Research & Reviews: Journal of Statistics and Mathematical Sciences

Some Improved Ratio Estimators for Estimating Mean of FinitePopulation

Muhammad Ijaz[1]*, Hameed Ali[2]

[1]Department of Statistics, University Of Peshawar, Pakistan

[2]Department of Statistics, Islamia College Peshawar, Pakistan

Research Article

Received date: 27/03/2018

Accepted date: 21/06/2018

Published date: 25/06/2018

***For Correspondence**

Department Of Statistics, University Of Peshawar, Pakistan

E-mail: ijaz.statisticstics@gmail.com

Keywords: Auxiliary information, Study variable, Bias, MSE, Percentage relative efficiency

ABSTRACT

In this paper we have proposed an efficient estimator for estimating the mean of finite population under simple random sampling schemes. We have proposed a modified ratio estimator whose efficiency is the same as of regression estimator. It is a well establish fact that linear regression estimator is more efficient than most of the ratio estimators. We have found the Bias and MSE up to first order of approximation. The conditions under which the proposed estimators perform well as compared to other estimators. These properties are supported by real data sets.

INTRODUCTION

The efficiency of an estimator can be increased largely, if we incorporate the auxiliary/benchmark variable(s) correlated with the study variable. Utilizing auxiliary information in a way so that the results become highly efficient. The use of auxiliary information is a challenging problem. Many statisticians use auxiliary information in their own way. It was Cochran who first uses auxiliary information in estimating the mean of finite population. Many other Statisticians make use of auxiliary information at estimation stages [1-6].

The Classical estimator of the mean of the finite population $\overline{y}$ is $\overline{y}$. This estimator is an unbiasedestimator of population mean and its variance is given by

$$V\left(\overline{y}\right) = \lambda C_y^2 \qquad (1.0)$$

Cochran [1] introduces the traditional ratio type estimator and is given by

$$\overline{y}_{cr} = \left(\frac{\overline{y}}{\overline{x}}\right)\overline{X} \qquad (1.1)$$

The Bias and MSE are

$$B\left(\overline{y}_{cr}\right) \cong \lambda \overline{Y}\left(C_x^2 - C_{yx}\right) \qquad (1.2)$$

And

$$MSE\left(\overline{y}_{cr}\right) \cong \lambda \overline{Y}^2\left(C_y^2 - C_x^2 - 2\rho C_y C_x\right) \qquad (1.3)$$

Where $C_y^2 = \frac{S_y^2}{\overline{Y}^2}$ the coefficient of variation of the study variable Y is, $C_{yx} = \frac{S_{yx}}{YX}$ is the coefficient ofvariation of the auxiliary variable X, $C_{yx} = \frac{S_{yx}}{YX}$ is the coefficient of covariance between the studyvariable and the auxiliary variable and $\rho = \frac{C_{yx}}{C_y C_x}$ is the coefficient of correlation between Y and X.Sisodia and Dwivedi introduces the ratio type estimator for estimating the mean of finitepopulation and is follows as,

$$\overline{y_{sd}} = \bar{y}\frac{\bar{X}+C_x}{\bar{x}+C} \qquad (1.4)$$

The mean square error is

$$MSE\left(\bar{y}_{sd}\right) \cong \bar{Y}^2\lambda\left\{C_y^2+\left(\frac{\bar{X}}{\bar{X}+C_x}\right)^2 C_x^2 - 2\left(\frac{\bar{X}}{\bar{X}+C_x}\right)\rho C_y C_x\right\} \qquad (1.5)$$

An exponential ratio type estimator due to Bhal and Tuteja[7] is given by

$$\bar{y}_{BT} = \bar{y}\exp\left(\left(\bar{X}-\bar{x}\right)/\left(\bar{X}+\bar{x}\right)\right) \qquad (1.6)$$

The Bias and MSE of (1.6) is given by

$$B\left(\bar{y}_{BT}\right) = \lambda\bar{Y}\left(\frac{3C_x^2}{8}-\frac{C_{yx}}{2}\right) \qquad (1.7)$$

&

$$MSE\left(\bar{y}_{BT}\right) = \lambda\bar{Y}^2\left(C_y^2+\left(\frac{C_x^2}{4}\right)^2 C_{yx}\right) \qquad (1.8)$$

Sing and tailor [8] proposed another estimator for estimating the mean of finite population for theknown value of correlation coefficient between the study variable and auxiliary variable. The estimatoris given by

$$\overline{y_{st}} = \bar{y}\frac{\bar{X}+\rho}{\bar{x}+\rho} \qquad (1.9)$$

The MSE is written as

$$MSE\left(\bar{y}_{st}\right) \cong \bar{Y}^2\lambda\left\{C_y^2+\left(\frac{\bar{X}}{\bar{X}+\rho}\right)^2 C_x^2 - 2\left(\frac{\bar{X}}{\bar{X}+\rho}\right)\rho C_y C_x\right\} \qquad (1.10)$$

PROPOSED ESTIMATORS

We suggest the following estimators

$$\bar{y}_{pr} = \left\{\omega_1\bar{y}+\left(1-\omega_1\right)\left(\bar{y}\frac{\bar{X}}{\bar{x}}\right)\right\} \qquad (2.0)$$

$$\bar{y}_{pr^2} = \left\{\omega_2\bar{y}+\left(1-\omega_2\right)\left(\bar{y}\exp\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)\right\} \qquad (2.1)$$

Where ω_1 & are constants or some functions of auxiliary information which is to be determined, sothat to get minimum MSE for the proposed estimator.

Properties of the First Proposed Estimator

We will come across through the following terms and notations to compute the Bias and MSE for theproposed estimator,

$$e_0 = \left(\frac{\overline{Y} - \overline{y}}{\overline{Y}}\right), e_1 = \left(\frac{\overline{X} - \overline{x}}{\overline{X}}\right)$$

$$E(e_1) = E(e_0) = 0, E(e_0^2) = \lambda C_y^2, E(e_1^2) = \lambda C_x^2 \;\&\; E(e_0 e_1) = \lambda C_{yx} \text{ or } E(e_0 e_1) = \lambda \rho C_y C_x$$

Then we can write (2.0) as follows

$$\overline{y}_{pr} = \overline{Y}(1+e_0)(1+e_1)^{-1} + \omega_1\left(\overline{Y}(1+e_0)(e_1 - e_1^2)\right) \qquad (2.2)$$

By neglecting the higher power terms, we have

$$\overline{y}_{pr} = -\overline{Y} \cong \overline{Y}\left\{e_0 - e_1 - e_0 e_1 + e_1^2 + \omega_1\left(e_1 - e_1^2 + e_0 e_1\right)\right\} \qquad (2.3)$$

The Bias corresponding to (2.0) is given by

$$Bias\left(\overline{y}_{pr}\right) \cong E\left(\overline{y}_{pr} - \overline{Y}\right)$$

Or

$$Bias\left(\overline{y}_{pr}\right) \cong \overline{Y}\left(C_x^2 - \rho C_y C_x - \omega_1\left(C_x^2 - \rho C_y C_x\right)\right) \qquad (2.4)$$

For MSE, Squaring and taking expectation of equation (2.3), we have

$$E\left(\overline{y}_{pr} - \overline{Y}\right)^2 \cong E\left\{\overline{Y}\left(e_0 - e_1 + \omega_1 e_1\right)\right\}^2 \qquad (2.5)$$

Since,

$$MSE\left(\overline{y}_{pr}\right) \cong E\left(\overline{y}_{pr} - \overline{Y}\right)^2$$

$$MSE\left(\overline{y}_{pr}\right) \cong \overline{Y}^2 \lambda\left[C_y^2 + C_x^2 - 2C_{yx} + \omega_1^2 C_x^2 + 2\omega_1 C_y^2 - 2\omega_1 C_x^2\right] \qquad (2.6)$$

We can find the optimum value of $\overline{y}_{pr}$ by minimizing the MSE of $\overline{y}_{pr}$ with respect to ω_1

Differentiating (2.6) w.r.to $\frac{\partial MSE\left(\overline{y}_{pr}\right)}{\partial \omega_1} = 0$ and equating to zero $\frac{\partial MSE\left(\overline{y}_{pr}\right)}{\partial \omega_1} = 0$ we get

$$\omega_{1opt} = 1 - \frac{\rho C_y}{C_x}$$

By substituting ω_{1opt} , in (2.4) and (2.6) we get

$$Bias\left(\overline{y}_{pr}\right) \cong \overline{Y}\lambda\rho C_y\left(C_x - \rho\right) \qquad (2.7)$$

$$MSE\left(\overline{y}_{pr}\right) \cong \overline{Y}^2 \lambda C_y^2 (1-\rho^2) \qquad (2.8)$$

Properties of the Second Proposed Estimator

$$\overline{y}_{pr^2} = \overline{Y}(1+e_0)\left[\exp\left\{\frac{-1}{2}e_1\left(1+\frac{e_1}{2}\right)^{-1}\right\} + \omega_2\left(1 - \exp\left\{\frac{-1}{2}e_1\left(1+\frac{e_1}{2}\right)^{-1}\right\}\right)\right] \qquad (2.9)$$

Terms with power higher than two is ignored, we have

$$\bar{y}_{pr^2} - \bar{Y} \cong \bar{Y}\left\{e_0 - \frac{1}{2}e_0e_1 + \omega_2\left(\frac{1}{2}e_1 - \frac{3}{8}e_1^2 + \frac{1}{2}e_0e_1\right)\right\} \quad (2.10)$$

We can write

$$Bias\left(\bar{y}_{pr^2}\right) \cong E\left(\bar{y}_{pr^2} - \bar{Y}\right)$$

or

$$Bias\left(\bar{y}_{pr^2}\right) \cong \bar{Y}\lambda\left(-\frac{1}{2}\rho C_y C_x + \omega_2\left(-\frac{3}{8}C_x^2 + \frac{1}{2}\rho C_y C_x\right)\right) \quad (2.11)$$

For MSE, Squaring and taking expectation of equation (3.3), we have

$$E\left(\bar{y}_{pr^2} - \bar{Y}\right) \cong E\left\{\bar{Y}\left(e_0^2 + \frac{1}{4}\omega_2^2 e_1^2 + \omega_2 e_1 e\right)\right\}^2 \quad (2.12)$$

Since,

$$MSE\left(\bar{y}_{pr2}\right) \cong E\left(\bar{y}_{pr2} - \bar{Y}\right)^2$$

$$MSE\left(\bar{y}_{pr2}\right) \cong \bar{Y}^2\lambda\left[C_y^2 + \frac{1}{4}\omega_2^2 C_x^2 + \omega_2 C_{yx}\right] \quad (2.13)$$

The optimum value of ω_2 can be find out by minimizing (3.6) with respect to the

Differentiating (3.6) w.r.to ω_2 and equating to zero $\omega_{2opt} = -2\frac{\rho C_y}{C_x}$ we get

$$\omega_{2opt} = -2\frac{\rho C_y}{C_x}$$

Substituting (3.4) (3.6) for ω_{2opt}, we get

$$Bias\left(\bar{y}_{pr^2}\right) \cong \bar{Y}\lambda\rho C_y\left(\frac{1}{4}C_x - \rho C_y\right) \quad (2.14)$$

$$MSE\left(\bar{y}_{pr}\right) \cong \bar{Y}^2\lambda C_y^2\left(1-\rho^2\right) \quad (2.15)$$

Theoretical Comparison of Proposed Estimators

Following are the conditions under which the suggested estimator performs well than the existingestimators considered here.

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}\right), i = 1,2 \text{ if}$$

$$\left(1-\rho^2\right) < 1$$

Which is always true if and only if $\rho \neq 0$

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{cr}\right), i = 1,2$$

If

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{sd}\right), i = 1,2$$

This is always true,

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{sd}\right), i = 1,2$$

If,

$$C_y^2\left(1-\rho^2\right) - \left(C_y^2 + \left(\frac{\bar{X}}{\bar{X}+C_x}\right)^2 C_x^2 - 2\left(\frac{\bar{X}}{\bar{X}+C_x}\right)\rho C_y C_x\right) < 0$$

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{BT}\right), i = 1,2$$

If

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{st}\right), i = 1,2$$

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{st}\right), i = 1,2$$

If

$$C_y^2\left(1-\rho^2\right) - \left(C_y^2 + \left(\frac{\bar{X}}{\bar{X}+\rho}\right)^2 C_x^2 - 2\left(\frac{\bar{X}}{\bar{X}+\rho}\right)\rho C_y C_x\right) < 0$$

$$MSE\left(\bar{y}_{pri}\right) < MSE\left(\bar{y}_{set}\right), i = 1,2$$

If

$$C_y^2\left(1-\rho^2\right) - \left\{C_y^2 + \left(\frac{\bar{X}}{\bar{X}+\beta_2}\right)^2 C_x^2 - 2\left(\frac{\bar{X}}{\bar{X}+\beta_2}\right)\rho C_y C_x\right\} < 0$$

Obviously the above conditions will always true when we apply it to real data sets.

$$rankF\left[M\left(T_n\right)\right] = \begin{cases} n^n - 0 \; if \; S \; is \; void, \\ n^n - \sum_{j=1}^{P} j! \; if \; S \; is \; nonvoid. \end{cases}$$

Applications in SRS

Here in this section we will apply our proposed estimator to different real data sets taken from variousfield of life. The table 2 below shows that our proposed estimator is best as compared to the existingestimators, discussed in the literature. The following data sets have been considered for the comparisonpurpose.

Table 1. Differeent data sets with their parameteres values.

Parameters	data set 1 Source: Murthy (1967),	Data set 2 Source: Murthy (1967),	Data set 3 Source: US Agriculture Statistics(2010)	Data set 4 Source: Koyuncu and Kadilar (2009)	Data set 5 Source : Pakistan MFA (2004)
N	108	80	69	923	97
n	16	20	17	180	25
$\bar{X}$	461.3981	11.2664	4505.16	11440.498	3050.28
$\bar{Y}$	172.704	51.8264	4514.9	436.43	3135.62

ρ	0.828315	0.3542	1.3756	1.718299	2.302173
ρ	0.6903	0.7507	1.18324	1.8645	2.327893
ρ	0.7896	0.9513	0.902327	0.9543	0.9871
β_1	1.3612	1.05	5.141563	3.9365	28.345
β_2	1.6307	-0.06339	29.77932	18.7208	50.32

Table 2. Percentage relative efficiency of the proposed estimators against some existing estimators

Estimators	Population 1	Population 2	Population 3	Population 4	Population 5
$\bar{y}$	100	100	100	100	100
$\bar{y}_{cr}$	263.83	66.28	439.899	939.7	3818.46
$\bar{y}_{sd}$	263.92	82.5	440.15	940.11	3823.8
$\bar{y}_{st}$	168.81	200.13	448.38	817.15	895.53
$\bar{y}_{st}$	263.93	87.067	440.05	939.91	3820.9
$\bar{y}_{set}$	264.03	65.05	448.34	943.8	3895.86
$\bar{y}_{pri}$	266.15	877.54	538.19	1119.7	3994.77

CONCLUSION

It is clear from above table that efficiency of our proposed estimators is optimum than all estimatorsconsidered in the literature, for all data sets. The conditions mentioned above also supported by the real data. Both estimators are equally efficient and give best results tan all others considered above. So we can modify some basic ratio estimators by assigning some suitable constants to them and hence their efficiency can be increases considerably.

REFERENCES

1. Cochran WG. The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. J Agric Sci. 1940;30:262-275
2. Hansen MH, et al. On the theory of sampling from finite populations.Ann Math Stat.1943;14:333-362.
3. MFA.Crops Area Production.Ministry of Food and Agriculture. Islamabad. Pakistan; 2004.
4. Gupta G, Shabbir J. On improvement in estimating the population mean in simple random sampling. J Appl Stat. 2008;35:559-566.
5. Srivastava SK, et al.A class of estimators of the population mean using multi-auxiliary information. Calc Statist Assoc Bull.1983;32:47-56.
6. Shabbir J, et al. On estimating finite population mean in simple and stratified random samplingCommun Stat Theory Methods. 2011;40:199-212.
7. Bahl S, et al. Ratio and product type exponential estimator. JIOS.1991;12:159-163.
8. Singh HP, etalAn estimator of Population mean using power transformation. J.I.S.A.S. (2004): 58(2), 223-230