International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 7 July 2022, pp: 956-959 www.ijaem.net ISSN: 2395-5252

Modified Exponential – Ratio Type Estimator for Estimating Population Mean In Simple Random Sampling Scheme

Abdullahi Sifawa[1] and Ran Vijay Kumar Singh[2]
[1] Assistant Lecturer, Sokoto State University, Sokoto State, Nigeria
[2] Professor and Head, Mathematics Department, Kebbi State University of Science and Technology, Aliero , Kebbi State, Nigeria

Submitted: 10-07-2022 Revised: 17-07-2022 Accepted: 21-07-2022

ABSTRACT: In this article, a modified exponential – ratio type estimator for estimating population of the characteristic under study in simple random sampling has been proposed. The expressions for the bias, mean square error and optimum mean square error are derived up to the first order of approximation and found that the optimum mean square error of the proposed estimator is equal to the mean square error of the linear regression estimator. The theoretical and empirical studies carried out reveals that the proposed estimator performs better than the estimators provided in the article.

Key words: ratio, product, bias, mean square error, efficiency, auxiliary information, exponential estimator.

I. INTRODUCTION

In sampling theory, the auxiliary information is used to improve the efficiencies of estimators. The traditional estimators that uses the information on auxiliary variable includes ratio, product, and difference and regression estimator. The ratio estimator proposed by Cochran (1940) is more efficient if the study variable Y is positively correlated with auxiliary variable X, while the product estimator proposed by Murthy (1954) is more efficient if the study variable Y is negatively correlated with auxiliary variable X. The linear regression is more efficient if the line of regression of Y on X is linear and passes through the origin. Later on, several authors modified these estimators with different transformations in order to improve their efficiencies. The main aspect of this article is to modify an estimator that is more efficient than the classical estimators and exponential estimators proposed by Bahl and Tuteja (1991).

Notation

Let $U = \{U_{1,..}, U_N\}$ be a finite population of size N and let (y_i, x_i) be the values of the study variable Y and an auxiliary variable X on the i^{th} unit U_i, I = 1, . . . , N. let a sample of size n be drawn from this population using a simple random sampling without replacement. The aim is to estimate the population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$. Let $S_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ be the population variance of the study variable Y. $\bar{X} = \sum_{i=1}^{N}\frac{X_i}{N}$ be the population mean $S_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$ be the population variance of the auxiliary variable X. $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ be the sample mean of the study variable. $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample mean of the auxiliary variable. $S_{YX} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})$ be the population covariance between the study variable and auxiliary variable. We assume that the population mean $\bar{X}$ and the population variance S_X^2 of an auxiliary variable are known. Let ρ_{yx} be the correlation coefficient between the study variable and the auxiliary variable. Also, assume $C_x = \frac{S_x}{\bar{X}}$ and $C_y = \frac{S_y}{\bar{Y}}$ are the co-efficient of variation between the study variable Y and an auxiliary variable X, and $C_{yx} = \frac{S_{yx}}{\bar{Y}\bar{X}}$ is the coefficient of variation between Y and X. Where $\theta = \frac{1-f}{n}$ and $f = \frac{n}{N}$ is the sampling fraction.

Existing Estimators

It is also known that the sample mean is an unbiased estimator of the population mean and its variance, is given by

$Var(\bar{y}) = \theta\bar{Y}^2 C_y^2$,(1)

Where $\theta = \frac{1-f}{n}$ and $f = \frac{n}{N}$ is the sampling fraction.

The ratio estimator of the population mean $\bar{Y}$ of the study variable as suggested by Cochran (1940) is given as:

$\hat{\bar{Y}}_R = \bar{y}\frac{\bar{X}}{\bar{x}}$(2)

The mean square error of ratio estimator up to the first order of approximation is given by:

$MSE\ (\hat{\bar{Y}}_R) = E(\hat{\bar{Y}}_R - \bar{Y})^2 = \theta\bar{Y}^2(C_x^2 + C_y^2 - 2\rho_{yx}C_xC_y)$ (3)

If the sample size is sufficiently large, then up to the first order of approximation, the ratio estimator will be more efficient than the sample mean estimator if

$\rho_{yx} > \frac{C_x}{2C_y}$(4)

The product estimator introduced by Murthy (1964) and is given by:

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

(5)

The mean square error of product estimator up to the first order of approximation is given by:

$MSE\ (\hat{\bar{Y}}_P) = E(\hat{\bar{Y}}_P - \bar{Y})^2 = \theta\bar{Y}^2(C_x^2 + C_y^2 + 2\rho_{yx}C_xC_y)$ (6)

If the sample size is sufficiently large, then up to the first order of approximation, the product estimator will be more efficient than the sample mean estimator if

$\rho_{yx} < -\frac{C_x}{2C_y}$

(7)

The ratio estimator provides better estimate of population parameter when the line of regression of variable of interest y on auxiliary variable x is linear and passes through the origin. Sometimes the line of regression of variable of interest y on auxiliary variable x is linear but may not pass through origin, in such situation, the difference and the regression estimator was introduced.

The difference is given by:

$\hat{\bar{Y}}_D = \bar{y} + \beta_0(\bar{X} - \bar{x})$ (8)

Where β_0 a preassigned constant, difference estimator is an unbiased estimator of the population mean and its variance is given by:

$MSE(\hat{\bar{Y}}_D) = \left(\frac{1}{n} - \frac{1}{N}\right)(S_y^2 + \beta_0^2S_x^2 - 2\beta_0\rho S_xS_y)$

(9)

The regression estimator is given by:

$\hat{\bar{Y}}_{Reg} = \bar{y} + \hat{\beta}_{yx}(\bar{X} - \bar{x})$ (10)

Where $\hat{\beta}_{yx} = \frac{s_{xy}}{s_x^2}$ is the sample regression coefficient between y and x

The mean square error of regression estimator up to the first order of approximation is:

$MSE\ (\hat{\bar{Y}}_{Reg}) = E\left(\hat{\bar{Y}}_{Reg} - \bar{Y}\right)^2 = \theta\bar{Y}^2\left(1 - \rho^2{}_{yx}\right)C_y^2$(11)

Where $\beta_{yx} = \frac{S_{xy}}{S_x^2}$ is the population regression coefficient between the study variable Y and the auxiliary variable X and $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^r(X_i - \bar{X})^s . s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is the sample variance of the auxiliary variable X and $s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})$ is the sample covariance between x and y.

The exponential ratio and product estimator were suggested to improve the efficiencies of the ratio and product type estimator by Bahl and Tuteja (1991) as:

$\hat{\bar{Y}}_{BT,R} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)$(12)

$\hat{\bar{Y}}_{BT,P} = \bar{y}\exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right)$ (13)

The mean square error of these estimators up to the first order of approximation are given by:

$MSE\ (\hat{\bar{Y}}_{BT,R}) = E(\hat{\bar{Y}}_{BT,R} - \bar{Y})^2 = \theta\bar{Y}^2(C_y^2 + \frac{1}{4}C_x^2 - \rho_{yx}C_xC_y)$ (14)

$MSE\ (\hat{\bar{Y}}_{BT,P}) = E(\hat{\bar{Y}}_{BT,P} - \bar{Y})^2 = \theta\bar{Y}^2(C_y^2 + \frac{1}{4}C_x^2 + \rho_{yx}C_xC_y)$ (15)

Proposed Estimator

Motivated from the work done by Gamze OzelKadilar (2016) and others, a modified exponential-ratio type estimator for estimating population mean in simple random sampling has been proposed as:

$\hat{\bar{Y}}_P = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha} exp\left(\frac{\sqrt{\bar{X}}-\sqrt{\bar{x}}}{\sqrt{\bar{X}}+\sqrt{\bar{x}}}\right)$

(16)

Bias and Mean Square Error

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

(17)

$E(e_0) = E(e_1) = 0,\ E(e_0^2) = \theta C_y^2,\ \ E(e_1^2) = \theta C_x^2$ and $E(e_0e_1) = \theta\rho C_xC_y$ (18)

Where $\theta = \left(\frac{1}{n} - \frac{1}{N}\right)$

To obtain the Bias of the proposed estimator, we Substitute $\bar{y}$ and $\bar{x}$ in the above expression

$$\hat{\bar{Y}}_P = \bar{Y}(1 + e_0)\left(\frac{\bar{X}(1 + e_1)}{\bar{X}}\right)^{\alpha} exp\left(\frac{\sqrt{\bar{X}} - \sqrt{(1 + e_1)\bar{X}}}{\sqrt{\bar{X}} + \sqrt{(1 + e_1)\bar{X}}}\right)$$

$$= \bar{Y}(1 + e_0)(1 + e_1\alpha + \frac{\alpha(\alpha - 1)}{2}e_1^2 + \cdots)\left(1 - \frac{1}{4}e_1 + \frac{5}{32}e_1^2\right)$$

$$= \bar{Y}(1 + e_0 + e_1\alpha + e_1 e_0\alpha + \frac{\alpha(\alpha-1)}{2}e_1^2)\left(1 - \frac{1}{4}e_1 + \frac{5}{32}e_1^2\right)$$

$$= \bar{Y}(1 + e_0 + e_1\alpha + e_1 e_0\alpha + \frac{\alpha(\alpha-1)}{2}e_1^2 - \frac{e_1}{4} - \frac{e_1 e_0}{4} - \frac{\alpha}{4}e_1^2 + \frac{5}{32}e_1^2)$$

Now, $\left(\hat{\hat{Y}}_P - \bar{Y}\right) = \bar{Y}(e_0 + e_1\alpha + e_1 e_0\alpha + \frac{\alpha(\alpha-1)}{2}e_1^2 - \frac{e_1}{4} - \frac{e_1 e_0}{4} - \frac{\alpha}{4}e_1^2 + \frac{5}{32}e_1^2)$ (20)

The bias of the proposed estimator can be obtained by taking the expectation of (20) and substituting the result obtain in (18)

$$B\left(\hat{\hat{Y}}_P - \bar{Y}\right) = \theta\bar{Y}\left[\frac{(4\alpha-1)\rho C_x C_y}{4} + \frac{(16\alpha^2-24\alpha+5)}{32}C_x^2\right]$$

(21)

Squaring (20) and taking the expectation of both sides, the Mean Square Error (MSE)of the proposed estimator $\hat{\hat{Y}}_P$ is

$$MSE\left(\hat{\hat{Y}}_P\right) = \theta\bar{Y}^2\left(C_y^2 + \frac{(4\alpha-1)^2}{16}C_x^2 + \frac{(4\alpha-1)}{2}\rho C_x C_y\right)$$

(22)

Optimum Value of α

To obtain the optimum value of α, we take a partial derivative of the $MSE\left(\hat{\hat{Y}}_P\right)$ with respect to α

$$\frac{\partial MSE(\hat{\hat{Y}}_P)}{\partial\alpha} = 0$$

$$2\alpha\theta\bar{Y}^2C_x^2 - \frac{1}{2}\theta\bar{Y}^2C_x^2 + 2\theta\bar{Y}^2\rho C_x C_y = 0$$

$$\alpha_{opt} = \frac{1/2 - 2\rho C_y}{2C_x}$$

Substituting the value of α_{opt} in (22)above, the $MSE_{min}\left(\hat{\hat{Y}}_P\right)$ is

$$MSE_{min}\left(\hat{\hat{Y}}_P\right) = \theta\bar{Y}^2(1-\rho^2)C_y^2$$

(23)

It shown from (23) that the proposed estimator $\hat{\hat{Y}}_P$at its optimum condition is equally efficient as that of the linear regression estimator.

Efficiency Comparison

In this section efficiency of the proposed estimator is compared with that of existing estimators and conditions are obtained under which the proposed estimator is more efficient.

MSE $(\hat{\hat{Y}}^*_{SS,P})$ < Var $(\bar{y})$ = $\theta\bar{Y}^2C_Y^2\rho^2 > 0$ (24)

MSE $(\hat{\hat{Y}}^*_{SS,P})$ < MSE $(\hat{\hat{Y}}_R)$ = $\theta\bar{Y}^2(C_x - \rho CY2>0$ (25)

MSE $(\hat{\hat{Y}}^*_{SS,P})$ < MSE $(\hat{\hat{Y}}_P)$ =$\theta\bar{Y}^2(C_x + \rho CY2>0$ (26)

MSE $(\hat{\hat{Y}}^*_{SS,P})$ < MSE $(\hat{\hat{Y}}_{BT,R})$ = $\theta\bar{Y}^2\left(\frac{1}{2}C_x - \rho C_Y\right)^2 > 0$(27)

MSE $(\hat{\hat{Y}}^*_{SS,P})$ < MSE $(\hat{\hat{Y}}_{BT,P})$ = $\theta\bar{Y}^2\left(\frac{1}{2}C_x + \rho C_Y\right)^2 > 0$(28)

In the light of the five estimators compared above, we can argue that the proposed estimator is better than the $\bar{y}, \hat{\hat{Y}}_R, \hat{\hat{Y}}_P, \hat{\hat{Y}}_{BT,R}, and \hat{\hat{Y}}_{BT,P}$ estimators, because the condition (24) to (28) are satisfied.

Empirical Study

The efficiency of the proposed estimator has been compared with existing estimators considered in this paper. The Descriptions of the population is given as

Population 1: [Source: Cochran (1977), pp. 196] Let y be the peach production in bushels in an orchard and x be the number of peach trees in the orchard in North Carolina in June 1946. The summary statistics for this data set are: N = 256, n = 100, $\bar{Y}$ = 56.47, $\bar{X}$= 44.45, C_y = 1.42, C_x= 1.40, ρ_{yx} = 0.887.

Population 2: [Source: Murthy (1977), pp. 228] Let y be the output and x be the number of workers. The summary statistics for this data set are: N = 80, n = 10, $\bar{Y}$ = 51.8264, $\bar{X}$ = 2.8513, C_y = 0.3542, C_x = 0.9484, ρ_{yx} = 0.915.

Population 3: [Source: Das (1988)] Let y be the number of agricultural labourers for 1971 and x be the number of agricultural labourers for 1961. The summary statistics for this data set are: N = 278, n = 25, $\bar{Y}$ = 39.068, $\bar{X}$ = 25.111, C_y = 1.4451, C_x = 1.6198, ρ_{yx} = 0.7213. **Population 4:** [Source: Steel, Torrie & Dickey (1960), pp. 282] Let y be the log of lef burn in sacs and x be the chlorine percentage. The summary statistics for this data set are: N = 30, n = 6, $\bar{Y}$ = 0.6860, $\bar{X}$ = 0.8077, C_y = 0.7001, C_x = 0.7493, ρ_{yx} = −0.4996.

Population 5: [Source: Maddala (1977), pp. 282] Let y be the consumption per capita and x be the deflated prices of veal. The summary statistics for this data set are: N = 16, n = 4, $\bar{Y}$ = 7.6375, $\bar{X}$ = 75.4343, C_y = 0.2278, C_x = 0.0986, ρ_{yx} = −0.6823.

Table 1: PREs of different estimators of population mean with respect $\bar{y}$

Population	$\bar{y}$	$\hat{\bar{Y}}_R$	$\hat{\bar{Y}}_P$	$\hat{\bar{Y}}_{BT,R}$	$\hat{\bar{Y}}_{BT,P}$	$\hat{\bar{Y}}_{Pr}$
1	100	448.399	26.874	270.527	47.225	468.975
2	100	468.975	7.651	292.078	19.075	614.345
3	100	156.397	25.817	197.785	47.112	208.452
4	100	31.106	92.932	54.912	133.042	133.261
5	100	56.243	1675.870	74.506	133.063	187.099

From the table above, it is shown that the proposed estimator has higher percentage relative efficiency than the estimators provided in this article.

II. CONCLUSION

The theoretical and empirical studies carried out reveals that the proposed estimator is better than the estimators considered in this article because the conditions are satisfied and has higher percentage relative efficiency, also for optimum value of α the proposed estimator is equally efficient as the linear regression estimator. Hence the proposed estimator is recommended for its practical use for estimating population mean when the auxiliary information is available in survey sampling.

REFERENCES

[1]. Enang E., Uket J., and Ekpenyong E. (2017). A Modified Class of Exponential Type Estimator of Population Mean in Simple Random Sampling Scheme International Journal of Advanced Statistics and Probability, Vol. 5, No. 2, 70-76

[2]. G. O. Kadilar, "A new exponential type estimator for the population mean in Simple Random Sampling" Journal of Modern Applied Statistical Methods Vol. No.2, Vol.15, 2016, 207-214

[3]. Gupta S., Shabbir J., Sousa R. and Corte-Real(2012). Estimation of the mean ofasensitive variableinthepresence of Auxiliary information.Communicationsin Statistics Theory and Methods, Vol.41 (13),2394-2404163

[4]. L. N. Upadhyaya, H.P. Singh, S. Chatterjee and R. Yadav, "Improved ratio and Product exponential type estimators", Journal of Stat. Theo. and Prac. 5(2): 285-302, (2011).

[5]. M.N. Murthy, "Product method of estimation", Samkhya, 26: 294-307(1964).

[6]. R. S. Solanki, H. P. Singh and A. Rathour, "An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys", ISRN Probability and Statistics. Volume 2012: doi:10.5402/2012/657682, (2012).

[7]. S. Bahl and R. K. Tuteja, "Ratio and product type exponential estimator". Information and Optimization Sciences. 12: 159-163, (1991).

[8]. S. Bahl and R. K. Tuteja, "Ratio and product type exponential estimator". Information and Optimization Sciences. 12: 159-163, (1991).

[9]. Singh R.V.K and Ahmed, A. (2015), Improved Exponential Ratio and Product Type Estimators for Finite Population Mean, International Journal of Engineering Science and Innovative Technology, Vol. 4, Issue 3, pp 317-322

[10]. W.J. Cochran, "The estimation of the yields of the cereals experiments by sampling For the ratio of grain to total produce" The journal of agricultural science, 30, 262-275(1940).