AN IMPROVED CLASS OF ATIOTYPE ESTIMATORS FOR FINITE POPULATION MEAN UNDER MAXIMUM AND MINIMUM VALUES

[1]Mursala Khan, [2,*]Sultan Hussain

[1,2] Department of Mathematics, COMSATS Institute of Information Technology, Abbottabad, 22060, Pakistan

mursala.khan@yahoo.com, tausef775650@yahoo.co.in

ABSTRACT. In this research article, a class of ratio-type estimators has been proposed for the estimation of finite population mean under simple random sampling scheme when there is maximum and minimum values. We have also found the approximate bias and mean squared error up to first order of approximation. Also some theoretical conditions are carry out under which the suggested class of ratio-type estimators have always been efficient than the usual unbiased, Sarndal [3], the classical ratio, Singh and Tailor [4], Sisodia and Dwivedi [5] and Kadilar and Cingi [1] estimators. Numerical study has also been carried out with the help of natural population, that the proposed class of estimators is more efficient than the previous work.

Key Words: Study variable, Auxiliary variable, Ratio estimators, Maximum and Minimum values, Simple random sampling, Bias, Mean squared error, Efficiency.

1 INTRODUCTION

The information on auxiliary variable plays a major role for the estimation of finite population mean in the field of survey sampling and have greatly increase the efficiency of the estimators. Auxiliary information is efficiently used in ratio, product and regression estimators, for the estimation of finite population mean of the variable of interest. Using auxiliary information Singh and Tailor [4] suggested a ratio estimator using the known knowledge of correlation coefficient, Sisodia and Dwivedi [5] proposed ratio estimator using the known knowledge coefficient of variation of an auxiliary variable. Kadilar and Cingi [1] propose a class of ratio estimators for finite population mean using the knowledge of the auxiliary variable. In this study we have proposed a class of ratio-type estimators for the estimation of finite population mean of the study variable.

Let us consider a finite population of size N of different units $U=\{U_1,U_2,U_3....,U_N\}$. Let y and x be the study and the auxiliary variable with corresponding values y_i and x_i respectively for the i^{th} unit $i=\{1,2,3....,N\}$ defined on a finite population U.

Let $\bar{Y}=(1/N)\sum_{i=1}^{N}y_i$ and $\bar{X}=(1/N)\sum_{i=1}^{N}x_i$ be the population means of the study and the auxiliary variable, respectively. And

$S_y^2=(1/N-1)\sum_{i=1}^{N}(y_i-\bar{Y})^2$, $S_x^2=(1/N-1)\sum_{i=1}^{N}(x_i-\bar{X})^2$ be the corresponding population mean square error of the study and the auxiliary variable, respectively and let $C_y=\frac{S_y}{\bar{Y}}$ and $C_x=\frac{S_x}{\bar{X}}$ be the coefficient of variation of the study as well as auxiliary variable respectively, and $\rho_{yx}=\frac{S_{yx}}{S_yS_x}$ be the population correlation coefficient between x and y.

In order to estimate the unknown population parameters we take a random sample of size n units from the finite population U by using simple random sample without replacement. Let $\bar{y}=(1/n)\sum_{i=1}^{n}y_i$ and $\bar{x}=(1/n)\sum_{i=1}^{n}x_i$ be the corresponding sample means of the study and the auxiliary variable, respectively and their corresponding sample variances are $\hat{S}_y^2=(1/n-1)\sum_{i=1}^{n}(y_i-\bar{y})^2$ and $\hat{S}_x^2=(1/n-1)\sum_{i=1}^{n}(x_i-\bar{x})^2$ respectively.

The usual unbiased estimator for population mean $\bar{Y}$ of the study variable is, given by

$$\bar{y}_0=\frac{\sum_{i=1}^{n}y_i}{n}. \tag{1.1}$$

The variance of the estimator $\bar{y}$ up to first order of approximation is, given by

$$Var(\bar{y}_0)=\theta S_y^2, \tag{1.2}$$

Where $\theta=\frac{1}{n}-\frac{1}{N}$.

In many real data sets, when we want to estimate the unknown characteristic of the population there exists some large ($y_{\max}$) or small values ($y_{\min}$) and to estimate the population quantities without considering these information is very sensitive in either the case the result will be either over estimated or under estimated. In order to tackle this situation Sarndal [3], proposed the following unbiased estimator for finite population mean

$$\bar{y}_S=\begin{cases}\bar{y}_0+c \text{ if sample contains } y_{\min} \text{ but not } y_{\max}\\ \bar{y}_0-c \text{ if sample contains } y_{\max} \text{ but not } y_{\min}\\ \bar{y}_0 \quad \text{for all other samples,}\end{cases}. \tag{1.3}$$

where c is a constant, whose values is to be find for minimum variance.

The minimum variance of the estimator $\bar{y}_S$ up to first order of approximation is, given as

$$\operatorname{var}\left(\bar{y}_{s}\right)_{\min}=\operatorname{var}\left(\bar{y}\right)-\frac{\theta\left(y_{\max}-y_{\min}\right)^{2}}{2\left(N-1\right)},\qquad(1.4)$$

where the optimum value of c_{opt} is $c_{opt}=\frac{\left(y_{\max}-y_{\min}\right)}{2n}$.

When the populations mean of the auxiliary variable is known the usual classical ratio estimator for finite population mean of the study variable is, given by

$$\bar{y}_{r}=\bar{y}\frac{\bar{X}}{\bar{x}}.\qquad(1.5)$$

The bias and mean square error of the estimator $\bar{y}_{r}$ up to first order of approximation are, given by

$$Bias\left(\bar{y}_{r}\right)=\frac{\theta}{\bar{X}}\left(RS_{x}^{2}-S_{yx}\right),\qquad(1.6)$$

$$MSE\left(\bar{y}_{r}\right)=\theta\left(S_{y}^{2}+R^{2}S_{x}^{2}-2RS_{yx}\right),\qquad(1.7)$$

where $R=\frac{\bar{Y}}{\bar{X}}$.

Singh and Tailor [4], suggested the following ratio estimator $\bar{y}_{st}$ for finite population mean of the study variable using the knowledge of the auxiliary variable

$$\bar{y}_{st}=\bar{y}\left(\frac{\bar{X}+\rho_{yx}}{\bar{x}+\rho_{yx}}\right).\qquad(1.8)$$

The bias and mean square error of the estimator $\bar{y}_{st}$ up to first order of approximation are, given by

$$Bias\left(\bar{y}_{st}\right)=\frac{\theta\alpha_{st}}{\bar{Y}}\left(\alpha_{st}S_{x}^{2}-S_{yx}\right),\qquad(1.9)$$

$$MSE\left(\bar{y}_{st}\right)=\theta\left(S_{y}^{2}+\alpha_{st}^{2}S_{x}^{2}-2\alpha_{st}S_{yx}\right),\qquad(1.10)$$

where $\alpha_{st}=\frac{\bar{Y}}{\bar{X}+\rho_{yx}}$.

When the population coefficient of variation of the auxiliary variable is known Sisodia and Dwivedi [5], suggested the following ratio estimator $\bar{y}_{sd}$ for the study variable *y* and is given by

$$\bar{y}_{sd}=\bar{y}\left(\frac{\bar{X}+C_{x}}{\bar{x}+C_{x}}\right),\qquad(1.8)$$

where C_x is the known value of the population coefficient variation of the auxiliary variable.

The bias and mean square error of the estimator $\bar{y}_{sd}$ up to first order of approximation are, given by

$$Bias\left(\bar{y}_{sd}\right)=\frac{\theta\alpha_{sd}}{\bar{Y}}\left(\alpha_{sd}S_{x}^{2}-S_{yx}\right),\qquad(1.9)$$

$$MSE\left(\bar{y}_{sd}\right)=\theta\left(S_{y}^{2}+\alpha_{sd}^{2}S_{x}^{2}-2\alpha_{sd}S_{yx}\right),\qquad(1.10)$$

where $\alpha_{sd}=\frac{\bar{Y}}{\bar{X}+C_{x}}$.

Kadilar and Cingi [1], propose the following class of ratio estimators $\bar{y}_{kci}$ for finite population mean of the study variable using the knowledge of the auxiliary variable,

$$\bar{y}_{kc1}=\bar{y}\left(\frac{\bar{X}C_{x}+\rho_{yx}}{\bar{x}C_{x}+\rho_{yx}}\right),\qquad(1.11)$$

$$\bar{y}_{kc2}=\bar{y}\left(\frac{\bar{X}\rho_{yx}+C_{x}}{\bar{x}\rho_{yx}+C_{x}}\right),\qquad(1.12)$$

$$\bar{y}_{kc3}=\bar{y}\left(\frac{\bar{X}\beta_{2x}+\rho_{yx}}{\bar{x}_{C_2}\beta_{2x}+\rho_{yx}}\right),\qquad(1.13)$$

$$\bar{y}_{kc4}=\bar{y}\left(\frac{\bar{X}\rho_{yx}+\beta_{2x}}{\bar{x}\rho_{yx}+\beta_{2x}}\right),\qquad(1.14)$$

where β_{2x} is the known value of the population coefficient of kurtosis of the auxiliary variable.

The bias and mean square error of the estimators $\bar{y}_{kci}$ up to first order of approximation are, given by

$$Bias\left(\bar{y}_{kci}\right)=\frac{\theta\alpha_{kci}}{\bar{Y}}\left(\alpha_{kci}S_{x}^{2}-S_{yx}\right),\qquad(1.15)$$

$$MSE\left(\bar{y}_{kci}\right)=\theta\left[\left(S_{y}^{2}+\alpha_{kci}^{2}S_{x}^{2}-2\alpha_{kci}S_{yx}\right)\right],\qquad(1.16)$$

for $i=1,2,3,4$.

where $\alpha_{kc1}=\frac{\bar{Y}C_{x}}{\bar{X}C_{x}+\rho_{yx}}$, $\alpha_{kc2}=\frac{\bar{Y}\rho_{yx}}{\bar{X}\rho_{yx}+C_{x}}$,

$\alpha_{kc3}=\frac{\bar{Y}\beta_{2x}}{\bar{X}\beta_{2x}+\rho_{yx}}$ and $\alpha_{kc4}=\frac{\bar{Y}\rho_{yx}}{\bar{X}\rho_{yx}+\beta_{2x}}$.

2. The Proposed Class of Ratio-Type Estimators

On the lines of Sarndal [3], we proposed a class of ratio-type estimators for the estimation of finite population mean of the study variable *y*, using the known knowledge of an auxiliary variable say *x*. Usually when the relationship between the study and the auxiliary variable is positive then the selection of the larger value of the auxiliary variable the larger the value of the study variable is to be expected, and the selection of the smaller value of the auxiliary variable the smaller the value of the study variable is to be expected. Using the above information we proposed the following class of ratio-type estimators

$$\bar{y}_{p1}=\bar{y}_{C_1}\left(\frac{\bar{X}+M_{x}}{\bar{x}_{C_2}+M_{x}}\right),\qquad(2.1)$$

$$\bar{y}_{p2}=\bar{y}_{C_1}\left(\frac{\bar{X}\rho_{yx}+M_{x}}{\bar{x}_{C_2}\rho_{yx}+M_{x}}\right),\qquad(2.2)$$

$$\bar{y}_{p3} = \bar{y}_{C_1}\left(\frac{\bar{X}C_x + M_x}{\bar{x}_{C_2}C_x + M_x}\right), \quad (2.3)$$

$$\bar{y}_{p4} = \bar{y}_{C_1}\left(\frac{\bar{X}\beta_{2x} + M_x}{\bar{x}_{C_2}\beta_{2x} + M_x}\right), \quad (2.4)$$

where M_x and β_{2x} are the known values of the population median and population coefficient of kurtosis of the auxiliary variable respectively and $\left(\bar{y}_{C_1} = \bar{y} + c_1, \quad \bar{x}_{C_2} = \bar{x} + c_2\right)$, where c_1 and c_2 are unknown constants whose values are to be find for optimality conditions.

To obtain the properties of $\bar{y}_{Pi}$ in terms of bias and Mean square error, we define the following relative error terms and their expectations.

$\zeta_0 = \frac{\bar{y}_{C_1} - \bar{Y}}{\bar{Y}}, \quad \zeta_1 = \frac{\bar{x}_{C_2} - \bar{X}}{\bar{X}}$, such that

Such that $E(\zeta_0) = E(\zeta_1) = 0.$

Also $E(\zeta_0^2) = \frac{\theta}{\bar{Y}^2}\left(S_y^2 - \frac{2nc_1}{N-1}(y_{\max} - y_{\min} - nc_1)\right),$

$$E(\zeta_1^2) = \frac{\theta}{\bar{X}^2}\left(S_x^2 - \frac{2nc_2}{N-1}(x_{\max} - x_{\min} - nc_2)\right),$$

and

$$E(\zeta_0\zeta_1) = \frac{\theta}{\bar{Y}\bar{X}}\left(S_{yx} - \frac{n}{N-1}\left(c_2(y_{\max} - y_{\min}) + c_1(x_{\max} - x_{\min}) - 2nc_1c_2\right)\right).$$

The biases and mean square errors of the estimators $\bar{y}_{pi}$ up to first order of approximation are, given by

$$Bias(\bar{y}_{pi}) = \frac{\theta\alpha_{pi}}{\bar{Y}}\left[\begin{array}{l}\alpha_{pi}\left(S_x^2 - \frac{2nc_2}{N-1}(x_{\max} - x_{\min} - nc_2)\right) - S_{yx} \\ + \frac{n}{N-1}\left(c_2(y_{\max} - y_{\min}) + c_1(x_{\max} - x_{\min}) - 2nc_1c_2\right)\end{array}\right], \quad (2.5)$$

$$MSE(\bar{y}_{pi})_{\min} = \theta\left[\begin{array}{l}\left(S_y^2 + \alpha_{pi}^2 S_x^2 - 2\alpha_{pi}S_{yx}\right) \\ -\frac{1}{2(N-1)}\left((y_{\max} - y_{\min}) - \alpha_{pi}(x_{\max} - x_{\min})\right)^2\end{array}\right], \quad (2.6)$$

where the optimum values of c_1 and c_2 are given by

$c_1 = \frac{(y_{\max} - y_{\min})}{2n}$ and $c_2 = \frac{(x_{\max} - x_{\min})}{2n}$.

Also $\alpha_{p1} = \frac{\bar{Y}}{\bar{X} + M_x}$, $\alpha_{p2} = \frac{\bar{Y}\rho_{yx}}{\bar{X}\rho_{yx} + M_x}$,

$\alpha_{p3} = \frac{\bar{Y}C_x}{\bar{X}C_x + M_x}$ and $\alpha_{p4} = \frac{\bar{Y}\beta_{2x}}{\bar{X}\beta_{2x} + M_x}$.

3. Comparison of Estimators

In this section, we have found some efficiency comparison conditions under which the proposed class of ratio-type estimators have perform better than the other existing estimators discussed in the literature.

(i) By (1.2) and (2.6), we have

$\left[MSE(\bar{y}_0) - MSE(\bar{y}_{pi})_{\min}\right] \geq 0$, if

$$\left[\begin{array}{l}\frac{1}{2(N-1)}\left\{(y_{\max} - y_{\min}) - \alpha_{Pi}(x_{\max} - x_{\min})\right\}^2 \\ -\alpha_{Pi}^2 S_x^2 + 2\alpha_{Pi}S_{yx}\end{array}\right] \geq 0,$$

(ii) By (1.4) and (2.6), we find

$\left[MSE(\bar{y}_r) - MSE(\bar{y}_{pi})_{\min}\right] \geq 0$, if

$$\left[\begin{array}{l}\frac{1}{2(N-1)}\left\{(y_{\max} - y_{\min}) - \alpha_{Pi}(x_{\max} - x_{\min})\right\}^2 \\ -S_x^2(\alpha_{Pi}^2 - R) + 2S_{yx}(\alpha_{Pi} - R)\end{array}\right] \geq 0,$$

(iii) By (1.10) and (2.6), we see that

$\left[MSE(\bar{y}_{sd}) - MSE(\bar{y}_{pi})_{\min}\right] \geq 0$, if

$$\left[\begin{array}{l}\frac{1}{2(N-1)}\left\{(y_{\max} - y_{\min}) - \alpha_{Pi}(x_{\max} - x_{\min})\right\}^2 \\ -S_x^2(\alpha_{Pi}^2 - \alpha_{sd}^2) + 2S_{yx}(\alpha_{Pi} - \alpha_{sd})\end{array}\right] \geq 0,$$

(iv) By (1.16) and (2.6), we also see that

$\left[MSE(\bar{y}_{kci}) - MSE(\bar{y}_{pi})_{\min}\right] \geq 0$, if

$$\left[\begin{array}{l}\frac{1}{2(N-1)}\left\{(y_{\max} - y_{\min}) - \alpha_{Pi}(x_{\max} - x_{\min})\right\}^2 \\ -S_x^2(\alpha_{Pi}^2 - \alpha_{kci}^2) + 2S_{yx}(\alpha_{Pi} - \alpha_{kci})\end{array}\right] \geq 0,$$

for *i*=1, 2, 3, 4.

4. Empirical Study

To analyze the performance of the proposed class of ratio-type estimators with some of the existing estimators discussed in the literature of survey sampling, we have considered a real data set. The description and the necessary data statistics of the population are given below.

For the percent relative efficiencies (*PRE's*) of the existing and the proposed class of ratio-type estimators, we have used the following expression for efficiency comparison

$$PRE(\bar{y}_g, \bar{y}_0) = \frac{MSE(\bar{y}_0)}{MSE(\bar{y}_g)} \times 100,$$

where *g=0,s*, *r*, *st*, *sd*, *kc1*, *kc2*, *kc3*, *kc4*, *p1*, *p2*, *p3* and *p4*.

Population-1:[Source: Agricultural Statistics (1999), [2] Washington, US]

Y: Amount (in \$000) of real estate farm loans in different states during 1997;

X: Amount (in $000) of non-real estate farm loans in different states during 1997.

$N=50,\quad n=12,\quad \bar{X}=878.16,\quad \bar{Y}=555.43,\quad S_y^2=342021.5,$

$S_x^2=1176526,\quad S_{yx}=509910.4,\quad y_{max}=2327.025,$

$y_{min}=1.611,\quad x_{max}=3928.732,\quad x_{min}=0.233,\quad \rho_{yx}=0.804,$

$M_y=322.305,\quad M_x=452.517,\quad C_y=1.0529,\quad C_x=1.2352,$

$\beta_{2x}=4.5247.$

The mean squared errors and the percent relative efficiencies of the existing and the proposed class of ratio-type mean estimators are shown in **Table-1.**

Table 1.*MSE* and *PRE* of the competing and the proposed class of ratio-type estimators

Estimator		Population 1	
		$MSE(\bar{y}_g)$	$PRE(\bar{y}_g, \bar{y}_0)$
Existing	$\bar{y}_0$	21649.9610	100.00
	$\bar{y}_S$ [2]	18157.1249	119.2367
	$\bar{y}_r$	10612.3030	204.0081
	$\bar{y}_{st}$ [4]	10596.5563	204.3113
	$\bar{y}_{sd}$ [5]	10586.7659	204.5002
	$\bar{y}_{kc1}$ [1]	10598.1592	204.2804
	$\bar{y}_{kc2}$ [1]	10582.9727	204.5735
	$\bar{y}_{kc3}$ [1]	10606.7051	204.1158
	$\bar{y}_{kc4}$ [1]	10494.6271	206.2957
Proposed	$\bar{y}_{p1}$ Proposed	7374.5707	**293.5759**
	$\bar{y}_{p2}$ Proposed	7412.2246	**292.0845**
	$\bar{y}_{p3}$ Proposed	7463.0311	**290.0961**
	$\bar{y}_{p4}$ Proposed	9002.3826	**240.4915**

5 CONCLUSION

In this study, we have suggested an improved class of ratio-type estimators for finite population mean, under maximum and minimum values using the knowledge of the of the auxiliary variable and have discussed their properties in simple random sampling. We have also found some theoretical conditions under which the suggested class of ratio-type estimators have always efficient than the usual unbiased, Sarndal [3], the classical ratio estimator, Singh and Tailor [4], Sisodia and Dwivedi [5] and Kadilar and Cingi [1] estimators. Theoretical results are also verified with the help of real data set which clearly indicates that the proposed class of ratio-type estimators has smaller mean squared error and higher percent relative efficiency from others the other estimators discussed in the literature. Thus the proposed strategy under maximum and minimum values may be preferred over the existing estimators for the use of practical applications.

REFERENCES

1. C. Kadilar, H. Cingi, An improvement in estimating the population mean by using the correlation coefficient, Hacettepe Journal of mathematics and Statistics, 35, (2006), 103-109.
2. http://www.nass.usda.gov/Publications/Ag_Statistics/1999/index.asp
3. Sarndal, C. E. Sample survey theory vs general statistical theory: Estimation of the population mean, International Statistical Institute, 40, (1972), 1-12.
4. H. P. Singh, R. Tailor, Use of known correlation coefficient in estimating the finite population mean, Statistics in Transition, 6, (2003), 555-560.
5. B. V. S. Sisodia, V. K. A. Dwivedi, Modified ratio estimator using coefficient of variation of auxiliary variable, Journal of the Indian Society of Agricultural Statistics, 33, (1981), 13-18.