Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online): 2229-5666
Vol. 12, Issue 1 (2019): 41-49

IMPROVEMENT FOR ESTIMATING POPULATION MEAN IN SIMPLE RANDOM SAMPLING SCHEME

Gagan Kumar[1], Priyanka Bharti[2], S.K. Yadav[3] and Surendra Kumar[*4]
[1]Department of Economics, Govt. Degree College, Pihani, Hardoi, India
[2]Department of Economics, DDU Govt. PG College, Sitapur, India
[3]Department of Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow, India
[*4]Department of Mathematics, Govt. Degree College Pihani, Hardoi, India
E Mail: *surendra.kumar776@gmail.com

Received October 09, 2018
Modified March 09, 2019
Accepted April 07, 2019

Abstract

This article, considers the improved population mean estimation without increasing the cost of the survey. An improved estimator for mean of population characteristic has been suggested which makes use of population median of main (study) variable. We study the sampling properties of the proposed estimator up to the approximation of order one. The least value of the mean squared error (MSE) for the optimum value of the constant of the suggested estimator has been obtained. The suggested estimator is compared with the competing estimators. The theoretical findings are justified with a numerical example. Through the numerical example we show the improvement of suggested estimator over other competing estimators.

Key Words: Main Variable, Secondary Variable, Exponential Ratio Estimators, Bias, MSE, Efficiency.

1. Introduction

It is essential to use sampling techniques when the population is very large to save the time and cost. Generally the corresponding statistics is used for estimating the parameter under consideration with the desirable properties of unbiasedness, consistency, efficiency and sufficiency etc. The mean per unit estimator of study variable is an appropriate estimator to be considered good for population mean and has unbiasedness property but it has a large amount of variation which is not desirable. Now we search for the estimator of population mean even biased but should have least MSE. The very purpose is achieved through using auxiliary variable and study variable as well which are in high correlation to each other. But this auxiliary information is collected on additional cost of the survey. Thus we further search for the improved estimators not using auxiliary information but makes use of known parameters of the main variable such as population median which is easily available with not influencing the survey cost.

In the present paper we use known population median of main variable for efficiently estimating the population mean of main variable and study its sampling properties up to approximation of degree one.

2. Review of Literature

Sample mean is the most natural estimator and is unbiased for population mean, given by,

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

Its variance is given by,

$$V(t_0) = \frac{1-f}{n} S_y^2 = \frac{1-f}{n}\bar{Y}^2 C_y^2 \tag{2}$$

$$C_y = \frac{S_y}{\bar{Y}},\ S_y^2 = \frac{1}{N-1}\sum_{i-1}^{N}(Y_i - \bar{Y})^2 = \frac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(\bar{y}_i - \bar{Y})^2,\ f = \frac{n}{N}.$$

Watson [14] given the conventional regression estimator as,

$$t_1 = \bar{y} + \beta(\bar{X} - \bar{x}) \tag{3}$$

Its variance is given by,

$$V(t_1) = \frac{1-f}{n}\bar{Y}^2 C_y^2 (1 - \rho_{yx}^2) \tag{4}$$

$$\rho_{yx} = \frac{Cov(x,y)}{S_x S_y}, Cov(x,y) = \frac{1}{N-1}\sum_{i-1}^{N}(Y_i - \bar{Y})(X_i - \bar{X}),$$

$$C_{yx} = \rho_{yx} C_y C_x,\ C_x = \frac{S_x}{\bar{X}},$$

$$S_x^2 = \frac{1}{N-1}\sum_{i-1}^{N}(X_i - \bar{X})^2 = \frac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(\bar{x}_i - \bar{X})^2.$$

Cochran [3] suggested traditional ratio estimator as,

$$t_2 = \bar{y}\frac{\bar{X}}{\bar{x}} \tag{5}$$

The MSE of t_2 is expressed as,

$$MSE(t_2) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + C_x^2 - 2C_{yx}] \tag{6}$$

Bahl and Tuteja [2] proposed the usual exponential ratio estimator as,

$$t_3 = \bar{y}\exp\left[\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right] \tag{7}$$

The MSE of the above estimator is as follows,

$$MSE(t_3) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + \frac{C_x^2}{4} - C_{yx}] \tag{8}$$

Kadilar and Cingi [4] presented the following ratio type estimator as,

$$t_4 = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^2 \tag{9}$$

The MSE of t_4 in (9) is given by,

$$MSE(t_4) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + 4C_x^2 - 4C_{yx}] \quad (10)$$

Srivastava [8] proposed a generalized ratio type estimator as,

$$t_5 = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha} \quad (11)$$

Where α is the arbitrary constant to be obtained that the MSE of t_5 is least.

The least MSE of above estimator for optimum $\alpha_{opt} = -C_{yx}\big/C_x^2$ is given by,

$$MSE_{\min}(t_5) = \frac{1-f}{n}\bar{Y}^2 C_y^2(1-\rho_{yx}^2) \quad (12)$$

Reddy [7] proposed a generalized ratio type estimator as,

$$t_6 = \bar{y}\left[\frac{\bar{X}}{\bar{X}+\alpha(\bar{x}-\bar{X})}\right] \quad (13)$$

The minimum MSE of above estimator for optimum $\alpha_{opt} = C_{yx}\big/C_x^2$ is given by,

$$MSE_{\min}(t_6) = \frac{1-f}{n}\bar{Y}^2 C_y^2(1-\rho_{yx}^2) \quad (14)$$

Subramani [10] suggested ratio type estimator as,

$$t_7 = \bar{y}\frac{M}{m} \quad (15)$$

Where M is population median and m is sample median of the main variable.

The MSE of t_7 is given by,

$$MSE(t_{12}) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + R_{12}^2 C_m^2 - 2R_{12}C_{ym}] \quad (16)$$

Where,

$$R_{12} = \frac{\bar{Y}}{M}, C_m = \frac{S_m}{M},\ S_{ym} = \frac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(\bar{y}_i - \bar{Y})(m_i - M),\ C_{ym} = \frac{S_{ym}}{\bar{Y}M} \text{ and}$$

$$S_m^2 = \frac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(m_i - M)^2 .$$

In the literature various authors including Kumar et al. [6], Subramani [9], Subramani and Kumarapandiyan [11-12], Tailor and Sharma [13], Yadav and Kadilar [16-17], Yadav and Mishra [18], Yadav et al. [19-21], Yan and Tian [22], Abid et al. [1] and Kumar et al. [7] used the auxiliary information and known population parameters of study variable as well form improved estimation of population mean.

3. Proposed Estimator

Getting motivated by Subramani [10], we proposed the new estimator as,

$$t=\bar{y}\left[\alpha\left(2-\frac{M}{m}\right)+(1-\alpha)\left(2-\frac{m}{M}\right)\right] \tag{17}$$

Where α is a characterizing scale and is obtained such that the MSE of t is minimum. We use the standard approximations given below for studying the properties of the proposed estimator as,

$\bar{y}=\bar{Y}(1+e_0)$ and $m=M(1+e_1)$ such that $E(e_0)=0$,

$E(e_1)=\frac{\bar{M}-M}{M}=\frac{Bias(m)}{M}$ and $E(e_0^2)=\frac{1-f}{n}C_y^2$, $E(e_1^2)=\frac{1-f}{n}C_m^2$,

$E(e_0e_1)=\frac{1-f}{n}C_{ym}$, where, $\bar{M}=\frac{1}{n}\sum_{i=1}^{n}m_i$

The suggested estimator t can be represented in the form of $e_i's$ ($i=1,2$) as,

$$t=\bar{Y}(1+e_0)\left[\alpha\left\{2-\frac{M}{M(1+e_1)}\right\}+(1-\alpha)\left(2-\frac{M(1+e_1)}{M}\right)\right]$$

$$=\bar{Y}(1+e_0)[\alpha\{2-(1+e_1)^{-1}\}+(1-\alpha)(1-e_1)]$$

$$=\bar{Y}(1+e_0)[\alpha\{2-(1-e_1+e_1^2+...)\}+(1-\alpha)(1-e_1)]$$

$=\bar{Y}(1+e_0)[\alpha(1+e_1-e_1^2)+(1-\alpha)(1-e_1)]$, up to approximation of order one

$$=\bar{Y}(1+e_0)[1-e_1+2\alpha e_1-\alpha e_1^2]$$

$$=\bar{Y}[1+e_0-e_1-e_0e_1+2\alpha e_1-2\alpha e_0e_1-\alpha e_1^2]$$

$$t-\bar{Y}=\bar{Y}[e_0-e_1-e_0e_1+2\alpha e_1-2\alpha e_0e_1-\alpha e_1^2]$$

$$t-\bar{Y}=\bar{Y}[e_0+(2\alpha-1)e_1+(2\alpha-1)e_0e_1-\alpha e_1^2] \tag{18}$$

Taking expectations on both sides of (18) and using standard results of expectations, we have the bias of t as,

$$B(t)=\bar{Y}\left[(2\alpha-1)\frac{B(m)}{M}+(2\alpha-1)\lambda C_{ym}-\alpha\lambda C_m^2\right]$$

$$=\bar{Y}\left[\alpha_1\frac{B(m)}{M}+\alpha_1\lambda C_{ym}-\alpha\lambda C_m^2\right],\ \alpha_1=(2\alpha-1) \tag{19}$$

From equation (2), up to approximation of degree one, we have,

$$t-\bar{Y}\approx\bar{Y}[e_0+(2\alpha-1)e_1]=\bar{Y}[e_0+\alpha_1e_1]$$

Squaring above equation and getting expectations both sides, we have the MSE of t as,

$$MSE(t) = \bar{Y}^2 E[e_0 + \alpha_1 e_1]^2$$
$$= \bar{Y}^2 E[e_0^2 + \alpha_1^2 e_1^2 + 2\alpha_1 e_0 e_1]$$
$$= \bar{Y}^2 [E(e_0^2) + \alpha_1^2 E(e_1^2) + 2\alpha_1 E(e_0 e_1)]$$

We get MSE of t by putting, $E(e_0^2)$, $E(e_1^2)$ and $E(e_0 e_1)$ as,

$$MSE(t) = \lambda \bar{Y}^2 [C_y^2 + \alpha_1^2 C_m^2 + 2\alpha_1 C_{ym}] \qquad (20)$$

The $MSE(t)$ is least for,

$$\frac{\partial MSE(t)}{\partial \alpha_1} = 0 \text{ gives,}$$

$$\alpha_1 C_m^2 + C_{ym} = 0 \text{ or,}$$

$$\alpha_{1(opt)} = -\frac{C_{ym}}{C_m^2} \qquad (21)$$

The minimum $MSE(t)$ for $\alpha_{1(opt)}$ is given by,

$$MSE_{\min}(t) = \frac{1-f}{n} \bar{Y}^2 \left[C_y^2 - \frac{C_{ym}^2}{C_m^2} \right] \qquad (22)$$

4. Efficiency Comparison

From equation (22) and (2), we have,

$$V(t_0) - MSE_{\min}(t) > 0 \text{, if}$$

$$\frac{C_{ym}^2}{C_m^2} > 0 \text{, or if } C_{ym}^2 > 0$$

Thus t are better than t_0.

From equation (22) and (4), we observe,

$$MSE(t_1) - MSE_{\min}(t) > 0 \text{ if}$$

$$\frac{C_{ym}^2}{C_m^2} - C_y^2 \rho_{yx}^2 > 0$$

Thus t are better than t_1 if above condition is met.

From equation (22) and (6), we observe,

$$MSE(t_2) - MSE_{\min}(t) > 0 \text{ if}$$

$$C_x^2 - 2C_{yx} + \frac{C_{ym}^2}{C_m^2} > 0 \text{, if}$$

$$C_x^2 + \frac{C_{ym}^2}{C_m^2} > 2C_{yx}$$

Thus t are better than t_2 if above condition is met.
From equation (22) and (8), we observe,

$$MSE(t_3) - MSE_{\min}(t) > 0 \text{ if}$$

$$\frac{C_x^2}{4} - C_{yx} + \frac{C_{ym}^2}{C_m^2} > 0, \text{ or}$$

$$\frac{C_x^2}{4} + \frac{C_{ym}^2}{C_m^2} > C_{yx}$$

Thus t are better than t_3 if above condition is met.
From equation (22) and (10), we observe,

$$MSE(t_4) - MSE_{\min}(t) > 0, \text{ if}$$

$$\frac{C_{ym}^2}{C_m^2} - C_y^2 \rho_{yx}^2 > 0$$

Thus t are better than t_4 if above condition is met.

Proposed estimator is better than Reddy [7] and Kadilar [5] estimators t_5 and t_6 respectively of population mean under the same condition as for Srivastava [8] estimator given in above equation.
From equation (22) and (16), we have,

$$MSE(t_7) - MSE_{\min}(t) > 0, \text{ if}$$

$$R_7^2 C_m^2 - 2R_7 C_{ym} + \frac{C_{ym}^2}{C_m^2} > 0, \text{ or}$$

$$R_7^2 C_m^2 + \frac{C_{ym}^2}{C_m^2} > 2R_7 C_{ym}$$

Thus t are better than t_7 if above condition is met.

5. Empirical Study

For the verification of the theoretical developments, we used two populations of Subramani [10]. Table-1 represents various parameters of these populations and Table-2 represents the MSEs of competing and suggested estimators respectively. Table-2 also depicts the percentage relative efficiency (PRE) of the suggested estimator over competing estimators given by,

$$PRE(t) = \frac{MSE(t_i)}{MSE(t)} \times 100, \; i = 1, 2, ..., 7$$

Parameter	Pop-1	Pop-2
N	34	20
n	5	5
${}^{N}C_{n}$	278256	15504
$\overline{Y}$	856.4118	41.5
$\overline{M}$	736.9811	40.0552
M	767.5	40.5
$\overline{X}$	208.8824	441.95
R_7	1.1158	1.0247
C_y^2	0.125014	0.008338
C_x^2	0.088563	0.007845
C_m^2	0.100833	0.006606
C_{ym}	0.07314	0.005394
C_{yx}	0.047257	0.005275
ρ_{yx}	0.4491	0.6522

Table 1: Constants and Parameters of two populations

Estimator	Pop-1	PRE	Popln-2	PRE
t_0	15640.97	173.76	2.15	228.72
t_1	12486.75	138.72	1.24	131.91
t_2	14895.27	165.48	1.48	157.45
t_3	12498.01	138.85	1.30	138.30
t_4	12486.75	138.72	1.24	131.91
t_5	12486.75	138.72	1.24	131.91
t_6	12486.75	138.72	1.24	131.91
t_7	10926.53	121.39	1.09	115.96
t	9001.36	100.00	0.94	100.00

Table 2: Mean squared error of various estimators and PRE of Proposed estimator over others

6. Results and Discussion

In Table 2, it is observed that for both the populations, the MSEs of the competing estimators are in between [10926.53 15640.97] and [1.09 2.15] while that of the suggested estimator 9001.36 and 0.94 respectively which are least MSE among all other estimators considered in completion in this manuscript, which was the aim of the present research that search for such estimator even biased but should have least MSE without enhancing the survey cost. The PRE of the suggested estimators for both the populations over the competing estimators are in between [121.39 173.76] and [115.96 228.72] respectively.

7. Conclusion

In the present study, we suggested a new ratio type estimator for enhancing the estimate of population mean using known value population median of main variable. We tried to develop the improved estimator without using auxiliary information, which is collected on some addition cost of the survey. Thus our aim was to find the improved estimator without increasing the cost of the survey. It is evident form Table 2 that the proposed estimator is best among other competing estimators of population mean. Almost all other estimators made use of supplementary information which increases the cost of the survey. Thus our aim of searching more improved estimator without increasing survey cost is achieved. Therefore it is recommended to use this estimator for efficient estimation of population mean of main variable without increasing the survey cost.

References

1. Abid, M., Abbas, N. Sherwani, R.A.K. and Nazir, H. Z. (2016). Improved Ratio Estimators for the population mean using non-conventional measure of dispersion, Pakistan Journal of Statistics and Operations Research, XII (2), p. 353-367.
2. Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimator, Information and Optimization Sciences, XII (I), p. 159-163.
3. Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, The Journal of Agric. Science, 30, p. 262-275.
4. Kadilar, C. and Cingi, H. (2003). A study on the chain ratio type estimator, Hacettepe Journal of Mathematics and Statistics, 32, p. 105-108.
5. Kadilar, G.O. (2016). A new exponential type estimator for the population mean in simple random sampling, Journal of Modern Applied Statistical Methods, 15(2), p. 207-214.
6. Kumar, G., Bharti, P., Kumar, S. and Yadav, S.K. (2018). General exponential ratio type estimator of population mean under simple random sampling scheme, International Journal of Creative Research Thoughts, 6(1), p. 1-8.
7. Kumar, B. Kumar, M. Rizvi, S.E.H. and Bhat, M. I. J. (2018). Estimation of population mean through improved class of ratio type estimators using auxiliary information, Journal of Reliability and Statistical Studies, 11(1), p. 01-07.
8. Reddy, V.N. (1974). On a transformed ratio method of estimation, Sankhya, C, 36(1), p. 59-70.

9. Srivastava, S.K. (1967). An estimator using auxiliary information in sample surveys, Cal. Statist. Assoc. Bull., 16, p. 62-63.
10. Subramani, J. (2013). Generalized modified ratio estimator of finite population mean, Journal of Modern Applied Statistical Methods, 12(2), p. 121–155.
11. Subramani, J. (2016). A new median based ratio estimator for estimation of the finite population mean, Statistics in Transition New Series, 17, 4, p. 1-14.
12. Subramani, J. and Kumarapandiyan, G. (2013). Estimation of population mean using deciles of an auxiliary variable, Statistics in Transition-New Series, 14 (1), p. 75–88.
13. Subramani, J. and Kumarapandiyan, G. (2013). A new modified ratio estimator of population mean when median of the auxiliary variable is known, Pakistan Journal of Statistics and Operation Research, Vol. 9 (2), p. 137–145.
14. Tailor, R. and Sharma, B. (2009). A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of kurtosis, Statistics in Transition-New Series, 10 (1), p. 15-24.
15. Watson, D.J. (1937). The estimation of leaf area in field crops, The Journal of Agricultural Science, 27, 3, p. 474-483.
16. Yadav, S.K. and Kadilar, C. (2013a). improved class of ratio and product estimators, Applied Mathematics and Computation, 219, p.10726-10731.
17. Yadav, S.K. and Kadilar, C. (2013b). Efficient family of exponential estimator for population mean, Hacettepe journal of Mathematics and Statistics, 42(6), p. 671 - 677.
18. Yadav, S.K. and Mishra, S.S. (2015). Developing improved predictive estimator for finite population mean using auxiliary information, Statistika, 95(1), p. 76-85.
19. Yadav, S.K., Misra, S., Mishra, S.S., Shukla, A.K. and Chaudhary, N.K. (2016). Efficient estimation of population mean using non-conventional measures of dispersion of auxiliary variable, International journal of Agricultural and Statistical Sciences, 12(2), p. 547-553.
20. Yadav, S. K., Singh, L., Mishra, S. S., Mishra, P. P., and Kumar, S. (2017). A median based regression type estimator of the finite population mean, International Journal of Agricultural and Statistical Sciences, 13(1), p. 265-271.
21. Yadav, D.K., Devi, M. and Yadav, S.K. (2018). Estimation of finite population mean using known coefficient of variation in the simultaneous presence of non - response and measurement errors under double sampling scheme, Journal of Reliability and Statistical Studies, 11(1), p. 51-66.
22. Yan, Z. and Tian, B. (2010). Ratio method to the mean estimation using coefficient of skewness of auxiliary variable, ICICA, Part II, CCIS, 106, p. 103-110.