Journal of Statistics Applications & Probability Letters
An International Journal

http://dx.doi.org/10.12785/jsapl/020107

An Efficient Class of Two-Phase Exponential Ratio and Product – Type Estimators for Estimating the Finite Population Mean

Housila P. Singh, Neeraj Sharma and Tanveer A. Tarray[*]
School of Studies in Statistics, Vikram University Ujjain, M.P., 456010 India.
*E-mail: *tanveerstat@gmail.com*

Received: 1 Dec. 2013, Revised: 5 Jan. 2014, Accepted: 9 Mar. 2014, Published online: 1 Jan. 2015

Abstract: This paper addresses the problem of estimating the population mean of the study variable using auxiliary information in double sampling. We have suggested a class of estimators of population mean in double sampling. It has been shown that the estimator due to Khatua and Mishra (2013) is a member of the proposed class. We have obtained the bias and mean square error (MSE) of the proposed class of estimators under large sample approximation. It is observed that the mean square error expression of the estimator t_{dge} obtained by Khatua and Mishra (2013) is not correct. So, we have obtained the correct expression of the mean square error of the estimator t_{dge} due to Khatua and Mishra (2013). We have compared the proposed class of estimators with that of usual unbiased estimator, usual double sampling ratio and product estimators, Singh and Vishwakarma (2007) estimators, usual regression estimator and usual double sampling ratio estimator and shown that the proposed estimator is better than existing estimators. Numerical illustrations are also given in support of the present study.

Keywords: Auxiliary variable, Exponential estimator, Mean squared error, Two- Phase sampling.

1. Introduction

In survey sampling, it is not uncommon to estimate the finite population mean $\overline{Y}$ of the study variable of y. When information on auxiliary variable x, highly correlated with y, is readily available on all units of population, it is well known that ratio estimator (for high positive correlation), product estimator (for high negative correlation) and regression estimator (for high correlation) can be used to increase the efficiency, incorporating the knowledge of population mean $\overline{X}$. However, in certain practical situation when the population mean of auxiliary variable $\overline{X}$ is not known a priori, the technique of two-phase sampling is successfully used in practice. The conventional method in such situation, a larger sample of size n' to furnish the good estimate of the population mean $\overline{X}$ while a sub-sample size n is selected from the first phase sample 'n' to observe the characteristic under study.

Consider a finite population U={1,2,3. ... ... .N}. Let y and x be two real variable assuming the value y_i and x_i on the ith unit (i = 1,2,3.........N). Now consider y be the study variable and x be the auxiliary variable. We consider simple random sampling without replacement (SRSWOR) design to draw samples in each phase of two-phase sampling set-up. The first phase sample $s'(s' \subset u)$ of fixed size n' is drawn to observe x only. The second phase sample s $(s \subset s')$ of fixed size n' is drawn to observe y and x for given s, (n<n').

Let, $\bar{x} = \frac{1}{n}\sum_{i \in s} x_i$ $\bar{y} = \frac{1}{n}\sum_{i \in s} y_i$ and $\bar{x}' = \frac{1}{n}\sum_{i \in s} x_i$

Now the usual two phase ratio, product and regression estimators are given by

$$t_{dr} = \frac{\bar{y}}{\bar{x}}\,\bar{x}', \tag{1.1}$$

$$t_{dp} = \frac{\overline{yx}}{\bar{x}'}, \tag{1.2}$$

$$t_{dlr} = \bar{y} + b_{yx}\,(\bar{x}' - \bar{x}), \tag{1.3}$$

where b_{yx} is the sample regression coefficient of 'y' on 'x', caiculated from the data based on second phase sample of size 'n',

The mean square error (MSE) of the estimators given in (1.1), (1.2) and (1.3) to first order of approximation are respectively given by

$$MSE(t_{dr}) = \bar{Y}^2\left[\lambda C_y^2 + (\lambda - \lambda')(C_x^2 - 2C_{yx})\right] \tag{1.4}$$

$$MSE(t_{dp}) = \bar{Y}^2\left[\lambda C_y^2 + (\lambda - \lambda')(C_x^2 - 2C_{yx})\right] \tag{1.5}$$

$$MSE(t_{dlr}) = \bar{Y}^2 C_y^2\left[\lambda(1-\rho^2) + \lambda'\rho^2\right] \tag{1.6}$$

where $\lambda = \left(\frac{1}{n} - \frac{1}{N}\right), \lambda' = \left(\frac{1}{n'} - \frac{1}{N}\right)$

$C_y^2 = \frac{S_y^2}{\bar{Y}^2}$, $C_y^2 = \frac{S_y^2}{\bar{Y}^2}$, $C_{yx} = \frac{S_{yx}}{\overline{yx}} = \rho C_y C_x$,

$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$, $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$,

$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X}) = \rho S_y S_x$,

ρ being the population correlation coefficient between 'y' and 'x'.

Singh and Vishwakarma (2007) have proposed modified exponential ratio and product estimate to estimate finite population mean $\bar{Y}$ of study variable y in presence of auxiliary variable x. As Singh and Vishwakarma (2007) assumed that the population mean $\bar{X}$ of auxiliary variable x is not known, they used the two phase mechanism to estimate the population mean $\bar{Y}$.

The modified exponential ratio and product estimator suggested by Singh and Vishwakarma (2007) to estimate population mean $\bar{Y}$ under two phase sampling are given by

$$t_{der} = \bar{y}.\exp\left(\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\right) \tag{1.7}$$

$$t_{dep} = \bar{y}.\exp\left(\frac{\bar{x} - \bar{x}'}{\bar{x} + \bar{x}'}\right) \tag{1.8}$$

The MSE of the estimators t_{der} and t_{dep} to first order of approximation respectively are

given by

$$MSE(t_{der}) = \bar{Y}^2\left[\lambda C_y^2 + (\lambda - \lambda')\left(\frac{C_x^2}{4} - C_{yx}\right)\right] \tag{1.9}$$

$$MSE(t_{dep}) = \bar{Y}^2\left[\lambda C_y^2 + (\lambda - \lambda')\left(\frac{C_x^2}{4} + C_{yx}\right)\right] \tag{1.10}$$

Khatua and Mishra (2013) have suggested a generalized exponential estimator to estimate finite population mean $\bar{Y}$ under two-phase sampling scheme with assumption that the population mean auxiliary variable x, $\bar{X}$ is not known.

$$t_{dge} = \bar{y}[d_1 + d_2(\bar{x}' - \bar{x})]\exp\left[\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\right], \tag{1.11}$$

where d_1 and d_2 are suitable chosen constants or suitable chosen constants or statistics and $d_1 + d_2$ are not necessarily equal to unity.

It is to be mentioned that the mean square error (MSE) expression of the estimator t_{dge} obtained by Khatua and Mishra(2013) is not correct. The correct expression of the bias and MSE of the Khatua and Mishra (2013) estimator t_{dge} are given in the following theorem.

Theoram 1.1 To the first degree of approximation, the bias and MSE of the Khatua and Mishra (2013) estimator t_{dge} are respectively given by

$$B\left(t_{dge}\right)=\bar{Y}\left[d_1\left\{1+\frac{(\lambda-\lambda')}{8}C_x^2(3-4k)\right\}+\frac{d_2\bar{X}(\lambda-\lambda')}{2}C_x^2(1-2k)-1\right] \tag{1.12}$$

and

$$\begin{aligned}MSE\left(t_{dge}\right)=\bar{Y}^2&\left[d_1^2\left\{1+\lambda C_y^2+(\lambda-\lambda')C_x^2(1-2k)\right\}+d_2^2(\lambda-\lambda')\bar{X}^2C_x^2+1\right.\\&+2d_1d_2(\lambda-\lambda')\bar{X}(1-2k)C_x^2-2d_1\left\{1+\frac{(\lambda-\lambda')}{8}C_x^2(3-4k)\right\}\\&\left.-d_2\bar{X}(\lambda-\lambda')(1-2k)C_x^2\right]\end{aligned} \tag{1.13}$$

where $k=\rho C_y/C_x$.

Proof. To obtain the bias and MSE of t_{dge} we write

$$\bar{y}=\bar{Y}(1+e_0),\ \bar{x}=\bar{X}(1-e_1),\ \bar{x}'=\bar{X}(1+e_1)$$

Such that

$$E(e_0)=E(e_1)=E(e_1)=0$$

and

$$E\left(e_0^2\right)=\lambda C_y^2, E\left(e_1^2\right)=\lambda C_x^2,\qquad E\left(e_1^2\right)=\lambda' C_x^2,\qquad E(e_0e_1)=\lambda k C_x^2,\qquad E(e_0e_1')=\lambda' k C_x^2,$$
$$E(e_1e_1')=\lambda' C_x^2$$

Now expressing (1.11) in terms of e's we have

$$\begin{aligned}t_{dge}&=\bar{Y}(1+e_0)\left[d_1+d_2\bar{X}(1+e_1'-1-e_1)\right]\exp\left\{\frac{1+e_1'-1-e_1}{1+e_1'+1+e_1}\right\}\\&=\bar{Y}(1+e_0)\left[d_1+d_2\bar{X}(e_1-e_1')\right]\exp\left\{-\frac{(e_1-e_1')}{2}\left(1+\frac{e_1+e_1'}{2}\right)^{-1}\right\}\\&=\bar{Y}(1+e_0)\left[d_1+d_2\bar{X}(e_1-e_1')\right]\left[1-\frac{(e_1-e_1')}{2}\left(1+\frac{e_1+e_1'}{2}\right)^{-1}+\frac{(e_1-e_1')^2}{8}\left(1+\frac{e_1+e_1'}{2}\right)^{-2}\cdots\right]\end{aligned}$$

$$=\overline{Y}\left[d_1(1+e_0)\left\{1-\frac{(e_1-e_1')}{2}\left(1+\frac{e_1+e_1'}{2}\right)^{-1}+\frac{(e_1-e_1')^2}{8}\left(1+\frac{e_1+e_1'}{2}\right)^{-2}\cdots\right\}\right.$$
$$\left.-d_2\overline{X}(1+e_0)(e_1-e_1')\left\{1-\frac{(e_1-e_1')}{2}\left(1+\frac{e_1+e_1'}{2}\right)^{-1}+\frac{(e_1-e_1')^2}{8}\left(1+\frac{e_1+e_1'}{2}\right)^{-2}\cdots\right\}\right]$$

$$=\overline{Y}\left[d_1(1+e_0)\left\{1-\frac{(e_1-e_1')}{2}\left(1-\frac{e_1+e_1'}{2}+\ldots\right)+\frac{(e_1-e_1')^2}{8}+\ldots\right\}\right.$$
$$\left.-d_2\overline{X}(1+e_0)(e_1-e_1')\left\{1-\frac{(e_1-e_1')}{2}\left(1-\frac{e_1+e_1'}{2}+\ldots\right)+\frac{(e_1-e_1')^2}{8}+\ldots\right\}\right]$$

$$=\overline{Y}\left[d_1(1+e_0)\left\{1-\frac{(e_1-e_1')}{2}+\frac{\left(e_1^2-e_1'^2\right)^2}{4}+\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{8}+\ldots\right\}\right.$$
$$\left.-d_2\overline{X}(1+e_0)(e_1-e_1')\left\{1-\frac{(e_1-e_1')}{2}+\frac{\left(e_1^2-e_1'^2\right)}{4}+\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{8}+\ldots\right\}\right]$$

$$=\overline{Y}\left[d_1\left\{1+e_0-\frac{(e_1-e_1')}{2}+\frac{(e_0e_1-e_0e_1')}{2}+\frac{\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{8}+\frac{\left(3e_0e_1^2-e_0e_1'^2-2e_0e_1e_1'\right)}{8}+\ldots\right\}\right.$$
$$\left.-d_2\overline{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{2}+\frac{(e_1-e_1')\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{8}+\ldots\right\}\right]$$

Neglecting terms of e’s having power greater than two we have

$$t_{dge}\cong\overline{Y}\left[d_1\left\{1+e_0-\frac{(e_1-e_1')}{2}+\frac{(e_0e_1-e_0e_1')}{2}+\frac{\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{8}\right\}\right.$$
$$\left.-d_2\overline{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{2}\right\}\right]$$

Subtracting $\overline{Y}$ from both side of the above expression we have

$$\left(t_{dge}-\overline{Y}\right)\cong\overline{Y}\left[d_1\left\{1+e_0-\frac{(e_1-e_1')}{2}+\frac{(e_0e_1-e_0e_1')}{2}+\frac{\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{8}\right\}\right.$$
$$\left.-d_2\overline{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{2}\right\}-1\right] \tag{1.14}$$

Taking expectation of both sides of (1.14) we get the bias of t_{dge} to the first degree of approximation as

$$B\left(t_{dge}\right)=E\left(t_{dge}-\overline{Y}\right)$$
$$=\overline{Y}\left[d_1\left\{1+\frac{(\lambda-\lambda')}{8}C_x^2(3-4k)\right\}\right)$$
$$\left.+d_2\overline{X}(\lambda-\lambda')C_x^2(1-2k)-1\right] \tag{1.15}$$

Squaring both sides of (1.14) and neglecting terms of e's having power greater than two we have

$$\left(t_{dge}-\overline{Y}\right)^2=\overline{Y}^2\left[d_1^2\left\{1+e_0^2+\frac{(e_1-e_1')^2}{4}+2e_0-(e_1-e_1')-(e_0e_1-e_0e_1')\right.\right.$$
$$\left.-(e_0e_1-e_0e_1')+\frac{\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{4}\right\}+d_2^2\overline{X}^2(e_1-e_1')^2+1$$
$$-2d_1d_2\overline{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{(e_1-e_1')^2}{2}\right.$$
$$\left.+(e_0e_1-e_0e_1')-\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{2}\right\}$$
$$-2d_1\left\{1+e_0-\frac{(e_1-e_1')}{2}-\frac{(e_0e_1-e_0e_1')}{2}+\frac{\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{8}\right\}$$
$$\left.+2d_2\overline{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{2}\right\}\right]$$

or

$$\left(t_{dge}-\overline{Y}\right)^2=\overline{Y}^2\Big[1+d_1^2\left\{1+2e_0-\left(e_1-e_1'\right)-2\left(e_0e_1-e_0e_1'\right)+\left(e_1^2-e_1e_1'\right)\right\}+d_2^2\overline{X}^2\left(e_1-e_1'\right)^2$$
$$-2d_1d_2\overline{X}\left\{\left(e_1-e_1'\right)+2\left(e_0e_1-e_0e_1'\right)-\left(e_1^2-2e_1e_1'+e_1'^2\right)\right\}$$
$$-2d_1\left\{1+e_0-\frac{\left(e_1-e_1'\right)}{2}-\frac{\left(e_0e_1-e_0e_1'\right)}{2}+\frac{\left(3e_1^2-e_1'^2-2e_1e_1'\right)}{8}\right\}$$
$$+2d_2\overline{X}\left\{\left(e_1-e_1'\right)+\left(e_0e_1-e_0e_1'\right)-\frac{\left(e_1^2-2e_1e_1'+e_1'^2\right)}{2}\right\}\Big] \tag{1.16}$$

Taking expectation of both sides of (1.16) we get the mean squared error of the estimator t_{dge} to the first degree of approximation as

$$MSE\left(t_{dge}\right)=\overline{Y}^2\Big[1+d_1^2\left\{1+\lambda C_y^2+\left(\lambda-\lambda'\right)C_x^2(1-2k)\right\}+d_2^2\overline{X}^2\left(\lambda-\lambda'\right)C_x^2$$
$$+2d_1d_2\overline{X}C_x^2\left(\lambda-\lambda'\right)(1-2k)$$
$$-2d_1\left\{1+\frac{\left(\lambda-\lambda'\right)C_x^2}{8}(3-4k)\right\}$$
$$-d_2\overline{X}\left(\lambda-\lambda'\right)C_x^2(1-2k)\Big]$$

$$=\overline{Y}^2\left[1+d_1^2A_{1(1)}+d_2^2A_{2(1)}+2d_1d_2A_{3(1)}-2d_1A_{4(1)}-d_2A_{3(1)}\right] \tag{1.17}$$

which is minimum when

$$\begin{bmatrix} A_{1(1)} & A_{3(1)} \\ A_{3(1)} & A_{2(1)} \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}=\begin{bmatrix} A_{4(1)} \\ A_{3(1)}/2 \end{bmatrix} \tag{1.18}$$

Solving (1.18) we get the optimum values of d_1 and d_2 as

$$d_1=\frac{\left(2A_{2(1)}A_{4(1)}-A_{3(1)}^2\right)}{2\left(A_{1(1)}A_{2(1)}-A_{3(1)}^2\right)}=d_{10} \tag{1.19}$$

$$d_2=\frac{A_{3(1)}\left(A_{1(1)}-2A_{4(1)}\right)}{2\left(A_{1(1)}A_{2(1)}-A_{3(1)}^2\right)}=d_{20} \tag{1.20}$$

where ,

$$A_{1(1)} = \left[1 + \lambda C_y^2 + (\lambda - \lambda')(1-2k)C_x^2\right],$$

$$A_{2(1)} = (\lambda - \lambda')\bar{X}^2 C_x^2,$$

$$A_{3(1)} = (\lambda - \lambda')\bar{X}(1-2k)C_x^2,$$

$$A_{4(1)} = \left[1 + \frac{(\lambda - \lambda')C_x^2}{8}(3-4k)\right]$$

Substituting (d_{10}, d_{20}) in place of (d_1, d_2) in (1.17) we get the minimum MSE of t_{dge} as

$$\min.\,MSE(t_{dge}) = \bar{Y}^2\left[1 - \left(\frac{4A_{2(1)}A_{4(1)}^2 + A_{1(1)}A_{3(1)}^2 - 4A_{3(1)}^2 A_{4(1)}}{4\left(A_{1(1)}A_{2(1)} - A_{3(1)}^2\right)}\right)\right] \qquad (1.21)$$

which is correct expression of MSE of t_{dge}.

If we set $d_1 = 1$ in (1.11), then the class of estimators t_{dge} reduces to:

$$t_{dge(1)} = \bar{y}\left[1 + d_2(\bar{x}' - \bar{x})\right]\exp\left[\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\right] \qquad (1.22)$$

Putting $d_1 = 1$ in (1.15) and (1.17) the bias and MSE of $t_{dge(1)}$ are respectively given by

$$B(t_{dge(1)}) = \frac{\bar{Y}(\lambda - \lambda')}{8}C_x^2\left[(3-4k) + 8d_2\bar{X}(1-2k)\right] \qquad (1.23)$$

and

$$MSE(t_{dge(1)}) = \bar{Y}^2\left[1 + A_{1(1)} - 2A_{4(1)} + d_2^2 A_{2(1)} + d_2 A_{3(1)}\right] \qquad (1.24)$$

The MSE $t_{dge(1)}$ is minimized for

$$d_2 = -\frac{A_{3(1)}}{2A_{2(1)}} = d_{20pt} \qquad (1.25)$$

Thus the resulting (minimum) MSE of $t_{dge(1)}$ is given by

$$MSE(t_{dge(1)}) = \bar{Y}^2\left[1 + A_{1(1)} - 2A_{4(1)} + \frac{A_{3(1)}^2}{4A_{2(1)}}\right]$$

$$= S_y^2\left[\lambda - (\lambda - \lambda')\rho^2\right] \quad (1.26)$$

which is equal to the approximate variance of the usual regression estimator

$$\bar{y}_{lrd} = \bar{y} + \hat{\beta}(\bar{x}' - \bar{x}) \quad (1.27)$$

where $\hat{\beta} = s_{xy} / s_x^2$ is the sample regression coefficient of y on x.

we note that the MSE expression of the estimator of $t_{dge(1)}$:

$$MSE(t_{dge(1)}) = \bar{Y}^2\left[1 + d_1^2\left\{1 + \lambda C_y^2 + \frac{(\lambda - \lambda')}{4}C_x^2(1 - 4k)\right\} + d_2^2\bar{X}^2(\lambda - \lambda')C_x^2 \right.$$
$$\left. + 2d_1d_2\bar{X}C_x^2(\lambda - \lambda')(1 - 2k) + 2d_1d_2\bar{X}C_x^2(\lambda - \lambda')(1 - 2k) - 2d_1\right] \quad (1.28)$$

obtained by Khatua and Mishra (2013) is not corrected. So the other results of the paper are also incorrect. This led authors to derive the correct expression of MSE of the estimator $t_{dge(1)}$ which is given in 1.28.

2. The Suggested Class of Estimators

For estimating the population mean $\bar{Y}$, we define a class of estimators as

$$t_s = \bar{y}[\delta_1 + \delta_2(\bar{x}' - \bar{x})]\left(\frac{\bar{x}'}{\bar{x}}\right)^{\alpha} \exp\left[\frac{\delta(\bar{x}' - \bar{x})}{(\bar{x}' + \bar{x})}\right], \quad (2.1)$$

where (δ_1, δ_2) are suitably chosen constants such that the MSE of t is minimum, (α, δ) are suitably chosen classes scalars which may assume real number or the functions of ρ, C_y, C_x etc.

It is to be noted that for $(\alpha, \delta) = (0,1)$ the proposed class of estimators t_s reduces to the class of estimators reported by Khatua and Mishra (2013).

To obtain the bias and MSE of t_s we express t_s in terms of e's we have

$$t_s = \bar{Y}(1 + e_0)[\delta_1 + \delta_2\bar{X}(e_1' - e_1)]\left(\frac{1 + e_1'}{1 + e_1}\right)^{\alpha} \exp\left[\frac{\delta(e_1' - e_1)}{2 + e_1 + e_1'}\right]. \quad (2.2)$$

We assume that $|e_1|<1, |e_1'|<1$ so that the right hand side of (2.2) is expandable. Expanding the right hand side, multiplying out and neglecting terms of e's having power greater than two we have some members of the proposed class of estimators t_s are given in Table 2.1

Table 2.1 Some members of the proposed class of estimators t_s

S.No.	Estimators	Values of scalars			
		δ_1	δ_2	α	δ
1	$t_1 = \bar{y}$	**1**	**0**	**0**	**0**
2	$t_2 = \bar{y}(\bar{x}'/\bar{x})$	**1**	**0**	**1**	**0**
3	$t_3 = \bar{y}(\bar{x}/\bar{x}')$	**1**	**0**	**-1**	**0**
4	$t_4 = \bar{y}(\bar{x}'/\bar{x})^{\alpha}$ **Srivastava (1970) estimator**	**1**	**0**	α	**0**
5	$t_5 = \bar{y}\exp\left(\frac{\bar{x}'-\bar{x}}{\bar{x}'+\bar{x}}\right)$ **Singh & Vishwakarma (2007) estimator**	**1**	**0**	**0**	**1**
6	$t_6 = \bar{y}\exp\left(\frac{\bar{x}-\bar{x}'}{\bar{x}+\bar{x}'}\right)$ **Singh & Vishwakarma (2007) estimator**	**1**	**0**	**0**	**-1**
7	$t_7 = \bar{y}\exp\left(\frac{\delta(\bar{x}'-\bar{x})}{\bar{x}'+\bar{x}}\right)$	**1**	**0**	**0**	δ
8	$t8 = \bar{y}[d_1 + d_2(\bar{x}'-\bar{x})]\exp\left(\frac{\bar{x}'-\bar{x}}{\bar{x}'+\bar{x}}\right)$ **Khatua and Mishra (2013) estimator**	d_1	d_2	**0**	**1**
9	$t_9 = \bar{y}[\delta_1 + \delta_2(\bar{x}'-\bar{x})]\left(\frac{\bar{x}}{\bar{x}'}\right)$	δ_1	δ_2	**-1**	**0**
10	$t_{10} = \bar{y}[\delta_1 + \delta_2(\bar{x}'-\bar{x})]\left(\frac{\bar{x}}{\bar{x}'}\right)\exp\left(\frac{(\bar{x}-\bar{x}')}{8(\bar{x}+\bar{x}')}\right)$	δ_1	δ_2	**-1**	**-1/8**
11	$t_{11} = \bar{y}[\delta_1 + \delta_2(\bar{x}'-\bar{x})]\left(\frac{\bar{x}}{\bar{x}'}\right)\exp\left(\frac{(\bar{x}-\bar{x}')}{4(\bar{x}'+\bar{x})}\right)$	δ_1	δ_2	**-1**	**-1/4**

$$t_s \cong \bar{Y}\left[\delta_1\left\{1+e_0-\frac{\theta(e_1-e_1')}{2}-\frac{\theta(e_0e_1-e_0e_1')}{2}+\frac{\theta(e_1^2-e_1'^2)}{4}+\frac{\theta^2(e_1-e_1')^2}{8}\right\}\right.$$
$$\left.-\delta_2\bar{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\theta(e_1-e_1')^2}{2}\right\}\right]$$

or

$$(t_s-\bar{Y})\cong \bar{Y}\left[\delta_1\left\{1+e_0-\frac{\theta(e_1-e_1')}{2}-\frac{\theta(e_0e_1-e_0e_1')}{2}+\frac{\theta(e_1^2-e_1'^2)}{4}+\frac{\theta^2(e_1-e_1')^2}{8}\right\}\right.$$
$$\left.-\delta_2\bar{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\theta(e_1-e_1')^2}{2}\right\}-1\right] \tag{2.3}$$

where $\theta=(2\alpha+\delta)$.

Taking expectation of both sides of (2.3) we get the bias of the proposed class of estimator t to the first degree of approximation as

$$B(t_s)=\bar{Y}\left[\delta_1\left\{1+\frac{(\lambda-\lambda')\theta C_x^2(\theta-4k+2)}{8}\right\}+\delta_2\frac{\bar{X}(\lambda-\lambda')C_x^2(\theta-2k)}{2}-1\right] \tag{2.4}$$

Squaring both sides of (2.3) and neglecting terms of e's having power greater than two we have

$$(t_s-\bar{Y})^2=\bar{Y}^2\left[1+d_1^2\left\{1+2e_0-\theta(e_1-e_1')-2\theta(e_0e_1-e_0e_1')+e_0^2\right.\right.$$
$$\left.+\frac{\theta}{2}\theta(e_1^2+e_1'^2-2e_1e_1')+e_1^2-e_1'^2\right\}+d_2^2\bar{X}^2(e_1-e_1')^2$$
$$-2d_1d_2\bar{X}\left\{(e_1-e_1')+2(e_0e_1-e_0e_1')-\theta(e_1-e_1')^2\right\}$$
$$-2d_1\left\{1+e_0-\frac{\theta(e_1-e_1')}{2}-\frac{\theta(e_0e_1-e_0e_1')}{2}+\frac{\theta}{8}\left(\theta(e_1-e_1')^2+2(e_1^2-e_1'^2)\right)\right\}$$
$$\left.+2d_2\bar{X}\left\{(e_1-e_1')+(e_0e_1-e_0e_1')-\frac{\theta}{2}(e_1-e_1')^2\right\}\right] \tag{2.5}$$

Taking expectation of both sides of (2.5) we get the mean squared error of t_s to the first degree of approximation as

$$MSE(t_s)=\bar{Y}^2\left[1+\delta_1^2A_1+\delta_2^2A_2+2\delta_1\delta_2A_3-2\delta_1A_4-\delta_2A_3\right] \tag{2.6}$$

where

$$A_1 = \left[1+\lambda C_y^2 + \frac{(\lambda-\lambda')\theta C_x^2(\theta-4k+1)}{2}\right]$$

$$A_2 = (\lambda-\lambda')\overline{X}^2 C_x^2,$$

$$A_3 = \overline{X}(\lambda-\lambda')(\theta-2k)C_x^2,$$

$$A_4 = \left[1+\frac{\theta(\lambda-\lambda')C_x^2}{8}(\theta-4k+2)\right]$$

The MSE of t_s is minimum when

$$\begin{bmatrix} A_1 & A_3 \\ A_3 & A_2 \end{bmatrix}\begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} A_4 \\ (A_3/2) \end{bmatrix} \tag{2.7}$$

Solving (2.7) we get the optimum values of δ_1 and δ_2 as

$$\delta_1 = \frac{\left(2A_2A_4 - A_3^2\right)}{2\left(A_1A_3 - A_3^2\right)} = \delta_{10} \tag{2.8}$$

$$\delta_2 = \frac{A_3(A_1 - 2A_4)}{2\left(A_1A_2 - A_3^2\right)} = \delta_{20} \tag{2.9}$$

Putting (2.8) and (2.9) in (2.6) we get the minimum MSE of t_s as

$$\min .\mathrm{MSE}(t_s) = \overline{Y}^2\left[1 - \frac{4A_2A_4^2 + A_1A_3^2 - 4A_3^2A_4}{4\left(A_1A_2 - A_3^2\right)}\right] \tag{2.10}$$

For $\delta_1 = 1$, the proposed class of estimators t_s reduces to :

$$t_{s(1)} = \overline{y}\left[1+\delta_2(\overline{x}'-\overline{x})\right]\left(\frac{\overline{x}'}{\overline{x}}\right)^{\alpha} \exp\left\{\frac{\delta(\overline{x}'-\overline{x})}{\overline{x}'-\overline{x}}\right\} \tag{2.11}$$

Putting $\delta_1 = 1$ in (2.4) and (2.6) we get the bias and MSE of $t_{s(1)}$ to the first degree of approximation are respectively given by

$$B\left(t_{s(1)}\right)=\overline{Y}\frac{(\lambda-\lambda')C_x^2}{8}\left[\theta(\theta-4k+2)+4\delta_2\overline{X}(\theta-2k)\right] \tag{2.12}$$

$$MSE\left(t_{s(1)}\right)=\overline{Y}^2\left[1+A_1-2A_4+\delta_2^2A_2+\delta_2A_3\right] \tag{2.13}$$

The MSE $t_{s(1)}$ is minimized for

$$\delta_2=-\frac{A_3}{2A_2}=d_{2(0pt)} \tag{2.14}$$

Thus resulting minimum MSE of $t_{s(1)}$ is given by

$$\min .MSE\left(t_{s(1)}\right)=\overline{Y}^2\left[1+A_1-2A_4-\frac{A_3^2}{4A_2}\right]$$

$$=S_y^2\left[\lambda-(\lambda-\lambda')\rho^2\right] \tag{2.15}$$

which is equal to the approximate variance of the usual regression estimator $\overline{y}_{lrd}$ in two phase sampling.

3. Efficiency comparisons

For $(\delta_1,\delta_2,\alpha,\delta)=(1,0,0,0)$ in (2.2.1), the class of estimators t_s reduces to the usual unbiased estimator

$$t_s^{(1)}=\overline{y} \tag{3.1}$$

The variance of $\overline{y}$ is given by

$$var(\overline{y})=MSE(\overline{y})=\lambda\overline{y}^2C_y^2=\lambda S_y^2 \tag{3.2}$$

From (2.1.14), (2.1.15), (2.1.6) and (2.3.2) we have

$$MSE(\overline{y})-MSE(t_{dlr})=(\lambda-\lambda')\overline{Y}^2C_x^2k^2\geq 0 \tag{3.3}$$

$$MSE(t_{dr})-MSE(t_{dlr})=(\lambda-\lambda')\overline{Y}^2C_x^2(1-k)^2\geq 0 \tag{3.4}$$

$$MSE(t_{dp})-MSE(t_{dlr})=(\lambda-\lambda')\overline{Y}^2C_x^2(1+k)\geq 0 \tag{3.5}$$

$$MSE(t_{der}) - MSE(t_{dlr}) = \frac{(\lambda - \lambda')\overline{Y}^2}{4} C_x^2 (1-2k)^2 \geq 0 \tag{3.6}$$

$$MSE(t_{dep}) - MSE(t_{dlr}) = \frac{(\lambda - \lambda')\overline{Y}^2}{4} C_x^2 (1+2k)^2 \geq 0 \tag{3.7}$$

It is clear from (3.3), to (3.7) that the usual two – phase regression estimator $\overline{y}_{dlr}$ is better than sample mean $\overline{y}$, ratio estimator t_{dr}, the product estimator t_{dp} in two phase sampling, two phase exponential ratio estimator t_{der} and the two phase exponential procedure estimator t_{dep}.

Further from (2.10) and (2.15) we have

$$MSE(t_{dlr}) - \min . MSE(t_s) = \overline{Y}^2\left[A_1 - 2A_4 - \frac{A_3^2}{4A_2} + \frac{\left(4A_2A_4^2 + A_1A_3^2 - 4A_3^2A_4\right)}{4\left(A_1A_2 - A_3^2\right)}\right]$$

$$= \frac{\overline{Y}^2\left(2A_1A_2 - A_3^2 - 2A_2A_4\right)^2}{4A_2\left(A_1A_2 - A_3^2\right)} \geq 0 \tag{3.8}$$

which shows that the proposed class of estimator's t_s is more efficient than the regression estimator t_{dlr}. Thus the proposed class of estimator's t_s is better than usual unbiased estimator $\overline{y}$, conventional two phase ratio estimator t_{dr} two – phase product estimator t_{dp} and the two phase regression estimator t_{dlr}.

4. Empirical Study

To have tangible idea about the performance of the members of the proposed class of estimators t_s over usual unbiased estimator $\overline{y}$ we have computed the percent relative efficiency (PRE) of the member of the suggested class of estimator's t_s with respect to $\overline{y}$ by using the formula:

$$PRE\left(t_s, \overline{y}\right) = \left[\frac{\lambda C_y^2}{\left[1 + \delta_1^2 A_1 + \delta_2^2 A_2 + 2\delta_1\delta_2 A_3 - 2\delta_1 A_4 - \delta_2 A_3\right]}\right] \times 100 \tag{4.1}$$

For better eye view, we have considered two natural population data set with positive correlation between y and x and two population data sets with negative correlation between y and x earlier considered by Khatua and Mishra (20103). The descriptions are given below:

Population I: Murthy (1967), P. 228

y: Output, x: Fixed Capital, N =80, n' =70 ,n=30

$\lambda = 0.02083, \lambda' = 0.00179,\qquad \bar{Y} = 5182.64, C_y^2 = 0.1255, C_x^2 = 0.5635, C_{yx} = 0.2503,$
$\rho = 0.9413,\ k = 0.44419$

Population II: Das (1988)

y: No. of agricultural labors for 1971, x: No of agricultural labors for 1961, N =278, n'=70, n=30

$\lambda = 0.0297, \lambda' = 0.01069,\qquad \bar{Y} = 39.0\quad 6, C_y^2 = 2.0883, C_x^2 = 2.6237, C_{yx} = 1.6883,$
$\rho = 0.7213,\ k = 0.64348$

Population III: Steel and Torrie (1960), P. 282

y: Log of leaf burn in sacs , x:Chlorine percentage, N =30, n'=12 ,n=4

$\lambda = 0.21667, \lambda' = 0.05,\qquad \bar{Y} = 0.6\quad 8, C_y^2 = 0.2306, C_x^2 = 0.5614, C_{yx} = -0.1798,$
$\rho = -0.4996,\ k = -0.32027$

Population IV: Gujurati(1999), P. 259

y: Year to year percentage change in the index of hourly earnings, x: The unemployment rate (%), N =12, n'=8 ,n=5,

$\lambda = 0.11667, \lambda' = 0.041667,\qquad \bar{Y} = 4.066, C_y^2 = 0.0977, C_x^2 = 0.0535, C_{yx} = -0.0519,$
$\rho = -0.718,\ k = -0.970$

We have computed the percent relative efficiencies of the different members of the proposed class of estimator's t_{dge} with respect to usual unbiased estimator $\bar{y}$ and findings are shown in Tables 2.4.1 and 2.4.2.

Table 2.4.1 Percent relative efficiency with respect to mean per estimator ($\bar{y}$) when ρ > 0

Population	Estimators						
	t_{dr}	t_{der}	t_{dlr}	t_{dge}	t_9	t_{10}	t_{11}
1	**68.57**	**493.04**	**526.58**	**530.93**	**1406.10**	**1948.66**	**3427.95**
2	**130.02**	**146.34**	**149.98**	**156.69**	**172.39**	**179.84**	**189.72**

Table 2.4.2 Percent relative efficiency with respect to mean per estimator ($\bar{y}$) when ρ < 0

Population	Estimators						
	t_{pr}	t_{dep}	t_{dlr}	t_{dge}	t_9	t_{10}	t_{11}
3	**59.76**	**115.15**	**123.76**	**126.25**	**127.78**	**129.96**	**132.96**
4	**149.47**	**133.95**	**149.56**	**149.65**	**149.81**	**149.79**	**149.77**

Tables 2.4.1 and 2.4.2 clearly show that the modified two-phase estimator is more efficient than mean per unit estimator, two- phase ratio estimator, two –phase ratio type exponential estimator, two-phase product estimator, two-phase product type exponential and two-phase regression estimator. Tables 2.4.1 and 2.4.2 also exhibits that the new generated estimators are more efficient than the existing once. Thus the proposed estimators are to be preferred in practice.

References

[1] Bhal S. and Tuteja R.K. (1991) Ratio and product type exponential estimator, Infor. Optimiz. Sci. 12, 159-163.

[2]Cochran W.G. (1977) Sampling Technique, 3rd Edition. New York: John Wiley and Sons, USA.

[3]Das A.K. (1977) Contribution to the theorey of sampling strategies bases on auxiliary information. Unpublised Ph.D. Thesis.

[4]Khatua P. and Mishra G.(2013)A modified generalized two-phase exponential estimator to estimate finite population mean. IOSR J. Math. Vol 8 No 6 pp. 43-47.

[5]Gujurati D. (1999) Essential of Economics, 2th Edition Mc Graw-Hill International.

[6]Muthy M.N. (1977)Sampling theory and methods. Statist. Pub. Soc. Calcutta.

[7] Solanki R. S., Singh H.P. and Pal S. K. (2013) Improved estimation of finite population mean sample surveys. J. Adv. Com. Vol 1 No 2 pp. 70-78.

[8]Singh H.P., Pal S. K. and Solanki R. S. (2013) Improved estimation of finite population variance using quartiles. Istatistik - J. Tur. Stat. Asso.,Vol6 No 3 pp. 166-121.

[9]Singh H.P. and Vishwakaram G.K. (2007) Modified exponential ratio and product estimators for finite population mean in double sampling. Aust. J. Stat. Vol 36 No 3 pp.217-225.

[10]Steel R.G.D. and Torrie J.H. (1960) Principle and procedures of statistics, MC – Grew Hill, Newyork.

[11]Sukhatme P.V., Sukhatme B.V., and Ahsok C. (1984) Sampling Theory of surveys with applications. Ind. Soc. Agri. Stat., New Delhi.