ISSN 2250-3153

Improved Ratio Estimators for Estimating Population Mean Using Auxiliary Information

Jamiu Olasunkanmi Muili[1], Ayodele Adebiyi[2], and Eric Ndukaku Agwamba[3]

[1,3]Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria
[2]Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria.

DOI: 10.29322/IJSRP.10.05.2020.p101111
http://dx.doi.org/10.29322/IJSRP.10.05.2020.p101111

Abstract: The study presents ratios estimators for the finite population mean. The new proposed estimators are based on Subzar *et al*. (2018) estimators. The characteristics of the proposed estimators, i.e. bias and mean square error, were derived up to the first approximation by the Taylor series expansion, and the conditions for its effectiveness were established relative to some existing estimators. The effectiveness of the proposed estimators shows a significant improvement over the estimators considered in the study. The results of the empirical study show that the proposed estimators are more effective than existing estimators based on measurements of the comparison criteria.

KEYWORDS: Auxiliary variable, Mean Square Error, Median, Ratio estimator.

1. INTRODUCTION

In the sample study, there are different ways to use ancillary information to improve the estimator of overall population or population mean. The use of ancillary information was generated by Cochran in 1940 and has made important contributions to modern sampling research by proposing methods for using ancillary information for assessments to maximize precisions. The intermediate (Median) plays an important role in improving the accuracy of the finite population mean. In a situation where the information about an auxiliary variable X is known and the relationship between the study variable and the auxiliary variable is positive, the ratio estimation method is useful. Then if the relationship is negative, the product estimation method can be used efficiently. An estimator that has a least value of variance or mean square error are considered to be the most efficient when comparing. Estimation of population mean has been used by many researchers in order to improve ratio, product and exponential estimators using study auxiliary information. For examples Kadilar and Cingi (2004, 2006) defined ratio estimators of population mean when the values of coefficient of correlation, kurtosis and coefficient of variation are known. Also, Singh and Tailor (2005), Upadhyaya and Singh (1999), etc. In this study, auxiliary information is used with the aim of obtaining highly efficient estimators of population mean in simple random sampling scheme.

Let a finite population $P=\{P_1,P_2,...,P_N\}$ having N units where each $P_i=(X_i, Y_i)$, $i=1,2,...,N$ 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. $\bar{y}$ and $\bar{x}$ are the sample means of the study and auxiliary variables respectively. S_y^2 and S_x^2 are the population mean squares of Y and X respectively and s_y^2 and s_x^2 be respective sample mean squares based on the random sample of size n drawn without replacement. These are the other notations used in this study.

n : Sample size

N : Population size

C_y, C_x : Coefficient of variations of study and auxiliary variables

MR : Population mid-range

Y : Study variable

β_1 : Coefficient of skewness of auxiliary variable

X : Auxiliary variable

ρ : Coefficient of correlation

HL : Hodges-Lehman

$\bar{y}, \bar{x}$: Sample means of study and auxiliary variables

Q_3 : The upper quartile

$\bar{Y}, \bar{X}$: Population means of study and auxiliary variables

Q_r : Inter-quartile range

TM : Tri-Mean

β_2 : Coefficient of kurtosis of auxiliary variable

M_d : Median of the auxiliary

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad \bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i, \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \gamma = \frac{1-f}{n},$$

$$TM = \frac{(Q_1 + 2Q_2 + Q_3)}{4}, s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2, \quad s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2,$$

$$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2, MR = \frac{X_{(1)} + X_{(N)}}{2}, HL = Median\left(\frac{(X_i + X_j)}{2}, 1 \le i \le j \le N\right)$$

$\bar{Y}$ is the population mean of the study variable (Y) whereby the ratio estimator for estimating is:

$$\hat{\bar{Y}}_{ratio} = \frac{\bar{y}}{\bar{x}}\bar{X} \quad (1)$$

$$Bias\left(\hat{\bar{Y}}_{ratio}\right) = \gamma \frac{1}{\bar{X}}\left(RS_x^2 - \rho S_x S_y\right) \quad (2)$$

$$MSE\left(\hat{\bar{Y}}_{ratio}\right) = \gamma\left(S_y^2 + R^2 S_x^2 - 2R\rho S_x S_y\right) \quad (3)$$

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

Kadilar and Cingi (2006) proposed mean estimators for finite population mean inculcated known values of coefficients of kurtosis (β_2), correlation (ρ) and variation (C_x) of auxiliary variables as:

$$\hat{\bar{Y}}_1 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x} + \rho\right)}\left(\bar{X} + \rho\right) \quad (4)$$

$$\hat{\bar{Y}}_2 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}C_x + \rho\right)}\left(\bar{X}C_x + \rho\right) \quad (5)$$

$$\hat{\bar{Y}}_3 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\rho + C_x\right)}\left(\bar{X}\rho + C_x\right) \quad (6)$$

$$\hat{\bar{Y}}_4 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + \rho\right)}\left(\bar{X}\beta_2 + \rho\right) \quad (7)$$

$$\hat{\bar{Y}}_5 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\rho + \beta_2\right)}\left(\bar{X}\rho + \beta_2\right) \quad (8)$$

$$Bias\left(\hat{\bar{Y}}_i\right) = \gamma \frac{S_x^2}{\bar{Y}} R_i^2, \qquad where\ i = 1,2,3,4,5 \quad (9)$$

$$MSE\left(\hat{\bar{Y}}_i\right) = \gamma\left(R_i^2 S_x^2 + S_y^2\left(1 - \rho^2\right)\right) \qquad where\ i = 1,2,3,4,5 \quad (10)$$

$$R_1 = \frac{\bar{Y}}{\bar{X} + \rho},\ R_2 = \frac{\bar{Y}C_x}{\bar{X}C_x + \rho},\ R_3 = \frac{\bar{Y}\rho}{\bar{X}\rho + C_x},\ R_4 = \frac{\bar{Y}\beta_{2x}}{\bar{X}\beta_{2x} + \rho},\ R_5 = \frac{\bar{Y}\rho}{\bar{X}\rho + \beta_{2x}}$$

Subzar *et al.* (2018) came up with ratio type estimators used values of coefficient of skewness $\left(\beta_1\right)$ and coefficient of Kurtosis $\left(\beta_2\right)$ of auxiliary variables as:

$$\hat{\bar{Y}}_6 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_1 + TM\right)}\left(\bar{X}\beta_1 + TM\right) \quad (11)$$

$$\hat{\bar{Y}}_7 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_1 + MR\right)}\left(\bar{X}\beta_1 + MR\right) \quad (12)$$

$$\hat{\bar{Y}}_8 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_1 + HL\right)}\left(\bar{X}\beta_1 + HL\right) \quad (13)$$

$$\hat{\bar{Y}}_9 = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + TM\right)}\left(\bar{X}\beta_2 + TM\right) \quad (14)$$

$$\hat{\bar{Y}}_{10} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + MR\right)}\left(\bar{X}\beta_2 + MR\right) \quad (15)$$

ISSN 2250-3153

$$\hat{\bar{Y}}_{11} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + HL\right)}\left(\bar{X}\beta_2 + HL\right) \quad (16)$$

$$Bias\left(\hat{\bar{Y}}_j\right) = \gamma \frac{S_x^2}{\bar{Y}} R_j^2, \qquad where \;\; j = 6,7,8,9,10,11 \quad (17)$$

$$MSE\left(\hat{\bar{Y}}_j\right) = \gamma\left(R_j^2 S_x^2 + S_y^2\left(1-\rho^2\right)\right) \qquad where \;\; j = 6,7,8,9,10,11 \quad (18)$$

$$R_6 = \frac{\bar{Y}\beta_1}{\bar{X}\beta_1 + TM}, \; R_7 = \frac{\bar{Y}\beta_1}{\bar{Y}\beta_1 + MR}, \; R_8 = \frac{\bar{Y}\beta}{\bar{X}\beta_1 + HL} R_9 = \frac{\bar{Y}\beta_2}{\bar{X}\beta_2 + TM}, \; R_{10} = \frac{\bar{Y}\beta_2}{\bar{X}\beta_2 + MR}, \; R_{11} = \frac{\bar{Y}\beta_2}{\bar{X}\beta_2 + HL}$$

2. PROPOSED ESTIMATORS

After critical studied on the work of Subzar *et al*. (2018), six new ratio estimators for estimating population mean were proposed using value of auxiliary information as:

$$\hat{\bar{Y}}_{p1} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_1 + \phi_1\right)}\left(\bar{X}\beta_1 + \phi_1\right) \quad (19)$$

$$\hat{\bar{Y}}_{p2} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_1 + \phi_2\right)}\left(\bar{X}\beta_1 + \phi_2\right) \quad (20)$$

$$\hat{\bar{Y}}_{p3} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_1 + \phi_3\right)}\left(\bar{X}\beta_1 + \phi_3\right) \quad (21)$$

$$\hat{\bar{Y}}_{p4} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + \phi_1\right)}\left(\bar{X}\beta_2 + \phi_1\right) \quad (22)$$

$$\hat{\bar{Y}}_{p5} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + \phi_2\right)}\left(\bar{X}\beta_2 + \phi_2\right) \quad (23)$$

$$\hat{\bar{Y}}_{p6} = \frac{\bar{y} + b\left(\bar{X} - \bar{x}\right)}{\left(\bar{x}\beta_2 + \phi_3\right)}\left(\bar{X}\beta_2 + \phi_3\right) \quad (24)$$

where $\phi_1 = Md \times TM, \;\; \phi_2 = Md \times MR, \; \phi_3 = Md \times HL$

deriving the properties of proposed estimators, define $e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}$ and $e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}}$ whereby

$\bar{y} = \bar{Y}\left(1 + e_0\right) \;\; and \;\; \bar{x} = \bar{X}\left(1 + e_1\right)$, from the definition of e_0 and e_1, we obtain (25)

$$Bias\left(\hat{\bar{Y}}_{Pi}\right) = \gamma \frac{S_x^2}{\bar{Y}} R_{Pi}^2, \qquad \left(i = 1,2,3,4,5,6\right) \quad (26)$$

$$MSE\left(\hat{\hat{\bar{Y}}}_{Pi}\right)=\gamma\left(R_{Pi}^{2}S_{x}^{2}+S_{y}^{2}\left(1-\rho^{2}\right)\right), \qquad \left(i=1,2,3,4,5,6\right) \qquad (27)$$

where $R_{p1}=\frac{\bar{Y}\beta_1}{\bar{X}\beta_1+\phi_1}$, $R_{p2}=\frac{\bar{Y}\beta_1}{\bar{Y}\beta_1+\phi_2}$, $R_{p3}=\frac{\bar{Y}\beta_1}{\bar{X}\beta_1+\phi_3}$ $R_{p4}=\frac{\bar{Y}\beta_2}{\bar{X}\beta_2+\phi_1}$,

$R_{p5}=\frac{\bar{Y}\beta_2}{\bar{X}\beta_2+\phi_2}$, $R_{p6}=\frac{\bar{Y}\beta_2}{\bar{X}\beta_2+\phi_3}$

2.1 EFFICIENCY COMPARISONS

We compare the effectiveness and efficiency of the proposed estimators with other estimators in this study.

The $\hat{\hat{\bar{Y}}}_{pi}$ of estimators of the finite population mean are better than $\hat{\hat{\bar{Y}}}_{r}$ when,

$$MSE\left(\hat{\hat{\bar{Y}}}_{pi}\right)<MSE\left(\hat{\hat{\bar{Y}}}_{r}\right) \qquad i=1,2,3,4,5,6$$

$$\left(R_{pi}^{2}S_{x}^{2}+S_{y}^{2}\left(1-\rho^{2}\right)\right)<\left(S_{y}^{2}+R^{2}S_{x}^{2}-2R\rho S_{x}S_{y}\right) \qquad (28)$$

The $\hat{\hat{\bar{Y}}}_{pi}$ of proposed estimators of the population mean are better than $\hat{\hat{\bar{Y}}}_{j}$ when,

$$MSE\left(\hat{\hat{\bar{Y}}}_{pi}\right)<MSE\left(\hat{\hat{\bar{Y}}}_{j}\right) \qquad i=1,2,3,4,5,6 \quad j=1,2,3,\ldots,11$$

2.2 NUMERICAL ILLUSTRATION

Table 1: Populations Characteristics [Subzar *et al*. (2018)]

Parameter	Population I	Population II
N	34	34
n	20	20
$\bar{Y}$	856.4117	856.4117
$\bar{X}$	199.4412	208.8823
ρ	0.4453	0.4491
S_y	733.1407	733.1407
C_y	0.8561	0.8561
S_x	150.2150	150.5059

C_x	0.7531	0.7205
β_2	1.0445	0.0978
β_1	1.1823	0.9782
M_d	142.5	150
MR	320	284.5
TM	165.562	162.25
HL	184	190

Table 2: Constant and Bias of existing and proposed estimators

Estimator	**Population I**			**Population II**		
	Constant	Bias	MSE	Constant	Bias	MSE
$\hat{\bar{Y}}_{ratio}$	4.294	4.9399	10961.1	4.100	4.2696	10540.0
$\hat{\bar{Y}}_1$	4.284	9.9578	17399.5	4.091	9.1147	16639.9
$\hat{\bar{Y}}_2$	4.281	9.9432	17387.1	4.088	9.0995	16626.9
$\hat{\bar{Y}}_3$	4.258	9.8348	17294.2	4.069	9.0149	16554.4
$\hat{\bar{Y}}_4$	4.285	9.9597	17401.1	4.011	8.7630	16338.7
$\hat{\bar{Y}}_5$	4.244	9.7711	17239.7	4.096	9.1349	16654.2
$\hat{\bar{Y}}_6$	2.523	3.4523	11828.4	2.285	2.8440	11269.8
$\hat{\bar{Y}}_7$	1.822	1.8003	10413.6	1.714	1.5994	10203.9
$\hat{\bar{Y}}_8$	2.412	3.1557	11574.4	2.124	2.4578	10939.1
$\hat{\bar{Y}}_9$	2.393	3.1052	11531.1	0.458	0.1145	8932.2
$\hat{\bar{Y}}_{10}$	1.693	1.5551	10203.6	0.275	0.0411	8869.3
$\hat{\bar{Y}}_{11}$	2.280	2.8202	11287.0	0.398	0.0863	8908.0
$\hat{\bar{Y}}_{p1}$	0.042	0.0010	8872.60	0.034	0.000635	8834.70
$\hat{\bar{Y}}_{p2}$	0.022	0.0003	8871.99	0.020	0.000208	8834.33
$\hat{\bar{Y}}_{p3}$	0.038	0.0008	8872.44	0.029	0.000464	8834.55
$\hat{\bar{Y}}_{p4}$	0.038	0.0008	8872.42	0.003	0.000006	8834.16
$\hat{\bar{Y}}_{p5}$	0.020	0.0002	8871.94	0.002	0.000002	8834.16
$\hat{\bar{Y}}_{p6}$	0.034	0.0006	8872.30	0.003	0.000005	8834.16

Percentage Relative Efficiency (PRE) of the estimators is given as:

$$PRE = \frac{V\left(\hat{\bar{Y}}_{ratio}\right)}{MSE\left(\hat{\bar{Y}}_i\right)} X100 \quad (30)$$

$\hat{\bar{Y}}_i$ are estimators in the study.

Table 3: PRE of the Proposed Estimators and existing Estimators for Population I

	$\hat{\bar{Y}}_{p1}$	$\hat{\bar{Y}}_{p2}$	$\hat{\bar{Y}}_{p3}$	$\hat{\bar{Y}}_{p4}$	$\hat{\bar{Y}}_{p5}$	$\hat{\bar{Y}}_{p6}$
$\hat{\bar{Y}}_{ratio}$	123.5388	123.5473	123.541	123.5413	123.548	123.5429
$\hat{\bar{Y}}_1$	196.1037	196.1172	196.1073	196.1077	196.1183	196.1104
$\hat{\bar{Y}}_2$	195.964	195.9775	195.9675	195.968	195.9786	195.9706
$\hat{\bar{Y}}_3$	194.9169	194.9303	194.9205	194.9209	194.9314	194.9235
$\hat{\bar{Y}}_4$	196.1218	196.1353	196.1253	196.1257	196.1364	196.1284
$\hat{\bar{Y}}_5$	194.3027	194.316	194.3062	194.3066	194.3171	194.3093
$\hat{\bar{Y}}_6$	133.3138	133.323	133.3162	133.3165	133.3237	133.3183
$\hat{\bar{Y}}_7$	117.3681	117.3761	117.3702	117.3705	117.3768	117.372
$\hat{\bar{Y}}_8$	130.4511	130.46	130.4534	130.4537	130.4608	130.4555
$\hat{\bar{Y}}_9$	129.963	129.972	129.9654	129.9657	129.9727	129.9674
$\hat{\bar{Y}}_{10}$	115.0012	115.0091	115.0033	115.0036	115.0098	115.0051
$\hat{\bar{Y}}_{11}$	127.2119	127.2206	127.2142	127.2144	127.2213	127.2162

Table 4: PRE of the Proposed Estimators and existing Estimators for Population II

	$\hat{\bar{Y}}_{p1}$	$\hat{\bar{Y}}_{p2}$	$\hat{\bar{Y}}_{p3}$	$\hat{\bar{Y}}_{p4}$	$\hat{\bar{Y}}_{p5}$	$\hat{\bar{Y}}_{p6}$
$\hat{\bar{Y}}_{ratio}$	119.3023	119.3073	119.3043	119.3096	119.3096	119.3096
$\hat{\bar{Y}}_1$	188.3471	188.355	188.3503	188.3586	188.3586	188.3586
$\hat{\bar{Y}}_2$	188.1999	188.2078	188.2031	188.2114	188.2114	188.2114
$\hat{\bar{Y}}_3$	187.3793	187.3872	187.3825	187.3908	187.3908	187.3908
$\hat{\bar{Y}}_4$	184.9378	184.9455	184.9409	184.9491	184.9491	184.9491
$\hat{\bar{Y}}_5$	188.5089	188.5168	188.5121	188.5205	188.5205	188.5205
$\hat{\bar{Y}}_6$	127.5629	127.5682	127.5651	127.5707	127.5707	127.5707
$\hat{\bar{Y}}_7$	115.498	115.5028	115.4999	115.505	115.505	115.505

$\hat{\bar{Y}}_8$	123.8197	123.8249	123.8218	123.8273	123.8273	123.8273
$\hat{\bar{Y}}_9$	101.1036	101.1078	101.1053	101.1098	101.1098	101.1098
$\hat{\bar{Y}}_{10}$	100.3916	100.3958	100.3933	100.3978	100.3978	100.3978
$\hat{\bar{Y}}_{11}$	100.8297	100.8339	100.8314	100.8358	100.8358	100.8358

3. RESULT AND DISCUSSION

Six new estimators for estimating population mean are proposed in the study by inculcating known values of population parameters of the auxiliary information. Real life populations were used to show the effect or performance of the proposed estimators over ratio and existing estimators. The features of the proposed estimators i.e. the mean square errors (MSE), constant and bias are established and compared to other estimators.

4. CONCLUSION

Going by the results in tables 2, 3, and 4, it shown and confirmed the proposed estimators are more efficiency and effectiveness than other existing estimators, due to the fact that they have the minimum mean square error (MSE) and the higher percentage relative error (PRE). Based on these qualifies possessed by the proposed estimators, therefore, we recommend its use in practical applications to estimate or calculate the population mean.

5. ACKNOWLEDGMENT

The authors hereby wish to acknowledge the reviewers and the Management Team for their service and recommendations.

6. CONFLICT OF INTEREST

There is no any atom of conflict of interest correlated with this work.

REFERENCES

[1] W.G. Cochran. "*Sampling Techniques*". 3rd Edition. Wiley Eastern Limited, India. 1940.

[2] C. Kadilar and H. Cingi. "Ratio Estimators in Simple Random Sampling". *Applied Mathematics and Computation* 151: 893 – 902. 2004.

[3] C. Kadilar and H. Cingi "An Improvement in Estimating the Population Mean By Using the Correlation Coefficient". *Hacettepe Journal of Mathematics and Statistics* 35: 103 – 109. 2006.

[4] H. P. Singh, and R. Tailor, "An Improved Estimation of Population Mean using Power Transformation". *Journal of the Indian Society of Agricultural Statistics*, 8 (2): 223–230. 2005.

[5] M. Subzar, S. Maqbool, T.A. Raja, and M. Abid, "Ratio Estimators for Estimating Population Mean in Simple Random Sampling Using Auxiliary Information". *Applied Mathematics and Information Sciences Letters*. 6(3), 126 – 130. 2018. DOI:10.18576/amisl/060304

[6] L.N. Upadhyaya, and H. P. Singh, "Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean". *Biometrical Journal.* 41:627 – 636. 1999.

AUTHORS

First Author – Jamiu Olasunkanmi Muili, PhD Student in Statistics, Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria and jamiunice@yahoo.com

Second Author – Ayodele Adebiyi, MS.c (Statistics), Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria and adebiyiayodele26@gmail.com

Third Author – **Eric Ndukaku Agwamba**, PhD Student in Statistics, Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria and agwambaeric@gmail.com

Correspondence Author – Jamiu Olasunkanmi Muili, jamiunice@yahoo.com, +2348062716248.