General exponential ratio type estimator of population mean under simple random sampling scheme

[1]GAGAN KUMAR, [2]PRIYANKA BHARTI, [3]SURENDRA KUMAR, [4]S. K. YADAV

Assistant Professor
[1]Department of Economics,
[1]Govt. Degree College Pihani, Hardoi - 241406 U.P., INDIA
[2]Department of Economics
[2]DDU Govt. P.G. College Sitapur- 261001 U.P., INDIA
[3]Department of Mathematics
[3]Govt. Degree College Pihani, Hardoi- 241406 U.P., INDIA
[4]Department of Statistics
[4]Babasaheb Bhimrao Ambedkar University, Lucknow-226025, U.P., INDIA

Abstract: **In present paper we propose a general exponential ratio type estimator for estimating population mean of study variable. An estimator utilizing information on population median of study variable has been developed for improved estimation of population mean. The expressions for the bias and mean squared error have been derive up to approximation of degree one. The proposed estimator is theoretically compared with other competing estimators of population mean. Theoretical developments have been judged through the numerical study. It has been shown through the numerical study that proposed estimator performs better than other competing estimators which make use of auxiliary information collected on additional cost of the survey.**

2010 AMS Classification: 62D05
***Keywords*: Study variable, auxiliary variable, exponential ratio estimators, bias, mean squared error, efficiency.**

I. INTRODUCTION

Sampling is a must whenever the population is very large because it very costly and time consuming to get the information on every unit of the population. The best way for estimating any parameter is its corresponding statistic. Thus most truly estimator for population mean of study variable Y is the sample mean. But its variance is reasonably large so we search for some improved estimator. Improved estimators are obtained through the use of auxiliary information collected on additional cost of the survey. Now our aim is to find the improved estimator without using auxiliary information. It is done through the use of known population parameter of study variable such as population median.

Let us consider a finite population for investigation containing *N* distinct and identifiable units and let $(x_i, y_i),\ i = 1, 2, ..., n$ be a bivariate sample of size n on bivariate set (*X*, Y) drawn using simple random sampling without replacement (SRSWOR) technique. Let $\bar{X}$ and $\bar{Y}$ represents the population means of auxiliary variable and the study variable respectively. Let $\bar{x}$ and $\bar{y}$ be the sample means of the auxiliary and study variables respectively. It is further to be noted that under SRSWOR, the sample means of auxiliary and study variables are unbiased estimators of respective population means $\bar{X}$ and $\bar{Y}$ respectively.

For the population under consideration, Table-I represents various estimators of population mean using and without using auxiliary information along with their mean squared errors up to the approximation of degree one.

Table-I: Various estimators along with their mean squared errors

S. No.	Estimator	Mean Squared Error
1.	$t_o = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$	$V(t_0) = \frac{1-f}{n} S_y^2 = \frac{1-f}{n}\bar{Y}^2 C_y^2$
2.	$t_1 = \bar{y} + \beta(\bar{X} - \bar{x})$ Watson (1937)	$V(t_1) = \frac{1-f}{n}\bar{Y}^2 C_y^2 (1-\rho_{yx}^2)$
3.	$t_2 = \bar{y}\frac{\bar{X}}{\bar{x}}$ Cochran (1940)	$MSE(t_2) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + C_x^2 - 2C_{yx}]$
4.	$t_3 = \bar{y}\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right]$ Bahl and Tuteja (1991)	$MSE(t_3) = \frac{1-f}{n}\bar{Y}^2[C_y^2 + \frac{C_x^2}{4} - C_{yx}]$

5.	$t_4=\bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha}$ Srivastava (1967)	$MSE(t_4)=\frac{1-f}{n}\bar{Y}^2[C_y^2+\alpha^2C_x^2+2\alpha C_{yx}]$ $MSE_{\min}(t_4)=\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{yx}^2)$ for $\alpha_{opt}=-C_{yx}\big/C_x^2$
6.	$t_5=\bar{y}\left[\frac{\bar{X}}{\bar{X}+\alpha(\bar{x}-\bar{X})}\right]$ Reddy (1974)	$MSE(t_5)=\frac{1-f}{n}\bar{Y}^2[C_y^2+\alpha^2C_x^2-2\alpha C_{yx}]$ $MSE_{\min}(t_5)=\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{yx}^2)$ for $\alpha_{opt}=C_{yx}\big/C_x^2$
7.	$t_6=\bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^{\delta}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)$ Kadilar (2016)	$MSE(t_6)=\frac{1-f}{n}\bar{Y}^2\left[C_y^2+\left(\delta^2+\delta+\frac{1}{4}\right)C_x^2+(2\delta+1)C_{yx}\right]$ $MSE_{\min}(t_6)=\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{yx}^2)$ for $\delta_{opt}=\left(\frac{1}{2}-\rho_{yx}C_y/C_x\right)$
8.	$t_7=\bar{y}\frac{M}{m}$ Subramani (2016)	$MSE(t_7)=\frac{1-f}{n}\bar{Y}^2[C_y^2+R_7^2C_m^2-2R_7C_{ym}]$

Where,

$$C_y=\frac{S_y}{\bar{Y}},\ S_y^2=\frac{1}{N-1}\sum_{i-1}^{N}(Y_i-\bar{Y})^2=\frac{1}{^NC_n}\sum_{i-1}^{^NC_n}(\bar{y}_i-\bar{Y})^2,\ C_{yx}=\rho_{yx}C_yC_x,\ f=\frac{n}{N},\ C_x=\frac{S_x}{\bar{X}},$$

$$S_x^2=\frac{1}{N-1}\sum_{i-1}^{N}(X_i-\bar{X})^2=\frac{1}{^NC_n}\sum_{i-1}^{^NC_n}(\bar{x}_i-\bar{X})^2,\ Cov(x,y)=\frac{1}{N-1}\sum_{i-1}^{N}(Y_i-\bar{Y})(X_i-\bar{X}),\ \rho_{yx}=\frac{Cov(x,y)}{S_xS_y},$$

$$R_7=\frac{\bar{Y}}{M},\ C_m=\frac{S_m}{M},\ S_{ym}=\frac{1}{^NC_n}\sum_{i-1}^{^NC_n}(\bar{y}_i-\bar{Y})(m_i-M),\ C_{ym}=\frac{S_{ym}}{\bar{Y}M}\ \text{and}\ S_m^2=\frac{1}{^NC_n}\sum_{i-1}^{^NC_n}(m_i-M)^2.$$

Many more authors in the literature have given different estimators, the latest can be referred as Subramani (2013), Subramani and Kumarapandiyan (2012, 2013), Tailor and Sharma (2009), Yan and Tian (2010), Yadav et al. (2014, 2015), Yadav et al. (2016) and Abid *et al.* (2016).

II. PROPOSED ESTIMATOR

Motivated by Subramani (2016), we have proposed the following general exponential ratio type estimator of population mean of study variable using population median of study variable as,

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

Where α is a characterizing scale and is obtained such that the MSE of t is minimum.

The approximations given below have been made 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 proposed estimator t can be expressed in terms of $e_i's$ ($i=1,2$) as,

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

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

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

$$= \bar{Y}(1+e_0)\left[\alpha + (1-\alpha)\exp\left(\frac{-e_1}{2}(1-\frac{e_1}{2})\right)\right]$$, up to approximation of degree one

$$= \bar{Y}(1+e_0)\left[\alpha + (1-\alpha)\left(1-\frac{e_1}{2}+\frac{e_1^2}{4}+\frac{e_1^2}{8}\right)\right]$$

$$= \bar{Y}(1+e_0)\left[\alpha + (1-\alpha)\left(1-\frac{e_1}{2}+\frac{3e_1^2}{8}\right)\right]$$

$$= \bar{Y}(1+e_0)\left[1-(1-\alpha)\frac{e_1}{2}+(1-\alpha)\frac{3e_1^2}{8}\right]$$

$$= \bar{Y}(1+e_0)\left[1-\alpha_1\frac{e_1}{2}+\alpha_1\frac{3e_1^2}{8}\right]$$, where $\alpha_1 = (1-\alpha)$

$$= \bar{Y}\left[1+e_0-\alpha_1\frac{e_1}{2}-\alpha_1\frac{e_0e_1}{2}+\alpha_1\frac{3e_1^2}{8}\right]$$

$$t-\bar{Y} = \bar{Y}\left[e_0-\alpha_1\frac{e_1}{2}-\alpha_1\frac{e_0e_1}{2}+\alpha_1\frac{3e_1^2}{8}\right] \quad (2)$$

Taking expectations on both sides of above equation and putting values of different expectations, we get the bias of t as,

$$B(t) = \bar{Y}\left[\frac{3}{8}\alpha_1\lambda C_m^2 - \frac{\alpha_1}{2}\lambda C_{ym} - \frac{\alpha_1}{2}\frac{B(m)}{M}\right] \quad (3)$$

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

$$t-\bar{Y} \approx \bar{Y}\left[e_0-\alpha_1\frac{e_1}{2}\right]$$

Squaring on both sides and taking expectation both sides, we get the mean squared error of t up to approximation of degree one as,

$$MSE(t) = \bar{Y}^2 E\left[e_0-\alpha_1\frac{e_1}{2}\right]^2$$

$$= \bar{Y}^2 E\left[e_0^2+\alpha_1^2\frac{e_1^2}{4}-\alpha_1 e_0 e_1\right]$$

$$= \bar{Y}^2\left[E(e_0^2)+\alpha_1^2\frac{E(e_1^2)}{4}-\alpha_1 E(e_0e_1)\right]$$

Putting values of different expectations, we get MSE of t as,

$$MSE(t) = \lambda\bar{Y}^2\left[C_y^2+\frac{\alpha_1^2}{4}C_m^2-\alpha_1 C_{ym}\right] \quad (4)$$

The $MSE(t)$ is minimum for,

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

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

$$\alpha_{1(opt)} = 2\frac{C_{ym}}{C_m^2} \tag{5}$$

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

$$MSE_{\min}(t) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right] \tag{6}$$

III. EFFICIENCY COMPARISON

From equation (6) and the mean squared error of estimator at serial number 1 in Table-1which is mean per unit estimator, 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 (6) the mean squared error of estimator at serial number 2 in Table-1which is Watson (1937) usual regression estimator, we have,

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

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

Thus under above condition, t are better than t_1.

From equation (5) or equivalently equation (8) and the mean squared error of estimator at serial number 3 in Table-1which is Cochran (1940) usual ratio estimator, we have,

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

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

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

Thus under above condition, t are better than t_2.

From equation (6) and the mean squared error of estimator at serial number 4 in Table-1which is Bahl and Tuteja (1991) usual exponential ratio estimator, we have,

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

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

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

Thus under above condition, t are better than t_3.

From equation (6) and the mean squared error of estimator at serial number 5 in Table-1which is Srivastava (1967) generalized ratio estimator of population mean, we have,

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

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

Thus under above condition, t are better than t_4.

Proposed estimator is better than Reddy (1974) and Kadilar (2016) estimators of population mean under the same condition as for Srivastava (1967) estimator mentioned in above equation.

From equation (6) and the mean squared error of estimator at serial number 8 in Table-1which is Subramani (2016) ratio type estimator of population mean, 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 under above condition, t are better than t_7 .

IV. NUMERICAL STUDY

We have considered various populations given in Subramani (2016) for judgment of the theory. Table-1 and Table-2 represent the parameter values for three populations along with constants and mean squared errors of existing and proposed estimators respectively.

Table-1: Parameter values and constants for three natural populations

Parameter	Population-1	Population-2	Population-3
N	34	34	20
n	5	5	5
${}^N C_n$	278256	278256	15504
$\bar{Y}$	856.4118	856.4118	41.5
$\bar{M}$	736.9811	736.9811	40.0552
M	767.5	767.5	40.5
$\bar{X}$	208.8824	199.4412	441.95
R_7	1.1158	1.1158	1.0247
C_y^2	0.125014	0.125014	0.008338
C_x^2	0.088563	0.096771	0.007845
C_m^2	0.100833	0.100833	0.006606
C_{ym}	0.07314	0.07314	0.005394
C_{yx}	0.047257	0.048981	0.005275
ρ_{yx}	0.4491	0.4453	0.6522

Table-2: Mean squared error of various estimators

Estimator	Popln-1	Popln-2	Popln-3
t_0	15640.97	15640.97	2.15
t_1	12486.75	12539.30	1.24
t_2	14895.27	15492.08	1.48
t_3	12498.01	12539.30	1.30
t_4	12486.75	12539.30	1.24
t_5	12486.75	12539.30	1.24
t_6	12486.75	12539.30	1.24
t_7	10926.53	10926.53	1.09
t	9001.52	9001.52	0.95

IV. RESULTS AND CONCLUSION

The present paper proposes a general ratio type estimator of population mean of study variable using population median of study variable. The bias and mean squared error are obtained up to first order of approximation. The proposed estimator is compared with other competing estimators of population mean. The theoretical findings are numerically justified through three different natural populations. It can be seen from Table-2 that the proposed estimator is having least mean squared error among all mentioned competing estimators of population mean which make use of auxiliary information which is collected on additional cost of the survey. Thus it is recommended to use the proposed estimator for improved estimation of the population mean of study variable without using auxiliary information.

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): 353-367.

[2] Bahl, S. and Tuteja, R.K. 1991. Ratio and product type exponential estimator, *Information and Optimization Sciences*, XII (I): 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*: 262-275.

[4] Subramani, J. 2013. Generalized modified ratio estimator of finite population mean, *Journal of Modern Applied Statistical Methods*, 12 (2): 121–155.

[5] Subramani, J. 2016. A new median based ratio estimator for estimation of the finite population mean, *Statistics in Transition New Series,* 17, 4:1-14.

[6] Subramani, J., Kumarapandiyan, G. 2013a. Estimation of population mean using deciles of an auxiliary variable, *Statistics in Transition-New Series*, 14 (1): 75–88.

[7] Subramani, J., Kumarapandiyan, G. 2013b. 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): 137–145.

[8] Watson, D.J. 1937. The estimation of leaf area in field crops, *The Journal of Agricultural Science*, 27(3): 474-483.

[9] Yadav S.K. and Kadilar, C. 2013a. Improved Class of Ratio and Product Estimators, *Applied Mathematics and Computation*, 219: 10726-10731.

[10] Yadav S.K. and Kadilar, C. 2013b. Efficient family of exponential estimator for population mean, *Hacettepe journal of Mathematics and Statistics*, Volume 42, 6: 671 - 677.

[11] Yadav, S.K. and Mishra, S.S. 2015. Developing Improved Predictive Estimator for Finite Population Mean Using Auxiliary Information, *Statistika*, 95, 1: 76-85.

[12] 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: 547-553.