ISSN: 2456-799X, Vol.06, No.(1-2) 2021, Pg. 26-31

Oriental Journal of Physical Sciences

www.orientaljphysicalsciences.org

Estimation of Finite Population Mean of Median Based Using Power Transformation

J. O. MUILI[1*] and A. AUDU[2]

[1]Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria.
[2]Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria.

Abstract

This paper deals with the assessment of finite population mean. An estimator is suggested for estimation of finite population mean of study variable. The purpose of this study is to evolve a ratio-type estimator to enhance the proficiency of the existing estimators considered in the study in sample random sampling without replacement using information of auxiliary variable. Expressions of the bias and mean square error (MSE) of the proposed estimator was derived by Taylor series method. The efficiency conditions under which the proposed ratio-type estimator is better than sample mean, ratio estimator, and other estimators considered in this study have been established. Theoretical and empirical findings are incentive and brace the robustness of the proposed estimator for mean estimation. The empirical results shown that the suggested estimator is more efficient than the sample mean, ratio estimator and other estimators.

Article History
Received: 6 October 2021
Accepted: 3 January 2022

Keywords
Auxiliary Variable; Efficiency; Median; Ratio Estimator; Study Variable.

Introduction

Median is one of the auxiliary variables aid in improving the precision of estimates of the finite population mean. Auxiliary variables associated with the study variables have been identified to be helpful in improving the efficiency of ratio, product and regression estimators both at planning and estimation stages. Cochran[1] invented the use of auxiliary information and developed a ratio estimator for population mean. Ratio type estimator provides effective estimate (minimum value of MSE) in comparison to simple mean estimator provided the variable of interest and auxiliary variable are positively associated. If the association between the study and auxiliary variables is positive, then ratio type estimator is applicable. Product estimator is useful where the association between the study variable and auxiliary variable is negative, and more efficient than sample mean. This concept has been utilized by several researchers in order to increase the precisions of ratio and product type estimators in estimating population mean of study

CONTACT J. O. Muili jamiunice@yahoo.com Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria.

© 2021 The Author(s). Published by Oriental Scientific Publishing Company
This is an Open Access article licensed under a Creative Commons license: Attribution 4.0 International (CC-BY).
Doi: http://dx.doi.org/10.13005/OJPS06.01-02.05

variable using auxiliary information for assessments to maximize precisions. Bahl and Tuteja[2] initiated exponential estimators with the used of exponential function in simple random sampling. Singh *et al.*[3] developed exponential ratio estimator with the used of known values of coefficient of variation, correlation coefficient and coefficient of kurtosis. Sanaullah *et al.*[4], Riaz *et al.*[5], Yadav and Adeware[6], Rashid *et al.*[7], and Kadilar[8] have developed different exponential estimators purposely to solve the problem of estimation of finite population mean. Few researchers have made used of median as the only auxiliary information in their works such as Subramani[9] and Kumar *et al.*[10] Other researchers have also used linear combinations of median and other auxiliary parameters for the estimation of population mean such as Subramani and Kumarapandiyan,[11,12,13] Subramani and Kumarapandiyan,[14] Yadav *et al.*[15], Subzar *et al.*[16], Muili *et al.*[17], and Muili *et al.*[18]

The purpose of this research is to evolve a ratio-type estimator to improve the precision of estimation of finite population mean in sample random sampling without replacement with the use of available known information of auxiliary variable.

Let a finite population $\Psi=\{\Psi_1, \Psi_2,...,\Psi_N\}$ having N units where each $\Psi_i = (X_i, Y_i)$, i=1,2,3,4,...,N has a pair of values. X is the auxiliary variable which Y is the study variable and is correlated with X, where y $=\{y_1, y_2,...,y_n\}$ and x $=\{x_1, x_2,...,x_n\}$ are the n sample values. $\bar{y}$ is the sample mean of the study variable and $\bar{x}$ is the sample mean auxiliary variable. Let S_y^2 and S_x^2 be the population mean squares of Y and X respectively and S_y^2 be sample mean square of study variable and S_x^2 be sample mean squares based on the random sample of size n drawn without replacement. N: Population size, Y: Study variable, $\bar{y}$ Sample mean of study variable and $\bar{x}$ sample mean of auxiliary variables, f: Sampling fraction, p_{xy}: Coefficient of correlation between X and Y, $c_y c_x$: Coefficient of variations of study and auxiliary variables, n : Sample size, M : Median of the study variable, ${}^{N}C_n$: Number of possible samples of size n from the population size N, m. Sample Median of the study variable, X : auxiliary variable, $\bar{M}$: Average of Sample Median of the study variable, $\bar{Y}$, $\bar{X}$: Population means of study and auxiliary variables.

$$\bar{X}=\frac{1}{N}\sum_{i=1}^{N}X_i,\ \bar{Y}=\frac{1}{N}\sum_{i=1}^{N}Y_i,\ \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i,\ \gamma=\frac{1-f}{n},\ f=\frac{n}{N}, C_y=\frac{S_y}{\bar{Y}}, C_{ym}=\frac{S_{ym}}{\bar{Y}M},$$

$$S_{ym}=\frac{1}{{}^{N}C_n}\sum_{i=1}^{{}^{N}C_n}(m_i-M)(y_i-\bar{Y}), C_m=\frac{S_m}{M}, S_m^2=\frac{1}{{}^{N}C_n}\sum_{i=1}^{{}^{N}C_n}(m_i-M)^2,\ S_y^2=\frac{1}{N-1}\sum_{i=1}^{N}(Y_i-\bar{Y})^2, C_x=\frac{S_x}{\bar{X}},$$

$$S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})^2,\ C_{yx}=\rho_{yx}C_yC_x,\ \rho_{yx}=\frac{Cov(y,x)}{S_yS_x},\ Cov(y,x)=\frac{1}{{}^{N}C_n}\sum_{i=1}^{{}^{N}C_n}(X_i-\bar{X})(Y_i-\bar{Y}).$$

Literature Review

Sample mean ($\bar{y}$) of simple random sampling is given as:

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i \quad ...(1)$$

And variance of $\bar{y}$ is given by:

$$V(\bar{y})=\gamma\bar{Y}^2C_y^2 \quad ...(2)$$

Watson[19] developed what is known as linear regression estimator of finite population mean using information of auxiliary variable that are highly correlation with variable of interest as

$$\hat{\bar{Y}}_l = \bar{y} + b_{yx}\left(\bar{X}-\bar{x}\right) \quad ...(3)$$

where $b_{yx}=\frac{S_{yx}}{S_x^2}$ is the regression coefficient of study variable on auxiliary variable.

$$V\left(\hat{\bar{Y}}_l\right)=\gamma\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right) \quad ...(4)$$

Cochran[1] ratio estimator is given in (5) for estimation of finite population mean of the study variable as:

$$\hat{\bar{Y}}_r = \frac{\bar{y}}{\bar{x}}\bar{X} \quad ...(5)$$

Bias and mean square error of (5) are:

$$Bias\left(\hat{\bar{Y}}_r\right)=\gamma\bar{Y}^2\left(C_x^2-C_{yx}\right) \quad ...(6)$$

$$MSE\left(\hat{\bar{Y}}_r\right)=\gamma\bar{Y}^2\left(C_y^2+C_x^2-2C_{yx}\right) \quad ...(7)$$

Bahl and Tuteja[2] pioneer exponential estimator of population for estimation of finite population mean given as:

$$\hat{\bar{Y}}_{BT} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \quad ...(8)$$

Bias and mean square error (MSE) of Bahl and Tuteja[2] are given in (9) and (10) respectively

$$Bias\left(\hat{\bar{Y}}_{BT}\right)=\gamma\bar{Y}^2\left(\frac{3}{8}C_x^2-\frac{1}{2}C_{yx}\right) \quad ...(9)$$

$$MSE\left(\hat{\bar{Y}}_{BT}\right)=\gamma\bar{Y}^2\left(C_y^2+\frac{1}{4}C_x^2-C_{yx}\right) \quad ...(10)$$

Subramani[9] developed a ratio-type estimator for estimation of finite population mean given by:

$$\hat{\bar{Y}}_S = \bar{y}\left(\frac{M}{m}\right) \quad ...(11)$$

$$Bias\left(\hat{\bar{Y}}_S\right) = \gamma \bar{Y}^2\left(C_m^2 - C_{ym}\frac{Bias(m)}{M}\right) \quad ...(12)$$

$$MSE\left(\hat{\bar{Y}}_S\right) = \gamma \bar{Y}^2\left(C_y^2 + R^2C_m^2 - 2RC_{ym}\right) \quad ...(13)$$

where $R = \frac{\bar{Y}}{M}$

Proposed Estimator

Motivated by Subramani,[9] of finite population mean based on the information on auxiliary variable for estimation of population mean of study variable is proposed by:

$$\hat{\bar{Y}}_p = \bar{y}\log\left(\frac{M}{m}\right)^{\alpha} \quad ...(14)$$

Derivation of Bias and Mean Square Error of the Proposed Estimator $\left(\hat{\bar{Y}}_p\right)$

Note that: $e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $e_1 = \frac{m-M}{M}$ such that

$\bar{y} = \bar{Y}\left(1+e_0\right)$ *and* $m = M\left(1+e_1\right)$,

$$\left.\begin{array}{l} E(e_0) = 0,\ E(e_1) = \frac{\bar{M}-M}{M} = \frac{Bias(m)}{M}, \\ E(e_0^2) = \gamma C_y^2,\ E(e_1^2) = \gamma C_m^2,\ E(e_0e_1) = \gamma C_{ym} \end{array}\right\} \quad ...(15)$$

Expressing (14) in terms of e_0 & e_1, we have

$$\hat{\bar{Y}}_p = \bar{Y}\left(1+e_0\right)\left(1+e_1\right)^{-\alpha} \quad ...(16)$$

$$\hat{\bar{Y}}_p = \bar{Y}\left(1+e_0\right)\left(1-\alpha e_1 + \frac{\alpha(\alpha+1)}{2}e_1^2\right) \quad ...(17)$$

Simplifying (17) gives (18) as:

$$\hat{\bar{Y}}_p = \bar{Y}\left(1+e_0-\alpha e_1-\alpha e_0e_1+\frac{\alpha(\alpha+1)}{2}e_1^2\right) \quad ...(18)$$

Subtracting $\bar{Y}$ from and taking expectation of both sides

$$E\left(\hat{\bar{Y}}_p - \bar{Y}\right) = \bar{Y}E\left(e_0-\alpha e_1-\alpha e_0e_1+\frac{\alpha(\alpha+1)}{2}e_1^2\right) \quad ...(19)$$

Applying the results of (15) in (19), gives the bias as:

$$Bias\left(\hat{\bar{Y}}_p\right) = \bar{Y}\left(\frac{\alpha(\alpha+1)}{2}\gamma C_m^2 - \alpha\frac{\bar{M}-M}{M} - \alpha\gamma C_{ym}\right) \quad ...(20)$$

Squaring and taking expectation of (19) as

$$MSE\left(\hat{\bar{Y}}_p\right) = \bar{Y}^2E\left(e_0-\alpha e_1\right)^2 \quad ...(21)$$

Expanding (21)

$$MSE\left(\hat{\bar{Y}}_p\right) = \bar{Y}^2E\left(e_0^2+\alpha^2e_1^2-2\alpha e_0e_1\right) \quad ...(22)$$

Applying the results of (15) to (22), gives

$$MSE\left(\hat{\bar{Y}}_p\right) = \gamma\bar{Y}^2\left(C_y^2+\alpha^2C_m^2-2\alpha C_{ym}\right) \quad ...(23)$$

Differentiation of (23) w.r.t α as:

$$\frac{\partial\left(MSE\left(\hat{\bar{Y}}_p\right)\right)}{\partial\alpha} = \gamma\bar{Y}^2\left(C_y^2+\alpha^2C_m^2-2\alpha C_{ym}\right) = 0 \quad ...(24)$$

Making α the subject of the formula, gives

$$\alpha = \frac{C_{ym}}{C_m^2} \quad ...(25)$$

Substituting (25) in (23), gives

$$MSE\left(\hat{\bar{Y}}_p\right)_{\min} = \gamma\bar{Y}^2\left(C_y^2-\frac{C_{ym}^2}{C_m^2}\right) \quad ...(26)$$

EFFICIENCY COMPARISON

Efficiency of the suggested estimator is compared with efficiencies of the existing estimators in the study

Proposed estimator $\left(\hat{\bar{Y}}_p\right)$ of the finite population mean is better than $\bar{y}$ if,

$$MSE\left(\hat{\bar{Y}}_p\right) < V\left(\bar{y}\right)$$

$$\gamma\bar{Y}^2\left(C_y^2-\frac{C_{ym}^2}{C_m^2}\right) < \gamma\bar{Y}^2C_y^2 \quad ...(27)$$

$$C_{ym}^2 > 0 \quad ...(28)$$

Proposed estimator $\left(\hat{\bar{Y}}_p\right)$ is better than Linear Regression[19] estimator $\left(\hat{\bar{Y}}_l\right)$ if,

$$MSE\left(\hat{\bar{Y}}_p\right) < V\left(\hat{\bar{Y}}_l\right)$$

$$\gamma\bar{Y}^2\left(C_y^2-\frac{C_{ym}^2}{C_m^2}\right) < \gamma\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right) \quad ...(29)$$

$$C_{ym}^2 > C_y^2C_m^2\rho_{yx}^2 \quad ...(30)$$

Proposed estimator $\left(\hat{\bar{Y}}_p\right)$ is better than Cochran[1] Ratio estimator $\left(\hat{\bar{Y}}_r\right)$ if,

$$MSE\left(\hat{\bar{Y}}_p\right) < MSE\left(\hat{\bar{Y}}_r\right)$$

$$\gamma\bar{Y}^2\left(C_y^2-\frac{C_{ym}^2}{C_m^2}\right) < \gamma\bar{Y}^2\left(C_y^2+C_x^2-2C_{yx}\right) \quad ...(31)$$

$$C_{ym}^2 > C_m^2\left(2C_{yx}-C_x^2\right) \quad ...(32)$$

Proposed estimator $\left(\hat{\bar{Y}}_p\right)$ is better than Bahl and Tuteja[2] estimator $\left(\hat{\bar{Y}}_{BT}\right)$ if,

$$MSE\left(\hat{\bar{Y}}_p\right) < MSE\left(\hat{\bar{Y}}_{BT}\right)$$

$$\gamma\bar{Y}^2\left(C_y^2 - \frac{C_{ym}^2}{C_m^2}\right) < \gamma\bar{Y}^2\left(C_y^2 + \frac{1}{4}C_x^2 - C_{yx}\right) \quad ...(33)$$

$$C_{ym}^2 > C_m^2\left(C_{yx} - \frac{1}{4}C_x^2\right) \quad ...(34)$$

Proposed estimator $\left(\hat{\bar{Y}}_p\right)$ is better than Subramani[9] estimator $MSE\left(\hat{\bar{Y}}_p\right) < MSE\left(\hat{\bar{Y}}_S\right)$

$$\gamma\bar{Y}^2\left(C_y^2 - \frac{C_{ym}^2}{C_m^2}\right) < \gamma\bar{Y}^2\left(C_y^2 + R^2C_m^2 - 2RC_{ym}\right) \quad ...(35)$$

$$C_{ym}^2 > RC_m^2\left(2C_{ym} - RC_m^2\right) \quad ...(36)$$

When conditions (28), (30), (32), (34), and (36) are satisfied, conclude that the proposed estimator is better (efficient) than sample mean, linear regression,[19] Cochran,[1] Bahl and Tuteja,[2] and Subramani[9] estimators considered in the study.

VERIFIABLE STUDY

To assess the performance of the suggested estimator, a natural population is used as: Source: Subramani[9]

Table 1: Characteristics of Populations

Parameter	Population 1	Population 2	Population 3
N	34	34	20
n	5	5	5
^NC_n	278256.0	278256.0	15504.0
$\bar{Y}$	856.412	856.412	41.50
$\bar{M}$	736.981	736.981	40.055
M	767.50	767.50	40.50
$\bar{X}$	208.882	199.441	441.950
R_7	1.1158	1.1158	1.0247
C^2_y	0.12501	0.12501	0.00834
C^2_X	0.08856	0.09677	0.00785
C^2_m	0.10083	0.10083	0.00661
C_{ym}	0.073140	0.073140	0.0053940
C_{yx}	0.04726	0.04898	0.00528
P_{yx}	0.449	0.445	0.652

Table 1 shows the values of the populations' parameters

Table 2: Mean Square Error of the Estimator

Estimator	Population 1	Population 2	Population 3
Sample Mean $(\bar{y})$	15641.31	15641.31	2.154018
Linear Regression[19] $\left(\hat{\bar{Y}}_l\right)$	12486.6	12539.76	1.237775
Ratio Estimator[1] $\left(\hat{\bar{Y}}_r\right)$	14896.74	15492.29	1.455215
Bahl and Tuteja[2] $\left(\hat{\bar{Y}}_{BT}\right)$	12498.85	12539.89	1.297952
Subramani[9] $\left(\hat{\bar{Y}}_S\right)$	10926.77	10926.77	1.090159
Proposed Estimator $\left(\hat{\bar{Y}}_p\right)$	9003.55	9003.55	1.016205

Table 2 shows MSE of the estimators using the three set of populations. The result revealed that the suggested estimator has minimum mean square error compared to the conventional estimators.

This implies that the proposed estimator is better and can produce better estimates of population mean than the conventional estimators, Bahl and Tuteja[2] and Subramani.[9]

Table 3: Percentage Relative Efficiency of the Estimators

Estimator	Population 1	Population 2	Population 3
Sample Mean ($\bar{y}$)	100	100	100
Linear Regression[19] $\left(\hat{\bar{Y}}_l\right)$	125.2648	124.7337	174.0234
Ratio Estimator[1] $\left(\hat{\bar{Y}}_r\right)$	104.9982	100.9619	148.0206
Bahl and Tuteja[2] $\left(\hat{\bar{Y}}_{BT}\right)$	125.142	124.7324	165.9551
Subramani[9] $\left(\hat{\bar{Y}}_S\right)$	143.1467	143.1467	197.5875
Proposed Estimator $\left(\hat{\bar{Y}}_p\right)$	173.7238	173.7238	211.9669

Table 3 shows PRE of the proposed and some existing estimators using the three set of populations. The result revealed that the suggested estimator has the highest value of PRE compared to the conventional estimators, Bahl and Tuteja[2] and Subramani.[9] This implies that the suggested estimator is more efficient and can produce better estimates of population mean than the conventional estimators, Bahl and Tuteja[2] and Subramani.[9]

Results and Discussion

Proficient ratio estimator of finite population mean is suggested. The properties of the suggested estimator were obtained. Table 2 shows MSE of the suggested and some existing estimators using the three set of populations. The result revealed that the suggested estimator has minimum MSE compared to the conventional estimators, Bahl and Tuteja[2] and Subramani.[9] Estimators. Table 3 shows PRE of the estimators using the three set of populations. The result revealed that the suggested estimator has highest PRE compared to the conventional estimators. This implies that the suggested estimator is more efficient and can produce better estimates of population mean than the conventional estimators, Bahl and Tuteja[2] and Subramani[9] estimators.

Conclusion

Based on the empirical study conducted on the efficiency comparison of the suggested estimator with related estimators, it is obtained that the suggested estimator is highly efficient and can produce better estimates of finite population mean than the conventional and existing estimators considered in the study. Future scope of this study can be study under different sampling schemes like stratified sampling, successive sampling or cluster sampling.

Funding

The authors received no financial support for the research, authorship and publication of this article.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Cochran W. G., 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, (1940).
2. Bahl S. and Tuteja R. K., Ratio and Product Type Exponential Estimator, *Information and Optimization Sciences,* XII (I), 59-163, (1991).
3. Singh R., Kumar M., Chaudhary M. K. and Kadilar C., Improved Exponential Estimator in Stratified Random Sampling. *Pakistan Journal of Statistical Operational Research,* 5(2), 67-82, (2009).
4. Sanuallah A., Khan H., Ali H. A. and Singh R., Improved Exponential Ratio Type Estimator in Survey Sampling. *Journal of Reliability and Scientific Research,* 5(2), 119-132, (2012).
5. Riaz N., Noo-rul-Amin M. and Hanif M., Regression-Cum-Exponential Ratio Type Estimators for the Population Mean.

Middle-East Journal of Scientific Research, 19(12), 1716-1721, (2014).

6. Yadav S.K. and Adewara A.A., On Improved Estimation of Population Mean Using Qualitative Auxiliary Information. *International Journal of Mathematical Theory and Modelling,* 3(11), 42-50, (2013).
7. Rashid R., Noor ul Amin M. and Hanif M., Exponential Estimators for Population Using Transformed Auxiliary Variables. *International Journal of Applied Mathematics and Sciences,* 9(4), 2107-2112, (2015).
8. Kadilar G.O., A New Exponential Type Estimator for the Population Mean in Simple Random Sampling, *Journal of Modern Applied Statistical Methods,* 15, 2, 207-214, (2016).
9. Subramani J., A New Median Based Ratio Estimator for Estimation of the Finite Population Mean, *Statistics in Transition New Series,* 17, 4, 1-14, (2016).
10. Kumar R., Yadav D. K., Misra S., and Yadav S. K., Estimating Population Mean Using Known Median of the Study Variable. *International Journal of Engineering Sciences and Research Technology,* 6(7), 15-21, (2017). doi:10.5281/zenodo.822944
11. Subramani J., and Kumarapandiyan G., Estimation of Population Mean Using Coefficient of Variation and Median of an Auxiliary Variable, *International Journal of Probability and Statistics,* 1 (4), 111–118, (2012a).
12. Subramani J., and Kumarapandiyan G., Variance Estimation Using Median of the Auxiliary Variable. *International Journal of Probability and Statistics,* 1, 6-40, (2012b).
13. Subramani J., and Kumarapandiyan G., Estimation of Population Mean Using Known Median and Coefficient of Skewness, *American Journal of Mathematics and Statistics,* 2 (5), 101–107, (2012c).
14. Subramani J., and Kumarapandiyan G., 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, (2013).
15. Yadav S. K, Mishra S. S. and Shukla A. K., Improved Ratio Estimators for Population Mean Based on Median Using Linear Combination of Population Mean and Median of an Auxiliary Variable. *American Journal of Operational Research,* 4, 2, 21-27, (2014).
16. Subzar M. Maqbool S. Raja T. A., Surya K. P. and Sharma P., Efficient Estimators of Population Mean Using Auxiliary Information Under Simple Random Sampling. *Statistics in Transition new series,* 19(2), 1 – 20, (2018).
17. Muili J.O., Audu A., Odeyale A.B., and Olawoyin I.O., Ratio Estimators for Estimating Population Mean Using Tri-mean, Median and Quartile Deviation of Auxiliary Variable. *Journal of Science and Technology Research,* 1(1), 91-102, (2019).
18. Muili J. O., Adebiyi A. and Agwamba E. N., Improved Ratio Estimators for Estimating Population Mean Using Auxiliary Information, *International Journal of Scientific and Research Publications,* 10, 5, (2020).
19. Watson D. J., The Estimation of Leaf Area in Field Crops, *The Journal of Agricultural Science,* 27, 3, 474-483, (1937).