© 2016 IJSRSET | Volume 2 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Science and Technology

Modified Ratio Estimators for Population Mean Using Size of the Sample, Selected From Population

Mohmmad Jerajuddin[1]*, Jai Kishun[2]

[1]Consultant MIS (RCH-II), National Institute of Health and Family Welfare, New Delhi, India

[2]Assistant Professor (Statistics & Demography), National Institute of Health and Family Welfare, New Delhi, India

ABSTRACT

This paper deals with a modified ratio estimator for the estimation of population mean of study variable using the size of the sample, selected from the population under SRSWOR. The bias and mean square error of the proposed estimator up to the first order of approximation is derived [Appendix]. The constants, biases and mean square errors (MSEs) are computed using the data from Murthy [1] and Mukhopadhyay [3]. The percent relative efficiencies (PREs) are also computed for both existed and proposed estimators and compare the results accordingly for justifying the betterment of the proposed estimators over other mentioned modified estimators.

Keywords: Auxiliary Variable, Sample size, Simple Random Sampling, Bias, Mean Square Error, Relative Standard Error (RSE), Percent Relative Efficiency (PRE).

Notations & Terminology Used in this paper:

N – Population Size

n – Sample Size

$f=\frac{n}{N}$ Sampling Fraction

X – Auxiliary Variable

Y – Study Variable

$\bar{X},\bar{Y}$ – Population Mean

$\bar{x},\bar{y}$ – Sample Mean

C_x,C_y – Co-efficient of Variations of X and Y respectively

S_x,S_y – Population Standard Deviations of X and Y respectively

S_{xy} – Population Covariance between X and Y

M_d – Median of the auxiliary variable

ρ – Correlation Co-efficient between X & Y

$\beta_1=\frac{N\sum_{i=1}^{N}(X_i-\bar{X})^3}{(N-1)(N-2)S_x^2}=\frac{\mu_3^2}{\mu_2^3}$, Co-efficient of Skewness of the auxiliary variable

$\beta_2=\frac{N(N+1)\sum_{i=1}^{N}(X_i-\bar{X})^4}{(N-1)(N-2)(N-3)S_x^4}-\frac{3(N-1)^2}{(N-2)(N-3)}=\frac{\mu_4}{\mu_2^2}$, Co-efficient of Kurtosis of the auxiliary variable.

$\hat{\bar{Y}}_j$ – Existed j[th] modified ratio estimator of $\bar{Y}$

$\hat{\bar{Y}}_p$ – Proposed modified ratio estimator of $\bar{Y}$

Bias(□) – Bias of the estimator

MSE(□) – Mean square error of the estimator

PRE(□) – Percent relative efficiency of the estimator

I. INTRODUCTION

Cochran (1940) had first made his contribution to introduce the ratio estimator in literature using known information of the auxiliary variable in improving the efficiency of the estimator of the population mean $\bar{Y}$ [4]. Assuming that the population mean of the auxiliary variable X is known, and correlation between study and auxiliary variable is positive (high) [2], [3]; an estimator $\hat{\bar{Y}}_R$ of population mean $\bar{Y}$ was introduced. The ratio estimator is given below.

$$\hat{\bar{Y}}_R=\left(\frac{\bar{y}}{\bar{x}}\right)\bar{X} = \hat{R}\bar{X}\text{, where } \hat{R}=\frac{\bar{y}}{\bar{x}}=\frac{y}{x} \quad (1.1)$$

Where $\bar{y}$ is the sample mean of the study variable Y, and it is assumed that the population mean $\bar{X}$ of auxiliary variable X is known. The ratio estimator defined in (1.1) is the classical ratio estimate [18].

The bias in the ratio estimator $\hat{\bar{Y}}_R$ of the population mean $\bar{Y}$, to the first order of approximation:

$$Bias\left(\hat{\bar{Y}}_R\right) = \left(\frac{1-f}{n}\right)\bar{Y}\left(C_x^2 - \rho_{xy}C_xC_y\right) \quad (1.2)$$

The mean squared error of the ratio estimator $\hat{\bar{Y}}_R$ of the population mean $\bar{Y}$, to the first order of approximation:

$$MSE\left(\hat{\bar{Y}}_R\right) = \left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_x^2 + C_y^2 - 2\rho C_xC_y\right) \quad (1.3)$$

The ratio estimator $\hat{\bar{Y}}_R$ is more efficient than the sample mean $\bar{y}$ if

$$\rho > \frac{1}{2}\cdot\frac{C_x}{C_y} \quad (1.4)$$

The ratio estimator $\hat{\bar{Y}}_R$ has a bias of order n^{-1}. Since the standard error of the estimate is of order $n^{-\frac{1}{2}}$, the bias is also of order $n^{-\frac{1}{2}}$ [2]. In other words the bias of the estimator decreases as the sample size n increases which shows that the ratio estimate is a consistent estimate [15]. That is the ratio estimate can be almost unbiased for large n [16]. In this regard, the criteria of unbiasedness and consistency perform as primary filters in assessing possible estimates [17]. Since efficiency of an estimator depends on its variance and it is logical that the variance as well as the mean square error of estimator are minimum when increasing the sample size [19]. Initially Cochran (1940) introduced the ratio estimator, thereafter various modifications have been proposed for improving the efficiency of the ratio estimators by using some known descriptive statistics such as population mean, population variance, coefficient of variation, coefficient of kurtosis, coefficient of skewness, population correlation coefficient, function of quartiles, and use of median etc [5],[7],[12]. All these descriptive statistics have been used from the early information of the auxiliary variable. We have observed that no efforts have been made on the modification of ratio estimators for improving their efficiencies by utilizing the size of the sample, selected from the population under study. A sample usually represents a subset of manageable size so the chance of occurrence of sampling error is also possible when selecting the sample from the population. This sampling error is minimized with a required degree of precision (also called margin of error), say 'E' of estimate. If we consider the components of confidence interval when estimating sample size, its width is determined by a reliability coefficient, an standard error $\sigma/\sqrt{n}$ of estimate and the use of finite correction factor $\sqrt{\frac{N-n}{N-1}}$. Algebraically

$$\mathrm{E} = \left(reliability\ coefficient\right)\cdot\frac{\sigma}{\sqrt{n}}\cdot\sqrt{\frac{N-n}{N-1}} \quad (1.5)$$

The formula (1.5) is used when the sample size n is more than 5% of the finite population N. Since σ is a fixed quantity, the only way to have a small standard error is to take a large sample (assuming reliability coefficient is also fixed). But how large a sample depends on the size of σ, the population standard deviation and the preferred degree of reliability (ie. 99.9%, 99%, 95%, 90% etc.) [20], [21]. Under SRSWOR scheme, we also need an appropriate sample size with a minimum relative standard error for a smallest margin of error. Algebraically, the minimum sample size under SRSWOR is given by

$$n \geq \left[\frac{1}{N} - \frac{\alpha^2\bar{Y}^2}{S_y^2}\right]^{-1} \quad (1.6)$$

Where α denote the relative standard error (RSE) [4], which should be minimum as possible for a more precise sample. We collect those samples which contain some information about the population and calculate some specific statistics from it, so that it may be helpful for making inferences about the population. This process of collecting information from the sample is referred to as sampling. That's why sampling theory is connected to samples and its estimate is also based on the sizes of the sample.

II. METHODS AND MATERIAL

Consider a finite population of N units $U=\{U_1,U_2,\ldots,U_N\}$. Let (X_i,Y_i) ; $i=1,2,\ldots,N$ be real valued function defined on the finite population U, where X_i denote the auxiliary variables which have full information about population and Y_i the variable under study and the objective is to estimate the population mean $\bar{Y}$.

Several modified ratio estimators, proposed by statisticians are available in the literature. These estimators are also biased but have minimum mean squared errors (MSE) as compared to that of classical ratio estimator as well as the other modified ratio estimators. When the population coefficient of variation of auxiliary variable is known, Sisodia & Dwivedi [6] have proposed a modified ratio estimator for estimating $\bar{Y}$ together with its bias and mean squared error defined below.

$$\hat{\bar{Y}}_1=\bar{y}\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right) \tag{2.1}$$

$$Bias\left(\hat{\bar{Y}}_1\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_1^2C_x^2-\delta_1\rho C_xC_y\right) \tag{2.2}$$

$$MSE\left(\hat{\bar{Y}}_1\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_1^2C_x^2-2\delta_1\rho C_xC_y\right) \tag{2.3}$$

Where $\delta_1=\dfrac{\bar{X}}{\bar{X}+C_x}$

Motivated by Sisosidia & Dwivedi [6] and Singh et. al. [9] developed a modified ratio estimator replacing the coefficient of variation with coefficient of kurtosis, given below.

$$\hat{\bar{Y}}_2=\bar{y}\left(\frac{\bar{X}+\beta_2}{\bar{x}+\beta_2}\right) \tag{2.4}$$

$$Bias\left(\hat{\bar{Y}}_2\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_2^2C_x^2-2\delta_2\rho C_xC_y\right) \tag{2.5}$$

$$MSE\left(\hat{\bar{Y}}_2\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_2^2C_x^2-2\delta_2\rho C_xC_y\right) \tag{2.7}$$

Where $\delta_2=\dfrac{\bar{X}}{\bar{X}+\beta_2}$

In addition, various research works have been done in earlier time under simple random sampling on the modification of ratio estimators such as Upadhyaya & Singh [8], Singh & Tailor [11], Yan & Tian [10], and Subramani & Kumarpandiyan [5] etc have used the population parameters of the auxiliary variable for improving the efficiency of the ratio estimators by minimizing the mean sum of square of the estimators. Some modified ratio estimators are listed in Table 1.1 below. It should be noted that it is not the overall list of modified ratio estimators available in the literature. We have taken only some of them to compare our result with the existing modified ratio estimators.

Let us denote the existing modified ratio estimators as follows:

$\hat{\bar{Y}}_1$ = Ratio Estimator by Sisosidia & Dwivedi

$\hat{\bar{Y}}_2$ = Ratio Estimator by Singh et. al.

$\hat{\bar{Y}}_3$ = Ratio Estimator by Upadhyay & Singh

$\hat{\bar{Y}}_4$ = Ratio Estimator by Yan & Tian

$\hat{\bar{Y}}_5$ = Ratio Estimator by Singh & Tailor

$\hat{\bar{Y}}_6$ = Ratio Estimator by Subramani & Kumarpandiya

Table 1.1 A list of existing modified ratio estimators are summarized with their biases, mean square errors, and the constant terms δ_j (j=1,2,…6)

Sr. No.	***Estimators***	***Constants (δ_j)***	***Bias ($\hat{\bar{Y}}_j$)***	***MSE ($\hat{\bar{Y}}_j$)***
1.	$\hat{\bar{Y}}_1=\bar{y}\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right)$	$\delta_1=\frac{\bar{X}}{\bar{X}+C_x}$	$\text{Bias}\left(\hat{\bar{Y}}_1\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_1^2C_x^2-\delta_1\rho C_xC_y\right)$	$\text{MSE}\left(\hat{\bar{Y}}_1\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_1^2C_x^2-2\delta_1\rho C_xC_y\right)$
2.	$\hat{\bar{Y}}_2=\bar{y}\left(\frac{\bar{X}+\beta_2}{\bar{x}+\beta_2}\right)$	$\delta_2=\frac{\bar{X}}{\bar{X}+\beta_2}$	$\text{Bias}\left(\hat{\bar{Y}}_2\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_2^2C_x^2-\delta_2\rho C_xC_y\right)$	$\text{MSE}\left(\hat{\bar{Y}}_2\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_2^2C_x^2-2\delta_2\rho C_xC_y\right)$
3.	$\hat{\bar{Y}}_3=\bar{y}\left(\frac{\bar{X}C_x+\beta_2}{\bar{x}C_x+\beta_2}\right)$	$\delta_3=\frac{\bar{X}C_x}{\bar{X}C_x+\beta_2}$	$\text{Bias}\left(\hat{\bar{Y}}_3\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_3^2C_x^2-\delta_3\rho C_xC_y\right)$	$\text{MSE}\left(\hat{\bar{Y}}_3\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_3^2C_x^2-2\delta_3\rho C_xC_y\right)$
4.	$\hat{\bar{Y}}_4=\bar{y}\left(\frac{\bar{X}+\beta_1}{\bar{x}+\beta_1}\right)$	$\delta_4=\frac{\bar{X}}{\bar{X}+\beta_1}$	$\text{Bias}\left(\hat{\bar{Y}}_4\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_4^2C_x^2-\delta_4\rho C_xC_y\right)$	$\text{MSE}\left(\hat{\bar{Y}}_4\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_4^2C_x^2-2\delta_4\rho C_xC_y\right)$

5.	$\hat{\bar{Y}}_5=\bar{y}\left(\frac{\bar{X}+\rho}{\bar{x}+\rho}\right)$	$\delta_5=\frac{\bar{X}}{\bar{X}+\rho}$	$\text{Bias}\left(\hat{\bar{Y}}_5\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_5^2C_x^2-\delta_5\rho C_xC_y\right)$	$\text{MSE}\left(\hat{\bar{Y}}_5\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_5^2C_x^2-2\delta_5\rho C_xC_y\right)$
6.	$\hat{\bar{Y}}_6=\bar{y}\left(\frac{\bar{X}+M_d}{\bar{x}+M_d}\right)$	$\delta_6=\frac{\bar{X}}{\bar{X}+M_d}$	$\text{Bias}\left(\hat{\bar{Y}}_6\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_6^2C_x^2-\delta_6\rho C_xC_y\right)$	$\text{MSE}\left(\hat{\bar{Y}}_6\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_6^2C_x^2-2\delta_6\rho C_xC_y\right)$

For a better intimation to the readers in a single sight, the Biases and MSEs of the modified ratio estimators $\hat{\bar{Y}}_j$, (j=1, 2… 6), summarized in Table 1.1 can be represented as follows:

$$Bias\left(\hat{\bar{Y}}_j\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_j^2C_x^2-2\delta_j\rho C_xC_y\right) \quad (2.8)$$

$$MSE\left(\hat{\bar{Y}}_j\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_j^2C_x^2-2\delta_j\rho C_xC_y\right) \quad (2.9)$$

Where

$\delta_1=\frac{\bar{X}}{\bar{X}+C_x}$, $\delta_2=\frac{\bar{X}}{\bar{X}+\beta_2}$, $\delta_3=\frac{\bar{X}C_x}{\bar{X}C_x+\beta_2}$, $\delta_4=\frac{\bar{X}}{\bar{X}+\beta_1}$, $\delta_5=\frac{\bar{X}}{\bar{X}+\rho}$ and $\delta_6=\frac{\bar{X}}{\bar{X}+M_d}$.

III. PROPOSED MODIFIED RATIO ESTIMATOR

The fact is that the auxiliary variable X is closely related to (having positive correlation with) the study variable Y and it is assumed that the population total and/or mean of X is known, we have made effort to introduce a new kind of ratio estimator based on size of the sample, selected from the population under SRSWOR which is more efficient as compared to the other existed modified ratio estimators. Some of those have been summarized in Table 1.1 to make comparison of biases and mean square errors in our study. It has been observed that the Bias and MSE of the proposed estimator up to the first order of approximation is least as compared to those of already existed in the literature.

The proposed modified ratio estimator for population mean $\bar{Y}$ is

$$\hat{\bar{Y}}_p=\bar{y}\left(\frac{\bar{X}+n}{\bar{x}+n}\right) \quad (3.1)$$

To the first degree of approximation, the biases and mean squared errors of $\hat{\bar{Y}}_p$ are given as:

$$Bias\left(\hat{\bar{Y}}_p\right)=\left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_p^2C_x^2-2\delta_p\rho C_xC_y\right) \quad (3.2)$$

$$MSE\left(\hat{\bar{Y}}_p\right)=\left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2+\delta_p^2C_x^2-2\delta_p\rho C_xC_y\right) \quad (3.3)$$

Where $\delta_p=\frac{\bar{X}}{\bar{X}+n}$

And $f=\frac{n}{N}$ is the sampling fraction also called finite population correction factor (f.p.c)

From (2.9) and (3.3), it is observed that the proposed modified ratio estimator $\hat{\bar{Y}}_p$ perform better than the existing estimators $\hat{\bar{Y}}_j$ (j=1,2,…,6) and that's why $\hat{\bar{Y}}_p$ is more efficient than $\hat{\bar{Y}}_j$ (j=1,2,…,6). So that

$$MSE\left(\hat{\bar{Y}}_p\right)\le MSE\left(\hat{\bar{Y}}_j\right) \text{ For j=1,2,…,6} \quad (3.4)$$

$$\text{If } \rho\ge\frac{\delta_p+\delta_j}{2}\cdot\frac{C_x}{C_y} \text{ is satisfied} \quad (3.5)$$

Where $\delta_p=\frac{\bar{X}}{\bar{X}+n}$ and δ_j is defined in Table 1.1.

The percent relative efficiency (PRE) of the modified ratio estimators $\hat{\bar{Y}}_p$ and $\hat{\bar{Y}}_j$ with respect to the usual estimator $\bar{y}$ is given by

$$PRE(\square)=\frac{V(\bar{y}_{srs})}{MSE(\square)}\text{x}100 \quad (3.6)$$

IV. EMPIRICAL STUDY

We take some natural population data sets and compute the biases and the mean square errors of the existing modified ratio estimators listed in Table 1.1 and compare the results with that of proposed modified ratio estimator under first order of approximation.

Data Set: We have considered four natural populations. First two set of populations (Population 1 and Population 2) have been taken from Murthy [1] in page no. 228 and another two populations (Population 3 and Population 4) have been taken from Mukhopadhyay [3] in page no. 168. The descriptive statistics (Population Parameters), obtained from these populations are summarized in the Table 1.2.

Murthy [1] "Table 6.9":

Population 1: (X_1) = Data on number of workers and (Y) = Output for 80 factories in a region.
Population 2: (X_2) = Fixed Capital and (Y) = Output for 80 factories in a region.

Mukhopadhyay [3] "Table E6.1":

Population 3: (X_1) = Data on number of workers and (Y) = Output for 40 factories in a region.
Population 4: (X_2) = Fixed Capital and (Y) = Output for 40 factories in a region.

Table 1.2 Computed Parameters form the Populations P1, P2, P3, and P4.

Population Parameters	Murthy [1]		Mukhopadhyay [3]	
	P1	P2	P3	P4
N	80	80	40	40
n	20	20	8	8
f	0.25	0.25	0.2	0.2
$\bar{Y}$	51.8264	51.8264	50.7858	50.7858
$\bar{X}$	11.2646	2.8513	2.3033	9.4543
ρ	0.9413	0.9150	0.8006	0.8349
S_y	18.3566	18.3566	16.7352	16.7352
C_y	0.3542	0.3542	0.3295	0.3295
C_y^2	0.1255	0.1255	0.1086	0.1086
S_x	8.4561	2.7043	1.9360	6.3869
C_x	0.7507	0.9485	0.8406	0.6756
C_x^2	0.5635	0.8996	0.7065	0.4564
β_2	-0.0634	0.6977	-0.5344	-0.4622
β_1	1.0500	1.3006	0.9740	0.8799
M_d	7.5750	1.4800	1.250	7.0700

Tables [1.3 to 1.7] show the comparison between existing and proposed modified estimators. The graphical representation for MSEs and PREs of existing and modified ratio estimators are also presented herewith in section 4.1 and section 4.2.

V. RESULTS AND DISCUSSION

Table 1.3: Constants of the Existing and proposed modified ratio estimators for the Populations P1, P2, P3, and P4.

$\hat{\bar{Y}}_j$	P1	P2	P3	P4
$\hat{\bar{Y}}_1$	0.9375	0.7504	0.7326	0.9333
$\hat{\bar{Y}}_2$	1.0057	0.8034	1.3021	1.0514
$\hat{\bar{Y}}_3$	1.0076	0.7949	1.3813	1.0780
$\hat{\bar{Y}}_4$	0.9147	0.6868	0.7028	0.9149
$\hat{\bar{Y}}_5$	0.9229	0.7571	0.7421	0.9189
$\hat{\bar{Y}}_6$	0.5979	0.6583	0.6482	0.5721
$\hat{\bar{Y}}_p$	***0.3603***	***0.1248***	***0.2235***	***0.5417***

Table 1.4: Biases of the Existing and proposed modified ratio estimators for the Populations P1, P2, P3, and P4.

$\hat{\bar{Y}}_j$	P1	P2	P3	P4
$\hat{\bar{Y}}_1$	0.5066	0.5361	1.1009	1.1379
$\hat{\bar{Y}}_2$	0.6185	0.6484	4.6177	1.5697
$\hat{\bar{Y}}_3$	0.6218	0.6297	5.2908	1.6759
$\hat{\bar{Y}}_4$	0.4715	0.4142	0.9809	1.0763
$\hat{\bar{Y}}_5$	0.4839	0.5497	1.1402	1.0895
$\hat{\bar{Y}}_6$	0.1007	0.3643	0.7777	0.2186
$\hat{\bar{Y}}_p$	***0.0331***	***0.0473***	***0.0724***	***0.1687***

Table 1.5: MSE of the Existing and proposed modified ratio estimators for the Populations P1, P2, P3, and P4.

$\hat{\bar{Y}}_j$	P1	P2	P3	P4
$\hat{\bar{Y}}_1$	15.2581	0.5066	42.0165	41.0566
$\hat{\bar{Y}}_2$	19.3383	0.6185	188.0469	57.3218

$\hat{Y}_3$	19.4592	0.6218	217.7027	61.4403
$\hat{Y}_4$	14.0113	0.4715	37.6265	38.8113
$\hat{Y}_5$	14.4503	0.4839	43.4736	39.2910
$\hat{Y}_6$	2.7825	0.1007	30.4301	11.6837
$\hat{Y}_p$	***1.8389***	***0.0331***	***11.5428***	***10.6098***

1.1. Graphical Representation of the Mean Square Errors of the Existing and Proposed Modified Ratio Estimators:

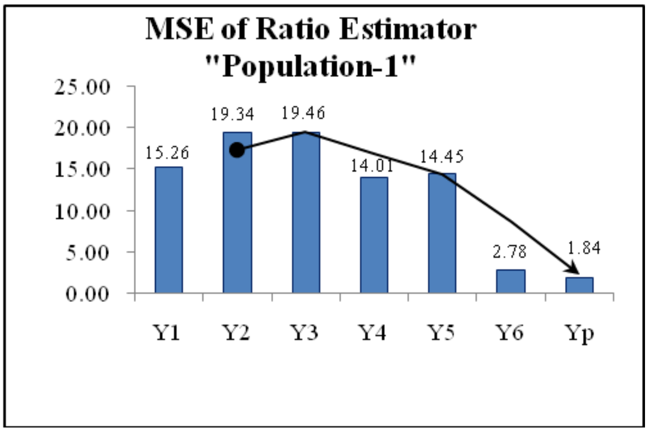

Figure 1: Mean Square Errors of Existing and Proposed Estimators for Population–1

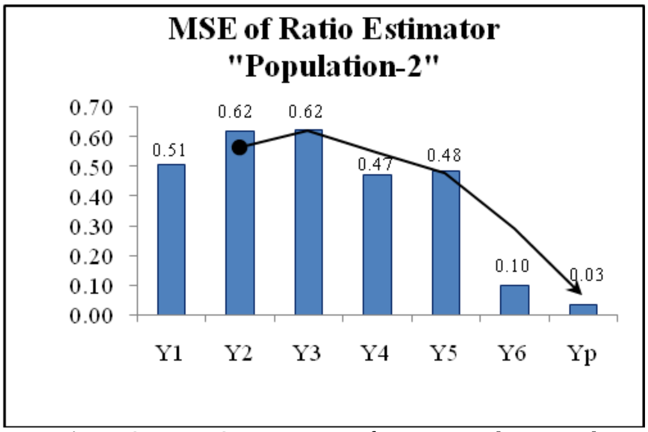

Figure 2: Mean Square Errors of Existing and Proposed Estimators for Population–2

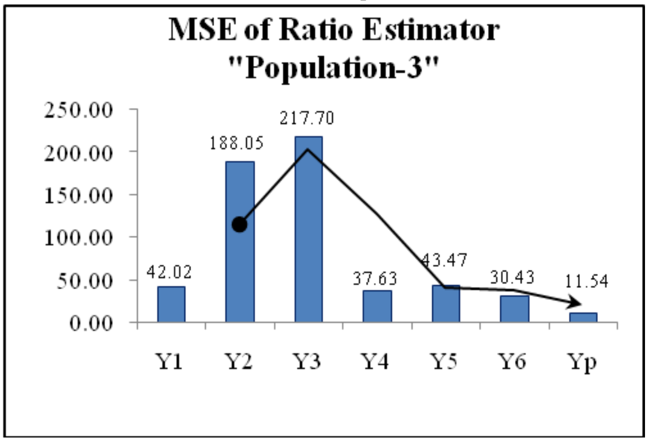

Figure 3: Mean Square Errors of Existing and Proposed Estimators for Population–3

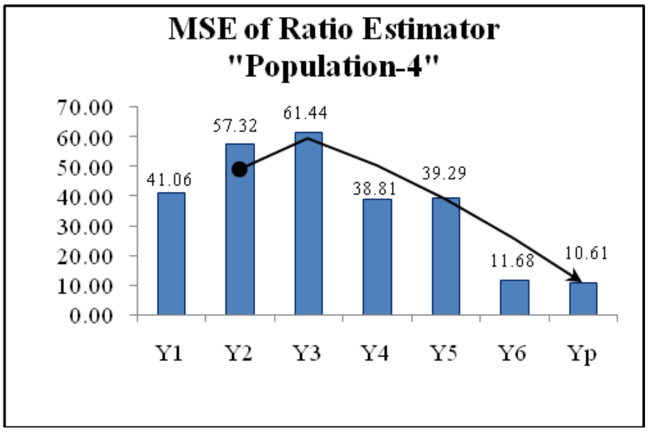

Figure 4: Mean Square Errors of Existing and Proposed Estimators for Population–4

Table 1.6: Comparison of Existing and Proposed Estimators with their Biases for the Populations P1, P2, P3, and P4.

$\hat{Y}_j$	P1	P2	P3	P4
	Bias	**Bias**	**Bias**	**Bias**
$\hat{Y}_1$	0.507	0.536	1.101	1.138
$\hat{Y}_2$	0.618	0.648	4.618	1.570
$\hat{Y}_3$	0.622	0.630	5.291	1.676
$\hat{Y}_4$	0.471	0.414	0.981	1.076
$\hat{Y}_5$	0.484	0.550	1.140	1.090
$\hat{Y}_6$	0.101	0.364	0.778	0.219
$\hat{Y}_p$	***0.033***	***0.047***	***0.072***	***0.169***

Table 1.7: Comparison of Existing and Proposed Estimators with their Mean Square Errors for the Populations P1, P2, P3, and P4.

$\hat{Y}_j$	P1	P2	P3	P4
	MSE	**MSE**	**MSE**	**MSE**
$\hat{Y}_1$	15.258	0.507	42.016	41.057
$\hat{Y}_2$	19.338	0.618	188.047	57.322
$\hat{Y}_3$	19.459	0.622	217.703	61.440
$\hat{Y}_4$	14.011	0.471	37.627	38.811
$\hat{Y}_5$	14.450	0.484	43.474	39.291
$\hat{Y}_6$	2.783	0.101	30.430	11.684
$\hat{Y}_p$	***1.839***	***0.033***	***11.543***	***10.610***

Table 1.8: Comparison of PREs for the estimators corresponding to all four populations

$\hat{Y}_j$	P1	P2	P3	P4
	PRE	PRE	PRE	PRE
$\hat{Y}_1$	82.8188	73.5193	66.6563	68.2147
$\hat{Y}_2$	65.3450	59.1433	14.8934	48.8586
$\hat{Y}_3$	64.9390	61.1606	12.8646	45.5835
$\hat{Y}_4$	90.1886	98.3961	74.4332	72.1610
$\hat{Y}_5$	87.4488	71.4541	64.4222	71.2800
$\hat{Y}_6$	454.138	113.469	92.0361	239.707
$\hat{Y}_p$	***687.179***	***199.9113***	***242.6326***	***263.9703***

1.2. Graphical Representation of the Percent Relative Efficiency of the Existing and Proposed Modified Ratio Estimators:

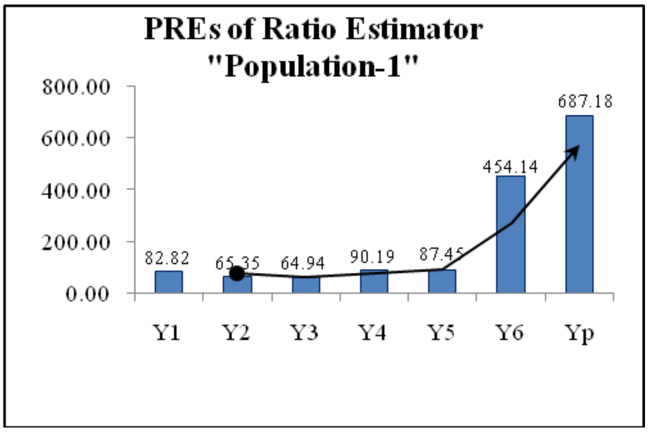

Figure 5: PRE of Existing and Proposed Estimators for Population–1

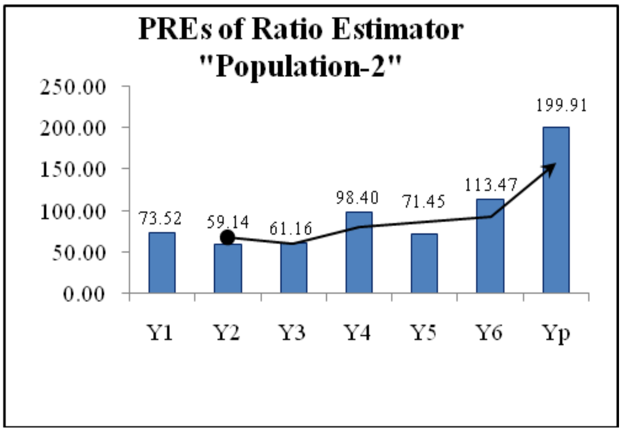

Figure 6: PRE of Existing and Proposed Estimators for Population–2

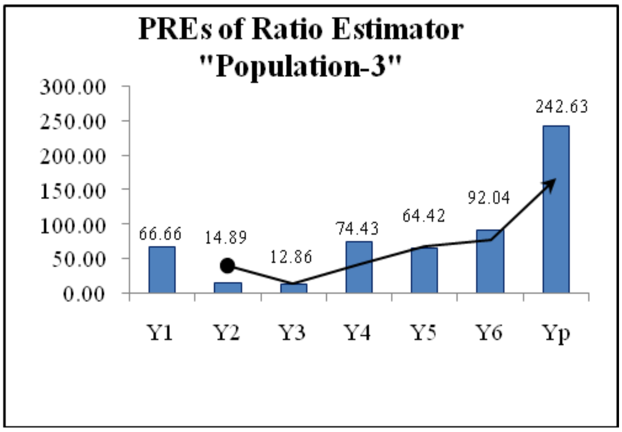

Figure 7: PRE of Existing and Proposed Estimators for Population–3

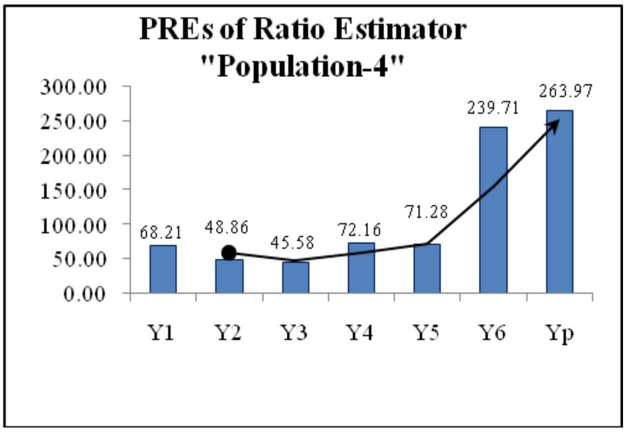

Figure 8: PRE of Existing and Proposed Estimators for Population–4

VI. RESULT and CONCLUSION

In the present paper, we have proposed a modified ratio estimator $\hat{Y}_p$ based on the size of the samples, selected from the population under SRSWOR for estimating the population mean of study variable Y when information of auxiliary variable X is available. We have found in our study that the proposed estimator is more efficient than the existing modified ratio estimators.

From the empirical study, we have observed that the bias and mean square error of the proposed modified ratio estimator is least as compared to those of existing modified estimators. Table 1.4 and Table 1.5 show the results of Biases and MSEs of both of the estimators (Proposed and Existing), from which we see that proposed estimator, performs better than those of existing estimators. The percent relative efficiency (PRE) in Table 1.7 also support the study as the proposed estimator is highly efficient than the existing estimators. The graphical representations (section 1.1 and 1.2) of MSEs and PREs legitimate the justification of the proposed modified ratio estimator over the existing ratio estimators. Thus if the auxiliary variable X is closely related to the study variable Y assuming that population total and/or mean of X is known, a modified ratio estimator based on the size of the samples, selected from the population under SRSWOR scheme, is recommended to estimate the population mean of the study variable.

VII. REFERENCES

[1] Murthy, M.N. (1967): Sampling Theory and Methods; Statistical Publishing Society, Calcutta.

[2] Cochran, W.G. (1977): Sampling Techniques; 3rd edition. Wiley Eastern, New Delhi.

[3] Mukhopadhyay, P. (2009): Theory and Methods of Survey Sampling; 2nd edition. PHI Learning Pvt. Ltd., New Delhi.

[4] Sarjinder, S. (2003): Advanced Sampling Theory with Applications; Kluwer Academic Publishers Volume. (1&2)

[5] Subramani, J. & Kumarpandiyan, G. (2013): New Modified Ratio Estimation of Population Mean when Median of Auxiliary Variable is Known; Pak.j.stat.oper.res. Vol.IX No.2 2013 pp137 – 145.

[6] Sisodia, B.V.S. and Dwivedi, V.K. (1981): A Modified Ratio Estimator using Coefficient of Variation of Auxiliary Variable; Jour. of Indian. Soc. of Agri. Stat., Vol. 33(1). Pp. 13–18.

[7] Khan, M. & Shabbir, J. (2013): A Ratio Type Estimator for the Estimation of Population Variance using Quartiles of an Auxiliary Variable; J.Stat. Appl. Pro.2, No. 3, 319 –325.

[8] Upadhyaya, L.N. & Singh, H.P. (1999): Use of transformed auxiliary variable in estimating the finite population means, Biometrical Journal 41 (5), 627–636.

[9] Singh, H.P. & Kakran, M.S. (1993): A Modified Ratio Estimator Using Known Coefficient of Kurtosis of an Auxiliary Character, (unpublished).

[10] Yan, Z. & Tian, B. (2010): Ratio Method to the Mean Estimation using Coefficient of Skewness of Auxiliary Variable; ICICA, Part-II, CCIS 106, pp. 103 – 110.

[11] Singh, H.P. & Tailor, R. (2003). Use of known Correlation Coefficient in Estimating the finite Population Means; Statistics in Transition 6 (4), 555-560.

[12] Yadav, S.K., Mishra, S.S. and Shukla, A.K. (2014): 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.

[13] Subramani, J., Kumarpandiyan, G. & Balamurli, S. (2014): Some Modified Linear Regression type Ratio Estimators for Estimation of Population Mean Using Known Parameters of an Auxiliary Variable; Journal of Building Research, 1(1): 28-42.

[14] Sangngam, P. (2014): Ratio Estimators Using Coefficient of Variation and Coefficient of Correlation Modern Applied Science; Vol. 8, No. 5.

[15] Sukhatme, P.V. (1957): Sampling Theory of Surves with Applications, The Indian Society of Agricultural Statistics, New Delhi, India.

[16] Gun, A.M., Gupta, M.K. & Dasgupta, B. (2008): Fundamentals of Statistics, Vol–2; The World Press Pvt. Ltd., Kolkata.

[17] Barnett, V. (1999): Comparative Statistical Inference; 3rd edition, JohnWiley & Sons Ltd.

[18] Kadilar, C. & Cingi, H. (2004): Ratio Estimators in Simple Random Sampling; Applied Mathematics and Computation 151, 893–902

[19] Allen L. Webster (2010): Applied Statistics for Business and Economics, An Essentials Version; 3rd edition, Tata McGraw–Hill Education Pvt. Ltd.

[20] David A. Kay (2005): CliffsAP Statistics, Wiley Publishing, Inc., Hoboken, NJ

[21] Daniel, W.W. & Cross, C.L. (2013): Biostatistics, A Foundation for Analysis in Health Sciences, 10th edition; JohnWiley & Sons, Inc.

APPENDIX

Let us define –

$e_0 = \frac{\bar{y}}{\bar{Y}} - 1$, $e_1 = \frac{\bar{x}}{\bar{X}} - 1$, $e_2 = \frac{s_y^2}{S_y^2} - 1$, $e_3 = \frac{s_x^2}{S_x^2} - 1$ and

$$e_4 = \frac{s_{xy}}{S_{xy}} - 1 \qquad \text{(A.1)}$$

With conditions

$E(e_j) = 0$ For all j=1, 2, 3, 4 (A.2)

And

$$E(e_0^2) = \frac{1-f}{n} C_y^2 \quad , \quad E(e_1^2) = \frac{1-f}{n} C_x^2 \quad ,$$

$$E(e_0 e_1) = \frac{1-f}{n} \rho C_x C_y \qquad \text{(A.3)}$$

1. We have proposed the modified ratio estimator in (3.1) as

$$\hat{\bar{\mathrm{Y}}}_\mathrm{p} = \bar{\mathrm{y}}\left(\frac{\bar{\mathrm{X}}+\mathrm{n}}{\bar{\mathrm{x}}+\mathrm{n}}\right)$$, Where $\delta_\mathrm{p} = \frac{\bar{\mathrm{X}}}{\bar{\mathrm{X}}+n}$ and

$f = \frac{\mathrm{n}}{\mathrm{N}}$ is the sampling fraction (3.1)

From (A.1), we obtain the relations $\bar{y} = \bar{Y}(1+e_0)$ and $\bar{x} = \bar{X}(1+e_1)$. Now substituting these values of $\bar{y}$ and $\bar{x}$ in the proposed modified ratio estimator (3.1), we have

$$\hat{\bar{\mathrm{Y}}}_\mathrm{p} = \bar{Y}(1+e_0)\left(\frac{\bar{\mathrm{X}}+\mathrm{n}}{\bar{X}(1+e_1)+\mathrm{n}}\right) = \bar{Y}(1+e_0)\left(1+\frac{\bar{\mathrm{X}}}{\bar{\mathrm{X}}+n}e_1\right)^{-1}$$

$\hat{\bar{\mathrm{Y}}}_\mathrm{p} = \bar{Y}(1+e_0)\left(1+\delta_\mathrm{p} e_1\right)^{-1}$, Where $\delta_\mathrm{p} = \frac{\bar{\mathrm{X}}}{\bar{\mathrm{X}}+n}$ (3.1.1)

Assuming $|e_1| < 1$ and using the binomial expansion of the term $(1+\delta_\mathrm{p} e_1)^{-1}$ of (3.1.1), we have

$$\hat{\bar{\mathrm{Y}}}_\mathrm{p} = \bar{Y}(1+e_0)\left(1-\delta_\mathrm{p} e_1 + \delta_p^2 e_1^2 + O(\mathrm{e}_1)\right) \qquad (3.1.2)$$

$$\hat{\bar{\mathrm{Y}}}_\mathrm{p} = \bar{Y}\left(1+e_0 - \delta_\mathrm{p} e_1 - \delta_\mathrm{p} e_0 e_1 + \delta_p^2 e_1^2 + O(\mathrm{e}_1)\right) \qquad (3.1.3)$$

Where $O(\mathrm{e}_1)$ denote the higher order terms of e_1. We also have assumed that $|e_1| < 1$, so that $\mathrm{e}_1^k \to 0$ as k increases or mathematically if k>1.

Thus the terms in (3.1.3) having higher powers of e_1 are negligible and therefore considered to be vanished.

Now taking expectations on both sides of (3.1.3) and using the results from (A.2) and (A.3), we have

$$E\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) = \bar{Y}\left(1+E(e_0) - \delta_\mathrm{p} E(e_1) - \delta_\mathrm{p} E(e_0 e_1) + \delta_p^2 E(e_1^2) + E\{O(\mathrm{e}_1)\}\right)$$

$$\Rightarrow E\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) = \bar{Y}\left(1-\delta_\mathrm{p}\frac{1-f}{n}\rho C_x C_y + \delta_p^2 \frac{1-f}{n} C_x^2 + O\left(\mathrm{n}^{-1}\right)\right)$$

$$\Rightarrow E\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) = \bar{Y}\left(1+\left(\frac{1-f}{n}\right)\left(\delta_p^2 C_x^2 - \delta_\mathrm{p}\rho C_x C_y\right) + O\left(\mathrm{n}^{-1}\right)\right)$$

$$\Rightarrow E\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) - \bar{Y} = \left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_p^2 C_x^2 - \delta_\mathrm{p}\rho C_x C_y\right) \qquad (3.1.4)$$

$$\Rightarrow Bias\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) = \left(\frac{1-f}{n}\right)\bar{Y}\left(\delta_\mathrm{p}^2 C_x^2 - \delta_\mathrm{p}\rho C_x C_y\right) \qquad (3.2)$$

Thus the bias in the proposed modified ratio estimator $\hat{\bar{\mathrm{Y}}}_\mathrm{p}$ to the first order of approximation is given by equation (3.2).

2. Consider equation (3.1.3) and on simplification we get

$$\hat{\bar{\mathrm{Y}}}_\mathrm{p} - \bar{Y} = \bar{Y}\left(e_0 - \delta_\mathrm{p} e_1 - \delta_\mathrm{p} e_0 e_1 + \delta_p^2 e_1^2 + O(\mathrm{e}_1)\right) \qquad (3.1.5)$$

Squaring both sides and taking expectations we have

$$E\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p} - \bar{Y}\right)^2 = \bar{Y}^2 E\left(e_0 - \delta_\mathrm{p} e_1 - \delta_\mathrm{p} e_0 e_1 + \delta_p^2 e_1^2 + O(\mathrm{e}_1)\right)^2 \qquad (3.1.6)$$

Again in the similar way, the higher order terms will be negligible. Now using the results from (A.2) and (A.3), the MSE to the first order of approximation is given by

$$MSE\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) = \bar{Y}^2 E\left(e_0^2 + \delta_p^2 e_1^2 - 2\delta_\mathrm{p} e_0 e_1\right) \qquad (3.1.7)$$

Thus we get the required result as given below

$$MSE\left(\hat{\bar{\mathrm{Y}}}_\mathrm{p}\right) = \left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2 + \delta_\mathrm{p}^2 C_x^2 - 2\delta_\mathrm{P}\rho C_x C_y\right) \qquad (3.3)$$

Hence the mean square error in the proposed ratio estimator $\hat{\bar{\mathrm{Y}}}_\mathrm{p}$ to the first order of approximation is given by (3.3)

3. The conditions for which the proposed estimator performs better than the sample mean under SRSWOR :

We know that in simple random sampling, the most suitable estimator of population mean $\bar{\mathrm{Y}}$ is the sample mean $\bar{\mathrm{y}}$. We also know that the variance of $\bar{\mathrm{y}}$, up to the first order of approximation is

$$V(\bar{y}_{srs}) = \left(\frac{1-f}{n}\right)\bar{Y}^2 C_y^2 \tag{3.1.8}$$

The modified ratio estimator $\hat{\bar{Y}}_p$ is more efficient than the sample mean $\bar{y}$ if the mean square error of the estimator $\hat{\bar{Y}}_p$ is less than the variance of sample mean $\bar{y}$. Algebraically

$$MSE\left(\hat{\bar{Y}}_p\right) \leq V(\bar{y}_{srs}) \tag{3.1.9}$$

$$\Rightarrow \left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2 + \delta_p^2 C_x^2 - 2\delta_P \rho C_x C_y\right) \leq \left(\frac{1-f}{n}\right)\bar{Y}^2 C_y^2$$

$$\Rightarrow \delta_p^2 C_x^2 - 2\delta_P \rho C_x C_y \leq 0$$

$$\Rightarrow \rho \geq \frac{\delta_P}{2} \cdot \frac{C_x}{C_y} \qquad \text{Where } \delta_p = \frac{\bar{X}}{\bar{X}+n} \tag{3.2.1}$$

Hence if the condition (3.2.1) is satisfied, the proposed modified ratio estimator is more efficient than the sample mean $\bar{y}$.

4. The condition for which the proposed estimator $\hat{\bar{Y}}_p$ performs better than the existing estimators $\hat{\bar{Y}}_j$:

The modified ratio estimator $\hat{\bar{Y}}_p$ is more efficient than the existing modified ratio estimators $\hat{\bar{Y}}_j$ if the MSE of $\hat{\bar{Y}}_p$ is less than MSE of $\hat{\bar{Y}}_j$. Algebraically

$$MSE\left(\hat{\bar{Y}}_p\right) \leq MSE\left(\hat{\bar{Y}}_j\right) \tag{3.4}$$

$$\Rightarrow \left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2 + \delta_p^2 C_x^2 - 2\delta_p \rho C_x C_y\right) \leq \left(\frac{1-f}{n}\right)\bar{Y}^2\left(C_y^2 + \delta_j^2 C_x^2 - 2\delta_j \rho C_x C_y\right)$$

$$\Rightarrow \delta_p^2 C_x^2 - 2\delta_P \rho C_x C_y \leq \delta_j^2 C_x^2 - 2\delta_j \rho C_x C_y$$

$$\Rightarrow \rho \geq \frac{\delta_P + \delta_j}{2} \cdot \frac{C_x}{C_y} \tag{3.5}$$

Where δ_j is defined in Table 1.1 and $\delta_p = \frac{\bar{X}}{\bar{X}+n}$

*****