Research Journal of Mathematical and Statistical Sciences ______________________________ISSN 2320-6047
Vol. **8(2),** 1-4, May **(2020)** *Res. J. Mathematical and Statistical Sci.*

Modified Regression-Cum-Dual to Ratio-Cum-Product Estimator under Double Sampling

Ishaq Olawoyin Olatunji[1], Ahmed Audu[2] and Rashida Abdulrahman[3]
[1]Department of Statistics, Kano University of Science and Technology, Wudil, Nigeria
[2]Department of Mathematics, Usmanu Danfodiyo University, P.M.B. 2346, Sokoto, Nigeria
[3]State College of Basic and Remedial Studies, Sokoto, Nigeria
babington4u@gmail.com

Available online at: www.iscamaths.com , www.isca.in , www.isca.me
Received 27th August 2018, revised 9th March 2020, accepted 10th April 2020

Abstract

Two estimators t_{R1}^{*} and t_{R2}^{*} for estimating finite population mean in double sampling are suggested. The existing estimator utilizes information on X and Z which can lead it to either over-estimation or under-estimation when they are weakly correlated with the study variable Y, that is, $\rho_{yx} \to 0$ and $\rho_{yz} \to 0$. To overcome the problem of over-estimation or under-estimation in existing estimator, sample means of the study variable of the estimator was transformed using linear regression technique leading two new estimators. The properties (Bias and MSE) of the suggested estimators were derived using Tailor's series expansion. The empirical study conducted revealed that the new estimators out-performed other related existing estimators considered in the study.

Keywords: Regression Estimators, dual to ratio-cum-product estimator, two-phase sampling, mean square error (MSE), efficiency.

Introduction

In theory, additional fact was widely used. The total of a study variables, the use of auxiliary variables in sample survey to obtain an improved design and achieve a better precision in the estimates of population parameters. Work of Neyman[1] can be considered as first work where auxiliary information is used to provide a better estimate of the estimator.

Cochran[2] presented the concept of ratio estimation that used the information obtained with single-phase sampling to develop the ratio estimator in order to estimate population mean. Murthy[3] suggested that the product estimator for negatively correlated study and auxiliary variables, which was similar to ratio estimator. Olkin[4] was the first to use several supplementary characteristics, positively associated with the study variable, using a linear combination of ratio estimator based on each auxiliary variable.

The Estimation of finite population mean with the population proportion of the additional variable was proposed by Shabbir and Gupta[5], while Jhajj *et al*[6] proposed a general family of estimators using information on auxiliary attribute. They used known information on the proportion of the population with a variable (strongly correlated with the variable (Y). This attribute is normally used when the auxiliary variables are not available, for example the quantity of milk produced and a particular cow breed or yield of wheat and a particular variety of wheat.

The use of additional information to derive efficiency of estimates of finite population mean has increased in the theory of sample surveys. We see that in the literature a number of research papers on various types of estimators using additional information are good illustrations in this context such as Handique[7], Adebola *et al*[8], Amoyedo and Adewara[9], Singh and Choudhury[10], Singh and Espejo[11], Kumar and Bahl[12] and Sitter[13]. The ratio and product estimators are more efficient than the regression estimators when the relationship between the study variable y and supplementary variable x. In many fields of application, this line does not go through the origin. Some authors suggested different types of linear transformations of supplementary variables to eliminate this situation. For recent developments on suggested estimators based on these transformations, several authors, including Singh *et al*[14] estimator, Kadilar and Cingi[15], Shabbir and Gupta[5] to estimate the population mean of the study variable y, the ratio estimator $\bar{y}_R$ has been widely used positive correlation between y and x.

When study the auxiliary variable y strongly correlated with auxiliary variable, the use of auxiliary information in ratio and product estimators may increase the accuracy of the estimates. To obtain the most efficient estimator, many authors proposed ratio and product estimators using standard deviation, coefficient of variation, skewness, kurtosis, correlation coefficient, and so on of the auxiliary variable. In this study, we propose a new estimator to estimate the population means of the study variable Y by modifying the estimators Singh et al[16] using the power transformation.

Proposed Estimators

With the study of estimator Singh *et al*[16], the following estimators are proposed.

$$t_{R1}^{*}=\left[\bar{y}+b_{xy}\left(\bar{x}-\bar{X}\right)\right]\left[\frac{\bar{x}^{*}}{\bar{X}}\right]^{\beta_1}\left[\frac{\bar{Z}}{\bar{z}^{*}}\right]^{\beta_2} \quad (1)$$

$$t_{R2}^{*}=\left[\bar{y}+b_{xy}\left(\bar{x}^{*}-\bar{X}\right)\right]\left[\frac{\bar{x}^{*}}{\bar{X}}\right]^{\phi_1}\left[\frac{\bar{Z}}{\bar{z}^{*}}\right]^{\phi_2} \quad (2)$$

Where

$$\bar{x}^{*}=\frac{N\bar{X}-n\bar{x}}{N-n}\ \bar{z}^{*}=\frac{N\bar{Z}-n\bar{z}}{N-n}$$

BIAS and MSE of t_{R1}^{*}

To get the bias and the MSE, let's define $e_0=\frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $e_1=\frac{\bar{x}-\bar{X}}{\bar{X}}$, $e_2=\frac{(\bar{z}-\bar{Z})}{\bar{Z}}$, and using this notations,

$$E(e_0)=E(e_1)=0,\ E(e_0^2)=\theta C_y^2,\ E(e_1^2)=\theta C_x^2,$$
$$E(e_0e_1)=\theta\rho_{xy}C_xC_y,\ E(e_0e_2)=\theta\rho_{YZ}C_yC_Z$$

Where

$$C_x=\frac{S_x}{\bar{X}},\ C_y=\frac{S_y}{\bar{Y}},\ S_x^2=\frac{\sum(X_i-\bar{X})^2}{N-1},$$
$$S_y^2=\frac{\sum(y_i-\bar{Y})^2}{N-1},\ \rho=\frac{S_{xy}}{S_xS_y},$$

$$S_{xy}=\frac{\sum(X_i-\bar{X})(y_i-\bar{Y})}{N-1},\ \theta=\frac{1-f}{n}$$

Bias and MSE of t_{R1}^{*} and t_{R2}^{*}: Equation (1) in terms of e's to second degree in terms of order becomes;

$$t_{R1}^{*}=\left[(1+e_0)\bar{Y}+b_{xy}(1+e_1)\bar{x}-\bar{X}\right]\left[\frac{N\bar{X}-n(1+e_1)\bar{X}}{(N-n)\bar{X}}\right]^{\beta_1}\left[\frac{(N-n)\bar{Z}}{N\bar{Z}-n(1+e_2)\bar{Z}}\right]^{\beta_2} \quad (3)$$

$$t_{R1}^{*}=\left[\bar{Y}+\bar{Y}e_0+b_{xy}\bar{X}e_1\right]\left[1-\lambda e_1\right]^{\beta_1}\left[1-\lambda e_2\right]^{-\beta_2} \quad (4)$$

Where $\lambda=\frac{n}{N-n}$

$$t_{R1}^{*}=\left[\bar{Y}+\bar{Y}e_0+b_{xy}\bar{X}e_1\right]\left[1-\lambda e_1\right]^{\beta_1}\left[1-\beta_2\left(-\lambda e_2\right)+\left(\frac{\beta_2\left(\beta_2+1\right)\lambda^2e_2^2}{2!}\right)\right] \quad (5)$$

$$t_{R1}^{*}-\bar{Y}=\bar{Y}\left[\beta_2\lambda_2e_2+\frac{\beta_2\left(\beta_2+1\right)}{2}\lambda e_2^2-B_1\lambda e_1-B_1B_2\lambda^2e_1e_2+\frac{\beta_1\left(\beta_1-1\right)\lambda^2e_1^2}{2}+e_0+\beta_2\lambda e_0e_2-\beta_1\lambda e_0e_1\right] \quad (6$$
$$+\bar{X}b_{xy}\left[e_0+e_1e_2\lambda e_2-\beta_1\lambda e_1^2\right]$$

Take expectation (6) and apply the results of (3), we obtained bias of t_{R1}^{*} as

$$Bias(t_{R1}^{*})=\bar{Y}\left[\frac{\beta_2(\beta_2+1)}{2}\lambda\theta C_z^2-\beta_1\beta_2\lambda^2\theta\rho_{xz}C_xC_z+\frac{\beta_1(\beta_1-1)}{2}\lambda^2\theta C_x^2+\beta_2\lambda\theta\rho_{yz}C_yC_z-\beta_1\lambda\theta\rho_{yx}C_yC_x\right] \quad (7)$$
$$+\bar{X}b_{xy}\left[\lambda\beta_2\theta\rho_{xz}C_xC_z-\beta_1\lambda\theta C_x^2\right]$$

Square (6), then take expectation and apply the result of (3), we obtain MSE of the t_{R1}^{*} as

$$MSE(t_{R1}^{*})=\bar{Y}^2\theta\left[C_y^2+\beta_1^2\lambda^2C_x^2+\beta_2^2\lambda^2C_z^2-2\beta_1\beta_2g^2\rho_{xz}C_xC_z+2\beta_2g\rho_{yz}C_yC_z-2\beta_1g\rho_{xy}C_yC_x\right] \quad (8)$$
$$+b_{yx}^2\bar{X}^2C_x^2+2\bar{Y}\bar{X}b_{xy}\theta\left[\beta_2\lambda\rho_{xz}C_xC_z-\beta_1\lambda C_x^2+\rho_{xy}C_xC_y\right]$$

In obtaining the optimum values of β_1 and β_2, we differentiate partially (8) w.r.t β_1 and β_2 and then we have;

$$\frac{\partial MSE\left(t_{R1}^{*}\right)}{\partial\beta_1}=\bar{Y}^2\theta\left(2\beta_1\lambda^2C_x^2-2\beta_2\lambda^2\rho_{xz}C_xC_z-2\lambda\rho_{yx}C_yC_x\right)-2\bar{Y}\bar{X}b_{yx}\theta\lambda C_x^2=0 \quad (9)$$

$$\beta_1=\frac{\bar{Y}\left(\beta_2\lambda\rho_{xz}C_z+\rho_{yx}C_y\right)+b_{xy}\bar{X}C_x}{\bar{Y}\lambda C_x} \quad (10)$$

$$\frac{\partial MSE\left(t_{R1}^{*}\right)}{\partial\beta_2}=\bar{Y}^2\theta\left(2\beta_2\lambda^2C_z^2-2\beta_1\lambda^2\rho_{xz}C_xC_z+2\lambda\rho_{yz}C_yC_z\right)+2\bar{Y}\bar{X}b_{yx}\theta\lambda\rho_{xz}C_xC_z=0 \quad (11)$$

$$\beta_2=\frac{\bar{Y}\left(\beta_1\lambda\rho_{xz}C_z+\rho_{yz}C_y\right)-b_{xy}\bar{X}C_x}{\bar{Y}\lambda C_z} \quad (12)$$

Solve (10) and (11) simultaneously, the expression for optimum values of β_1 and β_2 denoted by $\beta_{1_{opt}}$ and $\beta_{2_{opt}}$ are obtained as

$$\beta_{1_{opt}}=\frac{\left(\bar{Y}\rho_{yz}\rho_{xz}C_y-\bar{Y}\bar{X}b_{yx}\rho_{xz}^2C_x+\bar{Y}\rho_{yx}C_y+\bar{X}b_{yx}C_x\right)}{\bar{Y}\lambda C_x\left(1-\rho_{xz}^2\right)} \quad (13)$$

$$\beta_{2_{opt}}=\frac{\bar{Y}\rho_{yz}\rho_{xz}^2C_y-\bar{Y}\bar{X}b_{xy}\rho_{xz}^3C_x+\bar{Y}\rho_{yx}\rho_{xz}C_y+\bar{X}b_{xy}\rho_{xz}C_x+\left(1-\rho_{xz}^2\right)\left(\bar{Y}\rho_{yz}C_y+\bar{X}b_{xy}\rho_{xz}C_x\right)}{\bar{Y}\lambda C_z\left(1-\rho_{xz}^2\right)} \quad (14)$$

Bias and MSE of t^*_{R2}: Equation (2) in terms of e's to second degree in terms of order becomes;

$$t^*_{R2} = \left[\bar{Y}(1+e_0) + b_{xy}\left(\frac{N\bar{X} - n(1+e_1)\bar{X}}{N-n} - \bar{X}\right)\right]\left[\frac{N\bar{X} - n(1+e_1)\bar{X}}{(N-n)\bar{X}}\right]^{\phi_1}\left[\frac{\bar{Z}(N-n)}{n\bar{Z} - n(1+e_2)\bar{Z}}\right]^{\phi_2} \quad (15)$$

$$t^*_{R2} - \bar{Y} = \bar{Y}\left(e_0 - \lambda\phi_1 e_1 + \lambda\phi_2 e_2 - \frac{\phi_1(1-\phi_1)}{2}\lambda^2 e_1^2 + \frac{\phi_2(1-\phi_2)}{2}\lambda^2 e_2^2 - \phi_1\lambda e_0 e_1 + \phi_2\lambda e_0 e_2 - \phi_1\phi_2\lambda^2 e_1 e_2\right) - b_{xy}\lambda\bar{X}\left[e_1 + \phi_2\lambda e_1 e_2 - \phi_1\lambda e_1^2\right] \quad (16)$$

Take expectation of (16) and apply the results of (3) we obtain bias of t^*_{R2} as

$$Bias(t^*_{R2}) = \bar{Y}\theta\left(\frac{\phi_1(1-\phi_1)}{2}\lambda^2 C_x^2 + \frac{\phi_2(1+\phi_2)}{2}\lambda^2 C_z^2 - \phi_1\lambda\rho_{yx}C_yC_x + \phi_2\lambda\rho_{yz}C_yC_z - \phi_1\phi_2\lambda^2\rho_{xz}C_xC_z\right) - b_{xy}\bar{X}\lambda\theta\left(\lambda\phi_2\rho_{xz}C_xC_z - \lambda\phi_1 C_x^2\right) \quad (17)$$

Square (16), then take expectation and apply the result of (3), we obtain MSE of the t^*_{R2} as

$$MSE(t^*_{R2}) = \bar{Y}^2\theta\left(C_y^2 + \phi_1^2\lambda^2 C_x^2 + \phi_2^2\lambda^2 C_z^2 - 2\phi_1\lambda\rho_{yx}C_yC_x + 2\phi_2\lambda\rho_{yz}C_yC_z - 2\phi_1\phi_2\lambda^2\rho_{xz}C_xC_z\right) + b_{yx}^2\bar{X}\lambda^2\theta C_x^2 - 2\bar{Y}b_{yx}\bar{X}\lambda\theta\left(\rho_{yx}C_yC_x - \phi_1\lambda C_x^2 + \phi_2\lambda\rho_{xz}C_xC_z\right) \quad (18)$$

In obtaining the optimum values of ϕ_1 and ϕ_2 in (18), using partial differentiation with respect to ϕ_1 and ϕ_2 and then set to zero as;

$$\frac{\partial MSE(t^*_{R2})}{\partial\phi_1} = 0 \Rightarrow \bar{Y}^2\theta\left(2\phi_1\lambda^2 C_x^2 - 2\lambda\rho_{yx}C_yC_x - 2\phi_2\lambda^2\rho_{xz}C_xC_z\right) + 2\bar{Y}b_{yx}\bar{X}\lambda^2\theta C_x^2 = 0 \quad (19)$$

$$\phi_1 = \frac{\bar{Y}\left(\rho_{yx}C_y + \phi_2\rho_{xz}C_z\right) - b_{xy}\bar{X}\lambda C_x}{\bar{Y}\lambda C_x} \quad (20)$$

$$\frac{\partial MSE(t^*_{R2})}{\partial\phi_2} = 0 \Rightarrow \bar{Y}^2\theta\left(2\phi_2\lambda^2 C_z^2 + 2\lambda\rho_{yz}C_yC_z - 2\phi_1\lambda^2\rho_{xz}C_xC_z\right) - 2\bar{Y}b_{yx}\bar{X}\theta\lambda^2\rho_{xz}C_xC_z = 0 \quad (21)$$

$$\phi_2 = \frac{\bar{Y}\left[\phi_1\lambda\rho_{xz}C_x - \rho_{yz}C_y\right] + b_{xy}\bar{X}\lambda\rho_{xz}C_x}{\bar{Y}\lambda C_z} \quad (22)$$

Solve (20) and (22) simultaneously, the expression for optimum values of ϕ_1 and ϕ_2 denoted by $\phi_{1_{opt}}$ and $\phi_{2_{opt}}$ are obtained as

$$\phi_{1_{opt}} = \frac{\bar{Y}C_y\left(\rho_{xy}C_y - \rho_{xz}\rho_{yz}\right) + b_{xy}\bar{X}C_x\lambda\left(\rho_{xz}^2 - \lambda\right)}{\bar{Y}\lambda C_x\left(\lambda - \rho_{xz}^2\right)} \quad (23)$$

$$\phi_{2_{opt}} = \frac{C_y\left(\rho_{yx}\rho_{xz} - \rho_{yz}\right)}{C_z\left(\lambda - \rho_{xz}^2\right)} \quad (24)$$

Empirical Comparison

In order to study the efficiency of the proposed estimators compare to some existing estimators, the MSEs and PRE of the new estimators and associated estimators are computed using three different populations.

Data 1: Steel and Torrie[17]

N = 30, n = 6, $\bar{Y}$ = 0.6860, $\bar{X}$ =4.6537, $\bar{Z}$ =0.8077, C_y = 0.4803, C_x = 0.2295, C_z = 0.7493,

ρ_{yx} = 0.1794, ρ_{yz} =- 0.4996, ρ_{xz} = 0.4074

Data -2 Population II: Johnston[18]

N = 10, n = 4, $\bar{Y}$ = 52, $\bar{X}$ = 42, $\bar{Z}$ = 200, C_y^2 = 0.0244, C_x^2 = 0.5237, C_z^2 = 0.0021,

ρ_{yx} = 0.80, ρ_{yz} = -0.94, ρ_{xz} =- 0.073

Data 3: Singh[19]

N = 61, n =20, $\bar{Y}$ = 7.46, $\bar{X}$ = 5.31, $\bar{Z}$ = 179.00, C_y^2 =0.5046, C_x^2 = 0.5737, C_z^2 = 0.0633,

ρ_{yx} = 0.7737, ρ_{yz} = -0.2070, ρ_{xz} =- 0.003

Tables-1 to 3 shows the MSEs and PRE of the sample mean, Cochran[2], Murthy[3], linear regression, Singh et al[9] and proposed estimators. The results of the tables revealed that the new estimators have minimum mean square errors and higher PRE and, therefore, are more efficient than the other existing estimators considered.

Table-1: MSE and PRE of new estimators and some existing estimators using Data-1.

Estimators	MSE	PRE
Sample mean	0.01447479	100
Cochran[2]	0.01529802	94.61866
Murthy[3]	0.011244007	146.2008
Linear regression	0.01400892	103.3255
Singh *et al*[16]	0.01185507	122.0979
Proposed estimator (t^*_{R1})	0.003661867	395.2843
Proposed estimator (t^*_{R2})	0.00821608	176.1763

Table-2: MSE and PRE of proposed estimators and some existing estimator using Data 2.

Estimators	MSE	PRE
Sample mean	204.6658	100
Cochran[2]	151.4552	135.1329
Murthy[3]	63.58507	321.8771
Linear regression	73.67967	277.7778
Singh *et al*[16]	34.68806	590.0179
Proposed estimator (t^*_{R1})	24.50162	835.3153
Proposed estimator (t^*_{R2})	22.6830	902.2872

Table-3: MSE and PRE of proposed estimators and some existing estimator using data 3.

Estimators	MSE	PRE
Sample mean	0.9437325	100
Cochran[2]	0.4595863	205.3439
Murthy[3]	2.098533	44.97106
Linear regression	0.4588032	209.1353
Singh *et al*[16]	0.5566835	169.5277
Proposed estimator (t^*_{R1})	0.4458	211.6941
Proposed estimator (t^*_{R2})	0.339116	278.292

Conclusion

From the empirical study, the results of the tables above have shown that the new estimators are more efficient than the existing estimators considered. The new estimators are then recommended for use instead of sample mean, Cochran[2], Murthy[3], linear regression and Singh et al[9].

References

1. Neyman. J, (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, 33, 101–116.
2. 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 Agricultural Science*, 30(2), 262-275.
3. Murthy, M. N. (1967). Sampling Theory and Methods. *Statistical Publishing Society, Calcutta, India.*
4. Olkin, I. (1958). Multivariate Ratio Estimation for Finite Population. *Biometrika*, 45, 154-165.
5. Shabbir, J. and Gupta, S. (2007). On Estimating the Finite Population Mean with Population Proportion of an Auxiliary Variable. *Pakistan Journal of Statistics*, 23, 1-9.
6. Jhajj, H.S., Sharma, M.K. and Grover, L.K. (2006). A Family of Estimator of Population Mean Using Information on Auxiliary Attributes. *Pakistan Journal of Statistics*, 22, 43-50.
7. Handique. B. K (2012). A Class of Regression-Cum-Ratio Estimators in Two - Phase Sampling for Utilizing Information From High Resolution Satellite Data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, I(4), 71-76.
8. Adebola, F. B., Adegoke, N. A., & Sanusi, R. A. (2015). A class of regression estimator with cum-dual ratio estimator as intercept. *International Journal of Probability and Statistics*, 4(2), 42-50.
9. Amoyedo, F. E., & Adewara, A. A. (2019). Improved Modified Ratio Estimator for Estimating Population Mean in Double Sampling Using Information on Auxiliary Attribute. *Annals. Computer Science Series*, 17(2).
10. Singh, B. K., & Choudhury, S. (2012). Double sampling product-cum-dual to ratio estimators for finite population mean in sample surveys. *Int. J. Sci. Eng. Res*, 3(4), 1-7.
11. Singh. H and Espejo (2007). Double Sampling Ratio-product Estimator of a Finite Population Mean in Sample Surveys. *Journal of Applied Statistics*, 34(1), 75–81. https://doi.org/DOI: 10.1080/02664760600994562.
12. Kumar. M and Bahl. S. (2006). Class of dual to ratio estimators for double sampling, *Statistical Papers*, 47, 319-326.
13. Sitter. R. R. (1997). Variance estimation for the regression estimator in two-phase sampling. *Journal of the American Statistical Association*, 92(438), 780-787.
14. Singh, H.P., Singh, R., Espejo, M.R., Pineda, M.D and Nadarajah, S. (2005). On the efficiency of a dual to ratio-cum-product estimator in Sample surveys. *Mathematical Proceedings of the Royal Irish Academy*, 105A(2), 51-56.
15. Kadilar, C. and Cingi H. (2004). Ratio Estimator in simple random sampling. *Applied Mathematics and computation*, 151(3), 893-902.
16. Singh R. and Kumar M. (2011). A family of estimators of population variance using information on auxiliary attribute. Studies in sampling techniques and time series analysis. *Zip Publishing*, 63-70.
17. Steel, R. G. D. and Torrie, J. H. (1960). Principles and Procedures of Statistics. McGraw–Hill Book Co.
18. Johnston, J., (1972). Econometric methods (2nd edn.). McGraw–Hill Book Co., Tokyo.
19. Singh, M. P. (1969). Comparison of some ratio-cum-product estimators. *Sankhya Series B*, 31, 375-378.