

A Class of Exponential Regression Type Estimators for Population Variance in Two-Phase Sampling

A. Chatterjee

Department of Mathematics, Mallabhum Institute of Technology, Bishnupur-722122, India
ism.ananya@gmail.com

G.N. Singh

Department of Applied Mathematics, Indian School of Mines, Dhanbad-826004, India
gnsingh_ism@yahoo.com

A. Bandyopadhyay

Department of Mathematics, Asansol Engineering College, Asansol-713305, India
arnabbandyopadhyay4@gmail.com

Received 13 March 2012

Accepted 17 December 2016

Abstract

This article deals with the problems of efficient estimation of population variance in two-phase (double) sampling. Using information on two auxiliary variables, a class of chain exponential to regression type estimators has been proposed and its properties are studied under two different structures of two-phase sampling. Superiority of suggested class of estimators over some existing ones is established through numerical illustrations. Suitable recommendations to the survey statistician are also made.

Keywords: Double sampling, study variable, auxiliary variable, chain-type, Exponential, regression, bias, variance, efficiency.

Mathematics Subject Classification: 62D05.

1. Introduction

Auxiliary information plays an eminent role at planning, selection and estimation stages of a sample survey. Ratio, product and regression methods of estimation are well known examples of use of auxiliary information at estimation stage. Such estimation procedures take advantage of the correlation between the auxiliary variable and the study variable. These estimation procedures are widely used in estimation of population mean of study character. Variation is an inherent phenomenon of nature; it is present everywhere in our day to day life. For instance, a physician needs a full understanding of variations in the degree of human blood pressure, body temperature and pulse rate for adequate prescription. A manufacturer needs constant knowledge of the level of variations in people's reaction to his product to be able to know whether to reduce or increase his price or improve the quality of his product. An agriculturist needs an adequate understanding of the variations in climatic factors especially from place to place (or time to time) to be able to plan on when, how and where to plant his crop. Many more situations can be encountered in practice where the estimation of population variance of the study character assumes importance. The problem of estimation population variance was first

considered by Das and Tripathi (1978). Sometimes, information on auxiliary variable may be readily available for all the units of population; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation and number of beds available in different hospitals may be known well in advance in health care surveys. If complete information on auxiliary variables lacks, it is sometimes, relatively cheap to take a large preliminary sample where auxiliary variable alone is measured, such practice is applicable in two-phase (or double) sampling. Two-phase sampling happens to be a powerful and cost effective (economical) technique to generate the reliable estimates of the unknown population parameters of the auxiliary variables in first phase sample.

Some prominent authors including Isaki (1983), Turgut and Cingi (2008), Gupta and Shabbir (2008), Shabbir and Gupta (2010), Grover (2010), Singh et al. (1988, 2003, 2010), Tailor and Sharma (2012) and Singh and Solanki (2013) have contributed towards the improvement of the estimation procedure of population variance using information on auxiliary variable. It may be noted that most of the recent works on variance estimation procedures of are biased which become a serious drawbacks for their practical applications.

Motivated with the above arguments, we have proposed a class of unbiased chain exponential to regression type estimators of population variance and studied its properties under two different structures of two-phase sampling. Performances of the proposed estimator have been examined through empirical comparisons. Suitable recommendations to the survey statistician are made for practical application of the proposed work.

2. Sample Structure and Some Existing Estimation Procedures

Let y_k , x_k and z_k be the values of the study variable y , first auxiliary variable x and second auxiliary variable z respectively associated with the k^{th} unit of the finite population $U = (U_1, U_2, \dots, U_N)$. We wish to estimate the population variance S_y^2 of the study variable y in presence of auxiliary variables x and z , when the population variance S_x^2 of x is unknown but information on z is readily available for all the units of population.

To estimate S_y^2 a first phase sample $S' (S' \subset U)$ of fixed size n is drawn by simple random sampling without replacement scheme (SRSWOR) from the entire population U and observed for the auxiliary variable x to furnish the estimate of S_x^2 . Again a second - phase sample S of size m ($m \leq n$) is drawn according to the following rules by the method of SRSWOR to observe the study variable y .

Case I: Second phase sample is drawn as a subsample of the first phase sample (i. e. $S \subset S'$).

Case II: Second phase sample is drawn independently of the first phase sample.

The case where the second sample is drawn independently of the first one was also considered by Bose (1943).

Hence onwards, we use the following notations:

\bar{X} : Population mean of the auxiliary variable x .

\bar{Y} : Population mean of auxiliary variable y .

\bar{Z} : Population mean of the auxiliary variable z .

s_{ym}^2 : sample mean square error of the variable y based on the sample of size m .

s_{xm}^2 : sample mean square error of the variable x based on the sample of size m .

s_{xn}^2 : sample mean square error of the variable x based on the sample of size n .

s_{zn}^2 : sample mean square error of the variable z based on the sample of size n .

\bar{x} , \bar{y} , \bar{z} : Sample means of the respective variables based on the second phase sample of size m .

S_y^2 : Population variance of y.

S_x^2 : Population variance of x.

s_x^2 : Sample mean square error of the variable x based on the second phase sample of size m.

$\beta_{yx}, \beta_{xz}, \beta_{yz}$: Population regression coefficients between the variables shown in subscripts.

To estimate population variance of the study variable y, Isaki (1983) considered estimator for S_y^2 as

$$t_1 = s_{ym}^2 \left(\frac{S_x^2}{s_{xm}^2} \right) \quad (1)$$

When population variance S_x^2 is unknown, motivated by Isaki's (1983) estimator one may consider two-phase sampling estimator of S_y^2 as

$$t_2 = s_{ym}^2 \left(\frac{s_{xn}^2}{s_{xm}^2} \right) \quad (2)$$

Singh *et al.* (2011) considered the exponential type estimator of population variance in two-phase sampling as

$$t_3 = s_{ym}^2 \exp \left[\frac{s_{xn}^2 - s_{xm}^2}{s_{xn}^2 + s_{xm}^2} \right] \quad (3)$$

Tailor and Sharma (2012) considered the estimator of population variance of the study variable y in two-phase sampling as

$$t_3 = f s_{ym}^2 + (1-f) s_{ym}^2 \frac{s_{xn}^2}{s_{xm}^2} \quad (4)$$

The way in which the estimate of S_y^2 is improved using the auxiliary information on x can also be extended to improve the estimate of S_x^2 in the first-phase sample, if another auxiliary variable z closely related to x but remotely related to y is used. Thus, assuming that the population variance of the auxiliary variable z is known and motivated with the ways of Chand (1975) one may proposed a chain-type ratio estimator of population variance S_y^2 as

$$t_5 = s_{ym}^2 \frac{s_{xn}^2}{s_{xm}^2} \frac{S_z^2}{s_{zn}^2} \quad (5)$$

Utilizing the knowledge of coefficient of kurtosis ($\beta_2(z)$), Singh, Mathur and Chandra (2009) suggested a modified chain ratio type estimator for S_y^2 as

$$t_6 = s_{ym}^2 \frac{s_{xn}^2}{s_{xm}^2} \left\{ \frac{S_z^2 + \beta_2(z)}{s_{zn}^2 + \beta_2(z)} \right\} \quad (6)$$

3. Proposed Class of Estimators

Inspired with the above discussions, we have considered the following class of estimators of population variance as

$$t_{ac} = \sum_{i=-1,0,1} h_i s_{ym}^2 \exp \left\{ i \left(\frac{s_{xn}^2 + b_{xz} (S_z^2 - s_{zn}^2) - s_{xm}^2}{s_{xn}^2 + b_{xzn} (S_z^2 - s_{zn}^2) + s_{xm}^2} \right) \right\} \quad (7)$$

where $t_b = s_{xn}^2 + b (S_z^2 - s_{zn}^2)$ and b is a real constant whose optimum value may be derived by minimizing the variance of the estimator t_b as

$$b = \frac{\rho_{12} C_1 S_x^2}{C_2 S_z^2} . \quad (8)$$

Remark 3.1. The estimators t_{ac} is proposed under the following conditions:

- (i) The sum of the weights is one i. e., $\sum_{i=-1,0,1} h_i = 1$.
- (ii) The weights of the linear form are chosen so that approximate bias is zero.
- (iii) The approximate MSE attains minimum.

It may be noted that the proposed class of estimator t_{ac} reduces to

- (a) s_{ym}^2 when $i = 0$.
- (b) Exponential product type estimator when $i = -1$.
- (c) Exponential ratio type estimator when $i = 1$.

4. Properties of the estimator t_{ac}

It can be easily noted that from equation (7) that the proposed estimator is t_{ac} biased for S_y^2 . Following the remark (3.1), we find that it may be made unbiased for S_y^2 . The variance $V(\cdot)$ up-to first order approximations are derived under large sample approximations using the following transformations:

$$s_{ym}^2 = S_y^2 (1 + e_1), s_{xn}^2 = S_x^2 (1 + e_2), s_{zn}^2 = S_z^2 (1 + e_3), s_{xm}^2 = S_x^2 (1 + e_4) \quad (9)$$

such that $|e_i| < 1$, $\forall (i = 1, 2, 3, 4)$. Under the above transformations shown in (9) the estimator takes the following form

$$t_{ac} = \sum_{i=-1,0,1} h_i S_y^2 (1 + e_1) \left[\frac{1 + \frac{i}{2} \left\{ (e_2 - e_4) - \frac{b_{xz}}{S_x^2} S_z^2 e_3 - \frac{1}{2} (e_2^2 - e_4^2) - \frac{b_{xz}}{2S_x^2} S_z^2 e_3 (e_2 + e_4) - \frac{b_{xz}}{2S_x^2} S_z^2 e_3 (e_2 - e_4) - \frac{1}{2} \frac{b_{xz}^2 S_z^4}{S_x^4} e_3^2 \right\}}{2! 2^2} \left\{ (e_2 - e_4) - \frac{b_{xz}}{S_x^2} S_z^2 e_3 \right\}^2 \right] + \quad (10)$$

We have derived the expressions for bias and mean square error of the proposed class of estimators t_{ac} separately for the Cases I and II of the two-phase sampling structure defined in section 2 and presented below in sections 4.1 and 4.2 respectively.

4.1. CASE I: When the second phase sample is drawn as a subsample of the first phase sample

In this case we have the following expected values of the sample statistics:

$$\left\{ \begin{array}{l} E(e_1^2)=f_1C_0^2, E(e_2^2)=f_2C_1^2, E(e_3^2)=f_2C_2^2, E(e_4^2)=f_1C_1^2 \\ E(e_1e_2)=f_2\rho_{01}C_0C_1, E(e_1e_3)=f_2\rho_{02}C_0C_2, E(e_1e_4)=f_1\rho_{01}C_0C_1 \\ E(e_2e_3)=f_2\rho_{12}C_1C_2, E(e_2e_4)=f_2C_1^2, E(e_3e_4)=f_2\rho_{12}C_1C_2 \end{array} \right\} \quad (11)$$

where,

$$f_1 = \left(\frac{1}{m} - \frac{1}{N} \right), f_2 = \left(\frac{1}{n} - \frac{1}{N} \right), f_3 = \left(\frac{1}{m} - \frac{1}{n} \right)$$

$$\mu_{pqr} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^p (y_i - \bar{Y})^q (z_i - \bar{Z})^r; (p, q, r \geq 0) \text{ are integers.}$$

Expanding the expression binomially shown in equation (10) and using the results from equations (11), we have derived the expression of bias and mean square error of the estimator t_{ac} and retain the terms which are eligible for the first order approximations, we have

$$t_{ac} = \left[S_y^2 (1 + e_1) + \frac{\alpha}{2} \left\{ (e_2 - e_4) - \frac{b_{xz}}{S_x^2} S_z^2 e_3 \right\} \right] \quad (12)$$

$$\text{where, } \alpha = \sum_{i=-1, 0, 1} ih_i \quad (13)$$

Taking expectation on both sides of equation (12) and for large population size, ignoring finite population correction (f. p. c), we get bias of the estimator t_{ac} , to the first order of approximations as,

$$B(t_{ac}) = S_y^2 \left[h_{-1} \{ 7A' - B' + C' + D' \} + h_0 \times 0 + h_1 \{ 5A' + 3B' - C' - D' \} \right]$$

where,

$$A' = \frac{1}{8} f_2 \rho_{12}^2 C_1^2, B' = \frac{1}{8} f_3 C_1^2, C' = \frac{1}{2} f_3 \rho_{01} C_0 C_1, D' = \frac{1}{2} f_2 \rho_{01} \rho_{12} C_0 C_1 \quad (14)$$

Squaring both sides of the equation (12), neglecting the terms involving power of e 's greater than two and taking expectations, we get the mean square error (MSE) of the estimator t_{ac} , to the first order of approximations as,

$$M(t_{ac}) = S_y^2 \left[f_1 C_0^2 + \frac{\alpha^2}{4} \left\{ f_3 C_1^2 + \frac{b_{xz}^2 S_z^4}{S_x^4} f_2 C_2^2 \right\} - \alpha f_3 \rho_{01} C_0 C_1 - \alpha \frac{b_{xz} S_z^2}{S_x^2} f_2 \rho_{01} C_0 C_1 \right] \quad (15)$$

To minimize the MSE of t_{ac} , we differentiate the MSE of the estimator t_{ac} given in equation (15) with respect to α and equating to zero, thus, the optimum value of α is obtained as

$$\alpha_{opt} = -\frac{B}{A} \text{ where, } A = \frac{f_3 C_1^2 + \frac{b_{xz}^2 S_z^4}{S_x^4} f_2 C_2^2}{4}, B = \frac{-f_3 \rho_{01} C_0 C_1 - \frac{b_{xz} S_z^2}{S_x^2} f_2 \rho_{01} C_0 C_1}{2} \quad (16)$$

Substituting the optimum values of α in equation (15), we have the minimum MSE of the estimator t_{ac} as,

$$\min M(t_{ac}) = S_y^2 \left[f_1 C_0^2 - \frac{B^2}{A} \right] \quad (17)$$

From the equations (13) and (16) we have,

$$\alpha = \sum_{i=-1,0,1} ih_i, \quad (18)$$

and from Remark 3.1, we get

$$\sum_{i=-1,0,1} h_i = 1 \quad (19)$$

Hence, from the equations (18) and (19), we have three unknowns to be determined from two equations only. It is therefore, not possible to find unique solutions of the constants h_i 's ($i = -1, 0, 1$). Thus, in order to get unique solutions of the constants h_i 's ($i = -1, 0, 1$), we shall impose a linear constraint as:

$$B(t_{ac}) = 0 \quad (20)$$

Which follows from the equation (13) that

$$h_{-1} \{7A' - B' + C' + D'\} + h_0 \times 0 + h_1 \{5A' + 3B' - C' - D'\} = 0 \quad (21)$$

Combining the results of the equations (18), (19) and (21), we arrive to the following matrix from

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 7A' - B' + C' + D' & 0 & 5A' + 3B' - C' - D' \end{pmatrix} \begin{pmatrix} h_{-1} \\ h_0 \\ h_1 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{B}{A} \\ 0 \end{pmatrix} \quad (22)$$

Solving equation (22), we get the unique values of h_i 's ($i = -1, 0, 1$) as,

$$h_{-1} = E, h_1 = -\frac{B}{A} + E \text{ and } h_0 = 1 + \frac{B}{A} - 2E \quad (23)$$

where

$$E = \frac{-\frac{BC'}{A} - \frac{BD'}{A} + 3\frac{BB'}{A} + 5\frac{BA'}{A}}{(12A' + 2B')}.$$

Substituting the values of h_i ($i = -1, 0, 1$) from the equation (23), in equation (6), which yields the unbiased estimator and whose variance under case-I to the first degree approximation ignoring f. p. c is derived as,

$$V(t_{ac}) = S_y^2 \left[f_1 C_0^2 + \frac{\alpha^2}{4} \left\{ f_3 C_1^2 + \frac{b_{xz}^2 S_z^4}{S_x^4} f_2 C_2^2 \right\} - \alpha f_3 \rho_{01} C_0 C_1 - \alpha \frac{b_{xz} S_z^2}{S_x^2} f_2 \rho_{01} C_0 C_1 \right] \quad (24)$$

4.2. CASE II: When the second-phase sample is drawn independently of the first-phase sample

When second phase sample S is drawn independently of the first phase sample S' , we have derived the expected values of the sample statistics as

$$\left\{ \begin{array}{l} E(e_1 e_2) = E(e_1 e_4) = E(e_2 e_3) = E(e_3 e_4) = 0 \\ E(e_1^2) = f_1 C_0^2, E(e_2^2) = f_2 C_1^2, E(e_3^2) = f_2 C_2^2, E(e_4^2) = f_1 C_1^2 \\ E(e_1 e_3) = f_2 \rho_{02} C_0 C_2, E(e_2 e_4) = f_2 C_1^2 \end{array} \right\} \quad (25)$$

Proceeding as Case-1 we have the final form of unbiased estimator t_{ac} as

$$t_{ac} = \sum_{i=-1, 0, 1} h_i s_{ym}^2 \exp \left\{ i \left(\frac{s_{xn}^2 + b_{xz} (S_z^2 - s_{zn}^2) - s_{xm}^2}{s_{xn}^2 + b_{xzn} (S_z^2 - s_{zn}^2) + s_{xm}^2} \right) \right\}$$

$$\text{Where, } h_{-1} = E', h_1 = -\frac{B'}{A'} + E', h_0 = 1 + \frac{B'}{A'} - 2E' \text{ where } E' = \frac{3B'A'' - B'B'' - B'C''}{2(B'' + A'')} \quad (26)$$

$$\text{and } A'' = \frac{1}{8} f_3 C_1^2, B'' = \frac{1}{8} f_2 \rho_{12}^2 C_1^2, C'' = \frac{1}{2} f_2 \rho_{02} \rho_{12} C_0 C_1, A' = \frac{(f_2 C_1^2 + f_1 C_1^2) + \frac{b_{xz}^2 S_z^4}{S_x^4} f_2 C_2^2 - 2 \frac{b_{xz} S_z^2}{S_x^2} f_2 \rho_{12} C_1 C_2}{4}, B' = -\frac{f_1 \rho_{01} C_0 C_1}{2}$$

whose variance under case -II to the first degree approximation ignoring f. p. c. is derived as,

$$V(t_{ac}) = S_y^2 \left[f_1 C_0^2 + \frac{\alpha^2}{4} \left\{ (f_2 C_1^2 + f_1 C_1^2) + \frac{b_{xz}^2 S_z^4}{S_x^4} f_2 C_2^2 - 2 \frac{b_{xz} S_z^2}{S_x^2} f_2 \rho_{12} C_1 C_2 \right\} - \alpha f_1 \rho_{01} C_0 C_1 \right] \quad (27)$$

5. Efficiency Comparisons of the Proposed Class of Estimators t_{ac}

To examine the performances of the proposed class of estimators under two different cases of two-phase sampling set up as suggested in this paper, we have compared their efficiencies with some existing estimators of population variance such as S_y^2 (sample variance estimator) and t_i ($i = 1, 2, 3, 5$). The Variance/ M. S. E.s/ minimum M. S. E.s of the estimators t_i are obtained up to the first order of approximations under the Cases I and II of the two phase-sampling set up and presented below.

Case I:

$$M(t_1)_I = S_y^4 \left[f_1 C_0^2 + f_3 C_1^2 (1 - 2k_{01}) \right]$$

$$\text{Where } k_{01} = \frac{\rho_{01} C_0}{C_1}$$

$$M(t_2)_I = S_y^4 \left[f_1 C_0^2 + \frac{1}{4} f_3 C_1^2 (1 - 4k_{01}) \right]$$

$$M(t_3)_I = S_y^4 \left[f_1 C_0^2 - f_3 k_{01}^2 C_1^2 \right]$$

$$M(t_5)_I = S_y^4 \left[f_1 C_0^2 + f_3 C_1^2 (1 - 2k_{01}) + f_2 C_2^2 (1 - 2k_{02}) \right]$$

$$\text{where, } k_{02} = \frac{\rho_{02} C_0}{C_2}$$

$$M(t_6)_I = M(t_1) + \theta f_2 C_2^2 (\theta - 2k_{02})$$

$$\theta = \frac{S_z^2}{S_z^2 + \beta_2(z)}$$

Case2:

$$M(t_1)_{II} = S_y^4 \left[f_1 C_0^2 + (f_1 + f_2) C_1^2 \left(1 - 2 \frac{f_1}{f_1 + f_2} k_{01} \right) \right]$$

$$M(t_2)_{II} = S_y^4 \left[f_1 C_0^2 + \frac{1}{4} (f_1 + f_2) C_1^2 \left(1 - 4 \frac{f_1}{f_1 + f_2} k_{01} \right) \right]$$

$$M(t_3)_{II} = S_y^4 \left[f_1 C_0^2 - \frac{f_1^2}{f_1 + f_2} k_{01}^2 C_1^2 \right]$$

$$M(t_5)_{II} = S_y^4 \left[f_1 C_0^2 + (f_1 + f_2) C_1^2 \left(1 - 2 \frac{f_1}{f_1 + f_2} k_{01} \right) + f_2 C_2^2 (1 - 2A_{12}) \right]$$

$$\text{Where, } A_{12} = \rho_{12} \frac{C_1}{C_2}$$

$$M(t_6)_{II} = S_y^4 \left[f_1 C_0^2 + (f_1 + f_2) C_1^2 \left(1 - 2 \frac{f_1}{f_1 + f_2} k_{01} \right) + \theta f_2 C_2^2 (\theta - 2A_{12}) \right]$$

$$\theta = \frac{S_z^2}{S_z^2 + \beta_2(z)}$$

5.1. Numerical illustration

We have chosen four natural populations to illustrate the efficacious performance of our proposed classes of estimators. The source of the populations, the nature of the variables y, x, z and the values of the various parameters are as follows.

Population I-Source: Cochran (1977) p -182

y: Number of 'placebo' children.

x: Number of paralytic polio cases in the placebo group.

z: Number of paralytic polio cases in the 'not inoculated' group.

N=33, n=12, m=7, $C_0 = 1.0104$, $C_1 = 1.1780$, $C_2 = 1.0691$, $\rho_{01} = 0.1341$, $\rho_{02} = 0.4630$, $\rho_{12} = 0.3905$

Population II-Source: Murthy (228)

y: Output.

x: Fixed Capital

z: Number of workers.

$N=80, n=40, m=30, C_0=1.1255, C_1=1.6065, C_2=1.3662, \rho_{01}=0.7319, \rho_{02}=0.7940, \rho_{12}=0.9716$

Population III-Source: Anderson new (1958)

y: Head length of second son.

x: Head length of first son.

z: Head breadth of first son

$N=25, n=20, m=10, C_0=1.3512, C_1=1.4295, C_2=1.2853, \rho_{01}=0.5057, \rho_{02}=0.5683, \rho_{12}=0.4213$

Population IV-Source: Annetee . J .Dobson(1990) p-192

y :Survival Time

x: White blood cell count

z: White blood cell count at page number 74

$N=17, n=10, m=5, C_0=1.2223, C_1=1.4049, C_2=1.0818, \rho_{01}=-0.2413, \rho_{02}=0.2383, \rho_{12}=0.5790$

To have a tangible idea about the performance of the class of estimators t_{ac} , we have computed the percent relative efficiencies (PREs) of the estimators t_{ac} under their respective optimality conditions and existing estimators with respect to S_y^2 and the findings are displayed in Table 1 – 2 where PREs are designated as $PRE = \frac{V(S_y^2)}{M(T)} \times 100$ and $M(T)$ denote M. S. E./ Minimum M. S. E. of an estimator T.

Table 1. PRE of various estimators under Case I of two-phase sampling set up

Estimators	PRE			
	Pop-I	Pop-II	Pop-III	Pop-IV
t_1	64.3717	3.4926	96.0602	42.9452
t_2	91.1553	127.2506	127.0098	41.0854
t_3	100.9601	127.2743	127.0809	104.318
t_5	61.7539	3.5262	98.8506	41.0854
t_6	67.5328	3.5261	99.3928	43.1871
t_{ac}	101.1322	4201.2774	134.9688	108.2893

Table 2. PRE of various estimators under Case II of two-phase sampling set up

Estimators	PRE			
	Pop-I	Pop-II	Pop-III	Pop-IV
t_1	43.7756	46.07407	80.92	30.6666
t_2	41.3612	139.9393	91.57628	27.1839
t_5	43.25834	96.67503	80.30392	31.7887
t_6	43.33802	96.67494	36.61778	22.3825
t_{ac}	100.0000	158.0247	100.0001	105.1143

5.2. Artificial generated population

We have generated three sets of independent random numbers of size N ($N = 100$) namely x'_k , y'_k and z'_k ($k=1, 2, 3, \dots, N$) from a standard normal distribution with the help of R-software. Further, motivated by the artificial population generation techniques adopted by Singh and Deo (2003) and Singh *et al.* (2001), we have generated the following transformed variables of the population U with the values of $\rho_{xy} = 0.8$, $\rho_{xz} = 0.6$, $\sigma_y^2 = 100$, $\mu_y = 40$, $\sigma_x^2 = 225$, $\mu_x = 50$, $\sigma_z^2 = 25$ and $\mu_z = 30$ as

$$y_k = \mu_y + \sigma_y \left[\rho_{xy} x'_k + \left(\sqrt{1 - \rho_{xy}^2} \right) y'_k \right], \quad x_k = \mu_x + \sigma_x x'_k \quad \text{and} \quad z_k = \mu_z + \sigma_z \left[\rho_{xz} x'_k + \left(\sqrt{1 - \rho_{xz}^2} \right) z'_k \right]$$

Here we generate an artificial population, where:

Table 3. PRE of various estimators under Case I of two-phase sampling set up

Estimators	PRE
	Artificial-Pop
t_1	93.84974
t_2	105.3491
t_3	105.7069
t_5	64.95389
t_6	69.46762
t_{ac}	110.13

Table 4. PRE of various estimators under Case II of two-phase sampling set up

Estimators	PRE
	Artificial-Pop
t_1	53.6212
t_2	55.2369
t_3	110.7638
t_5	44.2081
t_6	46.09201
t_{ac}	110.8818

6. Conclusions

The following conclusions may be read-out from the present study.

- Table 1 exhibits that for high positive values of the correlation coefficients, the proposed class of estimators t_{ac} yield impressive gains in efficiency over the existing estimators t_i ($i=1, 2, 3, 5, 6$) and Table 2 exhibits that for high positive values of the correlation coefficients, the proposed class of estimators t_{ac} yield impressive gains in efficiency over the existing estimators t_i ($i=1, 2, 5, 6$). This

pattern indicates that proposed class of estimators are more efficient than the existing ones under different structures two – phase sampling set up.

- (b) Tables 3-4 exhibit that in presence of positively correlated auxiliary variables, proposed class of estimators t_{ac} yield impressive gains in efficiency over the existing estimators t_i ($i= 1, 2, 3, 5, 6$).

These phenomenon indicate that the proposed classes of estimators could perform more precisely, if information on high positively correlated auxiliary variables are available. Thus, it is clear that the use of auxiliary variables are highly rewarding in terms of the proposed classes of estimators. The proposition of the class of estimators in the present study is justified as it unifies several results including unbiased and efficient variance estimation strategy and hence the proposed work may be recommended to the survey practitioner for its practical applications.

References

- [1] Bahl, S. and Tuteja, R.K. (1991): Ratio and product type exponential estimator. *Information and Optimization Sciences*, 12, 159-163.
- [2] Chand, L. (1975): Some ratio type estimators based on two or more auxiliary variables. Unpublished Ph.D. dissertation. Iowa State University, Ames, Iowa (USA).
- [3] Isaki, C.T. (1983): Variance estimation using auxiliary information. *Journal of the American Statistical Association*, 78, 117-123.
- [4] Kadilar, C. and Cingi, H. (2006): Improvement in estimating the population mean in simple random sampling. *Applied Mathematics Letters*, 19, 7579.
- [5] Murthy, M.N. (1967): *Sampling Theory and Methods*. Statistical Publishing Society, Calcutta, India.
- [6] Nayak, R. and Sahoo, L. (2012): Some alternative predictive estimators of population variance, *Revista Colombiana de Estadística*, 35(3), 507–519.
- [7] Reddy, V.N. (1978): A study on the use of prior knowledge on certain population parameters in estimation, *Sankhya, Series C*, 40, 29-37
- [8] Singh, D. and Choudhary, F.S. (1986): *Theory and Analysis of Sample Survey Designs*. Wiley Eastern Limited, New Delhi, I Edition.
- [9] Singh, G.N. and Upadhyaya, L.N. (1995): A class of modified chain type estimators using two auxiliary variables in two phase sampling. *Metron*, LIII, 117-125.
- [10] Singh, G.N. (2001): On the use of transformed auxiliary variable in the estimation of population mean in two phase sampling. *Statistic in Transition*, 5(3), 405-416.
- [11] Singh, H. and Singh, R. (2001): Improved ratio-type estimator for variance using auxiliary information, *Journal of Indian Society of Agricultural Statistics*, 54(3), 276-287.
- [12] Singh, H. and Singh, R. (2003): Estimation of variance through regression approach in two phase sampling, *Aligarh Journal of Statistics*, 23, 13-30.
- [13] Subramani, J. and Kumarapandiyam, G. (2012): Variance estimation using median of auxiliary variable, *International Journal of Probability and Statistics*, 1(3), 36-40.
- [14] Upadhyaya, L., Singh, H., Chatterjee, S. and Yadav, R. (2011): Improved ratio and product exponential type estimators, *Journal of Statistical Theory and Practice*, 5(2), 285-302.
- [15] Upadhyaya, L.N. and Singh, H.P. (1999): An estimator for population variance that utilizes the kurtosis of an auxiliary variable in sample surveys. *Vikram Mathematical Journal*, 19, 14-17.