Hacettepe Journal of Mathematics and Statistics
Volume 47 (5) (2018), 1268 – 1280

Almost unbiased estimation procedures of population mean in two-occasion successive sampling

G. N. Singh*†, A. K. Singh‡ and Cem Kadilar§

Abstract

The objective of this paper is to construct some unbiased estimators of the current population mean in two-occasion successive sampling. Utilizing the readily available information on an auxiliary variable on both occasions, almost unbiased ratio and regression cum exponential type estimators of current population mean have been proposed. Theoretical properties of the proposed estimation procedures have been examined and their respective optimum replacement strategies are formulated. Performances of the proposed estimators are empirically compared with (i) the sample mean estimator, when no sample units were matched from the previous occasion and (ii) natural successive sampling estimator when no auxiliary information was used on any occasion. Empirical results are critically interpreted and suitable recommendations are made to the survey practitioners for their practical applications.

Keywords: Successive sampling, Auxiliary information, Bias, Mean square error, Optimum replacement strategy.

Mathematics Subject Classification (2010): 62D05

Received : 09/01/2016 *Accepted* : 08/11/2016 *Doi* : 10.15672HJMS.2016.391

*Department of Applied Mathematics, Indian Institute of Technology(Indian School of Mines), Dhanbad, India, Email: gnsingh_ism@yahoo.com

†Corresponding Author

‡Department of Applied Mathematics, Indian Institute of Technology(Indian School of Mines),Dhanbad, India, Email: aksingh.ism@gmail.com

§Department of Statistics, Hacettepe University, Turkey, Email: kadilar@hacettepe.edu.tr

1. Introduction

Change is the inherent phenomenon of the nature, if such change affects the human life, it is necessary to observe its behaviour and pattern over the period of time. If the change is to be observed for a large group of individuals (population), one time survey does not provide the relevant information over a period of time. Successive (rotation) sampling happens to be more reasonable statistical tool to generate the estimates of unknown population parameters on different points of time (occasions) and it is also capable of providing the information related to the patterns of variation in characteristics under study over the period of time. The problem of sampling on two successive occasions with a partial replacement of sampling units was first considered by Jessen (1942) in the analysis of a survey related to the farm data. The theory of successive sampling was further extended by Patterson (1950), Rao and Graham (1964), Gupta (1979), Das (1982) and Chaturvedi and Tripathi (1983) among others. Sen (1971, 1972, 1973) used the information on auxiliary variables from previous occasion and developed the estimators of current population mean in two occasions successive sampling. Later on Singh *et al.* (1991) and Singh and Singh (2001) used auxiliary information on current occasion in two occasions 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 transportation survey, many other examples may be cited where the information on auxiliary variables are available on both the occasion in two occasions successive sampling. Utilizing the auxiliary information on both occasions , Feng and Zou (1997), Birader and Singh (2001), Singh (2005), Singh and Priyanka (2008, 2010), Singh and Karna (2009), Singh and Vishwakarma (2009), Singh and Prasad (2013), Singh and Homa (2013), Singh and Pal (2015, 2016) and Srivastava and Srivastava (2016) among others have proposed varieties of estimators of population mean on current (second) occasions in two occasion successive sampling.

It is to be mentioned that above works describe the biased estimation procedures of current population mean in two occasions successive sampling. Bias is an important factor in degrading the performance of estimators, keeping this point in mind and motivated with the cited works we have suggested some unbiased estimation procedures of current population mean in two occasions successive sampling. Utilizing the information on readily available auxiliary information on both occasions, almost unbiased ratio and regression cum exponential type estimators of current population mean have been proposed and their properties are studied. The dominance of proposed estimation procedures have been shown over sample mean and natural successive sampling estimators through empirical studies. The results of empirical studies are critically analyzed and suitable recommendations are put forward to the survey practitioners.

2. Development of the Estimators

Consider a finite population $U = (U_1, U_2, ..., U_N)$ of N units which has been sampled over two occasions. The character under study is denoted by $x(y)$ on the first (second) occasion respectively. It is assumed that the information on an auxiliary variable z (stable over occasions) whose population mean is known, is readily available on both occasions and has positive correlation with x and y on the first and second occasions respectively. Let a simple random sample (without replacement) of size n be drawn on the first occasion. A random sub-sample of size $m = n\lambda$ is retained (matched) from the sample on first occasion for its use on the second occasion, while a fresh simple random sample (without replacement) of size $u = (n - m) = n\mu$ is drawn on the second occasion from the entire population so that the total sample size on this occasion is also n. Here

λ and μ $(\lambda+\mu=1)$ are the fractions of the matched and fresh samples, respectively, on the current (second) occasion. The values of λ or μ would be chosen optimally as they directly affect the cost of the survey.

The following notations are considered for further use:

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

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

$\rho_{yx},\rho_{yz},\rho_{xz}$: The population correlation coefficients between the variables shown in subscripts.

$S_x^2{=}(N-1)^{-1}\sum\limits_{i=1}^{N}(x_i-\bar{X})^2$: The population variance of the variable x.

S_y^2,S_z^2: The population variance of the variables y and z respectively.

$s_x^2(m)$: The sample variance of the variable x based on the matched sample of size m.

C_y,C_x,C_z: The coefficients of variation of the variables shown in subscripts.

$b_{yz}(u),b_{yx}(m)$: The sample regression coefficients between the variables shown in subscripts and based on the sample size shown in braces.

S_{yx},S_{xz},S_{yz}: The population covariances between the variables shown in subscripts.

$s_{yz}(u),s_{yx}(m)$: The sample covariances between the variables shown in subscripts and based on the sample sizes indicated in braces.

To estimate the population mean $\bar{Y}$ on the current (second) occasion, two sets of estimators are considered. The first set of estimators $S_u=\{T_{1u},T_{2u},T_{3u}\}$ is based on sample of size $u=n\mu$ drawn afresh on the second occasion and the second set of estimators $S_m=\{T_{1m},T_{2m}\}$ is based on the sample of size $m(=n\lambda)$ common with both the occasions, since, information on y is collected for the fresh sample of size u, therefore, we define following exponential type estimators of set S_u as

(2.1) $$T_{1u}{=}\bar{y}_u\sum_{i=1}^{3}a_ie^{\left\{i\left[\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right]\right\}}$$

(2.2) $$T_{2u}{=}\bar{y}_u^*\sum_{i=1}^{3}b_ie^{\left\{i\left[\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right]\right\}}$$

where $\bar{y}_u^*=\bar{y}_u+b_{yz}^{(u)}(\bar{Z}-\bar{z}_u)$

and

(2.3) $$T_{3u}{=}\bar{y}_u^{**}\sum_{i=1}^{3}c_ie^{\left\{i\left[\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right]\right\}}$$

where $\bar{y}_u^{**}=\frac{\bar{y}_u}{\bar{z}_u}\bar{Z}$, a_i, b_i and c_i (i=1,2,3) are suitably chosen scalars (weight) such that $\sum\limits_{i=1}^{3}a_i=\sum\limits_{i=1}^{3}b_i=\sum\limits_{i=1}^{3}c_i=1$ and $a_i,b_i,c_i\in R$(set of real numbers).

Since, the information on study variable y is also available for the matched sample of size m retained on second occasion from the sample on first occasion, therefore, again we suggest following ratio and regression cum exponential type estimators of population mean $\bar{Y}$ and belong to the set S_m

(2.4) $$T_{1m}{=}\bar{y}_m^*\sum_{i=1}^{3}\alpha_ie^{\left\{i\left[\frac{\bar{Z}-\bar{z}_m}{\bar{Z}+\bar{z}_m}\right]\right\}}$$

where $\bar{y}_m^*=\frac{\bar{y}_m}{\bar{x}_m}\bar{x}_n$

and

(2.5) $$T_{2m}{=}\bar{y}_m^{**}\sum_{i=1}^{3}\beta_ie^{\left\{i\left[\frac{\bar{Z}-\bar{z}_m}{\bar{Z}+\bar{z}_m}\right]\right\}}$$

where $\bar{y}_m^{**} = \bar{y}_m + b_{yx}^{(m)}(\bar{x}_n - \bar{x}_m)$

where α_i and β_i (i=1,2,3) are suitably chosen scalars such that $\sum_{i=1}^{3}\alpha_i = \sum_{i=1}^{3}\beta_i = 1$ and $\alpha_i, \beta_i \in R$

Considering the convex linear combinations of the estimators of the sets S_u and S_m , we have the final estimators of the population mean $\bar{Y}$ on the current (second) occasion as

$$(2.6) \qquad T_{ij} = \varphi_{ij}T_{iu} + (1-\varphi_{ij})T_{jm}(i = 1,2,3; j = 1,2)$$

where $\varphi_{ij}(i = 1,2,3; j = 1,2)$ are the unknown constants and to be determined under certain criterions.

2.1. Remark. The estimators $T_{iu}(i = 1,2,3)$ and $T_{jm}(j = 1,2)$ are proposed under the following conditions:

(1) The sums of their respective weights are one.

(2) The weights of the linear forms are chosen so that approximate biases are zero.

(3) The approximate mean square errors attain minimum.

3. Properties of the Proposed Estimators

3.1. Bias and variance of the proposed estimators. Since the estimators $T_{iu}(i = 1,2,3)$ and $T_{jm}(j = 1,2)$ defined in equations (2.1)-(2.5) are simple exponential, ratio and regression cum exponential type estimators, they are biased estimators of $\bar{Y}$. Following the Remark 2.1, the proposed estimators may be converted in form of almost unbiased estimators of $\bar{Y}$. The variances $V(.)$ up-to first order of sample sizes of these estimators are derived under large sample approximations using the following transformations:
$\bar{y}_u = (1+e_1)\bar{Y}, \bar{y}_m = (1+e_2)\bar{Y}, \bar{x}_m = (1+e_3)\bar{X}, \bar{x}_n = (1+e_4)\bar{X}, \bar{z}_u = (1+e_5)\bar{Z}, \bar{z}_m = (1+e_6)\bar{Z}, s_{yz}^{(u)} = (1+e_7)S_{yz}, s_z^2(u) = (1+e_8)S_z^2, s_{yx}^{(m)} = (1+e_9)S_{yx}$, $s_x^2(m) = (1+e_{10})S_x^2$ such that $E(e_k) = 0$ and $\mid e_k \mid < 1, \forall k = 1,2,...,10$ Under the above transformations the estimators $T_{iu}(i = 1,2,3)$ and $T_{jm}(j = 1,2)$ take the following forms:

$$(3.1) \qquad T_{1u} = \sum_{i=1}^{3} a_i\bar{Y}(1+e_1)expi\big[-\frac{e_5}{2}\big(1+\frac{e_5}{2}\big)^{-1}\big]$$

$$(3.2) \qquad T_{2u} = \sum_{i=1}^{3} b_i[\bar{Y}(1+e_1) - \bar{Z}\beta_{yz}e_5(1+e_7)(1+e_8)^{-1}]expi\big[-\frac{e_5}{2}\big(1+\frac{e_5}{2}\big)^{-1}\big]$$

$$(3.3) \qquad T_{3u} = \sum_{i=1}^{3} c_i\bar{Y}(1+e_1)(1+e_5)^{-1}expi\big[-\frac{e_5}{2}\big(1+\frac{e_5}{2}\big)^{-1}\big]$$

$$(3.4) \qquad T_{1m} = \sum_{i=1}^{3} \alpha_i\bar{Y}(1+e_2)(1+e_4)(1+e_3)^{-1}expi\big[-\frac{e_6}{2}\big(1+\frac{e_6}{2}\big)^{-1}\big]$$

and

$$(3.5) \qquad T_{2m} = \sum_{i=1}^{3} \beta_i[\bar{Y}(1+e_2) - \bar{X}\beta_{yx}(e_4-e_3)(1+e_9)(1+e_{10})^{-1}]expi\big[-\frac{e_6}{2}\big(1+\frac{e_6}{2}\big)^{-1}\big]$$

where β_{yz} and β_{yx} are population regression coefficients between the variables shown in subscripts.

To derived the bias and variances/mean square errors of the proposed estimators, the following expectations ignoring the finite population corrections are used
$E(e_1^2) = u^{-1}C_y^2$, $E(e_2^2) = m^{-1}C_y^2$, $E(e_3^2) = m^{-1}C_x^2$, $E(e_4^2) = n^{-1}C_x^2$, $E(e_5^2) = u^{-1}C_z^2$, $E(e_6^2) = m^{-1}C_z^2$, $E(e_1e_5) = u^{-1}\rho_{yz}C_yC_z$, $E(e_2e_4) = n^{-1}\rho_{yx}C_yC_x$,

$E(e_2e_3) = m^{-1}\rho_{yx}C_yC_x$, $E(e_3e_4) = n^{-1}C_x^2$, $E(e_2e_6) = m^{-1}\rho_{yz}C_yC_z$, $E(e_4e_6) = n^{-1}\rho_{xz}C_xC_z$, $E(e_3e_6) = m^{-1}\rho_{xz}C_xC_z$.
To derive the expressions of bias and mean square error of the estimator T_{1u} , we expand the right hand side of equation (3.1) binomially and exponentially and neglecting terms of $e^{'}s$ having power greater than two, we have

$$(3.6) \qquad T_{1u} - \bar{Y} = \bar{Y}\{e_1 - A\big(\frac{e_5}{2} + \frac{e_1e_5}{2} - \frac{e_5^2}{4}(1 + \frac{i}{2})\big)\}$$

where

$$(3.7) \qquad A = \sum_{i=1}^{3} ia_i$$

Taking expectation on both the sides of the equation (3.6) and for large population size, ignoring finite population correction (f.p.c), we get the bias of the estimator T_{1u} to the first order of approximations as

$$(3.8) \qquad B(T_{1u}) = E(T_{1u} - \bar{Y}) = \bar{Y}f_uA\big[\frac{1}{4}(1 + \frac{i}{2})C_z^2 - \rho_{yz}C_yC_z\big]$$

where $f_u = \frac{1}{u}$.
Squaring both sides of equation (3.6), neglecting the terms involving power of $e^{'}s$ greater than two and taking expectations, we get the mean square error (MSE) of the estimator T_{1u} to the first order of approximations as

$$(3.9) \qquad M(T_{1u}) = E(T_{1u} - \bar{Y})^2 = \bar{Y}^2\big[f_uC_y^2 + f_uA^2\frac{C_z^2}{4} - f_uA\rho_{yz}C_yC_z$$

To minimize the MSE of T_{1u} , we differentiate $M(T_{1u})$ given in equation (3.9) with respect to A and equating it to zero, we get the optimum value of A as,

$$(3.10) \qquad A = 2\rho_{yz}\frac{C_y}{C_z}$$

Substituting the optimum value of A in equation (3.9), we have the minimum MSE of the estimator T_{1u} as

$$(3.11) \qquad Min.MSE(T_{1u}) = \bar{Y}^2 f_uC_y^2\big[1 - \rho_{yz}^2\big]$$

From equation (3.7) and (3.10) we have

$$(3.12) \qquad A = \sum_{i=1}^{3} ia_i = 2\rho_{yz}\frac{C_y}{C_z}$$

From condition $\sum_{i=1}^{3}$ and equation (3.12), we have three unknowns to be determined from two equations only. It is therefore not possible to find unique solutions of the constants $a_i^{'}s(i = 1, 2, 3)$. Thus in order to get the unique solutions of the constants $a_i^{'}s(i = 1, 2, 3)$, we shall impose a linear constraint as:

$$(3.13) \qquad B(T_{1u}) = 0$$

which follows from equation (3.8) that

$$(3.14) \qquad a_1\big(\frac{1}{2}\rho_{yz}C_yC_z - \frac{3}{8}C_z^2\big) + 2a_2\big(\frac{1}{2}\rho_{yz}C_yC_z - \frac{1}{2}C_z^2\big) + 3a_3\big(\frac{1}{2}\rho_{yz}C_yC_z - \frac{5}{8}C_z^2\big) = 0$$

Condition $\sum_{i=1}^{3} = 1$ and equations (3.10) and (3.12) can be written in matrix form as

$$(3.15) \qquad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ \frac{1}{2}\rho_{yz}C_yC_z - \frac{3}{8}C_z^2 & 2(\frac{1}{2}\rho_{yz}C_yC_z - \frac{1}{2}C_z^2) & 3(\frac{1}{2}\rho_{yz}C_yC_z - \frac{5}{8}C_z^2) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2\rho_{yz}\frac{C_y}{C_z} \\ 0 \end{bmatrix}$$

Solving equation (3.15) and assuming $C_y \cong C_z$ we get the unique values of $a_i^{'}s(i=1,2,3)$ as:

(3.16) $a_1 = 4\rho_{yz}^2 - 7\rho_{yz} + 3, a_2 = 12\rho_{yz} - 8\rho_{yz}^2 - 3, a_3 = 1 + 4\rho_{yz}^2 - 5\rho_{yz}$

Substituting the values of a_1, a_2, and a_3 from equation (3.16) into equation (2.1), we have an almost unbiased optimum exponential type estimator say T_{1u}^* and written as

$$\begin{aligned} T_{1u}^* =&\{4\rho_{yz}^2 - 7\rho_{yz} + 3\}\bar{y_u} exp\Big[\frac{\bar{Z} - \bar{z_u}}{\bar{Z} + \bar{z_u}}\Big] + \{12\rho_{yz} - 8\rho_{yz}^2 - 3\}\bar{y_u} exp2\Big[\frac{\bar{Z} - \bar{z_u}}{\bar{Z} + \bar{z_u}}\Big] \\ &+ \{1 + 4\rho_{yz}^2 - 5\rho_{yz}\}\bar{y_u} exp3\Big[\frac{\bar{Z} - \bar{z_u}}{\bar{Z} + \bar{z_u}}\Big] \end{aligned} \tag{3.17}$$

whose variance to the first degree of approximation ignoring f. p. c. is given by

(3.18) $V(T_{1u}^*) = \bar{Y}^2 f_u C_y^2 \big[1 - \rho_{yz}^2\big]$

Proceeding as above, we have almost unbiased version of the estimator T_{2u} say T_{2u}^* as

$$\begin{aligned} T_{2u}^* =&\{4\delta + 3\}\bar{y}_u^* exp\Big[\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\Big] - \{8\delta + 3\}\bar{y}_u^* exp2\Big[\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\Big] \\ &+ \{4\delta + 1\}\bar{y}_u^* exp3\Big[\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\Big] \end{aligned} \tag{3.19}$$

with variance

(3.20) $V(T_{2u}^*) = \bar{Y}^2 f_u C_y^2 \big[1 - \rho_{yz}^2\big]$

where

$\delta = \rho_{yz} \frac{\bar{Z}}{S_z^2}\big(\frac{\zeta_{012}}{S_{yz}} - \frac{\zeta_{003}}{S_z^2}\big)$ and $E\big[(x_i - \bar{X})^r (y_i - \bar{Y})^s (z_i - \bar{Z})^t\big]$; (r,s,t are integer ≥ 0),

Similarly almost unbiased version of the estimator T_{3u} say T_{3u}^* is derived as

$$T_{3u}^* = c_1^* \bar{y}_u^{**} exp\Big[\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\Big] + c_2^* \bar{y}_u^{**} exp2\Big[\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\Big] + c_3^* \bar{y}_u^{**} exp3\Big[\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\Big] \tag{3.21}$$

with variance

(3.22) $V(T_{3u}^*) = \bar{Y}^2 f_u C_y^2 \big[1 - \rho_{yz}^2\big]$

where

$c_1 = c_1^* = \frac{-(4\rho_{yz}^2 - 9\rho_{yz} + 25)}{8\rho_{yz} - 1}$, $c_2 = c_2^* = \frac{-(8\rho_{yz}^2 - 24\rho_{yz} - 45)}{8\rho_{yz} - 1}$, $c_3 = c_3^* = \frac{(12\rho_{yz}^2 - 25\rho_{yz} - 21)}{8\rho_{yz} - 1}$

Similarly, the unbiased version of the estimators of the set S_m with their respective variances are derived as

$$T_{1m}^* = \alpha_1^* \bar{y}_m^* exp\Big[\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\Big] + \alpha_2^* \bar{y}_m^* exp2\Big[\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\Big] + \alpha_3^* \bar{y}_m^* exp3\Big[\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\Big] \tag{3.23}$$

with variance

(3.24) $V(T_{1m}^*) = \bar{Y}^2 C_y^2 \big[f_m(1 - F_1^2\rho_{yz}^2) + f_2(1 - 2\rho_{yz})\big]$

and

$$T_{2m}^* = \beta_1^* \bar{y}_m^{**} exp\Big[\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\Big] + \beta_2^* \bar{y}_m^{**} exp2\Big[\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\Big] + \beta_3^* \bar{y}_m^{**} exp3\Big[\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\Big] \tag{3.25}$$

with variance

(3.26) $V(T_{2m}^*) = \bar{Y}^2 C_y^2 \big[f_m(1 - F_2^2\rho_{yz}^2) + f_2\rho yx(\rho_{yx} - 2) - 2\rho_{yz}^2 F_2 f_1\big]$

where

$\alpha_1 = \alpha_1^* = \frac{6RF_1\rho_{yz} + 8QF_1\rho_{yz} - 6Q - 3R + S}{P - 3R - 4Q}$ $\qquad \beta_1 = \beta_1^* = 3 - 7F_2\rho_{yz} + \frac{4}{f_m}(P^* - Q^* f_2)$

$\alpha_2 = \alpha_2^* = \frac{-2F_1\rho_{yz}(P+3R+4Q)+3P-2S}{P-3R-4Q}$ $\quad\beta_2 = \beta_2^* = 3(4F_2\rho_{yz} - 1) - \frac{8}{f_m}(P^* - Q^* f_2)$

$\alpha_3 = \alpha_3^* = \frac{2F_1 P\rho_{yz}+2Q-2P+S}{P-3R-4Q}$ $\quad\beta_3 = \beta_3^* = 1 - 5F_2\rho_{yz} + \frac{4}{f_m}(P^* - Q^* f_2)$

$P = \frac{3}{4}f_m - f_1\rho_{yz},\ Q = \frac{1}{2}f_m - f_1\rho_{yz},\ R = \frac{5}{4}f_m - f_1\rho_{yz},\ S = -2f_2(1-\rho_{yx}),$

$F_1 = \frac{f_1}{f_m},\ P^* = \frac{1}{2}(f_m\rho_{yz} - f_2\rho_{yx}\rho_{yz}),\ Q^* = -\rho_{yx}\frac{\bar{X}}{S_x^2}\left(\frac{\zeta_{300}}{S_x^2} - \frac{\zeta_{210}}{S_{yx}}\right),\ F_2 = 1 - \rho_{yx}\frac{f_2}{f_m},$

$f_m = \frac{1}{m},\ f_1 = \frac{1}{n},\ f_2 = \left(\frac{1}{m} - \frac{1}{n}\right).$

3.1. Theorem. *The almost unbiased versions of the estimator T_{ij}^* are given as* $T_{ij}^* = \varphi_{ij}T_{iu}^* + (1-\varphi_{ij})T_{jm}^*(i = 1.2.3; j = 1,2)$

Proof. Since the estimators $T_{iu}^*(i = 1,2,3)$ and $T_{jm}^*(j = 1,2)$ derived in equations (3.17), (3.19), (3.21), (3.23) and (3.25) are almost unbiased estimators of $\bar{Y}$. The final estimator $T_{ij}^*(i = 1,2,3; j = 1,2)$ are the convex linear combinations of the estimators T_{iu}^* and T_{jm}^*, therefore they are also unbiased estimators of $\bar{Y}$. □

3.2. Theorem. *Variance of the estimator $T_{ij}^*(i = 1,2,3; j = 1,2)$ to the first order of approximations are obtained as*

$$V(T_{ij}^*) = \varphi_{ij}^2 V(T_{iu}^*) + (1-\varphi_{ij})^2 V(T_{jm}^*)(i = 1.2.3; j = 1,2) \tag{3.27}$$

Proof. It is obvious that the variance of the proposed estimators $T_{ij}^*(i = 1,2,3; j = 1,2)$ are given by
$V(T_{ij}^*) = E(T_{ij}^* - \bar{Y})^2 = E[\varphi_{ij}T_{iu}^* + (1-\varphi_{ij})T_{jm}^* - \bar{Y}]^2 = E[\varphi_{ij}(T_{iu}^* - \bar{Y}) + (1-\varphi_{ij})(T_{jm}^* - \bar{Y})]^2 = \varphi_{ij}^2 V(T_{iu}^*) + (1-\varphi_{ij}^2)V(T_{jm}^*) + 2\varphi_{ij}(1-\varphi_{ij})E[(T_{iu}^* - \bar{Y})(T_{jm}^* - \bar{Y})]$ □

where $V(T_{iu}^*)$ and $V(T_{jm}^*)$ are obtained in equations (3.18), (3.20), (3.22), (3.24) and (3.26). It should be noted that the estimators T_{iu}^* and T_{jm}^* are based on two non-overlapping samples of sizes u and m respectively, their covariance types of terms are of order N^{-1} and ignored for large population size.

3.3. Remark. The above results are derived under the following assumptions:
(i) Population size is sufficiently large($i.e., N \to \infty$) , therefore, finite population corrections (f.p.c.) are ignored.
(ii) "$\rho_{xz} = \rho_{yz}$", this is an intuitive assumptions, which has been also considered by Cochran (1977) and Feng and Zou (1997).
(iii) since x and y denote the same study variable over two occasions and z is an auxiliary variable correlated to x and y, therefore, looking on the stability nature of the coefficient of variation (Reddy 1978) and following Cochran (1977) and Feng and Zou (1997), the coefficients of variation of variables x, y and z are considered to be approximately equal i.e., $C_y \cong C_x \cong C_z$.

3.4. Remark. It is to be noted that the suitable choices of the weights a_i, b_i, c_i, α_i and $\beta_i(i = 1,2,3)$ derived earlier depend on unknown population parameters such as$\beta_{yz}, \beta_{yx}, C_x^2, S_{yz}, S_{yx}, \rho_{yx}, \rho_{yz}, \zeta_{003}, \zeta_{012}$ and ζ_{210}. Thus to make such estimators practicable one has to use guessed or estimated values of these parameters. Guessed values of population parameters may be obtained either from past data or experience gathered over time; see for instance Reddy (1978). If such guessed values are not available it is advisable to use sample data to estimate these parameters as suggested by Singh *et al.* (2007).

3.2. Minimum variances of the proposed estimators $T^*_{ij}(i = 1, 2, 3; j = 1, 2)$.
Since the variances of the estimators $T^*_{ij}(i = 1, 2, 3; j = 1, 2)$ obtained in equation (3.27) are the functions of the unknown constants $\varphi_{ij}(i = 1, 2, 3; j = 1, 2)$, therefore, to get the optimum values of φ_{ij} , the variances of the estimators T^*_{ij} are differentiated with respect to φ_{ij} and equated to zero and subsequently the optimum values of φ_{ij} are obtained as

$$(3.28) \qquad \varphi_{ij_{opt}} = \frac{V(T^*_{jm})}{V(T^*_{iu}) + V(T^*_{jm})}; (i = 1, 2, 3; j = 1, 2)$$

Substituting the optimum values of $\varphi_{ij}(i = 1, 2, 3; j = 1, 2)$ in equation (3.27), we get the optimum variances of the estimators T^*_{ij} as

$$(3.29) \qquad V(T^*_{ij})_{opt} = \frac{V(T^*_{iu}).V(T^*_{jm})}{V(T^*_{iu}) + V(T^*_{jm})}; (i = 1, 2, 3; j = 1, 2)$$

Since $V(T^*_{1u}) = V(T^*_{2u}) = V(T^*_{3u})$, therefore, we have $V(T^*_{11})_{opt} = V(T^*_{21})_{opt} = V(T^*_{31})_{opt}$ and $V(T^*_{12})_{opt} = V(T^*_{22})_{opt} = V(T^*_{32})_{opt}$. Further, substituting the expressions of variances of the estimators $T^*_{iu}(i = 1, 2, 3)$ and $T^*_{jm}(j = 1, 2)$ from the equations (3.18), (3.20), (3.22), (3.24) and (3.26) in equations (3.28) and (3.29), the simplified values of $\varphi_{ij_{opt}}(i = 1, 2, 3; j = 1, 2)$ and $V(T^*_{ij})_{opt}$ are obtained as

$$(3.30) \qquad \varphi_{11_{opt}} = \frac{\mu_{11}[A_3 - \mu^2_{11}A_2 + \mu_{11}A_5]}{A_1 - \mu_{11}A_6 - \mu^3_{11}A_2 + \mu^2_{11}A_5}$$

$$(3.31) \qquad V(T^*_{11})_{opt} = \frac{S^2_y}{n}\left[\frac{A_7 - \mu^2_{11}A_8 + \mu_{11}A_9}{A_1 - \mu_{11}A_6 - \mu^3_{11}A_2 + \mu^2_{11}A_5}\right]$$

$$(3.32) \qquad \varphi_{12_{opt}} = \frac{A_1 - \mu^2_{12}A_{15} - \mu_{12}A_{16}}{A_1 - \mu^3_{12}A_{15} - \mu^2_{12}A_{16}}$$

and

$$(3.33) \qquad V(T^*_{12})_{opt} = \frac{S^2_y}{n}\left[\frac{A_{17} - \mu^2_{12}A_{18} - \mu_{12}A_{19}}{A_1 - \mu^3_{12}A_{15} - \mu^2_{12}A_{16}}\right]$$

where
$A_1 = 1-\rho^2_{yz}$, $A_2 = \rho^2_{yz}$, $A_3 = 1-A_2$, $A_4 = 1-2\rho_{yx}$, $A_5 = 2A_2+A_4$, $A_6 = A_1-A_3$, $A_7 = A_1A_3$, $A_8 = A_1A_2$, $A_9 = A_1A_5$, $A_{10} = A_2A_8$, $A_{11} = A_2A_9$, $A_{12} = A_6A_8-A_5A_9+3A_2A_7$, $A_{13} = A_1A_8+A_5A_7$, $A_{14} = A_1A_9+A_7A_6$, $A_{15} = \rho^2_{yz}\rho^2_{yx}$, $A_{16} = \rho^2_{yx}-2\rho^2_{yz}\rho_{yx}$, $A_{17} = A^2_1$, $A_{18} = A_1A_{15}$, $A_{19} = A_1A_{16}$, $A_{20} = A_{15}A_{18}$, $A_{21} = A_{15}A_{19}$, $A_{22} = A_{16}A_{19} - 3A_{15}A_{17}$, $A_{23} = A_1A_{18}-A_{16}A_{17}$, $A_{24} = A_1A_{19}$ and μ_{1j} is the fraction of fresh sample required for estimator $T^*_{1j}(j = 1, 2)$.

3.5. Remark. Since the $V(T^*_{1u}) = V(T^*_{2u}) = V(T^*_{3u})$, therefore, we have $V(T^*_{11})_{opt}$ is similar to that of the $V(T^*_{21})_{opt}$, and $V(T^*_{31})_{opt}$, and $V(T^*_{12})_{opt}$ is similar to that of the $V(T^*_{22})_{opt}$ and $V(T^*_{32})_{opt}$. Hence onwards only the properties of the estimators T^*_{11} and T^*_{12} will be examined.

3.3. Optimum replacement strategies.
To determine the optimum values of $\mu_{1j}(j = 1, 2)$ so that the population mean $\bar{Y}$ may be estimated with maximum precision, we minimize $V(T^*_{1j})_{opt}(j = 1, 2)$ given in equations (3.31) and (3.33) respectively with respect to μ_{1j} which result in 4th degree equations in μ_1 and μ_{12} . The respective equations in μ_{1j} are obtained as

$$(3.34) \qquad \mu^4_{11}A_{10} - 2\mu^3_{11}A_{11} - \mu^2_{11}A_{12} + 2\mu_{11}A_{13} - A_{14} = 0$$

and

$$(3.35) \qquad \mu^4_{12}A_{20} + 2\mu^3_{12}A_{21} + \mu^2_{12}A_{22} + 2\mu_{12}A_{23} + A_{24} = 0$$

Solving equations (3.34) and (3.35) we get the solutions of $\mu_{1j}(j=1,2)$ say $\hat{\mu}_{1j}(j=1,2)$. While choosing the values of $\hat{\mu}_{1j}$, it should be remembered that $0 \leq \hat{\mu_{1j}} \leq 1$ and if more than one such admissible values of $\hat{\mu}_{1j}$ are obtained, the lowest one will be the best choice, because we have the same variance by replacing only the lowest fraction of total sample size on current occasion, which reduces the cost of the survey. All others values of $\hat{\mu}_{1j}(j=1,2)$ are inadmissible. Substituting the admissible values of $\hat{\mu}_{1j}$ say $\mu_{1j}^{(o)}(j=1,2)$ into the equations (3.31) and (3.33) respectively, we have the optimum values of $V(T_{1j}^{*})_{opt}(j=1,2)$, which are shown below

$$(3.36) \qquad V(T_{11}^{*})_{opt} = \frac{S_y^2}{n}\left[\frac{A_7 - \mu_{11}^{(o)2}A_8 + \mu_{11}^{(o)}A_9}{A_1 - \mu_{11}^{(o)}A_6 - \mu_{11}^{(o)3}A_2 + \mu_{11}^{(o)2}A_5}\right]$$

and

$$(3.37) \qquad V(T_{12}^{*})_{opt} = \frac{S_y^2}{n}\left[\frac{A_{17} - \mu_{12}^{(o)2}A_{18} - \mu_{12}^{(o)}A_{19}}{A_1 - \mu_{12}^{(o)3}A_{15} - \mu_{12}^{(o)2}A_{16}}\right]$$

4. Simulation Study

The percent relative efficiencies of the estimators $T_{ij}^{*}(i=1;j=1,2)$ with respect to (i) sample mean estimator $\bar{y}_n$ when there is no matching from previous occasion and (ii) natural successive sampling estimator $\hat{\bar{Y}} = \varphi^{*}\bar{y}_u + (1-\varphi^{*})\bar{y}_m^{*}$, where $\bar{y}_m^{*} = \bar{y}_m + b_{yx}^{(m)}(\bar{x}_n - \bar{x}_m)$, when no auxiliary information is used on any occasion have been computed for different choices of correlations. Since $\bar{y}_n$ is an unbiased estimator and $\hat{\bar{Y}}$ is a biased estimator of $\bar{Y}$, hence, following Sukhatme et.al (1984), the variance of $\bar{y}_n$ and the optimum mean square error of the estimator $\hat{\bar{Y}}$ for large N are respectively given by

$$(4.1) \qquad V(\bar{y}_n) = \frac{S_y^2}{n}$$

and

$$(4.2) \qquad M(\hat{\bar{Y}})_{opt} = \left[1 + \sqrt{(1-\rho_{yx}^2)}\right]\frac{S_y^2}{n}$$

4.1. Simulation results.

For different choices of ρ_{yz} and ρ_{yx} , Tables 1-2 present the optimum values of $\mu_{1j}^{(o)}(j=1,2)$ and percent relatives efficiencies E_1 and E_2 of the proposed estimators $T_{1j}^{*}(j=1,2)$ with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$, where $E_{1j} = \frac{V(\bar{y}_n)}{V(T_{1j}^{*})_{opt}} \times 100$ and $E_{2j} = \frac{M(\hat{\bar{Y}})_{opt}}{V(T_{1j}^{*})_{opt}} \times 100(j=1,2)$.

Table 1: Optimum values of μ_{11} and PRE's of the estimators T_{11}^{*} with respect to $\bar{y}_n$ and $\hat{\hat{\bar{Y}}}$.

ρ_{yz} →		0.5	0.6	0.7	0.8	0.9
ρ_{yx} ↓						
0.2	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.3760 105.43 104.37	0.3509 116.12 114.95	0.3198 133.18 131.83	0.2787 163.87 162.22	0.2158 237.78 235.38
0.3	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.3841 109.38 106.86	0.3586 120.47 117.69	0.3275 138.16 134.98	0.2861 170.03 166.12	0.2224 246.82 241.14
0.4	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.3897 113.93 109.17	0.3650 125.46 120.23	0.3347 143.89 137.88	0.2938 177.09 169.70	0.2296 257.15 246.42
0.5	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.3864 119.25 111.26	0.3670 131.31 122.51	0.3401 150.57 140.49	0.3011 185.31 172.90	0.2372 269.12 251.09
0.6	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.3492 125.51 112.96	0.3559 138.24 124.41	0.3405 158.51 142.66	0.3072 195.06 175.55	0.2453 283.25 254.93
0.7	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.1758 131.90 113.05	0.3040 146.36 125.44	0.3274 168.09 144.06	0.3092 206.88 177.31	0.2532 300.34 257.41
0.8	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.6842 147.16 117.73	**0.1502** 154.22 123.38	0.2789 179.47 143.57	0.3007 221.56 177.25	0.2599 321.63 257.30
0.9	$\mu_{11}^{(o)}$ E_{11} E_{21}	0.7173 165.23 118.63	0.7503 183.41 131.68	0.1631 190.98 137.11	0.2669 239.80 172.16	0.2617 349.21 250.71

Table 2: Optimum values of μ_{12} and PRE's of the estimators T_{12}^{*} with respect to $\bar{y}_n$ and $\hat{\hat{\bar{Y}}}$.

$\rho_{yz} \rightarrow$		0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\rho_{yx} \downarrow$ 0.3	$\mu_{12}^{(o)}$	0.5313	0.1822	0.4383	0.4467	0.4362	0.4108	0.3548
	E_{12}	111.13	118.97	131.20	150.00	180.99	239.19	387.53
	E_{22}	108.57	116.23	128.17	146.55	176.82	233.68	378.60
0.4	$\mu_{12}^{(o)}$	0.5337	0.5761	0.3320	0.4183	0.4180	0.3937	0.3357
	E_{12}	112.93	120.71	132.29	150.46	180.05	234.96	372.51
	E_{22}	108.22	115.68	126.77	144.18	172.54	225.15	356.96
0.5	$\mu_{12}^{(o)}$	0.5442	0.5651	0.6752	0.3589	0.3944	0.3777	0.3211
	E_{12}	115.64	123.29	135.04	152.06	180.74	233.36	363.56
	E_{22}	107.89	115.03	125.99	141.88	168.64	217.73	339.20
0.6	$\mu_{12}^{(o)}$	0.5616	0.5745	0.6212	0.2286	0.3568	0.3599	0.3091
	E_{12}	119.53	127.12	138.69	154.45	183.02	234.16	359.34
	E_{22}	107.57	114.41	124.82	139.00	164.72	210.74	323.41
0.7	$\mu_{12}^{(o)}$	0.5877	0.5964	0.6224	0.0389	0.2922	0.3363	0.2981
	E_{12}	125.13	132.76	144.34	156.18	186.74	237.36	359.34
	E_{22}	107.25	113.78	123.71	133.86	160.05	203.43	307.98
0.8	$\mu_{12}^{(o)}$	0.6279	0.6338	0.6497	0.6935	0.1965	0.3015	0.2862
	E_{12}	133.61	141.37	153.15	171.58	191.23	243.02	363.52
	E_{22}	106.89	113.10	122.52	137.26	152.98	194.42	290.82
0.9	$\mu_{12}^{(o)}$	0.6982	0.7018	0.7112	0.7340	**0.0862**	0.2512	0.2711
	E_{12}	148.23	156.30	168.58	187.76	194.94	250.99	372.29
	E_{22}	106.42	112.22	121.03	134.80	139.95	180.20	267.28

4.2. Interpretation of simulation results.
(1) From Table 1, it is clear that
(a) For smaller values of ρ_{yx} , the values of E_{11} and E_{21} are increasing and $\mu_{11}^{(o)}$ are decreasing while for the higher values of ρ_{yx} the values of E_{11} and E_{21} are increasing and the values of $\mu_{11}^{(o)}$ do not follow any definite pattern when the values of ρ_{yz} are increasing. This phenomenon indicates that smaller fraction of fresh sample on the current occasion is required, if information on a highly positively correlated auxiliary variable is used at the estimation stage.
(b) For the fixed values of ρ_{yz} , the values of E_{11} and E_{21} are increasing while no definite trends are visible in the values of $\mu_{11}^{(o)}$ with the increasing values of ρ_{yx} .
(c) Minimum value of $\mu_{11}^{(o)}$ is obtained as 0.1502, which indicates that only about 15 percent of the total sample size is to be replaced on the current (second) occasion for the corresponding choices of correlations.
(2) From Table 2, it is observed that
(a) For the fixed values of ρ_{yx} , the values of E_{12} and E_{22} are increasing while no definite patterns are seen in the values of $\mu_{12}^{(o)}$ with the increasing values of ρ_{yz}.
(b) For the fixed values of ρ_{yz} , the values of E_{12} and $\mu_{12}^{(o)}$ are increasing while the values of E_{22} are decreasing with the increasing values of ρ_{yx}.
(c) The minimum value of $\mu_{12}^{(o)}$ is obtained as 0.0862, which indicates that about 8 percent of the total sample size is to be replaced on the current (second) occasion for the corresponding choices of correlations.

5. Conclusions

The work presented in this paper provides some unbiased and efficient estimation procedures of current population mean in two-occasion successive sampling. The empirical results presented in Tables 1-2 and their subsequent analyses and interpretations vindicate that the use of auxiliary information on an auxiliary variable at estimation stage is highly rewarding in terms of the proposed estimators. It is also visible that if a highly correlated auxiliary variable is used at the estimation stage, relatively, only a small fraction of the sample on the current (second) occasion is required to be replaced by a fresh sample, which reduces the cost of the survey. Hence looking on the encouraging behaviours of the proposed estimators one may recommend them for their practical applications.

Acknowledgements Authors are thankful to the reviewer for their valuable suggestions which helped in enhancing the quality of this paper. Authors are also thankful to the Indian Institute of Technology (Indian School of Mines), Dhanbad for providing financial and infrastructural support in completion of this work.

References

[1] Biradar, R. S. and Singh, H. P. Successive sampling using auxiliary information on both occasions. *Calcutta Statistical Association Bulletin*, **5**1, 243-251, 2001.
[2] Chaturvedi, D. K. and Tripathi, T. P. Estimation of population ratio on two occasions using multivariate auxiliary information. *Journal of Indian Statistical Association*, **2**1,113-120, 1983.
[3] Cochran, W. G. Sampling Techniques, *Wiley Eastern Limited, New Delhi*, **I**II Edition, 1977.
[4] Das, A. K. Estimation of population ratio on two occasions.*Journal Indian Society Agricultures Statistics*, *3*4, 1-9, 1982

[5] Feng, S. and Zou, G. Sample rotation method with auxiliary variable. *Communications in Statistics-Theory and Methods*, **2**6, 6, 1497-1509, 1997.

[6] Gupta, P. C. Sampling on two successive occasions. *Journal Statistical Research*,**1**3, 7-16, 1979.

[7] Jessen, R.J. 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, 1942.

[8] Patterson, H. D. Sampling on successive occasions with partial replacement of units.*Journal of the Royal Statistical Society*,**1**2, 241-255, 1950.

[9] Rao, J. N. K. and Graham, J. E. Rotation design for sampling on repeated occasions.*Journal of the American Statististical Association*, **5**9, 492-509, 1964.

[10] Reddy, V. N. (1978): A study on the use of prior knowledge on certain population parameters in the estimation. *Sankhya, Series C*, **4**0, 29-37, 1978.

[11] Sen, A. R. Successive sampling with two auxiliary variables. *Sankhya, 33, Series B*,**3**71-378, 1971.

[12] Sen, A. R. Successive sampling with p (p$\geq$ 1) auxiliary variables. *The Annals Mathematical Statististics*, **4**3, 2031-2034, 1972.

[13] Sen, A. R. Theory and application of sampling on repeated occasions with several auxiliary variables.*Biometrics* **2**9, 381-385, 1973.

[14] 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. *Journal of Scientific Research* **4**1 C,149-159, 1991.

[15] Singh, G. N. On the use of chain-type ratio estimator in successive sampling. *Statistics in Transition*, **7**(1), 21-26, 2005.

[16] Singh, G. N. and Singh, V. K. On the use of auxiliary information in successive sampling. *Journal Indian Society Agricultures Statistics*, **5**4(1), 1-12, 2001.

[17] Singh, G. N. and Homa, F. Effective rotation patterns in successive sampling over two occasions. *Journal of Statistical Theory and Practice*, **7**(1), 146-155, 2013.

[18] Singh, G. N. and Karna, J. P. Estimation of population mean on current occasions in two-occasion successive sampling. *Metron*, **6**7(1), 69-85, 2009.

[19] Singh, G. N. and Prasad, S. (2013): Best linear unbiased estimators of population mean on current occasion in two-occasion successive sampling. *Statistics in Transition-New Series*, **1**4(1), 57-74, 2013.

[20] Singh, G.N. and Priyanka, K. On the use of auxiliary variables to improve the precision of estimates at current occasion. *Journal Indian Society of Agricultures Statistics*, **6**2(3), 253-265, 2008.

[21] Singh, G.N. and Priyanka, K. Estimation of population mean at current occasion in presence of several varying auxiliary variates in two – occasion successive sampling. *Statistics in Transition- new series*, **1**1(1), 105–126, 2010.

[22] Singh,H.P., Chandra,P., Joarder, A. H. and Singh, S. Family of estimators of mean,ratio and product of a finite population using random non-response. *Test*,**1**6, 565-597, 2007.

[23] Singh, H. P. and S. K. Pal On the estimation of population mean in current occasion in two-occasion rotation patterns. *Journal of Statistics Application and Probability*, **4**(2), 305-313, 2015.

[24] Singh, H. P. and S. K. Pal An efficient effective rotation pattern in successive sampling over two occasions. *Communications in Statistics-Theory and Methods*, **4**5(17), 5017-5027, 2016.

[25] Singh, H. P. And Vishwakarma, G. K. A general procedure for estimating population mean in successive sampling. *Communications in Statistics-Theory and Methods*, **3**8(2), 293-308, 2009.

[26] Srivastava, M. K. and Srivastava, N. Chain type ratio estimator in successive sampling using multiple-auxiliary information.*Communications in Statistics-Theory and Methods*, **4**5(16), 4798-4811, 2016.

[27] Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S. and Asok, C. Sampling theory of surveys with applications. *Iowa State University Press, Ames, Iowa (USA) and Indian Society of Agricultural Statistics, New Delhi*, 1984.