World Applied Sciences Journal 13 (10): 2216-2223, 2011
ISSN 1818-4952

Multivariate Estimators for Two Phase Sampling

[1] Nadeem Shafique Butt, [1] Shahid Kamal and [2] Muhammad Qaiser Shahbaz

[1] College of Statistical and Actuarial Sciences, University of the Punjab, Lahore, Pakistan
[2] Department of Mathematics, COMSATS Institute of IT, Lahore, Pakistan

Abstract: In this paper some new multivariate estimators for two phase sampling has been proposed. The proposed estimators use information on multiple quantitative variables and as well as multiple qualitative variables. Empirical study has been carried out to see the performance of proposed estimator over estimator proposed by Ahmed, Hussin [1].

Key words: Multivariate estimator . two phase sampling . multiple auxiliary variables . minimum variance

INTRODUCTION

The auxiliary information has always been a source of improvement in estimation of certain population characteristics. Several estimators have been developed in single and two phase sampling which utilizes information on auxiliary variables as well as auxiliary attributes. The classical estimators which use information on auxiliary variables are the ratio and regression estimators as given in Hansen, Hurwitz [2]. The classical regression estimator of population mean is given as:

$$\bar{y}_{lr} = \bar{y} + \beta\left(\bar{X} - \bar{x}\right) \quad (1.1)$$

The value of β for which the variance of (1.1) is minimum is $\beta = S_{xy}/S_x^2$. The minimum variance of (1.1) is given as:

$$\mathrm{Var}(\bar{y}_{lr}) = \theta S_y^2\left(1 - \rho_{yx}^2\right) \quad (1.2)$$

where $\theta = n^{-1}$-N^{-1} and ρ_{yx} is the correlation coefficient between X and Y. The estimator (1.1) in case of several auxiliary variables has been discussed by number of statisticians and the estimator in this case is given as:

$$\bar{y}_{mlr} = \bar{y} + \beta^{/}\left(\bar{\mathbf{X}} - \bar{\mathbf{x}}\right) \quad (1.3)$$

where $\bar{\mathbf{x}}$ is vector of sample means for auxiliary variables. The variance of (1.3); reported by Ahmed, Hanif [3] among others; is given as:

$$\mathrm{Var}\left(\bar{y}_{mlr}\right) = \theta S_y^2\left(1 - \rho_{y.\mathbf{x}}^2\right) \quad (1.4)$$

where $\rho_{y.\mathbf{x}}^2$ is the squared multiple correlation coefficient between Y and x. The classical regression estimator for two phase sampling is given by Hansen, Hurwitz [2] as:

$$\bar{y}_{1(2)} = \bar{y}_2 + \beta\left(\bar{x}_1 - \bar{x}_2\right) \quad (1.5)$$

where $\bar{x}_1$ and $\bar{x}_2$ are first phase and second phase means of auxiliary variable X and $\bar{y}_2$ is second phase mean of Y. The variance of (1.5) is given as:

$$\mathrm{Var}\left(\bar{y}_{1(2)}\right) = S_y^2\left\{\theta_2\left(1 - \rho_{yx}^2\right) + \theta_1\rho_{yx}^2\right\} \quad (1.6)$$

where

$$\theta_h = n_h^{-1} - N_h^{-1}$$

and n_h is sample size at h^{th} phase. Ahmed, Hanif [3] has extended the (1.6) the case of several variables. Sahoo, Sahoo [4] has proposed the regression type estimator using information of two auxiliary variables. The estimator proposed by Sahoo, Sahoo [4] is given as:

$$\bar{y}_{ssm} = \bar{y}_2 + \beta_1\left(\bar{x}_1 - \bar{x}_2\right) + \beta_2\left(\bar{Z} - \bar{z}\right) \quad (1.7)$$

The variance of (1.7) is:

$$\mathrm{Var}\left(\bar{y}_{ssm}\right) = S_y^2\left\{\theta_2\left(1 - \rho_{yx}^2\right) + \theta_1\left(\rho_{yx}^2 - \rho_{yz}^2\right)\right\} \quad (1.8)$$

where ρ_{yz}^2 is squared correlation coefficient between Y and Z.

Corresponding Author: Dr. Nadeem Shafique Butt, College of Statistical and Actuarial Sciences, University of the Punjab, Lahore, Pakistan

Jhajj, Sharma [5] have proposed a family of estimators in single and two phase sampling using information on auxiliary attributes. The variance of the proposed family depends upon the point bi-serial correlation coefficient. Samiuddin and Hanif [6] have also proposed several estimators in single and two phase sampling. A regression-in-ratio estimator proposed by Samiuddin and Hanif [6] is:

$$\bar{y}_{s1(2)} = \left[\bar{y}_2 + \beta_{yz}\left(\bar{z}_1 - \bar{z}_2\right)\right]\frac{\bar{X}}{\bar{x}_2} \tag{1.9}$$

The variance of (1.9) is:

$$\begin{aligned} MSE\left(\bar{y}_{s1(2)}\right) \approx \bar{Y}^2\Big[&\theta_2\left\{C_y^2\left(1-\rho_{xy}^2\right)+\left(C_x - C_y\rho_{xy}\right)^2\right\} \\ &+\left(\theta_2-\theta_1\right)\left\{C_x^2\rho_{xz}^2 - \left(C_y\rho_{yz} - C_x\rho_{xz}\right)^2\right\}\Big] \end{aligned} \tag{1.10}$$

Hanif, Haq [7] proposed a generalized family of estimators based on the information of "k" auxiliary attributes and discussed the estimator for full, partial and no information cases. Hanif, Haq [7] showed that the proposed family has smaller mean square error than given by Jhajj, Sharma [5]. Hanif, Haq [8] proposed some ratio estimators for single phase and two phase sampling by using information on multiple auxiliary attributes. The proposed estimators are generalization of the estimator proposed by Naik and Gupta [9]. Hanif, Haq [8] also drive the shrinkage version of the proposed estimators by using the method given Shahbaz and Hanif [10]

Hanif, Ahmed [11] proposed a number of generalized multivariate ratio estimators for two-phase and multi-phase sampling in the presence of multi-auxiliary variables for estimating population mean for a single variable and a vector of variables of interest. Hanif, Ahmed [11] proposed more general ratio estimator when information on all auxiliary variables are not available for population (No Information Situation), the estimator is:

$$T_{hk(1\times p)} = \left[\bar{y}_{(k)1}\prod_{i=1}^{q}\left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}}\right)^{a_{i1}} \bar{y}_{(k)2}\prod_{i=1}^{q}\left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}}\right)^{a_{i2}} \cdots \bar{y}_{(k)p}\prod_{i=1}^{q}\left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}}\right)^{a_{ip}}\right] \tag{1.11}$$

The variance-covariance matrix of the estimator is of the following form:

$$\Sigma_{T_{hk}(p\times p)} = q_k\,\Sigma_{y(p\times p)} - \left(q_k - q_h\right)\Sigma'_{y(p\times p)}\,\Sigma^{-1}_{x(q\times q)}\,\Sigma_{yx(q\times p)} \tag{1.12}$$

Ahmed, Hussin [1] proposed a number of generalized multivariate regression estimators for two-phase and multi-phase sampling in the presence of multi-auxiliary variables for estimating population mean for a single variable and a vector of variables of interest. The proposed estimator is of the following form:

$$T_{hk(1\times p)} = \left[\bar{y}_{(k)1}\,\bar{y}_{(k)2}\,\ldots\,\bar{y}_{(k)p}\right] + \left[\sum_{i=1}^{q}a_{i1}\left(\bar{x}_{(h)i}-\bar{x}_{(k)i}\right)\sum_{i=1}^{q}a_{i2}\left(\bar{x}_{(h)i}-\bar{x}_{(k)i}\right)\ldots\sum_{i=1}^{q}a_{ip}\left(\bar{x}_{(h)i}-\bar{x}_{(k)i}\right)\right] \tag{1.13}$$

The variance-covariance matrix of the estimator is of the following form:

$$\Sigma_{T_{hk}(p\times p)} = q_k\Sigma_{y(p\times p)} - \left(q_k - q_h\right)\Sigma_{yx(p\times q)}\Sigma^{-1}_{x(q\times q)}\Sigma'_{x(q\times p)} \tag{1.14}$$

In this paper we have proposed some multivariate regression estimators using information on several auxiliary variables and as well as auxiliary attributes.

NOTATIONS

In this section we define the notations used for the development of the multivariate estimators and variance covariance matrices. Let "w" and "x" be auxiliary variables and Y be the variable of interest. Let S_{xw} be the covariance between x and w, s_{yw} be the covariance between Y and w. Using these notations we define β_{xw} as regression coefficient between x and w for the i-th response variable and

$$\beta_{y_i x.w} = S_{xy.w} / S^2_{x.w}$$

as partial regression coefficient between Y_i and x keeping the w at constant level. Also $S_{y_i x.w}$ is partial covariance between Y_i and x after removing the effect of e, $S^2_{y_i.w}$ is the partial variance of T and $S^2_{x.w}$ is the partial variance of x. We also define

$$\rho^2_{y_i x.w} = S^2_{y_i x.w} \Big/ \left(S^2_{x.w} S^2_{y_i.w}\right)$$

as partial correlation coefficient between Y_i and x after removing the effect of w, $\rho^2_{y_i x.w}$ as squared multiple correlation coefficient between Y_i and combined effects of x and w, $\rho^2_{y_i w}$ as squared multiple correlation coefficient between Y_i and w.

MULTIVARIATE ESTIMATOR WITH QUANTITATIVE PREDICTORS

In this section the multivariate extension of Roy [12] estimator has been proposed. The multivariate extension has been proposed by using information on two auxiliary variables and can be used for simultaneous estimation of several variables. The multivariate extension is proposed below:

Suppose a first phase random sample of size n_1 is available and information on auxiliary variables X and W is recorded. Suppose further that a second phase random sample of size n_2 is available and information on auxiliary variables X and W has been collected alongside information of multiple response variables Y_1, Y_2,…, Y_p. Suppose further that $\overline{y}_2$ is mean vector of estimates based upon second phase sample, k is a vector of constants and A & B are diagonal matrices with diagonal entries α_i & β_i respectively. Based upon these information, the multivariate estimator is defined below:

$$\mathbf{t} = \overline{\mathbf{y}}_2 + \left(\overline{x}_1 - \overline{x}_2\right)\mathbf{k} + \left(\overline{W} - \overline{w}_1\right)\mathbf{Ak} + \left(\overline{W} - \overline{w}_2\right)\mathbf{Bk} \quad (3.1)$$

The *i*th component of (3.1) is given as:

$$t_i = \overline{y}_{i2} + k_i\left[\left\{\overline{x}_1 + a_i\left(\overline{W} - \overline{w}_1\right)\right\} - \left\{\overline{x}_2 + b_i\left(\overline{W} - \overline{w}_2\right)\right\}\right] \quad (3.2)$$

Using conventional transformation

$$\overline{w}_1 = \overline{W} - \overline{e}_{w_1};\ \overline{w}_2 = \overline{W} - \overline{e}_{w_2};\ \overline{y}_{i2} = \overline{Y}_i + \overline{e}_{y_{i2}}$$

$$\overline{x}_1 = \overline{X} - \overline{e}_{x_1};\ \overline{x}_2 = \overline{X} - \overline{e}_{x_2}$$

the estimator (3.2) can be written in the following form:

$$t_i - y_i = \overline{e}_{y_2} + k_i\left(\overline{e}_{x_1} - \overline{e}_{x_2}\right) - k_i a_i \overline{e}_{w_1} + k_i b_i \overline{e}_{w_2}$$

Squaring above equation:

$$\begin{aligned}(t_i - y_i)^2 = {} & \overline{e}^2_{y_{i2}} + k_i^2\left(\overline{e}_{x_1} - \overline{e}_{x_2}\right)^2 + k_i^2 a_i^2 \overline{e}^2_{w_1} + k_i^2 b_i^2 \overline{e}^2_{w_2} + 2k_i \overline{e}^2_{y_{i2}}\left(\overline{e}_{x_1} - \overline{e}_{x_2}\right) - 2k_i a_i \overline{e}_{y_{i2}} \overline{e}_{w_1} \\ & + 2k_i b_i \overline{e}_{y_{i2}} \overline{e}_{w_2} - 2k_i^2 a_i \overline{e}_{w_1}\left(\overline{e}_{x_1} - \overline{e}_{x_2}\right) + 2k_i^2 b_i \overline{e}_{w_2}\left(\overline{e}_{x_1} - \overline{e}_{x_2}\right) - 2k_i^2 a_i b_i \overline{e}_{w_1} \overline{e}_{w_2}\end{aligned}$$

By applying expectation, the mean square error of t_i is:

$$S_i = MSE(t_i) = E\left(t_i - \overline{y}_i\right)^2$$

or

$$\begin{aligned}S_i = {} & q_2 S^2_{y_i} + (q_2 - q_1) k_i^2 S_x^2 + q_1 k_i^2 a_i^2 S_w^2 + q_2 k_i^2 b_i^2 S_w^2 + 2(q_1 - q_2) k_i S_{xy_i} - 2q_1 k_i a_i S_{wy_i} \\ & + 2q_2 k_i b_i S_{wy_i} + 2(q_1 - q_2) k_i^2 b_i^2 S_{wx} - 2q_1 k_i^2 a_i b_i S_w^2\end{aligned} \quad (3.3)$$

Optimum values of α_i, β_i and k_i which minimize S_i can be obtain by differentiating (3.3) with respect to unknown quantities.

$$a_i = \frac{S_x}{S_w^2} = b_{xw} \quad (3.4)$$

$$b_i = \frac{S_{wx}}{S_w^2} - \frac{1}{k_i}\frac{S_{wy_i}}{S_w^2} = b_{xw} - \frac{1}{k_i} b_{y_i w} \quad (3.5)$$

$$k_i = \left(\frac{r_{xy_i} - r_{wx} r_{wy_i}}{1 - S_{wx}^2} \right) \frac{S_y}{S_x} = b_{y_i x.w} \tag{3.6}$$

Using the values of (3.4), (3.5) and (3.6) in (3.3); the MSE becomes

$$S_i = S_{y_i}^2 \left[q_2 \left(1 - r_{y_i.wx}^2 \right) + q_1 r_{xy_i.w}^2 \left(1 - r_{wy_i}^2 \right) \right] \tag{3.7}$$

The covariance between any two components of (3.1) is derived as under:

$$t_i = \overline{y}_{i2} + k_i \left[\left\{ \overline{x}_1 + a_i \left(\overline{w} - \overline{w}_1 \right) \right\} - \left\{ \overline{x}_2 + b_i \left(\overline{w} - \overline{w}_2 \right) \right\} \right]$$

$$t_j = \overline{y}_{j2} + k_j \left[\left\{ \overline{x}_1 + a_j \left(\overline{w} - \overline{w}_1 \right) \right\} - \left\{ \overline{x}_2 + b_j \left(\overline{w} - \overline{w}_2 \right) \right\} \right]$$

Using conventional transformations:

$$t_i - y_i = \overline{e}_{y_{i2}} + k_i \left(\overline{e}_{x_1} + \overline{e}_{x_2} \right) - k_i a_i \overline{e}_{w_1} + k_i b_i \overline{e}_{w_2}$$

Similarly

$$t_j - y_j = \overline{e}_{y_{j2}} + k_j \left(\overline{e}_{x_1} + \overline{e}_{x_2} \right) - k_j a_j \overline{e}_{w_1} + k_j b_j \overline{e}_{w_2}$$

Now

$$\begin{aligned} (t_i - y_i)(t_j - y_j) &= \overline{e}_{y_{i2}} \overline{e}_{y_{j2}} + k_i \overline{e}_{y_{j2}} \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right) - a_i k_i \overline{e}_{y_{j2}} \overline{e}_{w_1} + b_i k_i \overline{e}_{y_{j2}} \overline{e}_{w_2} + k_j \overline{e}_{y_{i2}} \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right) \\ &+ k_i k_j \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right)^2 - a_i k_i k_j \overline{e}_{w_1} \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right) + b_i k_i k_j \overline{e}_{w_2} \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right) - a_j k_j \overline{e}_{w_1} \overline{e}_{y_{i2}} - a_j k_i k_j \overline{e}_{w_1} \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right) \\ &+ a_i a_j k_i k_j \overline{e}_{w_1}^2 + b_i k_i a_j k_j \overline{e}_{w_1} \overline{e}_{w2} + b_j k_j \overline{e}_{w2} \overline{e}_{y_{i2}} + b_j k_i k_j \overline{e}_{w2} \left(\overline{e}_{x_1} - \overline{e}_{x_2} \right) - a_i k_i b_j k_j \overline{e}_{w_1} \overline{e}_{w2} + b_i b_j k_i k_j \overline{e}_{w2}^2 \end{aligned}$$

By applying expectation to above equation, the covariance is:

$$\begin{aligned} S_{ij} &= Cov\left(t_i, t_j \right) = E\left(t_i - \overline{y}_i \right)\left(t_j - \overline{y}_j \right) \\ S_{ij} &= q_2 S_{y_i y_j} + k_i \left(q_1 - q_2 \right) S_{xy_j} - q_1 a_i k_i S_{wy_j} + q_2 b_i k_i S_{wy_j} + k_j \left(q_1 - q_2 \right) S_{xy_i} + k_i k_j \left(q_2 - q_1 \right) S_x^2 \\ &+ \left(q_1 - q_2 \right) b_i k_i k_j S_{wx} - q_1 a_j k_j S_{wy_i} + q_1 a_i a_j k_i k_j S_w^2 + q_1 a_j b_i k_i k_j S_w^2 + q_2 b_j k_j S_{wy_i} \\ &+ \left(q_1 - q_2 \right) b_j k_i k_j S_{wx} - q_1 a_i b_j k_i k_j S_w^2 + q_2 b_i b_j k_i k_j S_w^2 \end{aligned} \tag{3.8}$$

Using (3.4), (3.5) and (3.6) in (3.8) we have:

$$\begin{aligned} S_{ij} = S_{y_i} S_{y_j} \Bigg[& q_2 \left\{ r_{y_i y_j} - \frac{r_{xy_i} r_{xy_j} + r_{wy_i} r_{wy_j} - r_{xy_i} r_{wy_j} r_{wx} - r_{xy_j} r_{wy_i} r_{wx}}{1 - r_{wx}^2} \right\} \\ & + q_1 r_{xy_i.w} r_{xy_j.w} \sqrt{1 - r_{wy_i}^2} \sqrt{1 - r_{wy_j}^2} \Bigg] \end{aligned} \tag{3.9}$$

The covariance matrix can be written by using (3.7) and (3.9)

MULTIVARIATE ESTIMATOR WITH QUALITATIVE PREDICTORS

In this section the multivariate extension of Roy [12] estimator has been proposed. The multivariate extension has been proposed by using information on two auxiliary a ttributes and can be used for simultaneous estimation of several variables. The multivariate extension is proposed as:

Suppose a first phase random sample of size n_1 is available and information on auxiliary attributes τ and W is recorded. Further a second phase random sample of size n_2 is available and information on auxiliary attributes τ and W has been collected alongside information of multiple response variables Y_1, Y_2,…, Y_p. Suppose that $\bar{y}_2$ is the mean vector of estimates based upon second phase, k is a vector of constants and A and B are diagonal matrices with diagonal entries γ_i and η_i respectively. Based upon these information, the multivariate estimator is defined below:

$$\mathbf{t} = \bar{\mathbf{y}}_2 + (t_1 - t_2)\mathbf{k} + \left(p_d - p_{d_1}\right)\mathbf{Ak} + \left(p_d - p_{d_2}\right)\mathbf{Bk} \qquad (4.1)$$

The ith component of (4.3.1) is given as

$$t_i = \bar{y}_{i2} + k_i\left[\left\{t_1 + g_i\left(p_d - p_{d_1}\right)\right\} - \left\{t_2 + h_i\left(p_d - p_{d_2}\right)\right\}\right] \qquad (4.2)$$

using conventional transformation

$$\bar{y}_{i2} = \bar{y}_i + \bar{e}_{y_{i2}} \; ; \; t_1 = t + \bar{e}_{t_1} \; ; t_2 = t + \bar{e}_{t_2} \; ; p_{d_1} = p_d - \bar{e}_{d_1}$$

and

$$p_{d_2} = p_d - \bar{e}_{d_2}$$

Using above representations, the estimator (4.3.2) can be put in the following form

$$t_i = \left(\bar{y}_i + \bar{e}_{y_{i2}}\right) + k_i\left[\left(t + \bar{e}_{t_1}\right) + g_i\left(p_d - p_d - \bar{e}_{d_1}\right) - \left\{\left(t + \bar{e}_{t_2}\right) + h_i\left(p_d - p_d - \bar{e}_{d_2}\right)\right\}\right]$$

or

$$t_i - y_i = \bar{e}_{y_{i2}} + k_i\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) - k_i g_i \bar{e}_{d_1} + k_i h_i \bar{e}_{d_2}$$

Squaring above equation:

$$\begin{aligned}\left(t_i - y_i\right)^2 = {} & \bar{e}_{y_{i2}}^2 + k_i^2\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right)^2 + k_i^2 g_i^2 \bar{e}_{d_1}^2 + k_i^2 h_i^2 \bar{e}_{d_2}^2 + 2k_i \bar{e}_{y_{i2}}^2\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) - 2k_i g_i \bar{e}_{y_{i2}} \bar{e}_{d_1} \\ & + 2k_i h_i \bar{e}_{y_{i2}} \bar{e}_{d_2} - 2k_i^2 g_i \bar{e}_{d_1}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) + 2k_i^2 h_i \bar{e}_{d_2}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) - 2k_i^2 g_i h_i \bar{e}_{d_1} \bar{e}_{d_2}\end{aligned}$$

By applying the expectation, the mean square error of t_i is:

$$S_i = MSE\left(t_i\right) = E\left(t_i - \bar{y}_i\right)^2$$

$$\begin{aligned}S_i = {} & q_2 S_{y_i}^2 + \left(q_2 - q_1\right)k_i^2 S_t^2 + q_1 k_i^2 g_i^2 S_d^2 + q_2 k_i^2 h_i^2 S_d^2 + 2\left(q_1 - q_2\right)k_i S_{t\,y_i} - 2q_1 k_i g_i S_{d\,y_i} \\ & + 2q_2 k_i h_i S_{d\,y_i} + 2\left(q_1 - q_2\right)k_i^2 h_i^2 S_{dt} - 2q_1 k_i^2 g_i h_i S_d^2\end{aligned} \qquad (4.3)$$

Optimum values of γ_i, η_i and k_i which minimize S_i can be obtained by differentiating (4.3) with respect to unknown quantities.

$$g_i = \frac{S_{dt}}{S_d^2} = b_{td} \qquad (4.4)$$

$$h_i = \frac{S_{dt}}{S_d^2} - \frac{1}{k_i}\frac{S_{d\,y_i}}{S_d^2} = b_{td} - \frac{1}{k_i} b_{y_i d} \qquad (4.5)$$

$$k_i = \left(\frac{r_{t\,y_i} - r_{dt}\, r_{d\,y_i}}{1 - S_{dt}^2}\right)\frac{S_y}{S_t} = b_{y_i t.d} \qquad (4.6)$$

Using the values of (4.4), (4.5) and (4.6) in (4.3); the MSE becomes

$$S_i = S_{y_i}^2\left[q_2\left(1 - r_{y_i.dt}^2\right) + q_1 r_{t\,y_i.d}^2\left(1 - r_{d\,y_i}^2\right)\right] \qquad (4.7)$$

The covariance between any two components of (4.3.1) is derived as under:

Table 1: Eigen values of the variance-covariance matrices of proposed estimator

		θ_1							
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
θ_2		λ_1							
	0.2	1.76							
	0.3	2.64	2.64						
	0.4	3.51	3.52	3.53					
	0.5	4.39	4.40	4.40	4.41				
	0.6	5.27	5.27	5.28	5.29	5.29			
	0.7	6.14	6.15	6.16	6.16	6.17	6.18		
	0.8	7.02	7.03	7.03	7.04	7.05	7.05	7.06	
	0.9	7.90	7.90	7.91	7.92	7.92	7.93	7.94	7.94
θ_2		λ_2							
	0.2	0.94							
	0.3	1.40	1.42						
	0.4	1.85	1.88	1.91					
	0.5	2.31	2.34	2.36	2.39				
	0.6	2.77	2.79	2.82	2.85	2.87			
	0.7	3.22	3.25	3.28	3.30	3.33	3.36		
	0.8	3.68	3.71	3.73	3.76	3.79	3.81	3.84	
	0.9	4.14	4.16	4.19	4.22	4.24	4.27	4.30	4.32
θ_2		λ_3							
	0.2	0.08							
	0.3	0.12	0.12						
	0.4	0.16	0.16	0.16					
	0.5	0.20	0.20	0.20	0.20				
	0.6	0.24	0.24	0.24	0.24	0.24			
	0.7	0.28	0.28	0.28	0.28	0.28	0.28		
	0.8	0.32	0.32	0.32	0.32	0.32	0.32	0.32	
	0.9	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35

$$t_i = \bar{y}_{i2} + k_i\left[\left\{t_1 + g_i\left(p_d - p_{d_1}\right)\right\} - \left\{t_2 + h_i\left(p_d - p_{d_2}\right)\right\}\right]$$

Using conventional transformations:

$$t_i - y_i = \bar{e}_{y_{i2}} + k_i\left(\bar{e}_{t_1} + \bar{e}_{t_2}\right) - k_i g_i \bar{e}_{d_1} + k_i h_i \bar{e}_{d_2}$$

Similarly:

$$t_j - y_j = \bar{e}_{y_{j2}} + k_j\left(\bar{e}_{t_1} + \bar{e}_{t_2}\right) - k_j g_j \bar{e}_{d_1} + k_j h_j \bar{e}_{d_2}$$

Now

$$\begin{aligned}\left(t_i - y_i\right)\left(t_j - y_j\right) = {} & \bar{e}_{y_{i2}}\bar{e}_{y_{j2}} + k_i\bar{e}_{y_{j2}}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) - g_i k_i \bar{e}_{y_{j2}}\bar{e}_{d_1} + h_i k_i \bar{e}_{y_{j2}}\bar{e}_{d_2} + k_j\bar{e}_{y_{i2}}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) \\ & + k_i k_j\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right)^2 - g_i k_i k_j \bar{e}_{d_1}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) + h_i k_i k_j \bar{e}_{d_2}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) - g_j k_j \bar{e}_{d_1}\bar{e}_{y_{i2}} - a_j k_i k_j \bar{e}_{d_1}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) \\ & + g_i g_j k_i k_j \bar{e}_{d_1}^2 + h_i k g_j k_j \bar{e}_{d_1}\bar{e}_{d2} + h_j k_j \bar{e}_{d2}\bar{e}_{y_{i2}} + h_j k_i k_j \bar{e}_{d_2}\left(\bar{e}_{t_1} - \bar{e}_{t_2}\right) - g_i k_i h_j k_j \bar{e}_{d_1}\bar{e}_{d2} + h_i h_j k_i k_j \bar{e}_{d2}^2\end{aligned}$$

By applying expectation to above equation we get:

Table 2: Eigen values of the variance-covariance matrices of estimator proposed by Ahmed, Hussin [1]

		θ_1							
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
θ_2		μ_1							
	0.2	2.79							
	0.3	3.68	5.31						
	0.4	4.85	5.58	7.88					
	0.5	6.03	6.28	8.07	10.45				
	0.6	7.22	7.36	8.37	10.63	13.03			
	0.7	8.42	8.51	8.97	10.84	13.19	15.60		
	0.8	9.61	9.69	9.93	11.16	13.38	15.76	18.17	
	0.9	10.80	10.88	11.03	11.70	13.62	15.94	18.33	20.75
θ_2		μ_2							
	0.2	2.31							
	0.3	2.74	3.56						
	0.4	2.90	4.62	4.78					
	0.5	3.05	5.24	5.90	5.98				
	0.6	3.19	5.49	6.93	7.13	7.19			
	0.7	3.33	5.66	7.65	8.23	8.34	8.40		
	0.8	3.48	5.81	8.01	9.24	9.47	9.55	9.60	
	0.9	3.62	5.95	8.23	10.01	10.55	10.69	10.76	10.81
θ_2	μ_3								
	0.2	0.11							
	0.3	0.16	0.17						
	0.4	0.21	0.22	0.23					
	0.5	0.25	0.27	0.28	0.28				
	0.6	0.29	0.32	0.33	0.34	0.34			
	0.7	0.32	0.37	0.38	0.39	0.40	0.40		
	0.8	0.36	0.41	0.43	0.44	0.45	0.45	0.46	
	0.9	0.39	0.45	0.48	0.49	0.50	0.51	0.51	0.51

Table 3: Relative efficiency of proposed estimator over estimator proposed by Ahmed, Hussin [1]

		θ_1							
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
θ_2		$\frac{\sum l_i}{\sum m_i}$							
	0.2	0.53							
	0.3	0.63	0.46						
	0.4	0.69	0.53	0.43					
	0.5	0.74	0.59	0.49	0.42				
	0.6	0.77	0.63	0.53	0.46	0.41			
	0.7	0.80	0.67	0.57	0.50	0.45	0.40		
	0.8	0.82	0.69	0.60	0.53	0.48	0.43	0.40	
	0.9	0.84	0.72	0.63	0.56	0.51	0.46	0.43	0.39

$$S_{ij} = Cov\left(t_i, t_j\right) = E\left(t_i - \overline{y}_i\right)\left(t_j - \overline{y}_j\right)$$

$$\begin{aligned} S_{ij} = {} & q_2 S_{y_i y_j} + k_i\left(q_1 - q_2\right) S_{t y_j} - q_1 g_i k_i S_{d y_j} + q_2 h_i k_i S_{d y_j} + k_j\left(q_1 - q_2\right) S_{t y_i} + k_i k_j\left(q_2 - q_1\right) S_t^2 \\ & + \left(q_1 - q_2\right) h_i k_i k_j S_{dt} - q_1 g_j k_j S_{d y_i} + q_1 g_i g_j k_i k_j S_d^2 + q_1 g_j h_i k_i k_j S_d^2 + q_2 h_j k_j S_{d y_i} \\ & + \left(q_1 - q_2\right) h_j k_i k_j S_{dt} - q_1 g_i h_j k_i k_j S_d^2 + q_2 h_i h_j k_i k_j S_d^2 \end{aligned} \quad (4.8)$$

Using (4.4), (4.5) and (4.6) in (4.8) we have

$$S_{ij} = S_{y_i} S_{y_j} \left[q_2 \left\{ r_{y_i y_j} - \frac{r_{ty_i} r_{ty_j} + r_{dy_i} r_{dy_j} - r_{ty_i} r_{dy_j} s_{dt} - r_{ty_j} r_{dy_i} r_{dt}}{1 - r_{dt}^2} \right\} + q_1 r_{ty_i.d} r_{ty_j.d} \sqrt{1 - r_{dy_i}^2} \sqrt{1 - r_{dy_j}^2} \right] \quad (4.9)$$

The covariance matrix can be written by using (4.7) and (4.9)

NUMERICAL STUDY

In this section empirical study is conducted to see the performance of the proposed estimator over the estimator proposed by Ahmed, Hussin [1]. Ratio of the Sum of Eigen values of variance-covariance matrices is used to calculate relative efficiencies of the proposed estimator for various values of θ_1 and θ_2.

Table 1 contains the Eigen values computed from the variance-covariance matrix of proposed estimator and Table 2 contain the Eigen values computed from the variance-covariance matrix of estimator proposed by Ahmed, Hussin [1]. Table 3 shows efficiency comparison of proposed multivariate estimator with estimator proposed by Ahmed, Hussin [1]. The entries of Table 3 clearly indicate that the proposed estimator is more efficient as compared with the estimator proposed by Ahmed, Hussin [1] for all combinations of θ_1 and θ_2.

REFERENCES

1. Ahmed, Z., A.G. Hussin and M. Hanif, 2010. Generalized multivariate regression estimators for multi-phase sampling using multi-auxiliary variables. Pak. J. Statist., 26 (4): 569-583.
2. Hansen, M.H., W.N. Hurwitz and W.G. Madow, 1953. Sample survey methods and theory. New York: Wiley.
3. Ahmed, Z., M. Hanif and M. Ahmad, 2009. Generalized Regression Cum-Ratio Estimators for Two-Phase Sampling Using Muliti-Auxiliary Variables. Pak. J. Statist., 25 (2): 93-106.
4. Sahoo, J., L.N. Sahoo and S. Mohanty, 1993. A regression approach to estimation in two-phase sampling using two auxiliary variables. Current Science, 65 (1): 73-75.
5. Jhajj, H.S., M.K. Sharma and L.K. Grover, 2006. A family of estimators of population mean using information on auxiliary attribute. Pak. J. Statist., 22 (1): 43-50.
6. Samiuddin, M. and M. Hanif, 2007. Estimation of Population Mean In Single and Two Phase Sampling with or without Additional Information. Pak. J. Statist., 23 (2): 99.
7. Hanif, M., I. Haq and M.Q. Shahbaz, 2009. On a new family of estimators using multiple auxiliary attributes. World Applied Sciences Journal, 7 (11): 1419-1422.
8. Hanif, M., I. Haq and M.Q. Shahbaz, 2010. Ratio Estimators using Multiple Auxiliary Attributes. World Applied Sciences Journal, 8 (1): 133-136.
9. Naik, V.D. and P.C. Gupta, 1996. A note on estimation of mean with known population proportion of an auxiliary character. Jour. Ind. Soc. Agr. Stat., 48 (2): 151-158.
10. Shahbaz, M.Q. and M. Hanif, 2009. A General Shrinkage Estimator in Survey Sampling. World Applied Sciences Journal, 7 (5): 593-596.
11. Hanif, M., Z. Ahmed and M. Ahmad, 2009. Generalized Multivariate Ratio Estimator using Multi-Auxiliary Variables for Multi-Phase Sampling. Pak. J. Statist., 25 (4): 615-629.
12. Roy, D.C., 2003. A regression type estimator in two phase sampling using two auxiliary variables. Pak J. Statist., 19 (3): 281-290.