Asian Journal of Probability and Statistics

7(4): 48-58, 2020; Article no.AJPAS.58543
ISSN: 2582-0230

Some Median Type Estimators to Estimate the Finite Population Mean

Waqar Hafeez[1*], Javid Shabbir[2], Muhammad Taqi Shah[2] and Shakeel Ahmed[2]

[1]School of Quantitative Sciences, Universiti Utara Malaysia, Malaysia.
[2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.

Authors' contributions

This work was carried out in collaboration among all authors. Author WH designed the study, performed the statistical analysis, wrote the protocol, wrote the first draft of the manuscript and managed the analyses of study under the supervision of author JS. Authors MTS and SA managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2020/v7i430192
Editor(s):
(1) Dr. Manuel Alberto M. Ferreira, Retired Professor, Lisbon University, Portugal.
Reviewers:
(1) Saad B. H. Farid, University of Technology, Iraq.
(2) Zeena N. Al-Kateeb, University of Mosul, Iraq.
(3) S. Shankar, Hindusthan College of Engineering and Technology, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/58543

Original Research Article

Received: 24 April 2020
Accepted: 02 July 2020
Published: 11 July 2020

Abstract

Researchers always appreciates estimators of finite population quantities, especially mean, with maximum efficiency for reaching to valid statistical inference. Apart from ratio, product and regression estimators, exponential estimators are widely considered by survey statisticians. Motivated from the idea of exponential type estimators, in this article, we propose some new estimators utilizing known median of the study variable with mean of auxiliary variable. Theoretical properties of the suggested estimators are studied up to first order of approximation. In addition, an empirical and simulation study the comparison of median based proposed class of estimators with sample mean, ratio and linear regression estimators are discussed. The results expose that the proposed estimators are more efficient than the existing estimators.

Keywords: Bias; efficiency; mean; median; exponential estimators.

**Corresponding author: E-mail: waqarhafeez78601@yahoo.com;*

1 Introduction

In case of homogenous population, the most suitable sampling technique to select a sample is simple random sampling (SRS). In this technique each unit has equal chance or equal probability of being selected as sample unit. In practice of selecting a sample by SRS, units are drawn one by one from a population. There are two methods for selecting a sample by SRS from a population; (i) with replacement and (ii) without replacement. If a selected unit is replaced back to the population before drawing the next unit, this method is called with replacement sampling. If the selected unit is not replaced back before drawing the next unit, it is called without replacement sampling. Various methods are available to select a sample from population in SRS, but random number table and lottery methods are two most commonly used in practice. For the estimation of finite population, Hafeez and Shabbir [1] introduced median based estimators in stratified sampling. Milton [2] work in simple random sampling for population variance and proposed ratio type estimators. Gruber [3] work on regression type estimators and use it in different techniques in his book. Onsongo, Otieno and Orwa [4] work on bias reduction technique for estimating finite population distribution function under simple random sampling by using without replacement sampling.

Consider a large population and draw a large sample of size *N* randomly, now consider it a complete population $U = \{\mu_1, \mu_2, \dots, \mu_N\}$ of size N. Let y_i and x_i be characteristics of the study variable y and the auxiliary variable x respectively Shakeel et al. [5]. We are interested to estimate population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ based on median information. In absence of the auxiliary variable we obtain median of the samples of the study variable y in our proposed estimators and subsequently precision improved. We draw all possible samples of size n from population U by using simple random sampling without replacement (SRSWOR) scheme.

Let $\bar{y}$ and $\bar{x}$ be the sample means; m be the sample median of the study variable; $\bar{M}$ be the average of sample medians; β be the population regression coefficient of y on x; and ρ_{yx} be the population correlation coefficient between y and x.

Also define: $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$, $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$, $V(m) = \frac{1}{\binom{N}{n}}\sum_{i=1}^{\binom{N}{n}}(m_i - \bar{M})^2$, $Cov(\bar{y}, \bar{x}) = \frac{1-f}{n(N-1)}\sum_{i=1}^{\binom{N}{n}}(\bar{y}_i - \bar{Y})(\bar{x}_i - \bar{X})$, $Cov(\bar{y}, m) = \frac{1}{\binom{N}{n}}\sum_{i=1}^{\binom{N}{n}}(\bar{y}_i - \bar{Y})(m_i - \bar{M})$, $f = \frac{n}{N}$ and $\lambda = \frac{1-f}{n}$.

In absence of the auxiliary variable under SRSWOR, the variance of $\bar{y}$, is given by

$$V(\bar{y}) = \frac{1-f}{n} S_y^2 \tag{1}$$

The usual ratio estimator is widely used when correlation between the study and the auxiliary variables is positive. it is given by:

$$\hat{\bar{Y}}_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \tag{2}$$

where $\bar{X}$ is the known population mean of x.

The bias and mean squared error, to first order approximation, are given by:

$$B\left(\hat{\bar{Y}}_R\right) \cong \lambda \bar{Y}\left(C_x^2 - C_{yx}\right) \tag{3}$$

and

$$MSE\left(\hat{\bar{Y}}_R\right) \cong \lambda \bar{Y}^2\left(C_y^2 + C_x^2 - 2C_{yx}\right) \tag{4}$$

where $C_y^2 = \frac{V(\bar{y})}{\bar{Y}^2}$, $C_x^2 = \frac{V(\bar{x})}{\bar{X}^2}$ and $C_{yx} = \frac{Cov(\bar{y},\bar{x})}{\bar{Y}\bar{X}}$.

Watson [6], suggested the usual regression estimator, is given

$$\bar{Y}_{lr} = \bar{y} + b(\bar{X} - \bar{x}) \tag{5}$$

where $b = \frac{s_{yx}}{s_x^2}$ is the least square estimate of $\beta = \frac{S_{yx}}{S_x^2}$.

The variance of $\hat{\bar{Y}}_{lr}$ is given

$$V\left(\hat{\bar{Y}}_{lr}\right) = V(\bar{y})\left(1 - \rho_{yx}^2\right) \tag{6}$$

Bahl and Tuteja [7], introduced the following exponential type ratio estimator.

$$\hat{\bar{Y}}_{BT} = \bar{y}exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{7}$$

The bias and MSE of $\hat{\bar{Y}}_{BT}$ to first order approximation are given below

$$B\left(\hat{\bar{Y}}_{BT}\right) \cong \lambda \bar{Y}\left(\frac{3C_x^2}{8} - \frac{C_{yx}}{2}\right) \tag{8}$$

and

$$MSE\left(\hat{\bar{Y}}_{BT}\right) \cong \lambda \bar{Y}^2\left(C_y^2 + \frac{C_x^2}{4} - C_{yx}\right) \tag{9}$$

Rao [8], suggested the following estimator:

$$\hat{\bar{Y}}_{Rao} = k_1\bar{y} + k_2(\bar{X} - \bar{x}) \tag{10}$$

where k_1 and k_2 are constants.

The bias and MSE of $\hat{\bar{Y}}_{Rao}$ are given below

$$B\left(\hat{\bar{Y}}_{Rao}\right) = (k_1 - 1)\bar{Y} \tag{11}$$

And

$$MSE\left(\hat{\bar{Y}}_{Rao}\right) = \frac{V(\bar{y})\left(1 - \rho_{yx}^2\right)}{1 + C_y^2\left(1 - \rho_{yx}^2\right)} \tag{12}$$

Grover and Kaur [9], suggested the following estimator

$$\hat{\bar{Y}}_{GK} = [k_1\bar{y} + k_2(\bar{X} - \bar{x})]exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{13}$$

where k_1 and k_2 are constants. The bias and MSE of $\hat{\hat{Y}}_{GK}$ are given below

$$B\left(\hat{\hat{Y}}_{GK}\right) = \bar{Y}\left[(k_1 - 1) + \lambda k_1 \frac{C_x}{2}\left(\frac{3}{4}C_x - \rho C_y\right)\right] + \lambda k_2 \bar{X}\frac{C_x^2}{2} \tag{14}$$

and

$$MSE\left(\hat{\hat{Y}}_{GK}\right) = \frac{\lambda \bar{Y}^2 C_y^2(1-\rho_{yx}^2)}{1+\lambda C_y^2(1-\rho_{yx}^2)} - \frac{\lambda^2 \bar{Y}^2 C_x^2\left(4C_y^2(1-\rho_{yx}^2) + \frac{C_x^2}{4}\right)}{16\left[1+\lambda C_y^2(1-\rho_{yx}^2)\right]} \tag{15}$$

2 Proposed Estimator

Most of the estimators based on sample means that may cause inefficient in some situations. Here we propose median based estimators that perform better than that estimators used sample means.

Bahl and Tuteja [7], estimator in terms of median of *y* is given by:

$$\hat{\hat{Y}}_{BTm} = \bar{y}exp\left(\frac{\bar{M} - m}{\bar{M} + m}\right) \tag{16}$$

To obtain the bias and MSE of $\hat{\hat{Y}}_{BTm}$, we define the following error terms.

Let $e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $e_1 = \frac{m-\bar{M}}{\bar{M}}$. Such that; $E(e_0) = E(e_1) = 0$, $E(e_0^2) = \frac{V(\bar{y})}{\bar{Y}^2}$, $E(e_1^2) = \frac{V(m)}{\bar{M}^2}$ and $E(e_0 e_1) = \frac{Cov(\bar{y},m)}{\bar{Y}\bar{M}}$.

To first order of approximation, (16) can be written as:

$$\hat{\hat{Y}}_{BTm} - \bar{Y} \cong \bar{Y}\left[e_0 - \frac{1}{2}e_1 + \frac{3}{8}e_1^2 - \frac{1}{2}e_0 e_1\right] \tag{17}$$

Using (17), the bias of $\hat{\hat{Y}}_{BTm}$, is given by

$$B\left(\hat{\hat{Y}}_{BTm}\right) \cong \bar{Y}\left(\frac{3C_m^2}{8} - \frac{C_{\bar{y}m}}{2}\right) \tag{18}$$

Squaring both sides of (17) and then taking expectation, we get MSE of $\hat{\hat{Y}}_{BTm}$, to first order approximation, as

$$MSE\left(\hat{\hat{Y}}_{BTm}\right) \cong \bar{Y}^2\left(C_{\bar{y}}^2 + \frac{C_m^2}{4} - C_{\bar{y}m}\right) \tag{19}$$

where $C_{\bar{y}}^2 = \frac{V(\bar{y})}{\bar{Y}^2}$, $C_m^2 = \frac{V(m)}{\bar{M}^2}$ and $C_{\bar{y}m} = \frac{Cov(\bar{y},m)}{\bar{Y}\bar{M}}$.

Another unbiased difference type median based estimator to estimate population mean $\bar{Y}$ by using medians of samples that are taken by using SRSWOR from the study variable is given by:

$$\hat{\hat{Y}}_{dm} = \bar{y} + k(\bar{M} - m) \tag{20}$$

where k is the constant.

The minimum MSE of $\hat{\bar{Y}}_{dm}$ at $k_{(opt)} = \frac{Cov(\bar{y},m)}{V(m)}$ is given by,

$$MSE\left(\hat{\bar{Y}}_{dm}\right)_{min} = V(\bar{y})\left(1-\rho_{ym}^2\right) \tag{21}$$

Rao [8] estimator in terms of median is given by

$$\hat{\bar{Y}}_{Raom} = k_1\bar{y} + k_2(\bar{M}-m) \tag{22}$$

where k_1 and k_2 are constants.

The bias and MSE of $\hat{\bar{Y}}_{Raom}$ to first order of approximation are given by

$$B\left(\hat{\bar{Y}}_{Raom}\right) = (k_1-1)\bar{Y} \tag{23}$$

and

$$MSE\left(\hat{\bar{Y}}_{Raom}\right) \cong (k_1-1)^2\bar{Y}^2 + k_1^2\,V(\bar{y}) + k_2^2\,V(m) - 2k_1k_2Cov(\bar{y},m) \tag{24}$$

The minimum MSE of $\hat{\bar{Y}}_{Raom}$ at optimum values of k_1 and k_2 i.e.

$$k_{1(opt)} = \frac{\bar{Y}^2}{\bar{Y}^2 + V(\bar{y}) - \frac{\left(Cov(\bar{y},m)\right)^2}{V(m)}} \quad \text{and} \quad k_{1(opt)} = k_{1(opt)}\frac{Cov(\bar{y},m)}{V(m)}.$$

or,

$$k_{1(opt)} = \frac{1}{1+C_{\bar{y}}^2\left(1-\rho_{ym}^2\right)} \quad \text{and} \quad k_{2(opt)} = \frac{V(\bar{y})\rho_{ym}^2}{1+C_{\bar{y}}^2\left(1-\rho_{ym}^2\right)}$$

is given by:

$$MSE\left(\hat{\bar{Y}}_{Raom}\right)_{min} \cong \frac{\bar{Y}^2C_{\bar{y}}^2\left(1-\rho_{ym}^2\right)}{1+C_{\bar{y}}^2\left(1-\rho_{ym}^2\right)} \tag{25}$$

On the lines of Grover and Kaur [9], we propose a new median based estimator for $\bar{Y}$ by using average of median of samples taken from the study variable under SRSWOR sampling scheme. The new estimator, is given

$$\hat{\bar{Y}}_{GKm} = [k_1\bar{y} + k_2(\bar{M}-m)]exp\left(\frac{\bar{M}-m}{\bar{M}+m}\right) \tag{26}$$

where k_1 and k_2 are constants.

Solving (26), by expanding right hand side to first order of approximation, we have

$$\hat{\bar{Y}}_{GKm} - \bar{Y} \cong (k_1-1)\bar{Y} + k_1\bar{Y}\left\{e_0 - \frac{1}{2}e_1 + \frac{3}{8}e_1^2 - \frac{1}{2}e_0e_1\right\} - k_2\bar{M}\left\{e_1 + \frac{1}{2}e_1^2\right\} \tag{27}$$

The bias of $\hat{\hat{Y}}_{GKm}$ is given as

$$B\left(\hat{\hat{Y}}_{GKm}\right) \cong \bar{Y}\left[(k_1-1)+\frac{3}{8}k_1C_m^2-\frac{1}{2}k_1C_{\bar{y},m}\right]+\frac{1}{2}k_2\bar{M}C_m^2 \tag{28}$$

Squaring on both sides of eq (27), we get MSE of the estimator $\hat{\hat{Y}}_{GKm}$, as

$$MSE\left(\hat{\hat{Y}}_{GKm}\right) \cong \bar{Y}^2\left[(k_1-1)^2+k_1^2C_{\bar{y}}^2\right]+\left[(k_1R_2+k_2)^2-\left(\frac{3}{4}k_1R_2+k_2\right)R_2\right]\bar{M}^2C_m^2+k_1(R_2-2k_1R_2-2k_2)\bar{Y}\bar{M}C_{\bar{y}m} \tag{29}$$

The optimum values of k_1 and k_2 are,

$$k_{1(opt)}=\frac{1-\frac{1}{8}C_m^2}{1+C_{\bar{y}}^2(1-\rho_{\bar{y}m}^2)} \quad \text{and} \quad k_{2(opt)}=\left[k_{1(opt)}\left(\frac{C_{\bar{y},m}}{C_m^2}-1\right)+\frac{1}{2}\right]R_2$$

with minimum *MSE* of $\hat{\hat{Y}}_{GKm}$ is given by

$$MSE\left(\hat{\hat{Y}}_{GKm}\right) \cong \frac{\bar{Y}^2C_{\bar{y}}^2(1-\rho_{\bar{y}m}^2)}{1+C_{\bar{y}}^2(1-\rho_{\bar{y}m}^2)}-\frac{\bar{Y}^2C_m^2\left(4C_{\bar{y}}^2(1-\rho_{\bar{y}m}^2)+\frac{C_m^2}{4}\right)}{16\left[1+C_{\bar{y}}^2(1-\rho_{\bar{y}m}^2)\right]} \tag{30}$$

3 Efficiency Conditions

In this section the proposed median based estimator is compared in terms of MSE with all estimators.

Condition 1: Using equation (1) and (30)

$$MSE(\hat{\hat{Y}}_{GKm})_{min} \le V(\bar{y})$$

If

$$\bar{Y}^2C_{\bar{y}}^2\rho_{\bar{y}m}^2+A+B \ge 0$$

Condition 2: Using equation (4) and (30)

$$MSE(\hat{\hat{Y}}_{GKm})_{min} \le MSE(\hat{\bar{Y}}_R)$$

If

$$\bar{Y}^2\lambda\left[C_x^2-2C_{xy}\right]+C+B \ge 0$$

Condition 3: Using equation (6) and (30)

$$MSE\left(\hat{\hat{Y}}_{GKm}\right)_{min} \le V\left(\hat{\bar{Y}}_{lr}\right)$$

If

$$\frac{\bar{Y}^2C_{\bar{y}}^2\left[\rho_{yx}^2+2\rho_{\bar{y}m}^2\right]}{1+C_{\bar{y}}^2\left[1-\rho_{\bar{y}m}^2\right]}-C+B \ge 0$$

Condition 4: Using equation (9) and (30)

$$MSE\left(\hat{\hat{Y}}_{GKm}\right)_{min} \le MSE\left(\hat{\bar{Y}}_{BT}\right)$$

If

$$\lambda\bar{Y}^2\left[\frac{C_x^2}{4}-C_{yx}\right]+C+B\geq 0$$

Condition 5: Using equation (12) and (30)

$$MSE\left(\hat{\bar{Y}}_{GKm}\right)_{min}\leq MSE\left(\hat{\bar{Y}}_{Rao}\right)$$

If

$$\frac{V(\bar{y})(1-\rho^2)}{1+C_{\bar{y}}^2(1-\rho^2)}-\frac{\bar{Y}^2C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)}{1+C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)}+B\geq 0$$

Condition 6: Using equation (15) and (30)

$$MSE(\hat{\bar{Y}}_{GKm})_{min}\leq MSE(\hat{\bar{Y}}_{GK})_{min}$$

If

$$\frac{\bar{Y}^2C_{\bar{y}}^2\left(\rho_{yx}^2+2\rho_{\bar{y}m}^2\right)}{1+C_{\bar{y}}^2\left(1-\rho_{\bar{y}m}^2\right)}-C+B\geq 0$$

Condition 7: Using equation (19) and (30)

$$MSE(\hat{\bar{Y}}_{GKm})_{min}\leq MSE(\hat{\bar{Y}}_{BTm})$$

If

$$\bar{Y}^2\left(\rho_{\bar{y}m}C_{\bar{y}}-\frac{C_m}{2}\right)^2+A+B\geq 0$$

Condition 8: Using equation (21) and (30)

$$MSE\left(\hat{\bar{Y}}_{GKm}\right)_{min}\leq MSE\left(\hat{\bar{Y}}_{dm}\right)$$

If

$$A+B\geq 0$$

Condition 9: Using equation (25) and (30)

$$MSE(\hat{\bar{Y}}_{GKm})_{min}\leq MSE(\hat{\bar{Y}}_{Raom})$$

If

$B\geq 0$

Where

$$A=\frac{\bar{Y}^2\left[C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)\right]^2}{1+C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)}\geq 0 \qquad B=\frac{\bar{Y}^2C_{\mathrm{m}}^2\left[C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)+\frac{C_{\mathrm{m}}^2}{16}\right]}{4\left[1+C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)\right]}\geq 0$$

$$C = \frac{\bar{Y}^2 C_{\bar{y}}^2 \left[C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)+\rho_{\bar{y}m}^2\right]}{1+C_{\bar{y}}^2\left(1+\rho_{\bar{y}m}^2\right)} \geq 0 \qquad D = \frac{\lambda^2 \bar{Y}^2 C_{x}^2 \left[C_{y}^2\left(1+\rho_{yx}^2\right)+\frac{C_x^2}{16}\right]}{4\left[1+\lambda C_{y}^2\left(1+\rho^2\right)\right]} \geq 0$$

Condition 1,7,8 and 9 are always true.

4 Empirical Study

We consider data sets given in Singh and Mangat [10]. The first data is related to project of feeding and management practice of cows, where the number of milch cows in a district from 1990 census is taken as auxiliary variable X and the number of milch cows in the same district from 1993 census is treated as a study variable Y. The second data is related to the list of ration depots. The information in the respect of family members are collected. The number of female is the auxiliary variable X and the number of male is the study variable Y. The summary results are given, in Table 1.

Table 2 show the values of MSE's of all existing and proposed estimators. We can observe that the proposed estimators have small MSE values than the existing estimators. In Table 3 the percentage relative efficiency has been shown with respect to variance of simple random sampling variance. These results explain that the proposed estimators are more efficient than the existing estimators.

Table 1. Summery statistics for different sample size

Parameters	Population 1		Population 2	
	***n*=3**	***n*=5**	***n*=3**	***n*=5**
N	24	24	27	27
Y	19.1667	19.1667	725.37	725.37
$\bar{M}$	19.4575	19.5895	731.92	734.08
$\bar{X}$	16.7917	16.7917	644.48	644.48
R_1	1.1414	1.1414	1.1255	1.1255
R_2	0.9851	0.9784	0.9911	0.9881
$V(\bar{y})$	4.7620	2.5839	9651.10	5306.35
$\mathrm{V}(m)$	7.1272	4.1052	16765.87	10880.47
$V(\bar{x})$	6.6477	3.6070	5214.85	2867.22
$Cov(\bar{y},m)$	5.0051	2.6678	11276.64	6582.336
$Cov(\bar{y}, \bar{x})$	3.0429	1.6511	5330.98	2931.071
ρ	0.5408	0.5408	0.7515	0.7515

Table 2. Variance/MSE of different estimators for different sample size

Estimator	Population - 1		Population - 2	
	n=3	**n=5**	**n=3**	**n=5**
$\bar{y}$	4.75966	2.58382	9647.80	5306.29
using auxiliary variable.				
$\hat{\bar{Y}}_R$	6.47667	3.51425	4256.99	2340.58
$\hat{\bar{Y}}_{lr}$	3.36919	1.82810	4201.41	2310.02
$\hat{\bar{Y}}_{BT}$	3.45235	1.87413	5300.72	2915.40
$\hat{\bar{Y}}_{Rao}$	3.26491	1.77283	4089.80	2249.39
$\hat{\bar{Y}}_{GK}$	3.31412	1.81228	4152.36	2295.53
using median of study variable.				
$\hat{\bar{Y}}_{BTm}$	1.56065	0.95611	2592.14	1458.07
$\hat{\bar{Y}}_{dm}$	1.24716	0.85010	2066.49	1324.25
$\hat{\bar{Y}}_{Raom}$	1.24294	0.84818	2058.41	1320.93
$\hat{\bar{Y}}_{GKm}$	1.23945	0.84695	2046.36	1315.92

Table 3. PRE of estimators with respect to $\bar{y}$

Estimator	Population - 1		Population - 2	
	n=3	n=5	n=3	n=5
$\bar{y}$	100	100	100	100
using auxiliary variable.				
$\hat{\bar{Y}}_R$	73.4893	73.5239	226.634	226.710
$\hat{\bar{Y}}_{lr}$	141.270	141.3366	229.632	229.709
$\hat{\bar{Y}}_{BT}$	137.8674	137.8674	182.009	182.009
$\hat{\bar{Y}}_{Rao}$	145.7822	145.7822	235.8993	235.899
$\hat{\bar{Y}}_{GK}$	143.6176	142.5725	232.3451	231.1574
using median of study variable.				
$\hat{\bar{Y}}_{BTm}$	304.979	270.2444	372.195	363.926
$\hat{\bar{Y}}_{dm}$	381.640	303.9286	466.868	400.701
$\hat{\bar{Y}}_{Raom}$	382.936	304.632	468.702	401.710
$\hat{\bar{Y}}_{GKm}$	384.014	305.072	471.461	403.239

Table 4. Summary of efficiency conditions proved in section 3

Efficiency Conditions	Population – 1		Population – 2	
	n=3	*n*=5	*n*=3	*n*=5
One	3.5252	1.7381	7608.795	3992.094
Two	5.2390	2.6685	2216.538	1026.353
Three	4.8811	2.4802	12961.65	6954.269
Four	2.2179	1.0285	3216.723	1601.205
Five	2.0281	0.9266	2047.493	935.1270
Six	4.8840	2.4802	12961.65	6954.269
Seven	0.3239	0.1104	549.8355	143.8099
Eight	0.0104	0.0045	24.19046	9.992613
Nine	0.0061	0.0024	16.10609	6.668079

5 Simulation Study

In simulation study data sets are generated from three different distributions. One of which is skewed that is chi-square, the remining two are symmetric Logistic and Uniform distributions. Different Parameters are used for all distributions given in the Tables 5, 6 and 7.

Table 5. Simulation from chi-square distribution

Parameters		Estimators									
	d.f	$\bar{y}$	$\hat{\bar{Y}}_R$	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{BT}$	$\hat{\bar{Y}}_{Rao}$	$\hat{\bar{Y}}_{GK}$	$\hat{\bar{Y}}_{BTm}$	$\hat{\bar{Y}}_{dm}$	$\hat{\bar{Y}}_{Raom}$	$\hat{\bar{Y}}_{GKm}$
n=3	5	100	62.3167	100.352	88.8017	138.336	114.77	346.985	401.965	413.865	429.822
	10	100	46.7525	99.9981	78.0237	119.611	108.197	312.57	369.461	375.605	381.400
	15	100	45.1302	100.015	77.1652	114.187	106.042	284.998	349.486	353.925	356.990
n=5	5	100	40.7211	102.28	69.2737	134.739	110.085	223.814	253.546	259.389	261.422
	10	100	53.7337	100.058	81.5573	120.318	104.524	212.774	251.807	255.454	256.441
	15	100	49.7526	100.439	77.9068	115.22	103.745	249.126	283.021	285.682	286.967

Table 6. Simulation from logistic distribution

Parameters		Estimators									
n=3		$\bar{y}$	$\hat{\bar{Y}}_R$	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{BT}$	$\hat{\bar{Y}}_{Rao}$	$\hat{\bar{Y}}_{GK}$	$\hat{\bar{Y}}_{BTm}$	$\hat{\bar{Y}}_{dm}$	$\hat{\bar{Y}}_{Raom}$	$\hat{\bar{Y}}_{GKm}$
μ=5	s=1	100	61.259	100.072	87.209	114.379	105.393	269.502	347.899	352.381	354.918
	s=3	100	45.514	100.777	74.421	252.038	170.025	248.966	331.811	379.205	403.626
	s=5	100	45.204	102.246	82.672	437.522	382.961	237.797	345.404	450.456	505.735
μ=10	s=1	100	49.480	100.778	82.874	103.959	102.110	287.309	366.679	367.675	368.358
	s=3	100	66.669	100.009	88.571	141.768	114.970	213.233	273.197	286.281	289.745
	s=5	100	43.679	100.125	76.853	155.989	125.849	306.849	408.639	426.142	441.944
n=5											
μ=5	s=1	100	47.037	102.241	83.946	111.859	104.710	330.590	363.274	365.006	367.035
	s=3	100	29.405	109.278	75.298	150.472	126.854	295.720	332.237	339.652	345.955
	s=5	100	15.392	107.679	37.683	237.257	195.711	331.582	378.036	401.359	431.071
μ=10	s=1	100	40.457	100.024	72.611	103.409	100.873	225.143	255.966	256.576	256.785
	s=3	100	64.872	100.932	91.316	139.612	109.290	206.963	249.597	256.559	258.256
	s=5	100	57.049	100.611	86.863	209.820	126.179	215.742	264.036	283.693	289.501

Table 7. Simulation from Uniform distribution

Parameters		Estimators									
a=0		$\bar{y}$	$\hat{\bar{Y}}_R$	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{BT}$	$\hat{\bar{Y}}_{Rao}$	$\hat{\bar{Y}}_{GK}$	$\hat{\bar{Y}}_{BTm}$	$\hat{\bar{Y}}_{dm}$	$\hat{\bar{Y}}_{Raom}$	$\hat{\bar{Y}}_{GKm}$
n=3	b=1	100	56.7893	104.642	93.1577	138.769	119.578	379.705	445.253	455.945	474.201
	b=3	100	42.8234	100.809	72.3695	124.724	110.85	371.135	465.54	473.032	484.548
	b=5	100	65.9306	102.662	94.6244	145.855	119.631	466.057	527.908	541.44	583.883
n=5	b=1	100	47.9911	102.575	74.5267	139.316	110.729	360.09	398.241	404.855	414.882
	b=3	100	108.814	122.943	122.512	165.977	132.221	386.304	415.468	423.214	438.196
	b=5	100	49.6328	100.512	77.6881	130.632	107.285	384.993	409.151	414.572	425.051

In simulation study data is generated by using chi-square, Logistic and Uniform distributions. For different sample size and different values of parameters for all distribution Tables 5, 6 and 7 summarized the values of percentage relative efficiency with respect to simple variance. The results show that proposed median based estimators perform much better than mean based existing estimators.

6 Conclusion

Survey sampling practitioners have been working on efficiency improvement and bias reduction in finite population parameter estimation. In this direction, Bahl and Tetuja [7] considered exponential type estimator for improving efficiency of the mean estimator. The article delineated some Bahl and Tetuja [7] type estimators based on known population median of the study variable for improving the efficiency of the mean estimator taking motivation from Grover and Kaur [9]. The novel median based estimators perform better than the usual mean, ratio and regression estimators in terms of efficiency. An empirical study is conducted to show the superiority of the novel median based estimator over its existing counter parts in presence of auxiliary data. The novel estimator can be extended to more complex sampling design.

Acknowledgements

The authors are thankful to the Editor, Journal editorial office and anonymous referees for their fruitful comments for the success of this work.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Hafeez W, Shabbir J. Estimation of the finite population mean, using median based estimators in stratified random sampling. Journal of Statistics Applications & Probability. 2015;4(3):367-374.

[2] Milton TK. Estimation of population variance using the Coe Cient of kurtosis and median of an auxiliary variable under simple random sampling (Doctoral dissertation, JKUAT); 2018.

[3] Gruber M. Improving efficiency by shrinkage: The James--stein and ridge regression estimators. Routledge; 2017.

[4] Onsongo WM, Otieno RO, Orwa GO. Bias reduction technique for estimating finite population distribution function under simple random sampling without replacement; 2018.

[5] Ahmed S, Shabbir J, Hafeez W. Hansen and Hurwitz Estimator with scrambled response on second call in stratified random sampling. Journal of Statistics Applications & Probability. 2016;5(2):1-13.

[6] Watson D. The estimation of leaf area in field crops. Journal of Agricultural Science. 1937;27(3):474-483.

[7] Bahl S, Tuteja R. Ratio and product type exponential estimators. Journal of Information and Optimization Sciences. 1991;12(1):159–164.

[8] Rao T. On certail methods of improving ration and regression estimators. Communications in Statistics-Theory and Methods. 1991;20(10):3325–3340.

[9] Grover LK, Kaur P. An improved estimator of the finite population mean in simple random sampling. Model Assisted Statistics and Applications. 2011;6(1):47–55.

[10] Singh R, Mangat NS. Elements of Survey Sampling, volume 15. Springer; 1996.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/58543