Multivariate Rotation Design for Population Mean in Sampling on Successive Occasions

Kumari Priyanka\textsuperscript{a}, Richa Mittal\textsuperscript{a}, Jong-Min Kim\textsuperscript{1, b}

\textsuperscript{a}Department of Mathematics, Shivaji College (University of Delhi), India; \textsuperscript{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

Authors are also thankful to UGC, New Delhi, India for providing financial assistance to carry out the present work.

\textsuperscript{1}Corresponding author: Statistics, Division of Science and Mathematics, University of Minnesota-Morris, 2380 Science Building, MN 56267, Morris, USA. E-mail: jongmink@morris.umn.edu

Published 30 September 2015 / journal homepage: http://csam.or.kr
© 2015 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.
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 \( \mu \) and \( \lambda (\mu + \lambda = 1; 0 \leq \mu, \lambda \leq 1) \) are the fractions of fresh and matched samples respectively at the second (current) occasion.

3. Formulation of the Proposed Estimator \( T_{lp} \)

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

\[
T_u = W_u T_{\exp}(u),
\]  

(3.1)
where

\[ W_u \] is a column vector of \( p \)-weights given by

\[ W_u = [w_{u_1}, w_{u_2}, \ldots, w_{u_p}]' \]

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

such that \( 1'W_u = 1 \), where \( 1 \) 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 \( m = n_1n_2 \) common to both occasions and given by

\[ T_m = W_mT_{\exp}(m, n), \quad (3.2) \]

where

\[ W_m \] is a column vector of \( p \)-weights as

\[ W_m = [w_{m_1}, w_{m_2}, \ldots, w_{m_p}]' \]

\[ T_{\exp}(m, n) = \begin{bmatrix} T(1, m, n) \\ T(2, m, n) \\ \vdots \\ T(p, m, n) \end{bmatrix}, \] where \( T(i, m, n) = (\tilde{y}^*(i, m)/\tilde{x}^*(i, m))\tilde{x}^*(i, n) \),

where

\[ \tilde{y}^*(i, m) = \tilde{y}_m \exp \left( \frac{\tilde{Z}_i - \tilde{z}_i(m)}{\tilde{Z}_i + \tilde{z}_i(m)} \right), \]

\[ \tilde{x}^*(i, m) = \tilde{x}_m \exp \left( \frac{\tilde{Z}_i - \tilde{z}_i(m)}{\tilde{Z}_i + \tilde{z}_i(m)} \right), \]

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

Such that \( 1'W_m = 1 \), where \( 1 \) 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

\[ T_{\varphi} = \varphi T_u + (1 - \varphi)T_m, \quad (3.3) \]

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 \( T_{\varphi} \).

**4. Properties of the Proposed Estimator** \( T_{\varphi} \)

The properties of the proposed estimator \( T_{\varphi} \) are derived under following large sample approximations:

\[ \tilde{y}_u = \tilde{Y}(1 + e_0), \tilde{y}_m = \tilde{Y}(1 + e_1), \tilde{x}_u = \tilde{X}(1 + e_2), \tilde{x}_u = \tilde{X}(1 + e_3), \tilde{z}_i(u) = \tilde{Z}_i(1 + e_4), \tilde{z}_i(m) = \tilde{Z}_i(1 + e_5) \]

and \( \tilde{z}_i(n) = \tilde{Z}_i(1 + e_6) \) such that \(|e_k| < 1 \ \forall \ k = 0, 1, 2, 3, 4, 5 \) and \( 6 \) and \(|e_k| < 1 \ \forall \ i = 1, 2, 3, \ldots, p \).
Under the above transformations, the estimators $T_u$ and $T_m$ take the following forms:

\[
T(i, u) = \frac{\hat{y}}{8} \left( 8 + 8e_0 - 4e_d - 4e_0e_d + 3e_d^2 \right), \quad \text{for } i = 1, 2, \ldots, p, \quad (4.1)
\]

\[
T(i, m, n) = \frac{\hat{y}}{8} \left( 8 + 8e_1 - 8e_2 + 8e_1 - 4e_d - 8e_1e_d + 8e_1e_3 - 4e_1e_d - 8e_2e_3 + 4e_2e_d - 4e_3e_d + 8e_1^2 + 3e_d^2 \right), \quad \text{for } i = 1, 2, \ldots, p. \quad (4.2)
\]

Thus we have the following theorems:

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

\[
B(T_{ip}) = \varphi B(T_u) + (1 - \varphi)B(T_m), \quad (4.3)
\]

\[
B(T_u) = \frac{1}{u} W^r UB_u, \quad (4.4)
\]

\[
B(T_m) = W^r_m \left( \frac{1}{m} B_{m1} + \frac{1}{n} B_{m2} \right), \quad (4.5)
\]

where

\[
B_u = (B_1(u), B_2(u), \ldots, B_p(u))^\prime, \quad B_i(u) = \frac{1}{u} \hat{y} \left( \frac{3 C_{002}}{8 Z_i^2} - \frac{1 C_{011}}{2 Y_i Z_i} \right), \quad \text{for } i = 1, 2, 3, \ldots, p.
\]

\[
B_{m1} = \hat{y} \left( \frac{C_{200}}{X^2} - \frac{C_{110}}{XY} \right),
\]

\[
B_{m2} = \left( B_{m21}, B_{m22}, \ldots, B_{m2p} \right),
\]

where

\[
B_{m2i} = \hat{y} \left( \frac{3 C_{002}}{8 Z_i^2} + \frac{C_{110}}{2 Y_i Z_i} \right) - \frac{1}{2} \frac{C_{011}}{Y_i Z_i} - \frac{C_{200}}{X^2}, \quad C_{est} = E \left[ (x_i - \bar{x})^\prime (y_i - \bar{y})^\prime (z_i - \bar{z})^\prime \right];
\]

\[(r, s, t) \geq 0 \text{ for } i = 1, 2, 3, \ldots, p.\]

**Proof:** The bias of the estimator $T_{ip}$ is given by

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

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 $T_{ip}$ is obtained as in equation (4.3). □

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

\[
M(T_{ip}) = \varphi^2 M(T_u) + (1 - \varphi^2) M(T_m) + 2\varphi (1 - \varphi) \text{cov}(T_u, T_m), \quad (4.6)
\]

\[
M(T_u) = W^r_u K_u W_u, \quad (4.7)
\]

\[
M(T_m) = (B) W^r_m K_m W_m + W^r_m K_u W_u. \quad (4.8)
\]
where
\[ W_u = \begin{bmatrix} W_{u_1} & W_{u_2} & \cdots & W_{u_p} \end{bmatrix}, \quad W_m = \begin{bmatrix} W_{m_1} & W_{m_2} & \cdots & W_{m_p} \end{bmatrix}, \]

\( E \) is a unit matrix of order \( p \times p \),

\[ B = \left( \frac{1}{m} - \frac{1}{N} \right) B_1, \quad B_1 = 2F^2(1 - \rho_{xy})C_0^2, \]
\[ K_u = \frac{1}{u - 1} K_u', \quad K_m = \frac{1}{u - 1} K_m', \]

where
\[
K_u' = \begin{bmatrix}
ku_{11} & ku_{12} & \cdots & ku_{1p} \\
ku_{21} & ku_{22} & \cdots & ku_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
ku_{p1} & ku_{p2} & \cdots & ku_{pp}
\end{bmatrix}_{p \times p}
\]
and
\[
K_m' = \begin{bmatrix}
km_{11} & km_{12} & \cdots & km_{1p} \\
km_{21} & km_{22} & \cdots & km_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
km_{p1} & km_{p2} & \cdots & km_{pp}
\end{bmatrix}_{p \times p},
\]

where
\[
ku_{ij} = \hat{Y}^2 \left( C_0^2 + \frac{1}{4} C_2^2 - \rho_{yz} C_0 C_2 \right),
\]
\[ km_{ij} = \hat{Y}^2 \left( C_0^2 \frac{1}{2} \rho_{yz} C_0 C_2 - \frac{1}{2} \rho_{yz} C_0 C_{zj} + \frac{1}{4} \rho_{yz} C_2 C_{zj} \right), \]
\[ km_{ij} = \hat{Y}^2 \left( C_0^2 (2\rho_{yz} - 1) - \rho_{yz} C_0 C_{zj} + \frac{1}{4} C_2^2 \right), \]
\[ km_{ij} = \hat{Y}^2 \left( C_0^2 (2\rho_{yz} - 1) - \frac{1}{2} \rho_{yz} C_0 C_{zj} - \frac{1}{2} \rho_{yz} C_0 C_{zj} + \frac{1}{4} \rho_{yz} C_2 C_{zj} \right) \quad \forall i \neq j = 1, 2, 3, \ldots, p. \]

**Proof:** The mean square error of the estimator \( T_{lp} \) is given by
\[
M(T_{lp}) = E \left[ T_{lp} - \hat{Y} \right]^2
= E \left[ \varphi (T_u - \hat{Y}) + (1 - \varphi) (T_m - \hat{Y}) \right]^2
= \varphi^2 M(T_u) + (1 - \varphi)^2 M(T_m) + 2\varphi(1 - \varphi) \text{cov}(T_u, T_m),
\]
where \( M(T_u) = E[T_u - \hat{Y}]^2 \) and \( M(T_m) = E[T_m - \hat{Y}]^2 \).

The estimators \( T_u \) and \( T_m \) are based on two independent samples of sizes \( u \) and \( m \) respectively, hence \( \text{cov}(T_u, T_m) = 0 \), considering the population is sufficiently large so using large sample approximations assumed in Section 4 and retaining terms up to the first order of approximations and also assuming \( C_z = C_y = C_2 \) (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 \( T_{lp} \) is obtained as in equation (4.6).

5. **Choice of Optimal Weights**

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

$$L_u = W_u'K_uW_u - \lambda_u (1'W_u - 1),$$  \hfill (5.1)

where $1$ 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

$$\frac{\partial L_u}{\partial W_u} = \frac{\partial}{\partial W_u} [W_u'K_uW_u - \lambda_u (1'W_u - 1)] = 0.$$

This implies that, $2K_uW_u - \lambda_u 1 = 0$, which yields

$$W_u = \frac{\lambda_u}{2} K_u^{-1} 1.$$  \hfill (5.2)

Now pre-multiplying equation (5.2) by $1'$, we get

$$\frac{\lambda_u}{2} = \frac{1}{1'K_u^{-1} 1}.$$  \hfill (5.3)

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

$$W_{u_{opt}} = \frac{K_u^{-1}}{1'K_u^{-1} 1}.$$  \hfill (5.4)

In similar manners, the optimal of the weight $W_m = [w_{m_1}, w_{m_2}, \ldots, w_{m_p}]'$ 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

$$L_m = (B)W_m'E W_m + W_m'K_m W_m - \lambda_m (1'W_m - 1),$$

where $\lambda_m$ is the Lagrange multiplier.

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

$$W_{m_{opt}} = \frac{K_m^{-1}}{1'K_m^{-1} 1}.$$  \hfill (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:

$$M(T_u)_{opt.} = \left(1 - \frac{1}{N}\right) \frac{1}{1'K_u^{-1} 1},$$  \hfill (5.6)

$$M(T_m)_{opt.} = \left(1 - \frac{1}{N}\right) B_1 + \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{1'K_m^{-1} 1}. \hfill (5.7)$$

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

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

$$M(T_{ip}) = \varphi^2 M(T_u)_{opt.} + (1 - \varphi)^2 M(T_m)_{opt.}.$$
Minimizing $M(T_p)$ with respect to $\varphi$ gives the optimum value of $\varphi$ as

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

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

$$M(T_p)_{\text{opt.}} = \frac{M(T_u)_{\text{opt.}}M(T_m)_{\text{opt.}}}{M(T_u)_{\text{opt.}} + M(T_m)_{\text{opt.}}}.$$ \hfill (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 $\varphi_{\text{opt.}}$ and $M(T_p)_{\text{opt.}}$ are obtained as

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

$$M(T_p)_{\text{opt.}} = \frac{1}{n} \left[ \frac{[\mu D_1 - D_2]}{\mu^2 C - \mu D_1 - A} \right],$$ \hfill (6.4)

where $A = 1/(Y'K_p^{-1}Y)$, $B_1 = 2\bar{Y}_1(1 - \rho_{\varphi_1})C_0$, $C = 1/(Y'K_p^{-1}Y)$, $D_1 = AC$, $D_2 = AB_1 + AC$, $D_3 = B_1 + C - A$ and $\mu$ 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 $M(T_p)_{\text{opt.}}$ in equation (6.4) with respect to $\mu$ as:

$$\frac{\partial M(T_p)_{\text{opt.}}}{\partial \mu} = 0, \quad \Rightarrow \quad \mu^2 G_1 - 2\mu G_2 + G_3 = 0.$$  

Thus the optimum value of $\mu$ so obtained is one of the two roots given by

$$\hat{\mu} = \frac{G_2 \pm \sqrt{G_2^2 - G_1 G_3}}{G_1},$$ \hfill (7.1)

where $G_1 = CD_1$, $G_2 = CD_2$ and $G_3 = AD_1 + D_2 D_3$.

The real value of $\hat{\mu}$ exist, if $G_2^2 - G_1 G_3 \geq 0$. 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 $\mu_{\text{opt.}}$ 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 $\varphi$ as well as $\mu$ which, is given as

$$M(T_p)_{\text{opt.}} = \frac{1}{n} \left[ \frac{[\mu_{\text{opt.}} D_1 - D_2]}{\mu_{\text{opt.}}^2 C - \mu_{\text{opt.}} D_3 - A} \right].$$ \hfill (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 \( M(T_p) \) ≥ \( M(T_q) \), i.e.

\[
M\left( T_p \right) - M\left( T_q \right) \geq 0
\]

\[
\frac{1}{n} \begin{bmatrix} \mu A_p C_p - A_p (B + C_p) \\ \mu^2 C_p - \mu (B + C_p + A_p) - A_p \end{bmatrix} - \frac{1}{n} \begin{bmatrix} \mu A_q C_q - A_q (B + C_q) \\ \mu^2 C_q - \mu (B + C_q + A_q) - A_q \end{bmatrix} \geq 0.
\]

On simplification, we get

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

This reduces to the condition

\[
(A_p - A_q) \geq 0.
\]

(8.2)

From Section 6 above, we get

\[
\frac{1}{1' K_p^{-1} 1} - \frac{1}{1' K_q^{-1} 1} \geq 0
\]

\[
1' K_q^{-1} 1 \geq 1' K_p^{-1} 1.
\]

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

\[
K_q = \begin{bmatrix} K_p & F' \\ F & G \end{bmatrix},
\]

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,

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

(8.3)

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

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

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

\[
= \begin{bmatrix} 1_p (K_p^{-1} + HJH') - 1_{q-p} JH' - 1_{q-p} HJ + 1_{q-p} J \end{bmatrix} \begin{bmatrix} 1_p \\ 1_{q-p} \end{bmatrix}
\]

\[
= 1_p (K_p^{-1} + HJH') 1_p - 1_{q-p} JH' 1_{q-p} - 1_{q-p} HJ 1_{q-p} + 1_{q-p} J 1_{q-p}.
\]
implies

\[1'K_p^{-1}1_1'(K_p^{-1})_1p = 1_p'HH'H1_p = 1_p'H1_p', 1_p'H1_p + 1_p'HJ1_q1_p',
\]

\[1'K_p^{-1}1_1'(K_p^{-1})_1p = [1_p 1_q] [-HH' -JH'] [1_p 1_q] = 1_p'H1_p + 1_p'HJ1_q1_p',
\]

\[1'K_p^{-1}1_1'(K_p^{-1})_1p = 1_p'H1_p + 1_p'HJ1_q1_p'
\]

The latter follows since \(J\) is positive definite so that \(R'JR \geq 0\) for all \(R\), where \(R = \begin{bmatrix} H & -I \end{bmatrix} 1\).

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. \(\rho_{ij} = 0 \forall i \neq j = 1, 2, \ldots, p\), the proposed multivariate exponential ratio type estimator is applicable and the optimum value of \(\mu\) say \(\hat{\mu}^0\) and the optimum value of the mean square error of the estimator \(T_p\) with respect to \(\varphi\) as well as \(\hat{\mu}^0\) is given by

\[\hat{\mu}^0 = \frac{G_2^0 + \sqrt{G_2^0 - G_1^0 G_3^0}}{G_1^0} \quad (9.1)\]

and

\[M(T_p)_{opt} = \frac{1}{n} \left[ \frac{\mu^0_{2p} D_1^0 - D_2^0}{\mu^0_{2p} C^0 - \mu^0_{1p} D_3^0 - A^0} \right] \quad (9.2)\]

where

\[G_1^0 = C^0 D_1^0, \quad G_2^0 = C^0 D_2^0, \quad G_3^0 = A^0 D_1^0 + D_2^0 D_3^0, \quad A^0 = \frac{1}{1'S^0_{\varphi} 1}, \quad C^0 = \frac{1}{1'S^0_{\varphi} 1}, \]

\[B_1^0 = 2\hat{y}^2(1 - \rho_{ij}) C^0_0, \quad D_1^0 = A^0 C^0, \quad D_2^0 = A^0 B_1^0 + A^0 C^0, \quad D_3^0 = B_1^0 + C^0 - A^0.\]
where

\[ S_{p'} = \begin{bmatrix} su_{11} & su_{12} & \ldots & su_{1p} \\ su_{21} & su_{22} & \ldots & su_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ su_{p1} & su_{p2} & \ldots & su_{pp} \end{bmatrix}_{p \times p} \text{ and } S_m = \begin{bmatrix} sm_{11} & sm_{12} & \ldots & sm_{1p} \\ sm_{21} & sm_{22} & \ldots & sm_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ sm_{p1} & sm_{p2} & \ldots & sm_{pp} \end{bmatrix}_{p \times p}, \]

where

\[ su_{ii} = \frac{s^2}{G^2_1} \left( C_0^2 + \frac{1}{4} C_i^2 - \rho_{y_i z_i} C_0 C_i \right), \quad su_{ij} = \frac{s^2}{G^2_1} \left( C_0^2 - \frac{1}{2} \rho_{y_i z_i} C_0 C_i - \frac{1}{4} \rho_{y_j z_j} C_0 C_j \right), \]

\[ sm_{ii} = \frac{s^2}{G^2_1} \left( C_0^2 (2 \rho_{y_i z_i} - 1) - \rho_{y_i z_i} C_0 C_i + \frac{1}{4} C_i^2 \right), \quad sm_{ij} = \frac{s^2}{G^2_1} \left( C_0^2 (2 \rho_{y_i z_i} - 1) - \rho_{y_j z_j} C_0 C_j - \frac{1}{2} \rho_{y_j z_j} C_0 C_j \right), \]

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

Case 2

The \( p \)-auxiliary variates are mutually correlated i.e. \( \rho_{z_i z_j} = 0 \) for \( i \neq j = 1, 2, \ldots, p \). 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 \( C_0 = C_{ij}; \forall i = 1, 2, 3, \ldots, p \) approximately and hence, the optimum value of \( \mu \) say \( \hat{\mu}^* \) and optimum value of mean square error \( M(T_{ip})_{opt}^* \) of the proposed estimator \( T_{ip} \) reduces to

\[
\hat{\mu}^* = \frac{G_1^* \pm \sqrt{G_1^*} - G_1^* \bar{G}_1^*}{G_1^*}, \quad (10.1)
\]

\[
M(T_{ip})_{opt}^* = \frac{1}{n} \left[ \mu_{ip}^* D_1^* - D_2^* \right], \quad (10.2)
\]

where

\[
G_1^* = C^* D_1^*, \quad G_2^* = C^* D_2^*, \quad G_3^* = A^* D_1^* + D_2^* D_3^*, \quad A^* = \frac{1}{1H_{opt}^* 1}, \quad C^* = \frac{1}{1H_{opt}^* 1},
\]

\[
B_1^* = 2(1 - \rho_{y_j z_j}) S_j^2, \quad D_1^* = A^* C^*, \quad D_2^* = A^* B_1^* + A^* C^*, \quad D_3^* = B_1^* + C^* - A^*,
\]

where

\[
H_{p'} = \begin{bmatrix} hu_{11} & hu_{12} & \ldots & hu_{1p} \\ hu_{21} & hu_{22} & \ldots & hu_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ hu_{p1} & hu_{p2} & \ldots & hu_{pp} \end{bmatrix}_{p \times p} \text{ and } H_m = \begin{bmatrix} hm_{11} & hm_{12} & \ldots & hm_{1p} \\ hm_{21} & hm_{22} & \ldots & hm_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ hm_{p1} & hm_{p2} & \ldots & hm_{pp} \end{bmatrix}_{p \times p}
\]
10.2. Comparison of the proposed estimator $T_{ip}$ with respect to estimator $T_S$ due to Singh (2005)

The estimator proposed by Singh (2005) is given as

$$T_S = \psi \bar{y}_u Z + (1 - \psi) \bar{y}_m \bar{x}_u Z,$$

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

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

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

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

$$E_{T_{ip}}^S = \frac{M(T_S)_{opt}}{M(T_{ip})_{opt}} \times 100.$$

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

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

$$T_{SP} = \xi \left[ \bar{y}_m + \beta_{yz}(Z - \bar{z}_m) \right] + (1 - \xi) \left[ \bar{y}_m + \beta_{yx}(Z - \bar{z}_m) \right],$$

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

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

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

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

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

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

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\mu^*_1 = 0.5355$</th>
<th>$\mu^*_2 = 0.5196$</th>
<th>$\mu^*_3 = 0.5137$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{TS}^p$</td>
<td>109.55</td>
<td>120.17</td>
<td>125.17</td>
</tr>
<tr>
<td>$E_{TSP}^p$</td>
<td>$\mu_{SP} = 0.5496$</td>
<td><strong>101.44</strong></td>
<td><strong>105.14</strong></td>
</tr>
</tbody>
</table>

** 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 $\psi = \xi$)

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$E_{TS}^p$</th>
<th>$E_{TSP}^p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.1$</td>
<td>1349.90</td>
<td>104.42</td>
</tr>
<tr>
<td>$0.2$</td>
<td>5895.10</td>
<td>102.05</td>
</tr>
<tr>
<td>$0.3$</td>
<td>12296.00</td>
<td>127.52</td>
</tr>
<tr>
<td>$0.4$</td>
<td>17524.00</td>
<td>171.80</td>
</tr>
<tr>
<td>$0.5$</td>
<td>36619.00</td>
<td>262.54</td>
</tr>
<tr>
<td>$0.6$</td>
<td>57157.00</td>
<td>375.88</td>
</tr>
<tr>
<td>$0.7$</td>
<td>72197.00</td>
<td>520.91</td>
</tr>
<tr>
<td>$0.8$</td>
<td>84832.00</td>
<td>638.86</td>
</tr>
<tr>
<td>$0.9$</td>
<td>84170.00</td>
<td>867.55</td>
</tr>
</tbody>
</table>

** denote 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 $i^{th}$ state of U.S.
- $x_i$: The total energy consumption during 2002 in the $i^{th}$ state of U.S.
- $z_{1i}$: The total energy consumption during 2001 in the $i^{th}$ state of U.S.
- $z_{2i}$: The total energy consumption during 2000 in the $i^{th}$ state of U.S.
- $z_{3i}$: The total energy consumption during 1999 in the $i^{th}$ state of U.S.

For the considered population, the values of $\mu^*_T (p = 1, 2$ and $3$) defined in equation (10.2) and percent relative efficiencies $E_{TS}^p$ and $E_{TSP}^p$ 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.
Calculate sample mean $\bar{\psi}$ and $\bar{\phi}$.

Select $\phi$.

Iterate the parameter $\psi$.

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

Table 3:

<table>
<thead>
<tr>
<th>$\psi$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_S(S)$</td>
<td>1335.20</td>
<td>5940.60</td>
<td>12400.00</td>
<td>1776.00</td>
<td>36883.00</td>
<td>57438.00</td>
<td>73062.00</td>
<td>85114.00</td>
<td>84732.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>104.44</td>
<td>102.84</td>
<td>128.60</td>
<td>172.31</td>
<td>264.43</td>
<td>377.72</td>
<td>527.15</td>
<td>640.98</td>
<td>873.34</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>117.80</td>
<td>11795.10</td>
<td>11935.00</td>
<td>1692.00</td>
<td>23019.00</td>
<td>31143.00</td>
<td>41964.00</td>
<td>61237.00</td>
<td>70719.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>920.63</td>
<td>3431.40</td>
<td>8553.50</td>
<td>12956.00</td>
<td>23109.00</td>
<td>31143.00</td>
<td>41964.00</td>
<td>61237.00</td>
<td>70719.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>612.36</td>
<td>2276.40</td>
<td>5673.30</td>
<td>8977.40</td>
<td>15037.00</td>
<td>19933.00</td>
<td>27637.00</td>
<td>40674.00</td>
<td>46957.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>416.00</td>
<td>1502.60</td>
<td>3846.80</td>
<td>5988.20</td>
<td>10221.00</td>
<td>13801.00</td>
<td>18936.00</td>
<td>26671.00</td>
<td>32009.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

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

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

Table 4:

<table>
<thead>
<tr>
<th>$\psi$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_S(S)$</td>
<td>1470.70</td>
<td>6252.00</td>
<td>12967.00</td>
<td>18284.00</td>
<td>38857.00</td>
<td>61313.00</td>
<td>76382.00</td>
<td>89262.00</td>
<td>88716.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>110.12</td>
<td>108.23</td>
<td>134.48</td>
<td>181.31</td>
<td>278.58</td>
<td>403.21</td>
<td>551.10</td>
<td>672.22</td>
<td>914.40</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>1222.20</td>
<td>5606.40</td>
<td>12300.00</td>
<td>17513.00</td>
<td>34076.00</td>
<td>47645.00</td>
<td>63630.00</td>
<td>85846.00</td>
<td>10116.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>934.23</td>
<td>3549.50</td>
<td>8812.80</td>
<td>13413.00</td>
<td>23967.00</td>
<td>32399.00</td>
<td>43198.00</td>
<td>63448.00</td>
<td>73073.00</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>111.27</td>
<td>108.23</td>
<td>134.48</td>
<td>181.31</td>
<td>278.58</td>
<td>403.21</td>
<td>551.10</td>
<td>672.22</td>
<td>914.40</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>$E_S(S)$</td>
<td>114.85</td>
<td>168.66</td>
<td>233.41</td>
<td>326.21</td>
<td>414.19</td>
<td>544.87</td>
<td>672.22</td>
<td>819.34</td>
<td>981.90</td>
</tr>
<tr>
<td>$E_S(SP)$</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

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

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

(vii) Iterate the parameter $\phi$ from 0.1 to 0.9 with a step of 0.1.

(viii) Iterate $\psi$ from 0.1 to 0.9 with a step of 0.1 within (vii).

(ix) Calculate the percent relative efficiencies of the proposed estimator $T_{p|=2}$ with respect to estimator due to Singh.
Table 5: For $p = 1$

<table>
<thead>
<tr>
<th>$\rho_{p1}$</th>
<th>$\rho_{p2}$</th>
<th>$\rho_{ps}$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{p1}$</td>
<td>$\rho_{ps}$</td>
<td>$\rho_{p2}$</td>
<td>$E_{yp}$</td>
<td>$E_{yp}$</td>
<td>$E_{yp}$</td>
</tr>
<tr>
<td>0.1</td>
<td>0.51</td>
<td>0.53</td>
<td>141.36</td>
<td>142.23</td>
<td>143.34</td>
</tr>
<tr>
<td>0.3</td>
<td>0.49</td>
<td>0.54</td>
<td>134.21</td>
<td>134.93</td>
<td>135.86</td>
</tr>
<tr>
<td>0.6</td>
<td>0.44</td>
<td>0.49</td>
<td>116.35</td>
<td>116.69</td>
<td>117.12</td>
</tr>
</tbody>
</table>

** denotes estimator $T_{ps1}$ does not perform better than $T_{yp}$ in terms of efficiency.

Table 6: For $p = 2$ and $\rho_{p12} = 0$

<table>
<thead>
<tr>
<th>$\rho_{p1}$</th>
<th>$\rho_{p2}$</th>
<th>$\rho_{ps}$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{p1}$</td>
<td>$\rho_{ps}$</td>
<td>$\rho_{p2}$</td>
<td>$E_{yp}$</td>
<td>$E_{yp}$</td>
<td>$E_{yp}$</td>
</tr>
<tr>
<td>0.3</td>
<td>0.48</td>
<td>0.53</td>
<td>158.92</td>
<td>160.26</td>
<td>161.99</td>
</tr>
<tr>
<td>0.2</td>
<td>0.47</td>
<td>0.53</td>
<td>168.30</td>
<td>169.91</td>
<td>171.98</td>
</tr>
<tr>
<td>0.5</td>
<td>0.56</td>
<td>0.55</td>
<td>180.81</td>
<td>182.79</td>
<td>185.35</td>
</tr>
<tr>
<td>0.3</td>
<td>0.47</td>
<td>0.54</td>
<td>149.03</td>
<td>150.13</td>
<td>151.54</td>
</tr>
<tr>
<td>0.4</td>
<td>0.46</td>
<td>0.52</td>
<td>156.62</td>
<td>158.10</td>
<td>159.79</td>
</tr>
<tr>
<td>0.5</td>
<td>0.45</td>
<td>0.52</td>
<td>167.23</td>
<td>168.83</td>
<td>170.91</td>
</tr>
</tbody>
</table>

** denotes estimator $T_{ps2}$ does not perform better than $T_{yp}$ in terms of efficiency.

Table 7: For $p = 2$ and $\rho_{p12} \neq 0$

<table>
<thead>
<tr>
<th>$\rho_{p1}$</th>
<th>$\rho_{p2}$</th>
<th>$\rho_{ps}$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{p1}$</td>
<td>$\rho_{ps}$</td>
<td>$\rho_{p2}$</td>
<td>$E_{yp}$</td>
<td>$E_{yp}$</td>
<td>$E_{yp}$</td>
</tr>
<tr>
<td>0.2</td>
<td>0.47</td>
<td>0.53</td>
<td>162.28</td>
<td>166.81</td>
<td>168.77</td>
</tr>
<tr>
<td>0.5</td>
<td>0.46</td>
<td>0.53</td>
<td>178.37</td>
<td>180.27</td>
<td>182.72</td>
</tr>
<tr>
<td>0.3</td>
<td>0.47</td>
<td>0.54</td>
<td>151.70</td>
<td>152.87</td>
<td>154.38</td>
</tr>
<tr>
<td>0.5</td>
<td>0.46</td>
<td>0.52</td>
<td>162.40</td>
<td>163.86</td>
<td>165.75</td>
</tr>
<tr>
<td>0.4</td>
<td>0.46</td>
<td>0.52</td>
<td>146.88</td>
<td>147.89</td>
<td>149.26</td>
</tr>
</tbody>
</table>

** denotes estimator $T_{ps2}$ does not perform better than $T_{yp}$ in terms of efficiency.

(2005) and Singh and Priyanka (2008) as

$$E_p(S) = \frac{\sum_{k=1}^{5000} \left[T_{ypk} - \hat{Y} \right]^2}{\sum_{k=1}^{5000} \left[T_{ypk} - \hat{Y} \right]^2} \times 100$$

and

$$E_p(SP) = \frac{\sum_{k=1}^{5000} \left[T_{ypk} - \hat{Y} \right]^2}{\sum_{k=1}^{5000} \left[T_{ypk} - \hat{Y} \right]^2} \times 100,$$

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, $T_{yp}$ (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 $\rho_{ij} = 0 \forall i \neq j = 1, 2, 3$

| $\rho_{11}$ | $\rho_{12}$ | $\rho_{13}$ | $\rho_{22}$ | $\rho_{23}$ | $\rho_{33}$ | $\mu_1^*$ | $\mu_2^*$ | $\mu_3^*$ | $E_{T|p=1}^2$ | $E_{T|p=2}^2$ |
|------------|------------|------------|------------|------------|------------|-----------|-----------|-----------|--------------|--------------|
| 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 $\rho_{ij} \neq 0 \forall i \neq j = 1, 2, 3$

| $\rho_{11}$ | $\rho_{12}$ | $\rho_{13}$ | $\rho_{22}$ | $\rho_{23}$ | $\rho_{33}$ | $\mu_1$ | $\mu_2$ | $\mu_3$ | $\mu_{SP}$ | $E_{T|p=1}^2$ | $E_{T|p=2}^2$ |
|------------|------------|------------|------------|------------|------------|--------|--------|--------|-----------|--------------|--------------|
| 0.6        | 0.2        | 0.3        | 0.5        | 0.5        | 0.46       | 0.46   | 0.55   | 0.53   | 177.71    | 111.83       |
| 0.7        | 0.4        | 0.7        | 0.6        | 0.3        | 0.48       | 0.48   | 0.58   | 0.56   | 182.62    | 132.40       |
| 0.7        | 0.3        | 0.7        | 0.5        | 0.4        | 0.60       | 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 $\varphi = 0.1$.

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 $\mu_1^*$, $\mu_2^*$ and $\mu_3^*$ 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
Kumari Priyanka, Richa Mittal, Jong-Min Kim

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 $T_{ip}$ 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 $T_{ip}$ 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 $T_{ip}$ 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 $\varphi$ (i.e. more attention is given to the estimator used at the first occasion) then the proposed estimators $T_{ip}$ ($p = 1, 2$ and 3) are superior than the estimator $T_S$ and abide by the theory while retaining higher weights for $\psi$ (i.e. more attention is given to the estimator used at the current occasion). This makes the proposed estimators $T_{ip}$ ($p = 1, 2$ and 3) more effective than the estimator $T_{SP}$ due to Singh (2005).

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

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 $T_{ip}$ 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 $T_{ip}$ 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

Authors are grateful to the honorable editors and honorable referees for their immensely valuable suggestions to improve the quality of this paper. Authors also acknowledge the free access of the data from Statistical Abstracts of the United States that is available on internet.

References


Received April 3, 2015; Revised August 4, 2015; Accepted August 17, 2015