 664–685.

FENG, S. and ZOU, G. (1997). Sample rotation method with auxiliary variable. Commun. Statist. Theo-Meth. 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. In: Iowa Agricultural Experiment Station Road Bulletin No. 304: 1–104, Ames, USA.

PATTERSON, H. D. (1950). Sampling on successive occasions with partial replacement of units. Jour. Royal Statist. Assoc., Ser. B, 12: 241–255.

PRIYANKA, K. (2008). On the use of model based techniques and development of estimation procedures in search of good rotation patterns on successive occasions and its applications. Unpublished Ph. D. thesis submitted to the Indian School of mines, Dhanbad.

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, Ser. B, 33: 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. and SINGH, V. K. (2001). On the use of auxiliary information in successive sampling. Jour. Ind. Soc. Agric. Statist. 54: 1–12.

SINGH, G. N. (2003). Estimation of population mean using auxiliary information on recent occasion in h occasions 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 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. BSE 1, A07: 42–60.

SINGH, G. N. and PRIYANKA, K. (2008 a). Search of good rotation patterns to improve the precision of estimates at current occasion. Commun. Statist. Theo-Meth 37, 3: 337–348.

SINGH, G. N. and PRIYANKA, K. (2008 b). On the use of several auxiliary variates to improve the precision of estimates at current occasion. Journal of Indian Society of Agricultural Statistics, 62, 3, 253–265.

SUKHATME, P. V., SUKHATME, B. V., SUKHATME, S. and ASOK, C. (1984). *Sampling theory of surveys with applications*. Iowa State University Press, Ames, Iowa (USA) and Indian Society of Agricultural Statistics, New Delhi (India).