EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2025, Vol. 18, Issue 2, Article Number 5922
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Generalized Class of Estimators for Median Estimation Using Auxiliary Information

Sohaib Ahmad[1,*], Saadia Masood[2], Manahil SidAhmed Mustafa[3], Elsiddig Idriss Mohamed[3], Elfarazdag M. M. Hussein[3]

[1] *Department of Statistics, Abdul Wali Khan University, Mardan, Pakistan*
[2] *Department of Mathematics and Statistics, PMAS University of Arid Agriculture, Rawalpindi, Pakistan*
[3] *Department of Statistics, Faculty of Science, University of Tabuk, Tabuk, Kingdom of Saudi Arabia*

Abstract. In this article, we propose a comprehensive class of estimator for estimation of population median under simple random sampling. To boost the efficiency of an estimator we utilize the auxiliary information. The numerical expression of the bias and mean squared error are consequent up to the first order of approximation. The proposed and existing estimators have been evaluated via real data sets and their performances were assessed using measurements of minimal mean square error and maximum percentage relative efficiency. The study showed that compared to some adopted existing estimators in this study, the proposed class of estimators performed better and efficient. We also visualize all the estimators using results of MSE and PRE. According to the results, the suggested estimator is superior to the adopted existing estimators. The significance and potential applications of our proposed class of estimators are highlighted by these results.

2020 Mathematics Subject Classifications: 62D05, 62f10

Key Words and Phrases: Median estimation, visualization, auxiliary information, MSE, efficiency

---

1. Introduction

The accuracy of parameter estimates for a population can be greatly enhanced with the incorporation of auxiliary infromation during the selection or estimation process, or both. Although many determinations have been made to progress estimator precision, estimating the population mean remains a persistent challenge in sampling. When the auxiliary information is taken into account, a huge number of alternatives become accessible. In

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v18i2.5922

Email addresses: sohaib_ahmad@awkum.edu.pk (S. Ahmad), saadia.masood@uaar.edu.pk (S. Masood), msida@ut.edu.sa (M. S. Mustafa), eidriss@ut.edu.sa (E. I. Mohamed), e.hussein@ut.edu.sa (E. M. M. Hussein)

order to deal with situations involving a large number of covariates, numerous estimators have been developed, consisting of incorporating aspects from ratio, product, and regression estimators. The variance, the coefficient of variation, and the kurtosis have all been used extensively by researchers for parameter estimation in populations. For this to work, we need to draw from a representative subset of the population. The population of interest is chosen with the help of simple random sampling. Using the ratio, product, or regression estimation approaches requires advanced knowledge of the population constraints of the auxiliary information. By adjusting the appropriate auxiliary information, many authors have provided alternative estimators. It is generally agreed that estimators who use auxiliary information in survey sampling can yield implicitly better estimates than those who do not. Sampling is a technique for efficiently and effectively collecting information about a population in a way that maximizes the precision of estimates with minimal outlay of resources. In order to make more precise estimates, sampling is done to gain insight into a population's characteristics with minimal effort. Since these methods yield reliable estimates, we employ the mean and standard least-squares approaches to estimate population parameters.

As a consequence, the data may not follow a normal distribution but exhibit severely skewed distributions (for example wages and consumption). Due to the mean sensitivity with outliers, it cannot be used to reliably compute these quantities. Therefore, the median is unaffected by outliers and extreme values, it can be used as a measure of center tendency. Attempts to provide a novel method for obtaining a reliable conclusion in such cases are notoriously challenging for academics. The use of auxiliary information in median estimation depends on the idea that related variables can enhance target variable estimation accuracy. Considering auxiliary information helps to both strengthen and stabilize the estimation results particularly when the original data has limited quantity or missing values. Through these methods the estimator accounts for new information facilitating both variance reduction and improve efficiency. A regression estimator can use auxiliary variables to recognize how the target variable associates with covariates thus enabling better median predictions. From a theoretical point of view the addition of auxiliary information enhances two important statistical criteria consisting of efficiency and consistency. The estimation method which takes auxiliary information into consideration delivers lower MSE results than methods working with observed data only. Some conditions allow auxiliary information to create estimators that deliver robust results particularly when outliers occur.

Several researchers have recommended estimators, either by making changes to the already existing estimators or by building completely novel estimators. Some notable work by these authors includes [3, 4, 5], [7–15], [16, 17, 18, 19, 20, 21, 22, 23] and [24]. The main purpose of this work is developed a new generalized class of estimators using auxiliary information for achieving precise median estimation. This work increases the efficiency of median estimation performance in multiple statistical models through deliberate use of auxiliary information. The study investigates both consistency and unbiasedness as statistical properties of the proposed estimators. The paper performs an analysis of new estimators versus standard methods while demonstrating their use in practical situations

to explain enhanced accuracy potential in estimation.
This paper makes a significant contribution toward enhancing median estimation accuracy through the use of auxiliary information. The proposed estimators provide flexibility which enables them to process diverse data types so they can be utilized throughout economic sectors and medical practices and social science applications. Median estimation improvements through auxiliary data lead to enhanced decision-making in both policy analysis and medical diagnostics along with precise statistical analysis of other fields. The paper includes a theoretical evaluation of these generalized estimators' statistical characteristics which includes unbiasedness and consistency and efficiency to prove their validity. The research has major implications because it enhances both statistical methods and field applications of median estimation techniques in empirical situations.
To account for situations in which the population distribution is not normally distributed, we generate estimators for estimation of population median in the current study.
The remaining of the article is arranged as follows:
Section 2, include the procedures and materials. The literature review of the available estimator for estimate of median under simple random sample is given in Section 3. In Section 4, we provide the suggested estimator. The numerical study is given in Section 5. In Section 6, the article findings and conclusions are outlined.

2. Methods and materials

Consider a population $\Omega = (\Omega_1, \Omega_2, \ldots, \Omega_N)$ consisting of N divergent units. Let Y_i and X_i be the i^{th} standards of the study and auxiliary variables. Consider a sample of n is chosen from Ω. The sample and population medians of the study and the auxiliary variable are denoted by M_y, M_x, and $\tilde{M}_y$ and $\tilde{M}_x$ with probability density functions of $f_y(M_y)$ and $f_x(M_x)$. The correlation coefficient between $\tilde{M}_y$ and $\tilde{M}_x$ are represented by ρyx and is defined as $\rho_{yx}(\tilde{M}_y, \tilde{M}_x) = 4P_{11}(y, x) - 1$, where $P_{11} = (y \leq M_y \cap x \leq M_x)$. To obtain the bias and MSE we used the following error terms:
$\xi_0 = (\frac{\tilde{M}_y - M_y}{M_y})$ and $\xi_1 = (\frac{\tilde{M}_x - M_x}{M_x})$,

$$E(\xi_0^2) = \lambda C_{My}^2, \quad E(\xi_1^2) = \lambda C_{Mx}^2, \quad E(\xi_0 \xi_1) = \lambda C_{Myx},$$

where

$$C_{My} = [M_y f_y(M_y)]^{-1}, \quad C_{Mx} = [M_x f_x(M_x)]^{-1}, \quad C_{Myx} = [\rho_{yx} C_{My} C_{Mx}], \quad \lambda \left(\frac{1}{4}(\frac{1}{n} - \frac{1}{N})\right)$$

3. Literature review

In this section, we have discussed existing estimators for population median which are given by:

(i) The [9] recommended a usual median estimator, which is given by:

$$\tilde{M}_U = \tilde{M}_y \tag{1}$$

The variance of $\tilde{M}_U$ is given by:

$$Var(M_U) = \lambda M_y^2 C_{My}^2 \tag{2}$$

(ii) The [12] recommended ratio estimator for median, is given by:

$$\tilde{M}_R = \tilde{M}_y \left(\frac{M_x}{\tilde{M}_x} \right) \tag{3}$$

The bias and MSE of M_R are given by:

$$Bias(M_R) = \lambda M_y [C_{My}^2 - C_{Myx}],$$

and

$$MSE(\tilde{M}_R) \approx \lambda M_y^2 [C_{My}^2 + C_{Mx}^2 - 2C_{Myx}]. \tag{4}$$

(iii) The [14] suggested the usual difference estimator $\tilde{M}_D$, which is given by:

$$\tilde{M}_D = \tilde{M}_y + d_1 (M_x - \tilde{M}_x) \tag{5}$$

The variance of $(\tilde{M}_D)$:

$$V(\tilde{M}_D) = \lambda M_y^2 C_{My}^2 (1 - \rho_{yx}^2) \tag{6}$$

(iv) The exponential ratio estimator developed by [24] is given by:

$$\tilde{M}_E = \tilde{M}_y \text{exp} \left(\frac{M_x - \tilde{M}_x}{M_x + \tilde{M}_x} \right), \tag{7}$$

The bias and MSE of $(\tilde{M}_E)$, are given by:

$$\text{Bias}(\tilde{M}E) \cong \lambda M_y \left[\frac{3}{4} C_{M_x}^2 - \frac{1}{2} C_{M_{yx}} \right],$$

$$\text{MSE}(\tilde{M}E) \cong \lambda M_y^2 \left[C_{My}^2 + \frac{1}{4} C_{Mx}^2 - C_{M_{yx}} \right]. \tag{8}$$

(v) The [13] proposed the difference-in-difference estimator, which is given by:

$$\tilde{M}_{RD_1} = d_2 \tilde{M}_y + d_3 (M_x - \tilde{M}_x), \tag{9}$$

the values of (d_2) and (d_3) are given by:

$$d_2 = \frac{1}{1 + \lambda C_{My}^2 (1 - \rho_{yx}^2)}$$

$$d_3 = \frac{M_y C_{My} \rho_{yx}}{M_x C_{Mx}\{1 + \lambda C_{My}^2(1 - \rho_{yx}^2)\}}.$$

The minimum MSE of $(\tilde{M}_{RD_1})$ is given by:

$$\text{MSE}(\tilde{M}RD_1)\text{min} = \frac{\lambda M_y^2 C_{My}^2 (1 - \rho_{yx}^2)}{[1 + \lambda C_{My}^2(1 - \rho_{yx}^2)]} \tag{10}$$

(vi) The [8] recommended the following difference-in-ratio type estimator, which is given as:

$$\tilde{M}_{RD_2} = (d_4 \tilde{M}_y + d_5(M_x - \tilde{M}_x)) \left(\frac{M_x}{\tilde{M}_x}\right) \tag{11}$$

The (d_4) and (d_5) are constants, which are given by:

$$d_4 = \left[\frac{1 - \lambda C_{Mx}^2}{1 - \lambda C_{Mx}^2 + \lambda C_{My}^2(1 - \rho_{yx}^2)}\right],$$

$$d_5 = \frac{M_y}{M_x}\left[\frac{M_y C_{My} \rho_{yx}}{M_x C_{Mx}\{1 + \lambda C_{My}^2(1 - \rho_{yx}^2)\}}\right].$$

Substituting the values of (d_4) and (d_5), are given by:

$$\text{MSE}(\tilde{M}_{RD_2}) = \frac{\lambda M_y^2 C_{My}^2 (1 - \rho_{yx}^2)(1 - \rho_{yx}^2)}{[(1 - \lambda C_{My}^2) + \lambda C_{My}^2(1 - \rho_{yx}^2)]} \tag{12}$$

(vii) The [15] given the generalized class of estimator:

$$\tilde{M}_s = \tilde{M}_y \exp\left(\frac{a(M_x - \tilde{M}_x)}{a(M_x + \tilde{M}_x) + 2b}\right) \tag{13}$$

The properties of $(\tilde{M}_S)$ is given by:

$$\text{Bias}(\tilde{M}s) = M_y\left[\frac{3}{8}\theta^2 CMy^2 - \frac{1}{2}\theta C_{Myx}\right],$$

$$\text{MSE}(\tilde{M}s) = \frac{M_y^2}{4}\left[4CMy^2 + \theta^2 C_{Mx}^2 - 4\theta C_{Myx}\right] \tag{14}$$

where $\theta = \left(\frac{aM_x}{aM_x + b}\right)$

4. Suggested generalized class of estimator

When the data is collected with the help of survey sampling, it is possible that the data follow a normal distribution. In order to get reliable estimates of population parameters, we employ the mean and standard least-squares methods. Sometimes the acquired data may not normally distributed (in terms of wage, consumption, etc.), but follow extremely skewed distributions. In such scenario, it is not reliable to calculate quantities like mean etc. In such circumstances, the median is used, because it is unaffected by outliers and extreme values. Median can be used as a measure of center tendency. By taking motivation from [5], we suggested the resulting improved class of estimators for population median. This work makes an essential contribution to statistical estimation because it introduces a new approach to median estimation through generalized class of estimators using auxiliary information. The framework developed by this research enables the integration of auxiliary data with the retention of essential median estimator features including robustness and consistency. A new approach enhances precision in estimation because researchers need exact median results in economics and medical research and social science studies especially when data is restricted. The paper verifies the statistical aspects of the proposed estimators while confirming both theoretical and practical proof for their application. This research develops an innovative statistical tool that gives practitioners and researchers improved ways to increase the accuracy of their analytical work through performance comparison with traditional methods.

$$\begin{aligned}\tilde{M}_{GP} = \psi_1 M_y \left(\frac{1}{4}\left(\frac{M_x}{\bar{M}_x} + \frac{\bar{M}_x}{M_x}\right) \times \left(\exp\left(\frac{M_x - \bar{M}_x}{M_x + \bar{M}_x}\right) + \exp\left(\frac{\bar{M}_x - M_x}{M_x + \bar{M}_x}\right)\right)\right) \\ + \psi_2 \left(M_x - \bar{M}_x\right) \exp\left(\frac{\alpha(M_x - \bar{M}_x)}{\alpha(M_x + \bar{M}_x) + 2\beta}\right) \qquad (15)\end{aligned}$$

ψ_1 and ψ_2 are the constants.

After simplification of $\tilde{M}_{GP}$, we have

$$\tilde{M}_{GP} = \left[\psi_1 M_y (1 + \xi_0)\left(1 + \frac{5}{8}\xi_1^2\right) - \psi_2 M_x \xi_1\right]\left[1 - \frac{1}{2}\theta\xi_1 + \frac{3\theta^2}{8}\xi_1^2\right] \qquad (16)$$

Expanding (16), we get

$$\tilde{M}_{GP} - M_y = M_y + M_y \left[(\psi_1 - 1) + \psi_1 \left\{\xi_0 - \frac{1}{2}\theta\xi_1 - \frac{1}{2}\theta\xi_0\xi_1 + \frac{1}{8}(5 + 3\theta^2)\xi_1^2\right\} - \psi_2 R\{\xi_1 - \frac{1}{2}\theta\xi_1^2\}\right] \qquad (17)$$

From (17), the bias of $\tilde{M}_{GP}$ is given by:

$$Bias(\tilde{M}GP) = M_y \left[(\psi_1 - 1) + \psi_1 \left\{\frac{1}{8}(5 + 3\theta^2)CMx - \frac{1}{2}\theta C_{Myx}\right\} + \psi_{19} R\lambda \frac{1}{2}\theta C_{Mx}^2\right] \qquad (18)$$

Squaring (17) and taking expectations:

$$\text{MSE}(\tilde{M}_{GP}) = M_y^2[1 + \psi_1^2 A_{11} + \psi_2^2 B_{11} - 2\psi_1 C_{11} - 2\psi_2 D_{11} + 2\psi_1\psi_2 E_{11}], \qquad (19)$$

where

$$A_{11} = 1+\lambda\left[C_{My} + \left(\frac{5}{4}+\theta^2\right)C_{Mx}^2 - 2\theta C_{Myx}\right], \quad B_{11} = R^2\lambda[C_{Mx}^2], \quad C_{11} = 1+\lambda\left[\left(\frac{5+3\theta^2}{8}\right)C_{Mx}^2 - \frac{1}{2}\theta C_{Myx}\right],$$

$$D_{11} = R\lambda\left[\frac{\theta C_{Mx}^2}{2}\right], \quad E_{11} = R\lambda[\theta C_{Mx}^2 - C_{Myx}].$$

Differentiate (19) w.r.t ψ_1 and ψ_2, the values of ψ_1 and ψ_2 as given by:

$$\psi_1(\text{opt}) = \frac{B_{11}C_{11} - D_{11}E_{11}}{A_{11}B_{11} - E_{11}^2},$$

and

$$\psi_2(\text{opt}) = \frac{A_{11}D_{11} - C_{11}E_{11}}{A_{11}B_{11} - E_{11}^2}.$$

Putting the optimum values of $\psi_1(\text{opt})$ and $\psi_2(\text{opt})$ in (19), we get the minimum MSE of $\tilde{M}_{GP}$ as given by:

$$\text{MSE}(\tilde{M}GP)\text{min} \approx M_y^2\left[1 - \frac{A_{11}D_{11}^2 + B_{11}C_{11}^2 - 2C_{11}D_{11}E_{11}}{A_{11}B_{11} - E_{11}^2}\right]. \quad (20)$$

5. Numerical study

To determine how effective our suggested class of estimators is, we perform a mathematical analysis with some actual data. We compare the effectiveness of our suggested class to that of existing estimators by using a percentage relative efficiency (PRE). The theoretical expression for PRE is given by:

PRE $= \frac{Var(\tilde{M}_U)}{MSE(\tilde{M}_i)} \times 100$

Where (i=R, D, E, RD1, RD2, S, GP)

Population 1: [Source: [14]]

Y = Fish caught in 1995,

X = fish caught in 1994.

$N = 69, \quad n = 17, \quad M_y = 2068, \quad M_x = 2011, \quad f_{m_y} = 0.00014,$

$f_{m_x} = 0.00014, \rho_{yx} = 0.1505.$

Population 2: [Source: [2]]

Y = U.S export in Singapore

X = money supply in Singapore.

$N = 67, \quad n = 23, \quad M_y = 4.8, \quad M_x = 7, \quad f_{m_y} = 0.48294,$

$f_{m_x} = 0.343079, \quad \rho_{yx} = 0.61194.$

Population 3: [Source: [1]]

Y = Oil price from 1996 to 2017

X = oil price in preceding from 1996 to 2017.

$N = 1134, \quad n = 210, \quad M_y = 48.5500, \quad M_x = 48.4900,$

$f_{m_y} = 0.00754, \quad f_{m_x} = 0.00754, \quad \rho_{yx} = 0.99530.$

Population 4: [Source: [6]]

Y = Master degree in 2007

X = Master degree in 2006.

$N = 51, \quad n = 11, \quad M_y = 25.8000, \quad M_x = 25.6000,$
$f_{m_y} = 0.07280, \quad f_{m_x} = 0.00754, \quad \rho_{yx} = 0.99530.$

Table 1: Members of the recommended class of estimators

α	β	$\tilde{M}_S$	$\tilde{M}_{GP}$
1	Cx	$\tilde{M}_{S1}$	$\tilde{M}_{GP1}$
1	$\beta_2(x)$	$\tilde{M}_{S2}$	$\tilde{M}_{GP2}$
$\beta_2(x)$	Cx	$\tilde{M}_{S3}$	$\tilde{M}_{GP3}$
Cx	$\beta_2(x)$	$\tilde{M}_{S4}$	$\tilde{M}_{GP4}$
1	ρ_{yx}	$\tilde{M}_{S5}$	$\tilde{M}_{GP5}$
Cx	ρ_{yx}	$\tilde{M}_{S6}$	$\tilde{M}_{GP6}$
ρ_{yx}	Cx	$\tilde{M}_{S7}$	$\tilde{M}_{GP7}$
$\beta_2(x)$	ρ_{yx}	$\tilde{M}_{S8}$	$\tilde{M}_{GP8}$
ρ_{yx}	$\beta_2(x)$	$\tilde{M}_{S9}$	$\tilde{M}_{GP9}$
1	NM_x	$\tilde{M}_{S10}$	$\tilde{M}_{GP10}$

Table 2: MSE using Population 1

$Estimators$	$Values$	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	565443.600	627047.800	373018.400
$\tilde{M}_R$	988372.800	626625.000	373125.800
$\tilde{M}_D$	552636.100	627371.200	372936.300
$\tilde{M}_E$	627420.200	626097.100	372981.000
$\tilde{M}_{RD1}$	489395.200	627404.400	372927.900
$\tilde{M}_{RD2}$	480458.300	627415.800	372925.000
		624987.400	373542.300
		627418.100	372924.400
		622322.700	374222.000
		564223.900	398752.300

Table 3: PRE using Population 1

$Estimators$	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	100	90.175	151.586
$\tilde{M}_R$	57.209	90.236	151.542
$\tilde{M}_D$	102.317	90.129	151.619
$\tilde{M}_E$	90.121	90.312	151.601
$\tilde{M}_{RD1}$	115.539	90.124	151.622
$\tilde{M}_{RD2}$	117.688	90.122	151.623
		90.472	151.373
		90.122	151.624
		90.860	151.098
		100.216	141.803

Table 4: MSE using Population 2

Estimators	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	0.0306078	0.0198950	0.019076
$\tilde{M}_R$	0.0229662	0.019150	0.019068
$\tilde{M}_D$	0.019146	0.019498	0.019073
$\tilde{M}_E$	0.019658	0.019477	0.019039
$\tilde{M}_{RD1}$	0.019130	0.020011	0.019076
$\tilde{M}_{RD2}$	0.019130	0.020535	0.019078
		0.020052	0.019077
		0.019430	0.019073
		0.019431	0.019059
		0.030343	0.0190869

Table 5: PRE using Population 2

$Estimators$	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	100	153.846600	160.449400
$\tilde{M}_R$	133.273200	159.825500	160.519400
$\tilde{M}_D$	159.864600	156.972900	160.470600
$\tilde{M}_E$	155.701300	157.144600	160.758400
$\tilde{M}_{RD1}$	159.997400	152.952900	160.444700
$\tilde{M}_{RD2}$	159.997600	149.047300	160.427900
		152.638500	160.443100
		157.528400	160.475800
		157.519800	160.588600
		100.871100	160.360200

Table 6: MSE using Population 3

Estimators	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	17.06228000	4.79872700	0.09149240
$\tilde{M}_R$	0.16061000	5.56053900	0.09486919
$\tilde{M}_D$	0.16000850	4.39771000	0.08942814
$\tilde{M}_E$	4.33531100	5.23074000	0.09153061
$\tilde{M}_{RD1}$	0.15999760	4.50701300	0.09001304
$\tilde{M}_{RD2}$	0.15999750	4.39849600	0.08943241
		4.80085300	0.09150277
		4.35807100	0.09150277
		5.56585900	0.09150277
		17.047300	0.09150277

Table 7: PRE using Population 3

Estimators	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	100	355.5584	18648.8400
$\tilde{M}_R$	10623.4200	306.8457	17985.0500
$\tilde{M}_D$	10663.3600	387.9809	19079.3100
$\tilde{M}_E$	393.5652	326.1924	18641.0600
$\tilde{M}_{RD1}$	10664.0800	378.5717	18955.3400
$\tilde{M}_{RD2}$	10664.0900	387.9116	19078.4000
		355.4009	18646.7300
		391.5098	19125.6100
		306.5524	17981.0000
		100.0879	15439.0300

Table 8: MSE using Population 4

Estimators	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	3.363368	1.497556	0.02741737
$\tilde{M}_R$	0.3659373	1.840453	0.0276705
$\tilde{M}_D$	0.02953252	1.47938	0.02740078
$\tilde{M}_E$	1.476601	1.751034	0.02785551
$\tilde{M}_{RD1}$	0.02953121	1.533249	0.02744888
$\tilde{M}_{RD2}$	0.02953121	1.625763	0.02752447
		1.497648	0.02741745
		1.48423	0.02740524
		1.841761	0.02767128
		3.319765	0.02803544

Table 9: PRE using Population 4

Estimators	Values	$\tilde{M}S$	$\tilde{M}GP$
$\tilde{M}_U$	100	224.5904	12267.29
$\tilde{M}_R$	919.1105	182.7468	12155.07
$\tilde{M}_D$	11388.6900	227.3499	12274.72
$\tilde{M}_E$	227.7777	192.079	12074.34
$\tilde{M}_{RD1}$	11389.2000	219.3621	12253.21
$\tilde{M}_{RD2}$	11389.3000	206.8794	12219.56
		224.5767	12267.25
		226.6069	12272.72
		182.617	12154.72
		101.3134	11996.84

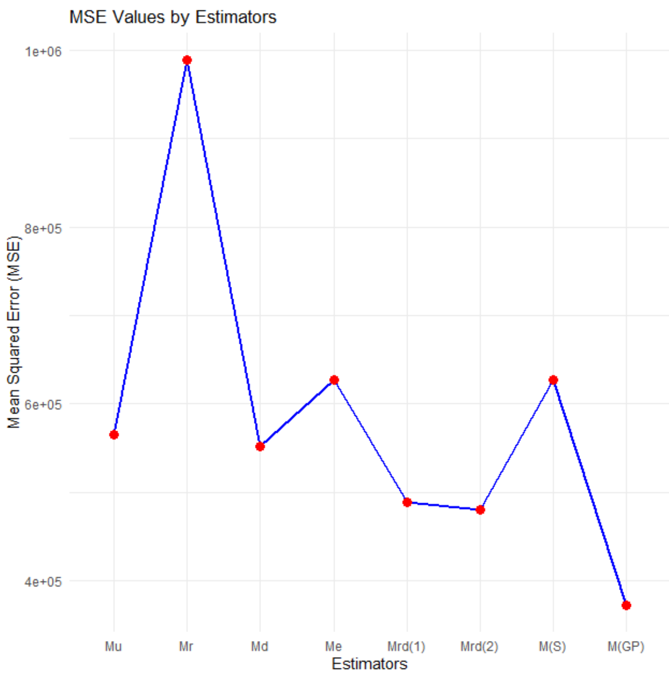

Figure 1: MSE of median estimators using population 1

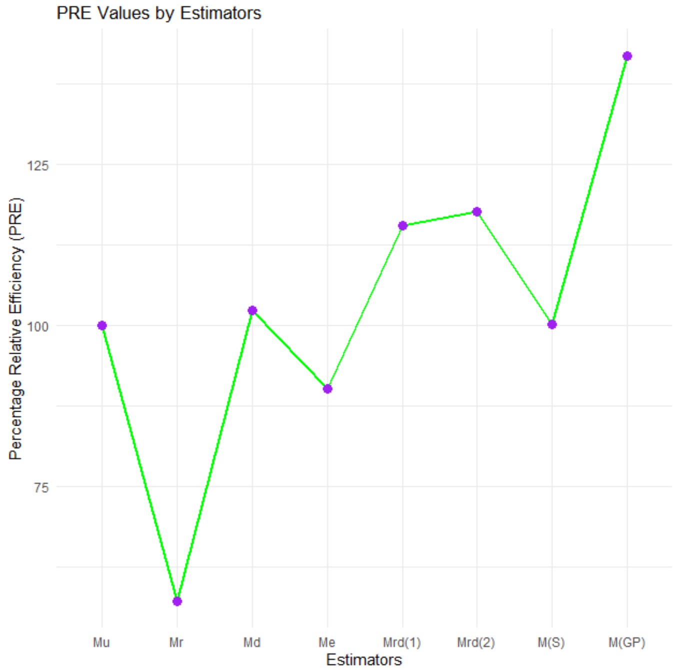

Figure 2: PRE of median estimators using population 1

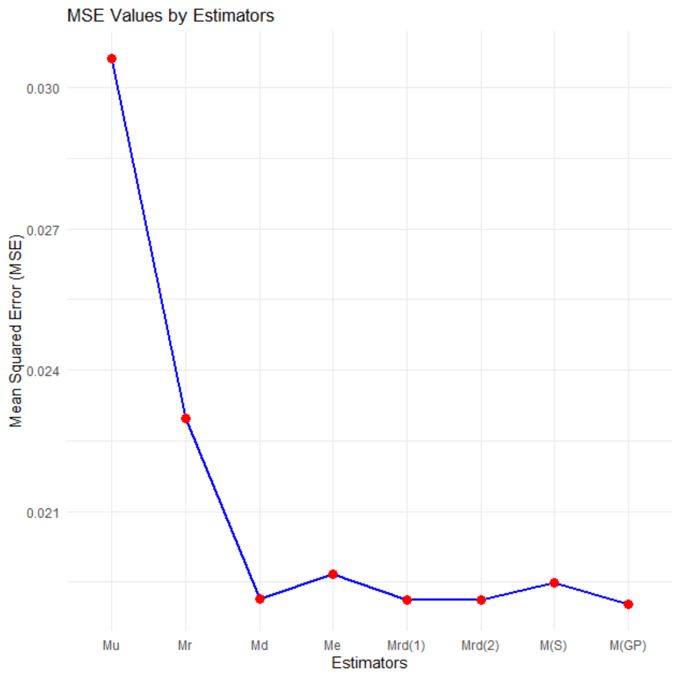

Figure 3: MSE of median estimators using population 2

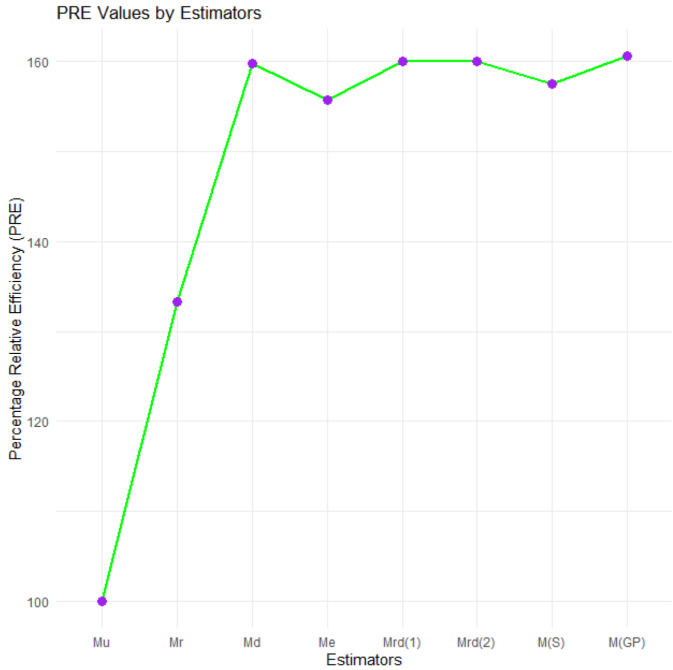

Figure 4: PRE of median estimators using population 2

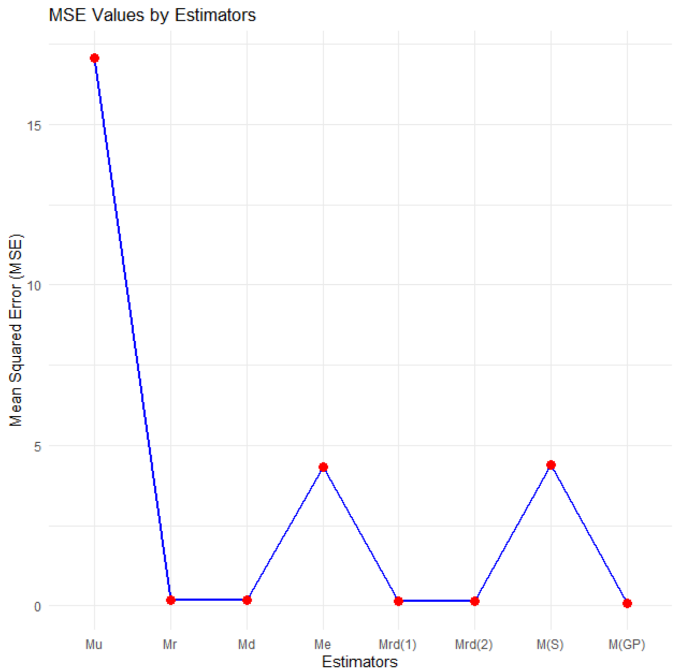

Figure 5: MSE of median estimators using population 3

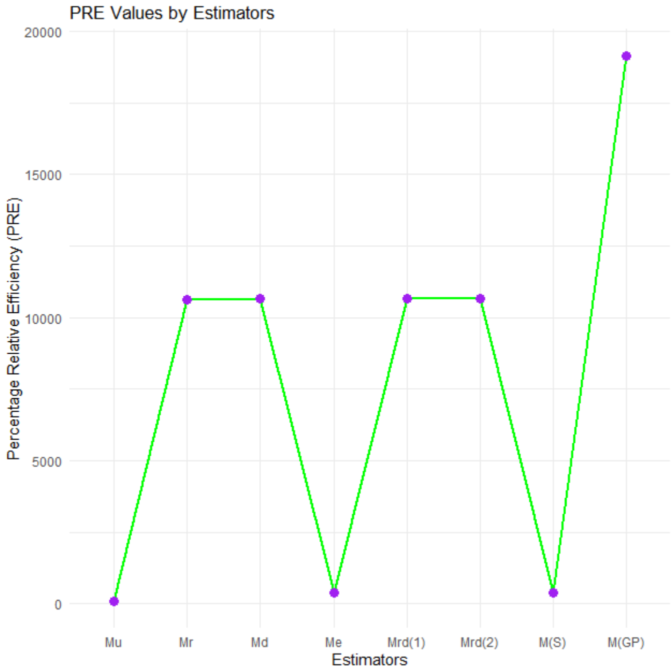

Figure 6: PRE of median estimators using population 3

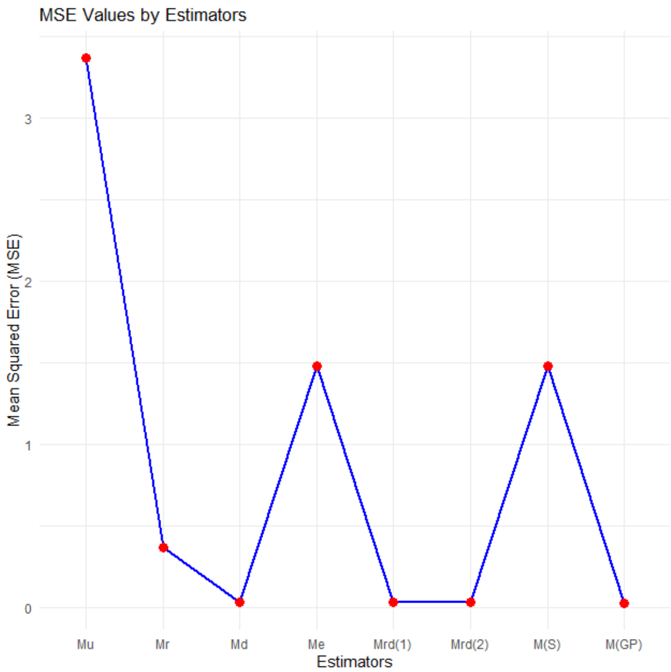

Figure 7: MSE of median estimators using population 4

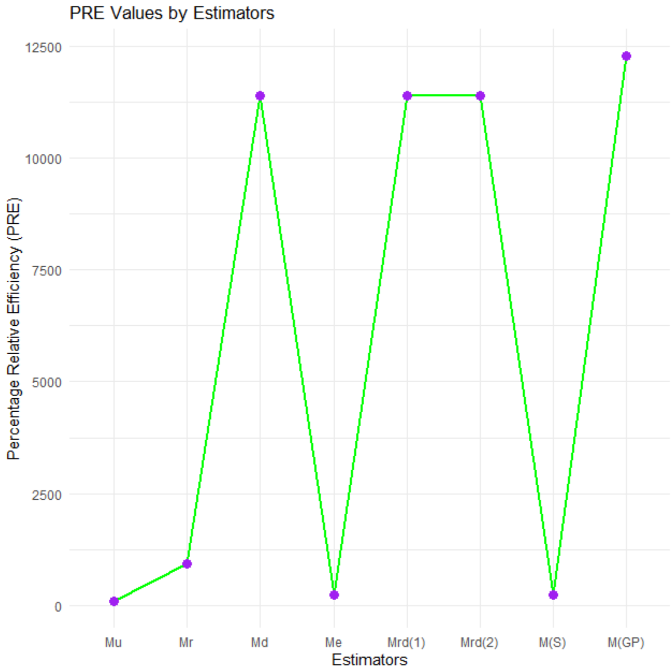

Figure 8: PRE of median estimators using population 4

6. Conclusion

In this article, we have suggested an improved generalized class of estimator for estimation of population median under simple random sampling. We examined the suggested estimator's bias and MSE up to the first approximation order. On four distinct datasets, the suggested generalized class of estimators is contrasted with existing estimators. Table 1 contains the members of the suggested class of estimators using different choices of α and β. Tables 2–9 show the mean squared error (MSE) and the PRE for the existing and the suggested class of estimators for four real data sets, respectively. The MSEs of various suggested and competing estimators for real Data sets I and IV are depicted graphically in Figures 1, 3, 5, and 7. The Data sets I to IV are shown graphically in Figures 2, 4, 6, and 8, which show the PREs of several suggested and competing estimators. When compared to other estimators, the proposed one has the best practical efficiency based on the results of real data sets, with the lowest MSE and maximum PRE. Using auxiliary information for population mean, distribution function and population proportion are discussed in detail by [21-24]. This work can be easily extended to estimate population median under systematic random sampling. Further extension of the current work is to develop improved estimators for estimation of median using neutrosophic data.
This work contributes its main innovation through the creation of a generalized class of estimators for median estimation which used auxiliary information effectively. This estimation method differs from standard median methods because it uses observed data and external information sources while related variables contribute to enhanced estimation precision. This work presents an adaptable framework that uses auxiliary data to deliver improved median estimations throughout diverse complex situations when observed data exists with defects or incompleteness. The paper thoroughly examines the statistical properties of these new estimators including unbiasedness, consistency along with efficiency in order to prove their theoretical validity. These estimators find practical applications across economics together with healthcare and social sciences to support crucial decision-making processes whenever accurate median estimation becomes essential. These findings establish better statistical inference approaches that demonstrate substantial progress in robust estimation research.

Competing interests

The authors declare no competing interests.

References

[1] Aamir, M., Shabri, A., & Ishaq, M. (2018). Improving forecasting accuracy of crude oil prices using decomposition ensemble model with reconstruction of IMFs based on ARIMA model. *Malaysian Journal of Fundamental and Applied Sciences*, *14*, 471-483.

[2] Aczel, A. D. (1996). *Complete business statistics*. Irwin Professional Publishing.

[3] Aladag, S., & Cingi, H. (2015). Improvement in estimating the population median in simple random sampling and stratified random sampling using auxiliary information. *Communications in Statistics-Theory and Methods*, *44*, 1013-1032.

[4] Bedi, P. K. (1996). Efficient utilization of auxiliary information at estimation stage. *Biometrical Journal*, *38*, 973-976.

[5] Baig, A., Masood, S., & Ahmed Tarray, T. (2020). Improved class of difference-type estimators for population median in survey sampling. *Communications in Statistics-Theory and Methods*, *49*, 5778-5793.

[6] Bandyopadhyay, A., Singh, G. N., & Das, P. (2016). Estimation of population median in presence of non-response under two-phase sampling. *Sri Lankan Journal of Applied Statistics*, *17*.

[7] Enang, E. I., Etuk, S. I., Ekpenyong, E. J., & Akpan, V. M. (2016). An alternative exponential estimator of population median. *International Journal of Statistics and Economics*, *17*, 85-97.

[8] Gupta, S., Shabbir, J., & Ahmad, S. (2008). Estimation of median in two-phase sampling using two auxiliary variables. *Communications in Statistics-Theory and Methods*, *37*, 18151822.

[9] Gross, S. (1980). Median estimation in sample surveys. In *Proceedings of the Section on Survey Research Methods* (Vol. 1814184). Alexandria, VA: American Statistical Association.

[10] Jhajj, H. S., Kaur, H., & Walia, G. (2014). Efficient family of ratio-product type estimators of median. *Model Assisted Statistics and Applications*, *9*, 277-282.

[11] Irfan, M., Javed, M., Shongwe, S. C., Zohaib, M., & Haider Bhatti, S. (2021). Estimation of population median under robust measures of an auxiliary variable. *Mathematical problems in Engineering*, *2021*, 1-14.

[12] Kuk, A. Y., & Mak, T. K. (1989). Median estimation in the presence of auxiliary information. *Journal of the Royal Statistical Society: Series B (Methodological)*, *51*, 261-269.

[13] Rao, T. J. (1991). On certail methods of improving ration and regression estimators. *Communications in Statistics-Theory and Methods*, *20*, 3325-3340.

[14] Singh, S. (2003). *Advanced Sampling Theory With Applications: How Michael"" Selected"" Amy* (Vol. 2). Springer Science & Business Media.

[15] Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2007). Improvement in estimating the population mean using exponential estimator in simple random sampling. *Auxiliary Information and a priori Values in Construction of Improved Estimators*, *33*.

[16] Shabbir, J., Gupta, S., & Hussain, Z. (2015). Improved estimation of finite population median under two-phase sampling when using two auxiliary variables. *Scientia Iranica*, *22*, 1271-1277.

[17] Ahmad, S., Hussain, S., & Ahmad, S. (2021). Finite population distribution function estimation using auxiliary information under simple random sampling. *Statistics, Computing and Interdisciplinary Research*, *3*, 29-38.

[18] Subzar, M., Lone, S. A., Ekpenyong, E. J., Salam, A., Aslam, M., Raja, T. A., & Almutlak, S. A. (2023). Efficient class of ratio cum median estimators for estimating the population median. *Plos one*, *18*, e0274690.

[19] Shabbir, J., & Gupta, S. (2017). A generalized class of di fference type estimators for population median in survey sampling. *Hacettepe Journal of Mathematics and Statistics*, *46*, 1015-1028.

[20] Hussain, I., Qureshi, M., Ismail, M., Iftikhar, H., Zywiołek, J., & López-Gonzales, J. L. (2024). Optimal features selection in the high dimensional data based on robust technique: Application to different health database. *Heliyon*, *10*.

[21] Ahmad, S., Qureshi, M., Iftikhar, H., Rodrigues, P. C., & Rehman, M. Z. (2025). An improved family of unbiased ratio estimators for a population distribution function. *AIMS Mathematics*, *10*, 1061-1084.

[22] Kumar, A., & Siddiqui, A. S. (2024). Enhanced estimation of population mean using simple random sampling. *Research in Statistics*, *2*, 2335949.

[23] Bhushan, S., & Kumar, A. (2023). Evaluating the performance of logarithmic type estimators using auxiliary attribute. *Life Cycle Reliability and Safety Engineering*, *12*, 285-292.

[24] Bahl, S and Tuteja, R. (1991). Ratio and Product type exponential estimators. *Journal of Information of Optimization Sciences,* 12: 159-164.