Scientia Iranica E (2015) 22(3), 1271–1277

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
www.scientiairanica.com

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

J. Shabbir[a,*], S. Gupta[b] and Z. Hussain[a]

a. *Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan.*
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.

Abstract. We propose an efficient estimator for population median under two-phase sampling when using two auxiliary variables on the lines of Diana [Diana, G. "A class of estimators of the population mean in stratified random sampling", *Statistica*, 53(1), pp. 59-66 (1993)] and Jhajj and Walia [Jhajj, H.S. and Walia, G.S. "A generalized difference-cum-ratio type estimator for the population variance in double sampling", *Communications in Statistics-Simulations and Computation*, 41, pp. 58-64 (2012)]. The expressions for mean squared errors are presented, correct to the first order of approximation. Both theoretical and numerical comparisons reveal that the proposed estimator performs better than the unbiased sample median estimator, ratio estimator, and estimators by Srivastava et al. [Srivastava, S.K., Rani, S., Khare, B.B., and Srivastava, S.R. "A generalized chain ratio estimator for mean of finite population", *Journal of the Indian Society of Agricultural Statistics*, 42(1), pp. 108-117 (1990)] and Gupta et al. [Gupta, S., Shabbir, J., Ahmad, S. "Estimation of median in two-phase sampling using two auxiliary variables", *Communications in Statistics- Theory and Methods*, 37(11), pp. 1815-1822 (2008)].

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 $(i = 1, 2, \cdots, N)$ be the values on the ith population unit for the study variable y and two auxiliary variables x and z, respectively. Also let y_i, x_i and z_i $(i = 1, 2, \cdots, n)$ be the values on the

*. *Corresponding author.*
E-mail addresses: js@qau.edu.pk (J. Shabbir); sngupta@uncg.edu (S. Gupta); zhlangh@yahoo.com (Z. Hussain)

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

	$y \leq M_y$	$y > M_y$	Total
$x \leq M_x$	$p_{11}(x,y)$	$p_{21}(x,y)$	$p_{.1}(x,y)$
$x > M_x$	$p_{12}(x,y)$	$p_{22}(x,y)$	$p_{.2}(x,y)$
Total	$p_{1.}(x,y)$	$p_{2.}(x,y)$	1

ith 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 $\hat{M}_y$, $\hat{M}_x$ and $\hat{M}_z$ 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 $\hat{M}'_y$, $\hat{M}'_x$ and $\hat{M}'_z$, 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 $\hat{M}_y$, $\hat{M}_x$ and $\hat{M}_z$. 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 tth quantile of Y then $\hat{M}_y = \hat{Q}_y(0.5)$. 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 $N \to \infty$, 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 $f_y(M_y)$, $f_x(M_x)$ and $f_z(M_z)$ be the probability density functions at M_y, M_x and M_z, respectively. Let $\rho_{yx} = 4p_{11}(x,y) - 1$, $\rho_{yz} = 4p_{11}(y,z) - 1$ and $\rho_{xz} = 4p_{11}(x,z) - 1$ be the population correlation coefficients between variables indicated by the respective subscripts. Let $e_0 = (\hat{M}_y - M_y)/M_y$, $e'_0 = (\hat{M}'_y - M_y)/M_y$, $e_1 = (\hat{M}_x - M_x)/M_x$, $e'_1 = (\hat{M}'_x - M_x)/M_x$, $e_2 = (\hat{M}_z - M_z)/M_z$ and $e'_2 = (\hat{M}'_z - M_z)/M_z$ 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]).

$$E(e_0^2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4\{M_y f_y(M_y)\}^2},$$

$$E(e_1^2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4\{M_x f_x(M_x)\}^2},$$

$$E(e_2^2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4\{M_z f_z(M_z)\}^2},$$

$$E({e'}_0^2) = \left(\frac{1}{n'} - \frac{1}{N}\right) \frac{1}{4\{M_y f_y(M_y)\}^2},$$

$$E({e'}_1^2) = E(e_1 e'_1) = \left(\frac{1}{n'} - \frac{1}{N}\right) \frac{1}{4\{M_x f_x(M_x)\}^2},$$

$$E({e'}_2^2) = \left(\frac{1}{n'} - \frac{1}{N}\right) \frac{1}{4\{M_z f_z(M_z)\}^2},$$

$$E(e_0 e_1) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{\{(4p_{11}(y,x) - 1\}\{M_x M_y f_x(M_x) f_y(M_y)\}},$$

$$E(e_0 e'_1) = E(e'_0 e_1) = E(e'_0 e'_1) = \left(\frac{1}{n'} - \frac{1}{N}\right) \frac{1}{\{(4p_{11}(y,x) - 1\}\{M_x M_y f_x(M_x) f_y(M_y)\}},$$

$$E(e_0 e_2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{\{(4p_{11}(y,z) - 1\}\{M_y M_z f_y(M_y) f_z(M_z)\}},$$

$$E(e_0 e'_2) = E(e'_0 e_2) = E(e'_0 e'_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) \frac{1}{\{(4p_{11}(y,z) - 1\}\{M_y M_z f_y(M_y) f_z(M_z)\}},$$

$$E(e_1 e_2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{\{(4p_{11}(x,z) - 1\}\{M_x M_z f_x(M_x) f_z(M_z)\}},$$

$$E(e'_1 e_2) = E(e_1 e'_2) = E(e'_1 e'_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) \frac{1}{\{(4p_{11}(x,z) - 1\}\{M_x M_z f_x(M_x) f_z(M_z)\}}.$$

2. Some existing estimators

In this section, we discuss some of the existing estimators of population median (M_y).

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

$$\mathrm{Var}\left(\hat{M}_y\right) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{4\{f_y(M_y)\}^2}. \tag{1}$$

Chand [14] suggested the chain-ratio type estimator for population median (M_y) under two-phase sampling. It is given by:

$$\hat{M}_R = \hat{M}_y \left(\frac{\hat{M}'_x}{\hat{M}_x}\right) \left(\frac{M_z}{\hat{M}'_z}\right), \tag{2}$$

where M_z is known. The MSE of $\hat{M}_R$, to first order of the approximation, is given by:

$$\mathrm{MSE}\left(\hat{M}_R\right) \cong \frac{1}{4\{f_y(M_y)\}^2}\left[\left(\frac{1}{n}-\frac{1}{N}\right)\right.$$
$$+\left(\frac{1}{n}-\frac{1}{n'}\right)\frac{M_y f_y(M_y)}{M_x f_x(M_x)}\left(\frac{M_y f_y(M_y)}{M_x f_x(M_x)}-2\rho_{yx}\right)$$
$$\left.+\left(\frac{1}{n'}-\frac{1}{N}\right)\frac{M_y f_y(M_y)}{M_z f_z(M_z)}\left(\frac{M_y f_y(M_y)}{M_z f_z(M_z)}-2\rho_{yz}\right)\right]. \quad (3)$$

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

$$\hat{M}_{\mathrm{SR}} = \hat{M}_y\left(\frac{\hat{M}'_x}{\hat{M}_x}\right)^{\alpha_1}\left(\frac{M_z}{\hat{M}'_z}\right)^{\alpha_2}, \quad (4)$$

where α_1 and α_2 are constants. The minimum MSE of $\hat{M}_{\mathrm{SR}}$, to first order of the approximation, at optimum values of α_1 and α_2, i.e.:

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

and:

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

is given by:

$$\mathrm{MSE}\left(\hat{M}_{\mathrm{SR}}\right)_{\min} \cong \frac{1}{4\{f_y(M_y)\}^2}$$
$$\left[\left(\frac{1}{n}-\frac{1}{N}\right)-\left(\frac{1}{n}-\frac{1}{n'}\right)\rho_{yx}^2-\left(\frac{1}{n'}-\frac{1}{N}\right)\rho_{yz}^2\right], \quad (5)$$

which is equal to the variance of the difference estimator:

$$\hat{M}_D = \hat{M}_y + d_1\left(\hat{M}'_x - \hat{M}_x\right) + d_2\left(M_z - \hat{M}'_z\right),$$

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:

$$\hat{M}_G = \hat{M}_y\left(\frac{\hat{M}'_x}{\hat{M}_x}\right)^{\gamma_1}\left(\frac{M_z+R_z}{\hat{M}'_z+R_z}\right)^{\gamma_2}\left(\frac{M_z+R_z}{\hat{M}'_z+R_z}\right)^{\gamma_3}, \quad (6)$$

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

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

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

and:

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

for:

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

is given by:

$$\mathrm{MSE}(\hat{M}_G)_{\min} \cong \frac{1}{4\{f_y(M_y)\}^2}$$
$$\left[\left(\frac{1}{n}-\frac{1}{N}\right)-\left(\frac{1}{n'}-\frac{1}{N}\right)\rho_{yz}^2-\left(\frac{1}{n}-\frac{1}{n'}\right)R_{y.xz}^2\right], \quad (7)$$

where:

$$R_{y.xz}^2 = \frac{\rho_{yx}^2+\rho_{yz}^2-2\rho_{yx}\rho_{yz}\rho_{xz}}{1-\rho_{xz}^2}.$$

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

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

where λ_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:

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

where θ and α are constants.

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

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

where δ, ε, η 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:

$$\hat{M}_P = \left[\hat{M}_y+\psi\left(\hat{M}'_y-\hat{M}_y\right)\right]\left[\psi+(1-\psi)\frac{\hat{M}_x}{\hat{M}'_x}\right]^{w_1}$$

$$\left[\psi+(1-\psi)\frac{\hat{M}_z}{\hat{M}'_z}\right]^{w_2}\left[\psi+(1-\psi)\frac{\hat{M}'_z}{M_z}\right]^{w_3}, \quad (8)$$

where ψ 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 in 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 $\hat{M}'_x$ 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 Jhajj and Walia [16] in estimating the finite population variance.

Solving Eq. (8) in terms of $e's$ to the first order of approximation, we have:

$$\begin{aligned}\hat{M}_P =& M_y[1+e_0+\psi(e'_0-e_0)]\\ &\left[1+w_1(1-\psi)\{(e_1-e'_1)+{e'}^2_1-e_1e'_1\}\right.\\ &\left.+\frac{w_1(w_1+1)}{2}(1-\psi)^2(e_1-e'_1)^2\right]\\ &\left[1+w_2(1-\psi)\{(e_2-e'_2)+{e'}^2_2-e_2e'_2\}\right.\\ &\left.+\frac{w_2(w_2+1)}{2}(1-\psi)^2(e_2-e'_2)^2\right]\\ &\left[1+w_3(1-\psi)e'_2+\frac{w_3(w_3+1)}{2}(1-\psi)^2{e'}^2_2\right].\end{aligned}$$

Hence, up to the first order of approximation:

$$\begin{aligned}\mathrm{MSE}(\hat{M}_P)\cong& M_y^2E[e_0+\psi(e'_0-e_0)\\ &+w_1(1-\psi)(e_1-e'_1)+w_2(1-\psi)(e_2-e'_2)\\ &+w_3(1-\psi)e'_2]^2.\end{aligned}$$

Squaring and taking expectations, the MSE of $\hat{M}_P$, to the first degree of approximation, is given by:

$$\begin{aligned}\mathrm{MSE}(\hat{M}_P)\cong&\frac{1}{4\{f_y(M_y)\}^2}\left[\left(\frac{1}{n}-\frac{1}{N}\right)\right.\\ &+(\psi^2-2\psi)\left(\frac{1}{n}-\frac{1}{n'}\right)\\ &+\left(\frac{1}{n'}-\frac{1}{N}\right)\left\{(1-\psi)^2w_3^2\left(\frac{M_yf_y(M_y)}{M_zf_z(M_z)}\right)^2\right.\\ &\left.+2(1-\psi)w_3\left(\frac{M_yf_y(M_y)}{M_zf_z(M_z)}\right)\rho_{yz}\right\}\\ &+(1-\psi^2)\left(\frac{1}{n}-\frac{1}{n'}\right)\left\{w_1^2\left(\frac{M_yf_y(M_y)}{M_xf_x(M_x)}\right)^2\right.\\ &\left.+w_2^2\left(\frac{M_yf_y(M_y)}{M_zf_z(M_z)}\right)^2\right\}\\ &+2(1-\psi)^2\left(\frac{1}{n}-\frac{1}{n'}\right)\left\{w_1\left(\frac{M_yf_y(M_y)}{M_xf_x(M_x)}\right)\rho_{yx}\right.\\ &\left.+w_2\left(\frac{M_yf_y(M_y)}{M_zf_z(M_z)}\right)\rho_{yz}\right\}\\ &+2w_1w_2(1-\psi)^2\left(\frac{1}{n}-\frac{1}{n'}\right)\\ &\left.\left(\frac{\{M_yf_y(M_y)\}^2}{\{M_xf_x(M_x)\}\{(M_zf_z(M_z)\}}\right)\rho_{xz}\right\}\Bigg]. \quad (9)\end{aligned}$$

Setting $\frac{\partial \mathrm{MSE}(\hat{M}_P)}{\partial w_i}=0$, $(i=1,2,3)$, we have:

$$w_{1(\mathrm{opt})}=\frac{M_xf_x(M_x)(\rho_{yz}\rho_{xz}-\rho_{yx})}{M_yf_y(M_y)(1-\rho_{xz}^2)},$$

$$w_{2(\mathrm{opt})}=\frac{M_zf_z(M_z)(\rho_{yx}\rho_{xz}-\rho_{yz})}{M_yf_y(M_y)(1-\rho_{xz}^2)},$$

and:

$$w_{3(\mathrm{opt})}=-\frac{M_zf_z(M_z)\rho_{yz}}{M_yf_y(M_y)(1-\psi)}.$$

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:

$$\begin{aligned}\mathrm{MSE}(\hat{M}_P)_{\min}\cong&\frac{1}{4\{f_y(M_y)\}^2}\left[\left(\frac{1}{n'}-\frac{1}{N}\right)(1-\rho_{yz}^2)\right.\\ &\left.+(1-\psi)^2\left(\frac{1}{n}-\frac{1}{n'}\right)(1-R_{y.xz}^2)\right]. \quad (10)\end{aligned}$$

Note that Jhajj and Walia [16] have shown that MSE is minimum for $\psi=1$. So, further minimizing Relation (10) with respect to ψ (i.e. taking $\psi=1$), we have:

$$\mathrm{MSE}(\hat{M}_P)_{\min}^{\psi=1}\cong\frac{1}{4\{f_y(M_y)\}^2}\left[\left(\frac{1}{n'}-\frac{1}{N}\right)(1-\rho_{yz}^2)\right]. \quad (11)$$

In Tables 2 and 3, MSE values and Percent Relative Efficiency (PRE) are given for different values of ψ, 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 ψ.

Estimator	Population 1	Population 2	Population 3
$\hat{M}_y$	565443.57	2.22	113343.27
$\hat{M}_R$	840264.22	1.01	180840.61
$\hat{M}_{SR}$	525744.59	0.87	110225.37
$\hat{M}_G$	506293.76	0.57	109805.56
$\hat{M}_P$			
$\psi = 0$	506293.76	0.57	109805.56
$\psi = 0.5$	360471.28	0.38	75308.90
$\psi = 1$	311863.78	0.31	63810.01
$\psi = 1.5$	360471.28	0.38	75308.90
$\psi = 2$	506293.76	0.57	109805.56

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

Estimator	Population 1	Population 2	Population 3
$\hat{M}_y$	100.000	100.000	100.000
$\hat{M}_R$	67.294	220.004	62.676
$\hat{M}_{SR}$	107.551	254.494	102.829
$\hat{M}_G$	111.683	390.314	103.222
$\hat{M}_P$			
$\psi = 0$	111.683	390.314	103.222
$\psi = 0.5$	156.862	587.840	150.504
$\psi = 1$	181.311	707.124	177.626
$\psi = 1.5$	156.862	587.840	150.504
$\psi = 2$	111.683	390.314	103.222

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), $\text{MSE}(\hat{M}_P)_{\min} < \text{Var}(\hat{M}_y)$ if:

$$\left(\frac{1}{n'} - \frac{1}{N}\right)\rho_{yz}^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left\{1 - (1-\psi)^2(1 - R_{y.xz}^2)\right\} > 0.$$

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

$$\left(\frac{1}{n} - \frac{1}{n'}\right)\left[\left(\frac{M_y f_y(M_y)}{M_x f_x(M_x)} - \rho_{yx}\right)^2 - (1-\psi)^2(1 - R_{y.xz}^2)\right] + \left(\frac{1}{n'} - \frac{1}{N}\right)\left[\left(\frac{M_y f_y(M_y)}{M_z f_z(M_z)} - \rho_{yz}\right)^2 - (1 - \rho_{yz}^2)\right] > 0.$$

Condition (iii)
By Relations (5) and (10):

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

if:

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

Condition (iv)
By Relations (7) and (10):

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

if:

$$\left(\frac{1}{n} - \frac{1}{n'}\right)(1 - R_{y.xz}^2)\psi(2-\psi) > 0.$$

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:

$$N = 69, \quad n' = 24, \quad n = 17,$$

$$M_y = 2068, \quad M_x = 2011, \quad M_z = 2307,$$

$$f_y(M_y) = 0.00014, \quad f_x(M_x) = 0.00014,$$

$$f_z(M_z) = 0.00013, \quad \rho_{yx} = 0.1505,$$

$$\rho_{yz} = 0.3166, \quad \rho_{xz} = 0.1431.$$

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:

$$N = 67, \quad n' = 23, \quad n = 15,$$

$$M_y = 4.8, \quad M_x = 7, \quad M_z = 151,$$

$$f_y(M_y) = 0.0763, \quad f_x(M_x) = 0.0526,$$

$$f_z(M_z) = 0.00024, \quad \rho_{yx} = 0.6624,$$

$$\rho_{yz} = 0.8624, \quad \rho_{xz} = 0.7592.$$

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:

$$N = 97, \quad n' = 46, \quad n = 33,$$

$$M_y = 1242, \quad M_x = 1233, \quad M_z = 1207,$$

$$f_y(M_y) = 0.00021, \quad f_x(M_x) = 0.00022,$$

$$f_z(M_z) = 0.00023, \quad \rho_{yx} = 0.2096,$$

$$\rho_{yz} = 0.1233, \quad \rho_{xz} = 0.1496.$$

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

$$\text{PRE} = \frac{\text{Var}\left(\hat{M}_y\right)}{\text{MSE}\left(\hat{M}_i\right) \text{ or MSE}\left(\hat{M}_i\right)_{\min}} \times 100,$$

$$i = y, R, SR, G, P.$$

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

The estimators M_i $(i = y, R, SR, G)$ are independent of ψ. 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 ψ. 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 ψ, their numerical results clearly show that optimal value of ψ 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] $(\hat{M}_{\text{SR}})$ 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 ψ, 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

1. Kadilar, C. and Cingi, H. "Ratio estimators in stratified sampling", *Biometrical Journal*, **45**, pp. 218-225 (2003).
2. Kadilar, C. and Cingi, H. "Anew estimator in stratified random sampling", *Communications in Statistics-Theory and Methods*, **34**, pp. 5967-602 (2005).
3. Shabbir, J. and Gupta, S. "A new estimator of population mean in stratified sampling", *Communications in Statistics-Theory and Methods*, **35**(7), pp. 1201-1209 (2006).
4. Koyuncu, N. and Kadilar, C. "Efficient estimators for the population mean", *Hacettepe Journal of Mathematics and Statistics*, **38**(2), pp. 217-225 (2009).

5. Koyuncu, N. and Kadilar, C. "On the family of estimators of population mean in stratified sampling", *Pakistan Journal of Statistics*, **26**(2), pp. 427-443 (2010).
6. Diana, G., Giordan, M. and Perri, P.F. "An improved class of estimators for the population mean", *Statistical Methods and Applications*, **20**(2), pp. 123-140 (2011).
7. Kuk, A.Y.C and Mak, T.K. "Median estimation in the presence of auxiliary information", *Journal of Royal Statistical Society*, B, **51**, pp. 261-269 (1989).
8. Singh, S., Singh H.P. and Upadhyaya, L.N. "Chain ratio and regression type estimators for median estimation in survey sampling", *Statistical Papers*, **48**, pp. 23-46 (2006).
9. Gupta, S., Shabbir, J. and Ahmad, S. "Estimation of median in two phase sampling using two auxiliary variable", *Communications in Statistics-Theory and Methods*, **37**(11), pp. 1815-1822 (2008).
10. Al, S. and Cingi, H. "New estimators for the population median in simple random sampling", *Proceedings of the Tenth Islamic Countries Conference on Statistical Sciences (ICCS-X)*, pp. 375-383 (2010).
11. Singh, H.P. and Solanki, R.S. "Some classes of estimators for the population median using auxiliary information", *Communications in Statistics-Theory and Methods*, **42**, pp. 4222-4238 (2013).
12. Khoshnevisan, M., Sexena, S., Singh, H.P., Singh, S. and Smaradache, F. "A general class of estimators of population median using two auxiliary variables in double sampling", pp. 26-43 in *Randomness and Optimal Estimation in Data Sampling*, 2nd Ed., pp. 1-63 (2002).
13. Gross, S.T. "Median estimation in sample surveys", *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pp. 181-184 (1980).
14. Chand, L. "Some ratio-estimators based on two or more auxiliary variables", PhD Dissertation, Iowa State university, Ames, Iowa (1975).
15. Srivastava, S.K., Rani, S., Khare, B.B. and Srivastava, S.R. "A generalized chain ratio estimator for mean of finite population", *Journal of Indian Society of Agricultural Statistics*, **42**(1), pp. 108-117 (1990).
16. Jhajj, H.S. and Walia, G.S. "A generalized difference-cum-ratio type estimator for the population variance in double sampling", *Communications in Statistics-Simulation and Computation*, **41**, pp. 58-64 (2012).
17. Diana, G. "A class of estimators of the population mean in stratified random sampling", *Statistica*, **53**(1), pp. 59-66 (1993).
18. Singh, S., *Advances Sampling Theory and Applications: How Michael 'Selected' Amy*, **I and II**, Kluwer Academic Publishers, The Netherlands (2003).
19. Aczel, A.D. and Sounderpandian, J., *Complete Business Statistics*, 5th Ed., New York, McGraw Hill (2004).
20. MFA, *Crops Area Production*, Ministry of Food, Agriculture and Livestock, Economic Wing, Islamabad, Pakistan (2004).

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.