Maejo Int. J. Sci. Technol. **2018**, *12*(03), 272-281

Maejo International
Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

On some new estimators of population variance in single and two-phase sampling

Ali Hussein Al–Marshadi, Abdullah Hamoud Alharby and Muhammad Qaiser Shahbaz*

Department of Statistics, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

* Corresponding author, e-mail: mkmohamad@kau.edu.sa; qshahbaz@gmail.com

Received: 7 August 2017 / Accepted: 7 December 2018 / Published: 17 December 2018

Abstract: We propose some estimators for the estimation of population variance in single- and two-phase sampling. The estimators are proposed based upon information of multiple auxiliary variables. We obtain the mean square error of the proposed estimators and compare them with some existing estimators. We have also conducted an empirical study for both single- and two-phase sampling.

Keywords: variance estimation, logarithmic estimators, mean square error, two-phase sampling

---

INTRODUCTION

The estimation of population variance has been an interesting area in statistics. The sample variance s_y^2 has been popularly used as an estimator of population variance S_y^2 from the early time of statistical science. The variance estimation has found significant application in designed experiments, especially the study of random-effect models as discussed by Montgomery [1]. The variance estimation in those experiments is model-specific and depends upon the underlying model.

The estimation of variance has gained significant attraction in the field of survey sampling and is done with the objective to achieve high efficiency in estimation. The classical estimator of population variance s_y^2 uses the information of study variables and hence may have lower efficiency in some cases. When the information on some auxiliary variables is available, then that information can also be utilised in the estimation of population variance just as the case in which we use auxiliary information for estimation of mean. Isaki [2] proposed the classical ratio and regression estimators of population variance by using the information of a single auxiliary variable. Singh and Singh [3] proposed an improved ratio-type estimator for population variance in single-phase sampling. Singh and Singh [4] have also proposed a regression-type estimator of population variance in two-phase sampling. Jhajj et al. [5] proposed a chain-ratio estimator of population variance in two-phase sampling using the information of multiple auxiliary variables.

The estimation of variance has also been studied in the context of stratified and two-phase sampling. Kadilar and Cingi [6] proposed a ratio-type estimator of population variance in simple and stratified random sampling. Singh and Vishwakarma [7] proposed some families of estimators for population variance in stratified random sampling.

The exponential estimators have also attracted several survey statisticians. Bhal and Tuteja [8] have proposed an exponential estimator of variance in single-phase sampling, which has been improved by Singh et al. [9] by using the information of two auxiliary variables. Asghar et al. [10] have proposed a generalised exponential estimator of population variance using the information of two auxiliary variables, which has been further extended by Shabbir and Gupta [11]. Some other notable references include Gupta and Shabbier [12], Upadhyaya et al. [13], Bansal et al. [14], Subramani and Kumarapandiyan [15], Yadav and Kadilar [16, 17] and Yadav et al. [18], among others.

SOME EXISTING ESTIMATORS

Several authors have proposed different estimators for population variance from time to time. The main objective in proposing these estimators is to improve the efficiency of estimates. In the following we discuss some popular estimators of population variance, alongside variance of their sampling distributions, but we first give some important terms which are necessary to understand these estimators. It is customary to say that S_Y^2 and S_X^2 represent population variances of study and auxiliary variables respectively. The corresponding sample variances are s_Y^2 and s_X^2. The sample and population variances are connected as $s_Y^2 = S_Y^2(1+e_Y)$ and $s_X^2 = S_X^2(1+e_X)$, where e_Y and e_X are much smaller as compared with s_Y^2 and s_X^2. In the case of two auxiliary variables we can use $s_Z^2 = S_Z^2(1+e_Z)$.

We have the following results for e_Y, e_X and e_Z:

$$\left.\begin{array}{l} s_Y^2 = S_Y^2(1+e_Y),\ s_X^2 = S_X^2(1+e_X), s_Z^2 = S_Z^2(1+e_Z) \\ E(e_Y) = E(e_X) = E(e_Z) = 0 \\ E\left(e_Y^2\right) = \gamma\varphi_{400}^*, E\left(e_X^2\right) = \gamma\varphi_{040}^*, E\left(e_Z^2\right) = \gamma\varphi_{004}^*, E(e_Y e_X) = \gamma\varphi_{220}^* \\ \gamma = n^{-1},\ \varphi_{rst}^* = (\varphi_{rst} - 1)\ and\ \varphi_{rst} = \mu_{rst} \big/ \left(\mu_{200}^{r/2}\mu_{020}^{s/2}\mu_{002}^{t/2}\right) \\ \mu_{rst} = N^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^r\left(x_i - \bar{X}\right)^s\left(z_i - \bar{Z}\right)^t \end{array}\right\}. \qquad (1)$$

In the case of two-phase sampling we have the following additional notations:

$$\left.\begin{array}{l} E\left(e_{X(1)}^2\right) = \gamma_1\varphi_{040}^*, E\left(e_{Z(1)}^2\right) = \gamma_1\varphi_{004}^*, E\left(e_{Y(2)}^2\right) = \gamma_2\varphi_{400}^*, E\left(e_{X(2)}^2\right) = \gamma_2\varphi_{040}^*, \\ E\left(e_{Z(2)}^2\right) = \gamma_2\varphi_{004}^*, E\left(e_{X(1)}e_{Z(1)}\right) = \gamma_1\varphi_{022}^*,\ E\left(e_{X(2)}e_{Z(2)}\right) = \gamma_2\varphi_{022}^*,\ E\left(e_{X(1)}e_{Z(2)}\right) = \gamma_1\varphi_{022}^* \end{array}\right\} \qquad (2)$$

In the case of single auxiliary variable the last subscript in the above notations is removed. The above notations can be easily extended to the case of several auxiliary variables. For example, if we have k auxiliary variables, then

$$\left.\begin{aligned} &E\left(e_Y^2\right)=\gamma\varphi^*_{\underset{k\ terms}{\underline{400...0}}}=\gamma\varphi^*_{\mathbf{40}},E\left(e_{X_j}^2\right)=\gamma\varphi^*_{\underset{jth\ place}{\underline{00...4...0}}}=\gamma\varphi^*_{04_j} \\ &E\left(e_Y e_{X_j}\right)=\gamma\varphi^*_{\underset{jth\ place}{\underline{20...2...0}}}=\gamma\varphi^*_{22_j}\ and\ E\left(e_{X_j}e_{X_h}\right)=\gamma\varphi^*_{\underset{(j,h)th\ place}{\underline{00...2...2...0}}}=\gamma\varphi^*_{02_j2_h} \end{aligned}\right\} \tag{3}$$

We now give some popular estimators of population variance with their mean square errors (MSEs), where MSE for an estimator $\hat{\theta}$ of parameter θ is defined as

$$MSE\left(\hat{\theta}\right)=E\left[\left(\hat{\theta}-\theta\right)^2\right].$$

The estimators, with their MSEs, are given below.

a. Isaki [2]:

Estimator: $$\hat{S}_I^2=s_Y^2\left(\frac{S_X^2}{s_X^2}\right) \tag{4}$$

$$MSE\left(\hat{S}_I^2\right)\cong\gamma S_Y^4\left[\varphi^*_{40}+\varphi^*_{04}-2\varphi^*_{22}\right] \tag{5}$$

b. Singh, Chauhan, Sawan and Smarandache [9]:

Estimator: $$\hat{S}_{SCSS}^2=s_Y^2\left[\alpha\exp\left\{\frac{S_X^2-s_X^2}{S_X^2+s_X^2}\right\}+\left(1-\alpha\right)\exp\left\{\frac{s_Z^2-S_Z^2}{s_Z^2+S_Z^2}\right\}\right] \tag{6}$$

$$MSE\left(\hat{S}_{SCSS}^2\right)\cong\gamma S_Y^4\left[\varphi^*_{400}+\frac{\alpha}{4}\varphi^*_{040}+\frac{\left(1-\alpha^2\right)}{4}\varphi^*_{004}-\alpha\varphi^*_{220}+\left(1-\alpha\right)\varphi^*_{202}-\frac{\alpha\left(1-\alpha\right)}{2}\varphi^*_{022}\right] \tag{7}$$

c. Yadav and Kadilar [16]:

Estimator: $$\hat{S}_{YK}^2=s_Y^2\exp\left[\frac{S_X^2-s_X^2}{S_X^2+\left(a-1\right)s_X^2}\right] \tag{8}$$

$$MSE\left(\hat{S}_{YK}^2\right)\cong\gamma S_Y^4\left[\varphi^*_{40}-\frac{\varphi^{*2}_{22}}{\varphi^*_{40}}\right] \tag{9}$$

d. Asghar, Sanaullah and Hanif [10]:

Estimator: $$\hat{S}_{ASH}^2=\lambda s_Y^2\exp\left[\frac{c\left(\bar{X}-\bar{x}\right)}{\bar{X}+\left(d-1\right)\bar{x}}\right] \tag{10}$$

$$MSE\left(\hat{S}_{ASH}^2\right)\cong\gamma S_Y^4\left[1-\frac{1}{\varphi_{40}-\varphi_{21}^2}\right] \tag{11}$$

e. Shabbir and Gupta [11]:

Estimator: $$\hat{S}_{SG}^2=\lambda s_Y^2\exp\left[\frac{c\left(\bar{X}-\bar{x}\right)}{\bar{X}+\left(d-1\right)\bar{x}}\right]\exp\left[\frac{e\left(S_X^2-s_x^2\right)}{S_X^2+\left(f-1\right)s_X^2}\right] \tag{12}$$

$$MSE\left(\hat{S}_{SG}^2\right)\cong\gamma S_Y^4\left[\varphi^*_{40}-\varphi_{21}^2-\frac{\left(\varphi^*_{22}-\varphi_{21}\varphi_{03}\right)^2}{\varphi^*_{04}-\varphi_{03}^2}\right] \tag{13}$$

We now propose some new estimators of population variance. We first propose estimators for single-phase sampling and then provide their counterparts in two-phase sampling.

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 272-281

SINGLE-PHASE SAMPLING ESTIMATORS

In this section we propose estimators of the variance for single-phase sampling. We first propose an estimator using the information of a single auxiliary variable and then extend the result to the case of multiple auxiliary variables. The proposed estimator using the information of a single variable is

$$\hat{S}^2_{MHS1} = s^2_Y + \ln\left(\frac{s^2_X}{S^2_X}\right)^{\alpha}, \tag{14}$$

where α is a constant and MSH1 is AlMarshadi-AlHarbi-Shahbaz first estimator. We can see that for $\alpha = 0$ the estimator (14) reduces to the classical variance estimator s^2_Y. In order to obtain the bias and MSE of (14) we use the notations given in (1). Using those notations we have

$$\hat{S}^2_{MHS1} = S^2_Y\left(1+e_Y\right) + \alpha\ln\left\{\frac{S^2_X\left(1+e_X\right)}{S^2_X}\right\} = S^2_Y\left(1+e_Y\right) + \alpha\ln\left(1+e_X\right).$$

Using the first-order expansion of $\ln\left(1+e_X\right)$ and applying the expectation, the bias of (14) is

$$Bias\left(\hat{S}^2_{MHS1}\right) = -2\gamma\frac{\alpha}{S^2_Y}\varphi^*_{04}. \tag{15}$$

The MSE of (14) is

$$MSE\left(\hat{S}^2_{MHS1}\right) = E\left(\hat{S}^2_{MHS1} - S^2_Y\right)^2 = \gamma S^4_Y\left(\varphi^*_{40} + \frac{\alpha^2}{S^4_Y}\varphi^*_{04} + 2\frac{\alpha}{S^2_Y}\varphi^*_{22}\right). \tag{16}$$

The value which minimises the MSE given in (16) is

$$\alpha = -S^2_Y\frac{\varphi^*_{22}}{\varphi^*_{04}}.$$

Using this value of α in (15) and (16), the minimum bias and minimum MSE of $\hat{S}^2_{MHS1}$ are

$$Bias\left(\hat{S}^2_{MHS1}\right) = 2\gamma\varphi^*_{22},$$

$$MSE_{\min}\left(\hat{S}^2_{MHS1}\right) = \gamma S^4_Y\left(\varphi^*_{40} - \frac{\varphi^{*2}_{22}}{\varphi^*_{04}}\right). \tag{17}$$

We can see that the minimum MSE of $\hat{S}^2_{MHS1}$ is the same as that of the estimator proposed by Yadav and Kadilar [16].

We now extend the estimator (14) to the case of several auxiliary variables. The proposed estimator is

$$\hat{S}^2_{MHS2} = s^2_Y + \ln\left[\prod_{j=1}^{k}\left(s^2_{X_j}/S^2_{X_j}\right)^{\alpha_j}\right], \tag{18}$$

where MHS2 is AlMarshadi-AlHarbi-Shahbaz second estimator. The estimator (18) can be written as

$$\hat{S}^2_{MHS2} = s^2_Y + \sum_{j=1}^{k}\alpha_j\ln\left(\frac{s^2_{X_j}}{S^2_{X_j}}\right).$$

Using the notations given in (3), we can write the above estimator as

$$\hat{S}^2_{MHS2} = S_Y^2\left(1+e_Y\right) + \sum_{j=1}^{k}\alpha_j \ln\left\{\frac{S^2_{X_j}\left(1+e_{X_j}\right)}{S^2_{X_j}}\right\} = S_Y^2\left(1+e_Y\right) + \sum_{j=1}^{k}\alpha_j \ln\left\{\left(1+e_{X_j}\right)\right\}.$$

Further, by using the notations given in (3), the MSE of (18) is

$$MSE\left(\hat{S}^2_{MHS2}\right) = \gamma S_Y^4\left[\varphi^*_{40} + \frac{2}{S_Y^2}\boldsymbol{\alpha}'\boldsymbol{\varphi} + \frac{1}{S_Y^4}\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}\right], \tag{19}$$

where $\boldsymbol{\alpha} = \begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_k\end{bmatrix}$, $\boldsymbol{\varphi} = \begin{bmatrix}\varphi^*_{22_1} & \varphi^*_{22_2} & \cdots & \varphi^*_{22_k}\end{bmatrix}$ and

$$\boldsymbol{\varphi} = \begin{bmatrix}\varphi^*_{22_1}\\ \varphi^*_{22_2}\\ \vdots\\ \vdots\\ \varphi^*_{22_k}\end{bmatrix} \text{ and } \boldsymbol{\Phi} = \begin{bmatrix}\varphi^*_{04_1} & \varphi^*_{02_12_2} & \varphi^*_{02_12_3} & \cdots & \varphi^*_{02_12_k}\\ \varphi^*_{02_22_1} & \varphi^*_{04_2} & \varphi^*_{02_22_3} & \cdots & \varphi^*_{02_22_k}\\ \vdots & \vdots & \vdots & \cdots & \vdots\\ \vdots & \vdots & \vdots & \cdots & \vdots\\ \varphi^*_{02_k2_1} & \varphi^*_{02_k2_2} & \varphi^*_{02_k2_3} & \cdots & \varphi^*_{04_k}\end{bmatrix}. \tag{20}$$

Now differentiating (19) and equating to zero, the value of $\boldsymbol{\alpha}$ which minimises (19) is

$$\boldsymbol{\alpha} = -S_Y^2\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}.$$

Using this value in (19), the minimum MSE of the estimator $\hat{S}^2_{MHS2}$ is

$$MSE_{\min}\left(\hat{S}^2_{MHS2}\right) = \gamma S_Y^4\left(\varphi^*_{40} - \boldsymbol{\varphi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}\right). \tag{21}$$

The MSE given in (21) can be computed for any number of auxiliary variables.

ESTIMATORS FOR TWO-PHASE SAMPLING

Let a random first-phase sample of size n_1 be drawn from a population of size *N*; then from the first-phase sample we have a second-phase sample of size n_2. Using the sample information, we propose the following estimator of population variance in two-phase sampling with single auxiliary variable:

$$\hat{S}^2_{MHS1(2)} = s^2_{Y(2)} + \ln\left(\frac{s^2_{X(2)}}{s^2_{X(1)}}\right)^{\alpha}, \tag{22}$$

where $s^2_{Y(2)}$ is the variance of *Y* based upon the second-phase sample, and $s^2_{X(1)}$ and $s^2_{X(2)}$ are variances of the auxiliary variable based upon the first-phase and second-phase samples respectively. Also, MHS1(2) is AlMarshadi-AlHarbi-Shahbaz first estimator for two-phase sampling. Using the notations given in (2), we can write above estimator as

$$\hat{S}^2_{MHS1(2)} = S_Y^2\left(1+e_{Y(2)}\right) + \alpha\ln\left\{\frac{\left(1+e_{X(2)}\right)}{\left(1+e_{X(2)}\right)}\right\}.$$

Expanding the above terms, squaring and applying expectation, we have

$$MSE\left(\hat{S}^2_{MHS1(2)}\right) = S_Y^4\left[\gamma_2\varphi^*_{40} + \left(\gamma_2-\gamma_1\right)\frac{\alpha^2}{S_Y^4}\varphi^*_{04} + 2\left(\gamma_2-\gamma_1\right)\frac{\alpha}{S_Y^2}\varphi^*_{22}\right]. \tag{23}$$

The value of α, which minimises (23), is $\alpha = -S_Y^2\left(\varphi^*_{22}/\varphi^*_{04}\right)$. Now using this value in (23), the minimum MSE of $\hat{S}^2_{MHS1(2)}$ is

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 272-281

$$MSE\left(\hat{S}^2_{MHS1(2)}\right) = S_Y^4\left[\gamma_2\varphi^*_{40} - (\gamma_2 - \gamma_1)\frac{\varphi^*_{22}}{\varphi^*_{04}}\right]. \tag{24}$$

The estimator (22) can be extended to the case of multiple auxiliary variables. In this case the proposed estimator for two-phase sampling is

$$\hat{S}^2_{MHS2(2)} = s^2_{Y(2)} + \ln\left[\prod_{j=1}^{k}\left(s^2_{X_j(2)}\Big/s^2_{X_j(1)}\right)^{\alpha_j}\right], \tag{25}$$

where MHS2(2) is AlMarshadi-AlHarbi-Shahbaz second estimator for two-phase sampling. The estimator (25) can be written as

$$\hat{S}^2_{MHS2(2)} = s^2_{Y(2)} + \sum_{j=1}^{k}\alpha_j \ln\left(s^2_{X_j(2)}\Big/s^2_{X_j(1)}\right).$$

Using the notations given in (2), we can write the above estimator as

$$\hat{S}^2_{MHS2(2)} = S_Y^2\left(1+e_{Y(2)}\right) + \sum_{j=1}^{k}\alpha_j \ln\left[\frac{\left(1+e_{X_j(2)}\right)}{\left(1+e_{X_j(1)}\right)}\right].$$

Expanding the above terms, squaring and applying expectation, the MSE of (25) is

$$MSE\left(\hat{S}^2_{MHS2(2)}\right) = S_Y^4\left[\gamma_2\varphi^*_{40} + (\gamma_2 - \gamma_1)\left\{\frac{2}{S_Y^2}\boldsymbol{\alpha}'\boldsymbol{\varphi} + \frac{1}{S_Y^4}\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}\right\}\right], \tag{26}$$

where $\boldsymbol{\varphi}$ and $\boldsymbol{\Phi}$ are defined in (20). The value of $\boldsymbol{\alpha}$, which minimises (26), is

$$\boldsymbol{\alpha} = -S_Y^2\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}.$$

Using this value of $\boldsymbol{\alpha}$ in (26), the minimum MSE of $\hat{S}^2_{MHS2(2)}$ is

$$MSE\left(\hat{S}^2_{MHS2(2)}\right) = S_Y^4\left[\gamma_2\varphi^*_{40} - (\gamma_2 - \gamma_1)\boldsymbol{\varphi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}\right]. \tag{27}$$

The MSE (27) can be computed for any number of auxiliary variables. It is the same as the MSE of t_{3g} given by Abu-Dayyeh and Ahmed [19].

NUMERICAL AND SIMULATION STUDY

The numerical and simulation study for the proposed estimators is conducted in this section. The numerical study is given for $\hat{S}^2_{MHS2}$ and is compared with $\hat{S}^2_0 = s^2_Y$ and $\hat{S}^2_{SCSS}$ (SCSS = Singh-Chauhan-Sawan-Smarandache estimator) for two auxiliary variables. The simulation study is conducted for the proposed estimator $\hat{S}^2_{MHS2}$. The numerical study for $\hat{S}^2_{MHS2(2)}$ is also conducted. We compute the relative efficiency (RE) of the estimators as

$$RE = \frac{MSE\left(\hat{S}_i^2\right)}{MSE\left(\hat{S}_0\right)}\times 100\,;\ i = MHS2, SCSS.$$

We use five populations for the numerical study. The first three populations were taken from Weisberg [20] and the last two populations were taken from Kutner et al. [21]. The summary measures for the five populations are given in Table 1. We next compute the MSEs of the different estimators. Their values and the relative efficiency are given in Tables 2 and 3 respectively.

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 272-281

Table 1. Summary measures for populations

Measure	Pop–I	Pop–II	Pop–III	Pop–IV	Pop–V
N	17	58	32	23	110
$\bar{Y}$	202.9529	13.1879	55.9063	61.3478	6.8317
$\bar{X}_1$	25.0588	31.8207	4.4222	39.6087	27.4273
$\bar{X}_2$	1.3961	27.7690	4.3238	50.7826	3.0680
$S_Y^2 = \mu_{200}$	31.2225	2.4276	239.7725	267.1834	5.2011
$S_{X_1}^2 = \mu_{020}$	8.5846	24.0482	2.0431	68.5860	275.8447
$S_{X_2}^2 = \mu_{002}$	0.0025	4.5625	1.8763	18.8658	2.2415
μ_{400}	1897.9720	17.8135	186348.8000	176009.6000	70.5461
μ_{040}	140.4802	1606.3240	11.9619	8990.6110	136845.6000
μ_{004}	0.000001	62.7807	10.6925	1168.5720	12.0884
μ_{220}	514.5816	71.7041	1115.7210	28465.5900	1367.4500
μ_{202}	0.1514	14.4173	1140.2090	11070.3200	24.0244
μ_{022}	0.0411	280.0175	11.0965	1610.9210	600.0912
φ_{400}^*	0.9469	2.0227	2.2414	1.4656	1.6078
φ_{040}^*	0.9062	1.7776	1.8657	0.9113	0.7985
φ_{004}^*	0.9128	2.0159	2.0371	2.2833	1.4061
φ_{220}^*	0.9199	0.2282	1.2776	0.5534	-0.0469
φ_{202}^*	0.9263	0.3017	1.5344	1.1962	1.0607
φ_{022}^*	0.9199	0.2282	1.2776	0.5534	-0.0469

Table 2. MSEs of different estimators

Estimator	Pop–I	Pop–II	Pop–III	Pop–IV	Pop–V
N	5	15	10	5	35
$\hat{S}_0^2$	184.6256	0.7947	12885.80	20924.52	1.2427
$\hat{S}_{SCSS}^2$	437.9805	1.0183	22898.12	21249.81	0.6526
$\hat{S}_{MHS2}^2$	0.9639	0.7769	5001.13	9062.38	0.6236

Table 3. RE of different estimators

Estimator	Pop–I	Pop–II	Pop–III	Pop–IV	Pop–V
N	5	15	10	5	35
$\hat{S}_0^2$	100	100	100	100	100
$\hat{S}_{SCSS}^2$	237.226	128.136	177.700	101.555	52.515
$\hat{S}_{MHS2}^2$	0.522	97.760	38.811	43.310	50.181

We can see that the proposed estimator $\hat{S}^2_{MHS2}$ clearly outperforms the competing estimators in the study. The estimator $\hat{S}^2_{SCSS}$ performs better than $\hat{S}^2_0$ in only one case.

Next , we conduct the empirical study of the two-phase sampling estimators. For this, we use the two-phase sampling version of Isaki estimator [2] and the following estimator proposed by Abu-Dayyeh and Ahmed [19]:

$$\hat{S}^2_{DA(2)} = t_1 = s^2_Y \frac{s^2_{X_1(2)}}{s^2_{X_1(1)}} \frac{s^2_{X_2(2)}}{s^2_{X_2(1)}} . \tag{28}$$

The MSE of (28) is

$$MSE\left(\hat{S}^2_{DA(2)}\right) = S^4_Y\left[\gamma_2\varphi^*_{400} - (\gamma_2 - \gamma_1)\left(\varphi^*_{040} + \varphi^*_{004} - 2\varphi^*_{220} - 2\varphi^*_{202} + 22\varphi^*_{022}\right)\right].$$

The results of the study are given in Tables 4 and 5. These results show that the proposed estimator outperforms other two-phase sampling estimators included in the study.

Table 4. MSEs of different estimators

Estimator	**Pop–I**	**Pop–II**	**Pop–III**	**Pop–IV**	**Pop–V**
n_1	10	30	12	12	60
n_2	5	12	5	6	25
$\hat{S}^2_{I(2)}$	275.624	2.604	30395.87	18600.05	1.776
$\hat{S}^2_{DA(2)}$	191.650	1.727	21900.67	16334.76	1.665
$\hat{S}^2_{MHS2(2)}$	92.795	0.980	16572.82	12494.54	1.234

Table 5. RE of different estimators

Estimator	**Pop–I**	**Pop–II**	**Pop–III**	**Pop–IV**	**Pop–V**
n_1	10	30	12	12	60
n_2	5	12	5	6	25
$\hat{S}^2_{I(2)}$	100	100	100	100	100
$\hat{S}^2_{DA(2)}$	69.533	66.321	72.051	87.821	93.750
$\hat{S}^2_{MHS2(2)}$	33.667	37.634	54.523	67.175	69.482

We also conducted a simulation study for the single-phase sampling estimator. For this simulation study we generated random populations of size 200 from normal populations. The population for X was generated from $N\left(50,3^2\right)$. The variable Y was generated from $Y = 2X_1 + 2.5e$, where e is $N\left(0,1\right)$. For each generated population, a random sample of size 80 was drawn and the estimators $\hat{S}^2_0$, $\hat{S}^2_I$ and $\hat{S}^2_{MHS1}$ was computed for each sample. The estimator $\hat{S}^2_{MHS1}$ was computed for $\alpha = 1.0, 1.5, 2.0$ and 2.5. The procedure was repeated 5000 times and the average and variance of the different estimators were then computed. The results are given in Table 6.

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 272-281

Table 6. Simulated results by different estimators

Estimator	Mean	Variance
$\hat{S}_0^2$	36.6211	27.6164
$\hat{S}_I^2$	37.3585	49.2630
$\hat{S}_{MHS1}^2 : \alpha = -2.5$	36.6533	27.6661
$\hat{S}_{MHS1}^2 : \alpha = -2.0$	36.6469	27.6418
$\hat{S}_{MHS1}^2 : \alpha = -1.5$	36.6404	27.6247
$\hat{S}_{MHS1}^2 : \alpha = -1.0$	36.6340	27.6148
$\hat{S}_{MHS1}^2 : \alpha = 1.0$	36.6082	27.6467
$\hat{S}_{MHS1}^2 : \alpha = 1.5$	36.6018	27.6726
$\hat{S}_{MHS1}^2 : \alpha = 2.0$	36.5953	27.7056
$\hat{S}_{MHS1}^2 : \alpha = 2.5$	36.5889	27.7458
$\hat{S}_{MHS1}^2 : \alpha = Optimum$	37.1864	25.5796

From Table 6, we can see that the estimator $\hat{S}_{MHS1}^2$ has a smaller variance as compared to the other competing estimators. Further, we can see that the variance of $\hat{S}_{MHS1}^2$ is minimum when the optimum value of α is used.

CONCLUSIONS

We have proposed some estimators for estimating the population variance. The estimators are proposed for single-phase and two-phase sampling using single and multiple auxiliary variables. In contrast to the available estimators, we have proposed the estimators by using the logarithm of auxiliary information. We have also conducted the empirical and simulation study of the proposed estimators.

We have found that the proposed estimator in single-phase sampling is equivalent to Yadav and Kadilar estimator but the construction of our estimator is much simpler. We have also found that our proposed estimator for two-phase sampling is equivalent to Abu-Dayyeh and Ahmed estimator but again the construction of our estimator is much simpler.

ACKNOWLEDGEMENTS

This project was funded by the Deanship of Scientific Research (DSR), at King Abdulaziz University, Jeddah, under Grant no. G-136/130/37. The authors, therefore, acknowledge with thanks DSR for technical and financial support.

REFERENCES

1. D. C. Montgomery, “Design and Analysis of Experiments”, 5th Edn., John Wiley and Son, New York, **2001**, pp.512-518.
2. C. T. Isaki, “Variance estimation using auxiliary information”, *J. Am. Stat. Assoc.*, **1983**, *78*, 117-123.

3. H. P. Singh and R. Singh, "Improved ratio-type estimator for variance using auxiliary information", *J. Indian Soc. Agric. Stat.*, **2001**, *54*, 276-287.
4. H. P. Singh and R. Singh, "Estimation of variance through regression approach in two-phase sampling", *Aligarh J. Stat.*, **2003**, *23*, 13-30.
5. H. S. Jhajj, M. K. Sharma and L. K. Grover, "An efficient class of chain estimators of population variance under sub-sampling scheme", *J. Japan Stat. Soc.*, **2005**, *35*, 273-286.
6. C. Kadilar and H. Cingi, "Ratio estimators for the population variance in simple and stratified random sampling", *App. Math. Comp.*, **2006**, *173*, 1047-1059.
7. H. P. Singh and G. K. Vishwakarma, "Some families of estimators of variance of stratified random sample mean using auxiliary information", *J. Stat. Theory Pract.*, **2008**, *2*, 21-43.
8. S. Bahl and R. K. Tuteja, "Ratio and product type exponential estimators", *J. Inform. Optim. Sci.*, **1991**, *12*, 159-164.
9. R. Singh, P. Chauhan, N. Sawan and F. Smarandache, "Improved exponential estimator for population variance using two auxiliary variables", *Octogon Math. Magaz.*, **2009**, *17*, 667-674.
10. A. Asghar, A. Sanaullah and M. Hanif, "Generalized exponential type estimator for population variance in survey sampling", *Rev. Colomb. Estad.*, **2014**, *37*, 213-224.
11. J. Shabbir and S. Gupta, "A note on generalized exponential type estimator for population variance in survey sampling", *Rev. Colomb. Estad.*, **2015**, *38*, 385-397.
12. S. Gupta and J. Shabbir, "Variance estimation in simple random sampling using auxiliary information", *Hacettepe J. Math. Stat.*, **2008**, *37*, 57-67.
13. L. N. Upadhyaya, H. P. Singh, S. Chatterjee and R. Yadav, "Improved ratio and product exponential type estimators", *J. Stat. Theory Pract.*, **2011**, *5*, 285-302.
14. M. Bansal, M. Javed and N. Khanna, "A class of estimators of variance of the regression estimator", *Int. J. Agric. Stat. Sci.*, **2011**, *7*, 275-279.
15. J. Subramani and G. Kumarapandiyan, "Variance estimation using median of the auxiliary variable", *Int. J. Prob. Stat.*, **2012**, *1*, 36-40.
16. S. K. Yadav and C. Kadilar, "Improved exponential type ratio estimator of population variance", *Rev. Colomb. Estad.*, **2013**, *36*, 145-152.
17. S. K. Yadav and C. Kadilar, "A two parameter variance estimator using auxiliary information", *Appl. Math. Comput.*, **2014**, *226*, 117-122.
18. S. K. Yadav, C. Kadilar, J. Shabbir and S. Gupta, "Improved family of estimators of population variance in simple random sampling", *J. Stat. Theory Pract.*, **2015**, *9*, 219-226.
19. W. Abu-Dayyeh and M. S. Ahmed, "Ratio and regression estimators for the variance under two-phase sampling", *Int. J. Stat. Sci.*, **2005**, *4*, 49-56.
20. S. Weisberg, "Applied Linear Regression", 2^{nd} Edition, John Wiley and Sons, New York, **1985**.
21. M. H. Kutner, C. J. Nachtsheim, J. Neter and W. Li, "Applied Linear Statistical Models", 5^{th} Edn., McGraw Hill, New York, **2004**.