Research article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

Some Exponential Produt Type Estimators Using Auxiliary Attributes.

Bulu Mahanty[1*] and Gopabandu Mishra [2]

[1*].P.G. Department of Statistics ,Utkal uinversiy, Bhubaneswar-751025, Odisha, India.
[2.] P.G. Department of Statistics, Utkal University, Bhubaneswar, -751004, Odisha, India.

ABSTRACT

In this paper some improved exponential product type estimators of finite population mean have been suggested in presence of auxiliary attributes. For construction of estimators we have used a *priori*/*a posteriori* knowledge of coefficient variationand auxiliary attributes .The efficiencies of these estimators are compared with the exponential estimator using auxiliary attribute suggested by Singh et.al.[1]and among themselves with regard to biases and mean square errors both theoretically and numerically

KEYWORDS: Auxiliary attributes, Population proportion, Simple random sampling, Exponential product type estimators, Bias, Mean square error, Efficiency.

*Corresponding author

Mr. Bulu Mahanty

P.G. Department of Community Medicine,

Hi-Tech Medical College,

Bhubaneswar-751025, Odisha, India.

Email. Id : bulu.mahanty@gmail.com., Mob No- +91-7978459396

1. INTRODUCTION

In theory of sampling judicious use of auxiliary information to develop efficient estimators is a long practice. Cochran [2] developed a ratio estimator, when the study variable y and auxiliary variable x are positively correlated. However when "y" and "x" are negatively correlated the ratio estimator does not perform better. Robson[3] and Murthy [4] have suggested a product estimator to estimate population mean $\bar{Y}$ which perform better than mean per unit estimator $\bar{y}$, when y and x are highly negatively correlated.

Singh et.al.[1] have suggested some exponential ratio type and product type estimators for positive correlation and negative correlation exit between study variable "y" and"a" (auxiliary attribute).

In this paper when *a priori or a posteriori* information on population coefficient of variation and auxiliary attributes are available we suggest some improved exponential product type estimators to estimate finite population mean $\bar{Y}$.

Let there be a finite population U consisting of N unit $U = (U_1, U_2, U_3 \ldots U_i \ldots U_N)$. The i[th]unit is indexed by a pair of real value (y_i, a_i) where y_i is the study variable and a_i is the auxiliary attribute. It is assumed that y_i and a_i are negatively correlated and the correlation coefficient between them is denoted by ρ.

2. PROPOSEDESTIMATORS

From the finite population U, a sample of size "n" is selected using simple random sampling without replacement (SRSWOR). We denote the sample mean of study variable $\bar{y}$ and sample proportion $\left(\frac{a}{n}\right) = p$ respectively.

Searls[5] suggested an estimator to estimate population mean $\bar{Y}$ using known population coefficient variation of study variable i.e. $C_y = \frac{S_y}{\bar{Y}}, S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$. The suggested estimator of Searls is given by $\hat{Y}_s = \frac{\bar{y}}{1+\theta_1 C_y^2}$, (2.1)

Where, $\theta_1 = \frac{1}{n} - \frac{1}{N}$.

Following Bhal and Tuteja[6], Singh et.al.[1] have proposed an exponential product type estimator of population mean using population proportion (i.e. in presence of auxiliary attributes), is given by

$$t_{EPP1} = \bar{y} \exp\left[\frac{p - P}{p + P}\right] \qquad (2.2)$$

Where, P and p are the population proportion and sample proportion respectively.

Now we proposed an improved exponential product type estimator of population mean when we have a priori knowledge of coefficient of variation of study variable i.e. C_y and presence of auxiliary attributes.

$$t_{EPP2} = \frac{\bar{y}}{1 + \theta_1 C_y^2} \exp\left[\frac{p - P}{p + P}\right] \tag{2.3}$$

Further, if the a priori knowledge of C_y is not known, we can still construct an improved estimator by considering the estimate of population coefficient of variation of study variable y from the sample. The estimator is given by

$$t_{EPP3} = \frac{\bar{y}}{1 + \theta_1 \hat{C}_y^2} \exp\left[\frac{p - P}{p + P}\right] \tag{2.4}$$

Where, $\hat{C}_y^2 = \dfrac{s_y^2}{\bar{y}^2}$ and $s_y^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$

Following Upadhyaya and Srivastava, a[7], andUpadhyaya and Srivastava, b[8] we suggested another estimator

$$t_{EPP4} = \bar{y}(1 + \theta_1 \hat{C}_y^2) \exp\left[\frac{p - P}{p + P}\right] \tag{2.5}$$

3. BIAS AND MSE OF DIFFERENT ESTIMATORS

Assuming the validity of Taylor's series expansion of $t_{EPP1}, t_{EPP2}, t_{EPP3}$ and t_{EPP4}, considering the expected value to $O\left(\frac{1}{n}\right)$,the bias of the different estimators are given as.

$$\text{B}\,(t_{EPP1}) = \text{E}\,(t_{EPP1}) - \bar{Y} = \theta_1 \bar{Y}\left[\frac{1}{2}C_{11} - \frac{1}{8}C_{20}\right] \tag{3.1}$$

$$\text{B}\,(t_{EPP2}) = \text{E}\,(t_{EPP2}) - \bar{Y} = \theta_1 \bar{Y}\left[\frac{1}{2}C_{11} - \frac{1}{8}C_{20} - C_{02}\right] \tag{3.2}$$

$$\text{B}\,(t_{EPP3}) = \text{E}\,(t_{EPP3}) - \bar{Y} = \theta_1 \bar{Y}\left[\frac{1}{2}C_{11} - \frac{1}{8}C_{20} - C_{02}\right] \tag{3.3}$$

$$\text{B}\,(t_{EPP4}) = \text{E}\,(t_{EPP4}) - \bar{Y} = \theta_1 \bar{Y}\left[\frac{1}{2}C_{11} - \frac{1}{8}C_{20} + C_{02}\right] \tag{3.4}$$

Where, $C_{rs} = \dfrac{\mu_{rs}(p, y)}{P^r \bar{Y}^s}$

Where, $\mu_{rs}(p, y)$ in the (r, s) $^{\text{th}}$ bivariate moments of p and y.

The Mean Square Error (MSE) of different estimators to $O\left(\frac{1}{n^2}\right)$ are given as

MSE (t_{EPP1}) = E ($t_{EPP1} - \bar{Y}$)^2

$$=\bar{Y}^2\left[\theta_1(C_{02}+\frac{1}{4}C_{20}+C_{11})+\theta_2(+C_{12}+\frac{1}{4}C_{21}-\frac{1}{8}C_{30}-\frac{5}{8}C_{11}C_{20}+\frac{7}{64}C_{20}^2)\right] \quad (3.5)$$

Where $\theta_2 = \left(\frac{1}{n^2}-\frac{1}{N^2}\right)$

MSE (t_{EPP2}) = E ($t_{EPP1} - \bar{Y}$)^2

$$=\bar{Y}^2\left[\begin{array}{l}\theta_1(C_{02}+\frac{1}{4}C_{20}+C_{11})+\theta_2(+C_{12}+\frac{1}{4}C_{21}-\frac{1}{8}C_{30}\\ -\frac{5}{8}C_{11}C_{20}+\frac{7}{64}C_{20}^2-3C_{11}C_{02}-\frac{1}{4}C_{20}C_{02}-C_{02}^2)\end{array}\right] \quad (3.6)$$

MSE (t_{EPP3}) = E ($t_{EPP3} - \bar{Y}$)^2

$$=\bar{Y}^2\left[\begin{array}{l}\theta_1(C_{02}+\frac{1}{4}C_{20}+C_{11})+\theta_2(+C_{12}+\frac{1}{4}C_{21}-\frac{1}{8}C_{30}\\ -\frac{5}{8}C_{11}C_{20}+\frac{7}{64}C_{20}^2-C_{11}C_{02}-\frac{1}{4}C_{20}C_{02}+C_{02}^2-2C_{02}C_{03}-C_{02}C_{12})\end{array}\right] \quad (3.7)$$

MSE (t_{EPP4}) = E ($t_{EPP4} - \bar{Y}$)^2

$$=\bar{Y}^2\left[\begin{array}{l}\theta_1(C_{02}+\frac{1}{4}C_{20}+C_{11})+\theta_2(+C_{12}+\frac{1}{4}C_{21}-\frac{1}{8}C_{30}\\ -\frac{5}{8}C_{11}C_{20}+\frac{7}{64}C_{20}^2+\frac{1}{4}C_{20}C_{02}+3C_{11}C_{02}+3C_{02}^2+2C_{02}C_{03}+C_{02}C_{12})\end{array}\right] \quad (3.8)$$

4. COMPARISON OF BIASES AND MEAN SQUARED ERRORS

When the sample is large enough the biases of the estimators t_{EPP1}, t_{EPP2}, t_{EPP3} and t_{EPP4} of $O\left(\frac{1}{n}\right)$ are negligible .

From the above equations (3.2) and (3.3) both the estimators to $O\left(\frac{1}{n}\right)$ the biases are same. i.e.

B (t_{EPP2}) = B (t_{EPP3})

However, the estimators t_{EPP2}, t_{EPP3} and t_{EPP4} are more biased than t_{EPP1}.

The mean square errors t_{EPP1}, t_{EPP2}, t_{EPP3} and t_{EPP4} to $O\left(\frac{1}{n}\right)$ are same.Thus for the purpose of comparison of efficiencies, the MSE of the estimators are considered up to $O\left(\frac{1}{n^2}\right)$.

The comparison of efficiencies of different estimators are made under two cases.

Case I: Under general condition

Case II:Under the Bivariate Symmetrical Distribution..

I. t_{EPP2} is more efficient than t_{EPP1} if

Case I : $$C_{11} > -\frac{1}{12}(C_{20} + 4C_{02}) \tag{4.1}$$

Case II : $$\rho < -\frac{1}{12W}(W^2 + 4) \tag{4.2}$$

Where, $W = \left(\frac{C_{20}}{C_{02}}\right)^{\frac{1}{2}}$

II. t_{EPP3} is more efficient than t_{EPP1} if

Case I : $$C_{11} < \frac{1}{C_{02}}\left(-\frac{1}{4}C_{20}C_{02} - C_{02}^2 - 2C_{03} - C_{12}\right) \tag{4.3}$$

Case II : $$\rho > -\frac{1}{4W}(W^2 - 4) \tag{4.4}$$

III. t_{EPP4} is more efficient than t_{EPP1} if

Case I : $$C_{11} < \frac{1}{3C_{02}}\left(\frac{1}{4}C_{20}C_{02} + 3C_{02}^2 + 2C_{02}C_{03} + C_{02}C_{12}\right) \tag{4.5}$$

Case II : $$\rho < \frac{1}{12W}(W^2 + 4) \tag{4.6}$$

IV. t_{EPP3} is more efficient than t_{EPP2} if

Case I : $$C_{11} < \frac{1}{2C_{02}}(2C_{02}^2 + 2C_{03} + C_{12}) \tag{4.7}$$

Case II : $$\rho < \frac{1}{W} \tag{4.8}$$

V. t_{EPP4} is more efficient than t_{EPP2} if

Case I : $$C_{11} < \frac{1}{4C_{02}}\left(\frac{1}{2}C_{02}C_{20} + 2C_{02}^2 - 4C_{03} - 2C_{12}\right) \tag{4.9}$$

Case II : $$\rho < \frac{1}{8W}(W^2 + 16) \tag{4.10}$$

VI. t_{EPP4} is more efficient than t_{EPP3} if

CaseI : $C_{11} < \frac{1}{6C_{02}}(\frac{1}{2}C_{02}C_{20} + 4C_{02}^{2} - 2C_{03} - 2C_{12})$ (4.11)

Case II : $\rho < \frac{1}{12W}(W^{2} + 8)$ (4.12)

5. NUMERICAL ILLUSTRATION

For comparison of biases and mean square errors, we consider four natural populations. One population is considered from Draper and Smith[9],one population considered from Swain[10], and two populations are consider from Daniel and Cross[11]. Biases are calculated considering terms up to first order of approximation and mean square errors are calculated considering second order of approximations. For calculations we consider four datasets showing N, n, $\bar{Y}$,P and $C_{rs}(p,y)$

Where, $C_{rs}(p,y) = \frac{\mu_{rs}(p,y)}{P^{r}\bar{Y}^{s}}$

Data Set-1

The data for the empirical analysis are taken from Natural Population dataset was considered by Draper and Smith [1998, P.40], Appendix 1A, stream plant data.

Y=Response

P= Predictors

$\bar{Y} = 52.6$, P= 0.92, N = 25, n = 6, $\rho = -0.3416$, $C_{20} = 0.0869$, $C_{02} = 0.1034$, $C_{11} = -0.0324$, $C_{30} = -0.0793$, $C_{03} = -0.0036$, $C_{12} = -0.0032$, $C_{21} = 0.0295$

Data Set-2

The data for the empirical analysis are taken from Natural Population dataset considered by Swain A.K.P.C. [2003, P.274],

Y = No. of Milk Cows , 1956

P = No. of Milk Cows in Rainy Season

$\bar{Y} = 67.3684$ $P = 0.4210$ N = 19, n = 7, $\rho = -0.5899$, $C_{20} = 1.375$, $C_{02} = 1.2385$, $C_{11} = -0.7699$, $C_{30} = 0.5156$, $C_{03} = 2.2061$, $C_{12} = -0.6252$, $C_{21} = 0.2887$

Data Set-3

The data for the empirical analysis are taken from Natural Population dataset considered by Daniel and Cross [2015, P.758], Tab No. 14.3.2

Y = Time (Months)

P = Vital Status (Censored and Dead)

$\overline{Y} = 63.0256, P = 0.5641,$ N = 39 , n = 14 , $\rho = -0.5362$ $C_{20} = 0.7727,$ $C_{02} = 1.2629,$ $C_{11} = -0.5297$ $C_{30} = -0.1756, C_{03} = 2.4308, C_{12} = -0.4395, C_{21} = 0.1204$

Data Set-4

The data for the empirical analysis are taken from Natural Population dataset considered by Daniel and Cross [2015, P.757], Tab No. 14.3.1

Y= Time (Months)

P= Tumor Grade (Low Grade & High Grade)

$\overline{Y} = 63.025, P = 0.3589,$ N =39, n = 20, $\rho = -0.4719$, $C_{20} = 1.7857,$ $C_{02} = 1.2629, C_{11} = -0.7087,$ $C_{30} = 1.4030, C_{03} = 2.4308, C_{12} = -0.5378, C_{21} = -0.5568$

TABLE. 1PERCENT OF RELATIVE BIAS OF ESTIMATORS $t_{EPP1}, t_{EPP2}, t_{EPP3}$ and t_{EPP4} $O\left(\frac{1}{n}\right)$

Data set No.	t_{EPP1}	t_{EPP2}	t_{EPP3}	t_{EPP4}
1	0.0312	0.1510	0.1517	0.0871
2	0.1844	0.5674	1.3146	0.1679
3	0.0804	0.3602	0.4631	0.1737
4	0.0898	0.2843	0.3190	0.0955

TABLE. 2MSE OF ESTIMATORS, $\overline{y}, t_{EPP1}, t_{EPP2}, t_{EPP3}$ and t_{EPP4} $O\left(\frac{1}{n^2}\right)$

Data set No.	$t_0 = \overline{y}$	t_{EPP1}	t_{EPP2}	t_{EPP3}	t_{EPP4}
1	36.2490	33.2524	33.1240	32.8448	34.0557
2	507.1931	336.6183	369.9062	68.9115	608.7238
3	229.7066	167.9355	169.3361	102.4355	224.1438
4	122.2039	97.5004	98.7411	78.4485	121.3179

6. CONCLUSIONS

1. It is observed that the suggested estimators t_{EPP2}, t_{EPP3} and t_{EPP4} are more biased than the estimator t_{EPP1}. However the biases are negligible if sample size is large.
2. Comparing the biases of t_{EPP1}, t_{EPP2}, t_{EPP3} and t_{EPP4} we observed

B(t_{EPP1}) < B(t_{EPP2}) < B(t_{EPP3}) < B(t_{EPP4}) for all populations.

3. Considering the value of MSE of mean per unit estimator ($\bar{y}$) , t_{EPP1} , t_{EPP2}, t_{EPP3} and t_{EPP4} we observed that the MSE of t_{EPP3} is most efficient for the all populations.

ACKNOWLEDGEMENT

We acknowledge Dr. P.R. dash HOD of P.G. department of statistics, Utkal University for giving some fruitful suggestion to improve the article.

REFERENCES

1. Singh, R.Chauhan. P. Sawan, N. and Smarandache, F.: Ratio-Product Type Exponential Estimator For Estimating Finite Population Mean Using Information On Auxiliary Attribute, Book on "Auxiliary information and a priori values in construction of improved estimators" ,Renaissance High Press, Ann, Arbor, USA,1st Edition , 2007;: 18-32.
2. Cochran, W. G.: The estimation of the yield of cereal experiments by sampling for the ratio gain to total procedure, *Jour. Agri. Sc.* 1940; 30(2): 262-275.
3. Robson, D.S. : Application of multivariate polykays to the theory of unbiased ratio-type estimation. Jour. Amer. Stat. Assoc.,1957; 52: 511 - 522.
4. Murthy, M.N.: Product method of estimation. Sankhya , 1964; 26(A): 67 – 74.
5. Searls, D. T.: "The utilization of known coefficient of variation in the estimation procedure", Journal of American Statistical Association, 1964; 59: 1125-1226,.
6. Bahl, S. and Tuteja, R.K.:" Ratio and Product type exponential estimator, Information and Optimization sciences", 1991; XII(I) : 159-163,.
7. Upadhyaya, L.N. and Srivastava, S.R. "A note on the use of coefficient of variation in estimating mean," Jour. Ind. Soc. Agri. Statist.,1976 a; 28(2):97-99..
8. Upadhyaya L.N. and Srivastava,S.R.: "An efficient estimator of mean when population variance is known," Jour. Ind. Soc. Agrl. Stat., 1976b; 28(1): 9-10.
9. Draper, N.R. and Smith, H. "Applied regression Analysis", 3rd Edition John Wiley and Sons, 1998:40, Appendix 1A.
10. Swain AKPC: "Finite population Sampling, theory and methods", 1st Edition, South Asian pub. Pvt. Ltd., 2003; 274.
11. Daniel W.W. and Cross L.C." Biostatistics: Basic concepts and Methodology for the health sciences", International student version, Wiley India Pvt. Ltd., 10thEdition, 2014; 757-758.