ISSN: 2456-799X, Vol.08, No.(1) 2023, Pg. 21-26

Oriental Journal of Physical Sciences

www.orientjphysicalsciences.org

A Sine Type Median Based Estimator for the Estimation of Population Mean

MOJEED ABIODUN YUNUSA[1], JAMIU OLASUNKANMI MUILI[2]* AHMED AUDU[1] and RAN VIJAY KUMAR SINGH[2]

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

Abstract

In the literature, there are numerous estimators for estimating population means when auxiliary information is provided. Subramani suggested ratio median based estimator when the median of the study variable is available and the regression estimator was shown to be significantly less efficient than the estimator. In this research, we suggested an estimator for the population mean of the studied variable based on a sine type median. Using Taylor series expansion, the bias and mean square error of the estimator were obtained up to the first order of approximation. The condition under which the proposed estimator is more efficient than the existing estimators was established. An empirical investigation was done to compare the suggested estimator's efficiency to that of the existing estimators, and the numerical findings showed that the proposed estimator is more efficient.

Article History

Received: 03 August 2022

Accepted: 22 November 2022

Keywords:

Auxiliary Variable; Median; MSE; Efficiency.

Introduction

One of the measures of central tendency is the median. It is employed to describe the distribution's center, when there is presence of outliers or extreme values in a dataset. Despite the fact that the mean is the most widely used indicator of central tendency but it has a weakness, that is, it sensitive to outliers or extreme values. The mean is susceptible to the influence of outliers in a dataset, hence, it is not a center-resisting measure. Contrarily, regardless of how significant these changes are, the median is unaffected or only marginally affected by changes in the numerical value of a tiny subset of the observations. A reliable indicator of the center is the median.

In sampling theory, auxiliary information is used in practice increasing the efficiency of estimators for the estimation of population mean of the study variable using the estimation strategies such as ratio, product and regression. When there is a positive correlation between the study and the

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

Doi: https://doi.org/10.13005/OJPS08.01.05

auxiliary variable, the ratio technique of estimation is applied. Cochran[3] was the first to propose the ratio method of estimation under this assumption. Bahl and Tuteja[2] developed the estimators of a finite population's exponential ratio and product type. The main importance of the exponential estimators is that they give more precise estimate than ratio method of estimation when there is little relationship between the study and the auxiliary variable. To address the issue of population mean estimation, authors including Singh *et al.,*[9] Sanaullah,[8] Riaz *et al.,*[7] Yadav and Adewara,[14] Audu and Singh[1] and Yunusa *et al.,*[16] Hussain *et al,*[17] Rather *et al.*[18] have developed various exponential estimators. Because of the resistant nature of median to outliers, some authors have used median as the auxiliary information in their works such as Subramani,[10] Srija *et al.,*[12] Subramani and Kumaranpandiyan,[11] Yadav *et al.,*[15] Muili *et al.,*[4] Muili *et al.,*[5] Muili and Audu,[6] Rather and Yousuf[19] and Zakari *et al.*[20]

In order to increase the accuracy of estimating the mean of a finite population, this study proposes a sine type estimator that can yield estimates that are more closely related to the true population mean of the study variable.

Consider a finite population $V=(V_1, V_2,...,V_N)$. Utilizing the simple random sampling without replacement (SRSWOR) strategy, we select a sample of size n from the population. Let y and x respectively be the study and the auxiliary variables and y_1 and x_1, respectively be the observations on the ith unit. $\bar{x}=n^{-1}\sum_{i=1}^{n}x_i$ and $\bar{y}=n^{-1}\sum_{i=1}^{n}y_i$ are the sample means. $\bar{Y}=N^{-1}\sum_{i=1}^{N}y_i$ and $\bar{X}=N^{-1}\sum_{i=1}^{N}x_i$, be the corresponding population means of the study and auxiliary variables respectively. $s_y^2=(n-1)^{-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$ and $s_x^2=(n-1)^{-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ are the sample variances and $S_y^2=(N-1)^{-1}\sum_{i=1}^{N}(y_i-\bar{Y})^2$ and $S_x^2=(N-1)^{-1}\sum_{i=1}^{N}(x_i-\bar{X})^2$, are the corresponding population variances. *p* is the correlation coefficient between y and x.

n : the sample size, *M*: the median, ${}^{N}C_{M}$: number of size n samples that can be drawn from a population of size N, $\bar{M}$: Average of Sample Median, *m*: Sample median,

Finally let $C_y=S_y\bar{Y}^{-1}$ and $C_x=S_x\bar{X}^{-1}$ respectively be the coefficients of variation for y and x.

$$C_{ym}=\frac{S_{ym}}{\bar{Y}M},\ S_{ym}=\frac{1}{{}^{N}C_{n}}\sum_{i=1}^{{}^{N}C_{n}}(m_i-M)(y_i-\bar{Y}),C_m=\frac{S_m}{M},\ S_m^2=\frac{1}{{}^{N}C_{n}}\sum_{i=1}^{{}^{N}C_{n}}(m_i-M)^2$$

$$\rho=\frac{Cov(y,x)}{S_yS_x},\ \lambda=\left(n^{-1}-N^{-1}\right),$$

$$C_{yx}=\rho C_yC_x,\ Cov(y,x)=\frac{1}{{}^{N}C_{n}}\sum_{i=1}^{{}^{N}C_{n}}(X_i-\bar{X})(Y_i-\bar{Y}).$$

Some Literature-Based Estimators

In this part, we look at a few existing finite population mean estimators in use.

It is said that the typical sample mean is:

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

The estimator's variance is provided by

$$V(\bar{y})=\lambda\bar{Y}^2C_y^2 \qquad ...(2)$$

Cochran[3] initiated the ratio method of estimation and it is defined as:

$$\bar{y}_R=\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \qquad ...(3)$$

As a first order approximation, the MSE of $\bar{y}_R$ is given by

$$MSE(\bar{y}_R)=\lambda\bar{Y}^2\left(C_y^2+C_x^2-2\rho C_yC_x\right) \qquad ...(4)$$

Watson[13] suggested linear regression estimators of $\bar{Y}$, and it is given as:

$$\bar{y}_{lr}=\bar{y}+\beta\left(\bar{X}-\bar{x}\right) \qquad ...(5)$$

$\beta=\frac{S_{xy}}{S_x^2}$ is the regression slope

The variance of $\bar{y}_{lr}$ to first order of approximation is given by

$$Var(\bar{y}_{lr})=\lambda\bar{Y}^2C_y^2\left(1-\rho^2\right) \qquad ...(6)$$

Yunusa *et al*[16] proposed an efficient exponential type estimators for $\bar{Y}$ as

$$T_M=2^{-1}\bar{y}\left(\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha}+\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)\right) \qquad ...(7)$$

The MSE of the estimator is given by

$$MSE(T_M)=\lambda\bar{Y}^2\left(C_y^2+\left(\frac{\alpha}{2}-\frac{1}{4}\right)C_x^2+2\left(\frac{\alpha}{2}-\frac{1}{4}\right)\rho C_yC_x\right) \qquad ...(8)$$

Where, $\alpha=\frac{1}{2}-\frac{2\rho C_y}{C_x}$ is the value of the unknown

in the estimator (T_M)

The estimator's minimum MSE is provided by

$$MSE(T_M)_{\min} = \lambda \bar{Y}^2 C_y^2 \left(1-\rho^2\right) \quad ...(9)$$

Subramani10 developed an efficient estimator of population mean as

$$\bar{y}_S = \bar{y}\left(\frac{M}{m}\right) \quad ...(10)$$

First-order approximation MSE of $\bar{y}_s$ is provided by

$$MSE(\bar{y}_S) = \lambda \bar{Y}^2 \left(C_y^2 + R^2 C_m^2 - 2RC_{ym}\right) \quad ...(11)$$

Where, $R = \frac{\bar{Y}}{M}$

Proposed Estimator

Motivated by the work of Subramani.[10] and Watson,[13] a sine type median-based estimator of the study variable's population mean is suggested by

$$t_d = 2^{-1}\bar{y}\left(\frac{M}{m}+\frac{m}{M}\right) + b\sin\left(\frac{M-m}{M}\right) \quad ...(12)$$

where, *b* is the unknown constant to be obtained by means of differentiating the MSE of (t_d)

Bias and Mean Square Error (MSE) of the proposed estimator's (t_d) Derivation

To derive the proposed estimator's bias and mean square error (t_d) in (12), we write

Let $e_1 = \bar{Y}^{-1}(\bar{y} - \bar{Y})$ and $e_2 = M^{-1}(m-M)$ such that $\bar{y} = \bar{Y}(1= e_1)$, $m=M(1=e_2)$.

$$\left.\begin{array}{l} E(e_1)=0, E(e_2)=\frac{\bar{M}}{M}-1=\frac{Bias(m)}{M}, \\ E(e_1^2)=\lambda C_y^2, E(e_2^2)=\lambda C_m^2, E(e_1e_2)=\lambda C_{ym} \end{array}\right\} \quad ...(13)$$

Expressing estimator t_d in terms of error terms in (13), we have

$$t_d = \frac{\bar{Y}(1+e_1)}{2}\left(\frac{M}{M(1+e_2)}+\frac{M(1+e_2)}{M}\right)+b\sin\left(\frac{M-M(1+e_2)}{M}\right) \quad ...(14)$$

$$t_d = \frac{\bar{Y}(1+e_1)}{2}\left((1+e_2)^{-1}+(1+e_2)\right)+b\sin(-e_2) \quad ...(15)$$

By expanding $(1+e_2)^{-1}$ and $\sin(-e_2)$ to first order of approximation, (15) becomes

$$t_d = \frac{\bar{Y}(1+e_1)}{2}\left(1-e_2+e_2^2+1+e_2\right)-be_2 \quad ...(16)$$

By simplifying and expanding (16) to first order approximation, we have

$$t_d = \bar{Y}\left(1+e_1+\frac{e_2^2}{2}\right)-be_2 \quad ...(17)$$

Subtract $\bar{Y}$ from both sides of (17), we have

$$t_d - \bar{Y} = \bar{Y}\left(e_1+\frac{e_2^2}{2}\right)-be_2 \quad ...(18)$$

Taking expectation of both sides of (18) and apply the results in (13) to obtain the bias of t_d, we have, to first order of approximation

$$Bias(t_d) = \bar{Y}\lambda\frac{C_m^2}{2} - b\left(\frac{\bar{M}}{M}-1\right) \quad ...(19)$$

Squaring and taking the expectation from both sides of equation (18) will gives the estimator's (t_d) MSE as follows

$$MSE(t_d) = \lambda\left(\bar{Y}^2C_y^2 + b^2C_m^2 - 2b\bar{Y}C_{ym}\right) \quad ...(20)$$

By differentiating (20) partially with respect to b, equate to zero and solve for b, we obtain

$$b = \frac{2\bar{Y}C_{ym}}{C_m^2} \quad ...(21)$$

By substituting (21) into (20), we obtain the minimum MSE of t_d as

$$MSE(t_d)_{\min} = \lambda\bar{Y}^2\left(C_y^2 - \frac{C_{ym}^2}{C_m^2}\right) \quad ...(22)$$

Efficiency Comparisons

The conditions under which the proposed estimator outperforms various other estimators were developed in this section.

Estimator t_d is more efficient than the estimator $\bar{y}$ if $Var(\bar{y}) > MSE_{min}(t_d)$

$$C_{ym}^2 > 0 \quad ...(23)$$

Estimator t_d is more efficient than the estimator $\bar{y}_R$ if $MSE(\bar{y}_R) > MSE(t_d)_{min}$

$$C_{ym} > C_m\sqrt{\left(2C_{yx} - C_x^2\right)} \quad ...(24)$$

Estimator t_d is more efficient than the estimator $\bar{y}_{lr}$ if Var $(\bar{y}_{lr})$> MSE $(t_d)_{min}$

$$C_{ym} > pC_y C_m \quad ...(25)$$

Estimator t_d is more efficient than the estimator T_M if MSE $(T_M)_{min}$> MSE $(t_d)_{min}$

$$C_{ym}pCYCm \quad ...(26)$$

Estimator t_d is more efficient than the estimator $\bar{y}_s$ if MSE $(\bar{y}_s)$> MSE $(t_d)_{min}$

$$C_{ym} > C_m\sqrt{R\left(2C_{yx} - RC_m^2\right)} \quad ...(27)$$

Results and Discussion

In this section, numerical analyses to assess the performance of the estimators are illustrated.

Population 1: [Source: Subramani[10]]

N= 34, n= 5, ^NC_n= 278256.0, $\bar{Y}$=856.412, $\bar{M}$ =736.981, M=767.50, $\bar{X}$= 208.882, p= 0.449, R = 1.115, C^2_y=0.12501, C^2_x=0.08856, C^2_m=0.10083, C_{ym}=0.073140, C_{yx}=0.04726

Population 2: [Source: Subramani[10]]

N= 34, n= 5, ^NC_n= 278256.0, $\bar{Y}$=856.412, $\bar{M}$ =736.981, M=767.50, $\bar{X}$= 199.441, p= 0.445, R = 1.1158, C^2_y=0.12501, C^2_x=0.09677, C^2_m=0.10083, C_{ym}=0.073140, C_{yx}=0.04898

Population 3: [Source: Subramani[10]]

N= 20, n= 5, ^NC_n= 15504.0, $\bar{Y}$= 41.50, $\bar{M}$ = 40.055, M=40.50, $\bar{X}$= 441.950, p= 0.652, R = 1.0247, C^2_y=0.00834, C^2_x=0.00785, C^2_m=0.00661, C_{ym}= 0.0053940, C_{yx}=0.00528

Table 1: Proposed and Existing Estimators' MSE and PRE

Estimators	Population 1		Population 2		Population 3	
	MSE	PRE	MSE	PRE	MSE	PRE
$\bar{y}$	15641.31	100.00	15641.31	100.00	2.1540	100.00
$\bar{y}_R$	14896.74	104.998	15492.29	100.962	1.4552	148.021
$\bar{y}_{lr}$	12486.6	125.265	12539.76	124.734	1.2378	174.023
T_M	12486.6	125.265	12539.76	124.734	1.2378	174.023
$\bar{y}_s$	10926.77	143.147	10926.77	143.147	1.0902	197.588
t_d (proposed)	9003.55	173.724	9003.55	173.724	1.0162	211.967

The MSEs and PREs of the proposed estimator and the existing estimators are displayed in Table 1. The results showed that the proposed estimator has minimum MSE and higher PRE among other estimators for the three populations considered in the study. These findings suggest that the proposed estimator outperforms the existing ones taken into account in this investigation.

Conclusion

In this study, a sine type median based estimator is proposed for the estimation of finite population mean. The empirical results revealed that the new proposed estimator outperformed estimators in literature considered in this study. Therefore, it is advised to apply the proposed estimator in actual cases.

Acknowledgements

The authors thank the Oriental Journal of Physical Sciences Editorial Board and anonymous reviewers for their insightful work and suggestions.

Funding

The authors received no any form of funding this research, authorship and publication.

Conflict of Interest

Authors have declared that no competing interest exist.

References

1. Audu A., and Singh R. V. K., Exponential-type regression CompromisedImputation Class of estimators. *Journal of Statistics and Management System,* 24(6), 1253-1266, (2021). (DOI;10.1080/09720510/2020.1814501).
2. Bahl, S. and Tuteja, R. K., Ratio and Product Type Exponential Estimators, *information and Optimization Sciences,* 12, 1, 159-163, (1991).
3. Cochran, W. G., The Estimation of Yields of the cereal Experiments by Sampling for the Ratio of Grain to Total Produce. *The Journal of Agric. Science,* 30,262-275, (1940).
4. 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, (2020a).
5. Muili J. O., Agwamba E. N., Erinola Y. A., Yunusa M. A., Audu A., and Hamzat M. A., Modified ratio-cum-product estimators of population mean using two auxiliary variables. *Asian Journal of Research in Computer Science,* 6(1), 55-65, (2020b). Article no. AIRCOS.59248, ISSN:25818260.
6. Muili J. O., and Audu A., Estimation of Finite Population Mean of Median Based Using Power Transformation, *Oriental Journal of Physical Sciences,* Vol. 06, No. (1-2), 26-31, (2021).
7. 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(2), 1716-1721, (2014)
8. 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).
9. 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).
10. 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).
11. Subramani J., and Kumaranpandiyan 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, (2012).
12. Srija R., Subramani., J., and Jayalakshmi S., Ratio-Cum-Median Based Modified Ratio Estimators for the Estimation of Finite Population Mean with Known Skewness, Kurtosis and Correlation Coefficient. *International Journal of Computers and Mathematical Sciences,* 5(7), 29-34, (2016)
13. Watson D. J., The Estimation of Leaf Area in Field Crops, *The Journal of Agricultural Science,* 27, 3, 474-483, (1983).
14. 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).
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 Auxiliary Variable. *American Journal of Operational Research,* 4, 2, 21-27, (2014).
16. Yunusa M. A., Audu A., Ishaq O. O., and Beki D. O., An efficient exponential type estimator for estimating finite population mean under simple random sampling. *Annals Computer Science Series,* Vol. 19, 1, 46-51, (2021).
17. Hussain S., Sharma M., and Bhat M. I. J., Optimum Exponential Ratio Type Estimators for Estimating the Population Mean, *Journal of Statistics Applications & Probability Letters*, 8, 2, 73-82 (2021).
18. Rather K. U. I, Jeelani M. I., Rizvi S. E. H., and Sharma M. A., New Exponential Ratio Type Estimator for the Population Means Using Information on Auxiliary Attribute. *Journal of Statistics Applications & Probability Letters,* 9, 1, 31-42, (2022).
19. Rather, K. U. I. and Yousuf, R., Improved Estimator for Estimating the Population Mean

In Simple Random Sampling. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 7(5), 16-18, (2020).
20. Zakari, Y., Muili, J. O., Tela, M. N., Danchadi, N. S., & Audu, A., Use of Unknown Weight to Enhance Ratio-Type Estimator in Simple Random Sampling. Lapai Journal of Applied and Natural Sciences, 5(1), 74-81, (2020).