American Journal of Operational Research 2014, 4(2): 21-27
DOI: 10.5923/j.ajor.20140402.02

Improved Ratio Estimators for Population Mean Based on Median Using Linear Combination of Population Mean and Median of an Auxiliary Variable

Subhash Kumar Yadav[1], Sant Sharan Mishra[1,*], Alok Kumar Shukla[2]

[1]Department of Mathematics and Statistics (A Centre of Excellence), Dr. RML Avadh University, Faizabad, India
[2]Department of Statistics, D.A-V College, Kanpur, India

Abstract In this manuscript the two efficient estimators of population mean using linear combination of population mean and median of auxiliary variable have been proposed. The expressions for the bias and mean square error (MSE) have been obtained up to the first order of approximation. A comparison has been made with the mentioned existing estimators of population mean. A numerical study has also been carried out to justify the improvement of the proposed estimators over other mentioned estimators of population mean of study variable.

Keywords Ratio estimator, Regression estimator, Bias, Mean square error, Efficiency

1. Introduction

The proper use of the auxiliary variable which is highly (positively or negatively) correlated with the main variable under study enhances the efficiency of the estimators of the parameters under considerations.

Let us consider a finite population having N distinct and identifiable units. Let Y_i ($i = 1, 2, ..., N$) be the observations on the characteristic under study for the main variable Y. Let X_i ($i = 1, 2, ..., N$) be the observations for the auxiliary variable X. The auxiliary variable is the variable about which we have full information.

When the parameters of the auxiliary variable X, such as population mean, population variance, coefficient of variation, coefficient of kurtosis, coefficient of skewness, median etc. are known, then the use of these parameters improves the efficiency of the estimator. Many estimators by using these parameters have been proposed in the literature for the improvement over traditional ratio estimator.

Following notations will be used in this manuscript,

N - Size of the population
n - Size of the sample
Y - Study variable
M - Median of the study variable
X - Auxiliary variable
M_d - Median of the auxiliary variable
ρ - Correlation coefficient between X and Y
$\overline{Y}$, $\overline{X}$ - Population means
$\bar{y}$, $\bar{x}$ - Sample means
$\overline{M}$ - Average of sample medians of Y
m - Sample median of Y
β - Regression coefficient of Y on X
$B(.)$ - Bias of the estimator
$V(.)$ - Variance of the estimator
$MSE(.)$ - Mean squared error of the estimator
$PRE(t_e, t_p) = \frac{MSE(t_e)}{MSE(t_p)} * 100$ - Percentage relative efficiency of the estimator t_p over t_e

Following formulae have been used in this paper,

$$V(\bar{y}) = \frac{1}{{}^{N}C_{n}} \sum_{i=1}^{{}^{N}C_{n}} (\bar{y}_i - \overline{Y})^2 ,$$

$$V(\bar{x}) = \frac{1}{{}^{N}C_{n}} \sum_{i=1}^{{}^{N}C_{n}} (\bar{x}_i - \overline{X})^2 ,$$

$$MSE(m) = V(m) = \frac{1}{{}^{N}C_{n}} \sum_{i=1}^{{}^{N}C_{n}} (m_i - M)^2 ,$$

* Corresponding author:
sant_x2003@yahoo.co.in (Sant Sharan Mishra)
Published online at http://journal.sapub.org/ajor
Copyright © 2014 Scientific & Academic Publishing. All Rights Reserved

$$Cov(\bar{y},\bar{x}) = \frac{1}{{}^{N}C_{n}} \sum_{i=1}^{{}^{N}C_{n}} (\bar{y}_i - \bar{Y})(\bar{x}_i - \bar{X}) = \frac{1-f}{n} \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})(x_i - \bar{X}),$$

$$Cov(\bar{y},m) = \frac{1}{{}^{N}C_{n}} \sum_{i=1}^{{}^{N}C_{n}} (\bar{y}_i - \bar{Y})(m_i - M),\ C_{xx} = \frac{V(\bar{x})}{\bar{X}^2},\ C_{mm} = \frac{V(m)}{M^2},\ C_{ym} = \frac{V(\bar{y},m)}{M\bar{Y}},\ C_{yx} = \frac{V(\bar{y},\bar{x})}{\bar{X}\bar{Y}}$$

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

As we know that for any population parameter the most appropriate estimator is its corresponding statistic. So the most appropriate estimator for population mean $\bar{Y}$ is the sample mean $\bar{y}$.

The Variance of $\bar{y}$, up to the first order of approximation is,

$$V(\bar{y}) = \frac{1-f}{n} S_y^2 \tag{1.1}$$

The variance given in (1.1) is sufficiently large, the aim is to reduce this variance. Cochran (1940) used the auxiliary variable to estimate the population mean of the main variable and proposed the traditional ratio estimator as,

$$t_R = \bar{y}\frac{\bar{X}}{\bar{x}} \tag{1.2}$$

The mean square error of this estimator, up to the first order of approximation is,

$$MSE(t_R) = V(\bar{y}) + R_1^2 V(\bar{x}) - 2R_1 Cov(\bar{y},\bar{x}), \text{ where } R_1 = \frac{\bar{Y}}{\bar{X}} \tag{1.3}$$

The traditional regression estimator using auxiliary variable is given by,

$$t_{lr} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x}) \tag{1.4}$$

The variance of the estimator t_{lr} , up to the first order of approximation is,

$$V(t_{lr}) = V(\bar{y})(1-\rho^2) \tag{1.5}$$

Subramani and Prabavathy (2014) proposed two estimators of population mean based on median of study variable using auxiliary information as,

$$\hat{\bar{Y}}_{SP1} = \bar{y}\left[\frac{M_d M + \bar{X}}{M_d m + \bar{X}}\right] \tag{1.6}$$

$$\hat{\bar{Y}}_{SP2} = \bar{y}\left[\frac{\bar{X}M + M_d}{\bar{X}m + M_d}\right] \tag{1.7}$$

The mean square errors of these estimators, up to the first order of approximation respectively are,

$$MSE(\hat{\bar{Y}}_{SPi}) = V(\bar{y}) + R_2^2\theta_i^2 V(m) - 2R_2\theta_i Cov(\bar{y},m),\ i = 1,2 \tag{1.8}$$

Where,

$$R_2 = \frac{\bar{Y}}{M},\quad \theta_1 = \frac{M_d M}{M_d M + \bar{X}},\ \theta_2 = \frac{\bar{X}M}{\bar{X}M + M_d}.$$

Many authors including Sisodia and Dwivedi (1981), Srivastava (1983), Mishra and Sahoo (1986), Pandey and Dubey (1988), Prasad (1989), Singh and Kakran (1993), Mohanty and Sahoo (1995), Agrawal and Sthapit (1997), Upadhyaya and Singh (1999), Singh (2003), Singh and Tailor (2003), Singhet.al (2004, 2008, 2010), Singh et al. (2004), Kadilar and Cingi (2004, 2006), Tailor and Sharma (2009), Yan and Tian (2010), Yadav (2011), Pandey *et al*. (2011), Subramani and

Kumarapandiyan (2012), Solanki *et al*. (2012), Onyeka (2012), Jeelani et.al (2013) and Yadav and Kadilar (2013) etc have used auxiliary information for improved estimation of population mean of study variable y.

2. Proposed Improved Ratio Estimators

In this paper, we have made our fresh attempt to extend the estimators of Subramani and Prabavathy (2014). Novelty of present work lies in the fact that when we introduce the constants kappas (κ_1 and κ_2) in the work of Subramani and Prabavathy (2014), we have been able to seek the new estimators which are found more efficient in the process of optimal search. We have derived the expressions for the biases and the mean square errors for the proposed estimators. The mean square errors of the proposed estimators contain the unknown constants kappas. These constants are obtained by minimizing the mean square errors of the proposed estimators using method of maxima and minima. When the optimum values of these estimators which are obtained by extending the previous work are put in the expressions of the mean square error, the mean square errors are minimized as compared to Subramani and Prabavathy (2014) and other existing estimators leading to desired optimum search of efficient estimators for population mean of the given population.

Motivated by Subramani and Prabavathy (2014) and Prasad (1989), we have proposed two improved ratio estimators of population mean as,

$$t_1 = \kappa_1 \bar{y}\left[\frac{M_d M + \bar{X}}{M_d m + \bar{X}}\right] \tag{2.1}$$

$$t_2 = \kappa_2 \bar{y}\left[\frac{\bar{X} M + M_d}{\bar{X} m + M_d}\right] \tag{2.2}$$

To study the large sample properties of proposed estimators t_1 and t_2, let us define,

$e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}$ and $e_1 = \frac{m - M}{M}$ such that $E(e_0) = 0$, $E(e_1) = \frac{Bias(m)}{M}$ and $E(e_0^2) = \frac{V(\bar{y})}{\bar{Y}^2}$, $E(e_1^2) = \frac{V(m)}{M^2}$, $E(e_0 e_1) = \frac{Cov(\bar{y}, m)}{\bar{Y}M}$.

Expressing estimator t_1 in terms of e_i's, we have

$$t_1 = \kappa_1 \bar{Y}(1 + e_0)\left[\frac{M_d M + \bar{X}}{M_d m + \bar{X}}\right]$$

$$= \kappa_1 \bar{Y}(1 + e_0)\left[\frac{1}{1 + \frac{M_d M}{M_d M + \bar{X}} e_1}\right]$$

$$= \kappa_1 \bar{Y}(1 + e_0)\left[\frac{1}{1 + \theta_1 e_1}\right]$$

$$= \kappa_1 \bar{Y}(1 + e_0)(1 - \theta_1 e_1)^{-1}$$

$$= \kappa_1 \bar{Y}(1 + e_0)(1 - \theta_1 e_1 + \theta_1^2 e_1^2 - \ldots)$$

$$= \kappa_1 \bar{Y}(1 + e_0 - \theta_1 e_1 - \theta_1 e_0 e_1 + \theta_1^2 e_1^2 - \ldots)$$

On subtracting $\bar{Y}$ both the sides of above equation, we have up to the first order of approximation,

$$t_1 - \bar{Y} = \kappa_1 \bar{Y}(1 + e_0 - \theta_1 e_1 - \theta_1 e_0 e_1 + \theta_1^2 e_1^2 - \ldots) - \bar{Y}$$

Taking expectation on both sides of above equation, we get bias of t_1 as,

$$B(t_1) = \bar{Y}[(\kappa_1 - 1) + \theta_1^2 E(e_1^2) - \theta_1 E(e_0 e_1) - \theta_1 E(e_1)]$$
$$= \bar{Y}[(\kappa_1 - 1) + \theta_1^2 \frac{V(m)}{M^2} - \theta_1 \frac{Cov(\bar{y}, m)}{\bar{Y}M} - \theta_1 \frac{Bias(m)}{M}] \quad (2.3)$$

The mean square error of t_1, up to the first order of approximation is,

$$MSE(t_1) = E[t_1 - \bar{Y}]^2$$
$$= E[\kappa_1 \bar{Y}(1 + e_0 - \theta_1 e_1 - \theta_1 e_0 e_1 + \theta_1^2 e_1^2) - \bar{Y}]^2$$

On simplifying, we have

$$MSE(t_1) = \bar{Y}^2[(\kappa_1 - 1)^2 + \kappa_1^2\{E(e_0^2) + 3\theta_1^2 E(e_1^2) - 4\theta_1 E(e_0 e_1)\} - 2\kappa_1\{\theta_1^2 E(e_1^2) - \theta_1 E(e_0 e_1)\}] \quad (2.4)$$

To find the minimum value of the $MSE(t_1)$, we use the maxima minima method. In this method, we differentiate $MSE(t_1)$ with respect to κ_1 and equate it to zero and solving this equation for κ_1, we find the optimum value of κ_1 which minimizes the value of $MSE(t_1)$. The optimum value of κ_1 is,

$$\kappa_1 = \frac{1 + \theta_1^2 E(e_1^2) - \theta_1 E(e_0 e_1)}{1 + E(e_0^2) + 3\theta_1^2 E(e_1^2) - 4\theta_1 E(e_0 e_1)} = \frac{1 + \theta_1^2 \frac{V(m)}{M^2} - \theta_1 \frac{Cov(\bar{y}, m)}{\bar{Y}M}}{1 + \frac{V(\bar{y})}{\bar{Y}^2} + 3\theta_1^2 \frac{V(m)}{M^2} - 4\theta_1 \frac{Cov(\bar{y}, m)}{\bar{Y}M}} = \frac{A_1}{B_1}$$

Where,

$$A_1 = 1 + \theta_1^2 \frac{V(m)}{M^2} - \theta_1 \frac{Cov(\bar{y}, m)}{\bar{Y}M}$$

$$B_1 = 1 + \frac{V(\bar{y})}{\bar{Y}^2} + 3\theta_1^2 \frac{V(m)}{M^2} - 4\theta_1 \frac{Cov(\bar{y}, m)}{\bar{Y}M}$$

Putting the optimum value of κ_1 in equation (2.4), we get the minimum MSE of t_1 as,

$$MSE_{\min}(t_1) = \bar{Y}^2\left[1 - \frac{A_1^2}{B_1}\right] \quad (2.5)$$

Similarly the bias and the mean square error of t_2, up to the first order of approximation respectively are,

$$B(t_2) = \bar{Y}[(\kappa_2 - 1) + \theta_2^2 \frac{V(m)}{M^2} - \theta_2 \frac{Cov(\bar{y}, m)}{\bar{Y}M} - \theta_2 \frac{Bias(m)}{M}] \quad (2.6)$$

$$MSE(t_2) = \bar{Y}^2[(\kappa_2 - 1)^2 + \kappa_2^2\{E(e_0^2) + 3\theta_2^2 E(e_1^2) - 4\theta_2 E(e_0 e_1)\} - 2\kappa_2\{\theta_2^2 E(e_1^2) - \theta_2 E(e_0 e_1)\}] \quad (2.7)$$

Which is minimum for,

$$\kappa_2 = \frac{1 + \theta_2^2 E(e_1^2) - \theta_2 E(e_0 e_1)}{1 + E(e_0^2) + 3\theta_2^2 E(e_1^2) - 4\theta_2 E(e_0 e_1)} = \frac{1 + \theta_2^2 \frac{V(m)}{M^2} - \theta_2 \frac{Cov(\bar{y}, m)}{\bar{Y}M}}{1 + \frac{V(\bar{y})}{\bar{Y}^2} + 3\theta_2^2 \frac{V(m)}{M^2} - 4\theta_2 \frac{Cov(\bar{y}, m)}{\bar{Y}M}} = \frac{A_2}{B_2}$$

Where,

$$A_2 = 1 + \theta_2^2 \frac{V(m)}{M^2} - \theta_2 \frac{Cov(\bar{y}, m)}{\bar{Y}M}$$

$$B_2 = 1 + \frac{V(\bar{y})}{\bar{Y}^2} + 3\theta_2^2 \frac{V(m)}{M^2} - 4\theta_2 \frac{Cov(\bar{y}, m)}{\bar{Y}M}$$

The minimum MSE of t_2 is given by,

$$MSE_{\min}(t_2) = \overline{Y}^2\left[1-\frac{A_2^2}{B_2}\right] \tag{2.8}$$

3. Efficiency Conditions

From equation (1.1) and (2.5), we have

$$MSE(\overline{y}) - MSE(t_1) > 0 \text{ , if}$$

$$\frac{(1-f)}{n}S_y^2 - \overline{Y}^2\left[1-\frac{A_1^2}{B_1}\right] > 0 \tag{3.1}$$

The proposed estimator in (2.5) is better than the ratio estimator in (1.2), if

$$MSE(t_R) - MSE(t_1) > 0 \text{ or,}$$

$$\left[V(\overline{y}) + R_1^2 V(\overline{x}) - 2R_1 Cov(\overline{y},\overline{x})\right] - \overline{Y}^2\left[1-\frac{A_1^2}{B_1}\right] > 0 \tag{3.2}$$

From equation (1.5) and (2.5), we have

$$MSE(t_{lr}) - MSE(t_1) > 0 \text{ or, } \left[V(\overline{y})(1-\rho^2)\right] - \overline{Y}^2\left[1-\frac{A_1^2}{B_1}\right] > 0 \tag{3.3}$$

From equation (1.8) and (2.5), we have

$$MSE(\hat{\overline{Y}}_{SPi}) - MSE(t_1) > 0 \text{ , } i = 1,2 \text{ , if}$$

$$\left[V(\overline{y}) + R_2^2\theta_i^2 V(m) - 2R_2\theta_i Cov(\overline{y},m)\right] - \overline{Y}^2\left[1-\frac{A_1^2}{B_1}\right] > 0 \text{ , } i = 1,2 \tag{3.4}$$

Note: Similar conditions are for the estimator t_2.

4. Numerical Example

To meet out the theoretical findings, we have considered the population from Singh and Chaudhary (1986) page 177. The data statistics is as follows,

Table 1. Population Parameters for sample size 3

$N = 34$	$n = 3$	$\overline{Y} = 856.4118$	$M = 767.50$
$\overline{X} = 208.8824$	$M_d = 150$	$\theta_1 = 0.9982$	$\theta_2 = 0.9991$
$V(m) = 101518.7738$	$R_1 = 4.0999$	$R_2 = 1.1158$	$Cov(\overline{y},m) = 90236.2939$
$V(\overline{y}) = 163356.4086$	$V(\overline{x}) = 6884.4455$	$\rho = 0.4491$	$Cov(\overline{y},\overline{x}) = 15061.4011$

Table 2. Population Parameters for sample size 5

$N = 34$	$n = 5$	$\overline{Y} = 856.4118$	$M = 767.50$
$\overline{X} = 208.8824$	$M_d = 150$	$\theta_1 = 0.9981$	$\theta_2 = 0.9990$
$V(m) = 59396.2836$	$R_1 = 4.0999$	$R_2 = 1.1158$	$Cov(\overline{y},m) = 48074.9542$
$V(\overline{y}) = 91690.3713$	$V(\overline{x}) = 3864.1726$	$\rho = 0.4491$	$Cov(\overline{y},\overline{x}) = 8453.8187$
$\overline{M} = 736.9811$			

Table 3. Mean square errors and percentage relative efficiencies of different estimators

Estimator	For sample size n=3		For sample size n= 5	
	MSE	PRE	MSE	PRE
$\overline{y}$	163356.4086	100.00	91690.3713	100.00
t_R	155577.8155	105.00	87324.3215	105.00
t_{lr}	130408.9222	125.27	73197.2660	125.27
$\hat{\overline{Y}}_{SP1}$	88286.3450	185.03	58283.5957	157.32
$\hat{\overline{Y}}_{SP2}$	88331.1066	184.94	58319.0161	157.22
t_1 (Proposed)	73464.9536	222.36	50889.93	180.17
t_2 (Proposed)	73468.7387	222.35	50905.30	180.12

5. Results and Conclusions

From the theoretical developments of the manuscript and the results in Table-3, we infer that the proposed estimators t_1 and t_2 of population mean have the minimum mean square errors than the mentioned estimators, the sample mean, ratio estimator, regression estimator and the estimators of Subramani and Prabavathy (2014). Thus the proposed estimators should be preferred for the estimation of population mean.

REFERENCES

[1] Agrawal, M.C. and Sthapit, A.B. (1997): Hierarchic predictive ratio-based and product-based estimators and their efficiency. Journal of Applied Statistics, 24, 1,97-104.

[2] Cochran, W.G. (1940): The estimation of the yields of the cereal experiments by sampling for the Ratio of grain to total produce, Jour. Agri. Sci. 59, 1225-1226.

[3] Jeelani, M.I., Maqbool, S. and Mir, S.A. (2013): Modified Ratio Estimators of Population Mean Using Linear Combination of Co-efficient of Skewness and Quartile Deviation. International Journal of Modern Mathematical Sciences, 6, 3, 174-183.

[4] Kadilar, C. and Cingi, H. (2004): Ratio estimators in simple random sampling, Applied Mathematics and Computation 151, 893-902.

[5] Kadilar, C. and Cingi, H. (2006): An improvement in estimating the population mean by using the correlation co-efficient, Hacettepe Journal of Mathematics and Statistics Volume 35 (1), 103-109.

[6] Mishra, G. and Sahoo, L.N. (1986): A Class of transformed ratio and product estimators in sample surveys. Journ. Stat. Research, 20, 81-85.

[7] Mohanty, S. and Sahoo, J. (1995): A note on improving the ratio method of estimation through linear transformation using certain known population parameters. Sankhya B, 57, 93-102.

[8] Murthy, M.N. (1967): Sampling Theory and Methods, Statistical .Publishing Society, Calcutta.

[9] Onyeka AC (2012) Estimation of population mean in poststratified sampling using known value of some population parameter(s). Statistics in Transition-New Series 13:65-78

[10] Prasad, B. (1989): Some improved ratio type estimators of population mean and ratio in finite population sample surveys, Communications in Statistics: Theory and Methods 18, 379–392,

[11] Pandey, B.N. and Dubey, V. (1988): Modified product estimator using coefficient of variation of auxiliary variate. Assam Stat. Review, 2, 64-66. 56

[12] Pandey H, Yadav SK, Shukla AK (2011) An improved general class of estimators estimating population mean using auxiliary information. International Journal of Statistics and Systems 6:1-7

[13] Robson, D.S. (1957): Applicationf multivariate polykays to the theory of unbiased ratio type estimation. J.Amer.stat. Assoc.,52, 411-422.

[14] Solanki RS, Singh HP, Rathour A (2012) An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. International Scholarly Research Network: Probability and Statistics, 2012:1-14

[15] Singh, H.P. and Kakran, M.S. (1993): A modified ratio estimator using coefficient of kurtosis of an auxiliary character(unpublished manuscript)

[16] Singh, G.N. (2003): On the improvement of product method of estimation in sample surveys. JISAS, 56, 267-275.

[17] Srivastava, S.K. (1983): Predictive estimation of finite population mean using product estimator. Metrika, 30, 93-99.

[18] Singh, D. and Chaudhary, F.S (1986). Theory and analysis of sample survey designs, New Age International Publisher.

[19] Singh, H.P. and Tailor, R. (2003): Use of known correlation co-efficient in estimating the finite population means, Statistics in Transition 6 (4), 555-560.

[20] Singh, H.P., Tailor, R., Tailor, R. and Kakran, M.S. (2004): An improved estimator of population mean using power transformation, Journal of the Indian Society of Agricultural Statistics 58(2), 223-230.

[21] Sisodia, B.V.S. and Dwivedi, V.K. (1981): A modified ratio estimator using co-efficient of variation of auxiliary variable, Journal of the Indian Society of Agricultural Statistics 33(1), 13-18.

[22] Subramani, J. and Kumarapandiyan, G. (2012): Estimation of Population Mean Using Co-Efficient of Variation and Median of an Auxiliary Variable. International Journal of Probability and Statistics 1(4): 111-118.

[23] Subramani, J. and Prabavathy, G. (2014). Median Based Modified Ratio Estimators with Linear Combinations of Population Mean and Median of an Auxiliary Variable, Journal of Reliability and Statistical Studies, 7, 1, 01-10.

[24] Upadhyaya, L.N. and Singh, H.P. (1999): Use of transformed auxiliary variable in estimating the finite population mean, Biometrical Journal 41 (5), 627-636.

[25] Yadav SK (2011) Efficient estimators for population variance using auxiliary information. Global Journal of Mathematical Sciences: Theory and Practical 3:369-376

[26] Yadav, S.K. and Kadilar, C. (2013): Improved class of ratio and product estimators, Applied Mathematics and Computation, 219, 10726–10731.

[27] Yan, Z. and Tian, B. (2010): Ratio method to the mean estimation using co-efficient of skewness of auxiliary variable, ICICA , Part II, CCIS 106, pp. 103–110.