Available online at www.elixirpublishers.com (Elixir International Journal)

Statistics

Elixir Statistics 89 (2015) 36451-36457

ISSN: 2229-712X

A New Exponential Ratio-Type Estimator for Population Variance with Linear Combination of Two Auxiliary Attributes

Nitesh K. Adichwal[1,*] Rajesh Singh[1] and Yashpal Singh Raghav[2]

[1]Department of Statistics, Banaras Hindu University, Varanasi, India.

[2]Department of Mathematics, Jazan University, Jazan, Kingdom of Saudi Arabia.

ARTICLE INFO

Article history:
Received: 7 September 2015;
Received in revised form:
15 November 2015;
Accepted: 20 November 2015;

Keywords
Auxiliary Information,
Auxiliary Attribute,
Simple Random Sampling,
Mean Square Error.

ABSTRACT

This paper suggests a new exponential ratio-type estimator for estimating the population variance using information of two auxiliary attributes in SRSWOR based on the adaption of the estimator presented by Lu et al. (2014). Problem is extended to the case of two phase sampling. The expressions for the mean square error of the proposed estimator have been derived. The proposed estimator has been compared theoretically with the usual unbiased estimator, usual ratio type estimator, estimator proposed by Chauhan et al. (2009) and the estimator proposed by Singh et al. (2009) and the conditions under which the proposed estimator are better than some existing estimators have also been given. An empirical study has also been carried out using two population data sets to demonstrate the efficiencies of the proposed estimator.

© 2015 Elixir All rights reserved.

Introduction

In the theory of sample surveys, the utilization of auxiliary information is frequently acknowledged to the higher accuracy of the estimation of population characteristics. Ratio, regression and product methods of estimation are good example of it. There exists many situations when the information is available in the form of attribute which is highly correlated with y. In literature of sampling survey many authors have suggested estimator based on auxiliary information. However in many situations of practical importance, instead of an auxiliary variable x there exists an attribute (say ϕ) which is highly correlated with study variable y. In these situations by taking the advantage of point bi-serial correlation between the study variable y and the auxiliary attribute ϕ , the efficient estimator of population parameters of interest can be constructed.

Sometimes it may possible that information of two auxiliary attributes are available instead of one auxiliary attribute. In such situations it is applicable to use two auxiliary attributes for estimation of population parameter of interest. While estimating the population variance S_y^2 of the study variable y, we can use the information of the two auxiliary attributes X_1 and X_2.

Singh and kumar (2011), Singh and Malik (2014), Adichwal et al. (2015) proposed some estimators of population variance using an auxiliary attribute under different situation. This paper suggests a new exponential ratio-type estimator for estimating the population variance using information of two auxiliary attributes in SRSWOR.

Let us consider a sample of size n is drawn by SRSWOR from a population of size N. Further let a study variate y and auxiliary attributes X_1 and X_2 taking the values y_i and x_{1i} and x_{2i} respectively, on the unit U_i (i=1,2,....N), It is assume that attributes X_1 and X_2 takes only two values viz, 0 and 1 according as x_{1i} and x_{2i}=1, if ith unit of the population possesses and attribute x_{1i} and x_{2i}=0, if otherwise.

To estimate $S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(y_i-\bar{Y})^2$, it is assumed that $S_{x1}^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(x_{1i}-\bar{X}_1)^2$ and $S_{x2}^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(x_{2i}-\bar{X}_2)^2$ are known. Assume that population of size N is large so that the finite population correlation terms are ignored.

Following notations are used for deriving the properties of the estimators:

$$E(s'^2_{x1}-S^2_{x1}) = \frac{S^4_{x1}}{n'}(\partial_{040}-1), \quad V(s^2_{x1}) = \frac{S^4_{x1}}{n}(\partial_{040}-1), \quad V(s^2_{y}) = \frac{S^4_{y}}{n}(\partial_{400}-1), \quad V(s^2_{x2}) = \frac{S^4_{x2}}{n}(\partial_{004}-1),$$

$$s'^2_{x1} = \frac{1}{n'-1}\sum_{i=1}^{n'}(x_{1i}-\bar{x}'_1)$$

Tele:
E-mail addresses: nnitesh139@gmail.com

© 2015 Elixir All rights reserved

$$Cov(S_{x1}^2, S_y^2) = \frac{S_{x1}^2 S_y^2}{n}(\partial_{220} - 1),\ Cov(S_{x2}^2, S_y^2) = \frac{S_{x2}^2 S_y^2}{n}(\partial_{202} - 1),\ Cov(S_{x1}^2, S_{x2}^2) = \frac{S_{x1}^2 S_{x2}^2}{n}(\partial_{022} - 1)$$

where,

$$\partial_{pqr} = \frac{\mu_{pqr}}{(\mu_{200}^{p/2}\mu_{020}^{q/2}\mu_{002}^{r/2})},$$

$$\mu_{pqr} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \overline{Y})^p (x_i - \overline{X})^q (z_i - \overline{Z})^r$$; p, q, r being the non-negative integers.

Estimators in Literature

In order to have an estimate of population variance of study variable y, Isaki (1983) proposed the following estimator

$$t_0 = s_y^2 \frac{S_y^2}{s_y^2} \tag{2.1}$$

Assuming the knowledge of auxiliary variables x and z, Singh et al. (2009) proposed the following estimators for estimating s_y^2.

$$t_1 = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right] \tag{2.2}$$

$$t = s_y^2\left[\alpha \exp\left\{\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right\} + (1-\alpha)\exp\left\{\frac{s_z^2 - S_z^2}{s_z^2 + S_z^2}\right\}\right] \tag{2.3}$$

Where α is a constant to be determined by

$$\alpha = \frac{\{\partial_{004} + 2(\partial_{220} + \partial_{202}) + \partial_{022} - 6\}}{(\partial_{040} + \partial_{004} + 2\partial_{022} - 4)}$$

Here we consider, the case when study variable y is a simple variable and auxiliary variables x and z are available in the form of attributes X_1 and X_2 respectively.

The MSE expressions of t_1, t_2 and t when auxiliary attribute is used, are respectively given by

$$MSE(t_1) = \frac{S_y^2}{n}\left[\partial_{400} + \frac{\partial_{040}}{4} - \partial_{220} + \frac{1}{4}\right] \tag{2.4}$$

$$MSE(t) = \frac{S_y^4}{n}\left[(\partial_{400} - 1) + \frac{\alpha^2}{4}(\partial_{040} - 1) + \frac{(1-\alpha)^2}{4}(\partial_{004} - 1) - \alpha(\partial_{220} - 1) + (1-\alpha)(\partial_{202} - 1)\right.$$
$$\left. - \frac{\alpha(1-\alpha)}{2}(\partial_{022} - 1)\right] \tag{2.5}$$

Proposed Estimator

Motivated by Lu et al. (2014), we have proposed a new exponential ratio-type estimator using linear combination of two auxiliary attributes X_1 and X_2 for estimating population variance of study variable y. The proposed estimator is

$$t = s_y^2 \exp\left[\frac{s_*^2 - S_*^2}{s_*^2 + S_*^2}\right] \tag{3.1}$$

where,

$s_*^2 = w_1 s_{x1}^2 + w_2 s_{x2}^2$, $S_*^2 = w_1 S_{x1}^2 + w_2 S_{x2}^2$; w_1, w_2 are weights that satisfy the condition: $w_1 + w_2 = 1$.

MSE of the estimator can be found using Taylor series method defined as

$$f(s_y^2, s_x^2, s_z^2) \cong f(S_y^2, S_x^2, S_z^2) + \left.\frac{\partial f}{\partial s_y^2}\right|_{(S_y^2, S_x^2, S_z^2)}(s_y^2 - S_y^2) + \left.\frac{\partial f}{\partial s_x^2}\right|_{(S_y^2, S_x^2, S_z^2)}(s_x^2 - S_x^2)$$
$$+ \left.\frac{\partial f}{\partial s_z^2}\right|_{(S_y^2, S_x^2, S_z^2)}(s_z^2 - S_z^2) \tag{3.2}$$

Where, $f(s_y^2, s_x^2, s_z^2) = t$

$$t_N - S_y^2 = (s_y^2 - S_y^2) - \frac{1}{2}w_1 R_*(s_{x1}^2 - S_{x1}^2) - \frac{1}{2}w_2 R_*(s_{x2}^2 - S_{x2}^2) \tag{3.3}$$

Where, $\dfrac{S_y^2}{w_1 S_{x1}^2 + w_2 S_{x2}^2} = R_*$

The MSE expression of this exponential ratio-type estimator is given by

$$MSE(t_N) = V(s_y^2) + \frac{1}{4} w_1^2 R_*^2 V(s_{x1}^2) + \frac{1}{4} w_2^2 R_*^2 V(s_{x2}^2) - w_1 R_* Cov(s_{x1}^2, s_y^2) - w_2 R_* Cov(s_{x2}^2, s_y^2)$$

$$+ \frac{1}{2} w_1 w_2 R_*^2 Cov(s_{x1}^2, s_{X2}^2)$$

$$= \frac{1}{n}\left\{ S_y^4(\delta_{400}-1) + \frac{1}{4} w_1^2 R_*^2 S_{x1}^4(\delta_{040}-1) + \frac{1}{4} w_2^2 R_*^2 S_{x2}^4(\delta_{004}-1) - w_1 R_* S_{x1}^2 S_y^2(\delta_{220}-1) \right.$$

$$\left. - w_2 R_* S_y^2 S_{x2}^2(\delta_{202}-1) + \frac{1}{2} w_1 w_2 R_*^2 S_{x1}^2 S_{X2}^2(\delta_{022}-1) \right\} \quad (3.4)$$

The optimum value of w_1 and w_2 are given by

$w_1^* = \dfrac{U}{U+V}$, $w_2^* = 1 - w_1^*$

Where $U = 2Cov(s_y^2, s_{x2}^2)S_{x1}^4 - Cov(s_{x1}^2, s_{x2}^2)S_{x1}^2 S_y^2 - 2Cov(s_y^2, s_{x1}^2)S_{x1}^2 S_{x2}^2 + V(s_{x1}^2)S_{x2}^2 S_y^2$

$V = V(s_{x2}^2)S_{x1}^2 S_y^2 - 2Cov(s_y^2, s_{x2}^2)S_{x1}^2 S_{x2}^2 - Cov(s_{x1}^2, s_{x2}^2)S_{x2}^2 S_y^2 + 2Cov(s_y^2, s_{x1}^2)S_{x2}^4$

The expression for minimum MSE of estimator t_N can be written as

$$MSE(t_N)_{min} = \frac{1}{n}\left\{ S_y^4(\delta_{400}-1) + \frac{1}{4} w_1^{*2} R_*'^2 S_{x1}^4(\delta_{040}-1) + \frac{1}{4} w_2^{*2} R_*'^2 S_{x2}^4(\delta_{004}-1) - w_1^* R_*' S_{x1}^2 S_y^2(\delta_{220}-1) \right.$$

$$\left. - w_2^* R_*' S_y^2 S_{x2}^2(\delta_{202}-1) + \frac{1}{2} w_1^* w_2^* R_*'^2 S_{x1}^2 S_{X2}^2(\delta_{022}-1) \right\} \quad (3.5)$$

where $\dfrac{S_y^2}{w_1^* S_{x1}^2 + w_2^* S_{x2}^2} = R_*'$

Efficiency Comparison

In this section we are comparing the minimum MSE of the proposed estimator t_N with other existing estimators. It is known that the variance of the s_y^2 under SRSWOR is given by

$$V(s_y^2) = \frac{S_y^4}{n}\left[\partial_{400} - 1\right].$$

$$V(s_y^2) - MSE(t_N)_{min} = \frac{S_y^4}{n}(\partial_{400}-1) - \frac{1}{n}\left\{ S_y^4(\delta_{400}-1) + \frac{1}{4} w_1^{*2} R_*'^2 S_{x1}^4(\delta_{040}-1) + \frac{1}{4} w_2^{*2} R_*'^2 S_{x2}^4(\delta_{004}-1) \right.$$

$$\left. - w_1^* R_*' S_{x1}^2 S_y^2(\delta_{220}-1) - w_2^* R_*' S_y^2 S_{x2}^2(\delta_{202}-1) + \frac{1}{2} w_1^* w_2^* R_*'^2 S_{x1}^2 S_{X2}^2(\delta_{022}-1) \right\} \geq 0$$

or,

$$w_1^{*2} R_*' S_{x1}^4(\delta_{040}-1) + w_2^{*2} R_*' S_{x2}^4(\delta_{004}-1) \geq 2\left\{ 2w_1^* S_{x1}^2 S_y^2(\delta_{220}-1) + 2w_2^* S_y^2 S_{x2}^2(\delta_{202}-1) \right.$$

$$\left. + 2w_1^* w_2^* R' S_{x1}^2 S_{X2}^2(\delta_{022}-1) \right\} \quad (4.1)$$

$$MSE(t_1) - MSE(t_N)_{min} = S_y^2\left[\frac{\partial_{040}}{4} - \partial_{220} + \frac{5}{4}\right] - \left\{ \frac{1}{4} w_1^{*2} R_*'^2 S_{x1}^4(\delta_{040}-1) + \frac{1}{4} w_2^{*2} R_*'^2 S_{x2}^4(\delta_{004}-1) \right.$$

$$\left. - w_1^* R_*' S_{x1}^2 S_y^2(\delta_{220}-1) - w_2^* R_*' S_y^2 S_{x2}^2(\delta_{202}-1) + \frac{1}{2} w_1^* w_2^* R_*'^2 S_{x1}^2 S_{X2}^2(\delta_{022}-1) \right\} \geq 0 \quad (4.2)$$

$$MSE(t) - MSE(t_N)_{min} = S_y^4\left[\frac{\alpha^2}{4}(\partial_{040}-1) + \frac{(1-\alpha)^2}{4}(\partial_{004}-1) - \alpha(\partial_{220}-1) + (1-\alpha)(\partial_{202}-1) \right.$$

$$-\frac{\alpha(1-\alpha)}{2}(\partial_{022}-1)\Bigg]-\Bigg\{\frac{1}{4}w_1^{*2}R_*'^2S_{x1}^4(\delta_{040}-1)+\frac{1}{4}w_2^{*2}R_*'^2S_{x2}^4(\delta_{004}-1)$$

$$-w_1^*R_*'S_{x1}^2S_y^2(\delta_{220}-1)-w_2^*R_*'S_y^2S_{x2}^2(\delta_{202}-1)+\frac{1}{2}w_1^*w_2^*R_*'^2S_{x1}^2S_{X2}^2(\delta_{022}-1)\Bigg\}\geq 0 \qquad (4.3)$$

When the condition (4.1) to (4.3) are satisfied, our suggested estimator t_N is more efficient as compared to t_1 and t respectively.

Empirical Study

To illustrate the performance of various estimators of S_y^2, we consider the following data sets:

Population I $\left[\text{Source : Singh and Chaudhary (1986), p.177}\right]$.

The population consists of 34 wheats farm in 34 villages in certain region of India. The variables are defined as:

Y :area under wheat crop (in acres) during 1974.

X_1 : proportion of farms under wheat crop which have more than 1000 acres land during 1971 and

X_2 : proportion of farms under wheat crop which have more than 400 acres land during 1973.

For this data we have

N=34, n=10, $n'=20$

$\partial_{400}=3.72569$, $\partial_{040}=1.81666$, $\partial_{004}=4.97241$, $\partial_{220}=1.64178$, $\partial_{202}=3.42213$, $\partial_{022}=2.15862$ $s_{x1}'^2=0.19737$,

$S_y^2=22564.55704$, $S_{x1}^2=0.21390$, $S_{x2}^2=0.12923$.

Population II $\left[\text{Source : Murthy(1967), p.266}\right]$.

y: output, x: proportion of workers more than 400, z: proportion of fixed capital more than 2000,

N=80, n=15, $n'=20$.

$\partial_{400}=2.26665$, $\partial_{040}=2.15301$, $\partial_{004}=3.52303$, $\partial_{220}=1.52506$, $\partial_{202}=1.94997$, $\partial_{022}=2.33036$

$s_{x1}'^2=0.2$, $S_y^2=3369642.209$, $S_{x1}^2=0.20213$, $S_{x2}^2=0.15428$.

Table 5.1. PRE's of various estimators w.r.t. S_y^2

Estimator	PRE	
	Population I	Population II
S_y^2	100.00	100.00
t_1	97.76	82.80
t	162.10	124.03
$t_{N(\min)}$	219.77	137.91

Table 5.1 reveals that the PRE of the proposed estimator $t_{N(\min)}$ along with the PRE's of the existing estimator with respect to S_y^2 for two given population data sets. From table it is clear that the estimator $t_{N(\min)}$ under optimum condition is more efficient than the ratio estimator and the other estimators discussed in this paper. Hence for observed choice of parameter the proposed estimator $t_{N(\min)}$ is more precise as compared to all other estimators discussed here.

Proposed Estimator in two phase sampling

In some situations when S_{x1}^2 is not known, the technique of two phase or double sampling is used. This scheme requires collection of information on x1 and x2 the first phase sample s' of size n' (n' <N) and on y for the second phase sample s of size n (n< n') from the first phase sample.

Jhajj and Walia (2011), Singh and Kumar (2015) proposed some estimators of population variance using two phase sampling. Chauhan et al. (2009) proposed the following estimator for estimating population variance.

$$t_{1d}=s_y^2\exp\left\{\frac{s_{x1}'^2-s_{x1}^2}{s_{x1}'^2+s_{x1}^2}\right\} \qquad (6.1)$$

$$t_{2d}=s_{x2}^2\exp\left\{\frac{s_{x2}'^2-s_{x2}^2}{s_{x2}'^2+s_{x2}^2}\right\} \qquad (6.2)$$

$$t_d=s_y^2\left[k\exp\left\{\frac{s_{x1}'^2-s_{x1}^2}{s_{x1}'^2+s_{x1}^2}\right\}+(1-k)\exp\left\{\frac{s_{x2}'^2-s_{x2}^2}{s_{x2}'^2+s_{x2}^2}\right\}\right] \qquad (6.3)$$

The MSE expressions of t_{1d}, t_{2d} and t_d are respectively given by

$$MSE(t_{1d}) = S_y^4\left\{\frac{1}{n}(\delta_{400}-1)+\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{040}-1)+\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{220}-1)\right\} \quad (6.4)$$

$$MSE(t_{2d}) = S_y^4\left\{\frac{1}{n}(\delta_{400}-1)+\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{004}-1)-\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{202}-1)\right\} \quad (6.5)$$

$$MSE(t_d)_{min} = S_y^4\left\{\frac{1}{n}(\delta_{400}-1)+\frac{{k_0}^2}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{040}-1)+\frac{({k_0}^2-1)}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{004}-1)\right.$$

$$\left.+k_0\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{220}-1)+(k_0-1)\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{202}-1)-\frac{k_0(k_0-1)}{2}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\} \quad (6.6)$$

where , $k_0 = \frac{\{\delta_{004}+2(\delta_{220}-1)+\delta_{022}-6\}}{\{\delta_{040}+\delta_{004}+2\delta_{022}-4\}}$

Following Lu et al. (2014), we have proposed the following exponential cum ratio type estimator when two auxiliary attributes are available. It is assume that the population variance S_{x1}^2 for the first auxiliary attribute X_1 is unknown, while it is known for the second auxiliary attribute X_2. The proposed estimator t_N in case of two phase sampling will take the following form.

$$t_M = s_y^2 \exp\left[\frac{s_p^2 - S_p^2}{s_p^2 + S_p^2}\right] \quad (6.7)$$

where,

$s_p^2 = w_1 s_{x1}^2 + w_2 s_{x2}^2$, $S_p^2 = w_1 s_{x1}'^2 + w_2 S_{z2}^2$; w_1, w_2 are weights that satisfy the condition: $w_1 + w_2 = 1$.

MSE of the estimator can be found using Taylor series method defined as

$$f\left(s_y^2, s_{x1}^2, s_{x2}^2\right) \cong f\left(S_y^2, S_{x1}^2, S_{x2}^2\right) + \left.\frac{\partial f}{\partial s_y^2}\right|_{\left(S_y^2, S_{x1}^2, S_{x2}^2\right)}\left(s_y^2 - S_y^2\right) + \left.\frac{\partial f}{\partial s_{x1}^2}\right|_{\left(S_y^2, S_{x1}^2, S_{x2}^2\right)}\left(s_{x1}^2 - s_{x1}'^2\right)$$

$$+ \left.\frac{\partial f}{\partial s_{x2}^2}\right|_{\left(S_y^2, S_{x1}^2, S_{x2}^2\right)}\left(s_{x2}^2 - S_{x2}^2\right)$$

Where $f\left(s_y^2, s_{x1}^2, s_{x2}^2\right) = t_M$

$$t_M - S_y^2 = (s_y^2 - S_y^2) - \frac{1}{2}w_1 T(s_{x1}'^2 - S_{x1}^2) - \frac{1}{2}w_2 T(s_{x2}^2 - S_{x2}^2) \quad (6.8)$$

Where $\frac{S_y^2}{w_1 s_{x1}'^2 + w_2 S_{x2}^2} = T$

The MSE expression of this exponential ratio-type estimator is given by

$$MSE(t_M) = E(s_y^2 - S_y^2) + \frac{1}{4}w_1^2 T^2 E(s_{x1}^2 - s_{x1}'^2) + \frac{1}{4}w_2^2 T^2 E(s_{x2}^2 - S_{x2}^2) - w_1 T E(s_{x1}^2 - s_{x1}'^2)(s_y^2 - S_y^2)$$

$$- w_2 T E(s_{x2}^2 - S_{x2}^2)(s_y^2 - S_y^2) + \frac{1}{2}w_1 w_2 T^2 E(s_{x1}^2 - s_{x1}'^2)((s_{X2}^2 - S_{x2}^2)$$

$$= \left\{\frac{1}{n}S_y^4(\delta_{400}-1) + \frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2 T^2 S_{x1}^4(\delta_{040}-1) + \frac{1}{4n}w_2^2 T^2 S_{x2}^4(\delta_{004}-1)\right.$$

$$\left.-\left(\frac{1}{n}-\frac{1}{n'}\right)w_1 T S_{x1}^2 S_y^2(\delta_{220}-1) - \frac{1}{n}w_2 T S_y^2 S_{x2}^2(\delta_{202}-1) + \frac{1}{2}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1 w_2 T^2 S_{x1}^2 S_{X2}^2(\delta_{022}-1)\right\} \quad (6.9)$$

The optimum value of w_1 and w_2 are given by

$$w_1^* = \frac{U}{U+V}, \quad w_2^* = 1 - w_1^*$$

Where,

$$U=\frac{2}{n}S_y^2S_{x2}^2s_{x1}'^4(\delta_{202}-1)-\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)s_{x1}'^2S_y^2S_{x1}^2S_{X2}^2-2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{220}-1)s_{x1}'^2S_{x2}^2S_{x1}^2S_y^2$$
$$+\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{040}-1)S_{x1}^4S_{x2}^2S_y^2$$

$$V=\frac{1}{n}(\delta_{004}-1)s_{x1}'^2S_{x2}^4S_y^2-\frac{2}{n}(\delta_{202}-1)s_{x1}'^2S_{x2}^4S_y^2-\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)S_{x1}^2S_{x2}^4S_y^2$$
$$+2\left(\frac{1}{n}-\frac{1}{n'}\right)S_{x1}^2S_y^2(\delta_{220}-1)S_{x2}^4$$

The expression for minimum MSE of estimator t_M is given by

$$MSE(t_M)_{min}=\left\{\frac{1}{n}S_y^4(\delta_{400}-1)+\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2T_*^2S_{x1}^4(\delta_{040}-1)+\frac{1}{4}w_2^2T_*^2S_{x2}^4(\delta_{004}-1)\right.$$
$$\left.-\left(\frac{1}{n}-\frac{1}{n'}\right)w_1T_*S_{x1}^2S_y^2(\delta_{220}-1)-w_2T_*S_y^2S_{x2}^2(\delta_{202}-1)+\frac{1}{2}w_1w_2T_*^2S_{x1}^2S_{X2}^2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\} \quad (6.10)$$

where,
$$\frac{S_y^2}{w_1^*s_{x1}'^2+w_2^*S_{x2}^2}=T_*$$

Efficiency Comparison

In this section we are comparing the minimum MSE of the proposed estimator t_M with other existing estimators. It is known that the variance of the s_y^2 under SRSWOR is given by

$$V(s_y^2)=\frac{S_y^4}{n}\left[\partial_{400}-1\right].$$

$$V(s_y^2)-MSE(t_M)_{min}=\frac{S_y^4}{n}(\partial_{400}-1)-\left\{\frac{1}{n}S_y^4(\delta_{400}-1)+\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2T_*^2S_{x1}^4(\delta_{040}-1)+\frac{1}{4}w_2^2T_*^2S_{x2}^4(\delta_{004}-1)\right.$$
$$\left.-\left(\frac{1}{n}-\frac{1}{n'}\right)w_1T_*S_{x1}^2S_y^2(\delta_{220}-1)-w_2T_*S_y^2S_{x2}^2(\delta_{202}-1)+\frac{1}{2}w_1w_2T_*^2S_{x1}^2S_{X2}^2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\}\geq 0$$

or,

$$\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2T_*^2S_{x1}^4(\delta_{040}-1)+\frac{1}{4}w_2^2T_*^2S_{x2}^4(\delta_{004}-1)\geq\left(\frac{1}{n}-\frac{1}{n'}\right)w_1T_*S_{x1}^2S_y^2(\delta_{220}-1)$$
$$\left.+w_2T_*S_y^2S_{x2}^2(\delta_{202}-1)-\frac{1}{2}w_1w_2T_*^2S_{x1}^2S_{X2}^2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1\right\}\geq 0 \quad (7.1)$$

$$MSE(t_{1d})-MSE(t_M)_{min}=S_y^4\left\{\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{040}-1)+\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{220}-1)\right\}-\left\{\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2T_*^2S_{x1}^4(\delta_{040}-1)\right.$$
$$+\frac{1}{4}w_2^2T_*^2S_{x2}^4(\delta_{004}-1)-\left(\frac{1}{n}-\frac{1}{n'}\right)w_1T_*S_{x1}^2S_y^2(\delta_{220}-1)-w_2T_*S_y^2S_{x2}^2(\delta_{202}-1)$$
$$\left.+\frac{1}{2}w_1w_2T_*^2S_{x1}^2S_{X2}^2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\}\geq 0 \quad (7.2)$$

$$MSE(t_{2d})-MSE(t_M)_{min}=S_y^4\left\{\frac{1}{n}(\delta_{400}-1)+\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{004}-1)-\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{202}-1)\right\}$$
$$-\left\{\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2T_*^2S_{x1}^4(\delta_{040}-1)+\frac{1}{4}w_2^2T_*^2S_{x2}^4(\delta_{004}-1)-\left(\frac{1}{n}-\frac{1}{n'}\right)w_1T_*S_{x1}^2S_y^2(\delta_{220}-1)\right.$$
$$\left.-w_2T_*S_y^2S_{x2}^2(\delta_{202}-1)+\frac{1}{2}w_1w_2T_*^2S_{x1}^2S_{X2}^2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\}\geq 0 \quad (7.3)$$

$$MSE(t_d)_{min} - MSE(t_M)_{min} = S_y^4\left\{\frac{k_0{}^2}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{040}-1)+\frac{(k_0{}^2-1)}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{004}-1)+k_0\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{220}-1)\right.$$

$$\left.+(k_0-1)\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{202}-1)-\frac{k_0(k_0-1)}{2}\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\}-\left\{\frac{1}{4}\left(\frac{1}{n}-\frac{1}{n'}\right)w_1^2T_*^2S_{x1}^4(\delta_{040}-1)\right.$$

$$+\frac{1}{4}w_2^2T_*^2S_{x2}^4(\delta_{004}-\left(\frac{1}{n}-\frac{1}{n'}\right)w_1T_*S_{x1}^2S_y^2(\delta-1)-w_2T_*S_y^2S_{x2}^2(\delta_{202}-1)$$

$$\left.+\frac{1}{2}w_1w_2T_*^2S_{x1}^2S_{X2}^2\left(\frac{1}{n}-\frac{1}{n'}\right)(\delta_{022}-1)\right\}\geq 0 \qquad (7.4)$$

When the condition (7.1) to (7.4) are satisfied, our suggested estimator t_M is more efficient as compared to t_{d1}, t_{d2} and t_d respectively.

Empirical Study

To illustrate the performance of various estimators of S_y^2, we perform our empirical study based on two given live data sets discussed above.

Table 8.1. PRE's of various estimators w.r.t. S_y^2

Estimator	PRE	
	Population I	**Population II**
S_y^2	100.00	100.00
t_{1d}	86.57	86.17
t_{2d}	135.53	106.72
$t_{d(min)}$	162.95	133.47
$t_{M(min)}$	218.74	138.90

From Table 8.1, it is clear that the estimator $t_{M(min)}$ under optimum condition is more efficient than the other estimators discussed here. Hence for observed choice of parameter the proposed estimator $t_{M(min)}$ is more precise as compared to all other estimators discussed in this paper.

Conclusion

In this article we have proposed an estimator for population variance of study variable when information is available on two auxiliary attributes in simple random sampling without replacement (SRSWOR) and the problem was extended to the case of two phase sampling. The expression for MSE was obtained and compared with some known estimators of population variance such as usual unbiased estimator, ratio and exponential ratio cum dual to ratio usual unbiased estimator and the other existing estimators were derived. The performance of the proposed estimator was assessed using two population data sets. From table 5.1 and 8.1 it is clear that the proposed estimator t_N and t_M performs better than conventional estimator s_y^2 and other estimators discussed here and it is recommended for use in practice.

References

Adichwal, N. K., Sharma, P., Verma, H. K., and Singh, R.(2015): Generalized class of estimators for estimating population variances using auxiliary attribute, International journal of applied and computational mathematics, 1-10.

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

Jhajj, H. S. And Walia, G. S. (2011): A generalized difference –cum-ratio type estimator for the population variance in double sampling, IAENG international journal of applied mathematics, 1-5.

Lu, j., Yan, Z., and Peng, X. (2014): A new exponential ratio-type estimator with linear combination of two auxiliary variables, Plos One, 9(12), 1-10.

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

Shabbir, J., and Gupta, S. (2010): Estimation of the finite population mean in two phase sampling when auxiliary variables are attributes, Hacettepe Journal of Mathematics and Statistics, 1-9.

Singh, D. and Chaudhary, F. S. (1967): Theory and analysis of sample survey designs. John Wiley & Sons, 1986.

Singh, R., Chauhan, P., Sawan, N., and Smarandache, F. (2009): Improved exponential estimator for population variance using two auxiliary variables. Studies in Statistical Inference, Sampling Techniques and Demography, pg 36-44..

Singh, R., and Malik. S.(2014): Improved estimator of population variance using information on auxiliary attribute in simple random sampling, Applied mathematics and computation, 235, 43-49.

Singh, R. and Singh., V. K. (2015): Improved class of estimator for variance under single and two phase sampling, Journal of statistics applications & probability, 1-11.