Efficient Class of Ratio cum Regression Estimators using Non-Conventional Parameters for Estimating Population mean

Asra Nazir[1], Rafia Jan[2], T. R. Jan[3]

1,2,3−Department of Statistics, University of Kashmir, J&K, India

E-mail: 1: asranazir44@gmail.com, 2: rafiajan836@gmail.com, 3: drtrjan@gmail.com

Abstract

The article is based on ratio cum regression estimators using non –conventional parameters The expression for MSE is derived. It has proved hypothetically as well as empirically by comparing MSE and PRE's that the revised class is enhanced than the existing estimator.

Keywords: Auxiliary variable, Bias, Mean Square Error (MSE), Percent Relative Efficiency (PRE's).

1.Introduction

Many estimators have been given by various authors for improved estimation procedure. The Ratio estimator was given by Cochran in (1940) while Murthy(1967) gave product estimator. The latest references on ratio cum regression is given by ,Kadilar and Cingi (2004),Yan(2010).Yadav, Mishra and Kumar(2014)(2015) (2016).

Consider a population of N distinct units and let (x_i, y_i), $i=1,2,3,\ldots\ldots n$ be a two variables of sample of size n .Let $\overline{X}$ and $\overline{Y}$ be the population means and let $\overline{x}$ and $\overline{y}$ be the sample means of $\overline{X}$ and $\overline{Y}$ respectively. Let ρ the correlation coefficient between the two variables

2. Notations

$N =$ Size of the population

$n =$ Size of the sample

$X =$ auxiliary variable

$Y =$ study variable

$\overline{X}, \vec{Y} =$ population means

$\overline{x}, \overline{y} =$ sample means

$S_x, S_y =$ standard deviation

$S_{xy} =$ covariance between x and y

$C_x, C_y =$ coefficient of variation

M_d =median of auxiliary variable

ρ = correlation coefficient between x and y

$b = \frac{S_{xy}}{S_x^2}$ Regression coefficient of y on x

$k = \frac{\overline{X}}{\overline{Y}}$

$B = \frac{S_{xy}}{S_x^2}$

$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{X})^3}{(N-1)(N-2)S_x^3}$ Coefficient of Skewness,

$\beta_2 = \frac{N(N+1)\sum_{i=1}^{N}(x_i - \overline{X})^4}{(N-1)(N-2)(N-3)S_x^4} - \frac{3(N-1)^2}{(N-1)(N-2)}$ Coefficient of Kurtosis,

$QD = \frac{Q_3 - Q_1}{2}$ Quartile Deviation,

$G = \frac{4}{(N-1)}\sum_{i=1}^{N}\left[\frac{2i - N - 1}{2N}\right]X_i$ - Gini means diffeence

$D = \frac{2\sqrt{\pi}}{(N-1)}\sum_{i=1}^{N}\left[i - \frac{N+1}{2}\right]X_i$ - Downtown parameters

$S_{pw} = \frac{\sqrt{\pi}}{N^2}\sum_{i=1}^{N}[2i - N - 1]X_i$ - probability weighted moments

$B(.)$ = Bias of the estimator

$V(.)$ = Variance of the estimator

$MSE(.)$ = Mean Square Errror of the estimator

3. Estimators in Literature

The estimator of population mean is given by,

$$t_o = \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

It is unbiased and variance is given by,

$$V(t_o) = \frac{1-f}{n} S_y^2 \qquad (1)$$

Cochran [2] gave the ratio estimator and is given as

$$t_R = \bar{y}\frac{\bar{X}}{\bar{x}}$$

The bias(B) and mean squared error(MSE) of this estimator,is given by,

$$Bias(t_R) == \frac{1-f}{n}\frac{1}{\bar{X}}\left(R_1 S_x^2 - \rho S_x S_y\right)$$

$$MSE(t_R) == \frac{1-f}{n}\left(S_y^2 + R_1 S_x^2 - 2R_1 \rho S_x S_y\right) \qquad (2)$$

Where $R_1 = \frac{\bar{Y}}{\bar{X}}$

Following Abid ,Abbas ,Sherwani and Nazir(2016) for estimation procedure, a specific parameter as ratio of correlation and skewness along with some non-traditional parameters is given as,

$$t_{p1} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\tau\bar{x} + G)}\left(\tau\bar{X} + G\right) \qquad (3)$$

$$t_{p2} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\tau\bar{x} + D)}\left(\tau\bar{X} + D\right) \qquad (4)$$

$$t_{p3} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\tau\bar{x} + S_{pw})}\left(\tau\bar{X} + S_{pw}\right) \qquad (5)$$

where, $\tau = \rho / B_1$

The biases(B.) and the mean squared errors(MSE) of proposed estimators are given by,

$$Bias(t_{pj}) == \frac{1-f}{n}\frac{S_x^2}{\bar{Y}} R_{pj}, \left(j = 1,2,3\right) \qquad (6)$$

$$MSE(t_{pj}) == \frac{1-f}{n}\left(R_{pj} S_x^2 + S_y^2(1-\rho^2)\right) \qquad (7)$$

Where $R_{p1} = \frac{\bar{Y}\tau}{\bar{X}\tau + G}$, $R_{p2} = \frac{\bar{Y}\tau}{\bar{X}\tau + D}$, $R_{p3} = \frac{\bar{Y}\tau}{\bar{X}\tau + S_{pw}}$

4.Proposed Estimator

We have extended the work given by Abid et .al (2016) by suggested three new improved class of Ratio cum Regression estimators.

$$t_1^* = \frac{\alpha\bar{y}(\tau\bar{X}+G)}{(\tau\bar{x}+G)} + \frac{\beta(\bar{y}+b(\bar{X}-\bar{x}))}{\bar{x}}\bar{X} \quad (8)$$

$$t_2^* = \frac{(\bar{y}+b(\bar{X}-\bar{x}))}{\bar{x}}\exp\left(\frac{(\tau\bar{X}+D_M)}{(\tau\bar{x}+D_M)}-1\right) \quad (9)$$

$$t_3^* = \frac{\bar{y}+b(\bar{X}-\bar{x})}{(\tau\bar{x}+MR)}\left(\tau\bar{X}+MR\right) \quad (10)$$

Such that $\alpha+\beta=1$

In order to get the bias(B.) and the mean squared error(MSE) for the proposed ratio estimator,

Let

$$\bar{x} = \bar{X}(1+e_1);\ \bar{y} = \bar{Y}(1+e_0) \quad (11)$$

$$E(e_o) = E(e_1) = 0$$

$$E(e_o^2) = \frac{1-f}{n}C_y^2, E(e_1^2) = \frac{1-f}{n}C_x^2,$$

$$E(e_0e_1) = \frac{1-f}{n}\rho C_x C_y$$

Therefore, expressing (8) in terms of (11),we obtain

$$y_{t1}^* = \frac{\alpha\bar{Y}(1+e_0)(\bar{X}+G)}{(\bar{x}+G)} + (1-\alpha)\left[\frac{\alpha\bar{Y}(1+e_0)+b[\bar{X}-\bar{X}(1+e_1)}{\bar{X}(1+e_1)}\right]\bar{X}$$

$$= \frac{\alpha\bar{Y}(1+e_0)(\bar{X}+G)}{(\bar{X}+G)+\bar{X}e_1} + (1-\alpha)\left[\bar{Y}(1+e_0)(1+e_1)^{-1} + b\bar{X}e_1 - (1+e_1)^{-1}\right]$$

$$= \alpha\bar{Y}(1+e_0)(1+\theta e_1)^{-1} + \bar{Y}(1-e_1+e_0-e_0e_1) - \alpha\bar{Y}(\alpha-\alpha e_1+\alpha e_0-\alpha e_0e_1) -$$
$$b\bar{X}e_1 + b\bar{X}e_1^2 + \alpha b\bar{X}e_1 - \alpha b\bar{X}e_1^{\ 2})$$

The bias is obtained as

$$B(\bar{y}_{t1}^{\ *}) = E(\bar{y}_{t1}^{\ *} - \bar{Y})$$

$$= E\left[\bar{Y}\left(1 + e_0 + (\alpha BK - BK - \alpha - \alpha\theta\right)e_1 + (\alpha - \alpha\theta - 1)e_0 e_y + (\alpha\theta + 1 - \alpha + BK - \alpha BK)e_1^{\ 2} - \bar{Y})\right]$$

$$= \frac{1-f}{n}\bar{Y}\left[(\alpha - \alpha\theta - 1)\rho C_y C_x + (\alpha\theta^2 + 1 - \alpha + BK - \alpha BK)C_x^2\right] \quad (12)$$

The MSE is given by:

$$MSE(\bar{y}_{t1}^{\ *}) = E(\bar{y}_{t1}^{\ *} - \bar{Y})^2$$

$$= E\left[\bar{Y}\left(1 + e_0 + (\alpha BK - BK - \alpha - \alpha\theta\right)e_1 + (\alpha - \alpha\theta - 1)e_0 e_y + (\alpha\theta + 1 - \alpha + BK - \alpha BK)e_1^{\ 2} - \bar{Y})\right]$$

$$= \frac{1-f}{n}\bar{Y}^2\left[C_y^2 + 2(\alpha BK - BK - \alpha - \alpha\theta^2)\rho C_y C_x + (\alpha BK - BK - \alpha - \alpha\theta)^2 C_x^2\right] \quad (13)$$

Differentiate eq.(13) with respect to α we get

$$\alpha = \frac{(BK+1)C_x^2 - \rho C_y C_x}{(BK + 1 - \theta)C_x^2}$$

Substituting gives the optimal MSE

$$MSE(t_1^{\ *}) = \frac{1-f}{n}[S_y^2(1-\rho^2)] \quad (14)$$

And,expressing (9) in terms of e's,we obtain

$$t_2^* = \bar{Y}(1+e_0) + b(\bar{X} - \bar{X}(1+e_1))\exp\left[\frac{\tau\bar{X} + D_M}{\tau\bar{X}(1+e_1) + D_M}\right] - 1$$

$$= \bar{Y}(1+e_0) + b\bar{X} - b\bar{X} - b\bar{X}e_1 \exp[1 + \theta e_1 - 1]^{-1}$$

$$= \left(\bar{Y}(1+e_0) - b\bar{X}e_1\right)[1 - \theta e_1] \quad (15)$$

Where $\theta = \dfrac{\tau\bar{X}}{\tau\bar{X} + D_M}$

Taking expectations and squaring above equation and ignoring higher terms more than two,we get

$$E(t_2^* - \bar{Y})^2 = \left(\bar{Y}(e_0 - \theta e_1) - b\bar{X}e_1\right)^2$$

$$= \left(\bar{Y}(e_0 - \theta e_1)\right)^2 + b^2\bar{X}^2 e_1^{\ 2} - 2\bar{Y}(e_0 - \theta e_1)(-b\bar{X}\ddot{e}_1)$$

$$= \left(\bar{Y}^2 e_0^{\ 2} + \bar{Y}^2\theta^2 e_1^{\ 2} - 2\bar{Y}^2\theta e_0 e_1 + b^2\bar{X}^2 e_1^{\ 2} + 2\bar{Y}b\bar{X}e_0 e_1 - 2\bar{Y}b\bar{X}\theta e_1^{\ 2}\right)$$

$$=\left(\overline{Y}^2 C_y^{\ 2}+\overline{Y}^2\theta^2 C_x^{\ 2}-2\overline{Y}^2\theta\rho C_y C_x+b^2\overline{X}^2 C_x^{\ 2}+2\overline{Y}b\overline{X}\rho C_y C_x-2\overline{Y}b\overline{X}\theta C_x^{\ 2}\right) \quad (16)$$

On further simplification of eq.(16),we get mean square error(*MSE)*

$$MSE(t_2^*)=\left(\overline{Y}^2 C_y^{\ 2}+\overline{Y}^2\theta^2 C_x^{\ 2}-2\overline{Y}^2\theta\rho C_y C_x+b^2\overline{X}^2 C_x^{\ 2}+2\overline{Y}b\overline{X}\rho C_y C_x-2\overline{Y}b\overline{X}\theta C_x^{\ 2}\right)$$

$$=\frac{1-f}{n}[R^2\theta^2 S_x^2+S_y^2-\frac{S_{xy}}{S_x^2}]$$

$$=\frac{1-f}{n}[R_{t2}^{*\ 2} S_x^2+S_y^2(1-\rho^2)] \quad (17)$$

$$R_{t2}^{\ *}=\frac{\overline{Y}\tau}{\overline{X}\tau+D_M}$$

Now ,expressing (10) in terms of (11),we obtain

$$t_3^*=\overline{Y}(1+e_0)+b(\overline{X}-\overline{X}(1+e_1))\left[\frac{\tau\overline{X}+MR}{\tau\overline{X}(1+e_1)+MR}\right]$$

$$=\overline{Y}(1+e_0)+b\overline{X}-b\overline{X}-b\overline{X}e_1\left[1+\theta e_1\right]^{-1}$$

$$=\left(\overline{Y}(1+e_0)-b\overline{X}e_1\right)\left[1-\theta e_1\right] \quad (18)$$

Where $\theta=\frac{\tau\overline{X}}{\tau\overline{X}+TM}$

Taking expectations and squaring on both sides of eq.(18) and ignoring higher terms more than two,we get

$$E(t_3^*-\overline{Y})^2=\left(\overline{Y}(e_0-\theta e_1)-b\overline{X}e_1\right)^2$$

$$=\left(\overline{Y}(e_0-\theta e_1)\right)^2+b^2\overline{X}^2 e_1^{\ 2}-2\overline{Y}(e_0-\theta e_1)(-b\overline{X}\ddot{e}_1)$$

$$=\left(\overline{Y}^2 e_0^{\ 2}+\overline{Y}^2\theta^2 e_1^{\ 2}-2\overline{Y}^2\theta e_0 e_1+b^2\overline{X}^2 e_1^{\ 2}+2\overline{Y}b\overline{X}e_0 e_1-2\overline{Y}b\overline{X}\theta e_1^{\ 2}\right)$$

$$=\left(\overline{Y}^2 C_y^{\ 2}+\overline{Y}^2\theta^2 C_x^{\ 2}-2\overline{Y}^2\theta\rho C_y C_x+b^2\overline{X}^2 C_x^{\ 2}+2\overline{Y}b\overline{X}\rho C_y C_x-2\overline{Y}b\overline{X}\theta C_x^{\ 2}\right) \quad (19)$$

On further simplification of eq.(19),we get mean square error(*MSE)*

$$MSE(t_3^*)=\left(\overline{Y}^2 C_y^{\ 2}+\overline{Y}^2\theta^2 C_x^{\ 2}-2\overline{Y}^2\theta\rho C_y C_x+b^2\overline{X}^2 C_x^{\ 2}+2\overline{Y}b\overline{X}\rho C_y C_x-2\overline{Y}b\overline{X}\theta C_x^{\ 2}\right)$$

$$=\frac{1-f}{n}[R^2\theta^2 S_x^2+S_y^2-\frac{S_{xy}}{S_x^2}]$$

$$= \frac{1-f}{n}[R_{t3}^{*\,2} S_x^2 + S_y^2(1-\rho^2)] \quad (20)$$

$$R_{t3}^{\ *} = \frac{\overline{Y}\tau}{\overline{X}\tau + MR}$$

5. Efficiency Comparison

The comparison is made between the mean square error(MSE) of the advised estimator with the classical .

$$V(t_o) - MSE(t_1^*) = \frac{1-f}{n} S_y^2 \rho^2 \geq 0$$

$$V(t_o) - MSE(t_2^*) = S_y^2 \rho^2 - R_{t2}^* S_x^2 \geq 0$$

$$V(t_o) - MSE(t_3^*) = S_y^2 \rho^2 - R_{t3}^* S_x^2 \geq 0$$

$$MSE(t_R) - MSE(t_1^*) = S_y^2 \rho^2 + R_1^2 S_x^2 - 2R_1 \rho S_x S_y \geq 0$$

$$MSE(t_R) - MSE(t_2^*) = [R_1^2 S_x^2 - 2R_1 \rho S_x S_y] - [R_{t2}^* S_x^2 - S_y^2 \rho^2] \geq 0$$

$$MSE(t_R) - MSE(t_3^*) = [R_1^2 S_x^2 - 2R_1 \rho S_x S_y] - [R_{t3}^* S_x^2 - S_y^2 \rho^2] \geq 0$$

$$MSE(t_{P1}) - MSE(t_1^*) = [R_{p1} S_x^2] \geq 0$$

$$MSE(t_{P2}) - MSE(t_2^*) = [R_{p2} S_x^2 - R_{t2}^{\ *} S_x^2] \geq 0$$

$$MSE(t_{P3}) - MSE(t_3^*) = [R_{p3} S_x^2 - R_{t3}^{\ *} S_x^2] \geq 0$$

6. Numerical illustration

For calculating MSE,we have taken data of Murthy [1967] in which fixed capital is denoted by X (auxiliary variable) and output of 80 factories are denoted by Y (study variable approved project).

Table 1: Data Statistics

Parameters	Population	Parameters	Population
N	34	C_x	0.7205
n	20	β_2	0.0978
$\overline{Y}$	856.4117	β_1	0.9782
$\overline{X}$	208.8823	TM	162.25
ρ	0.4491	MR	284.5

M_d	150	G	155.446
S_y	733.1407	D	140.891
C_y	0.8561	S_{pw}	199.961
S_x	150.5059	DM	234.82

7. Conclusion

From the below table, we infer that the recommended estimators having less mean squared errors(MSE)and higher PRE's as compared to all classical estimators presented in the literature. So, the suggested estimators are more competent than the other estimators.

Table 2.MSE of Proposed and PRE's of the estimators.

Estimators	***MSE***	***PRE's***
$t_o = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$	10749.8	100.00
$t_R = \bar{y}\frac{\bar{X}}{\bar{x}}$	10221.28	105.17
$t_{p1} = \frac{\bar{y}+b(\bar{X}-\bar{x})}{(\tau\bar{x}+G)}\left(\tau\bar{X}+G\right)$	9656.16	110.87
$t_{p2} = \frac{\bar{y}+b(\bar{X}-\bar{x})}{(\tau\bar{x}+D)}\left(\tau\bar{X}+D\right)$	9741.5	110.35
$t_{p3} = \frac{\bar{y}+b(\bar{X}-\bar{x})}{(\tau\bar{x}+S_{pw})}\left(\tau\bar{X}+S_{pw}\right)$	9347.3	115.00
$t_1^* = \frac{\alpha\bar{y}(\tau\bar{X}+G)}{(\tau\bar{x}+G)} + \frac{\beta(\bar{y}+b(\bar{X}-\bar{x}))}{\bar{x}}\bar{X}$	**8581.73**	**125.26**
$t_2^* = \frac{(\bar{y}+b(\bar{X}-\bar{x}))}{\bar{x}}\exp\left(\frac{(\tau\bar{X}+D_M)}{(\tau\bar{x}+D_M)}-1\right)$	9201.8	116.82
$t_3^* = \frac{\bar{y}+b(\bar{X}-\bar{x})}{(\tau\bar{x}+MR)}\left(\tau\bar{X}+MR\right)$	9043.8	118.86

References

[1] M.Abid, N. Abbas, R.A.K. Sherwani, and H.Z. Nazir, "Improved Ratio Estimators for the Population Mean Using Non-Conventional Measures of Dispersion". Pakistan Journal of Statistics and Operations Research, Vol.2**,(2016),** pp. 353-367.

[2] W.G. Cochran, Sampling Techniques. 3rd Edition, Wiley Eastern Limited, New Delhi (1940).

[3] S.K. Yadav, S.S. , Mishra and S. Kumar "Optimal Search for Efficient Estimator of Finite Population Mean Using Auxiliary Information". American Journal of Operational Research, Vol.4**, (2014)** .pp.28-34.

[4] S.K. Yadav, S.S. Mishra, A.K. Shukla, , S Kumar. and R.S. Singh, "Use of Non-Conventional Measures of Dispersion for Improved Estimation of Population Mean." American Journal of Operational Research, Vol.6**, (2016)**, pp. 69-75.

[5] S.K. Yadav and S.S. Mishra "Developing Improved Predictive Estimator for Finite Population Mean Using Auxiliary Information". Statistika, Vol.95, **(2015)** ,pp.76-85.

[6]C. Kadilar and H. Cing Ratio Estimators in Simple Random Sampling. A pplied Mathematics and Computation, Vol.151**, (2004)**,pp. 893-902.

[7] Z. Yan and B. Tian "Ratio Method to the Mean Estimation Using Coefficient of Skewness of Auxiliary Variable." In: Zhu, R., Zhang, Y., Liu, B. and Liu, C., Eds., Information Computing and Applications. ICICA 2010. Communications in Co mputer and Information Science , Vol. 106, Springer, Berlin, Heidelberg, (**2010)**.pp.103-110. https://doi.org/10.1007/978-3-642-16339-5_14

[8] M.N. Murthy"Product method of estimation," Sankhya, A, Vol.26, (**1967),** pp. 69- 74.