Multivariate Rotation Design for Population Mean in Sampling on Successive Occasions

Kumari Priyanka[a], Richa Mittal[a], Jong-Min Kim[1,b]

[a]Department of Mathematics, Shivaji College (University of Delhi), India;
[b]Statistics, Division of Science and Mathematics, University of Minnesota-Morris, USA

Abstract

This article deals with the problem of estimation of the population mean in presence of multi-auxiliary information in two occasion rotation sampling. A multivariate exponential ratio type estimator has been proposed to estimate population mean at current (second) occasion using information on p-additional auxiliary variates which are positively correlated to study variates. The theoretical properties of the proposed estimator are investigated along with the discussion of optimum replacement strategies. The worthiness of proposed estimator has been justified by comparing it to well-known recent estimators that exist in the literature of rotation sampling. Theoretical results are justified through empirical investigations and a detailed study has been done by taking different choices of the correlation coefficients. A simulation study has been conducted to show the practicability of the proposed estimator.

Keywords: population mean, successive sampling, bias, mean square error, optimum replacement strategy, multi-auxiliary information

1. Introduction

Longitudinal surveys are correlational research studies that involve repeated observations of the same variables over long periods. Longitudinal studies are often used in psychology to study developmental trends across life spans and in sociology to study generational or lifetime. Longitudinal studies track the same people and the observed differences in people are less likely to be the result of cultural differences across generations. Because of this benefit, longitudinal studies make observing changes more accurate, and they are applied in various other fields. In medicine, the design is used to uncover predictors of certain diseases. In advertising, the design identifies the changes that advertising has produced in the attitudes and behaviors of those within the target audience who have seen the advertising campaign.

Many researchers have taken advantage of longitudinal surveys, refer to Chaturvedi and Tripathi (1983), Das (1982), Gupta (1979), Jessen (1942), Patterson (1950) and Rao and Graham (1964).

Sometimes we sense that different variables are related to the study characteristics that may be helpful to estimate the study characteristics. For example many countries keep track of the population through a total population register that is often used as a sampling frame for individuals or households. The register contains a number of variables (some quantitative and some categorical) that may

serve as auxiliary information to identify the human development index. Age and the taxable income of individual are the quantitative auxiliary variables while sex of the individual, marital status and residential specification may be considered as categorical auxiliaries.

Utilizing the auxiliary information on both the occasions Biradar and Singh (2001), Feng and Zou (1997), Singh (2005), Singh and Singh (2001), Singh *et al.* (1991), Sen (1972, 1973) have successfully added some literature in the field of successive sampling. Singh and Karna (2009), Singh and Prasad (2010), Singh and Priyanka (2006, 2007, 2008) proposed a variety of estimators to estimate the population mean on current (second) occasion in two occasions successive sampling.

The linear regression estimator is more efficient than the ratio estimator except when the regression line y on x passes through the neighborhood of the origin; in this case the efficiencies of the estimators are almost equal. There are also many practical situations when the regression line does not pass through the neighborhood of the origin, in such cases the ratio estimator does not perform as good as the linear regression estimator. The proposed work utilizes multi-auxiliary information available on both occasions and stable overtime. Multi-auxiliary information are blended with exponential type structures and a multivariate exponential ratio type estimator has been proposed to estimate the population mean at the current occasion in two occasion rotation sampling. The properties of the proposed estimator are derived up-to the first order of approximation and optimum replacement strategies are discussed. The properties have been corroborated empirically. The proposed multivariate estimator is compared with recent literature in rotation sampling due to Singh (2005) and Singh and Priyanka (2008). A simulation study has been conducted to determine the working efficiency of the proposed estimator. It has been observed that the proposed exponential type structure works well even if the auxiliary variables have a low correlation with the study variable.

2. Sample Structure and Notations

Let $U = (U_1, U_2, \ldots, U_N)$ be the finite population of N units, which has been sampled over two occasions. We assumed that the size of the population remains unchanged but the values of units change over two occasions. The characters under study have been denoted by x and y on the first and second occasions respectively. It has been assumed that information on p-additional auxiliary variables, $z_1, z_2, \ldots, z_p$ whose population means are known, correlated to x and y, stable over the occasions and readily available on both occasions. Simple random sample (without replacement) of n units is taken on the first occasion. A random subsample of $m = n\lambda$ units is retained for use on the second (current) occasion. Now at the current occasion a simple random sample (without replacement) of $u = (n-m) = n\mu$ units is drawn fresh from the remaining $(N-n)$ units of the population so that the sample size on the second occasion is also n. Let μ and λ ($\mu + \lambda = 1$; $0 \le \mu, \lambda \le 1$) are the fractions of fresh and matched samples respectively at the second (current) occasion.

3. Formulation of the Proposed Estimator $T_{|p}$

To estimate the population mean $\bar{Y}$ on the current (second) occasion, utilizing p-additional auxiliary information which is stable over time and readily available on both the occasions, a multivariate weighted estimator T_u based on sample of the size $u = n\mu$ drawn fresh on the current (second) occasion is proposed as

$$T_u = W_u' T_{\exp}(u), \tag{3.1}$$

where

W_u is a column vector of p-weights given by $W_u = [w_{u_1} \ w_{u_2} \ \cdots \ w_{u_p}]'$,

$$T_{\exp}(u) = \begin{bmatrix} T(1,u) \\ T(2,u) \\ \vdots \\ T(p,u) \end{bmatrix}, \quad \text{where} \ \ T(i,u) = \bar{y}_u \exp\left(\frac{\bar{Z}_i - \bar{z}_i(u)}{\bar{Z}_i + \bar{z}_i(u)}\right) \quad \text{for} \ \ i = 1,2,3,\ldots,p,$$

such that $\mathbf{1}'W_u = 1$, where $\mathbf{1}$ is a column vector of order p.

The second estimator T_m is also proposed as weighted multivariate chain type ratio to the exponential ratio estimator based on sample size $m = n\lambda$ common to both occasions and given by

$$T_m = W'_m T_{\exp}(m,n), \tag{3.2}$$

where

W_m is a column vector of p-weights as $W_m = [w_{m_1} \ w_{m_2} \ \cdots \ w_{m_p}]'$,

$$T_{\exp}(m,n) = \begin{bmatrix} T(1,m,n) \\ T(2,m,n) \\ \vdots \\ T(p,m,n) \end{bmatrix}, \quad \text{where} \ \ T(i,m,n) = \left(\frac{\bar{y}^*(i,m)}{\bar{x}^*(i,m)}\right)\bar{x}^*(i,n),$$

where

$$\bar{y}^*(i,m) = \bar{y}_m \exp\left(\frac{\bar{Z}_i - \bar{z}_i(m)}{\bar{Z}_i + \bar{z}_i(m)}\right), \quad \bar{x}^*(i,m) = \bar{x}_m \exp\left(\frac{\bar{Z}_i - \bar{z}_i(m)}{\bar{Z}_i + \bar{z}_i(m)}\right),$$

$$\bar{x}^*(i,n) = \bar{x}_n \exp\left(\frac{\bar{Z}_i - \bar{z}_i(n)}{\bar{Z}_i + \bar{z}_i(n)}\right), \quad \text{for} \ \ i = 1,2,3,\ldots,p.$$

Such that $\mathbf{1}'W_m = 1$, where $\mathbf{1}$ is a column vector of order p.

The optimum weights W_u and W_m in T_u and T_m are chosen by minimizing mean square errors respectively.

Now a convex linear combination of the two estimators T_u and T_m has been considered to define the final estimator of population mean $\bar{Y}$ on the current occasion and is given as

$$T_{|p} = \varphi T_u + (1 - \varphi)T_m, \tag{3.3}$$

where $\varphi \ (0 \leq \varphi \leq 1)$ is an unknown constant to be determined so as to minimize the mean square error of the estimator $T_{|p}$.

4. Properties of the Proposed Estimator $T_{|p}$

The properties of the proposed estimator $T_{|p}$ are derived under following large sample approximations: $\bar{y}_u = \bar{Y}(1 + e_0)$, $\bar{y}_m = \bar{Y}(1 + e_1)$, $\bar{x}_m = \bar{X}(1 + e_2)$, $\bar{x}_n = \bar{X}(1 + e_3)$, $\bar{z}_i(u) = \bar{Z}_i(1 + e_{4i})$, $\bar{z}_i(m) = \bar{Z}_i(1 + e_{5i})$ and $\bar{z}_i(n) = \bar{Z}_i(1 + e_{6i})$ such that $|e_k| < 1 \ \forall \ k = 0,1,2,3,4,5$ and 6 and $|e_{ki}| < 1 \ \forall \ i = 1,2,3,\ldots,p$.

Under the above transformations, the estimators T_u and T_m take the following forms:

$$T(i, u) = \frac{\bar{Y}}{8}\left(8 + 8e_0 - 4e_{4i} - 4e_0e_{4i} + 3e_{4i}^2\right), \quad \text{for } i = 1, 2, \ldots, p, \tag{4.1}$$

$$T(i, m, n) = \frac{\bar{Y}}{8}\left(8 + 8e_1 - 8e_2 + 8e_3 - 4e_{6i} - 8e_1e_2 + 8e_1e_3 - 4e_1e_{6i} - 8e_2e_3\right.$$
$$\left. + 4e_2e_{6i} - 4e_3e_{6i} + 8e_2^2 + 3e_{6i}^2\right), \quad \text{for } i = 1, 2, \ldots, p. \tag{4.2}$$

Thus we have the following theorems:

Theorem 1. *The bias of the proposed estimator $T_{|p}$ to the first order of approximation is obtained as*

$$B(T_{|p}) = \varphi B(T_u) + (1 - \varphi)B(T_m), \tag{4.3}$$

$$B(T_u) = \frac{1}{u}W_u'\boldsymbol{B}_u, \tag{4.4}$$

$$B(T_m) = W_m'\left(\frac{1}{m}\boldsymbol{B}_{m1} + \frac{1}{n}\boldsymbol{B}_{m2}\right), \tag{4.5}$$

where

$$\boldsymbol{B}_u = \left(B_1(u), B_2(u), \ldots, B_p(u)\right)', \quad B_i(u) = \frac{1}{u}\bar{Y}\left(\frac{3}{8}\frac{C_{002}}{\bar{Z}_i^2} - \frac{1}{2}\frac{C_{011}}{\bar{Y}\bar{Z}_i}\right), \quad \text{for } i = 1, 2, 3, \ldots, p,$$

$$\boldsymbol{B}_{m1} = \bar{Y}\left(\frac{C_{200}}{\bar{X}^2} - \frac{C_{110}}{\bar{X}\bar{Y}}\right),$$

$$\boldsymbol{B}_{m2} = \left(Bm_{21}, Bm_{22}, \ldots, Bm_{2p}\right),$$

where

$$Bm_{2i} = \bar{Y}\left(\frac{3}{8}\frac{C_{002}}{\bar{Z}_i^2} + \frac{C_{110}}{\bar{X}\bar{Y}} - \frac{1}{2}\frac{C_{011}}{\bar{Y}\bar{Z}_i} - \frac{C_{200}}{\bar{X}^2}\right), \quad C_{rst} = E\left[\left(x_i - \bar{X}\right)^r\left(y_i - \bar{Y}\right)^s\left(z_i - \bar{Z}\right)^t\right];$$

$(r, s, t) \geq 0$ *for $i = 1, 2, 3, \ldots, p$.*

Proof: The bias of the estimator $T_{|p}$ is given by

$$B(T_{|p}) = E\left[T_{|p} - Y\right] = \varphi B(T_u) + (1 - \varphi)B(T_m),$$

where $B(T_u) = E[T_u - Y]$ and $B(T_m) = E[T_m - Y]$.

Using large sample approximations assumed in Section 4 and retaining terms upto the first order of approximations, the expression for $T(i, u)$ and $T(i, m, n)$ for $i = 1, 2, 3, \ldots, p$ are obtained as in equations (4.1) and (4.2) respectively and hence using equations (4.1) and (4.2) in equations (3.1) and (3.2) respectively the expression for $B(T_u)$ and $B(T_m)$ are obtained as in equations (4.4) and (4.5) respectively, hence the expression for bias of the estimator $T_{|p}$ is obtained as in equation (4.3). □

Theorem 2. *The mean square error of the estimator $T_{|p}$ is given by*

$$M(T_{|p}) = \varphi^2 M(T_u) + (1 - \varphi)^2 M(T_m) + 2\varphi(1 - \varphi)\operatorname{cov}(T_u, T_m), \tag{4.6}$$

$$M(T_u) = W_u'K_uW_u, \tag{4.7}$$

$$M(T_m) = (B)W_m'EW_m + W_m'K_mW_m, \tag{4.8}$$

where

$$\boldsymbol{W}_u = \begin{bmatrix} W_{u_1} & W_{u_2} \cdots W_{u_p} \end{bmatrix}', \qquad \boldsymbol{W}_m = \begin{bmatrix} W_{m_1} & W_{m_2} \cdots W_{m_p} \end{bmatrix}',$$

E is a unit matrix of order $p \times p$,

$$B = \left(\frac{1}{m} - \frac{1}{N}\right) B_1, \quad B_1 = 2\bar{Y}^2(1 - \rho_{yx})C_0^2,$$

$$\boldsymbol{K}_u = \left(\frac{1}{u} - \frac{1}{N}\right) \boldsymbol{K}_{u^*}, \qquad \boldsymbol{K}_m = \left(\frac{1}{u} - \frac{1}{N}\right) \boldsymbol{K}_{m^*},$$

where

$$\boldsymbol{K}_{u^*} = \begin{bmatrix} ku_{11} & ku_{12} & \dots & ku_{1p} \\ ku_{21} & ku_{22} & \dots & ku_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ku_{p1} & ku_{p2} & \dots & ku_{pp} \end{bmatrix}_{p\times p} \quad and \quad \boldsymbol{K}_{m^*} = \begin{bmatrix} km_{11} & km_{12} & \dots & km_{1p} \\ km_{21} & km_{22} & \dots & km_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ km_{p1} & km_{p2} & \dots & km_{pp} \end{bmatrix}_{p\times p},$$

where

$$ku_{ii} = \bar{Y}^2 \left(C_0^2 + \frac{1}{4}C_{z_i}^2 - \rho_{yz_i}C_0 C_{z_i} \right),$$

$$ku_{ij} = \bar{Y}^2 \left(C_0^2 - \frac{1}{2}\rho_{yz_i}C_0 C_{z_i} - \frac{1}{2}\rho_{yz_j}C_0 C_{z_j} + \frac{1}{4}\rho_{z_i z_j}C_{z_i}C_{z_j} \right),$$

$$km_{ii} = \bar{Y}^2 \left(C_0^2(2\rho_{yx}-1) - \rho_{yz_i}C_0 C_{z_i} + \frac{1}{4}C_{z_i}^2 \right),$$

$$km_{ij} = \bar{Y}^2 \left(C_0^2(2\rho_{yx} - 1) - \frac{1}{2}\rho_{yz_i}C_0 C_{z_i} - \frac{1}{2}\rho_{yz_j}C_0 C_{z_j} + \frac{1}{4}\rho_{z_i z_j}C_{z_i}C_{z_j} \right) \quad \forall\, i \neq j = 1, 2, 3, \dots, p.$$

Proof: The mean square error of the estimator $T_{|p}$ is given by

$$\begin{aligned} M(T_{|p}) &= E\left[T_{|p} - \bar{Y} \right]^2 \\ &= E\left[\varphi\left(T_u - \bar{Y}\right) + (1 - \varphi)\left(T_m - \bar{Y}\right) \right]^2 \\ &= \varphi^2 M(T_u) + (1 - \varphi)^2 M(T_m) + 2\varphi(1 - \varphi)\operatorname{cov}(T_u, T_m), \end{aligned}$$

where $M(T_u) = E[T_u - \bar{Y}]^2$ and $M(T_m) = E[T_m - \bar{Y}]^2$.

The estimators T_u and T_m are based on two independent samples of sizes u and m respectively, hence $\operatorname{cov}(T_u, T_m) = 0$, considering the population is sufficiently large so using large sample approximations assumed in Section 4 and retaining terms upto the first order of approximations and also assuming $C_x = C_y = C_0$ (following Cochran (1977)), the expression for $M(T_u)$ and $M(T_m)$ are obtained from equations (4.7) and (4.8) and hence the expression for mean square error of estimator $T_{|p}$ is obtained as in equation (4.6). $\qquad\square$

5. Choice of Optimal Weights

To find the optimization of the weight vector $\boldsymbol{W}_u = [w_{u_1} \; w_{u_2} \; \cdots \; w_{u_p}]'$, the mean square error $M(T_u)$ given in equation (4.7) is minimized subject to the condition $\boldsymbol{1}'\boldsymbol{W}_u = 1$ using the method of Lagrange multiplier explained as:

To find the extrema using Lagrange multiplier technique, we define L_u as

$$L_u = W'_u K_u W_u - \lambda_u(1'W_u - 1), \tag{5.1}$$

where $\mathbf{1}$ is a unit column vector of order p and $\lambda_u$is the Lagrange multiplier.

By differentiating equation (5.1) partially with respect to W_u and equating it to zero we have

$$\frac{\partial L_u}{\partial W_u} = \frac{\partial}{\partial W_u} \left[W'_u K_u W_u - \lambda_u(1'W_u - 1) \right] = 0.$$

This implies that, $2K_u W_u - \lambda_u \mathbf{1} = 0$, which yields

$$W_u = \frac{\lambda_u}{2} K_u^{-1} \mathbf{1}. \tag{5.2}$$

Now pre-multiplying equation (5.2) by $\mathbf{1}'$, we get

$$\frac{\lambda_u}{2} = \frac{1}{1' K_u^{-1} \mathbf{1}}. \tag{5.3}$$

Using equation (5.3) in equation (5.2), we obtain the optimal weight vector as

$$W_{u_{opt.}} = \frac{K_u^{-1}}{1' K_u^{-1} \mathbf{1}}. \tag{5.4}$$

In similar manners, the optimal of the weight $W_m = [w_{m_1} \ w_{m_2} \cdots w_{m_p}]'$ is obtained by minimizing $M(T_m)$ subject to the constraint $1'W_m = 1$ using the method of Lagrange multiplier, for this we define

$$L_m = (B)W'_m E W_m + W'_m K_m W_m - \lambda(1'W_m - 1),$$

where λ_m is the Lagrange multiplier.

Now, differentiating L_m with respect to W_m and equating to 0, we get

$$W_{m_{opt.}} = \frac{K_m^{-1}}{1' K_m^{-1} \mathbf{1}}. \tag{5.5}$$

Then substituting the optimum values of W_u and W_m in equations (4.7) and (4.8) respectively, the optimum mean square errors of the estimators are obtained as:

$$M(T_u)_{opt.} = \left(\frac{1}{u} - \frac{1}{N} \right) \frac{1}{1' K_{u^*}^{-1} \mathbf{1}}, \tag{5.6}$$

$$M(T_m)_{opt.} = \left(\frac{1}{m} - \frac{1}{N} \right) B_1 + \left(\frac{1}{n} - \frac{1}{N} \right) \frac{1}{1' K_{m^*}^{-1} \mathbf{1}}. \tag{5.7}$$

6. Minimum Mean Square Error of the Proposed Estimator $T_{|p}$

The mean square error of the proposed estimator $T_{|p}$ is given by

$$M(T_{|p}) = \varphi^2 M(T_u)_{opt.} + (1 - \varphi)^2 M(T_m)_{opt.}.$$

Minimizing $M(T_{|p})$ with respect to φ gives the optimum value of φ as

$$\varphi_{opt.} = \frac{M(T_m)_{opt.}}{M(T_u)_{opt.} + M(T_m)_{opt.}}. \tag{6.1}$$

Now substituting the above value of $\varphi_{opt.}$ in equation (4.6), we obtain the optimum mean square error of the estimators $T_{|p}$ as

$$M(T_{|p})^*_{opt.} = \frac{M(T_u)_{opt.}.M(T_m)_{opt.}}{M(T_u)_{opt.}+M(T_m)_{opt.}}. \tag{6.2}$$

Further, substituting the optimum values of the mean square errors of the estimators given in equations (5.6) and (5.7) in equations (6.1) and (6.2) respectively, the simplified values $\varphi_{opt.}$ and $M(T_{|p})^*_{opt.}$ are obtained as

$$\varphi_{opt.} = \frac{\mu[\mu C-(B_1+C)]}{[\mu^2 C - \mu(B_1+C - A)-A]}, \tag{6.3}$$

$$M(T_{|p})^*_{opt.} = \frac{1}{n} \frac{[\mu D_1-D_2]}{[\mu^2 C - \mu D_3-A]}, \tag{6.4}$$

where $A = 1/(\mathbf{1}'K_{u^*}^{-1}\mathbf{1})$, $B_1 = 2\bar{Y}^2(1-\rho_{yx})C_0^2$, $C = 1/(\mathbf{1}'K_{m^*}^{-1}\mathbf{1})$, $D_1 = AC$, $D_2 = AB_1 + AC$, $D_3= B_1+C - A$ and μ is the fraction of the sample drawn afresh at the current (second) occasion.

7. Optimum Replacement Strategy for the Estimator $T_{|p}$

The idea of longitudinal surveys is concerned with obtaining efficient estimates with minimal costs in conducting the survey. It is technically convenient to maintain a high overlap between repeats of the survey that provide advantages due to many sampled units location and use in the survey. Hence the decision of the optimum value should be made (fractions of samples to be drawn fresh on the current occasion) so that $\bar{Y}$ may be estimated with maximum precision and minimum cost, we minimize the mean square error $M(T_{|p})^*_{opt.}$ in equation (6.4) with respect to μ as:

$$\frac{\partial\left(M(T_{|p})^*_{opt.}\right)}{\partial\mu} = 0, \quad \Rightarrow \quad \mu^2 G_1 - 2\mu G_2 + G_3 = 0.$$

Thus the optimum value of μ so obtained is one of the two roots given by

$$\hat{\mu} = \frac{G_2 \pm \sqrt{G_2^2-G_1G_3}}{G_1}, \tag{7.1}$$

where $G_1 = CD_1$, $G_2 = CD_2$ and $G_3 = AD_1 + D_2D_3$.

The real value of $\hat{\mu}$ exist, iff $G_2^2 - G_1G_3 \geq 0$. For any situation, which satisfies this condition, two real values of $\hat{\mu}$ may be possible, hence choose a value of $\hat{\mu}$ such that $0 \leq \hat{\mu} \leq 1$. All other values of $\hat{\mu}$ are inadmissible. If both the real values of $\hat{\mu}$ are admissible, the lowest one will be the best choice because it reduces the total cost of the survey. Substituting the admissible value of $\hat{\mu}$ say $\mu_{T_{|p}}$ from (7.1) in to the equation (6.4), we get the optimum value of the mean square error of the estimator $T_{|p}$ with respect to φ as well as μ which, is given as

$$M\left(T_{|p}\right)^{**}_{opt.} = \frac{1}{n} \frac{\left[\mu_{T_{|p}}D_1-D_2\right]}{\left[\mu_{T_{|p}}^2 C - \mu_{T_{|p}}D_3-A\right]}. \tag{7.2}$$

8. Efficiency with Increased Number of Auxiliary Variables

We know that increasing the number of auxiliary variables typically increases the precision of the estimates. In this section we verify the property for the proposed estimator as under: Let $T_{|p}$ and $T_{|q}$ be two proposed estimators based on p and q auxiliary variables respectively such that $p < q$, then $M(T_{|p}) \geq M(T_{|q})$, i.e.

$$M\left(T_{|p}\right) - M\left(T_{|q}\right) \geq 0 \tag{8.1}$$

$$\frac{1}{n} \frac{\left[\mu A_p C_p - A_p \left(B + C_p\right)\right]}{\left[\mu^2 C_p - \mu \left(B + C_p + A_p\right) - A_p\right]} - \frac{1}{n} \frac{\left[\mu A_q C_q - A_q \left(B + C_q\right)\right]}{\left[\mu^2 C_q - \mu \left(B + C_q + A_q\right) - A_q\right]} \geq 0.$$

On simplification, we get

$$\left(A_p - A_q\right) \left[\left(\mu - 1\right)^2 \left(\mu C_p C_q + \frac{A_p A_q \left(C_p - C_q\right)}{\left(A_p - A_q\right)}\right) - \mu B \left(\left(C_p - C_q\right) \left(\mu - 1\right) - B\right)\right] \geq 0.$$

This reduces to the condition

$$\left(A_p - A_q\right) \geq 0. \tag{8.2}$$

From Section 6 above, we get

$$\frac{1}{\mathbf{1}' K_p^{-1} \mathbf{1}} - \frac{1}{\mathbf{1}' K_q^{-1} \mathbf{1}} \geq 0$$

$$\mathbf{1}' K_q^{-1} \mathbf{1} \geq \mathbf{1}' K_p^{-1} \mathbf{1}.$$

Following Rao (2006), the matrix K_q can be partitioned and can be written as

$$K_q = \begin{pmatrix} K_p & F \\ F' & G \end{pmatrix},$$

where F, F' and G are matrices deduced from K_q such that their order never exceeds $q - p$ and always greater than or equal to 1. Then,

$$K_q^{-1} = \begin{bmatrix} K_p^{-1} + HJH' & -HJ \\ -JH' & J \end{bmatrix}, \tag{8.3}$$

where $J = (G - F' K_p^{-1} F)^{-1}$ and $H = K_p^{-1} F$ (see Rao (2006) and Olkin (1958)).

Now rewriting $\mathbf{1}' K_q^{-1} \mathbf{1}$ by putting the value of K_q^{-1} from equation (8.3), we get

$$\mathbf{1}' K_q^{-1} \mathbf{1} = \begin{bmatrix} \mathbf{1}_p & \mathbf{1}_{q-p} \end{bmatrix}' \begin{bmatrix} K_p^{-1} + HJH' & -HJ \\ -JH' & J \end{bmatrix} \begin{bmatrix} \mathbf{1}_p \\ \mathbf{1}_{q-p} \end{bmatrix}$$

$$= \left(\mathbf{1}_p' \left(K_p^{-1} + HJH'\right) - \mathbf{1}_{q-p}' JH' \quad -\mathbf{1}_p' HJ + \mathbf{1}_{q-p}' J\right) \begin{bmatrix} \mathbf{1}_p \\ \mathbf{1}_{q-p} \end{bmatrix}$$

$$= \mathbf{1}_p' (K_p^{-1} + HJH') \mathbf{1}_p - \mathbf{1}_{q-p}' JH' \mathbf{1}_p - \mathbf{1}_p' HJ \mathbf{1}_{q-p} + \mathbf{1}_{q-p}' J \mathbf{1}_{q-p}$$

implies

$$\mathbf{1}'K_q^{-1}\mathbf{1}-\mathbf{1}'_p(K_p^{-1})\mathbf{1}_p = \mathbf{1}'_p(HJH')\mathbf{1}_p-\mathbf{1}'_{q-p}JH'\mathbf{1}_p-\mathbf{1}'_pHJ\mathbf{1}_{q-p}+\mathbf{1}'_{q-p}J\mathbf{1}_{q-p},$$

$$\mathbf{1}'K_q^{-1}\mathbf{1}-\mathbf{1}'_p(K_p^{-1})\mathbf{1}_p = \begin{bmatrix} \mathbf{1}_p & \mathbf{1}_{q-p} \end{bmatrix}' \begin{bmatrix} HJH' & -HJ \\ -JH' & J \end{bmatrix} \begin{bmatrix} \mathbf{1}_p \\ \mathbf{1}_{q-p} \end{bmatrix},$$

$$\mathbf{1}'K_q^{-1}\mathbf{1}-\mathbf{1}'_p(K_p^{-1})\mathbf{1}_p = \mathbf{1}' \begin{bmatrix} H \\ -I \end{bmatrix} J \begin{bmatrix} H & -I \end{bmatrix} \mathbf{1} \geq 0.$$

The latter follows since J is positive definite so that $R'JR \geq 0$ for all R, where $R = \begin{pmatrix} H & -I \end{pmatrix}\mathbf{1}$.
   Hence from equation (8.1), we have

$$M(T_{|p})-M(T_{|q}) \geq 0.$$

This leads to the result that utilizing more auxiliary variables provides more efficient estimates in terms of mean square error for the proposed estimator.

9. Special Cases

Case 1

There are several instances where the p-auxiliary variates are mutually uncorrelated but are correlated to study variates, for example in a survey of commercial products where the aim is to estimate the number of persons reading newspaper. Then in that case the numbers of copies produced by different newspaper companies are different and the number of copies produced by a particular newspaper company is uncorrelated to the number of copies produced by another newspaper, but both are correlated to study variates, i.e., number of persons reading the newspaper. Similarly, the total seating capacity of different airlines may be treated as auxiliary variates in a transportation survey where the aim is to estimate the number of persons traveling by air per year. The seating capacity of different airlines is different and they are mutually uncorrelated, but the information on this will contribute a lot in estimation of the number of persons traveling by air. Hence, for modelling such type of situations where the p-auxiliary variates are mutually uncorrelated, i.e. $\rho_{z_i z_j} = 0 \ \forall \ i \neq j = 1, 2, \ldots, p$, the proposed multivariate exponential ratio type estimator is applicable and the optimum value of μ say μ^0 and the optimum value of the mean square error of the estimator $T_{|p}$ with respect to φ as well as μ^0 is given by

$$\hat{\mu}^0 = \frac{G_2^0 \pm \sqrt{G_2^{0^2} - G_1^0 G_3^0}}{G_1^0} \tag{9.1}$$

and

$$M\left(T_{|p}\right)_{opt.}^{**} = \frac{1}{n} \frac{\left[\mu_{T_{|p}}^0 D_1^0 - D_2^0\right]}{\left[\mu_{T_{|p}}^{0^2} C^0 - \mu_{T_{|p}}^0 D_3^0 - A^0\right]}, \tag{9.2}$$

where

$$G_1^0 = C^0 D_1^0, \qquad G_2^0 = C^0 D_2^0, \qquad G_3^0 = A^0 D_1^0 + D_2^0 D_3^0, \qquad A^0 = \frac{1}{\mathbf{1}'S_{u^*}^{-1}\mathbf{1}}, \qquad C^0 = \frac{1}{\mathbf{1}'S_{m^*}^{-1}\mathbf{1}},$$

$$B_1^0 = 2\bar{Y}^2(1 - \rho_{yx})C_0^2, \qquad D_1^0 = A^0 C^0, \qquad D_2^0 = A^0 B_1^0 + A^0 C^0, \qquad D_3^0 = B_1^0 + C^0 - A^0,$$

where

$$
S_{u^*} = \begin{bmatrix} su_{11} & su_{12} & \dots & su_{1p} \\ su_{21} & su_{22} & \dots & su_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ su_{p1} & su_{p2} & \dots & su_{pp} \end{bmatrix}_{p \times p} \quad \text{and} \quad S_{m^*} = \begin{bmatrix} sm_{11} & sm_{12} & \dots & sm_{1p} \\ sm_{21} & sm_{22} & \dots & sm_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ sm_{p1} & sm_{p2} & \dots & sm_{pp} \end{bmatrix}_{p \times p},
$$

where

$$
su_{ii} = \bar{Y}^2 \left(C_0^2 + \frac{1}{4} C_{z_i}^2 - \rho_{yz_i} C_0 C_{z_i} \right), \quad su_{ij} = \bar{Y}^2 \left(C_0^2 - \frac{1}{2} \rho_{yz_i} C_0 C_{z_i} - \frac{1}{2} \rho_{yz_j} C_0 C_{z_j} \right),
$$

$$
sm_{ii} = \bar{Y}^2 \left(C_0^2 (2\rho_{yx} - 1) - \rho_{yz_i} C_0 C_{z_i} + \frac{1}{4} C_{z_i}^2 \right), \quad sm_{ij} = \bar{Y}^2 \left(C_0^2 (2\rho_{yx} - 1) - \frac{1}{2} \rho_{yz_i} C_0 C_{z_i} - \frac{1}{2} \rho_{yz_j} C_0 C_{z_j} \right),
$$

$\forall \, i \neq j = 1, 2, 3, \dots, p.$

Case 2

The p-auxiliary variates are mutually correlated i.e. $\rho_{z_i z_j} \neq 0 \, \forall \, i \neq j = 1, 2, \dots, p.$ In this case if there is high correlation between p-auxiliary variates, then such a problem can be addressed as a problem of multicollinearity in survey sampling.

10. Efficiency Comparison

To examine the performance of the proposed estimator with some recent estimators due to Singh (2005) and Singh and Priyanka (2008) in successive sampling, some assumptions have been considered for the proposed estimator to discuss the properties of estimators as estimators proposed by Singh (2005) and Singh and Priyanka (2008).

Following Artes Rodriguez and Gracia Luengo (2005), Olkin (1958), Raj (1965) and Singh *et al.* (2011) we consider $C_0 = C_{z_i}; \, \forall \, i = 1, 2, 3, \dots, p$ approximately and hence, the optimum value of μ say $\hat{\mu}^*$ and optimum value of mean square error $M(T_{|p})^{**}_{opt.}$ of the proposed estimator $T_{|p}$ reduces to

$$
\hat{\mu}^* = \frac{G_2^* \pm \sqrt{G_2^{*2} - G_1^* G_3^*}}{G_1^*}, \tag{10.1}
$$

$$
M(T_{|p})^{**}_{opt.} = \frac{1}{n} \frac{\left[\mu_{T_{|p}}^* D_1^* - D_2^* \right]}{\left[\mu_{T_{|p}}^{*2} C^* - \mu_{T_{|p}}^* D_3^* - A^* \right]}, \tag{10.2}
$$

where

$$
G_1^* = C^* D_1^*, \qquad G_2^* = C^* D_2^*, \qquad G_3^* = A^* D_1^* + D_2^* D_3^*, \qquad A^* = \frac{1}{\mathbf{1} H_{u^*}^{-1} \mathbf{1}}, \qquad C^* = \frac{1}{\mathbf{1} H_{m^*}^{-1} \mathbf{1}},
$$

$$
B_1^* = 2(1 - \rho_{yx}) S_y^2, \qquad D_1^* = A^* C^*, \qquad D_2^* = A^* B_1^* + A^* C^*, \qquad D_3^* = B_1^* + C^* - A^*,
$$

where

$$
H_{u^*} = \begin{bmatrix} hu_{11} & hu_{12} & \dots & hu_{1p} \\ hu_{21} & hu_{22} & \dots & hu_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ hu_{p1} & hu_{p2} & \dots & hu_{pp} \end{bmatrix}_{p \times p} \quad \text{and} \quad H_{m^*} = \begin{bmatrix} hm_{11} & hm_{12} & \dots & hm_{1p} \\ hm_{21} & hm_{22} & \dots & hm_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ hm_{p1} & hm_{p2} & \dots & hm_{pp} \end{bmatrix}_{p \times p},
$$

where

$$hu_{ii} = \left(\frac{5}{4} - \rho_{yz_i}\right)S_y^2, \qquad hu_{ij} = \left(1 - \frac{1}{2}\rho_{yz_i} - \frac{1}{2}\rho_{yz_j} + \frac{1}{4}\rho_{z_i z_j}\right)S_y^2,$$

$$hm_{ii} = \left(2\rho_{yx} - \rho_{yz_i} - \frac{3}{4}\right)S_y^2, \qquad hm_{ij} = \left(2\rho_{yx} - \frac{1}{2}\rho_{yz_i} - \frac{1}{2}\rho_{yz_j} + \frac{1}{4}\rho_{z_i z_j} - 1\right)S_y^2,$$

$\forall \ i \neq j = 1, 2, 3, \ldots, p.$

10.1. Comparison of the proposed estimator $T_{|p}$ with respect to estimator T_S due to Singh (2005)

The estimator proposed by Singh (2005) is given as

$$T_S = \psi\frac{\bar{y}_u}{\bar{z}_u}\bar{Z} + (1 - \psi)\frac{\bar{y}_m}{\bar{x}_m}\frac{\bar{x}_n}{\bar{z}_n}\bar{Z}, \tag{10.3}$$

and the optimum mean square error of this estimator T_S is given by

$$M(T_S)_{opt.} = \frac{[\alpha_1^2 + \alpha_1\alpha_2\mu_S]S_y^2}{n[\alpha_1 + \alpha_2\mu_S^2]}$$

with $\alpha_1 = 2(1 - \rho_{yz})$, $\alpha_2 = 2(\rho_{yz} - \rho_{yx})$ and $\mu_S = -(1 - \rho_{yz}) \pm \sqrt{(1 - \rho_{yz})(1 - \rho_{yx})}/(\rho_{yz} - \rho_{yx})$.

Hence, the percent relative efficiency of the proposed estimator with respect to T_S is given as

$$E_{T_{|p}}^S = \frac{M[T_s]_{opt.}}{M(T_{|p})_{opt.}^{**}} \times 100. \tag{10.4}$$

10.2. Comparison of the proposed estimator $T_{|p}$ with respect to estimator T_{SP} due to Singh and Priyanka (2008)

The proposed estimator $T_{|p}$ at optimum condition is also compared with respect to the estimator T_{SP} due to Singh and Priyanka (2008) given as

$$T_{SP} = \xi\left[\bar{y}_u + \beta_{yz}\left(\bar{Z} - \bar{z}_u\right)\right] + (1 - \xi)\left[\bar{y}_m^* + \beta_{yx}\left(\bar{x}_n^* - \bar{x}_m^*\right)\right], \tag{10.5}$$

where $\bar{y}_m^* = \bar{y}_m + \beta_{yz}(\bar{Z} - \bar{z}_m)$, $\bar{x}_n^* = \bar{x}_n + \beta_{xz}(\bar{Z} - \bar{z}_n)$, $\bar{x}_m^* = \bar{x}_m + \beta_{xz}(\bar{Z} - \bar{z}_m)$, β_{yz} and β_{xz} are the population regression coefficients of y on z and x on z respectively and ξ is constant so as to minimize the variance of the estimator T_{SP}.

The optimum variance of estimator T_{SP} is given as

$$V(T_{SP})_{opt.}^* = \frac{\kappa[\kappa + \mu_{SP}\chi]}{[\kappa + \mu_{SP}^2\chi]}\frac{S_y^2}{n},$$

where $\kappa = 1 - \rho_{yz}^2$, $\chi = 2\rho_{yz}^2\rho_{yx} - \rho_{yx}^2(1 + \rho_{yz}^2)$ and $\mu_{SP} = \{-\kappa \pm \sqrt{\kappa(\kappa + \chi)}\}/\chi$.

Hence, the percent relative efficiency $E_{T_{|p}}^{SP}$ for $(p = 1, 2, 3, \ldots)$ of the estimator $T_{|p}$ (under their respective optimum conditions) with respect to T_{SP} is given by

$$E_{T_{|p}}^{SP} = \frac{V[T_{SP}]_{opt.}^*}{M(T_{|p})_{opt.}^{**}} \times 100, \quad \text{for } p = 1, 2, 3, \ldots. \tag{10.6}$$

Table 1: Empirical comparison of the proposed estimator $T_{|p}$ ($p = 1, 2$ and 3) with respect to the estimators T_S and T_{SP} respectively at their optimum conditions for considered population

		$p = 1$ $\mu_1^* = 0.5355$	$p = 2$ $\mu_2^* = 0.5196$	$p = 3$ $\mu_3^* = 0.5137$
$E_{T_{|p}}^S$	$\mu_S = 0.5502$	109.55	120.77	125.17
$E_{T_{|p}}^{SP}$	$\mu_{SP} = 0.5496$	**	101.44	105.14

** denote estimator $T_{|p=1}$ does not perform better than T_{SP} in terms of efficiency.

Table 2: Monte Carlo Simulation results when the proposed estimator $T_{|p=1}$ is compared to T_S and T_{SP} respectively (considering $\psi = \xi$)

φ		ψ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	$E_1(S)$	1349.90	5895.10	12296.00	17324.00	36619.00	57157.00	72197.00	84832.00	84170.00
	$E_1(SP)$	104.42	102.05	127.52	171.80	262.54	375.88	520.91	638.86	867.55
0.3	$E_1(S)$	1168.90	4741.70	11796.00	16766.00	32572.00	45474.00	61039.00	81957.00	97409.00
	$E_1(SP)$	**	**	113.58	159.70	237.69	329.08	469.05	577.94	764.94
0.5	$E_1(S)$	894.80	3399.70	8454.40	12852.00	22913.00	30904.00	41579.00	60708.00	70099.00
	$E_1(SP)$	**	**	**	113.93	167.23	231.62	323.22	410.61	540.09
0.7	$E_1(S)$	608.00	2258.00	5620.80	8907.00	14926.00	19794.00	27410.00	40345.00	46610.00
	$E_1(SP)$	**	**	**	**	109.21	153.80	211.71	272.91	358.30
0.9	$E_1(S)$	413.45	1493.10	3817.90	5946.80	10156.00	13715.00	18796.00	26496.00	31819.00
	$E_1(SP)$	**	**	**	**	**	105.43	141.45	184.93	242.63

** denotes estimator $T_{|p=1}$ does not perform better than T_{SP} in terms of efficiency.

11. Empirical Illustrations and Monte Carlo Simulation

Population Source (Free access to the data by Statistical Abstracts of the United States)

To conduct the empirical study, the population of total electric consumption in different states of United States has been considered.

To conduct the numerical illustration we considered the case of three auxiliary information (i.e. $p = 3$) stable over time and available at both occasions. The population comprise of $N = 51$ states of the United States. Let

y_i: The total energy consumption during 2007 in the ith state of U.S.

x_i: The total energy consumption during 2002 in the ith state of U.S.

z_{1i}: The total energy consumption during 2001 in the ith state of U.S.

z_{2i}: The total energy consumption during 2000 in the ith state of U.S.

z_{3i}: The total energy consumption during 1999 in the ith state of U.S.

For the considered population, the values of $\mu_{T_{|p}}^*$ ($p = 1, 2$ and 3) defined in equation (10.2) and percent relative efficiencies $E_{T_{|p}}^S$ and $E_{T_{|p}}^{SP}$ defined in equation (10.4) and (10.6) of $T_{|p}$ ($p = 1, 2$ and 3) with respect to T_S and T_{SP} have been computed (Table 2).

11.1. Simulation algorithm

(i) Choose 5000 samples of size $n = 20$ using simple random sampling without replacement on first occasions for both the study and auxiliary variables.

Table 3: Monte Carlo Simulation results when the proposed estimator $T_{|p=2}$ are compared to T_S and T_{SP} respectively (considering $\psi = \xi$)

φ		ψ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	$E_2(S)$	1335.20	5940.60	12400.00	17376.00	36883.00	57438.00	73062.00	85114.00	84732.00
	$E_2(SP)$	104.44	102.84	128.60	172.31	264.43	377.72	527.15	640.98	873.34
0.3	$E_2(S)$	1178.30	11795.10	11935.00	16892.00	32827.00	45857.00	61624.00	82646.00	98297.00
	$E_2(SP)$	**	**	114.92	160.90	239.55	331.85	473.55	582.80	771.92
0.5	$E_2(S)$	902.63	3431.40	8553.50	12956.00	23109.00	31143.00	41964.00	61237.00	70719.00
	$E_2(SP)$	**	**	**	114.85	168.66	233.41	326.21	414.19	544.87
0.7	$E_2(S)$	612.36	2276.40	5673.30	8977.40	15037.00	19933.00	27637.00	40674.00	46957.00
	$E_2(SP)$	**	**	**	**	110.02	154.88	213.46	275.13	360.96
0.9	$E_2(S)$	416.00	1502.60	3846.80	5988.20	10221.00	13801.00	18936.00	26671.00	32009.00
	$E_2(SP)$	**	**	**	**	**	106.09	142.50	186.15	244.08

** denotes estimator $T_{|p=2}$ does not perform better than T_{SP} in terms of efficiency.

Table 4: Monte Carlo Simulation results when the proposed estimator $T_{|p=3}$ are compared to T_S and T_{SP} respectively (considering $\psi = \xi$)

φ		ψ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	$E_3(S)$	1407.70	6252.00	12967.00	18284.00	38857.00	61313.00	76382.00	89262.00	88716.00
	$E_3(SP)$	110.12	108.23	134.48	181.31	278.58	403.21	551.10	672.22	914.40
0.3	$E_3(S)$	1222.20	5606.40	12300.00	17513.00	34076.00	47645.00	63630.00	85846.00	10116.00
	$E_3(SP)$	104.40	104.34	118.44	166.82	248.67	344.79	488.95	605.36	794.41
0.5	$E_3(S)$	934.23	3549.50	8812.80	13413.00	23967.00	32399.00	43198.00	63448.00	73073.00
	$E_3(SP)$	**	**	**	118.90	128.23	242.82	335.81	429.14	563.01
0.7	$E_3(S)$	638.63	2371.60	5891.40	9397.20	15675.00	20776.00	28770.00	42331.00	490.21
	$E_3(SP)$	**	**	**	**	114.69	161.42	222.22	286.35	376.83
0.9	$E_3(S)$	436.36	1577.90	4023.40	6288.20	10711.00	14474.00	19833.00	27952.00	33593.00
	$E_3(SP)$	**	**	**	**	**	111.27	149.26	195.09	256.16

** denotes estimator $T_{|p=3}$ does not perform better than T_{SP} in terms of efficiency.

(ii) Calculate sample mean $\bar{x}_{n|k}$, $\bar{z}_{1_{n|k}}$, $\bar{z}_{2_{n|k}}$ and $\bar{z}_{3_{n|k}}$ for $k = 1, 2, \ldots, 5000$.

(iii) Retain $m = 17$ units out of each $n = 20$ sample units of the study and auxiliary variables at the first occasion.

(iv) Calculate sample mean $\bar{x}_{m|k}$, $\bar{z}_{1_{m|k}}$, $z_{2_{m|k}}$ and $\bar{z}_{3_{m|k}}$ for $k = 1, 2, \ldots, 5000$.

(v) Select $u = 3$ units using simple random sampling without replacement from $N - n = 31$ units of the population for study and auxiliary variables at the second (current) occasion.

(vi) Calculate sample mean $\bar{y}_{u|k}$, $\bar{y}_{m|k}$, $\bar{z}_{1_{u|k}}$, $z_{2_{u|k}}$ and $\bar{z}_{3_{u|k}}$ for $k = 1, 2, \ldots, 5000$.

(vii) Iterate the parameter φ from 0.1 to 0.9 with a step of 0.1.

(viii) Iterate ψ from 0.1 to 0.9 with a step of 0.1 within (vii).

(ix) Calculate the percent relative efficiencies of the proposed estimator $T_{|p}(p = 1, 2$ and 3) with the case $p = 1$, $p = 2$ and $p = 3$ (i.e. $T_{|p=1}$, $T_{|p=2}$ and $T_{|p=3}$) with respect to estimator due to Singh

Table 5: For $p = 1$

ρ_{yz_1}	ρ_{yx}														
	0.5					0.6					0.7				
	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$
0.1	0.51	0.57	0.53	141.36	**	0.54	0.60	0.55	142.23	**	0.58	0.63	0.58	143.34	**
0.3	0.49	0.54	0.52	134.21	**	0.52	0.56	0.54	134.93	**	0.55	0.60	0.57	135.86	**
0.6	0.44	0.47	0.49	116.35	**	0.47	0.50	0.51	116.69	**	0.51	0.53	0.53	117.12	**

** denotes estimator $T_{|p=1}$ does not perform better than T_{SP} in terms of efficiency.

Table 6: For $p = 2$ and $\rho_{z_1z_2} = 0$

ρ_{yz_1}	ρ_{yz_2}	ρ_{yx}														
		0.5					0.6					0.7				
		μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$
	0.3	0.48	0.55	0.53	158.92	100.00	0.51	0.58	0.55	160.26	102.09	0.54	0.62	0.57	161.99	104.06
0.2	0.4	0.47	0.55	0.53	168.30	105.91	0.50	0.58	0.55	169.91	108.23	0.53	0.62	0.57	171.98	110.48
	0.5	0.46	0.55	0.53	180.81	113.78	0.48	0.58	0.55	182.79	116.44	0.52	0.62	0.57	185.35	119.07
	0.3	0.47	0.54	0.52	149.03	**	0.50	0.56	0.54	150.13	101.65	0.53	0.60	0.57	151.54	103.74
0.3	0.4	0.46	0.54	0.52	156.62	104.62	0.49	0.56	0.54	158.10	107.05	0.53	060	0.57	159.79	109.38
	0.5	0.45	0.54	0.52	167.23	111.58	0.48	0.56	0.54	168.83	114.32	0.52	0.60	0.57	170.91	117.00
	0.3	0.46	0.52	0.52	139.32	**	0.49	0.55	0.53	140.19	100.22	0.53	0.58	0.56	141.30	102.44
0.4	0.4	0.45	0.52	0.52	145.60	102.27	0.48	0.55	0.53	146.63	104.83	0.52	0.58	0.56	147.95	107.27
	0.5	0.45	0.52	0.52	154.20	108.31	0.47	0.55	0.53	155.46	111.14	0.51	0.58	0.56	157.09	113.89

** denotes estimator $T_{|p=2}$ does not perform better than T_{SP} in terms of efficiency.

Table 7: For $p = 2$ and $\rho_{z_1z_2} \neq 0$

ρ_{yz_1} $= \rho_{z_1z_2}$	ρ_{yz_2}	ρ_{yx}														
		0.5					0.6					0.7				
		μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$	μ_1^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^{SP}_{T_{|p}}$
0.2	0.4	0.47	0.55	0.53	162.28	104.01	0.50	0.58	0.58	166.81	106.26	0.53	0.62	0.57	168.77	108.41
	0.5	0.46	0.55	0.53	178.37	112.24	0.49	0.58	0.58	180.27	114.83	0.52	0.62	0.57	182.73	117.38
0.3	0.4	0.47	0.54	0.52	151.70	101.22	0.50	0.56	0.54	152.87	103.51	0.53	0.60	0.57	154.38	105.68
	0.5	0.46	0.54	0.52	162.40	108.36	0.48	0.56	0.54	163.86	110.95	0.52	0.60	0.57	165.75	113.47
0.4	0.4	0.46	0.52	0.52	138.66	**	0.49	0.55	0.53	139.51	**	0.53	0.58	0.56	140.60	101.94
	0.5	0.45	0.52	0.52	146.88	103.13	0.48	0.55	0.53	147.89	100.73	0.52	0.58	0.56	149.26	108.21

** denotes estimator $T_{|p=2}$ does not perform better than T_{SP} in terms of efficiency.

(2005) and Singh and Priyanka (2008) as

$$E_p(S) = \frac{\sum_{k=1}^{5000}\left[T_{S|k} - \bar{Y}\right]^2}{\sum_{k=1}^{5000}\left[T_{|p|k} - \bar{Y}\right]^2} \times 100 \quad \text{and} \quad E_p(SP) = \frac{\sum_{k=1}^{5000}\left[T_{SP|k} - \bar{Y}\right]^2}{\sum_{k=1}^{5000}\left[T_{|p|k} - \bar{Y}\right]^2} \times 100,$$

for $k = 1, 2, \ldots, 5000$.

11.2. Performance of proposed estimators for various choices of correlation coefficients

To validate the applicability and performance of the proposed estimator, $T_{|p}$ (for $p = 1, 2$ and 3) has been compared with Singh (2005) and Singh and Priyanka (2008) at their respective optimum conditions for various combinations of correlation coefficients (Tables 5–7).

Table 8: For $p = 3$ and $\rho_{z_i z_j} = 0 \ \forall \ i \neq j = 1, 2, 3$

ρ_{yx}	ρ_{yz_1}	ρ_{yz_2}	ρ_{yz_3}	μ_3^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^S_{T_{|p}}$
0.6	0.5	0.6	0.7	0.45	0.52	0.52	152.12	114.09
0.7	0.5	0.6	0.7	0.49	0.56	0.55	153.67	117.30
0.7	0.4	0.6	0.5	0.50	0.58	0.56	169.65	122.99

Table 9: For $p = 3$ and $\rho_{z_i z_j} \neq 0 \ \forall \ i \neq j = 1, 2, 3$

ρ_{yx}	ρ_{yz_1}	ρ_{yz_2}	ρ_{yz_3}	$\rho_{z_1 z_2}$	$\rho_{z_1 z_3}$	$\rho_{z_2 z_3}$	μ_3^*	μ_S	μ_{SP}	$E^S_{T_{|p}}$	$E^S_{T_{|p}}$
0.6	0.2	0.5	0.3	0.5	0.5	0.4	0.46	0.55	0.53	177.71	111.83
0.7	0.4	0.7	0.6	0.3	0.3	0.4	0.48	0.58	0.56	182.62	132.40
0.7	0.3	0.7	0.5	0.3	0.4	0.2	0.48	0.60	0.57	206.51	141.37
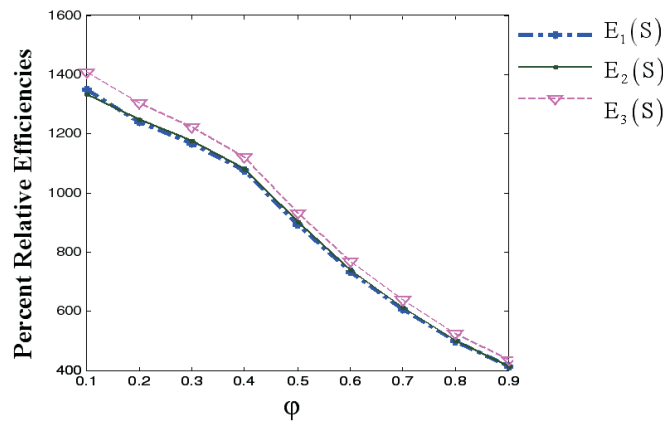
Figure 1: *Mutual comparison of proposed estimator $T_{|p=1}$, $T_{|p=2}$ and $T_{|p=3}$ with respect to the estimator T_S for $\psi = 0.1$.*

12. Mutual Comparison of the Estimators $T_{|p}$ ($p = 1, 2$ and 3)

The performances of the estimator $T_{|p}$ ($p = 1, 2$ and 3) have been elaborated empirically as well as through simulation studies in the above sections and obtained results (Table 1–9). In this section the mutual comparison of the estimators for the cases when $p = 1$, $p = 2$ and $p = 3$ has been elaborated graphically (Figures 1 and 2).

13. Rendition of Results

13.1. Results based on an empirical study for the considered population

(1) Table 1 indicates that the optimum values of μ_1^*, μ_2^* and μ_3^* exist for the considered population and $\mu_3^* < \mu_2^* < \mu_1^* < \mu_{SP} < \mu_S$. This indicates that a smaller fraction of fresh sample is required when more numbers of auxiliary variables is used and this fraction is lesser than the procedures given by Singh (2005) and Singh and Priyanka (2008). Hence, the total cost of the survey is reduced.

(2) The value of $E_{T|p=3} > E_{T|p=2} > E_{T|p=1}$ and justifies that efficiency increases highly when more numbers of auxiliary variates are considered and it results peachy in terms of cost as it decreases on increasing the number of auxiliary variables which also abide by Sukhatme (1984).

(3) Table 1 shows that the proposed estimator $T_{|p}$ is more efficient than the estimator T_S for all the
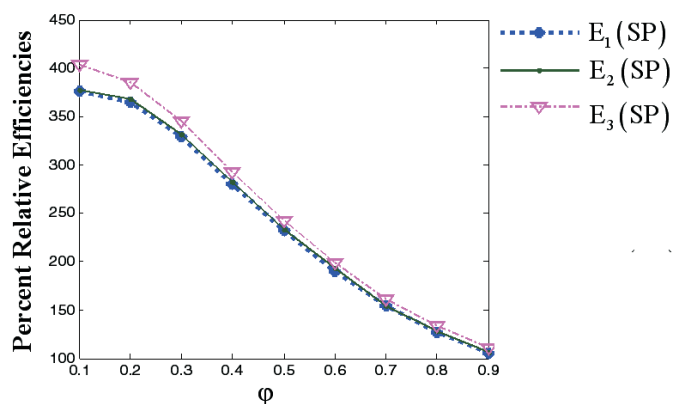
Figure 2: *Mutual comparison of proposed estimator $T_{|p=1}$, $T_{|p=2}$ and $T_{|p=3}$ with respect to the estimator T_{SP} for $\psi = 0.6$.*

considered values of p (i.e. $p = 1, 2$ and 3) and $T_{|p}$ is better than the estimator T_{SP} except for $p = 1$ as the number of auxiliary variables is increased, efficiency increases considerably, hence the estimator $T_{|p}$ is superior than the estimator T_{SP} due to Singh and Priyanka (2008) for $p = 2$ onwards in terms of efficiency but in terms of cost $T_{|p}$ is better than T_S and T_{SP} for every value of p.

13.2. Results based on simulation study

(1) From simulation results in Tables 2–4 we observe that if less attention is given to φ (i.e. more attention is given to the estimator used at the first occasion) then the proposed estimators $T_{|p}$ ($p = 1, 2$ and 3) are superior than the estimator T_S and abide by the theory while retaining higher weights for ψ (i.e. more attention is given to the estimator used at the current occasion). This makes the proposed estimators $T_{|p}$ ($p = 1, 2$ and 3) more effective than the estimator T_S due to Singh (2005).

(2) From Tables 2–4 it is vindicated that if less emphasis is supplied to φ (i.e. more attention is given to the estimator used at the first occasion) then the proposed estimators $T_{|p}$ ($p = 1, 2$ and 3) are better than the estimator T_{SP} and are in accordance with the theory while selecting a greater value for ξ (i.e. more attention is given to the estimator used at the current occasion). This makes the proposed estimators $T_{|p}$ ($p = 1, 2$ and 3) more effective than the estimator T_{SP} due to Singh and Priyanka (2008). We increase the value of φ, and the efficiency gets reduced for all choices of ξ.

13.3. Results extracted from general scenario i.e. by considering different choices of correlation coefficients

(1) Table 5 indicates that for a fixed value of correlation coefficient between the study variable at two occasions, if the correlation between the study and auxiliary variates increases then the proposed estimator $T_{|p}$ for $p = 1$ is efficient over the estimator T_S in terms of precision as well as cost but it is efficient over the estimator T_{SP} only in terms of cost. If the contribution of auxiliary variable increases the fraction of sample to be drawn on current occasion decreases.

(2) Tables 6–9 show that the proposed estimator $T_{|p}$ for $p = 2$ and $p = 3$ is efficient over the estimators T_S and T_{SP} even for very low correlation between study variable and auxiliary variable, which is a

positive point whether the auxiliary information utilized are mutually correlated or uncorrelated. The fraction of samples to be drawn fresh at the current occasion is least for the proposed estimator than the estimators due to Singh (2005) and Singh and Priyanka (2008); in addition, it reduces more as the contribution of auxiliary information increases.

14. Conclusion

The articulation of two structures (i.e. exponential ratio type and chain type ratio to exponential ratio type) is beneficial as summed with multi-auxiliary information which is stable in nature, pronto and does not to be highly correlated to study variables over the two occasions.The empirical study for the considered population, simulation study and the study by taking different choices of correlation coefficient suggest that the proposed estimator provides the lowest fraction of fresh sample drawn on the current occasion compared to well-known estimators available in the literature to estimate population mean, resulting in lowering the total cost of the survey. The proposed estimator $T_{|p}$ is better than the estimator T_{SP} for $p = 1$ in terms of cost only. As we increase the number of auxiliary variables, the proposed estimator becomes better than the estimator T_{SP} in terms of efficiency and cost. Hence, the proposed estimator may be recommended for its practical use by survey statisticians.

Acknowledgement

References

Artes Rodriguez, E. M. and Gracia Luengo, A. V. (2005). Multivariate indirect methods of estimation in successive sampling, *Journal of the Indian Society of Agricultural Statistics*, **59**, 97–103.

Biradar, R. S. and Singh, H. P. (2001). Successive sampling using auxiliary information on both the occasions, *Bulletin of the Calcutta Statistical Association*, **51**, 234–251.

Chaturvedi, D. K. and Tripathi, T. P. (1983). Estimation of population ratio on two occasions using multivariate auxiliary information, *Journal of Indian Statistical Association*, **21**, 113–120.

Cochran, W. G. (1977). *Sampling Techniques*, John Wiley & Sons, New York.

Das, A. K. (1982). Estimation of population ratio on two occasions, *Journal of the Indian Society of Agricultural Statistics*, **34**, 1–9.

Feng, S. and Zou, G. (1997). Sample rotation method with auxiliary variable, *Communication in Statistics - Theory and Methods*, **26**, 1497–1509.

Gupta, P. C. (1979). Sampling on two successive occasions, *Journal of Statistical Research*, **13**, 7–16.

Jessen, R. J. (1942). Statistical investigation of a sample survey for obtaining farm facts, *Iowa State College Research Bulletin*, **304**, 1–100.

Olkin, I. (1958). Multivariate ratio estimation for finite populations, *Biometrika*, **45**, 154–165.

Patterson, H. D. (1950). Sampling on successive occasions with partial replacement of units, *Journal of the Royal Statistical Society Series B (Methodological)*, **12**, 241–255.

Raj, D. (1965). On a method of using multi-auxiliary information in sample surveys, *Journal of the American Statistical Association*, **60**, 270–277.

Rao, C. R. (2006). *Linear Statistical Inference and Its Application (Second Edition)*, John Wiley and Sons (Asia), Singapore.

Rao, J. N. K. and Graham, J. E. (1964). Rotation designs for sampling on repeated occasions, *Journal of the American Statistical Association*, **59**, 492–509.

Sen, A. R. (1972). Successive sampling with p ($p \geq 1$) auxiliary variables, *Annals of Mathematical Statistics*, **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. (2005). On the use of chain-type ratio estimator in successive sampling, *Statistics in Transition*, **7**, 21–26.

Singh, G. N. and Karna, J. P. (2009). Search of efficient rotation patterns in presence of auxiliary information in successive sampling over two occasions, *Statistics in Transition New Series*, **10**, 59–73.

Singh, G. N. and Prasad, S. (2010). Some estimators of population mean in two-occasion rotation patterns, *Association of the Advanced Modeling & Simulation Techniques Enterprises*, **12**, 25–44.

Singh, G. N. and Priyanka, K. (2006). On the use of chain-type ratio to difference estimator in successive sampling, *International Journal of Applied Mathematics and Statistics*, **5**, 41–49.

Singh, G. N. and Priyanka, K. (2007). On the use of auxiliary information in search of good rotation patterns on successive occasions, *International Journal of Statistics & Economics*, **1**, 42–60.

Singh, G. N. and Priyanka, K. (2008). Search of good rotation patterns to improve the precision of the estimates at current occasion, *Communication in Statistics-Theory and Methods*, **37**, 337–348.

Singh, G. N. and Singh, V. K. (2001). On the use of auxiliary information in successive sampling, *Journal of Indian Society of Agricultural Statistics*, **54**, 1–12.

Singh, H. P., Kumar, S. and Bhougal, S. (2011). Multivariate ratio estimation in presence of non-response in successive sampling, *Journal of Statistical Theory and Practice*, **5**, 591–611.

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*, **41C**, 149–159.

Sukhatme, P. V. (1984). *Sampling Theory of Surveys with Applications*, Iowa State University Press and New Delhi India, Indian Society of Agricultural Statistics, Ames.