BULLETIN *of the*
Malaysian Mathematical
Sciences Society
http://math.usm.my/bulletin

Bull. Malays. Math. Sci. Soc. (2) **33**(2) (2010), 345–353

Regression Type Estimators of Finite Population Variance Under Multiphase Sampling

B. K. Pradhan
Department of Statistics, Utkal University, Bhubaneswar-751004, India
pradhanbkuu@rediffmail.com

Abstract. Different multivariate regression type of estimators for finite population variance under multiphase sampling set up in presence of two auxiliary variables have been suggested. These estimators are compared with estimators using no auxiliary variable or single auxiliary variable theoretically and with the help of numerical examples.

2000 Mathematics Subject Classification: 62D05

Key words and phrases: Finite population variance, auxiliary variables, regression type estimators, two phase sampling, three phase sampling.

1. Introduction and preliminaries

The problem of estimation of finite population variance of the study variable y was perhaps first focused through the writings of Evans [4] and Hansen, Hurwitz and Madow [5]. The finite population variance may be required to be estimated with a view to an idea about the variability exist in the population which is necessary for future surveys either to advocate stratification or for determination of sample size. In certain sampling designs like simple random sampling without replacement, the estimation of sampling variance of the sample mean of the study variable necessitates the estimation of finite population variance.

An exploratory work in this direction was initiated by Liu [8] in a general set up, i.e., under unequal probability sampling. Subsequently, Chaudhuri [1] suggested a series of non-negative estimates of the finite population variance. Liu and Thompson [9] have estimated the general problem of estimation of polynomial finite population parametric function in sample surveys.

Mukhopadhyay [12–14] has derived the optimum sampling strategies for estimating the finite population variance under a super population set up. Mukhopadhyay [15] also derived the asymptotic properties of a generalized predictor of finite population variance. Mishra [10], Mishra and Swain [11] have discussed an alternative

Communicated by Anton Abdulbasah Kamil.
Received: January 22, 2008; *Revised:* June 14, 2009.

method to derive Liu's generalized estimator of finite population variance and also suggested an alternative estimator for this purpose.

Using auxiliary information, Das and Tripathi [2] suggested a series of estimators to estimate the finite population variance of the study variable y. Srivastava and Jhajj [18] proposed a class of estimators and have shown that the estimators suggested by Das and Tripathi [2] belong to this class. Isaki [6] has discussed the multivariate ratio and regression estimators to estimate finite population variance. Mishra and Swain [11] also have suggested a regression type estimator for estimating finite population variance.

Situations may arise when the finite population variance S_x^2 of the auxiliary variable x is not known in advance. In order to obtain a more efficient estimator of S_y^2, the finite population variance of y, by using the relationship between auxiliary variable x and the variable of interest y, when the population variance S_x^2 of x is not known, Pradhan [16] and Diana and Tommasi [3] proposed a two phase sampling scheme. In the first phase, an initial simple random sample (without replacement) $s' \subset U$ of fixed size n' is selected to observe auxiliary variable x. In the second phase, a simple random sample (without replacement) s of fixed size n is drawn from s' to observe the variable of interest y. The regression type estimator of finite population variance S_y^2 in two phase sampling takes the form

$$\hat{S}^{*2}_{yreg} = s_y^2 + \beta_{22}(y,x)({s'}_x^2 - s_x^2) \tag{1.1}$$

where ${s'}_x^2$ and s_x^2 are estimates of finite population variance of x using first phase and second phase samples respectively, s_y^2 is an unbiased estimate of the finite population variance of y based on second phase sample and further

$$\beta_{22}(y,x) = \frac{\text{Cov}(s_y^2, s_x^2)}{\text{Var}(s_x^2)}.$$

Under bivariate normality of (y,x), $\beta_{22}(y,x) = \beta_{yx}^2$, where β_{yx} represents regression coefficient of y on x; and hence to first order of approximations,

$$V(\hat{S}^{*2}_{yreg}) \cong 2(1-\rho_{yx}^4)\frac{S_y^4}{n} + 2\rho_{yx}^4\frac{S_y^4}{n'} \tag{1.2}$$

where ρ_{yx} is the correlation coefficient between y and x.

2. Regression type estimators in two phase sampling using two auxiliary variables

Let there be two auxiliary variables under consideration to estimate the finite population variance S_y^2 of y. When the finite population variance S_x^2 of one of the auxiliary variables, say x, is not known but S_z^2 of z is known, we consider the following regression type estimators in two phase sampling following the techniques first suggested by Swain [20] and subsequently developed by Kiregyra [7] for the estimation of finite population mean in the presence of two auxiliary variables x and z. In the first phase a simple random sample $s'$of fixed size n' from the population U is drawn to observe both x and z. In the second phase a simple random sample s of fixed size n is drawn from s' to observe the study variable y. The sampling in both phases is carried out without replacement.

Assuming y, x and z to follow a trivariate normal distribution, regression type estimators for the finite population variance may be proposed as

$$(A) \qquad \hat{S}^2_{(1)} = s^2_y + \beta_{22}(y,x)\left(\hat{S}^2_x - s^2_x\right),$$

where

$$\hat{S}^2_x = {s'}^2_x + \beta'_{22}(x,z)\left(S^2_z - {s'}^2_z\right),$$

$$\beta_{22}(y,x) = \frac{\mathrm{Cov}(s^2_y, s^2_x)}{\mathrm{Var}(s^2_x)}$$

and

$$\beta'_{22}(x,z) = \frac{\mathrm{Cov}({s'}^2_x, {s'}^2_z)}{\mathrm{Var}({s'}^2_z)}.$$

Under bivariate normality of (y,x), $\beta_{22}(y,x) = \beta^2_{yx}$, where β_{yx} is the simple regression coefficient of y on x. Under bivariate normality of (x,z) $\beta'_{22}(x,z) = \beta^2_{xz}$ where β_{xz} is the simple regression coefficient of x on z, ${s'}^2_x$ and ${s'}^2_z$ be the estimates of S^2_x and S^2_z based on the first phase sample respectively and s^2_y and s^2_x are in usual sample estimates based on the second phase sample. Under trivariate normality of (y,x,z), assuming N to be sufficiently large and to the first order of approximations, the variance of $\hat{S}^2_{(1)}$ is given by

$$(2.1) \qquad V(\hat{S}^2_{(1)}) \cong 2\left(1-\rho^4_{yx}\right)\frac{S^4_y}{n} + 2\left(\rho^4_{yx} + \rho^4_{yx}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}\right)\frac{S^4_y}{n'}$$

where ρ_{yx}, ρ_{yz} and ρ_{xz} are simple correlation coefficients with usual notations. The outline of proof of (2.1) is given in Appendix.

$$(B) \qquad \hat{S}^2_{(2)} = s^2_y + \lambda_1(\hat{S}^2_x - s^2_x) + \lambda_2(S^2_z - s^2_z)$$

where $\hat{S}^2_x = {s'}^2_x + \beta'_{22}(x,z)(S^2_z - {s'}^2_z)$ and λ_1 and λ_2 are suitable constants to be determined so as to minimize $V(\hat{S}^2_{(2)})$.

The optimum values of λ_1 and λ_2 under the trivariate normality of (y,x,z) to the first order of approximations are given by

$$(2.2) \qquad \lambda_{1(opt)} = \frac{\rho^2_{yx} - \rho^2_{yz}\rho^2_{xz}}{1-\rho^4_{xz}}.\frac{S^2_y}{S^2_x} \quad \text{and} \quad \lambda_{2(opt)} = \frac{\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}}{1-\rho^4_{xz}}.\frac{S^2_y}{S^2_z}.$$

Thus under trivariate normality assuming N to be sufficiently large and to the first order of approximations

$$\begin{aligned} V_{opt}(\hat{S}^2_{(2)}) &\cong 2\left[1 - \frac{\rho^4_{yx} + \rho^4_{yz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1-\rho^4_{xz}}\right]\frac{S^4_y}{n} \\ (2.3) \qquad &+ 2\left[\frac{\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1-\rho^4_{xz}}\right]\frac{S^4_y}{n'}. \end{aligned}$$

Following Isaki [6], we may consider another estimator of S^2_y given by

$$(C) \qquad \hat{S}^2_{(3)} = s^2_y + \lambda'_1({s'}^2_x - s^2_x) + \lambda'_2({s'}^2_z - s^2_z) + \lambda'_3(S^2_z - {s'}^2_z),$$

where λ'_1, λ'_2 and λ'_3 are suitable constants to be determined so as to minimize $V(\hat{S}^2_{(3)})$.

Assuming trivariate normality of (y, x, z), the optimum values of λ'_1, λ'_2 and λ'_3 to the first order of approximations are

$$\lambda'_{1(opt)} = \frac{\rho^2_{yx} - \rho^2_{yz}\rho^2_{xz}}{1-\rho^4_{xz}} . \frac{S^2_y}{S^2_x}, \quad \lambda'_{2(opt)} = \frac{\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}}{1-\rho^4_{xz}} . \frac{S^2_y}{S^2_z}, \quad \lambda'_{3(opt)} = \rho^2_{yz}\frac{S^2_y}{S^2_z}. \tag{2.4}$$

Thus, to the first order of approximations,

$$V_{opt}(\hat{S}^2_{(3)}) \cong 2\left[1 - \frac{\rho^4_{yx} + \rho^4_{yz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1-\rho^4_{xz}}\right]\frac{S^4_y}{n} + 2\left[\frac{\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1-\rho^4_{xz}}\right]\frac{S^4_y}{n'}, \tag{2.5}$$

which is the same as the $V_{opt}(\hat{S}^2_{(2)})$ to same order of approximations.

$$(D) \qquad \hat{S}^2_{(4)} = s^2_y + \lambda_{(1)}\left(\hat{S}^2_x - s^2_x\right) + \lambda_{(2)}\left(S^2_z - {s'}^2_z\right)$$

where $\hat{S}^2_x = {s'}^2_x + \beta'_{22}(x,z)\,(S^2_z - {s'}^2_z)$.

Under bivariate normality of (x, z), $\beta'_{22}(x, z) = \beta^2_{xz}$. Under trivariate normality of (y, x, z) assuming N to be sufficiently large and to the first order of approximations, the optimum values of $\lambda_{(1)}$ and $\lambda_{(2)}$ are given by

$$\lambda_{(1)opt} = \rho^2_{yx}\frac{S^2_y}{S^2_x} \quad \text{and} \quad \lambda_{(2)opt} = \left(\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}\right)\frac{S^2_y}{S^2_z}. \tag{2.6}$$

Thus, to the first order of approximations,

$$V_{opt}\left(\hat{S}^2_{(4)}\right) \cong 2\left(1-\rho^4_{yx}\right)\frac{S^4_y}{n} + 2\left(\rho^4_{yx} - \rho^4_{yz}\right)\frac{S^4_y}{n'}. \tag{2.7}$$

The optimized constants in $\hat{S}^2_{(2)}$, $\hat{S}^2_{(3)}$ and $\hat{S}^2_{(4)}$ are functions of population parameters, which are usually not known. Hence, in practice we substitute the consistent estimators for the unknown parameters in the optimized constants for the purpose of estimation of variance.

3. Comparison of efficiency

(a) Since

$$V\left(\hat{S}^{*2}_{y\,reg}\right) - V_{opt}\left(\hat{S}^2_{(4)}\right) = 2\frac{\rho^4_{yz}}{n'} \geq 0,$$

we have

$$V_{opt}\left(\hat{S}^2_{(4)}\right) \leq V\left(\hat{S}^{*2}_{y\,reg}\right). \tag{3.1}$$

(b) Since

$$V\left(\hat{S}^2_{(1)}\right) - V_{opt}\left(\hat{S}^2_{(4)}\right) = \frac{2}{n'}\left(\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}\right)^2 \geq 0,$$

we have

$$V_{opt}\left(\hat{S}^2_{(4)}\right) \le V\left(\hat{S}^2_{(1)}\right). \tag{3.2}$$

(c) Since

$$V_{opt}\left(\hat{S}^2_{(4)}\right) - V_{opt}\left(\hat{S}^2_{(2)}\right) = 2\left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\left(\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}\right)^2}{1-\rho^4_{xz}} \ge 0,$$

we have

$$V_{opt}\left(\hat{S}^2_{(2)}\right) \le V_{opt}\left(\hat{S}^2_{(4)}\right). \tag{3.3}$$

(d) Since

$$V(\hat{S}^2_{(1)}) - V_{opt}(\hat{S}^2_{(2)}) \ge 2(\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz})^2\frac{S^2_y}{n'} \ge 0,$$

we have

$$V_{opt}(\hat{S}^2_{(2)}) \le V(\hat{S}^2_{(1)}). \tag{3.4}$$

Hence we conclude that $\hat{S}^2_{(2)}$ is more efficient estimator than $\hat{S}^{*2}_{y\,reg}$, $\hat{S}^2_{(1)}$ and $\hat{S}^2_{(4)}$. It may be noted that the estimators $\hat{S}^2_{(1)}$ and $\hat{S}^2_{(2)}$ due to Pradhan [16] belong to the class of estimators proposed by Diana and Tommasi [3].

4. Regression type estimators in three phase sampling using two auxiliary variables

In the case when the population variance of z, S^2_z is not known, we first select a large preliminary first phase sample s'' of size n'' from the finite population of size N and z is observed. Subsequently, in the second phase a sub-sample s' of size n' is drawn from n'' to observed x and finally in the third phase a sub-sample of size n is drawn from n' to observe the study variable y. The sampling designs in all these three phases are simple random sample without replacement.

Here, with usual notations $\beta_{22}(y,x)$ and $\beta'_{22}(x,z)$, we consider two estimators of finite population variance when y, x and z follow trivariate normality.

$$(A) \qquad \hat{S}^{*2}_{(1)} = s^2_y + \beta_{22}(y,x)(\hat{S}^2_x - s^2_x),$$

where $\hat{S}^2_x = {s'}^2_x + \beta'_{22}(x,z)({s''}^2_z - {s'}^2_z)$. Then under trivariate normality of (y,x,z) assuming N to be sufficiently large and to the first order of approximations, we find

$$\begin{aligned} V(\hat{S}^{*2}_{(1)}) &\cong 2(1-\rho^4_{yx})\frac{S^4_y}{n} + 2(\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz})\frac{S^4_y}{n'} \\ &\quad + 2(2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz} - \rho^4_{yx}\rho^4_{xz})\frac{S^4_y}{n'}. \end{aligned} \tag{4.1}$$

$$(B) \qquad \hat{S}^{*2}_{(2)} = s^2_y + \lambda^*_1(\hat{S}^2_x - s^2_x) + \lambda^*_2({s''}^2_z - s^2_z),$$

where $\hat{S}^2_x = {s'}^2_x + \beta'_{22}(x,z)({s''}^2_z - {s'}^2_z)$ and λ^*_1 and λ^*_2 are suitable constants to be determined under trivariate normality of (y,x,z). (See Appendix).

The optimized constants in $\hat{S}^{*2}_{(2)}$ are functions of population parameters, which are usually not known. Hence, in practice we substitute the consistent estimators for the unknown parameters in the optimized constants for the purpose of estimation of variance.

For sufficiently large N and under the trivariate normality the approximate variance of $\hat{S}^{*2}_{(2)}$ is given by

$$V_{opt}(\hat{S}^{*2}_{(2)}) \cong 2\left[1 - \frac{\rho^4_{yx} + \rho^4_{yz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1 - \rho^4_{xz}}\right]\frac{S^4_y}{n} + 2\left[\frac{\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1 - \rho^4_{xz}}\right]\frac{S^4_y}{n'}$$

$$+ 2\rho^4_{yz}\frac{S^4_y}{n''}. \tag{4.2}$$

The outline of proof of (4.2) is given in Appendix.

5. Comparison of efficiency

$$V(\hat{S}^{*2}_{(1)}) - V_{opt}(\hat{S}^{*2}_{(2)}) = \left(\frac{A}{n} + \frac{B}{n'} + \frac{C}{n''}\right) \times 2S^4_y,$$

where

$$A = \frac{\rho^4_{yx} + \rho^4_{yz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1 - \rho^4_{xz}} - \rho^4_{yx},$$

$$B = (\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}) - \left[\frac{\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{(1 - \rho^4_{xz})}\right]$$

and

$$C = 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz} - \rho^4_{yx}\rho^4_{xz} - \rho^4_{yz}.$$

Since

$$\left(\frac{A}{n} + \frac{B}{n'}\right) \geq \frac{A + B}{n'} = \frac{(\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz})^2}{n'},$$

we have

$$V(\hat{S}^{*2}_{(1)}) - V_{opt}(\hat{S}^{*2}_{(2)}) \geq 2\left(\frac{1}{n'} - \frac{1}{n''}\right)\left(\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}\right)^2 S^4_y \geq 0.$$

Hence we conclude that $V_{opt}(\hat{S}^{*2}_{(2)}) \leq V(\hat{S}^{*2}_{(1)})$.

6. Numerical illustrations

To observe the relative performance of different estimators discussed above, we consider two natural population data used earlier by others. These populations are described below.

Population-I (Sukhatme and Chand [19])
$N = 120$;
$y =$ bushels of apples harvested in 1964
$x =$ apple tree of bearing age in 1964
$z =$ bushels of apples harvested in 1959
$\rho_{yx} = 0.93, \quad \rho_{yz} = 0.84, \quad \rho_{xz} = 0.77$

Population-II (Srivastava [17])
$N = 50$;
y = yield per plant
x = height of the plant
z = base diameter
$\rho_{yx} = 0.7418, \quad \rho_{yz} = 0.5677, \quad \rho_{xz} = 0.2063$

Table 1. Relative efficiency of different estimators of population variance

Estimators	Auxiliary Variables Used	% Relative Efficiency	
		Pop^n. I ($n'' = 70$, $n' = 50$, $n = 20$)	Pop^n. II ($n'' = 30$, $n' = 20$, $n = 8$)
s_y^2	None	100	100
$\hat{S}_{(1)}^{*2}$	x, z	215.82	122.51
$\hat{S}_{(2)}^{*2}$	x, z	217.45	133.19

Remark 6.1. $\hat{S}_{(2)}^{*2}$ has substantial gain in efficiency compared to $\hat{S}_{(1)}^{*2}$ and s_y^2. The proposed estimators depend upon population regression coefficients, correlation coefficients and variances, which are generally not known. In practice, these population values are to be estimated from the given sample and as a result, the estimators become biased. However, in large samples, the biases are negligible and the variance expressions are asymptotically equivalent.

7. Appendix

Outline of proof of (2.1). Consider a regression estimator of population variance of the study variable y by

$$\hat{S}_{(1)}^2 = s_y^2 + \beta_{22}(y,x)\left(\hat{S}_x^2 - s_x^2\right),$$

where

$$\hat{S}_x^2 = {s'}_x^2 + \beta'_{22}(x,z)\left(S_z^2 - {s'}_z^2\right).$$

Now,

$$\begin{aligned} V\left(\hat{S}_{(1)}^2\right) &= V_1 E_2\left(\hat{S}_{(1)}^2\right) + E_1 V_2\left(\hat{S}_{(1)}^2\right) \\ &\cong \left[2\left(\frac{1}{n'} - \frac{1}{N}\right) S_y^4 + 2\left(\frac{1}{n'} - \frac{1}{N}\right)\rho_{yx}^4\rho_{xz}^4 S_y^4 - 4\left(\frac{1}{n'} - \frac{1}{N}\right)\rho_{yx}^2\rho_{yz}^2\rho_{xz}^2 S_y^4\right] \\ &\quad + \left[2\left(\frac{1}{n} - \frac{1}{n'}\right) S_y^4 + 2\left(\frac{1}{n} - \frac{1}{n'}\right)\rho_{yx}^4 S_y^4 - 4\left(\frac{1}{n} - \frac{1}{n'}\right)\rho_{yx}^4 S_y^4\right] \\ &\cong 2\left(1 - \rho_{yx}^4\right)\frac{S_y^4}{n} + 2\left(\rho_{yx}^4 + \rho_{yx}^4\rho_{xz}^4 - 2\rho_{yx}^2\rho_{yz}^2\rho_{xz}^2\right)\frac{S_y^4}{n'}, \end{aligned}$$

if N is sufficiently large. ∎

Outline of proof of (4.2). Consider a regression estimator of population variance of the study variable y by

$$\hat{S}^{*2}_{(2)} = s^2_y + \lambda^*_1(\hat{S}^2_x - s^2_x) + \lambda^*_2({s''}^2_z - s^2_z),$$

where

$$\hat{S}^2_x = {s'}^2_x + \beta'_{22}(x,z)({s''}^2_z - {s'}^2_z)$$

and

$$\beta'_{22}(x,z) = \frac{\mathrm{Cov}({s'}^2_x, {s'}^2_z)}{\mathrm{Var}({s'}^2_z)}$$

and λ^*_1 and λ^*_2 are preassigned constants to be estimated by minimizing $V(\hat{S}^{*2}_{(2)})$ under trivariate normality condition and for sufficiently large N. Now,

$$\begin{aligned}
V(\hat{S}^{*2}_{(2)}) &= V_1E_2E_3(\hat{S}^{*2}_{(2)}) + E_1V_2E_3(\hat{S}^{*2}_{(2)}) + E_1E_2V_3(\hat{S}^{*2}_{(2)}) \\
&= \left[2\left(\frac{1}{n''} - \frac{1}{N}\right)S^4_y\right] + \left[2\left(\frac{1}{n'} - \frac{1}{n''}\right)S^4_y + 2\left(\frac{1}{n'} - \frac{1}{n''}\right)\lambda^{*2}_1\rho^4_{xz}S^4_x\right. \\
&\quad + 2\left(\frac{1}{n'} - \frac{1}{n''}\right)\lambda^{*2}_2 S^4_z - 4\left(\frac{1}{n'} - \frac{1}{n''}\right)\lambda^*_1\rho^2_{xz}\rho^2_{yz}S^2_xS^2_y \\
&\quad \left. -4\left(\frac{1}{n'} - \frac{1}{n''}\right)\lambda^*_2\rho^2_{yz}S^2_yS^2_z + 4\left(\frac{1}{n'} - \frac{1}{n''}\right)\lambda^*_1\lambda^*_2\rho^2_{xz}S^2_xS^2_z\right] \\
&\quad + \left[2\left(\frac{1}{n} - \frac{1}{n'}\right)S^4_y + 2\left(\frac{1}{n} - \frac{1}{n'}\right)\lambda^{*2}_1 S^4_x + 2\left(\frac{1}{n} - \frac{1}{n'}\right)\lambda^{*2}_2 S^4_z\right. \\
&\quad - 4\left(\frac{1}{n} - \frac{1}{n'}\right)\lambda^*_1\rho^2_{yx}S^2_yS^2_x - 4\left(\frac{1}{n} - \frac{1}{n'}\right)\lambda^{*2}_2\rho^2_{yz}S^2_yS^2_z \\
&\quad \left. +4\left(\frac{1}{n} - \frac{1}{n'}\right)\lambda^*_1\lambda^*_2\rho^2_{xz}S^2_xS^2_z\right].
\end{aligned}$$

Applying the method of least square in order to minimize $V(\hat{S}^{*2}_{(2)})$, we find

$$\lambda^*_{1(opt)} = \frac{\rho^2_{yx} - \rho^2_{yz}\rho^2_{xz}}{1 - \rho^4_{xz}}.\frac{S^2_y}{S^2_x} \quad \text{and} \quad \lambda^*_{2(opt)} = \frac{\rho^2_{yz} - \rho^2_{yx}\rho^2_{xz}}{1 - \rho^4_{xz}}.\frac{S^2_y}{S^2_z}.$$

Substituting the values of $\lambda^*_{1(opt)}$ and $\lambda^*_{2(opt)}$ in $V(\hat{S}^{*2}_{(2)})$, we find

$$\begin{aligned}
V_{opt}(\hat{S}^{*2}_{(2)}) &\cong 2\left[1 - \frac{\rho^4_{yx} + \rho^4_{yz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1 - \rho^4_{xz}}\right]\frac{S^4_y}{n} \\
&\quad + 2\left[\frac{\rho^4_{yx} + \rho^4_{yz}\rho^4_{xz} - 2\rho^2_{yx}\rho^2_{yz}\rho^2_{xz}}{1 - \rho^4_{xz}}\right]\frac{S^4_y}{n'} + 2\rho^4_{yz}\frac{S^4_y}{n''}. \qquad \blacksquare
\end{aligned}$$

Acknowledgement. The author wishes to express his sincere gratitude to the referee for his valuable suggestions in improving the manuscript. The author also gratefully acknowledges the suggestions of Prof. A. K. P. C. Swain to improve the content of the paper.

References

[1] A. Chaudhuri, On estimating the variance of a finite population, *Metrika* **25** (1978), no. 2, 65–76.

[2] A. K. Das and T. P. Tripathi, Use of auxiliary information in estimating the finite population variances, *Sankhya*, *The Indian Journal of Statistics* **40** (1978), Ser. C, Pt. 2, 139–148.

[3] G. Diana and C. Tommasi, Estimation for finite population variance in double sampling, *Metron* **62** (2004), no. 2, 223–232.

[4] W. D. Evans, On the variance of estimates of the standard deviation and variance, *J. Amer. Statist. Assoc.* **46** (1951), 220–224.

[5] M. H. Hansen, W. N. Hurwitz and W. G. Madow, *Sample Survey Methods and Theory. Vol. I*, Reprint of the 1953 original, Wiley, New York, 1993.

[6] C. T. Isaki, Variance estimation using auxiliary information, *J. Amer. Statist. Assoc.* **78** (1983), no. 381, 117–123.

[7] B. Kiregyera, Regression-type estimators using two auxiliary variables and the model of double sampling from finite populations, *Metrika* **31** (1984), no. 3–4, 215–226.

[8] T. P. Liu, A general unbiased estimator for the variance of a finite population, *Sankhȳa Ser. C* **36** (1974), 23–32.

[9] T. P. Liu and M. E. Thompson, Properties of estimators of quadratic finite population functions: The batch approach, *Ann. Statist.* **11** (1983), no. 1, 275–285.

[10] G. Mishra, On estimation of finite population variance and coefficient of variation using auxiliary information, *Unpublished Ph.D. Thesis* (1991), Utkal University, Bhubaneswar, India.

[11] G. Mishra and A. K. P. C. Swain, A modified regression type estimator for estimating finite population variance, *Sankhyikee* (1994), Utkal University, Vol. I, 21–29.

[12] P. Mukhopadhyay, Estimating the variance of a finite population under a superpopulation model, *Metrika* **25** (1978), no. 2, 115–122.

[13] P. Mukhopadhyay, Optimum strategies for estimating the variance of a finite population under a superpopulation model, *Metrika* **29** (1982), no. 3, 143–158.

[14] P. Mukhopadhyay, Optimum estimation of finite population variance under generalised random permutation models, *Calcutta Statist. Assoc. Bull.* **33** (1984), no. 129–130, 93–106.

[15] P. Mukhopadhyay, On asymptotic properties of a generalised predictor of finite population variance, *Sankhyā Ser. B* **52** (1990), no. 3, 343–346.

[16] B. K. Pradhan, Some problems of estimation in multi-phase sampling, *Unpublished Ph.D. Statistics Thesis* (April 2000), Utkal University, Bhubaneswar, India.

[17] S. K. Srivastava, A generalized estimator for the mean of a finite population using multi-auxiliary information, *J. Amer. Stat. Assoc.* **66** (1971), 404–407.

[18] S. K. Srivastava and H. S. Jhajj, A class of estimators using auxiliary information for estimating finite population variance, *Sankhȳa Ser. C* **42** (1980), 87–96.

[19] B. V. Sukhatme and L. Chand, Multivariate ratio type estimator, Proceedings, Social Statistic Section, *Amer. Stat. Association* (1977), 927–931.

[20] A. K. P. C. Swain, A note on the use of multiple auxiliary variables in sample surveys, *Trabajos Estadíst.* **XXX** (1970), 135–141.