Review Article

Open Access CrossMark

Modified ratio cum product estimators for estimation of finite population mean with known correlation coefficient

Abstract

In this paper, a modified ratio cum product estimator for the estimation of finite population mean of the study variable using the known correlation coefficient of the auxiliary variable is introduced. The bias and mean squared error of the proposed estimator are also obtained. The relative performance of the proposed estimator along with some existing estimators is accessed for certain labeled and natural populations. The results show that the proposed estimator is to be more efficient than the existing estimators.

Keywords: bias, mean squared error, natural population, simple random sampling, linear regression estimator

Volume 4 Issue 6 - 2016

Jambulingam Subramani, Master Ajith S
Department of Statistics, Ramanujan School of Mathematical Sciences, Pondicherry University, India

Correspondence: Jambulingam Subramani, Department of Statistics, Ramanujan School of Mathematical Sciences, Pondicherry University, RV Nagar, Kalapet, Pondicherry, India,
Email drjsubramani@yahoo.co.in

Received: September 26, 2016 | **Published:** November 15, 2016

Introduction

In sampling theory, a wide variety of techniques is used to obtain efficient estimators for the population mean. The commonly used method to obtain the estimator for population mean is simple random sampling without replacement (SRSWOR) when there is no auxiliary variable available. There are methods that use the auxiliary information of the study characteristics. If there exists an auxiliary variable X which is correlated with the study variable Y, then a number of estimators such as ratio, product, modified ratio, modified product, regression estimators and their modifications are widely available for estimation of population mean of the study variable Y.

Consider a finite population $U = \{U_1, U_2, U_3 \ldots U_N\}$ of N distinct and identifiable units. Let Y be the study variable which takes the values $Y = \{Y_1, Y_2, Y_3, \ldots Y_N\}$. Here the problem is to estimate the population mean $\bar{Y} = \frac{1}{N}\Sigma_{i=1}^{N} Y_i$ on the basis of a random sample selected from the population U.

Before discussing further the various estimators, the notations to be used in this article are listed here.

N	-	Population size
n	-	Sample size
$f = \frac{n}{N}$	-	Sampling fraction
Y	-	Study variable
X	-	Auxiliary variable
$\bar{X}, \bar{Y}$	-	Population means
$\bar{x}, \bar{y}$	-	Sample means
S_x, S_y	-	Population standard deviations
s_x, s_y	-	Sample standard deviations
C_x, C_y	-	Coefficient of variations
ρ	-	Correlation coefficient between x and y
β_1	-	Coefficient of skewness
β_2	-	Coefficient of kurtosis
B(.)	-	Bias of estimators
MSE(.)	-	Mean squared error of estimators

In simple random sampling without replacement, the estimator $\bar{y}_{srs}$ is an unbiased estimator for the population mean $\bar{Y}$ and its variance is given by

$$V(\bar{y}_{srs}) = \delta S_y^2 \tag{1}$$

Where $\delta = \left(\frac{1-f}{n}\right)$

Cochran,[1] use auxiliary information for the estimation of population mean of the variable under study and proposed the ratio estimator of the population mean $\bar{Y}$ of the study variable,

$$\hat{Y}_R = \frac{\bar{y}}{\bar{x}}\bar{X} = \hat{R}\bar{X}$$

The bias and mean squared error of the ratio estimator are given by

$$B\left(\hat{\bar{Y}}_R\right) = \delta\bar{Y}\left[C_x^2 - \rho C_x C_y\right]$$

$$MSE\left(\hat{\bar{Y}}_R\right) = \delta\bar{Y}^2\left[C_y^2 + C_x^2 - 2\rho C_x C_y\right] \quad (2)$$

The linear regression estimator and its variance are given by

$$\bar{y}_{lr} = \bar{y} + b\left(\bar{X} - \bar{x}\right)$$

$$V\left(\bar{y}_{lr}\right) = \delta S_y^2\left(1 - \rho^2\right) \quad (3)$$

where b is the regression coefficient Y on X

Murthy[2] proposed the product estimator to estimate the population mean of the study variable when there is a negative correlation between the study variable Y and auxiliary variable X as

$$\hat{\bar{Y}}_p = \bar{y}\frac{\bar{x}}{\bar{X}}$$

The bias and the mean squared error of the product estimator are given by

$$B\left(\bar{Y}_p\right) = \delta\bar{Y}\left[\rho C_x C_y\right]$$

$$MSE\left(\hat{\bar{Y}}_p\right) = \delta\bar{Y}^2\left[C_y^2 + C_x^2 + 2\rho C_x C_y\right] \quad (4)$$

Singh and Tailor[3] introduced the modified ratio estimator for the population mean with known population correlation coefficient ρ of the auxiliary variable and is given by

$$\hat{\bar{Y}}_{MR} = \bar{y}\left(\frac{\bar{X}+\rho}{\bar{x}+\rho}\right)$$

The bias and mean squared error of this modified ratio estimator are given by

$$B\left(\hat{\bar{Y}}_{MR}\right) = \delta\bar{Y}\left[\theta^2 C_x^2 - \theta\rho C_x C_y\right]$$

$$MSE\left(\hat{\bar{Y}}_{MR}\right) = \delta\bar{Y}^2\left[C_y^2 + \theta^2 C_x^2 - 2\theta\rho C_x C_y\right] \quad (5)$$

where $\theta = \frac{\bar{X}}{\bar{X}+\rho}$

The modified product estimator with known correlation coefficient of the auxiliary variable when there is a negative correlation between the study variable Y and auxiliary variable X is given as

$$\hat{\bar{Y}}_{Mp} = \bar{y}\left(\frac{\bar{x}+\rho}{\bar{X}+\rho}\right)$$

The bias and mean squared error of the modified product estimator are given by

$$B\left(\hat{\bar{Y}}_{Mp}\right) = \delta\bar{Y}\left[\theta\rho C_x C_y\right]$$

$$MSE\left(\hat{\bar{Y}}_{Mp}\right) = \delta\bar{Y}\left[C_y^2 + \theta^2 C_x^2 + 2\theta\rho C_x C_y\right] \quad (6)$$

where $\theta = \frac{\bar{X}}{\bar{X}+\rho}$

In literature, several estimators are available with auxiliary variables. However the problem is that the best estimator in terms of bias and efficiency are not fully addressed. In this paper, we attempt to solve such type of problems. The existing estimators are biased but the percentage relative efficiency is better than that of simple random sampling, ratio and product estimators. These points are motivated us to introduce a new class of improved ratio cum product estimators for the estimation of the population mean of the study variable.

Proposed estimators

For estimating population mean $\bar{Y}$ we have proposed a class of ratio cum product estimators[4] for the population mean by using the known population correlation coefficient of the auxiliary variable and is given by

$$\hat{\bar{Y}}_{pr} = \alpha\lambda_1\bar{y}\left(\frac{\bar{X}+\rho}{\bar{x}+\rho}\right) + (1-\alpha)\lambda_2\bar{y}\left(\frac{\bar{x}+\rho}{\bar{X}+\rho}\right)$$

Here, $\lambda_1 = \frac{S_y}{S_y + \gamma_1 C_y}$ and $\lambda_2 = \frac{S_y}{S_y + \gamma_2 C_y}$,

$$\gamma_2 = B\left(\hat{\bar{Y}}_{Mp}\right)$$

Bias and mean squared error of the proposed estimators

The detailed derivation of the bias and mean squared error are given in the appendix whereas the procedures to obtain the bias and mean squared error of the proposed estimators are briefly outlined below:

Consider, $e_\theta = \frac{\bar{y}-\bar{Y}}{\bar{Y}}, e_1 = \frac{\bar{x}-\bar{X}}{\bar{X}} \quad \theta = \frac{\bar{X}}{\bar{X}+\rho}$

$$E\left(e_\theta\right) = E\left(e_1\right) = 0 \ ,$$

$$E\left(e_0{}^2\right) = \delta \bar{Y}^2 C_y^2 \ , E\left(e_1{}^2\right) = \delta \bar{X}^2 C_x^2 \ , E\left(e_0 e_1\right) = \delta \rho C_x C_y$$

Substitute these values in equation (9) and neglecting the high order expressions, we get

$$B\left(\hat{\bar{Y}}_{\mathrm{Pr}}\right) = E\left(\hat{\bar{Y}}_{\mathrm{Pr}} - \bar{Y}\right)$$

$$B\left(\hat{\bar{Y}}_{\mathrm{Pr}}\right) = \bar{Y}\left(\alpha\lambda_1+(1-\alpha)\lambda_2-1\right) + \delta\bar{Y}\left\{\alpha\lambda_1\theta^2C_x^2-\theta\rho C_xC_y\left(\alpha\lambda_1-(1-\alpha)\lambda_2\right)\right\}$$

$$MSE\left(\hat{\bar{Y}}_{\mathrm{Pr}}\right) = \bar{Y}^2\left(A-1\right)^2 + \delta\bar{Y}^2\left\{C_y{}^2\left(\alpha\lambda_1+(1-\alpha)\lambda_2\right)^2+\theta^2C_x{}^2\left(3\alpha^2\lambda_1{}^2+(1+\alpha)^2\lambda_2{}^2-2\alpha\lambda_1\right)+2\theta\rho C_xC_y\left(\alpha\lambda_1-(1-\alpha)\lambda_2\right)-2\left(\alpha^2\lambda_1{}^2-(1-\alpha)^2\lambda_2{}^2\right)\right\}$$

$$MSE\left(\hat{\bar{Y}}_{\mathrm{Pr}}\right) = \bar{Y}^2\left(A-1\right)^2 + \delta\bar{Y}^2\left\{A^2C_y{}^2+\theta^2C_x{}^2\left(A^2+\left(A+B\right)\left(B-1\right)\right)-2\theta\rho C_xC_yB\left(2A-1\right)\right\}$$

where $A = \left(\alpha\lambda_1+(1-\alpha)\lambda_2\right)$, $B = \left(\alpha\lambda_1-(1-\alpha)\lambda_2\right)$ and $\theta = \frac{\bar{X}}{\bar{X}+\rho}$

The optimal value of α is determined by minimizing the MSE ($\hat{\bar{Y}}_{pr}$) with respect to α. For this differentiate MSE with respect to α and equate to zero.[5]

$$\frac{\partial MSE}{\partial \alpha} = 0 \text{ , and we get the value of } \alpha \text{ , as}$$

$$\alpha = \frac{(\lambda_2-1)(\lambda_2-\lambda_1)-\delta\left[C_y{}^2\lambda_2(\lambda_1-\lambda_2)-\theta^2C_x{}^2\left(\lambda_1+\lambda_2{}^2\right)+\theta\rho C_xC_y\left(\lambda_1+\lambda_2-4\lambda_2{}^2\right)\right]}{(\lambda_1-\lambda_2)^2+\delta\left[(\lambda_1-\lambda_2)^2C_y{}^2+\theta^2C_x{}^2\left(3\lambda_1{}^2+\lambda_2{}^2\right)4\theta\rho C_xC_y\left(\lambda_2{}^2-\lambda_1{}^2\right)\right]}$$

Efficiency comparison

The efficiencies of the proposed estimators with that of the existing estimators are obtained algebraically and are as follows:

Comparison of proposed estimator and simple random sampling (SRSWOR) estimator

The proposed estimator is more efficient than simple random sampling estimator,

$$V\left(\bar{y}_{lr}\right) \geq MSE\left(\hat{\bar{Y}}_{\mathrm{Pr}}\right) \quad \text{if}$$

$$C_y^2 \geq \frac{\left\{(A-1)^2+\delta^2\left\{\theta^2C_x^2\left(A^2+\left(A+B\right)\left(B-1\right)\right)-2\theta\rho C_xC_yB\left(2A-1\right)\right\}\right\}}{\delta^2\left(1-A^2\right)}$$

Comparison of proposed estimator and linear regression estimator

The proposed estimator is more efficient than linear regression estimator,

$$V\left(\bar{y}_{lr}\right) \geq MSE\left(\hat{\bar{Y}}_{\mathrm{Pr}}\right) \quad \text{if}$$

$$C_y^2 \geq \frac{\left\{(A-1)^2+\delta\left[\theta^2C_x{}^2\left(A^2+\left(A+B\right)\left(B-1\right)\right)-2\theta C_xC_y\left(B\left(2A-1\right)-1\right)\right]\right\}}{\delta\left(1-\rho^2-A^2\right)}$$

Comparison of proposed estimator and ratio estimator

The proposed estimator is more efficient than ratio estimator

$$MSE\left(\hat{\bar{Y}}_P\right) \geq MSE\left(\hat{\bar{Y}}_{\Pr}\right) \text{ if}$$

$$C_y^2 \geq \frac{\left\{(A-1)^2+\delta\left\{C_x^2\left(\theta^2\left(A^2+(A+B)(B-1)\right)-1\right)-2\rho C_x C_y\left(\theta B(2A-1)-1\right)+1\right\}\right\}}{\delta\left(1-A^2\right)}$$

Comparison of proposed estimator and product estimator

The proposed estimator is more efficient than ratio estimator,[6]

$$MSE\left(\hat{\bar{Y}}_P\right) \geq MSE\left(\hat{\bar{Y}}_{\Pr}\right) \text{ if}$$

$$C_y^2 \geq \frac{\left\{(A-1)^2+\delta\left\{C_x^2\left(\theta^2\left(A^2+(A+B)(B-1)\right)-1\right)-2\rho C_x C_y\left(\theta B(2A-1)+1\right)\right\}\right\}}{\delta\left(1-A^2\right)}$$

Comparison of proposed estimator and modified ratio estimator

The proposed estimator is more efficient than modified ratio estimator[7]

$$MSE\left(\hat{\bar{Y}}_{MR}\right) \geq MSE\left(\hat{\bar{Y}}_{\Pr}\right)$$

$$C_y^2 \geq \frac{\left\{(A-1)^2+\delta\left\{\theta^2 C_x^2\left(A^2+(A+B)(B-1)-1\right)-2\theta\rho C_x C_y\left(\theta B(A-1)-1\right)-1\right\}\right\}}{\delta\left(1-A^2\right)}$$

Comparison of proposed estimator and modified product estimator

The proposed estimator is more efficient than modified product estimator,

$$MSE\left(\hat{\bar{Y}}_{Mp}\right) \geq MSE\left(\hat{\bar{Y}}_{\Pr}\right) \text{ if}$$

$$C_y^2 \geq \frac{\left\{(A-1)^2+\delta\left\{\theta^2 C_x^2\left(A^2+(A+B)(B-1)-1\right)-2\theta\rho C_x C_y\left(B(2A-1)-1\right)\right\}\right\}}{\delta\left(1-A^2\right)}$$

Numerical study

In this section, we consider the four natural populations population 1 Khoshnevisan et al.,[8] Population 2 Cochran[9] (page 325) population 3 and 4 Singh and Chaudhary,[10] (page 177) and are used to compare the percentage relative efficiency of proposed estimator with that of the existing estimators such as SRSWOR sample mean, linear regression estimator, ratio estimator, product estimator, modified ratio estimators, and modified product estimators.

Conclusion

We have proposed a class of modified ratio cum product estimators for finite population[11] mean of the study variable Y with known correlation coefficient of the auxiliary variable X. The bias and mean squared error of the proposed estimators are obtained and compared with that of the simple random sampling without replacement, regression, ratio, product, modified ratio, modified product estimators by both algebraically and numerically. We support this theoretical result with numerical examples. We have shown that the proposed estimator is more efficient than other existing estimators under the optimum values of α. Table 1&2 shows that the bias and MSE of the proposed estimators are smaller than the other competing estimators. Table 3 shows that the percentage relative efficiency of the proposed estimator with respect to the existing estimators,

Table 1 The computed values of constants and parameters from different populations

Parameters	Population 1	Population 2	Population 3	Population 4
N	20	10	34	34
n	8	3	3	5
$\bar{Y}$	19.55	101.1	856.4117	856.4117
$\bar{X}$	18.8	58.8	208.8823	208.8823
ρ	-0.9199	0.6515	0.4491	0.4491
	6.9441	15.4448	733.1407	733.1407
	0.3552	0.1527	0.8561	0.8561
	7.4128	7.9414	150.5059	150.5059
	0.3943	0.1351	0.7205	0.7205
	3.0613	0.2363	2.9123	2.9123
	0.5473	2.2388	0.9781	0.9781
	1.0514	0.989	0.9978	0.9965
	0.4506	0.1072	625915	35.1319
	-0.1986	0.3136	71.947	40.3832
	0.9774	0.9989	0.9319	0.9605
	1.0102	0.9969	0.9225	0.9549
	0.1055	0.8717	0.7614	0.7639

Table 2 Bias and MSE of proposed and existing estimators from different population

Estimator	Population 1		Population 2		Population 3		Population 4	
	Bias	MSE	**Bias**	MSE	**Bias**	MSE	`**Bias**	MSE
Proposed	1.14e-06	0.5463	1.13e-15	31.9319	-0.00274	109092.8	-0.0015	66145.84
$\bar{y}_{srs}$	-	3.6166	-	55.6603	-	163356.4	-	91690.37
$\bar{y}_{lr}$	-	0.5561	-	32.0343	-	130408.9	-	73197.27
$\hat{\bar{Y}}_{R}$	0.4168	15.4595	0.1132	35.0447	63.0193	155580.6	35.3721	87325.9
$\hat{\bar{Y}}_{p}$	-0.1889	0.6869	0.3171	163.283	72.0984	402564.2	40.4681	225955.4
$\hat{\bar{Y}}_{MR}$	0.4506	16.3099	0.1072	34.7991	62.5915	155359.1	35.1319	87201.54
$\hat{\bar{Y}}_{MP}$	-0.1986	0.7774	0.3163	161.6321	71.947	401824.2	40.3832	225540

In fact, the PRE is ranging from

I. 138.6185 to 661.9810 in case of SRSWOR sample mean

II. 100.3205 to 119.4555 in case of Linear Regression Estimator

III. 109.7481 to 2829.7520 in case of Ratio estimator

IV. 125.7468 to 511.3465 in case of Product estimator

V. 108.9790 to 2985.4080 in case of Modified Ratio estimator and

VI. 142.2864 to 506.1764 in case of Product estimator

From this, we have observed that the proposed estimator is performed better than that of other existing estimators and hence we recommend the proposed estimators for the practical problems.

Table 3 Percentage relative efficiency of the proposed estimator

Estimators	Population 1	Population 2	Population 3	Population 4
$\bar{y}_{srs}$	661.981	174.3092	149.6356	138.6185
$\bar{y}_{lr}$	101.802	100.3205	119.4555	110.6604
$\hat{\bar{Y}}_{R}$	2829.752	109.7481	142.5216	132.0283
$\hat{\bar{Y}}_{p}$	125.7468	511.3465	368.7708	341.6196
$\hat{\bar{Y}}_{MR}$	2985.408	108.979	142.183	131.8322
$\hat{\bar{Y}}_{MP}$	142.2864	506.1764	367.6544	340.9739

Acknowledgement

None.

Conflict of interest

None.

References

1. Cochran WG. The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*. 1940;30(2):262–275.
2. Murthy MN. Product method of estimation. Sankhyā:. *The Indian Journal of Statistics*. 1964;26(1):69–74.
3. Singh HP, Tailor R. Use of known correlation coefficient in estimating the finite population means. Statistics in Transition. 2003;6(4):555–560.
4. Ekaette Inyang Enang, Victoria Matthew Akpan, Emmanuel John Ekpenyong. Alternative ratio estimator of population mean in simple random sampling, *Journal of Mathematics Research*. 2014;6(3).
5. Housila P Singh, Surya K Pal, Vishal Mehta. A generalized class of dual to product–cum–dual to ratiotype estimators of finite population mean in sample surveys. *Appl Math Inf Sci Lett*. 2016;4(1):25–33.
6. J Subramani G Kumarapandiyan. A class of almost unbiased modified ratio estimators for population mean with known population parameters. *Elixir Statistics*. 2012;44:7411–7415.
7. Murthy MN. Sampling theory and methods. Statistical Publishing Society, Calcutta, India. 1967.
8. Khoshnevisan M, Singh R, Chauhan P, et al. A general family of estimators for estimating population mean using known value of some population parameter(s). *Far East Journal of Theoretical Statistics*. 2007;22:181–191.
9. Cochran WG. Sampling Techniques. (3rd edn), Wiley Eastern Limited, India. 1977;pp. 448.
10. Singh D, Chaudhary FS. Theory and analysis of sample survey designs. (1st edn), New Age International Publisher, India, 1986;pp. 332.
11. Subramani J. "Generalized modified ratio estimator for estimation of finite population mean". *Journal of Modern Applied Statistical Methods*. 2013;12(2):pp.121–155.