

ABSTRACT

In the present article, motivated by Jeelani et al. (2013), we have made an attempt to develop an improved ratio type estimator of population mean using predictive method of estimation by using linear combination of coefficient of skewness and the quartile deviation of auxiliary variable. The mathematical expressions for the bias and mean squared error (MSE) of the proposed estimator up to the first order approximation have been derived. Theoretical efficiency comparison of proposed estimator with the usual ratio estimator, usual product estimator, and Singh et al. (2014) estimators is also made. To amply corroborate the theoretical findings, an empirical study has also been carried out. The suitability of the proposed estimator can be established and appreciated as it has lesser mean squared error, when compared to other widely used estimators.

KEYWORDS: Predictive Approach, Ratio Estimator, Auxiliary Variable, Bias, Mean Squared Error.

I. INTRODUCTION

Technique of drawing sample from the population is known as sampling. Sampling is generally preferred to complete enumeration or census in the instances where the population under reference is potentially infinite and considering complete population for study is exceedingly time consuming, tedious and uneconomical. Theory and practice of sampling primarily deals with the estimation of population parameters. Suitable estimators are used to estimate for the estimation purpose. As is manifest, the efficiency of estimators can be further amplified by the appropriate use of auxiliary information. Auxiliary variable and the main variable under study are highly positively and negatively correlated with each other. In the cases where the auxiliary variable is found to be highly positively correlated with the main variable, the ratio type estimators are resorted for the estimation of parameters. On the contrary, if the auxiliary variable is highly negatively correlated with the main variable under study, product type estimators are preferred for the estimation of population parameters.

In certain situations, the estimation of the population mean of the study variable has received a considerable attention from experts engaged in survey-statistics. For example, in agriculture, the average production of crop is required for further planning or in manufacturing industries and pharmaceutical laboratories, the average life of their products is a necessity for their quality control.

Although, in literature, a great variety of techniques have been used mentioning the use of auxiliary information by means of ratio, product and regression methods for estimating population mean and other parameters. However, some efforts in this direction are reported by many authors. Agrawal and Roy (1999) proposed the efficient estimators of population variance using ratio and regression type predictive estimators, Upadhyaya and Singh (1999) used transformed auxiliary variable and proposed the estimator of population mean, Singh (2003), suggested the improved product type estimator of population mean for negative correlated auxiliary variable, Singh and Tailor (2003), utilized the correlation coefficient of auxiliary and main variable and proposed the improved estimator of population mean, Singh et al. (2004, 2014), proposed improved estimators using power

transformation and the predictive exponential estimators of population mean, respectively, Kadilar and Cingi (2004, 2006), utilized the different parameters of auxiliary variable and proposed improved ratio type estimators of population mean, Yan and Tian (2010), suggested the estimators of population mean using coefficient of skewness of auxiliary variable, Yadav (2011), proposed an efficient ratio estimator of population variance using auxiliary variable, Subramani and Kumarapandiyan (2012), suggested the efficient estimator of population mean using coefficient of variation and the median of auxiliary variable, Solanki et al. (2012), proposed an alternative estimator of population mean using variable for improved estimation of population mean, Onyeka (2012), proposed improved estimator of population mean in post stratified random sampling scheme, Jeelani et al. (2013), suggested modified ratio estimators of population mean using linear combination of coefficient of skewness and the quartile deviation of auxiliary variable, Saini (2013), proposed a predictive class of estimators on two stage random sampling, Yadav and Kadilar (2013) proposed the improved class of ratio and product estimators of population mean, Yadav et al. (2014) suggested the improved ratio estimators for population mean based on median using linear combination of population mean and median of an auxiliary variable and Yadav and Mishra (2015) proposed an improved estimator for estimating population mean under predictive modeling approach using linear combination of estimators proposed by Singh et al.(2014).They have thoroughly used auxiliary information for improved estimation of population mean through ratio and product type estimators for the main characteristic under study.

In sampling theory the use of auxiliary information is common in practice for improving the efficiency of estimators. A vast variety of approaches is available in literature to construct more and more efficient estimators for the population mean including design based and model based methods. Many authors have discussed the estimation of population mean using auxiliary information on design based methods. Herein, we have considered the use of auxiliary information on model based method also known as predictive method of estimation of population mean of study variable. The model-based approach or the predictive method of estimation in sampling theory is based on super population models. This approach assumed that the population under consideration was a realization of super-population random variables containing a super population model. Under this super population model the prior information about the population is formalized and used to predict the non-sampled values of the population that is the finite population quantities, mean and other parameters of the study variable.

Some of the advantages of super population model approach lie in the fact as: the statistical inferences about the population parameters of the study variable may be drawn using the predictive estimation theory of survey sampling or model based theory. The very well-known and popular estimators of population mean in the classical theory are the ratio, regression and product estimator and other estimators. These estimators can be used as predictors in a general prediction theory under a special model. Many authors have used ratio, product and regression type estimators of population parameters for predictive estimation. In this paper we have proposed a predictive estimator of population mean for simple random sampling design using auxiliary variable to construct ratio type estimator by utilizing coefficient of skewness and quartile deviation of auxiliary variable.

Let Y_i ($i = 1, 2, \dots, N$) be the real value taken by the variable under study from the finite population of U of size N . Here, the population parameter to be estimated is the population mean on the basis of observed values of y in an ordered sample of the finite population U of size N . Let S denote the collection of all possible samples from the finite population U . Let $w(s)$ denote the effective sample size, for any given $s \in S$ and \bar{s} denote the collection of all those units of U which are not in S .

We now denote:

$$\bar{y}_s = \frac{1}{w(s)} \sum_{i \in s} y_i$$

$$\bar{y}_{\bar{s}} = \frac{1}{N - w(s)} \sum_{i \in \bar{s}} y_i$$

We have,

$$\bar{Y} = \frac{w(s)}{N} \bar{y}_s + \frac{N - w(s)}{N} \bar{y}_{\bar{s}}$$



Basu (1971), asserted that in the representation of \bar{Y} above the sample mean \bar{y}_s being based on the observed y values on units in the sample s is known, therefore the statistician should attempt a prediction of the mean \bar{y}_s of the unobserved units of the population on the basis of observed units in s .

For $s \in S$ under simple random sampling without replacement (SRSWOR) with sample size $w(s) = n$ and $\bar{y}_s = \bar{y}$, the population mean \bar{Y} is given by

$$\bar{Y} = \frac{n}{N}\bar{y}_s + \frac{(N-n)}{N}\bar{y}_s \quad (1)$$

In view of equation (1) above, an appropriate estimator of population mean \bar{Y} is obtained as

$$\hat{y} = \frac{n}{N}\bar{y} + \frac{(N-n)}{N}T$$

where T is taken as the predictor of \bar{y}_s .

Let x_i ($i = 1, 2, \dots, N$) denote the i^{th} observation of the auxiliary variable x and X_i ($i = 1, 2, \dots, N$) be the values of x on the i^{th} unit of the population U . Auxiliary variable x is correlated with the variable under study y .

Let

$$\bar{X} = \frac{1}{N}\sum_{i=1}^N X_i$$

and

$$\bar{x} = \frac{1}{N}\sum_{i \in s} X_i$$

$$\hat{y} = \frac{n}{N}\bar{y} + \frac{(N-n)}{N}T \quad (2)$$

II. REVIEW OF SOME EXISTING ESTIMATORS UNDER PREDICTIVE MODELING APPROACH

i. Classical Ratio estimator

The usual ratio estimator is given as

$$t_R = \bar{y} \left(\frac{\bar{X}}{\bar{x}} \right) \quad (3)$$

Srivastava (1983), has shown that above estimator is used as predictive estimator of \bar{Y}_s , then the estimator (3) under predictive modelling approach can be written as

$$t_{Re} = \bar{y} \left(\frac{\bar{X}_s}{\bar{x}} \right) \quad (4)$$

$$\text{Bias}(t_{Re}) = \theta \bar{Y} C_x^2 (1 - C)$$

$$\text{MSE}(t_{Re}) = \theta C_x^2 \bar{Y}^2 [C_y^2 + C_x^2 + (1 - 2C)]$$

ii. Classical Product estimator

The usual product estimator is given as

$$t_p = \bar{y} \left(\frac{\bar{x}}{\bar{X}} \right) \quad (5)$$

Srivastava (1983) has used the estimator (5) for estimation of population mean under predictive modelling approach. The above estimator can be written as

$$t_p = \bar{y} \left(\frac{\bar{x}}{\bar{X}_s} \right) \quad (6)$$

$$\text{Bias}(\bar{y}_p) = \theta \bar{Y} C_x^2 C$$

$$\text{MSE}(\bar{y}_p) = \theta \bar{Y}^2 [C_y^2 + C_x^2 + (1 + 2C)]$$

iii. Singh et al.(2014), proposed the following ratio and product type exponential estimator of population estimator of population mean \bar{Y} using Bahl and Tuteja (1991), ratio and product types exponential estimator of population mean as the predictive estimator of \bar{Y}_s respectively as

$$t = t_{Re} = \left[\frac{n}{N}\bar{y} + \left(\frac{N-n}{N} \right) \bar{y} \exp \left(\frac{\bar{X}_s - \bar{x}}{\bar{X}_s + \bar{x}} \right) \right] = \frac{n}{N}\bar{y} + \left(\frac{N-n}{N} \right) \bar{y} \exp \left(\frac{N(\bar{X} - \bar{x})}{N(\bar{X} - \bar{x}) - 2n\bar{x}} \right)$$

$$t = t_{pe} = \left[\frac{n}{N} \bar{y} + \left(\frac{N-n}{N} \right) \bar{y} \exp \left(\frac{\bar{x} - \bar{X}_s}{\bar{x} + \bar{X}_s} \right) \right] = \frac{n}{N} \bar{y} + \left(\frac{N-n}{N} \right) \bar{y} \exp \left(\frac{N(\bar{x} - \bar{X})}{N\bar{X} + (N-2n)\bar{x}} \right)$$

$$\begin{aligned} \text{Bias}(t_{Re}) &= \frac{\theta}{8} \bar{Y} C_x^2 [3 - 4(C + f)] \\ \text{Bias}(t_{pe}) &= \frac{\theta}{8} \bar{Y} C_x^2 \left[4C - \frac{1}{(1-f)} \right] \\ \text{MSE}(t_{Re}) &= \theta \bar{Y}^2 \left[C_y^2 + \frac{C_x^2}{4} (1 - 4c) \right] \\ \text{MSE}(t_{pe}) &= \theta \bar{Y}^2 \left[C_y^2 + \frac{C_x^2}{4} (1 + 4c) \right] \end{aligned}$$

Note: The mean square errors of above Singh et al.(2014) estimators are equal to the mean square error of Bahl and Tuteja(1991) estimator.

Proposed Estimator for estimation of population mean

Jeelani et al. (2013) had suggested an estimator of population mean for simple random sampling using auxiliary variable wherein coefficient of skewness and quartile deviation of auxiliary variable have been utilised.

$$\bar{y} = \left(\frac{\bar{X}_s \beta_1 + QD}{\bar{x} \beta_1 + QD} \right)$$

Motivated by Jeelani et al.(2013), we have proposed the following estimator for estimating population mean using coefficient of skewness and quartile deviations of auxiliary variable under predictive modelling approach taking $T = \left(\frac{\bar{X}_s \beta_1 + QD}{\bar{x} \beta_1 + QD} \right)$ as a predictor in equation (1).

$$\tilde{y} = \frac{n}{N} \bar{y} + \frac{(N-n)}{N} \bar{y} \left(\frac{\bar{X}_s \beta_1 + QD}{\bar{x} \beta_1 + QD} \right) \tag{7}$$

(where $\bar{X}_s = \frac{1}{N-n} \sum_{i \in S} x_i = \frac{N\bar{X} - n\bar{x}}{N-n}$) as predictor for T in equation (1).

Let us consider

$$\bar{y} = \bar{Y}(1 + e_0), \quad \bar{x} = \bar{X}(1 + e_1)$$

$$\text{such that } E(e_0) = E(e_1) = 0, E(e_0^2) = \lambda C_y^2, E(e_1^2) = \lambda C_x^2, E(e_0 e_1) = \lambda C_{yx}, \lambda = \frac{1-f}{n} = \frac{(N-n)}{Nn}, C = \rho \left(\frac{C_y}{C_x} \right)$$

Equation (3) can also be written as

$$\tilde{y} = f\bar{y} + (1-f)\bar{y} \left[\frac{\bar{X}_s \beta_1 + QD}{\bar{x} \beta_1 + QD} \right] \tag{8}$$

$$\text{Let } t_1 = \bar{y} \left[\frac{\bar{X}_s \beta_1 + QD}{\bar{x} \beta_1 + QD} \right] \tag{9}$$

Putting values of \bar{y} and \bar{x} and \bar{X}_s in equation (9)

$$\begin{aligned} t_1 &= \bar{Y}(1 + e_0) \left[\frac{\frac{N\bar{X} - n\bar{X}(1+e_1)}{N-n} \beta_1 + QD}{\bar{X}(1 + e_1) \beta_1 + QD} \right] \\ &= \bar{Y}(1 + e_0) \left[\frac{\frac{N\bar{X} - n\bar{X}(1+e_1)}{N-n} \beta_1 + QD}{\bar{X} \beta_1 + \bar{X} \beta_1 e_1 + QD} \right] \\ &= \bar{Y}(1 + e_0) \left[\frac{\left(\frac{X(N-n) - n\bar{X}e_1}{N-n} \right) + QD}{(\bar{X} \beta_1 + QD) + \bar{X} \beta_1 e_1} \right] \\ &= \bar{Y}(1 + e_0) \left[\frac{\left(1 - \frac{n}{N-n} e_1 \right)}{(1 + \theta e_1)} \right] \quad \left(\text{where } \theta = \frac{\bar{X} \beta_1}{\bar{X} \beta_1 + QD} \right) \end{aligned}$$

[Kumar* *et al.*, 6(6): June, 2017]
ICTM Value: 3.00

$$\begin{aligned} &= \bar{Y}(1 + e_0)\left(1 + \frac{1}{(1-f)}\theta e_1\right)(1 + \theta e_1)^{-1} \\ &= \bar{Y}(1 + e_0)(1 + f_1\theta e_1)(1 - \theta e_1 + \theta^2 e_1^2 + \dots) \quad (\text{where } f_1 = \frac{1}{(1-f)}) \\ &= \bar{Y}(1 + e_0)[1 + (f_1 - 1)\theta e_1 - (f_1 - 1)\theta^2 e_1^2] \quad (10) \end{aligned}$$

Substituting the value of t_1 from (10) in (8), we get
 $\tilde{y} = f\bar{y} + (1 - f)\bar{Y}(1 + e_0)[1 + (f_1 - 1)\theta e_1 - (f_1 - 1)\theta^2 e_1^2]$

Taking first order approximation and taking expectation on both sides
 $\tilde{y} - \bar{Y} = f\bar{Y}e_0 + (1 - f)\bar{Y}[(f_1 - 1)\theta e_1 - (f_1 - 1)\theta^2 e_1^2 + e_0 + (f_1 - 1)\theta e_0 e_1]$ (11)

Taking expectation on both sides,
 $E(\tilde{y} - \bar{Y}) = f\bar{Y}E(e_0) + (1 - f)\bar{Y}[(f_1 - 1)\theta E(e_1) - (f_1 - 1)\theta^2 E(e_1^2) + E(e_0) + (f_1 - 1)\theta E(e_0 e_1)]$ (12)

Substituting the values of $E(e_0)$, $E(e_1)$, $E(e_1^2)$ and $E(e_0 e_1)$, we get Bias(\tilde{y}),
 $\text{Bias}(\tilde{y}) = \lambda f \bar{Y} [\theta C_{yx} - C_x^2]$

Squaring equation (11) and taking expectation on both sides we get the MSE(\tilde{y})
 $\begin{aligned} \text{MSE}(\tilde{y}) &= E[f\bar{Y} + (1 - f\bar{Y}e_0 + f\theta e_1 \bar{Y})]^2 \\ &= \bar{Y}^2 E[e_0 + \theta f e_1]^2 \\ &= \bar{Y}^2 E[e_0^2 + \theta^2 f^2 e_1^2 + 2\theta f e_0 e_1] \\ &= \bar{Y}^2 [E(e_0^2) + \theta^2 f^2 E(e_1^2) + 2\theta f E(e_0 e_1)] \end{aligned}$

Substituting values of $E(e_0^2)$, $E(e_1^2)$ and $E(e_0 e_1)$, we get
 $\text{MSE}(\tilde{y}) = \lambda \bar{Y}^2 [C_y^2 + \theta^2 f^2 C_x^2 + 2\theta f C_{yx}]$ (13)

III. THEORETICAL EFFICIENCY COMPARISON

In this section we are making a theoretical efficiency comparison among the proposed estimator and usual ratio estimator, usual product estimator and Singh et al (2014) estimator.

1. Proposed estimator will be more efficient than the usual ratio estimator if
 $\text{MSE}(\tilde{y}) < \text{MSE}(t_R)$, gives
 $\lambda [C_y^2 + \theta^2 f^2 C_x^2 + 2\theta f C_{yx}] < \theta [C_y^2 + C_x^2 (1 - 2c)]$ (14)

2. Proposed estimator will be more efficient than the usual product estimator if,
 $\text{MSE}(\tilde{y}) < \text{MSE}(\bar{y}_p)$, gives
 $\lambda [C_y^2 + \theta^2 f^2 C_x^2 + 2\theta f C_{yx}] < \theta [C_y^2 + C_x^2 (1 + 2c)]$ (15)

3. Proposed estimator will be more efficient than the Singh et al.(2014) estimator if,
 $\text{MSE}(\tilde{y}) < \text{MSE}(t_{Re})$, gives
 $\lambda [C_y^2 + \theta^2 f^2 C_x^2 + 2\theta f C_{yx}] < \theta [C_y^2 + \frac{C_x^2}{4} (1 - 4c)]$ (16)

and

$\text{MSE}(\tilde{y}) < \text{MSE}(t_{pe})$, gives
 $\lambda [C_y^2 + \theta^2 f^2 C_x^2 + 2\theta f C_{yx}] < \theta [C_y^2 + \frac{C_x^2}{4} (1 + 4c)]$ (17)

IV. EMPIRICAL STUDY

To see the performance of the proposed estimator and the existing estimators and to verify the conditions under which the proposed estimator performs better than other existing estimators of population mean, the numerical example given by Subramani (2016) has been used.

$N = 22, n = 5, \bar{Y} = 22.62, \bar{X} = 1467.546, \rho = 0.9022, s_y = 33.0469, C_y = 1.4609, C_x = 1.7459, S_{xx} = 2562.145, \beta_2 = 13.36, \beta_1 = 3.39, M_d = 534.5, Q_d = 1035, \theta = 0.8278$

Using the above values mean squared error (MSE) of existing estimators and the proposed estimator have been computed and tabulated below:

Table: 1. MSEs of existing and proposed estimators

S.No.	Estimator	MSE
1	t_R	245.74
2	\bar{y}_p	4144.91
3	t_{Re}	252.10
4	t_{pe}	2201.69
5	\tilde{y}	233.19

From table 1, it is observed that the proposed estimator \tilde{y} is more efficient than the existing estimators, as it has lesser MSE.

The percentage relative efficiencies (PRE) of the proposed estimator with respect to existing estimator have also been computed:

Table-2. PRE of the proposed estimator \tilde{y} with respect to existing estimators

S.No.	Estimator	PRE
1	t_R	105.38
2	\bar{y}_p	1777.48
3	t_{Re}	108.11
4	t_{pe}	943.86

From table 2, it is observed that proposed estimator performs better than the usual ratio estimator, usual product estimator, Singh et al.(2014) estimators under predictive modelling approach. Since both the variables y and x are positively correlated to each other, therefore use of product estimator for this data set is not advisable under predictive modeling approach.

V. RESULTS AND CONCLUSIONS

The present research article deals with a new improved ratio estimator for the estimation of population mean under predictive modeling approach. Estimator proposed by Jeelani (2013) has been used as a predictor. The conditions have also been derived for which the proposed estimator is more efficient than the existing estimators. Further, it has also been observed that our proposed estimator is more efficient than the usual ratio estimator, usual product estimator and Singh et al. (2014) estimators under predictive modeling approach. Thus, data sets which satisfy optimal conditions stated at (14),(15), (16),(17)& for which study variable and auxiliary variable are positively correlated, the proposed estimator may be recommended for practical applications under predictive modeling approach

VI. REFERENCES

- [1] Cochran, W.G.(1977): Sampling Techniques. John Wiley and Sons, New York.
- [2] Murthy, M.N. (1977): Sampling Theory & Methods. Statistical Publishing society, Calcutta.
- [3] Sheela Misra, R K Singh and Ashish Kumar Shukla (2013) "Partial Predictive Estimation of Finite Population Variance Using Product Type Estimator" jointly authored to International Journal of Statistics and Analysis (IJS), ISSN 2248-9959 Number 2 (2013) pp. 211-219. Hald, A. (1952). Statistical theory with engineering applications. John Wiley and Sons, Inc, New York.
- [4] [Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok C.(1984): Sampling Theory of Survey with Application. New Delhi.



- [5] Agrawal, M. C., Roy, D. C. Efficient estimators of population variance with regression-type and ratio-type predictor-inputs. *Metron*, 1999, 57(3), pp. 4–13.
- [6] Jeelani, M. I., Maqbool, S., Mir, S. A., Modified Ratio Estimators of Population Mean Using Linear Combination of Coefficient of Skewness and Quartile Deviation. *International Journal of Modern Mathematical Sciences*, 2013, 6, 3, pp. 174–183.
- [7] Kadilar, C., Cingi, H, An improvement in estimating the population mean by using the correlation coefficient. *Hacettepe Journal of Mathematics and Statistics*, Vol. 2006, 35 (1), pp. 103–109.

CITE AN ARTICLE

Misra, S., Kumar, B., Yadav, S. K., Yadav, D. K., & Shukla, A. K. (2017). AN IMPROVED RATIO TYPE PREDICTIVE ESTIMATOR FOR ESTIMATING FINITE POPULATION MEAN USING AUXILIARY INFORMATION. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY*, 6(6), 524-530. doi:10.5281/zenodo.817934