STATISTICA, anno LXXIV, n. 1, 2014

ESTIMATION OF FINITE POPULATION VARIANCE USING AUXILIARY INFORMATION IN SAMPLE SURVEYS

Housila P. Singh
School of Studies in Statistics, Vikram University, Ujjain-456010, M.P., India.
Alok K. Singh
School of Studies in Statistics, Vikram University, Ujjain-456010, M.P., India
Ramkrishna S. Solanki
School of Studies in Statistics, Vikram University, Ujjain-456010, M.P., India

1. INTRODUCTION

The problem of estimating the finite population mean in the presence of an auxiliary variable has been widely discussed in the finite population sampling literature. However in practice, the problem of estimation of population variance also assumes importance. Like finite population mean estimators, few authors including (Das and Tripathi 1978); (Isaki, 1983); (Srivastava and Jhajj, 1980); (Biradar and Singh, 1988); (Singh *et al.*,1988, 2003); (Cebrian and Garcia ,1997); (Kadilar and Cingi, 2007); (Grover, 2007); (Shabbir and Gupta, 2007); (Singh and Vishwakarma, 2008); (Singh and Solanki, 2009, 2010, 2013a, 2013b,); (Solanki and Singh, 2013) and others have paid their attention towards the estimation of population variance in the presence of auxiliary information.

Consider the finite population $U:\{U_1,U_2,...,U_N\}$ consisting of N units. Let Y and X denote the study variable and auxiliary variable taking values y_i and x_i respectively on the i^{th} unit U_i of the population U. The problem is to estimate the population variance S_y^2 of the study variable Y using information on the population variance S_x^2 of the auxiliary variable X which is highly correlated with the study variable Y. Let a sample of size n be drawn by simple random sampling without replacement (SRSWOR) from the population U. It is assumed that the population size N is large as compared to n so that finite population correction term is ignored.

The usual unbiased estimator of the population variance of the study variable Y is defined by

$$t_0 = s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i-\overline{y})^2 \tag{1.1}$$

where $\overline{y} = n^{-1}\sum_{i=1}^{n} y_i$ is the sample mean of study variable Y.

Using information on the population variance S_x^2 of the auxiliary variable X, Isaki (1983) suggested a ratio-type estimator of as

$$t_1 = s_y^2\left(S_x^2 / s_x^2\right) \tag{1.2}$$

where is an unbiased estimator of the population variance with and $\bar{X} = N^{-1}\sum_{i=1}^{N} x_i$.
Motivated by Bahl and Tuteja (1991), Singh et al. (2009) suggested a ratio-type exponential estimator of the population variance as

$$t_2 = s_y^2 \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right) \tag{1.3}$$

It is very well known that the mean squared error (*MSE*) or variance (*Var*) of the usual unbiased estimator s_y^2 under SRSWOR is approximately, for large n,

$$MSE\left(s_y^2\right) = Var\left(s_y^2\right) = n^{-1}S_y^4\left(\delta_{40} - 1\right) \tag{1.4}$$

where $\delta_{pq} = \left(\mu_{pq} / \mu_{20}^{p/2}\mu_{02}^{q/2}\right)$, $\mu_{pq} = N^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^p\left(x_i - \bar{X}\right)^q$ (p, q being non-negative integers) and $\bar{Y} = N^{-1}\sum_{i=1}^{N} y_i$.

To the first degree of approximation the biases and mean squared errors (*MSE*s) of t_1 and t_2 are respectively given by

$$Bias\left(t_1\right) = n^{-1}S_y^2\left(\delta_{04} - 1\right)\left(1 - c\right) \tag{1.5}$$

$$Bias\left(t_2\right) = \frac{S_y^2}{8n}\left(\delta_{04} - 1\right)\left(3 - 4c\right) \tag{1.6}$$

$$MSE\left(t_1\right) = n^{-1}S_y^4\left[\left(\delta_{40} - 1\right) + \left(\delta_{04} - 1\right)\left(1 - 2c\right)\right] \tag{1.7}$$

$$MSE\left(t_2\right) = n^{-1}S_y^4\left[\left(\delta_{40} - 1\right) + \left(1/4\right)\left(\delta_{04} - 1\right)\left(1 - 4c\right)\right] \tag{1.8}$$

where $c = \left[\left(\delta_{22} - 1\right) / \left(\delta_{04} - 1\right)\right]$.

In this paper motivated by Singh and Vishwakarma (2009), we have proposed some estimators of population variance of the study variable Y based on arithmetic mean, geometric mean and harmonic mean of the estimators $\left(t_0, t_1\right)$, $\left(t_0, t_2\right)$, $\left(t_1, t_2\right)$ and $\left(t_0, t_1, t_2\right)$. Properties of the suggested estimators are studied under large sample approximation. An empirical study is carried out in support of the present study.

2. SUGGESTED ESTIMATORS

In this section we have suggested some estimators of population variance based on usual unbiased estimator, usual ratio estimators and estimator due to Singh et al. (2009). The properties of suggested estimators have been obtained up to first degree of approximation

2.1 The estimators based on t_0 and t_1

Taking the arithmetic mean (*AM*), geometric mean (*GM*) and harmonic mean (*HM*) of the estimators t_0 and t_1 we get the following estimator of the population variance respectively as

$$t_3^{(AM)} = \frac{1}{2}\left(t_0 + t_1\right) = \left(\frac{s_y^2}{2}\right)\left(1 + \frac{S_x^2}{s_x^2}\right) \tag{2.1}$$

$$t_3^{(GM)} = \left(t_0 t_1\right)^{1/2} = s_y^2\left(S_x / s_x\right) \tag{2.2}$$

$$t_3^{(HM)} = \frac{2}{\left(\frac{1}{t_0} + \frac{1}{t_1}\right)} = \frac{2s_y^2}{\left(1 + \frac{s_x^2}{S_x^2}\right)} \tag{2.3}$$

To the first degree of approximation, the biases and mean squared errors of $t_3^{(AM)}$, $t_3^{(GM)}$ and $t_3^{(HM)}$ are respectively given by

$$Bias\left(t_3^{(AM)}\right) = \left(S_y^2 / 2n\right)\left(\delta_{04} - 1\right)\left(1 - c\right) \tag{2.4}$$

$$Bias\left(t_3^{(GM)}\right) = \left(S_y^2 / 8n\right)\left(\delta_{04} - 1\right)\left(3 - 4c\right) \tag{2.5}$$

$$Bias\left(t_3^{(HM)}\right) = \left(S_y^2 / 4n\right)\left(\delta_{04} - 1\right)\left(1 - 2c\right) \tag{2.6}$$

$$MSE\left(t_3^{(AM)}\right) = MSE\left(t_3^{(GM)}\right) = MSE\left(t_3^{(HM)}\right) = n^{-1}S_y^4\left[\left(\delta_{40} - 1\right) + \frac{\left(\delta_{04} - 1\right)}{4}\left(1 - 4c\right)\right] \tag{2.7}$$

2.2 The estimators based on t_0 and t_2

The estimators of S_y^2 based on *AM*, *GM* and *HM* of the estimators t_0 and t_2 are respectively defined as

$$t_4^{(AM)} = \frac{1}{2}\left(t_0 + t_2\right) = \left(\frac{s_y^2}{2}\right)\left(1 + \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right) \tag{2.8}$$

$$t_4^{(GM)} = \left(t_0 t_2\right)^{1/2} = s_y^2 \exp\left(\frac{1}{2}\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right) \tag{2.9}$$

$$t_4^{(HM)} = \frac{2}{\left(\frac{1}{t_0} + \frac{1}{t_2}\right)} = \frac{2s_y^2}{\left\{1 + \exp\left(\frac{s_x^2 - S_x^2}{S_x^2 + s_x^2}\right)\right\}} \tag{2.10}$$

To the first degree of approximation, the biases and *MSE*s of $t_4^{(AM)}$, $t_4^{(GM)}$ and $t_4^{(HM)}$ are respectively given by

$$Bias\left(t_4^{(AM)}\right) = \left(S_y^2 / 16n\right)\left(\delta_{04} - 1\right)\left(3 - 4c\right) \tag{2.11}$$

$$Bias\left(t_4^{(GM)}\right) = \left(S_y^2 / 32n\right)\left(\delta_{04} - 1\right)\left(5 - 8c\right) \tag{2.12}$$

$$Bias\left(t_4^{(HM)}\right) = \left(S_y^2 / 8n\right)\left(\delta_{04} - 1\right)\left(1 - 2c\right) \tag{2.13}$$

$$MSE\left(t_4^{(AM)}\right) = MSE\left(t_4^{(GM)}\right) = MSE\left(t_4^{(HM)}\right) = n^{-1} S_y^4 \left[\left(\delta_{40} - 1\right) + \frac{\left(\delta_{04} - 1\right)}{16}\left(1 - 8c\right)\right] \tag{2.14}$$

2.3 The estimators based on t_1 *and* t_2

We propose the following estimators of S_y^2 based on *AM*, *GM* and *HM* of the estimators respectively t_1 and t_2 as

$$t_5^{(AM)} = \frac{1}{2}\left(t_1 + t_2\right) = \left(\frac{s_y^2}{2}\right)\left(\frac{S_x^2}{s_x^2} + \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right) \tag{2.15}$$

$$t_5^{(GM)} = \left(t_1 t_2\right)^{1/2} = s_y^2 \left(\frac{S_x}{s_x}\right) \exp\left(\frac{1}{2}\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right) \tag{2.16}$$

$$t_5^{(HM)} = \frac{2}{\left(\frac{1}{t_1}+\frac{1}{t_2}\right)} = \frac{2s_y^2}{\left\{\frac{s_x^2}{S_x^2}+\exp\left(\frac{S_x^2-s_x^2}{S_x^2+s_x^2}\right)\right\}} \tag{2.17}$$

To the first degree of approximation, the biases and *MSE*s of $t_5^{(AM)}$, $t_5^{(GM)}$ and $t_5^{(HM)}$ are respectively given by

$$Bias\left(t_5^{(AM)}\right) = \left(S_y^2/16n\right)\left(\delta_{04}-1\right)\left(11-12c\right) \tag{2.18}$$

$$Bias\left(t_5^{(GM)}\right) = \left(3S_y^2/32n\right)\left(\delta_{04}-1\right)\left(7-8c\right) \tag{2.19}$$

$$Bias\left(t_5^{(HM)}\right) = \left(S_y^2/8n\right)\left(\delta_{04}-1\right)\left(5-6c\right) \tag{2.20}$$

$$MSE\left(t_5^{(AM)}\right) = MSE\left(t_5^{(GM)}\right) = MSE\left(t_5^{(HM)}\right) = n^{-1}S_y^4\left[\left(\delta_{40}-1\right)+\frac{3\left(\delta_{04}-1\right)}{16}\left(3-8c\right)\right] \tag{2.21}$$

2.4 The estimators based on t_0, t_1 and t_2

We define the following estimators of S_y^2 based on *AM*, *GM* and *HM* of the estimators t_0, t_1 and t_2 respectively as

$$t_6^{(AM)} = \frac{1}{2}\left(t_0+t_1+t_2\right) = \left(\frac{s_y^2}{3}\right)\left(1+\frac{S_x^2}{s_x^2}+\exp\left(\frac{S_x^2-s_x^2}{S_x^2+s_x^2}\right)\right) \tag{2.22}$$

$$t_6^{(GM)} = \left(t_0t_1t_2\right)^{1/2} = s_y^2\left(\frac{S_x}{s_x}\right)^{2/3}\exp\left(\frac{1}{3}\left(\frac{S_x^2-s_x^2}{S_x^2+s_x^2}\right)\right) \tag{2.23}$$

$$t_6^{(HM)} = \frac{3}{\left(\frac{1}{t_0}+\frac{1}{t_1}+\frac{1}{t_2}\right)} = \frac{3s_y^2}{\left\{1+\frac{s_x^2}{S_x^2}+\exp\left(\frac{S_x^2-s_x^2}{S_x^2+s_x^2}\right)\right\}} \tag{2.24}$$

To the first degree of approximation, the biases and *MSE*s of $t_6^{(AM)}$, $t_6^{(GM)}$ and $t_6^{(HM)}$ are respectively given by

$$Bias\left(t_6^{(AM)}\right)=\left(S_y^2/24n\right)\left(\delta_{04}-1\right)\left(11-12c\right) \tag{2.25}$$

$$Bias\left(t_6^{(GM)}\right)=\left(S_y^2/8n\right)\left(\delta_{04}-1\right)\left(3-4c\right) \tag{2.26}$$

$$Bias\left(t_6^{(HM)}\right)=\left(S_y^2/24n\right)\left(\delta_{04}-1\right)\left(7-12c\right) \tag{2.27}$$

$$MSE\left(t_6^{(AM)}\right)=MSE\left(t_6^{(GM)}\right)=MSE\left(t_6^{(HM)}\right)=n^{-1}S_y^4\left[\left(\delta_{40}-1\right)+\frac{\left(\delta_{04}-1\right)}{4}\left(1-4c\right)\right] \tag{2.28}$$

3. EFFICIENCY COMPARISONS

From (1.4), (1.7), (1.8), (2.7), (2.14), (2.21) and (2.28), we have

$$MSE\left(t_1\right)-MSE\left(s_y^2\right)=\left(S_y^4/n\right)\left(\delta_{04}-1\right)\left(1-2c\right) \tag{3.1}$$

$$MSE\left(t_2\right)-MSE\left(s_y^2\right)=\left(S_y^4/4n\right)\left(\delta_{04}-1\right)\left(1-4c\right) \tag{3.2}$$

$$MSE\left(t_3^{(j)}\right)-MSE\left(s_y^2\right)=\left(S_y^4/4n\right)\left(\delta_{04}-1\right)\left(1-4c\right) \tag{3.3}$$

$$MSE\left(t_4^{(j)}\right)-MSE\left(s_y^2\right)=\left(S_y^4/16n\right)\left(\delta_{04}-1\right)\left(1-8c\right) \tag{3.4}$$

$$MSE\left(t_5^{(j)}\right)-MSE\left(s_y^2\right)=\left(3S_y^4/16n\right)\left(\delta_{04}-1\right)\left(3-8c\right) \tag{3.5}$$

$$MSE\left(t_6^{(j)}\right)-MSE\left(s_y^2\right)=\left(S_y^4/4n\right)\left(\delta_{04}-1\right)\left(1-4c\right) \tag{3.6}$$

where (j = AM, GM, HM).

It is observed from (3.1)-(3.6) that the estimators t_1, t_2 and $t_i^{(j)}$, (i = 3, 4, 5, 6; j = AM, GM, HM) are better than usual unbiased estimator s_y^2 respectively if

$$c>\left(1/2\right), \tag{3.7}$$

$$c>\left(1/4\right), \tag{3.8}$$

$$c>\left(1/4\right), \tag{3.9}$$

$$c > (1/8), \tag{3.10}$$

$$c > (3/8), \tag{3.11}$$

$$c > (1/4). \tag{3.12}$$

From (3.7)-(3.12) it follows that the sufficient condition for the estimators t_1, t_2 and $t_i^{(j)}$, (i = 3, 4, 5, 6; j = *AM*, *GM*, *HM*) to be better than the usual unbiased estimator s_y^2 is $c > (1/2)$.
From (1.7), (1.8), (2.7), (2.14), (2.21) and (2.28), we have

$$MSE(t_2) - MSE(t_1) = (S_y^4/n)(\delta_{04} - 1)\left(c - \frac{3}{4}\right), \tag{3.13}$$

$$MSE\left(t_3^{(j)}\right) - MSE(t_1) = (S_y^4/n)(\delta_{04} - 1)\left(c - \frac{3}{4}\right), \tag{3.14}$$

$$MSE\left(t_4^{(j)}\right) - MSE(t_1) = (S_y^4/n)(\delta_{04} - 1)\left(c - \frac{5}{8}\right), \tag{3.15}$$

$$MSE\left(t_5^{(j)}\right) - MSE(t_1) = (S_y^4/2n)(\delta_{04} - 1)\left(c - \frac{7}{8}\right), \tag{3.16}$$

$$MSE\left(t_6^{(j)}\right) - MSE(t_1) = (S_y^4/n)(\delta_{04} - 1)\left(c - \frac{3}{4}\right), \tag{3.17}$$

where (j = AM, GM, HM).
From (3.13)-(3.17), we get that the estimators .t_2. and $t_i^{(j)}$, (i = 3, 4, 5, 6; j = *AM*, *GM*, *HM*) are better than the usual ratio estimator t_1 respectively if

$$c < (3/4), \tag{3.18}$$

$$c < (3/4) \tag{3.19}$$

$$c < (5/8) \tag{3.20}$$

$$c < (7/8) \tag{3.21}$$

$$c < (3/4) \tag{3.22}$$

Thus it follows from (3.18)-(3.22) that the sufficient condition that the estimators t_2 and $t_i^{(j)}$, (i =

3, 4, 5, 6; j = *AM*, *GM*, *HM*) to be better than the usual ratio estimator t_1 if $c < (5/8)$.
From (1.8), (2.7), (2.14), (2.21) and (2.28), we have

$$MSE\left(t_3^{(j)}\right) - MSE\left(t_2\right) = 0, \tag{3.23}$$

$$MSE\left(t_4^{(j)}\right) - MSE\left(t_2\right) = \left(S_y^4 / 16n\right)\left(\delta_{04} - 1\right)\left(8c - 3\right), \tag{3.24}$$

$$MSE\left(t_5^{(j)}\right) - MSE\left(t_2\right) = \left(S_y^4 / 16n\right)\left(\delta_{04} - 1\right)\left(5 - 8c\right), \tag{3.25}$$

$$MSE\left(t_6^{(j)}\right) - MSE\left(t_2\right) = 0, \tag{3.26}$$

where (j = *AM*, *GM*, *HM*).
It is observed from (3.23) and (3.26) that the estimators t_2, $t_3^{(j)}$ and $t_6^{(j)}$, (j = *AM*, *GM*, *HM*) are equally efficient. Further it is also observed from (3.24) and (3.25) that the estimators $t_4^{(j)}$ and $t_5^{(j)}$, (j = *AM*, *GM*, *HM*) are better than the estimators t_2 respectively if

$$c < (3/8), \tag{3.27}$$

$$c > (5/8). \tag{3.28}$$

From (2.7), (2.14) and (2.21), we have

$$MSE\left(t_4^{(j)}\right) - MSE\left(t_3^{(j)}\right) = \left(S_y^4 / 16n\right)\left(\delta_{04} - 1\right)\left(8c - 3\right), \tag{3.29}$$

$$MSE\left(t_5^{(j)}\right) - MSE\left(t_3^{(j)}\right) = \left(S_y^4 / 16n\right)\left(\delta_{04} - 1\right)\left(5 - 8c\right), \tag{3.30}$$

where (j = *AM*, *GM*, *HM*).
We note from (3.29) and (3.30) that the estimators $t_4^{(j)}$ and $t_5^{(j)}$ are better than the estimators $t_3^{(j)}$, (j = *AM*, *GM*, *HM*) respectively if

$$c < (3/8), \tag{3.31}$$

$$c > (5/8). \tag{3.32}$$

From (2.14), (2.21) and (2.28), we have

$$MSE\left(t_5^{(j)}\right) - MSE\left(t_4^{(j)}\right) = \left(S_y^4 / 2n\right)\left(\delta_{04} - 1\right)\left(1 - 2c\right), \tag{3.33}$$

$$MSE\left(t_6^{(j)}\right) - MSE\left(t_4^{(j)}\right) = \left(S_y^4 / 16n\right)\left(\delta_{04} - 1\right)\left(3 - 8c\right), \tag{3.34}$$

where (j = *AM*, *GM*, *HM*).

Expressions (3.33) and (3.34) indicate that estimators $t_5^{(j)}$ and $t_6^{(j)}$ are better than $t_4^{(j)}$, (j = *AM*, *GM*, *HM*) respectively if

$$c > (1/2), \tag{3.35}$$

$$c > (3/8). \tag{3.36}$$

It is obvious from (3.35) and (3.36) that $c > (1/2)$ is a sufficient condition for the estimators $t_5^{(j)}$ and $t_6^{(j)}$ to be more efficient than the estimators $t_4^{(j)}$, (j = *AM*, *GM*, *HM*).
Further from (2.21) and (2.28), we have

$$MSE\left(t_5^{(j)}\right) - MSE\left(t_6^{(j)}\right) = \left(S_y^4 / 16n\right)\left(\delta_{04} - 1\right)\left(5 - 8c\right), \tag{3.37}$$

which is positive if

$$c < (5/8). \tag{3.38}$$

Thus the estimators $t_6^{(j)}$ are better than $t_5^{(j)}$, (j = *AM*, *GM*, *HM*) if $c < (5/8)$.

4. EMPIRICAL STUDY

To have tangible idea about the absolute relative efficiency of the proposed estimators over their competitors numerically, we consider three natural population data sets. The descriptions of the population data sets are given in Table 1.

TABLE 1:
Population data sets.

Population	N	n	δ_{40}	δ_{04}	δ_{22}	c
I: Murthy (1967, p. 226) Y : Output. X : Number of workers.	80	10	2.2667	3.65	2.3377	0.5408
II: Murthy (1967, p. 127) Y : Cultivated area (in acres). X : Area in square miles.	80	10	2.3730	2.0193	1.6757	0.6630
III: Sukhatme and Sukhatme (1970, p. 185) Y : Area under wheat in 1937 (in acres). X : Cultivated area in 1931.	80	10	3.5469	3.2816	2.6601	0.7276

We have computed the percent absolute relative biases (*PARB*s) of different estimators of S_y^2 by using the formula:

$$PARB(t) = \left|\frac{Bias(t)}{\left(S_y^2 / n\right)}\right| \times 100, \quad (4.1)$$

where $t = t_1, t_2$ and $t_i^{(j)}$, (i = 2, 3, 4, 5, 6; j = *AM*, *GM*, *HM*).
Further we have computed the percent relative efficiencies (*PRE*s) of the various estimators of S_y^2 with respect to usual unbiased estimator s_y^2 by using the formula:

$$PRE\left(t, s_y^2\right) = \frac{MSE\left(s_y^2\right)}{MSE(t)} \times 100, \quad (4.2)$$

where $t = t_1, t_2$ and $t_i^{(j)}$, (i = 2, 3, 4, 5, 6; j = *AM*, *GM*, *HM*) and findings are summarized in Tables 2 and 3.
It is observed from Tables 2 and 3 that

(i) The proposed estimator $t_4^{(AM)}$ has least *PARB* bias comparatively to other estimators in all the three populations. It follows that the proposed estimator $t_4^{(AM)}$ is less biased among all the estimators.

(ii) The proposed estimators $t_3^{(j)}$ and $t_6^{(j)}$ have largest *PRE* (= 214.15) at par with the (Singh *et al*., 2009) estimator t_2 in the population I. However the *PARB* of the estimator $t_3^{(HM)}$ is less than t_1, t_2, $t_3^{(AM)}$, $t_3^{(GM)}$, $t_4^{(GM)}$, $t_5^{(j)}$ and $t_6^{(j)}$, (j = *AM*, *GM*, *HM*). So the estimator $t_3^{(HM)}$ is to be preferred over s_y^2, t_1, t_2, $t_3^{(AM)}$, $t_3^{(GM)}$, $t_4^{(j)}$, $t_5^{(j)}$ and $t_6^{(j)}$, (j = *AM*, *GM*, *HM*) in the population data set I.

(iii) The proposed estimators $t_5^{(j)}$ have larger efficiency than the estimators s_y^2, t_1, t_2, $t_3^{(j)}$, $t_4^{(j)}$ and $t_6^{(j)}$, (j = *AM*, *GM*, *HM*) in both population data sets II and III.

TABLE 2:
The PARBs of different estimators.

Estimator	Population		
	I	II	III
	PARB	PARB	PARB
t_1	131.23	34.36	62.15
t_2	32.49	35.51	20.44
$t_3^{(AM)}$	65.62	17.18	31.08
$t_3^{(GM)}$	32.49	4.44	2.56
$t_3^{(HM)}$	0.64	8.30	25.97
$t_4^{(AM)}$	0.16	0.02	0.01
$t_4^{(GM)}$	7.96	0.97	5.85
$t_4^{(HM)}$	0.32	4.15	12.98
$t_5^{(AM)}$	81.86	19.40	32.35
$t_5^{(GM)}$	73.58	16.21	25.22
$t_5^{(HM)}$	65.30	13.03	18.09
$t_6^{(AM)}$	54.57	12.93	21.57
$t_6^{(GM)}$	32.49	4.44	2.56
$t_6^{(HM)}$	10.41	4.06	16.46

TABLE 3:
The PREs of different estimators.

Estimator	Population		
	I	II	III
	PRE	PRE	PRE
s_y^2	100.00	100.00	100.00
t_1	102.05	131.91	168.86
t_2	214.15	144.20	174.78
$t_3^{(j)}$	214.15	144.20	174.78
$t_4^{(j)}$	165.91	124.95	136.97
$t_5^{(j)}$	168.72	147.19	190.05
$t_6^{(j)}$	214.15	144.20	174.78

$(j = AM, GM, HM)$

5. CONCLUSION

In this article we have suggested some estimators of population variance which are based on arithmetic mean, geometric mean and harmonic mean of the usual unbiased, usual ratio and (Singh *et al.,* 2009) estimators. The expressions of bias and mean squared error of proposed estimators have been derived up to first degree of approximation. The theoretical conditions under which the proposed estimators are more efficient than usual unbiased, usual ratio and (Singh *et al.,* 2009) estimators have been obtained. The empirical study shows that in the performances of proposed estimators are better than the other existing estimators i.e. the proposed estimators have larger *PRE*s and least *PARB*s in comparisons of others. Thus we recommend the proposed estimators in practice. However this conclusion can not be extrapolated due to limited empirical study.

ACKNOWLEDGEMENT

The authors are thankful to the learned referee for his constructive suggestions regarding improvement of the present paper.

REFERENCES

S. BAHL, R. K, TUTEJA (1991). *Ratio and product type exponential estimators.* Journal of Information and Optimization Sciences, 12(1): 159-164.

R. S. BIRADAR, H. P. SINGH (1998). *Predictive estimation of finite population variance.* Calcutta Statistical Association Bulletin, 48: 229-235.

A. A. Cebrian, M. R. Garcia (1997). *Variance estimation using auxiliary information-an almost unbiased multivariate ratio estimator*. Metrika, 45: 171-178.

A. K. DAS, T. P. TRIPATHI (1978). *Use of auxiliary information in estimating the finite population variance.* Sankhya, C 40: 139-148.

L. K. GROVER (2007). *A wide and efficient class of estimators of population variance under sub-sampling scheme*. Model Assisted Statistics and Applications, 2(1): 41-51.

C. T. ISAKI (1983). *Variance estimation using auxiliary information.* Journal of the American Statistical Association, 78(381): 117-123.

C. KADILAR, H. CINGI (2007). *Improvement in Variance Estimation in Simple Random Sampling.* Communications in Statistics-Theory and Methods, 36(11): 2075-2081.

M. N. MURTHY (1967). *Sampling: Theory and Methods*. Statistical Publishing Society, Calcutta, India.

J. SHABBIR, S. GUPTA (2007). *On improvement in variance estimation using auxiliary information.* Communications in Statistics- Theory and Methods, 36(12): 2177-2185.

H. P. SINGH, R. S. SOLANKI (2009-2010). *Estimation of finite population variance using auxiliary information in presence of random non-response.* Gujarat Statistical Review, 36 & 37(1-2): 46-58.

H. P. SINGH, R. S. SOLANKI (2013a). *A new procedure for variance estimation in simple random sampling using auxiliary information.* Statistical Papers, 54(2): 479-497.

H. P. SINGH, R. S. SOLANKI. (2013b). *Improved estimation of finite population variance using auxiliary information*. Communications in Statistics-Theory & Methods, 42(15): 2718-2730.

H. P. SINGH, G. K. VISHWAKARMA (2008). *Some families of estimators of variance of stratified random sample mean using auxiliary information.* Journal of Statistical Theory and Practice, 2(1): 21-43.

H. P. SINGH, G. K. VISHWAKARMA. (2009). *Some estimators of finite population mean using auxiliary information in sample surveys.* Journal of Applied Statistical Science, 16(4): 397-409.

H. P. SINGH, P. CHANDRA, S. SINGH (2003). *Variance estimation using multiauxiliary information for random non-response in survey sampling.* Statistica, 63(1): 23-40.

H. P. SINGH, L. N. UPADHYAYA, U.D. NAMJOSHI (1988). *Estimation of finite population variance.* Current Science, 57(24): 1331-1334.

R. SINGH, P. CHAUHAN, N. SAWAN, F. SMARANDACHE (2009). *Improved exponential estimator for population variance using two auxiliary variables.* Octogon Mathematical Magazine, 17(2): 667-674.

R. S. SOLANKI, H. P. SINGH (2013). *An improved class of estimators for the population variance.* Model Assisted Statistics and Applications, (in press).

S. K. SRIVASTAVA, H. S. JHAJJ (1980). *A class of estimators using auxiliary information for estimating finite population variance.* Sankhya, C 42: 87-96.

P. V. SUKHATME, B. V. SUKHATME (1970). *Sampling Theory of Surveys with Applications.* Iowa State University Press, Ames, Iowa.

SUMMARY

Estimation of finite population variance using auxiliary information in sample surveys

This paper addresses the problem of estimating the finite population variance using auxiliary information in sample surveys. Motivated by (Singh and Vishwakarma, 2009) some estimators of finite population variance have been suggested along with their properties in simple random sampling. The theoretical conditions under which the proposed estimators are more efficient than usual unbiased, usual ratio and (Singh *et al.*, 2009) estimators have been obtained. Numerical illustrations are given in support of the present study.

Keywords: Study variable; Auxiliary variable; Arithmetic mean; Geometric mean; Harmonic mean; Bias; Mean squared error

APPENDIX I

To obtain the biases and mean squared errors of the estimators $t_i^{(AM)}$, $t_i^{(GM)}$ and $t_i^{(HM)}$, (i = 3, 4, 5, 6) we define the following classes of estimators of population variance S_y^2 as

$$\begin{aligned} t_{AM} &= \left(\alpha_0 t_0 + \alpha_1 t_1 + \alpha_2 t_2\right), \quad \left(\alpha_0 + \alpha_1 + \alpha_2\right) = 1 \\ &= s_y^2\left[\alpha_0 + \alpha_1\left(\frac{S_x^2}{s_x^2}\right) + \alpha_2 \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right] \end{aligned} \tag{I.1}$$

$$\begin{aligned} t_{GM} &= \left(t_0^{\alpha_0} t_1^{\alpha_1} t_2^{\alpha_2}\right), \quad \left(\alpha_0 + \alpha_1 + \alpha_2\right) = 1 \\ &= s_y^2\left(\frac{S_x^2}{s_x^2}\right)^{\alpha_1} \exp\left(\frac{\alpha_1\left(S_x^2 - s_x^2\right)}{S_x^2 + s_x^2}\right) \end{aligned} \tag{I.2}$$

$$\begin{aligned} t_{HM} &= \frac{1}{\left(\dfrac{\alpha_0}{t_0} + \dfrac{\alpha_1}{t_1} + \dfrac{\alpha_2}{t_2}\right)}, \quad \left(\alpha_0 + \alpha_1 + \alpha_2\right) = 1 \\ &= \frac{s_y^2}{\left[\alpha_0 + \alpha_1\left(\dfrac{s_x^2}{S_x^2}\right) + \alpha_2 \exp\left(\dfrac{s_x^2 - S_x^2}{S_x^2 + s_x^2}\right)\right]} \end{aligned} \tag{I.3}$$

We write

$$s_y^2 = S_y^2\left(1 + e_0\right) \text{ and } s_x^2 = S_x^2\left(1 + e_1\right)$$

such that

$$E\left(e_0\right) = E\left(e_1\right) = 0$$

and to the first degree of approximation (ignoring finite population correction term)

$$\left.\begin{aligned} E\left(e_0^2\right) &= n^{-1}\left(\delta_{40} - 1\right) \\ E\left(e_1^2\right) &= n^{-1}\left(\delta_{04} - 1\right) \\ E\left(e_0 e_1\right) &= n^{-1}\left(\delta_{04} - 1\right)c \end{aligned}\right\},$$

where

$$c = \left(\delta_{22} - 1\right)\left(\delta_{04} - 1\right)^{-1}, \quad \delta_{pq} = \left(\mu_{pq} / \mu_{20}^{p/2}\mu_{02}^{q/2}\right), \quad \mu_{pq} = N^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^p\left(x_i - \bar{X}\right)^q$$

(p, q being non-negative integers).

Expressing (I.1) in terms of *e*'s we have

$$t_{AM}=S_y^2\left(1+e_0\right)\left[\alpha_0+\alpha_1\left(1+e_1\right)^{-1}+\alpha_2\exp\left(\frac{-e_1}{2+e_1}\right)\right]$$

$$=S_y^2\left(1+e_0\right)\left[\alpha_0+\alpha_1\left(1-e_1+e_1^2-\ldots\right)+\alpha_2\left\{1-\frac{e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}+\frac{e_1^2}{8}\left(1+\frac{e_1}{2}\right)^{-2}-\ldots\right\}\right]$$

$$=S_y^2\left(1+e_0\right)\left[1+\alpha_1\left(-e_1+e_1^2-\ldots\right)+\alpha_2\left\{-\frac{e_1}{2}\left(1-\frac{e_1}{2}+\ldots\right)+\frac{e_1^2}{8}\left(1+\frac{e_1}{2}\right)^{-2}-\ldots\right\}\right]$$

$$=S_y^2\left(1+e_0\right)\left[1-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1+\left(\alpha_1+\frac{\alpha_2}{4}+\frac{\alpha_2}{8}\right)e_1^2-\ldots\right]$$

$$\cong S_y^2\left[1+e_0-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_0e_1+\left(\alpha_1+\frac{3\alpha_2}{8}\right)e_1^2\right]$$

or

$$\left(t_{AM}-S_y^2\right)\cong S_y^2\left[e_0-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_0e_1+\left(\alpha_1+\frac{3\alpha_2}{8}\right)e_1^2\right]. \tag{I.4}$$

Taking expectation of both sides of (I.4), we get the bias of t_{AM} to the first degree of approximation as

$$Bias\left(t_{AM}\right)=n^{-1}S_y^2\left(\delta_{04}-1\right)\left[\left(\alpha_1+\frac{3\alpha_2}{8}\right)-\left(\alpha_1+\frac{\alpha_2}{2}\right)c\right] \tag{I.5}$$

Squaring both sides of (I.4) and neglecting terms of *e*'s having power greater than two we have

$$\left(t_{AM}-S_y^2\right)^2=S_y^4\left[e_0^2+\left(\alpha_1+\frac{\alpha_2}{2}\right)^2e_1^2-2\left(\alpha_1+\frac{\alpha_2}{2}\right)e_0e_1\right] \tag{I.6}$$

Taking expectation of both sides of (I.6) we get the mean squared error of t_{AM} to the first degree of approximation as

$$MSE\left(t_{AM}\right)=n^{-1}S_y^4\left[\left(\delta_{40}-1\right)+\left(\alpha_1+\frac{\alpha_2}{2}\right)\left(\delta_{04}-1\right)\left(\alpha_1+\frac{\alpha_2}{2}-2c\right)\right] \tag{I.7}$$

Putting $\left(\alpha_0,\alpha_1,\alpha_2\right)=\left(\frac{1}{2},\frac{1}{2},0\right),\left(\frac{1}{2},0,\frac{1}{2}\right),\left(0,\frac{1}{2},\frac{1}{2}\right)$ and $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ in (I.5) and (I.7) one can easily get the biases and mean squared errors of the estimators $t_i^{(AM)}$, (i = 3, 4, 5, 6) respectively given in ((2.4), (2.7)); ((2.11), (2.14));((2.18), (2.21)) and ((2.25), (2.28)).

Expressing the class of estimators t_{GM} in terms of *e*'s, we have

$$t_{GM} = S_y^2\left(1+e_0\right)\left(1+e_1\right)^{-\alpha_1}\exp\left(\frac{-\alpha_2 e_1}{2+e_1}\right)$$

$$= S_y^2\left(1+e_0\right)\left(1+e_1\right)^{-\alpha_1}\exp\left(\frac{-\alpha_2 e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}\right)$$

$$= S_y^2\left(1+e_0\right)\left\{1-\alpha_1 e_1+\frac{\alpha_1\left(\alpha_1+1\right)}{2}e_1^2-\ldots\right\}\left\{1-\frac{\alpha_2 e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}+\frac{\alpha_2^2 e_1^2}{8}\left(1+\frac{e_1}{2}\right)^{-2}-\ldots\right\}$$

$$= S_y^2\left(1+e_0\right)\left\{1-\alpha_1 e_1+\frac{\alpha_1\left(\alpha_1+1\right)}{2}e_1^2-\ldots\right\}\left\{1-\frac{\alpha_2 e_1}{2}\left(1-\frac{e_1}{2}+\ldots\right)+\frac{\alpha_2^2 e_1^2}{8}\left(1-e_1+\ldots\right)\right\}$$

$$= S_y^2\left[1+e_0-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_0 e_1+\left(\frac{\alpha_1\alpha_2}{2}+\frac{\alpha_1\left(\alpha_1+1\right)}{2}+\frac{\alpha_2\left(2+\alpha_2\right)}{8}\right)e_1^2+\ldots\right.$$

or

$$\left(t_{GM}-S_y^2\right)\cong S_y^2\left[e_0-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_0 e_1+\left(\frac{\alpha_1\alpha_2}{2}+\frac{\alpha_1\left(\alpha_1+1\right)}{2}+\frac{\alpha_2\left(2+\alpha_2\right)}{8}\right)e_1^2\right] \quad (I.8)$$

Taking the expectation of both sides of (I.8), we get the bias of t_{GM} to the first degree of approximation as

$$Bias\left(t_{GM}\right)=\left(S_y^2/8n\right)\left(\delta_{04}-1\right)\left[4\alpha_1\left(\alpha_1+1\right)+4\alpha_1\alpha_2+\alpha_2\left(2+\alpha_2\right)-8\left(\alpha_1+\frac{\alpha_2}{2}\right)c\right]$$
$$=\left(S_y^2/2n\right)\left(\delta_{04}-1\right)\left(\alpha_1+\frac{\alpha_2}{2}\right)\left(\alpha_1+\frac{\alpha_2}{2}+1-2c\right) \quad (I.9)$$

Squaring both sides of (I.8) and neglecting terms of *e*'s having power greater than two we have

$$\left(t_{GM}-S_y^2\right)^2=S_y^4\left[e_0^2+\left(\alpha_1+\frac{\alpha_2}{2}\right)^2 e_1^2-2\left(\alpha_1+\frac{\alpha_2}{2}\right)e_0 e_1\right] \quad (I.10)$$

Taking expectation of both sides of (I.10), we get the mean squared error of t_{GM} to the first degree of approximation as

$$MSE\left(t_{GM}\right)=\left(S_y^4/n\right)\left[\left(\delta_{40}-1\right)+\left(\delta_{04}-1\right)\left(\alpha_1+\frac{\alpha_2}{2}\right)\left(\alpha_1+\frac{\alpha_2}{2}-2c\right)\right] \quad (I.11)$$

Substituting the values of $\left(\alpha_0,\alpha_1,\alpha_2\right)=\left(\frac{1}{2},\frac{1}{2},0\right),\left(\frac{1}{2},0,\frac{1}{2}\right),\left(0,\frac{1}{2},\frac{1}{2}\right)$ and $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ in (I.9) and (I.11) one can easily get the biases and mean squared errors of the estimators $t_i^{(GM)}$, (i = 3, 4,

5, 6) respectively given in ((2.5), (2.7)); ((2.12), (2.14));((2.19), (2.21)) and ((2.26), (2.28)).
Expressing the class of estimators t_{HM} in terms of e's we have

$$t_{HM} = \frac{S_y^2(1+e_0)}{\left[\alpha_0 + \alpha_1(1+e_1) + \alpha_2 \exp\left(\frac{e_1}{2+e_1}\right)\right]}$$

$$= \frac{S_y^2(1+e_0)}{\left[\alpha_0 + \alpha_1(1+e_1) + \alpha_2 \exp\left(\frac{e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}\right)\right]}$$

$$= \frac{S_y^2(1+e_0)}{\left[\alpha_0 + \alpha_1(1+e_1) + \alpha_2\left\{1+\frac{e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1} + \frac{e_1^2}{8}\left(1+\frac{e_1}{2}\right)^{-2} + \dots\right\}\right]}$$

$$= \frac{S_y^2(1+e_0)}{\left[1+\alpha_1 e_1 + \alpha_2\left\{\frac{e_1}{2}\left(1-\frac{e_1}{2}+\dots\right) + \frac{e_1^2}{8}(1-e_1+\dots)\right\}\right]}$$

$$= \frac{S_y^2(1+e_0)}{\left[1+\alpha_1 e_1 + \alpha_2\left(\frac{e_1}{2} - \frac{e_1^2}{8} + \dots\right)\right]}$$

$$= S_y^2(1+e_0)\left[1+\left\{\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1 - \frac{\alpha_2}{8}e_1^2 + \dots\right\}\right]^{-1}$$

$$= S_y^2(1+e_0)\left[1-\left\{\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1 - \frac{\alpha_2}{8}e_1^2 + \dots\right\} + \left(\alpha_1+\frac{\alpha_2}{2}\right)^2 e_1^2 + \dots\right]$$

$$= S_y^2(1+e_0)\left[1-\left(\alpha_1+\frac{\alpha_2}{2}\right)e_1 + \left\{\frac{\alpha_2}{8} + \left(\alpha_1+\frac{\alpha_2}{2}\right)^2\right\}e_1^2 + \dots\right]$$

$$\cong S_y^2\left[1+e_0 - \left(\alpha_1+\frac{\alpha_2}{2}\right)e_1 - \left(\alpha_1+\frac{\alpha_2}{2}\right)e_0 e_1 + \left\{\frac{\alpha_2}{8} + \left(\alpha_1+\frac{\alpha_2}{2}\right)^2\right\}e_1^2\right]$$

or

$$\left(t_{HM} - S_y^2\right) \cong S_y^2\left[e_0 - \left(\alpha_1+\frac{\alpha_2}{2}\right)e_1 - \left(\alpha_1+\frac{\alpha_2}{2}\right)e_0 e_1 + \left\{\frac{\alpha_2}{8} + \left(\alpha_1+\frac{\alpha_2}{2}\right)^2\right\}e_1^2\right] \quad \text{(I.12)}$$

Taking expectation of both sides of (I.12), we get the bias of t_{HM} to the first degree of approximation as

$$Bias(t_{HM}) = (S_y^2 / n)(\delta_{04} - 1)\left[\frac{\alpha_2}{8} + \left(\alpha_1 + \frac{\alpha_2}{2}\right)^2 - \left(\alpha_1 + \frac{\alpha_2}{2}\right)c\right]$$

$$= (S_y^2 / n)(\delta_{04} - 1)\left[\frac{\alpha_2}{8} + \left(\alpha_1 + \frac{\alpha_2}{2}\right)\left(\alpha_1 + \frac{\alpha_2}{2} - c\right)\right] \quad (I.13)$$

Squaring both sides of (I.12) and neglecting terms of e's having power greater than two we have

$$\left(t_{HM} - S_y^2\right)^2 = S_y^4\left[e_0^2 + \left(\alpha_1 + \frac{\alpha_2}{2}\right)^2 e_1^2 - 2\left(\alpha_1 + \frac{\alpha_2}{2}\right)e_0 e_1\right] \quad (I.14)$$

Taking expectation of both sides of (I.14), we get the mean squared error of t_{HM} to the first degree of approximation as

$$MSE(t_{HM}) = (S_y^4 / n)\left[(\delta_{40} - 1) + (\delta_{04} - 1)\left(\alpha_1 + \frac{\alpha_2}{2}\right)\left(\alpha_1 + \frac{\alpha_2}{2} - 2c\right)\right] \quad (I.15)$$

Substituting the values of $(\alpha_0, \alpha_1, \alpha_2) = \left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{2}, 0, \frac{1}{2}\right), \left(0, \frac{1}{2}, \frac{1}{2}\right)$ and $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ in (I.13) and (I.15) one can easily get the biases and mean squared errors of the estimators $t_i^{(GM)}$, (i = 3, 4, 5, 6) respectively given in ((2.6), (2.7)); ((2.13), (2.14));((2.20), (2.21)) and ((2.27), (2.28)).