Improvement on Estimating of Median in Two-Phase Sampling Using Two Auxiliary Variables

Muhammad Aamir [1], Zahoor Ul Haq [2], Muhammad Tariq[1], Shahid Jan[1], Muhammad Jahangir[1]

[1.] Department of Management Sciences, Abdul Wali Khan University Mardan, KPK, Pakistan.
[2.] Chairman Department of Management Sciences, Abdul Wali Khan University Mardan, KPK, Pakistan.
aamirkhan@awkum.edu.pk

Abstract: In this paper, we have proposed an estimator which is the combination of a difference and exponential type of estimator using two auxiliary variables for estimating median in two-phase sampling. Under simple random sampling without replacement scheme, the expressions for the bias and mean square error are given for the suggested estimator. The Bias and MSE of the suggested estimator is less than all the other median estimators existing in the literature. The MSE and Bias comparison are provided using four different data sets.

[Muhammad Aamir, Zahoor Ul Haq, Muhammad Tariq, Shahid Jan, Muhammad Jahangir. **Improvement on Estimating of Median in Two-Phase Sampling Using Two Auxiliary Variables.** *Life Sci J* 2014;11(4s):98-104]. (ISSN:1097-8135). http://www.lifesciencesite.com. 14

Key words: Auxiliary Variables, Median, Range, Correlation Coefficient, Bias, Mean Square error, Two-Phase sampling, Relative efficiency

1. Introduction

The main objective of this paper is to propose more efficient estimator for median estimation in two-phase sampling using two auxiliary variables. In survey sampling we deals with the highly skewed distribution so in this kind of situation we prefer to use median as an average for skewed distributions. In sampling literature Gross (1980), Kuk and Mak (1989), Singh et al. (2001, 2006), Allen et al. (2002) etc have studied the estimation of median. At the end Gupta et al. (2008) have proposed a class of estimators, in their study of estimation of median, but they couldn't provide an estimator which has less MSE than the other existing estimator Singh et al. (2006) but improved the bias factor. In this paper we proposed an estimator which has more efficient and less bias than the Singh et al. and Gupta et al. (2008) and the other median estimator.

Consider a finite population *U= {1,2,3…,i,…N).* Let assume that Y be the study variable and the one auxiliary variable is representing by X and the other by Z. Let y_i and x_i , z_i $_{(i=1,2,3,\ldots)}$ denotes the values of the units included in sample s_m of size m selected from a population by the method of simple random sampling without replacement. We assume that y and x are strongly related and information on population median M_x is not available, but the information on the second auxiliary variable Z (closely related with the auxiliary variable X but remotely related with the study variable Y), is available on all units of the population. The two-phase sampling scheme is given below:

(i) The first phase sample s_n ($s_n \subset U$) of fixed size n is drawn to observe only x in order to furnish an estimate of M_x.

(ii) Given s_n, the second phase sample s_m ($s_m \subset s_n$) of fixed size m is drawn to observe y, x, and z.

Let M_y , M_x , and M_z be the population medians of their respective subscripts variables indicated by the subscripts. Let $\widehat{M}_y$, $\widehat{M}_x$, and $\widehat{M}_z$ be the corresponding sample medians for the second phase. Let $\widehat{M}'_x$ and $\widehat{M}'_z$ be the first-phase sample medians. Let $f(M_y)$, $f(M_x)$, and $f(M_z)$ be the probability density functions of M_y , M_x , and M_z , respectively. Let ρ_{xy} be the coefficient of correlation between sampling distribution of $\widehat{M}_y$ and $\widehat{M}_x$ which is defined as $\rho_{xy} = \rho_{(\widehat{M}_x,\widehat{M}_y)} = 4\{P_{11}(x,y) - 1\}$, where $P_{11}(x,y) = P(X \le M_x \cap Y \le M_y)$. It is assumed that as $N \longrightarrow \infty$, the distribution of the bivariate variable (X , Y) approaches a continuous distribution with marginal densities for X and Y, respectively. Similarly, we can define ρ_{xz} as : $\rho_{xz} = \rho_{(\widehat{M}_x,\widehat{M}_z)} = 4\{P_{11}(x,z) - 1\}$, where $P_{11}(x,z) = P(X \le M_x \cap Z \le M_z)$ and $\rho_{yz} = \rho_{(\widehat{M}_y,\widehat{M}_z)} = 4\{P_{11}(y,z) - 1\}$, where $P_{11}(y,z) = P(X \le M_y \cap Z \le M_z)$ respectively.

Suppose that $y_{(1)}, y_{(2)}, \ldots , y_{(n)}$, are the sample y values in the ascending order. Let t be an integer such that $Y_{(t)} \le M_y \le Y_{(t+1)}$ and let $P = t/m$ be the proportion of Y, values in the sample that are less than or equal to the value M_y (see Singh et al., 2006).

2. Estimators Suggested by Various Authors

We now discuss briefly a few estimators for median.

(i) Median per Unit Estimator.

Gross (1980) suggested the following estimator:

$$\widehat{M}_{GR} = \widehat{M}_y$$

The variance of $\widehat{M}_{GR}$ is given by:

$$Var(\widehat{M}_{GR}) = \frac{\left(\frac{1}{m} - \frac{1}{N}\right)}{4\left\{f(M_y)\right\}^2}$$

(ii) *Ratio Estimator in Two-Phase Sampling.*

Singh et al. (2001) suggested the following ratio estimator: $\widehat{M}_{SA} = \frac{\widehat{M}_y}{\widehat{M}_x}\widehat{M}'_x$

The bias and MSE of $\widehat{M}_{SA}$ are as follows:

$$B(\widehat{M}_{SA}) \cong \frac{M_y\left(\frac{1}{m}-\frac{1}{N}\right)}{4\{M_x f(M_x)\}^2}\left\{1-\rho_{xy}\left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)\right\}$$

$$MSE(\widehat{M}_{SA}) \cong \frac{1}{4\left\{f(M_y)\right\}^2}\left[\left(\frac{1}{m}-\frac{1}{N}\right) + \left(\frac{1}{m}-\frac{1}{n}\right)\left(\frac{M_y f(M_y)}{M_x f(M_x)}\right) \times \left\{\left(\frac{M_y f(M_y)}{M_x f(M_x)}\right) - 2\rho_{xy}\right\}\right]$$

(iii) Median Ratio-Type Estimator.

Srivastava (1970), a ratio-type estimator is given by: $\widehat{M}_{SR} = \widehat{M}_y\left(\frac{\widehat{M}'_x}{\widehat{M}_x}\right)^{\delta}$

where δ is constant.

The bias and MSE of $\widehat{M}_{SR}$ are as follows:

$$B(\widehat{M}_{SR}) \cong \frac{\left(\frac{1}{m}-\frac{1}{n}\right)\rho_{xy}}{8\,M_x\{f(M_x)f(M_y)\}^2}\left\{1-\rho_{xy}\left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)\right\}$$

$$\mathrm{MSE}(\widehat{M}_{SR}) \cong \frac{1}{4\left\{f(M_y)\right\}^2}\left\{\left(\frac{1}{m}-\frac{1}{N}\right) - \left(\frac{1}{m}-\frac{1}{n}\right)\rho_{xy}^2\right\}$$

for $\delta = \rho_{xy}\left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)$

(iv) Chain Ratio-Type Estimator

Chand (1975), a chain-type ratio estimator is given by: $\widehat{M}_{CH} = \widehat{M}_y\left(\frac{\widehat{M}'_x}{\widehat{M}_x}\right)\left(\frac{\widehat{M}_z}{\widehat{M}'_z}\right)$

The bias and MSE of $\widehat{M}_{CH}$ are as follows:

$$B(\widehat{M}_{CH}) \cong \frac{M_y}{4}\left\{\frac{\left(\frac{1}{m}-\frac{1}{n}\right)}{\{M_x f(M_x)\}^2}\left(1-\rho_{xy}\left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)\right) + \frac{\left(\frac{1}{n}-\frac{1}{N}\right)}{\{M_z f(M_z)\}^2}\left(1-\rho_{yz}\left(\frac{M_z f(M_z)}{M_y f(M_y)}\right)\right)\right\}$$

$$MSE(\widehat{M}_{CH}) \cong \frac{1}{4\left\{f(M_y)\right\}^2}\left[\left(\frac{1}{m}-\frac{1}{N}\right) + \left(\frac{1}{m}-\frac{1}{n}\right)\left(\frac{M_y f(M_y)}{M_x f(M_x)}\right) \times \left\{\left(\frac{M_y f(M_y)}{M_x f(M_x)}\right) - 2\rho_{xy}\right\} + \left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{M_y f(M_y)}{M_z f(M_z)}\right) \times \left\{\left(\frac{M_y f(M_y)}{M_z f(M_z)}\right) - 2\rho_{yz}\right\}\right]$$

(v) Power Chain-Type Ratio Estimator

Srivastava et al. (1990), a power chain-type ratio estimator is given by: $\widehat{M}_{SRA} = \widehat{M}_y\left(\frac{\widehat{M}'_x}{\widehat{M}_x}\right)^{\lambda_1}\left(\frac{\widehat{M}_z}{\widehat{M}'_z}\right)^{\lambda_2}$

where λ_i $(i = 1,2)$ are constants.

The bias and MSE of $\widehat{M}_{SRA}$ are as follows:

$$B(\widehat{M}_{SRA}) \cong \frac{1}{8f(M_y)}\left[\frac{\left(\frac{1}{m}-\frac{1}{n}\right)\rho_{xy}}{M_x f(M_x)}\left(1-\rho_{xy}\left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)\right) + \frac{\left(\frac{1}{n}-\frac{1}{N}\right)\rho_{yz}}{M_z f(M_z)}\left(1-\rho_{yz}\left(\frac{M_z f(M_z)}{M_y f(M_y)}\right)\right)\right]$$

$$MSE(\widehat{M}_{SRA}) \cong \frac{1}{4\left\{f(M_y)\right\}^2}\left[\left(\frac{1}{m}-\frac{1}{N}\right) - \left(\frac{1}{m}-\frac{1}{n}\right)\rho_{xy}^2 - \left(\frac{1}{n}-\frac{1}{N}\right)\rho_{yz}^2\right]$$

for $\lambda_1 = \rho_{xy}\left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)$ *and* $\lambda_2 = \rho_{yz}\left(\frac{M_z f(M_z)}{M_y f(M_y)}\right)$

The MSE of $\widehat{M}_{SRA}$ is equal to the variance of linear regression estimator in two-phase sampling using two auxiliary variables as given below:

$$\widehat{M}_d = \widehat{M}_y + d_1\left(\widehat{M}'_x - \widehat{M}_x\right) + d_2\left(\widehat{M}_z - \widehat{M}'_z\right)$$

where d_i $(i = 1,2)$ are constants.

(vi) Singh et al. Estimator

Singh et al. (2006) considered the following ratio-type estimator:

$$\widehat{M}_S = \widehat{M}_y\left(\frac{\widehat{M}'_x}{\widehat{M}_x}\right)^{\alpha_1}\left(\frac{M_z}{\widehat{M}'_z}\right)^{\alpha_2}\left(\frac{M_z}{\widehat{M}_z}\right)^{\alpha_3}$$

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

The bias and MSE of $\widehat{M}_S$ are as follows:

$$\hat{M}_S \cong \frac{M_y}{8\left\{M_y f(M_y)\right\}^2 (1-\rho_{xz}^2)^2} \times \left[\left(\frac{1}{m}-\frac{1}{n}\right)\left\{\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)^2 - 2\rho_{xy}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2) + \left(\frac{M_y f(M_y)}{M_x f(M_x)}\right)\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2)\right\} + \left(\frac{1}{n}-\frac{1}{N}\right)\left\{\rho_{xz}^2\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)^2 - 2\rho_{yz}\rho_{xz}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2) + \left(\frac{M_y f(M_y)}{M_z f(M_z)}\right)\rho_{xz}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2)\right\} + \left(\frac{1}{m}-\frac{1}{N}\right)\left\{\left(\rho_{yz}-\rho_{xy}\rho_{xz}\right)^2 + 2\rho_{xz}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)\left(\rho_{yz}-\rho_{xy}\rho_{xz}\right) + \left(\frac{M_y f(M_y)}{M_z f(M_z)}\right)\left(\rho_{yz}-\rho_{xy}\rho_{xz}\right)(1-\rho_{xz}^2)\right\}\right]$$

$$\text{MSE}(\hat{M}_S)\,\text{min} = \frac{1}{4\left\{f(M_y)\right\}^2}\left[\left(\frac{1}{m}-\frac{1}{N}\right) - \left(\frac{1}{n}-\frac{1}{N}\right)\rho_{yz}^2 - \left(\frac{1}{m}-\frac{1}{n}\right)R_{y.xz}^2\right]$$

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

The optimum values of α's are given as follows:

$$\alpha_1 = \left(\frac{M_x f(M_x)}{M_y f(M_y)}\right)\left(\frac{\rho_{yz}\rho_{xz} - \rho_{xy}}{\rho_{xz}^2 - 1}\right)$$

$$\alpha_2 = \left(\frac{M_z f(M_z)}{M_y f(M_y)}\right)\rho_{xz}\left(\frac{\rho_{yz}\rho_{xz} - \rho_{xy}}{\rho_{xz}^2-1}\right) \qquad \alpha_3 = \left(\frac{M_z f(M_z)}{M_y f(M_y)}\right)\left(\frac{\rho_{xy}\rho_{xz} - \rho_{yz}}{\rho_{xz}^2-1}\right)$$

(vii) Gupta et al. Estimator

Gupta et al. (2008), Consider an extension of the Singh et al. (2006) estimator by using a transformation involving the range of the second auxiliary variable Z. The estimator is: $\hat{M}_P = \hat{M}_y \left(\frac{\hat{M}_x'}{\hat{M}_x}\right)^{\gamma_1}\left(\frac{M_z+R_z}{\hat{M}_z'+R_z}\right)^{\gamma_2}\left(\frac{M_z+R_z}{\hat{M}_z+R_z}\right)^{\gamma_3}$

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

The bias and MSE of $\hat{M}_P$ are as follows:

$$B(\hat{M}_P) \cong \frac{M_y}{8\left\{M_y f(M_y)\right\}^2 (1-\rho_{xz}^2)^2} \times \left[\left(\frac{1}{m}-\frac{1}{n}\right)\left\{\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)^2 - 2\rho_{xy}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2) + \left(\frac{M_y f(M_y)}{M_x f(M_x)}\right)\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2)\right\} + \left(\frac{1}{n}-\frac{1}{N}\right)\left\{\rho_{xz}^2\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)^2 - 2\rho_{yz}\rho_{xz}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2) + \left(\frac{M_y f(M_y)}{M_z f(M_z)}\right)\left(\frac{M_z}{M_z+R_z}\right)\rho_{xz}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)(1-\rho_{xz}^2)\right\} + \left(\frac{1}{m}-\frac{1}{N}\right)\left\{\left(\rho_{yz}-\rho_{xy}\rho_{xz}\right)^2 + 2\rho_{xz}\left(\rho_{xy}-\rho_{yz}\rho_{xz}\right)\left(\rho_{yz}-\rho_{xy}\rho_{xz}\right) + \left(\frac{M_y f(M_y)}{M_z f(M_z)}\right)\left(\frac{M_z}{M_z+R_z}\right)\left(\rho_{yz}-\rho_{xy}\rho_{xz}\right)(1-\rho_{xz}^2)\right\}\right]$$

$$\text{MSE}(\hat{M}_P)\text{min} = \frac{1}{4\left\{f(M_y)\right\}^2}\left[\left(\frac{1}{m}-\frac{1}{N}\right) - \left(\frac{1}{n}-\frac{1}{N}\right)\rho_{yz}^2 - \left(\frac{1}{m}-\frac{1}{n}\right)R_{y.xz}^2\right]$$

The optimum values of γ 's are given as follows:

$$\gamma_1 = \alpha_1 \;,\; \gamma_2 = \alpha_2 \frac{1}{\left(\frac{M_z}{M_z+R_z}\right)}$$

$$and \quad \gamma_3 = \alpha_3 \frac{1}{\left(\frac{M_z}{M_z + R_z}\right)}$$

The MSE of $\widehat{M}_P$ is exactly the same as for the Singh et al.(2006) estimator. However, the bias expressions for the two estimators $\widehat{M}_S$ and $\widehat{M}_P$ are slightly different.

(viii) Nursel, K. Exponential Type of Estimator

Nursel(2011) consider the following estimator:

$$\widehat{M}_N = \left[k_1 \widehat{M}_y + k_2\left(\widehat{M}'_x - \widehat{M}_x\right)\right] exp\left(\frac{M_z - \widehat{M}'_z}{M_z + \widehat{M}'_z}\right)$$

where k_i $(i = 1,2)$ are constants.

The bias and MSE of $\widehat{M}_N$ are as follows:

$$B\left(\widehat{M}_N\right) = (k_1 - 1)M_y - \frac{k_1}{8}\left(\frac{N-n}{Nn}\right)\left\{\frac{\rho_{yz}}{f(M_y)M_z f(M_z)} - \frac{3M_y}{4\left(M_z f(M_z)\right)^2}\right\}$$

$$\text{MSE}(\widehat{M}_N)\text{min} = M_y^2 - \frac{M_y^4\left(f(M_y)\right)^2}{4\, M_y^2\, A\left(f(M_y)\right)^2 + \rho_{xy}^2\left(\frac{m-n}{mn}\right)} B^2$$

The optimum values of k's are given as follows:

$$k_1 = \frac{-B}{2A\left[1 + \left(\frac{m-n}{mn}\right)\frac{\rho_{xy}^2}{4\, M_y^2\, A\left(f(M_y)\right)^2}\right]}$$

and

$$k_2 = \frac{-B\, \rho_{xy} f(M_x)}{2A\, f(M_y)\left[1 + \left(\frac{m-n}{mn}\right)\frac{\rho_{xy}^2}{4M_y^2\, A\left(f(M_y)\right)^2}\right]}$$

where

$$\text{A} = \left[1 + \left(\frac{N-m}{4mN}\right)\frac{1}{\left(M_y f(M_y)\right)^2} - \left(\frac{N-n}{2nN}\right) \times \left\{\frac{\rho_{yz}}{M_y f(M_y) M_z f(M_z)} + \frac{1}{2\left(M_z f(M_z)\right)^2}\right\}\right]$$

$$\text{B} = \left[\frac{1}{4}\left(\frac{N-n}{nN}\right)\frac{\rho_{yz}}{M_y f(M_y) M_z f(M_z)} - \frac{3}{16}\left(\frac{N-n}{nN}\right)\frac{1}{\left(M_z f(M_z)\right)^2} - 2\right]$$

3. The Proposed Estimator

We have proposed the new estimator for the estimation of median in two- phase sampling which is the combination of a difference and exponential type of estimator. The proposed estimator is as follows:

$$\widehat{M}_A = t_1 \widehat{M}_y + t_2\left(\widehat{M}'_x - \widehat{M}_x\right) + t_3 \exp\left[\left(\frac{d\left(\widehat{M}_z - M_z\right)}{2a + d\left(\widehat{M}_z + M_z\right)}\right) + \left(\frac{d\left(\widehat{M}'_z - M_z\right)}{2a + d\left(\widehat{M}'_z + M_z\right)}\right)\right]$$

where t_i $(i = 1,2,3)$ are constants, a and d are either constants or function of known population parameters of second auxiliary variable such as skewness, kurtosis, correlation coefficient ρ_{xz} , Range etc.

$$\widehat{M}_A = t_1 M_y(1 + e_0) + t_2 M_x(e_2 - e_1) + t_3 \exp\left[\left(\frac{dM_z e_3}{2a + 2dM_z + dM_z e_3}\right) + \left(\frac{dM_z e_4}{2a + 2dM_z + dM_z e_4}\right)\right]$$

taking, $b = \frac{dM_z}{2\,(a + dM_z)}$, Therefore

$$\widehat{M}_A = t_1 M_y(1 + e_0) + t_2 M_x(e_2 - e_1) + t_3 \exp\left[\left(\frac{be_3}{1 + be_3}\right) + \left(\frac{be_4}{1 + be_4}\right)\right]$$

After simplification we get,

$$\widehat{M}_A = t_1 M_y + t_1 M_y e_0 + t_2 M_x e_2 - t_2 M_x e_1 + t_3 + t_3 be_4 - \frac{t_3 b^2 e_4^2}{2} + t_3 be_3 + t_3 b^2 e_3 e_4 - \frac{t_3 b^2 e_3^2}{2}$$

$$\widehat{M}_A - M_y = (t_1 - 1)M_y + t_1 M_y e_0 + t_2 M_x e_2 - t_2 M_x e_1 + t_3 + t_3 be_4 - \frac{t_3 b^2 e_4^2}{2} - \frac{t_3 b^2 e_3^2}{2} + t_3 be_3 + t_3 b^2 e_3 e_4 \quad (1)$$

Taking expectation on both sides of (1) we get the bias of $\widehat{M}_A$ to the first degree of approximation as

$B(\widehat{M}_A)= (t_1-1)M_y + t_3 + \frac{1}{2}t_3b^2f_2C_z^2 - \frac{1}{2}t_3b^2f_1C_z^2$

Taking squaring and expectation on both sides of (1), ignoring higher order terms means e_i's having power greater than two, thus we have

$MSE(\widehat{M}_A)$ =2 $t_1M_yt_3 + t_3^2 + t_1^2M_y^2 - 2t_1M_y^2 + M_y^2 - 2M_yt_3 - t_2^2M_x^2f_2C_x^2 + 4t_3^2b^2f_2C_z^2 - M_yt_3b^2f_2C_z^2 + M_yt_3b^2f_1C_z^2 + t_1M_yt_3b^2f_2C_z^2 - t_1M_yt_3b^2f_1C_z^2 + t_1^2M_y^2f_1C_y^2 + t_2^2M_x^2f_1C_x^2 + 2t_1M_yt_2M_xf_2\rho_{xy}C_xC_y + 2t_1M_yt_3bf_2\rho_{yz}C_yC_z - 2t_1M_yt_2M_xf_1\rho_{xy}C_xC_y + 2t_1M_yt_3bf_1\rho_{yz}C_yC_z + 2t_2M_xt_3bf_2\rho_{xz}C_xC_z - 2t_2M_xt_3bf_2\rho_{xz}C_xC_z$ (2)

To obtain the values of $t_1, t_2 and\ t_3$. We differentiate (2) with respect to $t_1, t_2 and\ t_3$ and equate to zero then solve the equations for $t_1, t_2 and\ t_3$. i.e.

$$\frac{d\mathrm{MSE}(\widehat{M}_A)}{dt_1} = 0, \frac{d\mathrm{MSE}(\widehat{M}_A)}{dt_2} = 0, \frac{d\mathrm{MSE}(\widehat{M}_A)}{dt_3} = 0$$

The minimum bias and MSE of $\widehat{M}_A$ are as follows:

$$B(\widehat{M}_A) \cong \frac{M_y}{A_1}\left[\left(A_2bM_yf(M_y) - A_1\right) - A_4\left(2M_z^2f(M_z)^2 + b^2\left(\frac{1}{4n} - \frac{1}{4m}\right)\right)\right] \quad (3)$$

$$\mathrm{MSE}(\widehat{M}_A)\min \cong M_y^2\left[\frac{1}{A_1^2}\left\{\left\{A_2bM_yf(M_y)\left(A_2bM_yf(M_y) - 2A_1 - b^2A_4\left(\frac{1}{2n} - \frac{1}{2m}\right)\right)\right\} + 4A_4M_z^2f(M_z)^2\left\{A_1 + b^2A_4\left(\frac{1}{n} - \frac{1}{N}\right) + A_4M_z^2f(M_z)^2 - A_2bM_yf(M_y)\right\} + b^2\left\{\left(2A_1A_4 + 2A_2A_3\rho_{xy} - A_3^2\right)\left(\frac{1}{4n} - \frac{1}{4m}\right) + A_2^2\left(\frac{1}{4m} - \frac{1}{4N}\right) - A_4M_zf(M_z)\left(A_2\rho_{yz}\left(\frac{1}{n} + \frac{1}{m} - \frac{2}{N}\right) + A_3\rho_{xz}\left(\frac{1}{n} - \frac{1}{m}\right)\right)\right\}\right\} + 1\right] \quad (4)$$

The optimum values of *t's* are as follows:

$$t_1 = \frac{A_2bM_yf(M_y)}{A_1}, t_2 = \frac{A_3bM_yf(M_x)}{A_1}$$
$$and\ t_3 = -\frac{2A_4M_yM_z^2f(M_z)^2}{A_1}$$

where

$$b = \frac{dM_z}{2(a + dM_z)}$$

$$\begin{aligned} A_1 = {} & 4M_yf(M_y)M_zf(M_z)b\left(b^2\left(\rho_{yz}(f_2^2 - f_1^2) + \rho_{xy}\rho_{xz}(f_2 - f_1)^2\right)\right. \\ & \left. + 2M_z^2f(M_z)^2\left(\rho_{yz}(f_2 + f_1) + \rho_{xy}\rho_{xz}(f_2 - f_1)\right)\right) + M_y^2f(M_y)^2b^4(f_2 - f_1)^2 \\ & - 4M_z^4f(M_z)^4\left(\rho_{xy}^2(f_2 - f_1) + f_1\right) \\ & + 4M_z^2f(M_z)^2b^2\left(\left(\rho_{xz}^2(f_1^2 - f_1f_2) + \rho_{yz}^2(f_2 + f_1)^2 + 4\rho_{xy}^2(f_1f_2 - f_2^2) - 4f_1f_2\right.\right. \\ & \left.\left. + 2\rho_{xy}\rho_{xz}\rho_{yz}(f_2^2 - f_1^2)\right) - M_y^2f(M_y)^2\left(f_1 + 3f_2 + \rho_{xz}^2(f_2 - f_1)\right)\right) \end{aligned}$$

$$\begin{aligned} A_2 = {} & 2M_zf(M_z)\left(b^2\left(\rho_{xy}\rho_{xz}(f_2 - f_1)^2 + \rho_{yz}(f_2^2 - f_1^2)\right) + 2M_z^2f(M_z)^2\left(\rho_{xy}\rho_{xz}(f_2 - f_1) + \rho_{yz}(f_2 + f_1)\right)\right) \\ & + M_yf(M_y)\left(b^3(f_2 - f_1)^2 - 4M_z^2f(M_z)^2b\left(f_1 + 3f_2 + \rho_{xz}^2(f_2 - f_1)\right)\right) \end{aligned}$$

$$\begin{aligned} A_3 = {} & 4M_yf(M_y)M_z^2f(M_z)^2b\left(\rho_{xz}\rho_{yz}(f_2 + f_1) - \rho_{xy}(f_1 + 3f_2)\right) \\ & + 2M_zf(M_z)b^2\left(\rho_{xy}(f_1^2 - f_1f_2) + \rho_{xy}\rho_{yz}(f_2^2 - f_1^2)\right) - 4M_z^3f(M_z)^3\left(\rho_{xz}f_1 - \rho_{xy}\rho_{yz}(f_2 + f_1)\right) \\ & + M_yf(M_y)\rho_{xy}b^3(f_2 - f_1)^2 \end{aligned}$$

$$\begin{aligned} A_4 = {} & 2M_z^2f(M_z)^2\left(f_1 + \rho_{xy}^2(f_2 - f_1)\right) - b^2\left(f_1^2 - f_1f_2 - \rho_{xy}^2(f_2 - f_1)^2\right) \\ & - 2bM_yf(M_y)M_zf(M_z)\left(\rho_{xy}\rho_{xz}(f_2 - f_1) + \rho_{yz}(f_2 + f_1)\right) \end{aligned}$$

$$f_1 = \frac{1}{4m} - \frac{1}{4N} \quad \text{and} \quad f_2 = \frac{1}{4n} - \frac{1}{4N}$$

4. Numerical Comparison

To check the relative efficiency and relative bias of the proposed estimator over the other estimators exist in the literature we use the following data sets which have been taken from different sources.

Data set 1(Source: Singh, 2003). *Y*: The number of fish caught by marine recreational fisherman in 1995; *X*: The number of fish caught by marine recreational fisherman in 1994; *Z*: The number of fish caught by marine recreational fisherman in 1993.

$$N = 69,\quad n = 24,\quad m = 17,\quad \rho_{xy} = 0.1505,\quad \rho_{yz} = 0.3166,\quad \rho_{xz} = 0.1431$$
$$M_y = 2068,\quad M_x = 2011,\quad M_z = 2307,\quad f(M_y) = 0.00014,\quad f(M_x) = 0.00014$$
$$f(M_z) = 0.00013,\quad R_z = 34025$$

Data set 2(Source: Aczel and Sounderpandian, 2004). *Y*: The U.S exports to Singapore in billions of Singapore dollars; *X*: The money supply figures in billions of Singapore dollars; *Z*: The local supply in U.S dollars.

$$N = 67,\quad n = 23,\quad m = 15,\quad \rho_{xy} = 0.6624,\quad \rho_{yz} = 0.8624,\quad \rho_{xz} = 0.7592,$$
$$M_y = 4.8,\quad M_x = 7.0,\quad M_z = 151,\quad f(M_y) = 0.0763,\quad f(M_x) = 0.0526$$
$$f(M_z) = 0.0024,\quad R_z = 116$$

Data set 3(Source: MFA, 2004). *Y*: District-wise tomato production (tones) in 2003; *X*: District-wise tomato production (tones) in 2002; *Z*: District-wise tomato production (tones) in 2001.

$$N = 97,\quad n = 46,\quad m = 33,\quad \rho_{xy} = 0.2096,\quad \rho_{yz} = 0.1233,\quad \rho_{xz} = 0.1494,$$
$$M_y = 1242,\quad M_x = 1233,\quad M_z = 1207,\quad f(M_y) = 0.00021,\quad f(M_x) = 0.00022$$
$$f(M_z) = 0.00023,\quad R_z = 16738$$

Data set 4(Source: Nursel, 2011). *Y*: The apple production in 1999 in Turkey; *X*: The number of trees in 1999 in Turkey; *Z*: Apple production in 1998 in Turkey.

$$N = 854, n = 410,\quad m = 290,\quad \rho_{xy} = 0.770491803,\quad \rho_{yz} = 0.8969555504,$$
$$M_y = 300,\quad M_x = 10000,\quad M_z = 303,\quad f(M_y) = 0.00002306,\quad f(M_x) = 0.00000271$$
$$\rho_{xz} = 0.793911007,\quad R_z = 126,\quad f(M_z) = 0.0000246$$

Using the above all data sets we have calculated MSE's and Biases of median estimators, the findings are in the following tables. It is observed from the following tables that the suggested estimator far better than the other estimator exist in the literature i.e Gupta et al.(2008), Singh et al. (2006) and Nursel (2011). It is observed that the suggested estimator has much smaller MSE and Bias than the other median estimators. We got the most efficient estimator for all data set above using d as R_z (Range of the second auxiliary variable) and $(1 - \rho_{xz})$.

Table 1: Numerical Comparison of *MSE*

Estimator	Data 1	Data 2	Data 3	Data 4
$\hat{M}_{GR}$	556443.5	2.22	113343.2	107052.5
$\hat{M}_{SA}$	717528.7	1.89	138663.1	914674.3
$\hat{M}_{SR}$	551589.5	1.78	111208.0	788875.5
$\hat{M}_{CH}$	826933.5	1.01	180828.4	435673.7
$\hat{M}_{SRA}$	517381.2	0.87	110223.9	309300.4
$\hat{M}_{S}$	463271.2	0.57	106785.4	204881.4
$\hat{M}_{P}$	463271.2	0.57	106785.4	204881.4
$\hat{M}_{N}$	518730.9	1.76	104869.1	87037.5
$\hat{M}_{A}$	1049.2	0.04	224.9	23814.2

Table 2: Numerical Comparison of *BIASES*

Estimator	Data 1	Data 2	Data 3	Data 4
$\hat{M}_{SA}$	5540.8	2.99	3335.3	208.1
$\hat{M}_{SR}$	7.0	0.02	3.1	313.7
$\hat{M}_{CH}$	200.5	0.11	68.3	150.8
$\hat{M}_{SRA}$	24.2	0.04	5.7	286.1
$\hat{M}_{S}$	58.6	0.44	8.1	2451.9
$\hat{M}_{P}$	20.2	0.28	3.6	1060.2
$\hat{M}_{N}$	218.6	0.23	79.8	290.1
$\hat{M}_{A}$	0.5	0.01	0.2	79.4

5. Conclusion

In this study we observe that our proposed estimator is more efficient than all other estimators present in literature and also it has less bias than all other existing median estimators for two-phase sampling. Thus we have recommended that the suggested estimator which is the combination of a difference and exponential type of estimator is used in practice.

Corresponding Author:
Muhammad Aamir
Department of Management Sciences
Abdul Wali Khan University Mardan
KPK, Pakistan.
E-mail: aamirkhan@awkum.edu.pk

References

1. Chand, L. (1975). Some Ratio-Estimators Based on Two or More Auxiliary Variables, Ph.D dissertation, Iwoa State University, Ames, Iowa. Unpublished.
2. Gross, S.T. (1980). Median estimation in sample surveys, Survey Research Method Section, American Statistical Association, 181-184.
3. Gupta, S., Shabbir, J. and Ahmad, S. (2008). Estimation of median in two-phase sampling using two auxiliary variables, Communications in Statistics-Theory and Methods, 37, 1815-1822.
4. Gupta, S. and Shabbir, J. (2007). On the use of transformed auxiliary variables in estimating the finite population mean using two auxiliary variables, Journal of Statistical Planning and Information, 137, 1606-1611.
5. Indian Horticulture Database (2011). State Wise Production of Major Spices in India, Ministry of agriculture, Govt of India, http//:www.nhb.gov.in.
6. Kuk, Y.C.A. and Mak, T.K. (1989).Median estimation in the presence of auxiliary information, Journal of Royal Statistical Society, B, 2, 261-269.
7. MFA. (2004).Crops Area Production, Ministry of Food and Agriculture,Islamabad, Pakistan.
8. Nursel, K. (2011). On improvement in estimating median in two phase sampling using two auxiliary information, Preprints.
9. Shabbir, J. and Gupta, S. (2005). Improved ratio estimators in stratified sampling, American Journal of Mathematics and Management Sciences, 25, 293-311.
10. Singh, S., Joarder, A.H. and Tracy, D.S. (2001). Median estimation using double sampling, Australian and Newzeland Journal of Statistics, 43, 1, 33-46.
11. Singh, S., Singh, H.P. and Upadhyaya, L.N. (2006). Chain ratio and regression type estimators for median estimation in survey sampling, Statistical Papers,48, 23-46.
12. Singh, S. (2003). Advanced Sampling Theory with Applications, Kluwer Academic Publishers, London.
13. Srivastava, S.K. (1980). A class of estimators using auxiliary information in sample surveys, Canadian Journal of Statistics, 8, 253-254.
14. Srivastava, S.K. (1971). A generalized estimator for the mean of a finite population using multi auxiliary information, Journal of the American Statistical Association, 66, 404-407.
15. Srivastava, S.K. and Jhajj, H.S. (1981). A class of estimators of the population mean in survey sampling using auxiliary information, Biometrika, 68, 341-343.
16. Srivastava, S.K. (1970). A two phase sampling estimator in sample surveys, Australian Journal of Statistics, 12, 1, 23-27.

3/1/2014