Bulletin of Business and Economics, 12(3), 473-484
https://bbejournal.com
https://doi.org/10.61506/01.00056

Advancing Sampling Techniques: Multivariate Ratio Estimation for Variance Vector in Two-Phase Sampling

Amber Asghar[1], Aamir Sanaullah[2], Azhar Mehmood Abbasi[3], Muhammad Hanif[4]

Abstract
The problem of variance vector estimation using multi-auxiliary variables is not quite often considered in the literature. In the present study, we propose a multivariate ratio estimation approach for estimating a vector of variances in two-phase sampling using a vector of l variables, following the situation if the desired parameters are available only for some of the auxiliary variables. Some special cases of the proposed generalized multivariate variance estimator have also been discussed. Expressions of the bias, and generalized variance has been derived. With the help of real-life data, the applicability of the proposed multivariate variance estimator has been given, and it is shown that the proposed multivariate estimator is more efficient than the modified versions of multivariate variance estimators. A simulation study has also been carried out to show that the proposed estimator surpasses the modified version of multivariate estimators.
Keywords: variance-covariance matrix; two-phase sampling; bias; ratio estimator; multi-auxiliary variables; simulation

1. Introduction

We ruminate a finite population $\left[\Omega = \left\{\Omega_1, \Omega_2, ...\Omega_N\right\}\right]$. Let Y_j be the variable under study where $j = 1, 2, ..., m$ and X_j be the auxiliary variable where $k = 1, 2, ..., l$. Let $\bar{Y}_j$ and $\bar{X}_k$ be the population means and $S^2_{y_j}$ and $S^2_{x_k}$ be the population variances of the study and auxiliary variable respectively. Further, let $C^2_{x_k}$, $C^2_{y_j}$ be the coefficients of disparity and $\rho_{y_j x_k}$ indicates the correlation coefficient between study and ancillary variables. We are considering two-phase sampling design; where the first-phase sample of size n_1 is particular from the population N units and the second-phase sample of n_2 units are selected from $n_1 (n_2 \subset n_1)$. Now an assumption is made that only l_1 auxiliary variables are available with their known population means and variances while l_2 represents the unknown auxiliary information.

Let $s^2_{y_{(2)j}} = S^2_{y_j}(1+\varepsilon_{y_{(2)j}})$, $s^2_{x_{(2)k}} = S^2_{x_k}(1+\varepsilon_{x_{(2)k}})$ and $\bar{x}_{(1)k} = \bar{X}_k(1+\bar{e}_{x_{(1)k}})$ defined the relative errors and $\varepsilon_{y_{(2)j}}, \varepsilon_{x_{(2)k}}$ and $\bar{e}_{x_{(1)k}}$ be the sampling errors. To understand the properties of an estimator, some necessary expectations are shown by,

$$E\left(\varepsilon_{x_{(1)k}}\right) = E\left(\varepsilon_{x_{(2)k}}\right) = 0,$$

$$E\left(\varepsilon_{y_{(2)j}}\right) = 0,\ E\left(\bar{e}_{x_{(1)k}}\right) = E\left(\bar{e}_{x_{(2)k}}\right) = 0,\ E_2\left(\varepsilon^2_{y_{(2)j}}\right) = \gamma_2 \mathrm{A}_y = \gamma_2\left[\beta_2(y)-1\right], E_2\left(\varepsilon^2_{x_{(2)k}}\right) = \gamma_2 \mathrm{A}_x = \gamma_2\left[\beta_2(x)-1\right],$$

$$E\left(\varepsilon_{y_{(2)}}\varepsilon_{x_{(1)}}\right) = \gamma_1 \mathrm{A}_{yx} = \gamma_1\left[\varphi_{220}-1\right], E_1E_2\left(\bar{e}_{x_{(1)k}}\varepsilon_{y_{(2)}}\right) = \gamma_1 \mathrm{A}_{yx_d} = \gamma_1\varphi_{210}C_{x_k},$$

$$E_1\left(\bar{e}^2_{x_{(1)k}}\right) = \gamma_1 \mathrm{A}_{x_d} = \gamma_1 C^2_{x_k}, E_1\left(\bar{e}^2_{x_{(2)k}}\right) = \gamma_1 \mathrm{A}_{x_{d_1}} = \gamma_1 C^2_{x_k}\ E_1(\bar{e}_{x_{(1)k}}\bar{e}_{x_{(1)h}}) = \gamma_1 A_{x_{dk}x_h} = \gamma_1 C_{x_k}C_{x_h}\rho_{x_k x_h},$$

where $k \neq h$

where E_1 and E_2 denotes the potentials over first and second phase and $\beta_2(y)$, $\beta_2(x)$ are Kurtosis of the study and auxiliary variable respectively.

$$\varphi_{p_1....p_m q_1....q_l} = \frac{\mu_{p_1....p_m q_1....q_l}}{\mu_{200}^{\ p_1....p_m/2}\ \mu_{020}^{\ q_1....q_l/2}}$$

$$\varphi_{p_1....p_m q_1....q_l} = \frac{1}{N}\sum_{i=1}^{N}(Y_{i1}-\bar{Y}_1)^{p_1}(Y_{i2}-\bar{Y}_2)^{p_2}...(Y_{ij}-\bar{Y}_1)^{p_m}(X_{i1}-\bar{X}_1)^{q_1}(X_{i2}-\bar{X}_2)^{q_2}...(X_{ik}-\bar{X}_k)^{q_l}.$$

We use the following results using simple random sampling without replacement (SRSWOR) for the derivation of bias and variance-covariance (Var-Cov) matrices under two-phase sampling in multivariate situations,

[1] Virtual University of Pakistan, Lahore, Pakistan, zukhruf10@gmail.com
[2] COMSATS University Islamabad Lahore Campus, Pakistan, chaamirsanaullah@yahoo.com
[3] Corresponding Author, Virtual University of Pakistan, Lahore, Pakistan, abbasiqau2007@yahoo.com
[4] NCBA&E, Lahore, Pakistan, drmianhanif@gmail.com

Let

$$\Delta_y = \left[\varepsilon_{y_{(2)1}} \quad \varepsilon_{y_{(2)2}} \quad . \quad . \quad \varepsilon_{y_{(2)m}} \right],$$

$$\Delta_x = \left[\varepsilon_{x_{(1)1}} \quad \varepsilon_{x_{(1)2}} \quad . \quad . \quad \varepsilon_{x_{(1)l1}} \right], \Delta_{x_1} = \left[\varepsilon_{x_{(2)1}} \quad \varepsilon_{x_{(2)2}} \quad . \quad . \quad \varepsilon_{x_{(2)l2}} \right],$$

$$\bar{D}_x = \left[\mathrm{e}_{x_{(1)1}} \quad \mathrm{e}_{x_{(1)2}} \quad . \quad . \quad \mathrm{e}_{x_{(1)l_1}} \right], \bar{D}_{x_1} = \left[\mathrm{e}_{x_{(2)1}} \quad \mathrm{e}_{x_{(2)2}} \quad . \quad . \quad \mathrm{e}_{x_{(2)l2}} \right],$$

$$E_1E_{2/1}\left(\Delta_x{}'\Delta_x\right) = \gamma_1 \Sigma_{x_{(l_1\times l_1)}}, E_1E_{2/1}\left(\Delta_y{}'\Delta_y\right) = \gamma_2 \Sigma_{y_{(m\times m)}}, E_1E_{2/1}\left(\Delta_x{}'\Delta_y\right) = \gamma_1 \Sigma_{xy_{(l_1\times m)}},$$

$$E_1E_{2/1}\left(\bar{D}_x{}'\bar{D}_x\right) = \gamma_1 \Sigma_{x_{d(l_1\times l_1)}}\ E_1E_{2/1}\left(\Delta_y{}'\bar{D}_x\right) = \gamma_1 \Sigma_{yx_{d(m\times l_1)}}\ , E_1E_{2/1}\left(\bar{D}_x{}'\Delta_y\right) = \gamma_1 \Sigma_{x_d y_{(l_1\times m)}},$$

$$E_1E_{2/1}\left(\bar{D}_{x_1}{}'\bar{D}_{x_1}\right) = \left(\gamma_2 - \gamma_1\right)\Sigma_{x_{d1(l_2\times l_2)}},\ E_1E_{2/1}\left(\Delta_{x_1}{}'\Delta_{x_1}\right) = \left(\gamma_2 - \gamma_1\right)\Sigma_{x_{1(l_2\times l_2)}},$$

$$E_1E_{2/1}\left(\Delta_{x_1}{}'\Delta_y\right) = \gamma_1 \Sigma_{x_1 y_{(l_2\times m)}}$$

and

$\gamma_1 = \dfrac{1}{n_1}$ and $\gamma_2 = \dfrac{1}{n_2}$, Σ represents the variance-covariance matrix.

Ahmed (2000) proposed the two estimators for assessing finite population variance using multivariate assisting variables when assisting variables are known and found that their estimators are more efficient than Isaki (1983). Abu-Dayyehand Ahmed (2002) proposed the ratio and regression estimators when assisting variables are not known for the population but known for the large samples with moderate cost. They proposed the variance estimator using two auxiliary variables under two-phase sampling. Kadilar and Cingi (2007) also suggested the regression estimator using an auxiliary variable and exposed that their estimator is more efficient than Isaki's (1983) ratio and regression estimator under assured situations. There are some other vital studies in the fiction such as Kim and Sitter (2003) Singh et al. (2003), Singh et al. (2010), Singh and Solanki (2013), Olufadi and kadilar (2014), Asghar et al. (2014), Singh and Singh (2015), Yadav et al. (2015), Singh and Pal (2016),Sanaullah et al. (2016), and recently Ismail et al. (2018), Abid et al. (2019), Naz et al. (2020), Zaman et al. (2021), Daraz et al. (2021) and Niaz et al. (2022) developed the ratio estimators for estimating population variance. Further, Cekim and Kadilar (2020) proposed the ln-type regression estimators for variance estimation under SRS and Haq et al. (2021) worked on the variance estimators under stratified sampling.

In this paper, we have considered the multivariate situation when some of the assisting variables are known, and the others are unknown, and we have adopted the two-phase sampling for estimating population variance. The goal of this study is to show how we can deal with the multiple situations using known and unknown auxiliary variables for estimating the variance vector. The arrangement of the paper is as follows, in Section 2, we modify some existing estimators into multivariate cases and find their Var-Cov matrices. The bias and Var-Cov matrix of multivariate ratio type estimator is given in Section 3. A real-life example has been taken to exhibit the routine of our suggested estimator in Section 4. Section 5 is based on the simulation studies and lastly, conclusions are presented in Section 6.

2. Modify form of some existing estimators

An unbiased estimator of population variance is sample variance i.e., $t_0 = s_y^2$, We modify it in multivariate case using multi-auxiliary variables as $t_{0d} = \left[t_{0d1} \quad t_{0d2} \cdots \quad t_{0dj} \cdots \quad t_{0dm} \right]$,

$$t_{0dj} = s^2_{y_{(2)j}}, \qquad j = 1, 2, ..., m, \tag{1}$$

and variance of unbiased estimator is $\mathrm{Var}\left(t_0\right)$,

Isaki (1983) suggested ratio estimator for estimating population variance is t_r.

Modify it in multivariate form as $t_r = \left[t_{r1} \quad t_{r2} \cdots \quad t_{rj} \cdots \quad t_{rm} \right]$,

$$t_{rj} = s^2_{y_{(2)j}} \prod_{k=1}^{l_1} \left(\frac{S^2_{x_{(1)k}}}{s^2_{x_{(1)k}}} \right) \prod_{k=l_1+1}^{l_1+l_2=l} \left(\frac{s^2_{x_{(1)k}}}{s^2_{x_{(2)k}}} \right), \ \mathrm{j} = 1, 2, ..., \mathrm{m}\ ,\ \mathrm{k} = 1, 2, ..., \mathrm{l}, \tag{2}$$

and its Var-Cov matrix is given,

$$\sum_{tr_{(m\times m)}} = S'S\begin{pmatrix} \gamma_2 I'_{(m\times m)} \sum_{y_{(m\times m)}} I_{(m\times m)} + \gamma_1 I'_{(m\times l_1)} \sum_{x_{(l_1\times l_1)}} I_{(l_1\times m)} \\ +(\gamma_2-\gamma_1) I'_{(m\times l_2)} \sum_{x_{1(l_2\times l_2)}} I_{(l_2\times m)} \\ -2\gamma_1 (I'_{(m\times m)} \sum_{yx_{(m\times l_1)}} I_{(l_1\times m)}) \\ -2(\gamma_1-\gamma_2)(I'_{(m\times m)} \sum_{yx_{1(m\times l_2)}} I_{(l_2\times m)}) \end{pmatrix}. \quad (3)$$

Upadhaya and Singh (1999) anticipated the following ratio estimator and we modify it in multivariate case,

$$t_u = \begin{bmatrix} t_{u1} & t_{u2}\cdots & t_{uj}\cdots & t_{um} \end{bmatrix}$$

$$t_{uj} = s^2_{y_{j(2)}} \prod_{k=1}^{l_1} \left(\frac{\bar{X}_{(1)k}}{\bar{x}_{(1)k}} \right) \prod_{k=l_1+1}^{l_1+l_2=l} \left(\frac{\bar{x}_{(1)k}}{\bar{x}_{(2)k}} \right), \; j=1,2,...,m, k=1,2,....,l, \quad (4)$$

and its Var-Cov matrix is given,

$$\sum_{tu_{(m\times m)}} = S'S\begin{pmatrix} \gamma_2 I'_{(m\times m)} \sum_{y_{(m\times m)}} S_{(m\times m)} + \gamma_1 I'_{(m\times l_1)} \sum_{x_{d(l_1\times l_1)}} I_{(l_1\times m)} + (\gamma_2-\gamma_1) I'_{(m\times l_2)} \sum_{x_{d1(l_2\times l_2)}} I_{(l_2\times m)} \\ -2\gamma_1 (I'_{(m\times m)} \sum_{yx_{d(m\times l_1)}} I_{(l_1\times m)}) - 2(\gamma_1-\gamma_2)(I'_{(m\times m)} \sum_{yx_{d1(m\times l_2)}} I_{(l_2\times m)}) \end{pmatrix}, \quad (5)$$

where I represents the identity matrix.

Das and Tripathi (1978) proposed the following ratio estimators using single auxiliary variable as,

$$t_{D1} = s_y^2 \left(\frac{\bar{X}}{\bar{x}_1} \right)^{\alpha_0}, t_{D2} = s_y^2 \left(\frac{S_x^2}{s_{x_1}^2} \right)^{\alpha_1}, \; t_{D3} = s_y^2 \left(\frac{\bar{X}}{\bar{X} + \alpha_2 (\bar{x}_1 - \bar{X})} \right).$$

We derived ratio type multivariate estimator following Das and Tripathi (1978) t_{D1},

$$t_d = \begin{bmatrix} t_{d1} & t_{d2}\cdots & t_{dj}\cdots & t_{d(dm)} \end{bmatrix}$$

$$t_{dj} = s^2_{y_{j(2)}} \prod_{k=1}^{l_1} \left(\frac{\bar{X}_{(1)k}}{\bar{x}_{(1)k}} \right)^{\alpha_1} \prod_{k=l_1+1}^{l_1+l_2=l} \left(\frac{\bar{x}_{(1)k}}{\bar{x}_{(2)k}} \right)^{\alpha_2}. \quad (6)$$

Its Var-Covmatrix is given,

$$\sum_{td_{(m\times m)}} = S'S\begin{bmatrix} \gamma_2 \sum_{y_{(m\times m)}} - \gamma_1 \sum_{yx_{d(m\times l_1)}} \sum^{-1}_{v(l_1\times l_1)} \sum_{x_d y_{(l_1\times m)}} \\ -(\gamma_2-\gamma_1) \sum_{yx_{d1(m\times l_2)}} \sum^{-1}_{x_{d1(l_2\times l_2)}} \sum_{x_{d1} y_{(l_2\times m)}} \end{bmatrix}, \quad (7)$$

where α_0 is used for minimizing.

3. Proposed estimator for Variance Estimation

Adapting the estimator of Das and Tripathi (1978), given in Section (1), we develop the following variance estimator for estimating the finite population variance in the multivariate case using multi-auxiliary variables,

$$t_p = \begin{bmatrix} t_{pj} \end{bmatrix}_{(1xm)},$$

where

$$t_{pj} = s^2_{y_{j(2)}} \left(\prod_{k=1}^{l_1} \left[\frac{S^2_{x_k}}{\alpha_{kj} s^2_{x(1)k} + (1-\alpha_{kj}) S^2_{x_k}} \right]^{g_{kj}} \prod_{k=l_1+1}^{l_1+l_2=l} \left[\frac{s^2_{x(1)k}}{\beta_{kj} s^2_{x(2)k} + (1-\beta_{kj}) s^2_{x(1)k}} \right]^{g^*_{kj}} \right), \quad (8)$$

and g_{kj} and $g_{kj}*$ be the known constants which may takes the values $(-1,0,1)$ to attain global product-type, unbiased variance and ratio-type estimators respectively, α_{kj} and β_{kj} are anonymous constants; estimated by minimizing the asymptotic Var-Cov matrix of the estimator. Up to first order of approximation, (8) may be written as in $\varepsilon_{i's}$,

$$t_{pj} = S_{y_j}^2 (1+\varepsilon_{y_{j(2)}}) \left(\prod_{k=1}^{l_1} \left[\frac{1}{(1+\alpha_{kj}\varepsilon_{x_{k(1)}})} \right]^{g_{kj}} \prod_{k=l_1+1}^{l_1+l_2=l} \left[\frac{(1+\varepsilon_{x_{k(1)}})}{1+\varepsilon_{x_{k(1)}} + \beta_{kj}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}})} \right]^{g_{kj}^*} \right), \tag{9}$$

or

$$t_{pj} = S_{y_j}^2 (1+\varepsilon_{y_{j(2)}}) \left(\begin{array}{l} \prod_{k=1}^{l_1} \left[(1+\alpha_{kj}\varepsilon_{x_{k(1)}})^{-1} \right]^{g_{kj}} \\ \prod_{k=l_1+1}^{l_1+l_2=l} \left[(1+\varepsilon_{x_{k(1)}}) \left(1 + \left\{ \varepsilon_{x_{k(1)}} + \beta_{kj}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \right\} \right)^{-1} \right]^{g_{kj}^*} \end{array} \right).$$

Ignoring the higher order terms and expanding up to first order estimation, we have

$$t_{pj} = S_{y_j}^2 (1+\varepsilon_{y_{j(2)}}) \left(\begin{array}{l} \prod_{k=1}^{l_1} \left[(1-g_{kj}\alpha_{kj}\varepsilon_{k(1)}) \right] \\ \prod_{k=l_1+1}^{l_1+l_2=l} \left[(1+\varepsilon_{k(1)}) \left(1 - g_{kj}^* \left\{ \varepsilon_{k(1)} + \beta_{kj}(\varepsilon_{k(2)} - \varepsilon_{k(1)}) \right\} \right) \right] \end{array} \right), \tag{10}$$

or

$$t_{pj} = S_{y_j}^2 (1+\varepsilon_{y_{j(2)}}) \left(1 - \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}\varepsilon_{k(1)} + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj} \left(\varepsilon_{k(1)} - \varepsilon_{k(2)} \right) \right).$$

After some simplifications, we may get,

$$t_{pj} = S_{y_j}^2 \left(1 + \varepsilon_{y_{j(2)}} - \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}\varepsilon_{k(1)} - \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj} \left(\varepsilon_{k(2)} - \varepsilon_{k(1)} \right) \right). \tag{11}$$

On simplification, we may have,

$$t_{pj} = S_{y_j}^2 \left(\begin{array}{l} 1 + \varepsilon_{y_{j(2)}} - \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}\varepsilon_{x_{k(1)}} + \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}^2\varepsilon_{x_{k(1)}}^2 + \sum_{k=1}^{l_1} g_{kj}^2\alpha_{kj}^2\varepsilon_{x_{k(1)}}^2 \\ + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^{*2}\beta_{kj}^2(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}})^2 - \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \\ + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}^2(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}})^2 + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}\varepsilon_{x_{k(1)}}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \end{array} \right). \tag{12}$$

Subtracting $S_{y_j}^2$ from (12) and taking expectation we have,

$$E_1E_{2/1}\left(t_{pj} - S_{y_j}^2\right) = S_{y_j}^2 E_1E_{2/1} \left(\begin{array}{l} \varepsilon_{y_{j(2)}} - \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}\varepsilon_{y_{j(2)}}\varepsilon_{x_{k(1)}} - \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}\varepsilon_{x_{k(1)}} \\ + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}\varepsilon_{x_{k(1)}}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) + \sum_{k=1}^{l_1} g_{kj}\alpha_{kj}^2\varepsilon_{x_{k(1)}}^2 \\ + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^{*2}\beta_{kj}^2(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}})^2 + \sum_{k=1}^{l_1} g_{kj}^2\alpha_{kj}^2\varepsilon_{x_{k(1)}}^2 \\ + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}^2(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}})^2 \\ - \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \\ + \sum_{k=l_1+1}^{l_1+l_2=l} g_{kj}^* \beta_{kj}\varepsilon_{y_{j(2)}}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \end{array} \right). \tag{13}$$

Since for bias (13) may be extended up to order more than one,

$$t_{pj} = S_{y_j}^2 (1+\varepsilon_{y_{j(2)}}) \begin{pmatrix} \prod_{k=1}^{l_1} \left[(1-\alpha_{kj}\varepsilon_{x_{k(1)}} + \alpha_{kj}^2 \varepsilon_{x_{k(1)}}^2) \right]^{g_{kj}} \\ \prod_{k=l_1+1}^{l_1+l_2=l} \left[(1+\varepsilon_{x_{k(1)}}) \begin{pmatrix} 1-\left\{ \varepsilon_{x_{k(1)}} + \beta_{kj}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \right\} \\ +\left\{ \varepsilon_{x_{k(1)}} + \beta_{kj}(\varepsilon_{x_{k(2)}} - \varepsilon_{x_{k(1)}}) \right\}^2 \end{pmatrix} \right]^{g_{kj}^*} \end{pmatrix}. \tag{14}$$

After some simplification, the bias of our proposed estimator is,

$$Bias\left(t_p\right)_{(1\times m)} = S_y^2 \begin{pmatrix} \gamma_1 A_{x_{(1\times l_1)}} \left[g_{kj}\alpha_{kj}^2 \right]_{(l_1\times m)} + \gamma_1 A_{x_{(1\times l_1)}} \left[g_{kj}^2 \alpha_{kj}^2 \right]_{(l_1\times m)} \\ -\gamma_1 A_{yx_{(1\times l_1)}} \left[g_{kj}\alpha_{kj} \right]_{(l_1\times m)} \\ +\left(\gamma_2-\gamma_1\right) A_{yx_{1(1\times l_2)}} \left[g_{kj}^* \beta_{kj} \right]_{(l_2\times m)} \\ +\left(\gamma_2-\gamma_1\right) A_{yx_{1(1\times l_2)}} \left[g_{kj}^* \beta_{kj}^2 \right]_{(l_2\times m)} \\ -\left(\gamma_2-\gamma_1\right) A_{yx_{1(1\times l_2)}} \left[g_{kj}^{*2} \beta_{kj}^2 \right]_{(l_2\times m)} \end{pmatrix}. \tag{15}$$

Using (15), we may proceed for Var-Covmatrix,

$$\begin{aligned} \sum_{t_{p(m\times m)}} &= E_1 E_{2/1}\left(\mathbf{t}_{p(1\times m)} - \mathrm{S}_{(1\times m)}\right)'\left(\mathbf{t}_{p(1\times m)} - \mathrm{S}_{(1\times m)}\right) \\ &= S_y^4 E_1 E_{2/1}\left(\Delta_{y_{(1\times m)}} - \Delta_{x_{(1\times n_1)}} \varpi_{(n_1\times m)} + \Delta_{x_{1(1\times n_2)}} \omega_{(n_2\times m)}\right)' \left(\Delta_{y_{(1\times m)}} - \Delta_{x_{(1\times l_1)}} \varpi_{(l_1\times m)} + \Delta_{x_{1(1\times l_2)}} \omega_{(l_2\times m)}\right), \end{aligned} \tag{16}$$

where $\varpi_{(l_1\times m)} = \left[g_{kj}\alpha_{kj} \right]_{(l_1\times m)}$ and $\omega_{(l_2\times m)} = \left[g_{kj}\beta_{kj} \right]_{(l_2\times m)}$, the matrix of regression coefficients and usually these coefficients are unknown, and their estimated values are used to minimize the Var-Cov matrix of multivariate ratio type estimator.

Using the results of expectations (11), (16) may be written as:

$$\sum_{t_{p(m\times m)}} = S'S \begin{pmatrix} \gamma_2 \sum_{y_{(m\times m)}} + \gamma_1 \varpi'_{(m\times l_1)} \sum_{x_{(l_1\times l_1)}} \varpi_{(l_1\times m)} - \gamma_1 \varpi'_{(m\times l_1)} \sum_{yx_{(l_1\times m)}} \\ +\left(\gamma_1-\gamma_2\right)\omega'_{(l_2\times m)} \sum_{yx_{1(m\times l2)}} + \left(\gamma_1-\gamma_2\right)\sum_{yx_{1(m\times l2)}} \omega_{(l_2\times m)} \\ +\left(\gamma_2-\gamma_1\right)\omega'_{(m\times l2)} \sum_{x_{1(l_2\times l_2)}} \omega_{(l_2\times m)} - \gamma_1 \sum_{yx_{(m\times l_1)}} \varpi_{(l_1\times m)} \end{pmatrix}. \tag{17}$$

We distinguish the above appearance with respect to ϖ and ω and get the optimal values as:

$$\varpi_{opt_{(l_1\times m)}} = \Sigma^{-1}_{x_{(l_1\times l_1)}} \Sigma_{yx_{(l_1\times m)}} \text{ and } \omega_{opt_{(l_2\times m)}} = \Sigma^{-1}_{x_{1(l_2\times l_2)}} \Sigma_{yx_{1(l_2\times m)}}.,$$

By means of the optimum value of ϖ and ω in (17) and we may get the minimum value of Var-Cov matrix of t_p,

$$\min \Sigma_{t_{p(m\times m)}} = S'S \begin{pmatrix} \gamma_2 \Sigma_{y_{(m\times m)}} - \gamma_1 \Sigma_{yx_{(m\times l_1)}} \Sigma^{-1}_{x_{(l_1\times l_1)}} \Sigma_{yx_{(l_1\times m)}} \\ -\left(\gamma_2-\gamma_1\right)\Sigma_{yx_{1(m\times l_2)}} \Sigma^{-1}_{x_{1(l_2\times l_2)}} \Sigma_{yx_{1(l_2\times m)}} \end{pmatrix}. \tag{18}$$

Remark 1:

By putting $g_{kj}^* = 0$ and $g_{kj} = 0$ directly in (8), we may get the multivariate ratio type estimator for full and no information cases. Similarly, we may also get the bias and Var-Cov matrices for full and no information from (15) and (18) respectively.

Remark 2:

One may get a univariate estimator based on different auxiliary variables from (8) taking $j = 1$ and $k = 1, 2, 3\cdots, l$ as:

$$t_{pu} = \left[t_{puj} \right]_{(1\times 1)},$$

$$t_{pj} = s^2_{y_{1(2)}} \left(\prod_{k=1}^{l_1} \left[\frac{S^2_{x_k}}{\alpha_k s^2_{x(1)\mathrm{k}} + (1-\alpha_k) S^2_{x_k}} \right]^{g_k} \prod_{k=l_1+1}^{l_1+l_2=l} \left[\frac{s^2_{x(1)\mathrm{k}}}{\beta_k s^2_{x(2)\mathrm{k}} + (1-\beta_k) s^2_{x(1)\mathrm{k}}} \right]^{g^*_k} \right),$$

and the Var-Cov matrix for the estimator in (17) may be acquired as:

$$\Sigma_{t_{p(1\times 1)}} = S'S \begin{pmatrix} \gamma_2 \Sigma_{y_{(1\times 1)}} + \gamma_1 \varpi'_{(1\times l_1)} \Sigma_{x_{(l_1\times l_1)}} \varpi_{(l_1\times 1)} + (\gamma_2 - \gamma_1) \omega'_{(1\times l_2)} \Sigma_{x_{1\,(l_2\times l_2)}} \omega_{(l_2\times 1)} \\ -\gamma_1 \Sigma_{yx_{(1\times l_1)}} \varpi_{(l_1\times 1)} + (\gamma_1 - \gamma_2) \omega'_{(l_2\times 1)} \Sigma_{yx_{1\,(1\times l2)}} + (\gamma_1 - \gamma_2) \Sigma_{yx_{1\,(1\times l2)}} \omega_{(l_2\times 1)} - \gamma_1 \varpi'_{(1\times l_1)} \Sigma_{yx_{(l_1\times 1)}} \end{pmatrix}.$$

Special cases:

For some specific $\alpha, \beta, \mathrm{g}\,\mathrm{and}\,\mathrm{g}^*$ the estimator in (9) provides some existing estimators in the following table,

Value of $\alpha, \beta, \mathrm{g}\,\mathrm{and}\,\mathrm{g}^*$	Estimators
$\alpha = 1, g = 1 \,\&\, \mathrm{g}^* = 0$	$t_R = s^2_{y_{(2)}} \frac{S^2_x}{s^2_x}$ (Ratio estimator)
$\alpha = 1, g = -1 \,\&\, \mathrm{g}^* = 0$	$t_{pr} = s^2_{y_{(2)}} \frac{s^2_x}{S^2_x}$ (Product estimator)
$\alpha = 1, g_{opt} \,\&\, \mathrm{g}^* = 0$	$t_R = s^2_{y_{(2)}} \left(\frac{S^2_x}{s^2_x} \right)^g$ (Das and Tripathi (1978))

4. A Real-Life Application-Based Results and Discussion

To show the implementation of the MRCE estimator, the Canadian climate data is used as a real-life application, which is published by the National Oceanic and Atmospheric Administration (NOAA). The weather data considered for this study, is only for the month of May 2017 that are recorded on daily basis, for 37247 different weather Stations of Canada, and then the data is converted into per week information taking week as Week-1 (average temp. for May, 2017), Week-2 (average temp. for May, 2017), Week-3 (average temp. for May, 2017), Week-4 (average temp. for May, 2017). There were many stations where the required data is not recorded for full week, so such Stations are dropped, and only those weeks where the data is recorded for full week. The monthly average temperature data recorded in the last four years say, 2017, 2016, 2015, & 2013, are considered as the auxiliary variables, and for these years, such stations were dropped where temperature record is not present for full week. Finally, only 704 Stations which had the data regarding temperature for full week, are left. For this purpose, the tidyverse Package (2016) is used to detect, and eliminate missing values, and to transform other variables mutate some more functions.

Temperature averages (TAVGs) for the three weeks of May, 2017 are treated as the study variable Y_i, where *i*=1, 2, 3, and in the previous three years (2016, 2015 and 2013) monthly TAVGs are taken as the auxiliary variables X_i, where *i*=1, 2, 3.

A finite population approach is used to model the variance and covariance of weekly temperature across Canada using the data published by National Oceanic and Atmospheric Administration (NOAA), and description over the population characteristics is presented in table (A), covariance matrix is presented in table (B) and table (C) is for correlation matrix among variables in appendix. Now, from the given population of size $N = 704$, a first-phase sample of size $n_1 = 211$ is selected, and necessary calculations are made based on the first-phase sample. Using SRSWOR, another sample i.e., second-phase sample of size $n_2 = 105$ is selected from the already selected first-phase sample. Some necessary calculations based on second-phase sample are also computed. The statistics computed based on the first phase, and the second-phase samples, are then incorporated into the proposed MV estimator, and the mentioned existing MV estimators for getting estimate about PV. This procedure is repeated for 1000 times. Finally, determinants for the Var-Cov matrix of each MV estimator are computed. Var-Cov matrices are computed in Table 1 for the proposed estimator, and for the altered existing estimators. The percentage relative efficiencies (PRE) of each estimator as compared to the sample variance estimator is presented using the following formula:

$$PRE = \frac{\left| \Sigma_{t_0} \right|}{\left| \Sigma_{t_*} \right|} \times 100.$$

We ruminate four study variables denoted by Y's and four auxiliary variables denoted by Xs for the multivariate situation (where, X_3 and X_4 are considered as known auxiliary; X_1 and X_2 are considered as unknown auxiliary variables). We also have shown our results for univariate situations using the same four X's as auxiliary variables. We compared our suggested estimator

(t_p) with the unbiased estimator (t_0), the Isaki (1983) estimator (t_r), Upadhaya and Singh (1999) ratio type estimator (t_u) and Das and Tripathi (1978) ratio type estimator (t_d).

Table 1: Variance-covariance matrices for our proposed Estimator

Variance-covariance Matrix of Proposed Estimator				
	$\mathbf{t}_{p1}$	$\mathbf{t}_{p2}$	$\mathbf{t}_{p3}$	$\mathbf{t}_{p4}$
$\mathbf{t}_{p1}$	0.019369890	4.905212e-03	-0.002014820	-3.803579e-03
$\mathbf{t}_{p2}$	0.004905212	1.653949e-02	0.001372994	-4.379532e-05
$\mathbf{t}_{p3}$	-0.002014820	1.372994e-03	0.010945712	5.472590e-03
$\mathbf{t}_{p4}$	-0.003803579	-4.379532e-05	0.005472590	2.098642e-02
Variance-covariance Matrix of $\mathbf{t}_0$ Estimator				
	$\mathbf{t}_{01}$	$\mathbf{t}_{02}$	$\mathbf{t}_{03}$	$\mathbf{t}_{04}$
$\mathbf{t}_{01}$	0.07672491	0.05630260	0.04004221	0.03830438
$\mathbf{t}_{02}$	0.05630260	0.06388211	0.03917459	0.03796022
$\mathbf{t}_{03}$	0.04004221	0.03917459	0.04337219	0.03747599
$\mathbf{t}_{04}$	0.03830438	0.03796022	0.03747599	0.05273034
Variance-covariance Matrix of $\mathbf{t}_d$ Estimator				
	$\mathbf{t}_{d_1}$	$\mathbf{t}_{d_2}$	$\mathbf{t}_{d_3}$	$\mathbf{t}_{d_4}$
$\mathbf{t}_{d_1}$	0.05563582	0.03759989	0.02706011	0.02310313
$\mathbf{t}_{d_2}$	0.03899938	0.04860736	0.02822920	0.02613081
$\mathbf{t}_{d_3}$	0.02928598	0.02923263	0.03569492	0.03007688
$\mathbf{t}_{d_4}$	0.02455991	0.02634051	0.02936588	0.04341882
Variance-covariance Matrix of $\mathbf{t}_r$ Estimator				
	$\mathbf{t}_{r_1}$	$\mathbf{t}_{r2}$	$\mathbf{t}_{r3}$	$\mathbf{t}_{r_4}$
$\mathbf{t}_{r_1}$	0.02253536	-0.003942296	0.04004221	0.03830438
$\mathbf{t}_{r2}$	0.01325979	0.025422865	0.03917459	0.03796022
$\mathbf{t}_{r_3}$	-0.05560742	-0.062276857	0.04337219	0.03747599
$\mathbf{t}_{r_4}$	-0.05750892	-0.062667890	0.03747599	0.05273034
Variance-covariance Matrix of $\mathbf{t}_u$ Estimator				
	$\mathbf{t}_{u_1}$	$\mathbf{t}_{u_2}$	$\mathbf{t}_{u_3}$	$\mathbf{t}_{u_4}$
$\mathbf{t}_{u_1}$	0.08124260	0.06187189	0.04004221	0.03830438
$\mathbf{t}_{u_2}$	0.05990383	0.06829672	0.03917459	0.03796022
$\mathbf{t}_{u_3}$	0.04338338	0.04319781	0.04337219	0.03747599
$\mathbf{t}_{u_4}$	0.04250808	0.04290596	0.03747599	0.05273034

Table 1 shows our multivariate cases; Var-Covmatrices have been computed from (2), (4), (6), (8) and (25) to illustrate the performance of our multivariate ratio estimator. We get the (4×4) matrix and observed that the diagonal elements of each matrix show the variances for $(Y_1,..,Y_4)$. It is also detected that the diagonal elements of our anticipated estimator have the lower variances among the other diagonal elements of existing estimators. It is noted that the estimators who are using population variance as auxiliary variable are more effective than the estimators having population mean as auxiliary variable. We can also see that our proposed and modified Isaki (1983) estimators are having lesser variances on the diagonals than Upadhaya and Singh (1999) and Das and Tripathi (1978).

Table 2: Determinant and PRE's of our proposed and literature estimators

Estimator	$\mathbf{t}_p$	$\mathbf{t}_0$	$\mathbf{t}_r$	$\mathbf{t}_d$	$\mathbf{t}_u$
Determinants	4.77554e-08	6.22207e-07	3.99522e-07	5.284e-07	5.4245e-06
PRE's for Univariate cases using four auxiliary variables					
	Weekly temperature				
Estimator	Y1	Y2	Y3	Y4	
$\mathbf{t}_p$	396.10400	386.00	396.24822	251.00	
$\mathbf{t}_r$	137.90559	131.42477	121.508018	121.44581	
$\mathbf{t}_d$	340.46454	251.27817	100	100	
$\mathbf{t}_u$	94.43925	93.536135	100	100	

Table 2 summaries the results for multivariate as well as for univariate cases. By finding the determinants of the Var-Covmatrices, we find the MSE of our proposed and literature estimators and it is observed that our anticipated estimator ($\mathbf{t}_p$) has the determinant 4.77554e-08 which is less than the determinant of $\mathbf{t}_0$, $\mathbf{t}_r$, $\mathbf{t}_d$ and $\mathbf{t}_u$. We also computed the numerical results by finding PRE's for every Y using the four auxiliary variables.

5. Simulation Study

The simulation study has also been computed to assess the presentation of our anticipated multivariate ratio type estimator for estimating population variance under a two-phase sampling scheme. We used a model for generating the finite populations by using the R- Statistical package. We consider four study variables (Y_1, Y_2, Y_3, Y_4) along with four auxiliary variables (X_1, X_2, X_3, X_4) for multivariate situation (where, X_3 and X_4 are considered as known auxiliary; X_1 and X_2 are considered as unknown auxiliary variables) and normal distribution is used for generating our auxiliary variables $X's$ as:

$$X_1 \sim N(12,4),\ X_2 \sim N(15,3),\ X_3 \sim N(18,6),\ X_4 \sim N(16,5).$$

And the study variable is simulated as:

Model-I $Y_{kj} = \sum_{k=1}^{l} X + \varepsilon.$

We generated a finite population for $N = 10,000$ and nominated a sample of size $n_1 = 2000$ units by SRSWOR in the first-phase and then we designated a small sample from n_1 of $n_2 = 600$ units by SRSWOR. This technique is continual q=1000 times to compute the ideals of our proposed and existing estimators.

The code for the simulation was carried out using the R statistical package (2016) and the David Robinson (2016) broom and Hadley Wickham (2016) Tidy verse packages.

Table 3: Variance-covariance matrix for ratio estimator

Variance-covariance Matrix of Proposed Estimator				
	t_{p1}	t_{p2}	t_{p3}	t_{p4}
t_{p1}	1.4261477	1.2066771	0.5101474	0.5941651
t_{p2}	1.2066771	1.4124647	0.4883920	0.6627137
t_{p3}	0.5101474	0.4883920	0.2882620	0.3918415
t_{p4}	0.5941651	0.6627137	0.3918415	0.5995745
Variance-covariance Matrix of t_0 Estimator				
	t_{01}	t_{02}	t_{03}	t_{04}
t_{01}	1.5426375	1.3363918	0.5462599	0.6488495
t_{02}	1.3363918	1.6529650	0.5400216	0.7535844
t_{03}	0.5462599	0.5400216	0.3189577	0.4450744
t_{04}	0.6488495	0.7535844	0.4450744	0.7043424
Variance -covariance Matrix of t_r Estimator				
	t_{r_1}	t_{r2}	t_{r3}	t_{r_4}
t_{r_1}	3.733230	3.673615	1.4686897	1.986607
t_{r2}	3.673615	4.165533	1.5256706	2.188923
t_{r_3}	1.468690	1.525671	0.7053697	1.004908
t_{r_4}	1.986607	2.188923	1.0049082	1.517789
Variance-covariance Matrix of t_u Estimator				
	t_{u_1}	t_{u_2}	t_{u_3}	t_{u_4}
t_{u_1}	1.6760171	1.4843051	0.6114653	0.7431973
t_{u_2}	1.4843051	1.8177848	0.6127588	0.8596508
t_{u_3}	0.6114653	0.6127588	0.3508144	0.4914471
t_{u_4}	0.7431973	0.8596508	0.4914471	0.7720322
Variance-covariance Matrix of t_d Estimator				
	t_{d_1}	t_{d_2}	t_{d_3}	t_{d_4}
t_{d_1}	1.5840520	1.8696998	0.5453600	0.7087673
t_{d_2}	1.8696998	8.1537205	0.5015866	0.8158619
t_{d_3}	0.5453600	0.5015866	0.3312526	0.5736758
t_{d_4}	0.7087673	0.8158619	0.5736758	1.8997817

Table 3 shows our multivariate cases; Var-Covmatrices have been computed from the artificial population to observe the performance of ratio type estimator. We can see variances on the diagonals of (4×4) matrices and observe that the variance of our proposed estimator for (Y_1, Y_2, Y_3, Y_4) are less than the variances of t_0 , t_r , t_d and t_u . It shows that our planned estimator (t_p) performs better than the other existing estimators.

Table 4: Determinant of our proposed and literature estimators for Model II

Estimator	$\mathbf{t}_p$	$\mathbf{t}_0$	$\mathbf{t}_r$	$\mathbf{t}_u$	$\mathbf{t}_d$
Determinants	0.00200573	0.0043825	0.0124250	0.0048802	1.067941

We compared our multivariate ratio type estimator with existing ratio type estimators. It is observed that our proposed estimator has less MSE in all univariate cases as shown in Table 5.1. In Table 5.2, our proposed estimator ($\mathbf{t}_p$) has lesser determinant (0.00200573, which shows the MSE for the multivariate case) among others.

6. Concluding Remarks

We proposed the multivariate ratio type estimator when some auxiliary variables are known and some are unknown. We found out the bias and the Var-Covmatrix for our proposed work. Further, we discussed the bivariate and univariate form of our proposed estimator using multi-auxiliary variables. We computed empirical results using Canadian climate data (NOAA) in Section 4. In Table 1, the Var-Covmatrices are given for our proposed estimator along with existing estimators. Table 2 demonstrated the determinants for the multivariate cases and PRE's for the univariate cases. It is observed that our proposed estimator has a smaller determinant among other standing multivariate estimators which provided indication that our planned estimator is more effectual. We performed a simulation study in Section 5 and selected a model to check the efficiency of our proposed estimator. The usual ratio estimator behaves well when the simulation model generates symmetric error distributions and linear associations between the auxiliary variables and the outcome. We used the normal distribution for generating our four auxiliary variables and simulated the data for 1000 times. In Table 3 and Table 4, it is shown that our proposed estimator $\mathbf{t}_p$ is more efficient than unbiased variance estimator $\left(\mathbf{t}_0\right)$, Isaki (1983) $\left(\mathbf{t}_0\right)$, Upadhaya and Singh (1999) ($\mathbf{t}_u$) and Das and Tripathi (1978) ($\mathbf{t}_d$) and we gained the lesser MSE for univariate cases. In future, we hope to discuss the effects of covariance terms in formation of a variance-covariance matrix. We also want to propose this work under stratified sampling.

References

Abu-Dayyeh, W. and Ahmed, M.S. (2002). General Estimator for a finite population variance using multivariate auxiliary variable. *Journal of Statistical Theory and Applications*, 1(2), 143-147.

Ahmed, M.S., Raman, M.S., and Hossain, M.I., (2000). Some competitive estimators of finite population variance Multivariate Auxiliary Information. *Information and Management Sciences*. 11 (1), 49-54.

Ahmad, Z., Hussain I. and Hanif, M., (2016). Estimation of Finite Population Variance in Successive Sampling Using Multi-Auxiliary Variables. *Communication in Statistics- Theory and Methods*. 45 (3), 553-565.

Asghar, A., Sanaullah, A. and Hanif, M., (2014). Generalized exponential type estimator for population variance in survey sampling. *Revista Colombiana de Estadística*a, 37 (1), 213-224.

Abid, M., Ahmed, S., Tahir, M., Zafar Nazir, H., & Riaz, M., (2019). Improved ratio estimators of variance based on robust measures. *Scientia Iranica*, 26(4), 2484-2494.

Cekim, O., H., Kadilar, C. (2020). ln-Type Variance Estimators in Simple Random Sampling. *Pakistan Journal of Statistics and Operation Research*, 689-696.

Naz, F. A. R. A. H., Abid, M., Nawaz, T., & Pang, T., (2020). Enhancing efficiency of ratio-type estimators of population variance by a combination of information on robust location measures. *Scientia Iranica*, 27(4), 2040-2056.

Niaz, I., Sanaullah, A., Saleem, I., & Shabbir, J., (2022). An improved efficient class of estimators for the population variance. *Concurrency and Computation: Practice and Experience*, 34(4), e6620.

Das, A.K. and Tripathi, T.P., (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya*, 40, 139-148.

David Robinson (2016). broom: Convert Statistical Analysis Objects into Tidy Data Frames. R package version 0.4.1.

Graunt, J., (1662). Natural and political observations upon the Bill of Mortality London: John Martyn.

Haq, A, Usman, M., Khan, M., (2021). Estimation of finite population variance under stratified random sampling. Communications in Statistics - Simulation and Computation, 1-17.

Hadley Wickham (2016). tidyverse: Easily Install and Load 'Tidyverse' Packages. R package version 1.0.0. CRAN.R-project.org/package=tidyverse.

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

Kadilar, C., Cingi, H., (2006). Improvement in variance using auxiliary Information}, Hecettepe *Journal of Mathematics and Statistics*, 35 (1), 111-115.

Kim, J.K., Sitter, R.R., (2003). Efficient variance estimation for two-phase sampling. *Statistic Sinica*, 13, 641-653, 2003.

Olufadi, Y. and Kadilar, C., (2014). A study on the chain ratio type estimator of finite population variance. *Journal of Probability and Statistics*.

Olkin, I., (1958). Multivariate ratio estimation for finite populations. *Biometrika*, 45, 154-165.

Singh, H. P., Singh, S., Kim, J. M. (2003). Variance estimation using multiauxiliary information for random non-response in survey sampling. *Statistica*, 63 (1), 23-40.

Singh H. P., Chandra P., Singh S., (2010). Efficient use of auxiliary variables in estimating finite population variance in two-phase sampling. *Communications for Statistical Applications and Methods*, 17 (2), 165-181.
Singh, H. P., Solanki, R. S., (2013). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical Papers*, 54 (2), 479-497.
Singh, R., Singh, V. K., (2015). Improved Class of Estimators for Variance under Single and Two-Phase Sampling. *Journal of Statistics Applications and Probability*, 4 (1), 161-165.
Sanaullah, A., Hanif, M. Asghar, A., (2016). Generalized Exponential Estimators for Population Variance under Two-Phase. *International Journal of Applied Computational Mathematics*, 2 (1), 75-84, 2016.
Singh, H.P., Pal, S.K., (2016). Improved estimation of finite population variance using auxiliary information in presence of measurement error. *Revista Investigacion Operacional*, 37 (2), 147-162.
Ismail, M., Kanwal, N., Shahbaz, M. Q., (2018). Generalized ratio-product-type estimator for variance using auxiliary information in simple random sampling, *Kuwait Journal of Science*, 45 (1), 79-88.
Upadhyaya, L. N., Singh, H. P., (1999). An estimator for population variance that utilizes the kurtosis of an auxiliary variable in sample surveys. *Vikram Mathemat*, 19, 4-17.
Daraz, U., Khan, M., (2021). Estimation of variance of the difference-cum-ratio-type exponential estimator in simple random sampling. *RMS: Research in Mathematics & Statistics*, 8(1), 1899402.
Yadav, S.K., Kumar, R., Kumar, S., Verma, S., Kumar, S., (2015). Efficient Estimator of Population Variance Using of Coefficient of Kurtosis and Population Mean of Auxiliary Variable. *International Journal of Mathematics and its Application*, 3 (4), 85-91.
Zaman, T., Bulut, H., (2021). An efficient family of robust-type estimators for the population variance in simple and stratified random sampling. *Communications in Statistics-Theory and Methods*, 1-15.

Appendix

Table A: Details of Variables for Population

Population	NOAA
Y_1	Week-1 (average temp. for May, 2017)
Y_2	Week-2 (average temp. for May, 2017)
Y_3	Week-3 (average temp. for May, 2017)
Y_4	Week-4 (average temp. for May, 2017)
X_1	Year-2016 (average temp. for May)
X_2	Year-2015 (average temp. for May)
X_3	Year-2014 (average temp. for May)
X_4	Year-2013 (average temp. for May)

Table B: Covariance and Correlation

Population	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4
Y_1	16.4344	11.7119	11.8797	11.2634	5.0995	2.0234	5.6778	0.5231
Y_2	11.7119	10.0879	9.3665	9.8238	1.3889	2.5090	5.0955	0.7491
Y_3	11.8797	9.3666	9.7386	9.3817	2.8759	2.5086	3.8010	1.3053
Y_4	11.2634	9.8238	9.3817	9.7610	1.3933	3.0050	4.4580	0.9370
X_1	5.0996	1.3888	2.8759	1.3933	7.4839	0.0406	-0.0779	-0.0538
X_2	2.0234	2.5090	2.5086	3.0050	0.0406	4.6808	-0.0304	0.0080
X_3	5.6777	5.0955	3.8009	4.4571	-0.0779	-0.0304	6.6511	-0.0093
X_4	0.5231	0.7490	1.3052	0.9369	-0.0538	0.0080	-0.0093	2.0420

Table C: Correlation Matrix

Population	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4
Y_1	1.0000	0.8869	0.7494	0.7654	0.9026	0.8528	0.8650	0.9122
Y_2	0.8869	1.0000	0.7215	0.7535	0.8397	0.7789	0.8273	0.8513
Y_3	0.7494	0.7215	1.0000	0.7959	0.7917	0.8683	0.8342	0.7632
Y_4	0.7654	0.7535	0.7959	1.0000	0.8335	0.8118	0.8301	0.8193
X_1	0.9026	0.8397	0.7917	0.8335	1.0000	0.9420	0.9561	0.9453
X_2	0.8528	0.7789	0.8683	0.8118	0.9420	1.0000	0.9667	0.9208
X_3	0.8650	0.8273	0.8342	0.8301	0.9561	0.9667	1.0000	0.9257
X_4	0.9122	0.8513	0.7632	0.8193	0.9453	0.9208	0.9257	1.0000