Improved estimation of finite population median under two-phase sampling when using two auxiliary variables

J. Shabbir\textsuperscript{a,*}, S. Gupta\textsuperscript{b} and Z. Hussain\textsuperscript{a}

\textsuperscript{a} Department of Statistics, Quaid-i-Azam University, Islamabad {\textsuperscript{15320}, Pakistan.}
\textsuperscript{b} Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro, NC 27412, USA.

Received 14 December 2013; received in revised form 21 April 2014; accepted 2 September 2014

KEYWORDS
Auxiliary variables; Mean Squared Error (MSE); Median; Efficiency.


© 2015 Sharif University of Technology. All rights reserved.

1. Introduction

Several authors including Kadilar and Cingi [1,2], Shabbir and Gupta [3], Koyuncu and Kadilar [4,5] and Diana et al. [6] have developed estimators for the finite population mean under different sampling schemes. However lesser degree of attention has been paid to estimation of population median. In many situations, median is a more appropriate measure of location than mean, particularly when the variable of interest follows a highly skewed distribution. Common examples of such variables are salaries, expenditure, and production quality. Kuk and Mak [7] introduced median ratio estimator that makes use of the auxiliary information. Singh et al. [8] suggested an estimator for population median under two-phase sampling scheme using two auxiliary variables. Gupta et al. [9] have suggested a class of estimators for population median using two auxiliary variables. Singh et al. [8] and Gupta et al. [9] estimators are equally efficient in the sense of MSE, but Gupta et al. [9] estimator is generally preferable because of its lower bias in most situations. Al and Cingi [10] and Singh and Solanki [11] introduced some classes of median estimators when using single auxiliary variable. In this paper, we consider a problem of median estimation and propose an estimator that makes use of two auxiliary variables under two-phase sampling scheme.

Consider a finite population with $N$ units. Let $y_i$, $x_i$ and $z_i$ ($i = 1, 2, \ldots, N$) be the values on the $i$th population unit for the study variable $y$ and two auxiliary variables $x$ and $z$, respectively. Also let $y_i$, $x_i$ and $z_i$ ($i = 1, 2, \ldots, n$) be the values on the
with population unit included in the sample of size $n$ drawn by simple random sampling without replacement (SRSWOR). Let $M_y$, $M_x$ and $M_z$ respectively, be the unknown population medians and $\bar{M}_y$, $\bar{M}_x$ and $\bar{M}_z$ be the sample medians for $y$, $x$ and $z$, respectively. When population median of the auxiliary variable is not known, we draw a preliminary large sample of size $n'$ according to SRSWOR (i.e. $n'<N$) and compute $\bar{M}_y$, $\bar{M}_x$ and $\bar{M}_z$, the sample medians of the study variable and the two auxiliary variables respectively. Further, we draw a subsample of size $n$ from the initial sample of size $n'$ (i.e. $n<n'$) by SRSWOR and compute $\bar{M}_y$, $\bar{M}_x$ and $\bar{M}_z$. Let $y_{(1)} \leq y_{(2)} \leq \cdots \leq y_{(n)}$ be the ordered sample values for the study variable $y$. Let $t$ be an integer, such that $y_{(t)} \leq M_y \leq y_{(t+1)}$ and $p = t/n$ be the proportion of the $y$ values that are less than or equal to $M_y$. If $Q_y(t)$ denotes the $t$th quantile of $Y$ then $\bar{M}_y = Q_y(0.5)$. Kuk and Mak [7] introduced the following matrix of proportion $p_{ij}(x, y)$ in Table 1.

Similarly, we can define the matrices of proportions $p_{ij}(x, z)$ and $p_{ij}(y, z)$. It is assumed that as $N \rightarrow \infty$, the distribution of the trivariate variables $(x, y, z)$ approaches a continuous distribution with marginal densities $f_x(x)$, $f_y(y)$ and $f_z(z)$ of $x$, $y$ and $z$, respectively. Let $p_{xy} = 4p_{11}(x, y) - 1$, $p_{yz} = 4p_{11}(y, z) - 1$ and $p_{xz} = 4p_{11}(x, z) - 1$ be the population proportion coefficients between variables indicated by the respective subscripts. Let $e_0 = (\bar{M}_y - M_y)/M_y$, $e'_0 = (\bar{M}_y - M_y)/M_y$, $e_1 = (\bar{M}_x - M_x)/M_x$, $e'_1 = (\bar{M}_x - M_x)/M_x$, $e_2 = (\bar{M}_z - M_z)/M_z$ and $e'_2 = (\bar{M}_z - M_z)/M_z$ such that $E(e_i) = E(e'_i) = 0$, $i = 0, 1, 2$.

The following expected values are correct to first degree of approximation (see [12]).

\[
E(e''_0) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4(M_y f_y(M_y))^2},
\]
\[
E(e''_1) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4(M_x f_x(M_x))^2},
\]
\[
E(e''_2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4(M_z f_z(M_z))^2},
\]
\[
E(e''_0) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4(M_y f_y(M_y))^2}.
\]

Table 1. A matrix of proportions $p_{ij}(x, y)$.

<table>
<thead>
<tr>
<th>$y \leq M_y$</th>
<th>$y &gt; M_y$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x \leq M_x$</td>
<td>$p_{11}(x, y)$</td>
<td>$p_{21}(x, y)$</td>
</tr>
<tr>
<td>$x &gt; M_x$</td>
<td>$p_{12}(x, y)$</td>
<td>$p_{22}(x, y)$</td>
</tr>
<tr>
<td>Total</td>
<td>$p_{1}(x, y)$</td>
<td>$p_{2}(x, y)$</td>
</tr>
</tbody>
</table>

2. Some existing estimators

In this section, we discuss some of the existing estimators of population median ($\bar{M}_y$).

The variance of the usual sample median estimator ($\bar{M}_y$) by Gross [13] is given by:

\[
\text{Var}(\bar{M}_y) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4(f_y(M_y))^2}.
\]  

(1)

Chand [14] suggested the chain-ratio type estimator for population median ($\bar{M}_y$) under two-phase sampling. It is given by:

\[
\hat{M}_R = \bar{M}_y \left(\frac{\bar{M}_x}{\bar{M}_y} \right) \left(\frac{M_z}{\bar{M}_z} \right).
\]

(2)

\[
E(e''_0) = E(e'_0 e'_1) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4(M_x f_x(M_x))^2}.
\]
where $M_z$ is known. The MSE of $\hat{M}_R$, to first order of the approximation, is given by:

$$\text{MSE} \left( \hat{M}_R \right) = \frac{1}{4\{f_y(M_y)\}^2} \left[ \left( \frac{1}{n} - \frac{1}{N} \right) \right]$$

$$+ \left( \frac{1}{n} - \frac{1}{n'} \right) \frac{M_y f_y(M_y)}{M_x f_x(M_x)} \left( \frac{M_y f_y(M_y)}{M_x f_x(M_x)} - 2 \rho_{yx} \right)$$

$$+ \left( \frac{1}{n} - \frac{1}{N} \right) \frac{M_y f_y(M_y)}{M_z f_z(M_z)} \left( \frac{M_y f_y(M_y)}{M_z f_z(M_z)} - 2 \rho_{yz} \right) \right]$$

(3)

Srivastava et al. [15] suggested the following power-chain-ratio type estimator:

$$\hat{M}_{SR} = \hat{M}_y \left( \frac{\hat{M}_y}{\hat{M}_x} \right)^{\alpha_1} \left( \frac{M_z}{\hat{M}_z} \right)^{\alpha_2},$$

(4)

where $\alpha_1$ and $\alpha_2$ are constants. The minimum MSE of $\hat{M}_{SR}$, to first order of the approximation, at optimum values of $\alpha_1$ and $\alpha_2$, i.e.:

$$\alpha_{1\text{opt}} = \rho_{yx} \frac{M_x f_x(M_x)}{M_y f_y(M_y)},$$

and:

$$\alpha_{2\text{opt}} = \rho_{yz} \frac{M_y f_y(M_y)}{M_z f_z(M_z)},$$

is given by:

$$\text{MSE} \left( \hat{M}_{SR} \right)_{\text{min}} \approx \frac{1}{4\{f_y(M_y)\}^2}$$

$$\left[ \left( \frac{1}{n} - \frac{1}{N} \right) - \left( \frac{1}{n'} - \frac{1}{n} \right) \rho_{yx}^2 - \left( \frac{1}{n} - \frac{1}{N} \right) \rho_{yz}^2 \right] \right]$$

(5)

which is equal to the variance of the difference estimator:

$$\hat{M}_D = \hat{M}_y + d_1 \left( \hat{M}_x - \hat{M}_z \right)$$

$$+ d_2 \left( M_z - \hat{M}_z \right),$$

where $d_1$ and $d_2$ are constants.

Gupta et al. [9] suggested the following estimator by utilizing the range of the second auxiliary variable ($z$), i.e. $R_z$ as:

$$\hat{M}_D = \hat{M}_y \left( \frac{\hat{M}_y}{\hat{M}_x} \right)^{\gamma_1} \left( \frac{M_z + R_z}{\hat{M}_z + R_z} \right)^{\gamma_2} \left( \frac{M_z + R_z}{\hat{M}_z + R_z} \right)^{\gamma_3},$$

(6)

where $\gamma_i$ ($i = 1, 2, 3$) are constants. The minimum MSE of $\hat{M}_D$, to first order of the approximation, at optimum values of $\gamma_i$ ($i = 1, 2, 3$), i.e.:

$$\gamma_{1\text{opt}} = \frac{M_x f_x(M_x)}{M_y f_y(M_y)} \left( \frac{\rho_{yx} \rho_{xz} - \rho_{yx}}{\rho_{xz}^2 - 1} \right),$$

$$\gamma_{2\text{opt}} = \frac{M_z f_z(M_z)}{M_y f_y(M_y)} \left( \frac{\rho_{yz} \rho_{xz} - \rho_{yz}}{\rho_{xz}^2 - 1} \right),$$

and:

$$\gamma_{3\text{opt}} = \frac{M_z f_z(M_z)}{M_y f_y(M_y)} \left( \frac{\rho_{yz} \rho_{xz} - \rho_{yz}}{\rho_{xz}^2 - 1} \right),$$

for:

$$g = \frac{M_z}{M_z + R_z}$$

is given by:

$$\text{MSE}(\hat{M}_D)_{\text{min}} \approx \frac{1}{4\{f_y(M_y)\}^2}$$

$$\left[ \left( \frac{1}{n} - \frac{1}{N} \right) - \left( \frac{1}{n'} - \frac{1}{n} \right) \rho_{yx}^2 - \left( \frac{1}{n} - \frac{1}{N} \right) \rho_{yz}^2 \right] \right]$$

(7)

where:

$$R^2_{y,x} = \frac{\rho_{yx}^2 + \rho_{yz}^2 - 2 \rho_{yx} \rho_{yz} \rho_{xz}}{1 - \rho_{xz}^2}.$$

The expression in Relation (7) is equal to minimum MSE of Singh et al. [8] estimator, given by:

$$\hat{M}_S = \hat{M}_y \left( \frac{\hat{M}_y}{\hat{M}_x} \right)^{\lambda_1} \left( \frac{M_z}{\hat{M}_z} \right)^{\lambda_2} \left( \frac{M_z}{\hat{M}_z} \right)^{\lambda_3},$$

where $\lambda_i$ ($i = 1, 2, 3$) are constants.

3. Proposed estimator

Jhajj and Walia [16] suggested the following estimator for population mean under two-phase sampling when using the single auxiliary variable:

$$\bar{y}_{JW} = \left[ g + \theta (\bar{g} - \bar{y}) \right] \left[ \frac{\bar{g}'}{\bar{g} + \theta (\bar{g} - \bar{y})} \right]^\alpha,$$

where $\theta$ and $\alpha$ are constants.

Diana [17] introduced a family of estimators for the population mean in stratified sampling given by:

$$\bar{y}_{D} = \bar{y}_{St} \left( \frac{\bar{x}}{X} \right)^\delta \left[ d + (1 - d) \left( \frac{\bar{x}}{X} \right)^\gamma \right]^\eta,$$

where $\delta$, $\varepsilon$, $\eta$ and $d$ are constants. By using these four parameters one can generate many estimators.

On the lines of Jhajj and Walia [16] and Diana [17], we propose a generalized difference-cum-ratio type estimator for population median under two phase sampling scheme. The proposed estimator is given by:

$$\hat{M}_P = \left[ \hat{M}_y + \psi \left( \hat{M}_y - \hat{M}_y \right) \right] \left[ \psi + (1 - \psi) \left( \frac{M_z}{\hat{M}_z} \right)^{\lambda_1} \right]$$
\[
\psi + (1 - \psi) \frac{\hat{M}_1}{\hat{M}_2} \] 
\[
\psi' + (1 - \psi') \frac{\hat{M}'_1}{\hat{M}'_2}
\]
where \( \psi \) and \( w_i \) (\( i = 1, 2, 3 \)) are constants.

The proposed estimator in Eq. (8) is different from the Gupta et al. [9] estimator given in Eq. (6), in the sense that the former given in Eq. (8), we measured \( \hat{M}'_y \), \( \hat{M}'_x \) and \( \hat{M}'_z \) at first phase, whereas in the latter, in Eq. (6), we measured \( \hat{M}_y \), \( \hat{M}_x \) and \( \hat{M}_z \) at the first phase but at the second phase we measured \( \hat{M}_y \), \( \hat{M}_x \) and \( \hat{M}_z \). This idea is discussed in detail by Jhaajj and Walla [16] in estimating the finite population variance.

Solving Eq. (8) in terms of \( \epsilon' \)s to the first order of approximation, we have:
\[
\hat{M}_p = M_y[1 + e_0 + \psi(\epsilon'_0 - e_0)]
\]
\[
[1 + w_1(1 - \psi)\{(e_1 - e'_1) + \epsilon'^2 - \epsilon_1\epsilon'_1\} + \frac{w_1(w_1 + 1)}{2}(1 - \psi)^2(e_1 - \epsilon'_1)^2] 
\]
\[
1 + w_2(1 - \psi)\{(e_2 - e'_2) + \epsilon'^2 - \epsilon_2\epsilon'_2\} + \frac{w_2(w_2 + 1)}{2}(1 - \psi)^2(e_2 - \epsilon'_2)^2 
\]
\[
1 + w_3(1 - \psi)e_2' + \frac{w_3(w_3 + 1)}{2}(1 - \psi)^2e_2'^2
\]
Hence, up to the first order of approximation:
\[
MSE(\hat{M}_p) \approx M'_y E[\epsilon_0 + \psi(\epsilon'_0 - e_0)] + w_1(1 - \psi)(e_1 - \epsilon'_1) + w_2(1 - \psi)(e_2 - \epsilon'_2) + w_3(1 - \psi)e_2'.
\]
Squaring and taking expectations, the MSE of \( \hat{M}_p \), to the first degree of approximation, is given by:
\[
MSE(\hat{M}_p) \approx \frac{1}{4\{f_y(M_y)\}^2} \left[ \left( \frac{1}{n} - \frac{1}{N} \right) \right] + (1 - \psi)^2 \left( \frac{1}{n} - \frac{1}{n'} \right) 
\]
\[
+ \left( \frac{1}{n'} - \frac{1}{N} \right) \left\{ (1 - \psi)^2 w_3^2 \left( \frac{M_y f_y(M_y)}{M_z f_z(M_z)} \right)^2 \right\} + 2(1 - \psi)w_3 \left( \frac{M_y f_y(M_y)}{M_z f_z(M_z)} \right) \rho_{yz}. \]
\]
\[
(9)
\]
Setting \( \frac{\partial MSE(\hat{M}_p)}{\partial w_i} = 0 \) (\( i = 1, 2, 3 \)), we have:
\[
w_{1(\text{opt})} = \frac{M_z f_z(M_z)\rho_{yx} + \rho_{yz}}{M_y f_y(M_y)(1 - \rho_{yz}^2)} ,
\]
\[
w_{2(\text{opt})} = \frac{M_z f_z(M_z)\rho_{yx} + \rho_{yz}}{M_y f_y(M_y)(1 - \rho_{yz}^2)},
\]
and:
\[
w_{3(\text{opt})} = -\frac{M_z f_z(M_z)\rho_{yz}}{M_y f_y(M_y)(1 - \psi)}
\]
Substituting the optimum values of \( w_i \) (\( i = 1, 2, 3 \)) in Relation (9), we get the minimum MSE of \( \hat{M}_p \), given by:
\[
MSE(\hat{M}_p)_{\text{min}} \approx \frac{1}{4\{f_y(M_y)\}^2} \left[ \left( \frac{1}{n} - \frac{1}{N} \right)(1 - \rho_{yz}^2) + (1 - \psi)^2 \left( \frac{1}{n} - \frac{1}{n'} \right)(1 - \rho_{yz}^2) \right].
\]
(10)
Note that Jhaajj and Walla [16] have shown that MSE is minimum for \( \psi = 1 \). So, further minimizing Relation (10) with respect to \( \psi \) (i.e. taking \( \psi = 1 \)), we have:
\[
MSE(\hat{M}_p)_{\psi = 1}^{\text{min}} \approx \frac{1}{4\{f_y(M_y)\}^2} \left[ \left( \frac{1}{n} - \frac{1}{N} \right)(1 - \rho_{yz}^2) \right].
\]
(11)
In Tables 2 and 3, MSE values and Percent Relative Efficiency (PRE) are given for different values of \( \psi \), i.e. 0, 0.5, 1, 1.5, 2. For \( \psi = 1 \), the proposed estimator \( \hat{M}_p \) performs well.
Table 2. MSE values of different estimators with respect to $\hat{M}_y$ for different values of $\psi$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_y$</td>
<td>565443.57</td>
<td>2.22</td>
<td>113343.27</td>
</tr>
<tr>
<td>$M_R$</td>
<td>840264.22</td>
<td>1.01</td>
<td>180840.61</td>
</tr>
<tr>
<td>$M_{SR}$</td>
<td>525744.59</td>
<td>0.87</td>
<td>110225.37</td>
</tr>
<tr>
<td>$M_G$</td>
<td>506293.76</td>
<td>0.57</td>
<td>109805.56</td>
</tr>
<tr>
<td>$\hat{M}_P$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi = 0$</td>
<td>506293.76</td>
<td>0.57</td>
<td>109905.56</td>
</tr>
<tr>
<td>$\psi = 0.5$</td>
<td>360471.28</td>
<td>0.38</td>
<td>75308.90</td>
</tr>
<tr>
<td>$\psi = 1$</td>
<td>311863.78</td>
<td>0.31</td>
<td>63810.01</td>
</tr>
<tr>
<td>$\psi = 1.5$</td>
<td>360471.28</td>
<td>0.38</td>
<td>75308.90</td>
</tr>
<tr>
<td>$\psi = 2$</td>
<td>506293.76</td>
<td>0.57</td>
<td>109905.56</td>
</tr>
</tbody>
</table>

Table 3. PRE of different estimators with respect to $\hat{M}_y$ for different values of $\psi$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_y$</td>
<td>100.000</td>
<td>100.000</td>
<td>100.000</td>
</tr>
<tr>
<td>$M_R$</td>
<td>67.294</td>
<td>220.004</td>
<td>62.676</td>
</tr>
<tr>
<td>$M_{SR}$</td>
<td>107.551</td>
<td>254.494</td>
<td>102.829</td>
</tr>
<tr>
<td>$M_G$</td>
<td>111683</td>
<td>390.314</td>
<td>103222</td>
</tr>
<tr>
<td>$\hat{M}_P$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi = 0$</td>
<td>111683</td>
<td>390.314</td>
<td>103222</td>
</tr>
<tr>
<td>$\psi = 0.5$</td>
<td>156862</td>
<td>587.840</td>
<td>150504</td>
</tr>
<tr>
<td>$\psi = 1$</td>
<td>181311</td>
<td>707.124</td>
<td>177626</td>
</tr>
<tr>
<td>$\psi = 1.5$</td>
<td>156862</td>
<td>587.840</td>
<td>150504</td>
</tr>
<tr>
<td>$\psi = 2$</td>
<td>111683</td>
<td>390.314</td>
<td>103222</td>
</tr>
</tbody>
</table>

4. Efficiency comparison

In this section, we compare the proposed estimator $\hat{M}_P$ with other existing estimators.

**Condition (i)**

By Relations (1) and (10), $\text{MSE}(\hat{M}_P)_{\text{min}} < \text{Var}(\hat{M}_y)$ if:

$$\left(\frac{1}{m'} - \frac{1}{N}\right) \rho_{yx}^2 + \left(\frac{1}{n} - \frac{1}{n'}\right) \left\{ 1 - (1 - \psi)^2 (1 - R^2_{y,xt}) \right\} > 0.$$

**Condition (ii)**

By Relations (3) and (10), $\text{MSE}(\hat{M}_P)_{\text{min}} < \text{MSE}(\hat{M}_R)$ if:

$$\left(\frac{1}{n} - \frac{1}{n'}\right) \left[ \left(\frac{M_y f_y(M_y)}{M_x f_x(M_x)} - \rho_{yx}\right)^2 + \left(\frac{1}{n'} - \frac{1}{N}\right) \left(\frac{M_y f_y(M_y)}{M_x f_x(M_x)} - \rho_{yx}\right)^2 \right]$$

$$- (1 - \psi)^2 (1 - R^2_{y,xt}) > 0.$$

**Condition (iii)**

By Relations (5) and (10):

$$\text{MSE}(\hat{M}_P)_{\text{min}} < \text{MSE}(\hat{M}_{SR})_{\text{min}}$$

if:

$$\left(\frac{1}{n} - \frac{1}{n'}\right) \left[ (1 - \rho_{yx})^2 - (1 - \psi)^2 (1 - R^2_{y,xt}) \right] > 0.$$

**Condition (iv)**

By Relations (7) and (10):

$$\text{MSE}(\hat{M}_P)_{\text{min}} < \text{MSE}(\hat{M}_G)_{\text{min}}$$

if:

$$\left(\frac{1}{n} - \frac{1}{n'}\right) (1 - R^2_{y,xt}) \psi (2 - \psi) > 0.$$

Conditions in Comparisons (i)-(iv) will always be true for $\psi = 1$, and our proposed estimator will perform better than the estimators $M_i$ ($i = y, R, SR, G$), as seen in Table 2.
5. Empirical study

In this section, we consider three populations to perform numerical comparisons of different estimators.

Population 1: Source: Singh [18]
Let Y, X, Z, respectively, be the number of fish caught by the marine recreational fishermen in year 1995, 1994 and 1993. The descriptive statistics are given below:

\[ N = 69, \quad n' = 24, \quad n = 17, \]
\[ M_Y = 2068, \quad M_X = 2111, \quad M_z = 2307, \]
\[ f_Y(M_Y) = 0.00014, \quad f_X(M_X) = 0.00014, \]
\[ f_z(M_z) = 0.00013, \quad \rho_{yx} = 0.1505, \]
\[ \rho_y = 0.3166, \quad \rho_{xz} = 0.1431. \]

Population 2: Source: Aczel and Sounderpandian [19]
Let Y be the US exports to Singapore in billions of Singapore dollars, X be the money supply figures in billions of Singapore dollars and Z be the local prices in US dollars.

The descriptive statistics are given below:

\[ N = 67, \quad n' = 23, \quad n = 15, \]
\[ M_Y = 4.8, \quad M_X = 7, \quad M_z = 151, \]
\[ f_Y(M_Y) = 0.0763, \quad f_X(M_X) = 0.0526, \]
\[ f_z(M_z) = 0.00014, \quad \rho_{yx} = 0.6624, \]
\[ \rho_y = 0.8624, \quad \rho_{xz} = 0.7592. \]

Population 3: Source: MFA [20]
Let Y, X, Z, respectively, represent the district-wise tomato production (tonnes) in Pakistan in year 2003, 2002 and 2001.

The descriptive statistics obtained from the population are given below:

\[ N = 97, \quad n' = 46, \quad n = 33, \]
\[ M_Y = 1242, \quad M_X = 1233, \quad M_z = 1207, \]
\[ f_Y(M_Y) = 0.00021, \quad f_X(M_X) = 0.00022, \]
\[ f_z(M_z) = 0.00023, \quad \rho_{yx} = 0.2096, \]
\[ \rho_y = 0.1233, \quad \rho_{xz} = 0.1496. \]

We use the following expression to obtain the Percent Relative Efficiency (PRE) as:

\[
PRE = \frac{\text{Var}(\hat{M}_y)}{\text{MSE}(\hat{M}_i)} \times 100, \quad i = y, R, SR, G, P.
\]

The MSE values and percent relative efficiencies are given in Tables 2 and 3, respectively.

The estimators \( M_i \) \((i = y, R, SR, G, P)\) are independent of \( \psi \). Based on the results in Tables 2 and 3, it is observed that the proposed estimator \( \hat{M}_P \) outperforms other competing estimators for different values of \( \psi \). The ratio estimator \( \hat{M}_R \) shows poor performances in Populations 1 and 3 because of weaker correlation between the study variable and auxiliary variables.

Although Jhajj and Walia [16] have presented results for various values of \( \psi \), their numerical results clearly show that optimal value of \( \psi \) is 1, a fact observed in this study as well.

6. Conclusion

We propose an improved estimator for population median on the lines of Jhajj and Walia [16] and Diana [17]. Both theoretical and numerical comparisons with other estimators show that the proposed estimator \( \hat{M}_P \) is more efficient than sample median estimator \( \hat{M}_y \), ratio estimator \( \hat{M}_R \), Srivastava et al. estimator [15] \( \hat{M}_{SR} \) and Gupta et al. estimator [9] \( \hat{M}_G \) for \( 0 < \psi < 2 \). For \( \psi = 0, 2 \), estimators \( \hat{M}_P \) and \( \hat{M}_G \) are equally efficient. Among different values of \( \psi \), maximum gain in precision occurs at \( \psi = 1 \).

Acknowledgments

Authors are thankful to the referees for their valuable remarks which helped the research paper improve.

References


Biographies

Javid Shabbir obtained his PhD degree in Statistics from the University of Kent, UK. He has published more than 70 research papers in international research journals. His research interests include: Survey sampling and randomized response models.

Sat Gupta is a professor of statistics at University of North Carolina-Greensboro, USA. He has earned PhD degrees in Mathematics (University of Delhi) and Statistics (Colorado State University). He has over 90 refereed publications and 4 edited book volumes, and has guided research students at all levels of the curriculum including undergraduate and PhD students. He has received external funding from United States agencies like the National Science Foundation (NSF), the National Institute of Health (NIH), and the Mathematical Association of America (MAA).

Zawar Hussain graduated from Quaid-i-Azam University, Islamabad, Pakistan. He has published around 30 research papers in reputed journals. His research interest include: Sampling techniques, randomized response models, and quality control.