STATISTICS IN TRANSITION-new series, August 2011
Vol. 12, No. 1, pp. 25—44

SOME ROTATION PATTERNS IN TWO-PHASE SAMPLING

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

ABSTRACT

A problem related to the estimation of population mean on the current occasion using two-phase successive (rotation) sampling on two occasions has been considered. Two-phase ratio, regression and chain-type estimators for estimating the population mean on current (second) occasion have been proposed. Properties of the proposed estimators have been studied and their respective optimum replacement policies are discussed. Estimators are compared with the sample mean estimator, when there is no matching and the natural optimum estimator, which is a linear combination of the means of the matched and unmatched portions of the sample on the current occasion. Results are demonstrated through empirical means of comparison and suitable recommendations are made.

Key words: Two-phase, successive sampling, auxiliary information, chain-type, bias, mean square error, optimum replacement policy.

Mathematics Subject Classification: 62D05

1. Introduction

In many social, demographic, industrial and agricultural surveys, the same population is sampled repeatedly and the same study variable is measured on each occasion, so that development over time can be followed. For example, labour force surveys are conducted monthly to estimate the number of people in employment, data on prices of goods are collected monthly to determine the consumer price index, political opinion surveys are conducted at regular intervals to know the voter's preferences, etc. In such studies, successive (rotation) sampling plays an important role to provide the reliable and the cost effective estimates of real life (practical) situations at different successive points of time (occasions). It also provides the effective (in terms of cost and precision) estimates of the patterns of change over a period of time.

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

The problem of successive (rotation) sampling with a partial replacement of sampling units was first considered by Jessen (1942) in the analysis of survey data related to agriculture farm. He pioneered using the entire information collected in previous investigations (occasions). The theory of successive (rotation) sampling was further extended by Patterson (1950), Rao and Graham (1964), Gupta (1979), Das (1982) and Chaturvedi and Tripathi (1983), among others. Sen (1971) applied this theory with success in designing the estimator for the population mean on the current occasion using information on two auxiliary variables available on previous occasion. Sen (1972, 1973) extended his work for several auxiliary variables. Singh *et al.* (1991) and Singh and Singh (2001) used the auxiliary information available only on the current occasion and proposed estimators for the current population mean in two-occasion successive (rotation) sampling. Singh (2003) generalized his work for h-occasions successive sampling. Feng and Zou (1997) and Biradar and Singh (2001) used the auxiliary information on both the occasions for estimating the current population mean in successive sampling.

In many situations, information on the auxiliary variable may be readily available on the first as well as on the second occasion; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation, number of beds in different hospitals may be known in hospital surveys, number of polluting industries and vehicles are known in environmental surveys, nature of employment status, educational status, food availability and medical aids of a locality are well known in advance for estimating the various demographic parameters in demographic surveys. Many other situations in life sciences could also be explored to show the benefits of the present study. Utilizing the auxiliary information on both the occasions, Singh (2005), Singh and Priyanka (2006, 2007, 2008), Singh and Karna (2009a,b) have proposed several estimators for estimating the population mean on current (second) occasion in two-occasion successive (rotation) sampling. It is worth to be mentioned that almost all the above recent works have assumed that the population means of the auxiliary variables are known, which may not often be the case. In such situations, it is more generously advisable to go for two-phase successive (rotation) sampling. Two-phase sampling is a well tested scheme to provide the estimates of unknown population parameters related to the auxiliary variables in first-phase sample. Motivated with this argument and utilizing the information on a stable auxiliary variable with unknown population mean, we have proposed some estimators under two-phase sampling scheme for estimating the current population mean in two-occasion successive (rotation) sampling. Behaviours of the proposed estimators are examined through empirical means of comparison and subsequently suitable recommendations are made.

2. Sample structures and notations on two occasions

Let U = (U_1, U_2, - - -, U_N) be the finite population of N units and the one 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 occasion) whose population mean is unknown on both the occasions is available and it is closely related (positively correlated) to x and y on the first and second occasions respectively. To furnish a good estimate of the population mean of the auxiliary variable z on the first occasion, a preliminary sample of size n' is drawn from the population by the method of simple random sampling without replacement (SRSWOR) and information on z is collected. Further, a second-phase sample of size n ($n' > n$) is drawn from the first-phase (preliminary) sample by the method of SRSWOR and henceforth the information on study character x is gathered. A random sub-sample of size m = nλ is retained (matched) from the second-phase sample selected on the first occasion for its use on the second occasion. Once again to furnish a fresh estimate of the population mean of the auxiliary variable z on the second occasion, a preliminary (first-phase) sample of size u' is drawn from the non-sampled units of the population by the method of SRSWOR and information on z is collected. A second-phase sample of size u = (n-m) = nμ ($u' > u$) is drawn from the first-phase (preliminary) sample by the method of SRSWOR and the information on study variable y is gathered. It is obvious that the sample size on the second occasion is also n. λ and μ (λ+μ =1) are the fractions of the matched and fresh samples, respectively, on the second (current) occasion. Hence onward, we consider the following notations for their further use:

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

$\bar{x}_n, \bar{x}_m, \bar{y}_u, \bar{y}_m, \bar{z}_u, \bar{z}_n, \bar{z}_m$: The sample means of the respective variables based on the sample sizes shown in suffices $\bar{z}'_n, \bar{z}'_u$: The sample means of the auxiliary variable z and based on the first-phase samples of sizes n' and u' respectively ρ_{yx}, ρ_{yz}, ρ_{xz}: The correlation coefficients between the variables shown in suffices.

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

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

3. Formulation of the estimators

To estimate the population mean $\bar{Y}$ on the current (second) occasion, two different sets of estimators are considered. One set of estimators $S_u = \{T_{1u}, T_{2u}\}$ based on sample of size u (= nμ) drawn afresh on the current (second) occasion and the second set of estimators $S_m = \{T_{1m}, T_{2m}\}$ based on the matched sample of size m (= nλ), which is common with both the occasions. Estimators of the sets S_u and S_m are defined as:

$$T_{1u} = \frac{\bar{y}_u}{\bar{z}_u}\bar{z}'_u \qquad , T_{2u} = \bar{y}_u + b_{yz}(u)\left(\bar{z}'_u - \bar{z}_u\right), \qquad T_{1m} = \frac{\bar{y}_m}{\bar{x}_m}\frac{\bar{x}_n}{\bar{z}_n}\bar{z}'_n,$$

$$T_{2m} = \bar{y}^*_m + b_{yx}(m)\left(\bar{x}^*_n - \bar{x}^*_m\right)$$

where $\bar{y}^*_m = \bar{y}_m + b_{yz}(m)\left(\bar{z}'_n - \bar{z}_m\right)$, $\bar{x}^*_n = \bar{x}_n + b_{xz}(n)\left(\bar{z}'_n - \bar{z}_n\right)$ and $\bar{x}^*_m = \bar{x}_m + b_{xz}(m)\left(\bar{z}'_n - \bar{z}_m\right)$.

$b_{yz}(u)$, $b_{yx}(m)$, $b_{yz}(m)$, $b_{xz}(n)$ and $b_{xz}(m)$ are the sample regression coefficients between the variables shown in suffices and based on the sample sizes shown in braces.

Combining the estimators of sets S_u and S_m, we have the following estimators of population mean $\bar{Y}$ on the current (second) occasion:

$$T_{ij} = \varphi_{ij}T_{iu} + \left(1-\varphi_{ij}\right)T_{jm} \quad (i, j = 1, 2) \tag{1}$$

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

Remark 3.1.

For estimating the population mean on each occasion the estimators T_{iu} (i = 1, 2) are suitable, which implies that more belief on T_{iu} could be shown by choosing $\varphi_{ij}(i, j = 1, 2)$ as 1 (or close to 1), while for estimating the change over the occasions, the estimators T_{jm} (j = 1, 2) could be more useful and hence φ_{ij} might be chosen as 0 (or close to 0). For asserting both the problems simultaneously, the suitable (optimum) choices of φ_{ij} are required.

4. Properties of the estimators t_{ij} (i, j = 1, 2)

Since $T_{iu} (i = 1, 2)$ and $T_{jm} (j = 1, 2)$ are ratio, simple linear regression, chain-type ratio and regression estimators, they are biased for the population mean $\bar{Y}$, therefore the resulting estimators T_{ij} (i, j = 1, 2) defined in equation (1) are also biased estimators of $\bar{Y}$. The bias B (.) and mean square errors M (.) are derived up to $o(n^{-1})$ under the large sample approximations and using the following transformations:

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

$\bar{z}_u = \bar{Z}(1+e_5)$, $\bar{z}_u^{'} = \bar{Z}(1+e_6)$,

$\bar{z}_n = \bar{Z}(1+e_7)$, $\bar{z}_n^{'} = \bar{Z}(1+e_8)$, $s_{yz}(u) = S_{yz}(1+e_9)$, $s_z^2(u) = S_z^2(1+e_{10})$,

$s_{yx}(m) = S_{yx}(1+e_{11})$, $s_x^2(m) = S_x^2(1+e_{12})$, $s_{yz}(m) = S_{yz}(1+e_{13})$,

$s_z^2(m) = S_z^2(1+e_{14})$, $s_{xz}(n) = S_{xz}(1+e_{15})$, $s_z^2(n) = S_z^2(1+e_{16})$,

$s_{xz}(m) = S_{xz}(1+e_{17})$ and $\bar{z}_m = \bar{Z}(1+e_{18})$ such that $E(e_k) = 0$ and $|e_k| < 1 \;\forall\; k = 1, 2, ---, 18.$

Under the above transformations T_{iu} (i = 1, 2) and T_{jm} (j = 1, 2) take the following forms:

$$T_{1u} = \bar{Y}(1+e_1)(1+e_6)(1+e_5)^{-1} \tag{2}$$

$$T_{2u} = \bar{Y}(1+e_1) + \bar{Z}\beta_{yz}(1+e_9)(e_6-e_5)(1+e_{10})^{-1} \tag{3}$$

$$T_{1m} = \bar{Y}(1+e_2)(1+e_4)(1+e_8)(1+e_3)^{-1}(1+e_7)^{-1} \tag{4}$$

and

$$T_{2m} = \bar{Y}(1+e_2) + \bar{Z}\beta_{yz}(1+e_{13})(e_8-e_{18})(1+e_{14})^{-1} + \beta_{yx}(1+e_{11})(1+e_{12})^{-1}(I_1 - I_2) \tag{5}$$

where

$$I_1 = \bar{X}(1+e_4) + \bar{Z}\beta_{xz}(1+e_{15})(1+e_{16})^{-1}(e_8 - e_7)$$

$$I_2 = \bar{X}(1+e_3) + \bar{Z}\beta_{xz}(1+e_{17})(1+e_{14})^{-1}(e_8 - e_{18})$$

Thus, we have the following theorems:

Theorem 4.1. Bias of the of estimators T_{ij} (i, j = 1, 2) to the first order of approximations are obtained as

$$B(T_{ij}) = \varphi_{ij} B(T_{i\,u}) + (1-\varphi_{ij}) B(T_{j\,m})\ ;\ (i, j = 1, 2) \tag{6}$$

where

$$B(T_{1u}) = \bar{Y}\left(\frac{1}{u}-\frac{1}{u'}\right)\left(C_z^2 - \rho_{yz} C_y C_z\right) \tag{7}$$

$$B\left(T_{2u}\right) = \beta_{yz}\left(\frac{1}{u}-\frac{1}{u'}\right)\left(\frac{\alpha_{003}}{S_z^2}-\frac{\alpha_{102}}{S_{yz}}\right) \tag{8}$$

$$B\left(T_{1m}\right) = \bar{Y}\left[\left(\frac{1}{m}-\frac{1}{n}\right)\left(C_x^2 - \rho_{yx} C_y C_x\right) + \left(\frac{1}{n}-\frac{1}{n'}\right)\left(C_z^2 - \rho_{yz} C_y C_z\right)\right] \tag{9}$$

$$B\left(T_{2m}\right) = \beta_{yz}\left(\frac{1}{m}-\frac{1}{n'}\right)\left(\frac{\alpha_{003}}{S_z^2}-\frac{\alpha_{102}}{S_{yz}}\right) + \beta_{yx}\left(\frac{1}{m}-\frac{1}{n}\right)\left\{\left(\frac{\alpha_{030}}{S_x^2}-\frac{\alpha_{120}}{S_{yx}}\right) + \beta_{xz}\left(\frac{\alpha_{012}}{S_{xz}}-\frac{\alpha_{003}}{S_z^2}-\frac{\alpha_{021}}{S_x^2}+\frac{\alpha_{111}}{S_{yx}}\right)\right\} \tag{10}$$

where

$$\alpha_{rst} = E\left[\left(y-\bar{Y}\right)^r \left(x-\bar{X}\right)^s \left(z-\bar{Z}\right)^t\right];\ (r, s, t \geq 0) \text{ are integers.}$$

Proof: The bias of the estimators T_{ij} (i, j = 1, 2) are given by

$$B(T_{ij}) = E\left[T_{ij}-\bar{Y}\right] = \varphi_{ij} E(T_{iu}-\bar{Y}) + (1-\varphi_{ij})\, E(T_{jm}-\bar{Y})$$

$$= \varphi_{ij} B(T_{iu}) + \left(1-\varphi_{ij}\right) B(T_{jm}) \tag{11}$$

where

$$B(T_{i\,u}) = E\left[T_{i\,u}-\bar{Y}\right] \text{ and } B(T_{j\,m}) = E\left[T_{j\,m}-\bar{Y}\right].$$

Substituting the values of T_{1u}, T_{2u}, T_{1m} and T_{2m} from equations (2) – (5) in the equation (11), expanding the terms binomially and taking expectations up to o (n^{-1}), we have the expressions for the bias of the estimators T_{ij} (i, j = 1, 2) as described in equation (6).

Theorem 4.2. Mean square errors of the estimators T_{ij} (i, j = 1, 2) to the first order of approximations are obtained as

$$M(T_{ij}) = \varphi_{ij}^2\, M\left(T_{iu}\right) + \left(1-\varphi_{ij}\right)^2 M\left(T_{jm}\right) + 2\varphi_{ij}\left(1-\varphi_{ij}\right)C_{ij}\ ;\ (i, j = 1, 2) \tag{12}$$

where

$$M(T_{1u}) = \left[2\left(\frac{1}{u}-\frac{1}{N}\right)\left(1-\rho_{yz}\right)-\left(\frac{1}{u'}-\frac{1}{N}\right)\left(1-2\rho_{yz}\right)\right] S_y^2 \tag{13}$$

$$M(T_{2u}) = \left[\left(\frac{1}{u}-\frac{1}{N}\right)-\left(\frac{1}{u}-\frac{1}{u'}\right)\rho_{yz}^2\right]S_y^2 \tag{14}$$

$$M(T_{1m}) = \left[\left(\frac{1}{m}-\frac{1}{N}\right)2\left(1-\rho_{yx}\right)+\left(\frac{1}{n}-\frac{1}{N}\right)2\left(\rho_{yx}-\rho_{yz}\right)-\left(\frac{1}{n'}-\frac{1}{N}\right)\left(1-2\rho_{yz}\right)\right]S_y^2$$

$$M(T_{2m}) = \left[\left(\frac{1}{m}-\frac{1}{N}\right)-\left(\frac{1}{m}-\frac{1}{n'}\right)\rho_{yz}^2+\left(\frac{1}{m}-\frac{1}{n}\right)\rho_{yx}\left\{\rho_{yz}^2\left(2-\rho_{yx}\right)-\rho_{yx}\right\}\right]S_y^2 \tag{16}$$

$$C_{11} = -\frac{S_y^2}{N} \tag{17}$$

$$C_{12} = -\frac{S_y^2}{N} \tag{18}$$

$$C_{21} = -\frac{S_y^2}{N} \tag{19}$$

and

$$C_{22} = -\frac{S_y^2}{N} \tag{20}$$

Remark 4.1.

The above results are derived under the assumptions that the coefficients of variations of the variables x, y and z are approximately equal and $\rho_{xz} = \rho_{yz}$, which are intuitive assumptions considered by Cochran (1977) and Feng and Zou (1997).

Proof: It is obvious that the mean square errors of the estimators T_{ij} (i, j = 1, 2) are given by

$$M(T_{ij}) = E\left[T_{ij}-\bar{Y}\right]^2 = E\left[\varphi_{ij}\left(T_{iu}-\bar{Y}\right)+\left(1-\varphi_{ij}\right)\left(T_{jm}-\bar{Y}\right)\right]^2$$

$$= \varphi_{ij}^2 M(T_{iu})+\left(1-\varphi_{ij}\right)^2 M\left(T_{jm}\right)+2\varphi_{ij}\left(1-\varphi_{ij}\right)C_{ij} \qquad (21)$$

where

$$M\left(T_{iu}\right)=E\left[T_{iu}-\bar{Y}\right]^2,\ M(T_{jm}) = E\left[T_{jm}-\bar{Y}\right]^2 \text{ and } C_{ij} = E\left[\left(T_{iu}-\bar{Y}\right)\left(T_{jm}-\bar{Y}\right)\right]$$

Substituting the expressions of T_{iu} (i = 1, 2) and T_{jm} (j = 1, 2) given in equations (2)-(5) in equation (21), expanding the terms binomially and taking expectations up to $o(n^{-1})$, we have the expressions of the mean square errors of T_{ij} as given in equation (12).

5. Minimum mean square errors of the estimators t_{ij} (i, j = 1, 2)

Since the mean square errors of the estimators T_{ij} (i, j = 1, 2) in equation (12) are the functions of the unknown constants φ_{ij} (i, j = 1, 2), therefore they are minimized with respect to φ_{ij} and subsequently the optimum values of φ_{ij} are obtained as

$$\varphi_{ij_{opt}} = \frac{M(T_{jm})-C_{ij}}{M(T_{iu})+M(T_{jm})-2C_{ij}};\ (i, j = 1, 2) \qquad (22)$$

Now, substituting the values of $\varphi_{ij\ opt}$ in equation (12), we get the optimum mean square errors of T_{ij} as

$$M(T_{ij})_{opt} = \frac{M(T_{iu}).M(T_{jm})-C_{ij}^2}{M(T_{iu})+M(T_{jm})-2C_{ij}};\ (i, j = 1, 2) \qquad (23)$$

Further, substituting the values from equations (13)-(20) in equations (22) and (23), the simplified values of $\varphi_{ij_{opt}}$ and M $(T_{ij})_{opt}$ are given as:

$$\varphi_{11_{opt}} = \frac{\mu_{11}\left[A_7+\mu_{11}A_8\right]}{\left[A_1+\mu_{11}A_{10}+\mu_{11}^2A_{11}\right]} \tag{24}$$

$$M(T_{11})_{opt} = \frac{1}{n}\left[\frac{A_{13}+\mu_{11}A_{14}+\mu_{11}^2A_{15}}{A_1+\mu_{11}A_{10}+\mu_{11}^2A_{11}}\right]S_y^2 \tag{25}$$

$$\varphi_{12_{opt}} = \frac{\mu_{12}\left[A_{23}+\mu_{12}A_{24}\right]}{\left[A_1+\mu_{12}A_{25}+\mu_{12}^2A_{26}\right]} \tag{26}$$

$$M(T_{12})_{opt} = \frac{1}{n}\left[\frac{A_{30}+\mu_{12}A_{31}+\mu_{12}^2A_{32}}{A_1+\mu_{12}A_{25}+\mu_{12}^2A_{26}}\right]S_y^2 \tag{27}$$

$$\varphi_{21_{opt}} = \frac{\mu_{21}\left[A_7+\mu_{21}A_8\right]}{\left[A_{18}+\mu_{21}A_{16}+\mu_{21}^2A_{17}\right]} \tag{28}$$

$$M(T_{21})_{opt} = \frac{1}{n}\left[\frac{A_{19}+\mu_{21}A_{20}+\mu_{21}^2A_{21}}{A_{18}+\mu_{21}A_{16}+\mu_{21}^2A_{17}}\right]S_y^2 \tag{29}$$

$$\varphi_{22_{opt}} = \frac{\mu_{22}\left[A_{23}+\mu_{22}A_{24}\right]}{\left[A_{18}+\mu_{22}A_{33}+\mu_{22}^2A_{34}\right]} \tag{30}$$

$$M(T_{22})_{opt} = \frac{1}{n}\left[\frac{A_{38}+\mu_{22}A_{39}+\mu_{22}^2A_{40}}{A_{18}+\mu_{22}A_{33}+\mu_{22}^2A_{34}}\right]S_y^2 \tag{31}$$

where

$$A_1 = 2\left(1-\rho_{yz}\right),\ A_2 = \left(1-2\rho_{yz}\right),\ A_3 = 2\left(1-\rho_{yx}\right),\ A_4 = 2\left(\rho_{yx}-\rho_{yz}\right),\ A_5 = \rho_{yz}^2$$

$$A_6 = A_3+A_4-A_2-1,\ A_7 = A_3+A_4-f'A_2-f\,A_6,\ A_8 = f'A_2-A_4+f\,A_6,$$

$$A_9 = A_1-2A_2+A_3+A_4-2,\ A_{10} = A_3-A_1-A_2\left(t+f'\right)+A_4-fA_9,$$

$$A_{11} = A_2\left(t+f'\right)-A_4+f\,A_9,\ A_{12} = \left(A_1-A_2\right)\left(A_3+A_4-A_2\right)-1,$$

$$A_{13} = A_1\left(A_3+A_4-f'A_2\right)-f\,A_1\left(A_3+A_4-A_2\right),$$

$$A_{14} = A_1\left(f'A_2-A_4\right)-\left(tA_2+f\left(A_1-A_2\right)\right)\left(A_3+A_4-f'A_2\right)+f\left(A_3+A_4-A_2\right)\left(A_1+tA_2\right)+f^2A_{12}$$

$$A_{15} = \left(tA_2 + f\left(A_1 - A_2\right)\right)\left(A_4 - f' A_2\right) - f\left(tA_2\left(A_3 + A_4 - A_2\right) - f A_{12}\right),$$

$$A_{16} = A_7 - 1 + A_5\left(t+1\right),$$

$$A_{17} = A_8 - tA_5,\ A_{18} = 1 - A_5,\ A_{19} = A_{18}\left(A_7 - f\right),$$

$$A_{20} = A_{18}\left(A_8 + f\right) + tA_5\left(A_7 - f\right) - fA_7,\ A_{21} = tA_5\left(A_8 - f\right) - fA_8,$$

$$A_{22} = \rho_{yx}\left\{\rho_{yz}^2\left(2 - \rho_{yx}\right) - \rho_{yx}\right\},\ A_{23} = A_{18} + A_5 f',\ A_{24} = A_{22} - f' A_5,$$

$$A_{25} = A_{23} - A_1 - tA_2 - \left(A_1 - A_2 - 1\right)f,\ A_{26} = tA_2 + A_{24} + (A_1 - A_2 - 1)f,\ A_{27} = A_1 A_{23},$$

$$A_{28} = A_1 A_{24} - tA_2 A_{23},\ A_{29} = -tA_2 A_{24},\ A_{30} = A_{27} - fA_1,$$

$$A_{31} = A_{28} - f\left\{\left(A_1 - A_2\right)A_{23} - \left(tA_2 + A_1\right)\right\} + f^2\left(A_1 - A_2 - 1\right),$$

$$A_{32} = A_{29} - f\left\{\left(A_1 - A_2\right)A_{24} + tA_2\right\} - f^2\left(A_1 - A_2 - 1\right),\ A_{33} = tA_5 - A_{18} + A_{23},$$

$$A_{34} = A_{24} - tA_5,\ A_{35} = A_{18}\left(A_{18} + f' A_5\right),\ A_{36} = A_{18}\left(A_{22} - f' A_5 + tA_5\right) + f' t\, A_5^2,$$

$$A_{37} = tA_5\left(A_{22} - f' A_5\right),\ A_{38} = A_{35} - f A_{18},\ A_{39} = A_{36} - f A_5\left(f' + t\right),$$

$$A_{40} = A_{37} - f\left\{A_{22} - A_5\left(t + f'\right)\right\},\ f = \frac{n}{N},\ f' = \frac{n}{n'}\ \text{and}\ t = \frac{n}{u'}.$$

6. Optimum replacement policy

To determine the optimum values of μ_{ij} (i, j =1, 2) (fraction of samples to be drawn afresh on the current (second) occasion) so that population mean $\bar{Y}$ may be estimated with the maximum precision, we minimize mean square errors of T_{ij} (i, j =1, 2) given in equations (25), (27), (29) and (31) respectively with respect to μ_{ij}, which result in quadratic equations in μ_{ij} and respective solutions of μ_{ij} , say $\hat{\mu}_{ij}$ (i, j =1, 2), are given below:

$$Q_1\mu_{11}^2 + 2Q_2\mu_{11} + Q_3 = 0 \tag{32}$$

$$\hat{\mu}_{11} = \frac{-Q_2 \pm \sqrt{Q_2^2 - Q_1 Q_3}}{Q_1} \tag{33}$$

$$Q_7\mu_{12}^2+2Q_8\mu_{12}+Q_9=0 \tag{34}$$

$$\hat{\mu}_{12}=\frac{-Q_8\pm\sqrt{Q_8^2-Q_7Q_9}}{Q_7} \tag{35}$$

$$Q_4\mu_{21}^2+2Q_5\mu_{21}+Q_6=0 \tag{36}$$

$$\hat{\mu}_{21}=\frac{-Q_5\pm\sqrt{Q_5^2-Q_4Q_6}}{Q_4} \tag{37}$$

$$Q_{10}\mu_{22}^2+2Q_{11}\mu_{22}+Q_{12}=0 \tag{38}$$

$$\hat{\mu}_{22}=\frac{-Q_{11}\pm\sqrt{Q_{11}^2-Q_{10}Q_{12}}}{Q_{10}} \tag{39}$$

where

$Q_1 = A_{10}A_{15}-A_{14}A_{11}$, $Q_2 = A_1A_{15}-A_{13}A_{11}$, $Q_3 = A_1A_{14}-A_{13}A_{10}$,

$Q_4 = A_{21}A_{16}-A_{20}A_{17}$, $Q_5 = A_{21}A_{18}-A_{19}A_{17}$, $Q_6 = A_{18}A_{20}-A_{19}A_{16}$,

$Q_7 = A_{25}A_{32}-A_{31}A_{26}$, $Q_8 = A_1A_{32}-A_{30}A_{26}$, $Q_9 = A_1A_{31}-A_{30}A_{25}$,

$Q_{10} = A_{33}A_{40}-A_{39}A_{34}$, $Q_{11} = A_{18}A_{40}-A_{38}A_{34}$, $Q_{12} = A_{18}A_{39}-A_{38}A_{33}$

From equations (33), (35), (37) and (39), it is obvious that the real values of $\hat{\mu}_{ij}$ (i, j =1, 2) exist if the quantities under square roots are greater than or equal to zero. For any combinations of correlations ρ_{yx} and ρ_{yz} , which satisfy the conditions of real solutions, two real values of $\hat{\mu}_{ij}$ are possible. Hence, while choosing the values of $\hat{\mu}_{ij}$, it should be remembered that $0 \le \hat{\mu}_{ij} \le 1$. All the other values of μ_{ij} (i, j =1, 2) are inadmissible. Substituting the admissible values of $\hat{\mu}_{ij}$, say $\mu_{ij}^{(0)}$ (i, j =1, 2), from equations (33), (35), (37) and (39) into equations (25), (27), (29) and (31) respectively, we have the following optimum values of mean square errors of T_{ij} (i, j =1, 2).

$$M(T_{11}^0)_{opt} = \frac{1}{n}\left[\frac{A_{13}+\mu_{11}^{(0)}A_{14}+\mu_{11}^{(0)^2}A_{15}}{A_1+\mu_{11}^{(0)}A_{10}+\mu_{11}^{(0)^2}A_{11}}\right]S_y^2 \tag{40}$$

$$M(T_{12}^0)_{opt} = \frac{1}{n}\left[\frac{A_{30}+\mu_{12}^{(0)}A_{31}+\mu_{12}^{(0)^2}A_{32}}{A_1+\mu_{12}^{(0)}A_{25}+\mu_{12}^{(0)^2}A_{26}}\right]S_y^2 \tag{41}$$

$$M(T_{21}^{0})_{opt} = \frac{1}{n}\left[\frac{A_{19}+\mu_{21}^{(0)}A_{20}+\mu_{21}^{(0)^2}A_{21}}{A_{18}+\mu_{21}^{(0)}A_{16}+\mu_{21}^{(0)^2}A_{17}}\right]S_y^2 \tag{42}$$

$$M(T_{22}^{0})_{opt} = \frac{1}{n}\left[\frac{A_{38}+\mu_{22}^{(0)}A_{39}+\mu_{22}^{(0)^2}A_{40}}{A_{18}+\mu_{22}^{(0)}A_{33}+\mu_{22}^{(0)2}A_{34}}\right]S_y^2 \tag{43}$$

7. Efficiency comparison

The percent relative efficiencies of the estimators $T_{ij}(i, j = 1, 2)$ with respect to (i) $\bar{y}_n$, when there is no matching and (ii) $\hat{\bar{Y}}=\varphi^*\bar{y}_u+(1-\varphi^*)\bar{y}'_m$, when no auxiliary information is used on 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 ρ_{yx} and ρ_{yz}. Since $\bar{y}_n$ and $\hat{\bar{Y}}$ are unbiased estimators of $\bar{Y}$, therefore, following Sukhatme *et al.* (1984), the variance of $\bar{y}_n$ and the optimum variance of $\hat{\bar{Y}}$ are given by

$$V(\bar{y}_n) = \left(\frac{1}{n}-\frac{1}{N}\right)S_y^2 \tag{44}$$

$$V\left(\hat{\bar{Y}}\right)_{opt^*} = \left[1+\sqrt{1-\rho_{yx}^2}\right]\frac{S_y^2}{2n}-\frac{S_y^2}{N} \tag{45}$$

For N = 5000, n = 100, u' = 500 and n' = 500, the percent relative efficiencies $E_{ij}^{(l)}\left(i, j, l = 1, 2\right)$ of the estimators $T_{ij}(i, j = 1, 2)$ are computed with respect to the estimators $\bar{y}_n$ and $\hat{\bar{Y}}$ and shown in the tables 1- 4; where

$$E_{ij}^{(1)} = \frac{V(\bar{y}_n)}{M\left(T_{ij}^{0}\right)_{opt}}\times 100 \text{ and } E_{ij}^{(2)} = \frac{V\left(\hat{\bar{Y}}\right)_{opt^*}}{M\left(T_{ij}^{0}\right)_{opt}}\times 100$$

Table 1. Optimum values of μ_{11} and percent relative efficiencies of T_{11} with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$

ρ_{yz}	ρ_{yx}	0.4	0.5	0.6	0.7	0.8	0.9
0.4	$\mu_{11}^{(o)}$	*	0.5228	0.5505	0.5858	0.6340	0.7101
	$E_{11}^{(1)}$	-	**	**	**	108.12	121.75
	$E_{11}^{(2)}$	-	**	**	**	**	**
0.5	$\mu_{11}^{(o)}$	0.4772	*	0.5279	0.5635	0.6126	0.6910
	$E_{11}^{(1)}$	**	-	105.69	112.99	123.08	139.28
	$E_{11}^{(2)}$	**	-	**	**	**	**
0.6	$\mu_{11}^{(o)}$	0.4495	0.4721	*	0.5359	0.5858	0.6667
	$E_{11}^{(1)}$	110.12	115.67	-	131.29	143.51	163.33
	$E_{11}^{(2)}$	105.43	107.76	-	112.14	114.22	116.32
0.7	$\mu_{11}^{(o)}$	0.4142	0.4365	0.4641	*	0.5505	0.6340
	$E_{11}^{(1)}$	131.67	138.55	147.05	-	173.46	198.70
	$E_{11}^{(2)}$	126.06	129.08	132.05	-	138.06	141.51
0.8	$\mu_{11}^{(o)}$	0.3660	0.3874	0.4142	0.4495	*	0.5858
	$E_{11}^{(1)}$	167.11	176.18	187.44	202.08	-	256.93
	$E_{11}^{(2)}$	160.00	164.14	168.31	172.61	-	182.98
0.9	$\mu_{11}^{(o)}$	0.2899	0.3090	0.3333	0.3660	0.4142	*
	$E_{11}^{(1)}$	242.00	255.47	272.22	294.11	325.12	-
	$E_{11}^{(2)}$	231.70	238.00	244.44	251.21	258.77	-

Note: "" indicate* $\mu_{11}^{(0)}$ *do not exist and "**" indicate no gain.*

Table 2. Optimum values of μ_{12} and percent relative efficiencies of T_{12} with respect to $\bar{y}_n$ and $\hat{\hat{\bar{Y}}}$

ρ_{yz}	ρ_{yx}	0.6	0.7	0.8	0.9
0.4	$\mu_{12}^{(o)}$	0.0583	0.2875	0.4441	0.5938
	$E_{12}^{(1)}$	115.10	117.89	124.79	138.70
	$E_{12}^{(2)}$	103.35	100.70	**	**
0.5	$\mu_{12}^{(o)}$	0.1270	0.3257	0.4631	0.5992
	$E_{12}^{(1)}$	126.05	129.91	138.10	154.13
	$E_{12}^{(2)}$	113.19	110.96	109.92	109.77
0.6	$\mu_{12}^{(o)}$	0.1797	0.3538	0.4741	0.5980
	$E_{12}^{(1)}$	142.57	147.57	157.37	176.31
	$E_{12}^{(2)}$	128.02	126.05	125.25	125.57
0.7	$\mu_{12}^{(o)}$	0.2223	0.3729	0.4760	0.5879
	$E_{12}^{(1)}$	168.37	174.77	186.81	210.05
	$E_{12}^{(2)}$	151.19	149.28	148.68	149.59
0.8	$\mu_{12}^{(o)}$	0.2712	0.3836	0.4643	0.5626
	$E_{12}^{(1)}$	212.72	221.18	236.74	267.06
	$E_{12}^{(2)}$	191.02	188.93	188.43	190.20
0.9	$\mu_{12}^{(o)}$	0.3344	0.3720	0.4190	0.4998
	$E_{12}^{(1)}$	308.63	320.57	342.77	386.86
	$E_{12}^{(2)}$	277.13	273.81	272.81	275.52

*Note: "**" indicate no gain.*

Table 3. Optimum values of μ_{21} and percent relative efficiencies of T_{21} with respect to $\bar{y}_n$ and $\hat{\bar{Y}}$

ρ_{yz}	ρ_{yx}	0.4	0.5	0.6	0.7	0.8	0.9
0.4	$\mu_{21}^{(o)}$	*	*	0.8737	0.7693	0.7444	0.7717
	$E_{21}^{(1)}$	-	-	116.12	121.27	130.14	145.31
	$E_{21}^{(2)}$	-	-	104.27	103.58	103.58	103.49
0.5	$\mu_{21}^{(o)}$	*	*	0.9061	0.7457	0.7150	0.7471
	$E_{21}^{(1)}$	-	-	126.37	132.06	142.42	160.21
	$E_{21}^{(2)}$	-	-	113.48	112.80	113.36	114.10
0.6	$\mu_{21}^{(o)}$	0.1373	*	*	0.7230	0.6805	0.7180
	$E_{21}^{(1)}$	118.57	-	-	147.91	160.34	181.85
	$E_{21}^{(2)}$	113.52	-	-	126.34	127.62	129.51
0.7	$\mu_{21}^{(o)}$	0.2730	0.1813	*	0.7091	0.6383	0.6825
	$E_{21}^{(1)}$	140.53	146.69	-	172.42	187.91	215.13
	$E_{21}^{(2)}$	134.54	136.66	-	147.28	149.56	153.21
0.8	$\mu_{21}^{(o)}$	0.3022	0.2906	0.1686	0.7479	0.5831	0.6363
	$E_{21}^{(1)}$	174.83	184.07	194.49	214.67	235.01	272.04
	$E_{21}^{(2)}$	167.39	171.49	174.65	183.36	187.05	193.74
0.9	$\mu_{21}^{(o)}$	0.2631	0.2678	0.2402	0.7922	0.4931	0.5662
	$E_{21}^{(1)}$	247.85	261.68	278.64	305.50	336.31	394.07
	$E_{21}^{(2)}$	237.29	243.79	250.21	260.95	267.68	280.65

Note: "" indicate* $\mu_{21}^{(0)}$ *do not exist.*

Table 4. Optimum values of μ_{22} and percent relative efficiencies of T_{22} with respect to $\bar{y}_n$ and $\hat{\hat{Y}}$

ρ_{yz}	ρ_{yx}	0.4	0.5	0.6	0.7	0.8	0.9
0.4	$\mu_{22}^{(o)}$	0.5089	0.5210	0.5390	0.5656	0.6065	0.6786
	$E_{22}^{(1)}$	119.31	122.17	126.41	132.68	142.33	159.37
	$E_{22}^{(2)}$	114.23	113.82	113.51	113.33	113.29	113.50
0.5	$\mu_{22}^{(o)}$	*	0.5109	0.5279	0.5536	0.5941	0.6667
	$E_{22}^{(1)}$	-	132.60	136.98	143.62	154.04	172.68
	$E_{22}^{(2)}$	-	123.54	123.00	122.67	122.60	122.98
0.6	$\mu_{22}^{(o)}$	0.4870	0.4962	0.5118	0.5365	0.5764	0.6495
	$E_{22}^{(1)}$	145.58	148.27	152.82	160.01	171.58	192.64
	$E_{22}^{(2)}$	139.38	138.13	137.23	136.67	136.56	137.20
0.7	$\mu_{22}^{(o)}$	0.4671	0.4741	0.4880	0.5113	0.5504	0.6240
	$E_{22}^{(1)}$	170.47	172.88	177.66	185.70	199.05	223.93
	$E_{22}^{(2)}$	163.21	161.06	159.53	158.61	158.43	159.48
0.8	$\mu_{22}^{(o)}$	0.4345	0.4386	0.4500	0.4715	0.5092	0.5834
	$E_{22}^{(1)}$	213.86	215.65	220.73	230.16	246.55	277.97
	$E_{22}^{(2)}$	204.75	200.91	198.21	196.59	196.23	197.97
0.9	$\mu_{22}^{(o)}$	0.3705	0.3707	0.3789	0.3973	0.4323	0.5054
	$E_{22}^{(1)}$	308.76	308.89	314.40	326.56	349.07	393.64
	$E_{22}^{(1)}$	295.61	287.77	282.31	278.93	277.83	280.35

Note: "" indicates* $\mu_{22}^{(o)}$ *does not exist.*

8. Conclusion

From Table 1, it is clear that

(a) For fixed values of $\rho_{yx,}$ the values of $\mu_{11}^{(0)}$ are decreasing while the values of $E_{11}^{(1)}$ and $E_{11}^{(2)}$ are increasing with the increasing values of ρ_{yz}. This behaviour is highly desirable, since it concludes that if highly correlated auxiliary character is available, it pays in terms of enhanced precision of estimates as well as it reduces the cost of survey.

(b) For fixed values of $\rho_{yz,}$ the values of $\mu_{11}^{(0)}$, $E_{11}^{(1)}$ and $E_{11}^{(2)}$ are increasing with the increasing values of $\rho_{yz.}$

(c) The minimum value of $\mu_{11}^{(0)}$ is 0.2899, which indicates that only about 29 percent of the total sample size is to be replaced on the current (second) occasion for the corresponding choices of correlations.

From Table 2, it is visible that

(a) For fixed values of ρ_{yx}, the values of $\mu_{12}^{(0)}$ are increasing for some choices of ρ_{yz} while decreasing pattern may also be seen for few choices of ρ_{yz} and the values of $E_{12}^{(1)}$ and $E_{12}^{(2)}$ are increasing with increase in the values of ρ_{yz}.

(b) For fixed values of ρ_{yz}, the values of $\mu_{12}^{(0)}$ and $E_{12}^{(1)}$ are increasing with the increasing values of ρ_{yx} while the values of $E_{12}^{(2)}$ are decreasing for some choices of $\rho_{yx.}$ This behaviour is in agreement with Sukhatme et al. (1984) results, which explains that the more value of ρ_{yx}, the more fraction of fresh sample is required on the current (second) occasion.

(c) The minimum value of $\mu_{12}^{(0)}$ is 0.0503, which indicates that the fraction of fresh sample to be replaced on current occasion is as low as about 5 percent of the total sample size, which leads to an appreciable reduction in the cost of the survey.

From Table 3, it is observed that

(a) For fixed values of ρ_{yx}, $\mu_{21}^{(0)}$ do not follow any definite pattern when the value of ρ_{yz} is increased while $E_{21}^{(1)}$ and $E_{21}^{(2)}$ are increasing with the increasing values of ρ_{yz}.

(b) For fixed values of ρ_{yz}, $\mu_{21}^{(0)}$ and $E_{21}^{(2)}$ do not follow any definite trend when the value of ρ_{yx} is increased while $E_{21}^{(1)}$ are increasing with the increase in the values of ρ_{yx}.

(c) The minimum value of $\mu_{21}^{(0)}$ is 0.1373, which indicates that the fraction of the fresh sample to be replaced on current occasion is as low as about 14 percent of the total sample size, leading to an appreciable reduction in the cost.

From Table 4, it can be seen that

(a) For fixed values of $\rho_{yx,}$ the values of $\mu_{22}^{(0)}$ are decreasing while the values of $E_{22}^{(1)}$ and $E_{22}^{(2)}$ are increasing with the increasing values of ρ_{yz}. This behaviour is highly desirable, since it concludes that if highly correlated auxiliary character is available, it pays in terms of enhanced precision of estimates as well as it reduces the cost of the survey.

(b) For fixed values of $\rho_{yz,}$ the values of $\mu_{22}^{(0)}$ and $E_{22}^{(1)}$ are increasing with the increasing values of ρ_{yx} while the values of $E_{22}^{(2)}$ are decreasing for some choices of $\rho_{yx.}$ This behaviour is in agreement with Sukhatme et al. (1984) results, which explains that the more value of ρ_{yx} , the more fraction of fresh sample is required on the current (second) occasion.

(c) The minimum value of $\mu_{22}^{(0)}$ is 0.3705, which indicates that the fraction of fresh sample to be replaced on current occasion is as low as about 37 percent of the total sample size, which leads to an appreciable reduction in the cost of the survey.

Thus, it is clear that the use of information on the auxiliary variable is highly rewarding in terms of the proposed estimators. It is also clear that if a highly correlated auxiliary variable is used, only a relatively smaller 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.

Acknowledgement

Authors are thankful to the referee for his valuable and inspiring suggestions. Authors are also thankful to the Indian School of Mines, Dhanbad for providing financial and infrastructural support to carry out the present work.

REFERENCES

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

COCHRAN, W.G., (1977): Sampling Techniques. 3rd ed. New York: Wiley.Chaturvedi, D. K. and Tripathi, T. P. (1983): Estimation of population ratio on two occasions using multivariate auxiliary information. *Jour. Ind. Statist. Assoc*., 21, 113-120.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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).*