Generalized Exponential Estimators for Population Variance Under Two-Phase Sampling

Aamir Sanaullah¹ · Muhammad Hanif² ·
Amber Asghar³

Published online: 7 April 2015
© Springer India Pvt. Ltd. 2015

Abstract In this study, two-phase sampling is considered for estimating the population variance of study variable taking two auxiliary variables. The proposed generalized estimator and class of estimators are the exponential function of auxiliary variables. The mean square errors and biases equations have been obtained for the proposed estimators. The conditions for which proposed estimators are more efficient as compared to other estimators have been discussed. The empirical study showed that proposed estimators are more efficient as compared to the unbiased sample variance estimator, double sampling version of Isaki (J Am Stat Assoc 78:117–123, 1983) and Singh et al. (Ital. J. Pure Appl. Math. 28(N):101–108, 2011) generalized estimator.

Keywords Population variance · Generalized exponential estimator · Two-phase sampling · Auxiliary information

Introduction

Neyman [14] is the pioneer of two-phase sampling designs. Rao [16] used it for stratification and non-response problems. Cochran [4–6, 17], and Khan et al. [12], developed some estimators for two-phase sampling. The auxiliary information has been effectively used in two-phase sampling to estimate an unknown population characteristic. Hanif et al. [9] and

✉ Aamir Sanaullah
chaamirsanaullah@yahoo.com; draamirsanaullah@ciitlahore.edu.pk

Muhammad Hanif
drmianhanif@gmail.com

Amber Asghar
zukhruf10@gmail.com

¹ Department of Statistics, COMSATS Institute of IT, Lahore, Pakistan

² Department of Statistics, NCBA&E, Lahore, Pakistan

³ Department of Statistics, Virtual University of Pakistan Lahore, Lahore, Pakistan

[10] proposed regression type estimators for the situation if auxiliary information is available at second-phase. Bahl and Tuteja [3] suggested the exponential estimator under simple random sampling without replacement. Singh et al. [21] proposed exponential type estimators under two-phase sampling. Sanaullah et al. [18, 19] proposed some generalized exponential estimators for stratified two-phase sampling.

Isaki [11] first suggested the ratio estimator for population variance. Prasad and Singh [15] proposed a ratio type estimator for estimating population variance which was an improved version of [11] estimator. Many more authors, including [1, 8, 20, 24, 25, 27, 28], proposed some ratio-type estimators of population variance. Yadav and Kadilar [31] proposed an exponential ratio-type estimator for the estimation of population variance and [2] proposed some generalized exponential-type estimators using mean of auxiliary information for estimating the population variance.

Consider a finite population consisting of N units $V = (V_1, V_2, \dots, V_N)$. Let y and (x, z) be the variable of interest and auxiliary characteristics respectively with population means $(\bar{Y}, \bar{X}, \bar{Z})$ and population variances (S_y^2, S_x^2, S_z^2) respectively. Let a simple random sample without replacement (SRSWOR) $S_1 (\subset V)$ of size $n_1 (\subset N)$ units are drawn at first-phase and from these units information on auxiliary characteristics say (x', z') are collected. At second-phase another sample $S_2 (\subset S_1)$ of size $n_2 (\subset n_1)$ is selected from the first sample S_1 and information related to the variable of interest y and auxiliary information (x, z) are gathered. Let $(\bar{x}', \bar{z}')$ be means of two auxiliary variables based on first-phase sample and $(\bar{y}, \bar{x}, \bar{z})$ be the means based on the information taken from second-phase sample.

The following notations have been considered in order to estimate the bias and mean square error expressions,

$$s_y^2 = S_y^2 (1 + e_0), \quad \bar{x}' = \bar{X} (1 + e'_1), \quad \bar{x} = \bar{X} (1 + e_1), \quad \bar{z}' = \bar{Z} (1 + e'_2), \quad \bar{z} = \bar{Z} (1 + e_2).$$

We may obtain the following results under SRSWOR if finite population correction factor is ignored and assume that the population size N is large.

$$\begin{aligned} E(e_0) &= E(e'_1) = E(e'_2) = E(e_2) = 0, \\ E(e_0^2) &= \frac{1}{n_2} (\delta_{400} - 1), \quad E(e_1'^2) = \frac{C_x^2}{n_1}, \quad E(e_1'^2) = \frac{C_x^2}{n_2}, \\ E(e_2'^2) &= \frac{C_z^2}{n_1}, \quad E(e_2'^2) = \frac{C_z^2}{n_2}, \quad E(e_0 e'_1) = \frac{\delta_{210} C_x}{n_1}, \quad E(e_0 e_1) = \frac{\delta_{210} C_x}{n_2}, \\ E(e_0 e'_2) &= \frac{\delta_{201} C_z}{n_1}, \quad E(e_0 e_2) = \frac{\delta_{201} C_z}{n_2}, \quad E(e_1 e_2) = \frac{\rho_{xz} C_x C_z}{n_2}, \end{aligned}$$

where

$$\delta_{pqr} = \frac{\mu_{pqr}}{\mu_{200}^{p/2} \mu_{020}^{q/2} \mu_{002}^{r/2}}, \quad \mu_{pqr} = \frac{1}{N} \sum_i^N (Y_i - \bar{Y})^p (X_i - \bar{X})^q (Z_i - \bar{Z})^r$$

(p, q, r) be the non-negative integers, μ_{pqr} and δ_{pqr} are the moments and moment's ratio of order p, q and r respectively further μ_{200}, μ_{020} , and μ_{002} , are the second order moments and $C_x = S_x/\bar{X}$ and $C_z = S_z/\bar{Z}$ are the coefficient of variations for auxiliary variables X and Z respectively.

Some Available Estimators

In this section we reproduce some estimators for the population variance available in literature.

An unbiased estimator for population variance is defined as,

$$t_0 = s_y^2. \quad (1)$$

The variance of t_0 is

$$\text{Var}(t_0) = \left(\frac{S_y^4}{n_2} \right) [\delta_{400} - 1]. \quad (2)$$

A two-phase sampling version of [11] for no-information case of ratio estimator is,

$$t_I = s_y^2 \frac{s_{x_1}^2}{s_{x_2}^2} \frac{s_{z_1}^2}{s_{z_2}^2}. \quad (3)$$

The mean square error of t_I is,

$$\text{MSE}(t_I) = S_y^4 \left[\frac{1}{n_2} (\delta_{400} - 1) + \left(\frac{1}{n_2} - \frac{1}{n_1} \right) (\delta_{040} - 1) (1 - 2C) \right]. \quad (4)$$

Singh et al. [21] suggested a generalized exponential ratio-product type estimator for population variance in two-phase sampling as,

$$t_{SE} = s_y^2 \left[k \exp \left\{ \frac{s_{x_1}^2 - s_{x_2}^2}{s_{x_1}^2 + s_{x_2}^2} \right\} + (1 - k) \exp \left\{ \frac{s_{z_1}^2 - s_{z_2}^2}{s_{z_1}^2 + s_{z_2}^2} \right\} \right]. \quad (5)$$

The bias and mean square errors of t_{se} is,

$$\text{Bias}(t_{SE}) \approx S_y^2 \left[\frac{1}{4} \left(\frac{1}{n_2} - \frac{1}{n_1} \right) (\delta_{040} - 1) + \left(\frac{1}{n_1} - \frac{1}{n_2} \right) (\delta_{220} - 1) \right], \quad (6)$$

and

$$\text{MSE}(t_{SE}) \approx S_y^4 \left[\begin{aligned} & \frac{1}{n_2} (\delta_{400} - 1) + \frac{k^2}{4} \left(\frac{1}{n_2} - \frac{1}{n_1} \right) (\delta_{040} - 1) \\ & + \frac{(k^2 - 1)}{4} \left(\frac{1}{n_2} - \frac{1}{n_1} \right) (\delta_{004} - 1) + k \left(\frac{1}{n_2} - \frac{1}{n_1} \right) (\delta_{220} - 1) \\ & + (k - 1) \left(\frac{1}{n_1} - \frac{1}{n_2} \right) (\delta_{202} - 1) - \frac{k(k-1)}{2} \left(\frac{1}{n_1} - \frac{1}{n_2} \right) (\delta_{022} - 1) \end{aligned} \right], \quad (7)$$

where

$$C = \frac{(\delta_{220} - 1)}{(\delta_{040} - 1)} \quad \text{and} \quad k = \frac{\{\delta_{004} + 2(\delta_{220} - 1) + \delta_{022} - 6\}}{(\delta_{040} + \delta_{004} + 2\delta_{022} - 4)}.$$

The motivation is to look up some generalized exponential-type estimators for estimating the population variance. A vital role of estimating variance in survey sampling is to dig up information on the accuracy of the estimator and set aside confidence intervals. In this research paper, an approach of utilizing sample means of the two auxiliary variables in two-phase sampling has been made for estimating the population variance instead of utilizing variances of auxiliary variables. This approach gives such a form that is simpler and can be implemented into computer software.

Generalized Exponential Type Ratio and Product Estimator

The proposed estimator is defined as,

$$t_{GE} = \lambda s_y^2 \exp \left[\alpha \left(\frac{a \bar{x}_1}{(a-1)\bar{x}_1 + \bar{x}_2} - 1 \right) \right] \exp \left[\beta \left(\frac{b \bar{z}_1}{(b-1)\bar{z}_1 + \bar{z}_2} - 1 \right) \right], \quad (8)$$

where (a, b) and λ (principal) are some constants and assumed to be estimated such that the mean square error of t_{GE} is minimum, and $\alpha, \beta (-\infty, \infty)$ are some constants for generalization.

It is observed that for different values of λ, α, β and a, b in (8), we may get various exponential ratio-cum-ratio taking positive values of $\alpha, \beta (> 0)$ and product-cum-product estimators taking negative values of $\alpha, \beta (< 0)$ as new family of t_{GE} .

In order to obtain the bias and MSE, (8) may be expressed in the form of e 's as,

$$t_{GE} = \lambda S_y^2 (1 + e_0) \exp \left[\frac{\alpha (e'_1 - e_1)}{a} \left\{ 1 + \left\{ e'_1 - \frac{(e'_1 - e_1)}{a} \right\}^{-1} \right\} \right] \\ \times \exp \left[\frac{\beta (e'_2 - e_2)}{b} \left\{ 1 + \left\{ e'_2 - \frac{(e'_2 - e_2)}{b} \right\}^{-1} \right\} \right]. \quad (9)$$

Expanding the exponentials and ignoring terms in e_0 and e_1 of the order more than one, we may have on simplification,

$$t_{GE} = \lambda S_y^2 \left[1 + e_0 + \frac{\alpha(e'_1 - e_1)}{a} + \frac{\beta(e'_2 - e_2)}{b} + \frac{\alpha e_0(e'_1 - e_1)}{a} + \frac{\beta e_0(e'_2 - e_2)}{b} - \left(\frac{\alpha}{a^2} + \frac{\alpha^2}{2a^2} \right) e_1'^2 \right. \\ \left. + \left(\frac{\alpha}{a^2} + \frac{\alpha^2}{2a^2} \right) e_1'^2 - \left(\frac{\beta}{b^2} + \frac{\beta^2}{2b^2} \right) e_2'^2 + \left(\frac{\beta}{b^2} + \frac{\beta^2}{2b^2} \right) e_2'^2 + \frac{\alpha\beta(e'_1 - e_1)(e'_2 - e_2)}{ab} \right]. \quad (10)$$

In order to obtain the bias, subtracting S_y^2 and taking expectation of (10), and after some simplification, we get the bias as,

$$Bias(t_{GE}) \\ = S_y^2 E \left[\lambda \left\{ 1 + e_0 + \frac{\alpha(e'_1 - e_1)}{a} + \frac{\beta(e'_2 - e_2)}{b} + \frac{\alpha e_0(e'_1 - e_1)}{a} + \frac{\beta e_0(e'_2 - e_2)}{b} - \left(\frac{\alpha}{a^2} + \frac{\alpha^2}{2a^2} \right) e_1'^2 + \left(\frac{\alpha}{a^2} + \frac{\alpha^2}{2a^2} \right) e_1'^2 - \left(\frac{\beta}{b^2} + \frac{\beta^2}{2b^2} \right) e_2'^2 + \left(\frac{\beta}{b^2} + \frac{\beta^2}{2b^2} \right) e_2'^2 \right. \right. \\ \left. \left. + \frac{\alpha\beta(e'_1 - e_1)(e'_2 - e_2)}{ab} \right\} - 1 \right].$$

or

$$Bias(t_{GE}) = S_y^2 \left[\lambda \left\{ 1 + \left(\frac{\alpha}{a^2} + \frac{\alpha^2}{2a^2} \right) \left(\frac{1}{n_2} - \frac{1}{n_1} \right) C_x^2 + \left(\frac{\beta}{b^2} + \frac{\beta^2}{2b^2} \right) \left(\frac{1}{n_2} - \frac{1}{n_1} \right) C_z^2 \right. \right. \\ \left. \left. - \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\alpha\delta_{210}C_x}{a} - \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\beta\delta_{201}C_z}{b} + \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\alpha\beta\rho C_x C_z}{ab} \right\} - 1 \right]. \quad (11)$$

In order to develop expression of the minimum mean square error, it would not be possible to estimate the optimal values of the constants (a, b) if we retain the terms of order one in e 's. It is therefore we are considering (10) again and ignoring the terms in e 's with power two or more as,

$$t_{GE} - S_y^2 \approx S_y^2 E \left[\lambda \left\{ 1 + e_0 + \frac{\alpha e'_1}{a} - \frac{\alpha e_1}{a} + \frac{\beta e'_2}{b} - \frac{\beta e_2}{b} \right\} - 1 \right]. \quad (12)$$

Squaring both sides and taking the expectation we get the MSE of t_{GE}

$$E \left(t_{GE} - S_y^2 \right)^2 \approx S_y^4 E \left[\lambda \left\{ 1 + e_0 + \frac{\alpha e'_1}{a} - \frac{\alpha e_1}{a} + \frac{\beta e'_2}{b} - \frac{\beta e_2}{b} \right\} - 1 \right]^2. \quad (13)$$

On simplifying we may get,

$$MSE(t_{GE}) \approx S_y^4 \left[1 + \lambda^2 \left\{ 1 + \frac{(\delta_{400}-1)}{n_2} + \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\alpha^2 C_x^2}{a^2} + \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\beta^2 C_z^2}{b^2} \right. \right. \\ \left. \left. - 2 \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\alpha \delta_{210} C_x}{a} - 2 \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\beta \delta_{201} C_z}{b} \right. \right. \\ \left. \left. + 2 \left(\frac{1}{n_2} - \frac{1}{n_1} \right) \frac{\alpha \beta \rho C_x C_z}{ab} \right\} - 2\lambda \right]. \quad (14)$$

Or

$$MSE(t_{GE}) \approx S_y^4 \left[1 + \lambda^2 \left\{ 1 + \frac{(\delta_{400}-1)}{n_2} + \theta w_1^2 C_x^2 + \theta w_2^2 C_z^2 - 2\theta w_1 \delta_{210} C_x \right. \right. \\ \left. \left. - 2\theta w_2 \delta_{201} C_z + 2\theta w_1 w_2 \rho C_x C_z \right\} - 2\lambda \right], \quad (15)$$

where

$$w_1 = \alpha/a \quad \text{and} \quad w_2 = \beta/b.$$

In order to minimize $MSE(t_{GE})$, we differentiate the $MSE(t_{GE})$ partially with respect to λ , w_1 and w_2 , and equating the results to zero. We get the optimum values of λ , w_1 and w_2 as

$$w_1 = (\delta_{210} - \delta_{201} \rho) \left((1 - \rho) C_x \right)^{-1}, \quad w_2 = (\delta_{201} - \delta_{210} \rho) \left((1 - \rho) C_z \right)^{-1},$$

$$\lambda = \frac{1}{(1 + V_{400} + M)},$$

where

$$M = \theta \left\{ w_1^2 C_x^2 + w_2^2 C_z^2 - 2w_1 C_x \delta_{210} - 2w_2 C_z \delta_{201} + 2\rho w_1 w_2 C_x C_z \right\},$$

$$\theta = \frac{1}{n_2} - \frac{1}{n_1} \quad \text{and} \quad V_{400} = \frac{\delta_{400} - 1}{n_2}.$$

The minimum $MSE(t_{GE})$ is obtained as,

$$MSE_{\min}(t_{GE}) \approx S_y^4 \left(1 - \frac{1}{1 + V_{400} + M} \right), \quad (16)$$

On substituting the optimal values of $\lambda = (1 + V_{400} + M)^{-1}$, α , β , a and b into (8), we may get the asymptotically optimal estimator as,

$$t_{asym} = \frac{s_y^2}{(1 + V_{400} + M)} \left\{ \frac{\exp \left(\frac{(\delta_{210} - \delta_{201} \rho) \{\bar{x}_1 - \bar{x}_2\}}{(C_x - C_x \rho^2 - 1)(\bar{x}_1 + \bar{x}_2)} \right)}{\exp \left(\frac{(\delta_{201} - \delta_{210} \rho) \{\bar{z}_1 - \bar{z}_2\}}{(C_z - C_z \rho^2 - 1)(\bar{z}_1 + \bar{z}_2)} \right)} \right\}. \quad (17)$$

The values of λ , α , β , a and b can be obtained in prior from the previous surveys, for case in point, see [1, 13, 22, 26, 30, 31].

Table 1 Family of the estimator t_{GE}

Exponential-type ratio-cum-r ratio & product-cum-product estimators	λ	α	β	a	b
$t_{1E} = s_y^2 \exp\left[\frac{\bar{x}_1 - \bar{x}_2}{\bar{x}_1 + \bar{x}_2}\right] \exp\left[\frac{\bar{z}_1 - \bar{z}_2}{\bar{z}_1 + \bar{z}_2}\right]$	1	1	1	2	2
$t_{2E} = s_y^2 \exp\left[\frac{\bar{x}_1 - \bar{x}_2}{\bar{x}_2}\right] \exp\left[\frac{\bar{z}_1 - \bar{z}_2}{\bar{z}_2}\right]$	1	1	1	1	1
$t_{3E} = s_y^2 \exp\left[-\left\{\frac{\bar{x}_1 - \bar{x}_2}{\bar{x}_1 + \bar{x}_2}\right\}\right] \exp\left[-\left\{\frac{\bar{z}_1 - \bar{z}_2}{\bar{z}_1 + \bar{z}_2}\right\}\right]$	1	-1	-1	2	2
$t_{4E} = s_y^2 \exp\left[-\left\{\frac{\bar{x}_1 - \bar{x}_2}{\bar{x}_2}\right\}\right] \exp\left[-\left\{\frac{\bar{z}_1 - \bar{z}_2}{\bar{z}_2}\right\}\right]$	1	-1	-1	1	1

In some situations, for the practitioner it is not possible to presume the values of λ , α , β , a and b by employ all the resources, it is worth sensible to replace λ , α , β , a and b (18) by their consistent estimates as,

$$\hat{\lambda} = \frac{1}{(1 + \hat{V}_{400} + M)}, \hat{w}_1 = (\hat{\delta}_{210} - \delta_{201}\hat{\rho}) \left((1 - \hat{\rho}^2) \hat{C}_x \right)^{-1},$$

$$\hat{w}_2 = (\hat{\delta}_{201} - \hat{\delta}_{210}\hat{\rho}) \left((1 - \hat{\rho}^2) \hat{C}_z \right)^{-1}. \quad (18)$$

$\hat{\delta}_{210}$, $\hat{\delta}_{201}$ and $\hat{C}_x$, $\hat{C}_z$ respectively are the consistent estimates of δ_{210} , δ_{201} and C_x , C_z .

As a result, the estimator in (18) may be obtained as,

$$\hat{t}_{asym} = \frac{s_y^2}{(1 + \hat{V}_{400} + \hat{M})} \left\{ \frac{\exp\left(\frac{(\hat{\delta}_{210} - \hat{\delta}_{201}\hat{\rho})(\bar{x}_1 - \bar{x}_2)}{(\hat{C}_x - \hat{C}_x\hat{\rho}^2 - 1)(\bar{x}_1 + \bar{x}_2)}\right)}{\exp\left(\frac{(\hat{\delta}_{201} - \hat{\delta}_{210}\hat{\rho})(\bar{z}_1 - \bar{z}_2)}{(\hat{C}_z - \hat{C}_z\hat{\rho}^2 - 1)(\bar{z}_1 + \bar{z}_2)}\right)} \right\}. \quad (19)$$

Thus in practical situation we may use the estimator $\hat{t}_{asym}$. The bias and MSE expressions for the new family of t_{GE} reported in Table 1, may be obtained by putting different values of $(\lambda, \alpha, \beta, a, b)$ in (11) and (14) respectively, such as,

For $\lambda = 1, \alpha = \beta = 1$ and $a = b = 2$ in (11), bias of t_{1E} is,

$$Bias(t_{1E}) \approx \theta S_y^2 \left[\frac{3}{8} C_x^2 + \frac{3}{8} C_z^2 - \frac{1}{2} \delta_{210} C_x - \frac{1}{2} \delta_{201} C_z + \frac{1}{4} \rho C_x C_z \right]. \quad (20)$$

For $\lambda = 1, \alpha = \beta = 1$ and $a = b = 1$ in (11), bias of t_{2E} is,

$$Bias(t_{2E}) \approx \theta S_y^2 \left[\frac{3}{2} C_x^2 + \frac{3}{2} C_z^2 - \delta_{210} C_x - \delta_{201} C_z + \rho C_x C_z \right]. \quad (21)$$

For $\lambda = 1, \alpha = \beta = -1$ and $a = b = 2$ in (11), bias of t_{3E} is,

$$Bias(t_{3E}) \approx \theta S_y^2 \left[\frac{3}{8} C_x^2 + \frac{3}{8} C_z^2 + \frac{1}{2} \delta_{210} C_x + \frac{1}{2} \delta_{201} C_z + \frac{1}{4} \rho C_x C_z \right]. \quad (22)$$

For $\lambda = 1, \alpha = \beta = -1$ and $a = b = 1$ in (11), bias of t_{4E} is,

$$Bias(t_{4E}) \approx \theta S_y^2 \left[\frac{3}{2} C_x^2 + \frac{3}{2} C_z^2 + \delta_{210} C_x + \delta_{201} C_z + \rho C_x C_z \right]. \quad (23)$$

For $\lambda = 1, \alpha = \beta = 1$ and $a = b = 2$ in (14), MSE of t_{1E} is,

$$MSE(t_{1E}) \approx S_y^4 \left[\frac{\delta_{400} - 1}{n_2} + \theta \left(\frac{1}{4} C_x^2 + \frac{1}{4} C_z^2 - \delta_{210} C_x - \delta_{201} C_z + \frac{1}{2} \rho C_x C_z \right) \right]. \quad (24)$$

For $\lambda = 1, \alpha = \beta = 1$ and $a = b = 1$ in (14), MSE of t_{2E} is,

$$MSE(t_{2E}) \approx S_y^4 \left[\frac{\delta_{400} - 1}{n_2} + \theta (C_x^2 + C_z^2 - 2\delta_{210}C_x - 2\delta_{201}C_z + 2\rho C_x C_z) \right]. \quad (25)$$

For $\lambda = 1, \alpha = \beta = -1$ and $a = b = 2$ in (14), MSE of t_{3E} is,

$$MSE(t_{3E}) \approx S_y^4 \left[\frac{\delta_{400} - 1}{n_2} + \theta \left(\frac{1}{4}C_x^2 + \frac{1}{4}C_z^2 + \delta_{210}C_x + \delta_{201}C_z + \frac{1}{2}\rho C_x C_z \right) \right]. \quad (26)$$

For $\lambda = 1, \alpha = \beta = -1$ and $a = b = 1$ in (14), MSE of t_{4E} is,

$$MSE(t_{4E}) \approx S_y^4 \left[\frac{\delta_{400} - 1}{n_2} + \theta (C_x^2 + C_z^2 + 2\delta_{210}C_x + 2\delta_{201}C_z + 2\rho C_x C_z) \right]. \quad (27)$$

Relative Performance of the Estimators

We find the efficiency conditions under which the generalized exponential-type estimator t_{GE} in (17) performs better than different existing estimators as,

1. The estimator t_{GE} will be precise estimator over unbiased variance estimator when $Var(t_0) - \min MSE(t_{GE}) > 0$ and this will be true if

$$\lambda > \kappa_1, \quad \text{where} \quad \kappa_1 = 1 - V_{400}.$$

2. The estimator t_{GE} will be precise estimator over double version of [11] ratio estimator when $MSE(t_I) - \min MSE(t_{GE}) > 0$ and this will be hold if

$$\lambda > \kappa_1 - \theta [(\delta_{040} - 1)(1 - 2C)], \quad \text{where} \quad C = \frac{(\delta_{220} - 1)}{(\delta_{040} - 1)}.$$

3. The estimator t_{GE} will be precise estimator than [21] generalized estimator when $MSE(t_{SE}) - \min MSE(t_{GE}) > 0$ and this may be true if

$$\begin{aligned} \lambda > \kappa_1 - \theta_2 L \quad \text{where} \\ L = \theta \left[\frac{k^2 (\delta_{040} - 1)}{4} + \frac{(k^2 - 1)}{4} (\delta_{004} - 1) + k (\delta_{220} - 1) \right. \\ \left. - (k - 1) (\delta_{202} - 1) + \frac{k(k - 1)}{2} (\delta_{022} - 1) \right]. \end{aligned}$$

4. The estimator t_{GE} will be precise estimator over t_{1E} estimator when $MSE(t_{1E}) - \min MSE(t_{GE}) > 0$ and this will true if

$$\lambda > \kappa_1 - \theta L_1, \quad \text{where} \quad L_1 = \left[\frac{1}{4}C_x^2 + \frac{1}{4}C_z^2 - \delta_{210}C_x - \delta_{201}C_z + \frac{1}{2}\rho C_x C_z \right].$$

5. The estimator t_{GE} will be precise estimator over t_{2E} when $MSE(t_{2E}) - \min MSE(t_{GE}) > 0$ and this is true if

$$\lambda > \kappa_1 - \theta L_2, \quad \text{where} \quad L_2 = \left[\frac{1}{4}C_x^2 + \frac{1}{4}C_z^2 - 2\delta_{210}C_x - 2\delta_{201}C_z + 2\rho C_x C_z \right].$$

Table 2 Description of variables for population 1–2

Population	Y	X	Z
1	Food cost	Size	Income
2	Average miles per gallon	Top seed, miles per hour	Engine horsepower

Table 3 Parameters of populations

Parameter	Population number	
	1	2
N	33	81
n_1	9	21
n_2	4	9
$\bar{Y}$	27.49091	33.8465
$\bar{X}$	3.7272	112.4568
$\bar{Z}$	72.5454	117.4691
C_y	0.362887	0.29523
C_x	0.40250	0.1248
C_z	0.143577	0.4831
ρ_{yx}	0.42378	−0.691135
ρ_{yz}	0.252166	−0.7948
ρ_{xz}	−0.065989	0.967
δ_{400}	5.72	3.59
δ_{210}	0.6305	0.05137
δ_{201}	0.5506	0.1209
δ_{040}	2.380	6.820
δ_{004}	2.14	3.79
δ_{220}	1.430	2.110
δ_{202}	2.287	1.595
δ_{022}	1.492	2.172

6. The estimator t_{GE} will be precise estimator over t_{3E} when $MSE(t_{3E}) - \min MSE(t_{GE}) > 0$ and this will true if

$$\lambda > \kappa_1 - \theta L_3, \quad \text{where } L_3 = \left[\frac{1}{4}C_x^2 + \frac{1}{4}C_z^2 + \delta_{210}C_x + \delta_{201}C_z + \frac{1}{2}\rho C_x C_z \right].$$

7. The estimator t_{GE} will be precise estimator over t_{4E} when $MSE(t_{4E}) - \min MSE(t_{GE}) > 0$ and this will be true if

$$\lambda > \kappa_1 - \theta L_4, \quad \text{where } L_4 = \left[\frac{1}{4}C_x^2 + \frac{1}{4}C_z^2 + 2\delta_{210}C_x + 2\delta_{201}C_z + 2\rho C_x C_z \right].$$

Empirical Results and Discussions

In order to examine the performance of proposed generalized estimator and class of estimator under two-phase sampling, we have taken two different real populations data. The sources of

Table 4 Percentage relative efficiency (PRE) for the estimators under two-phase sampling designs

Estimator	Population (#)	
	1	2
t_0	100	100
t_I	94.4	*
t_{SE}	88.945	
$t_{GE_{\min}}$	227.692	130.719
t_{1E}	103.520	*
t_{2E}	106.13	*
t_{3E}	*	96.67
t_{4E}	*	91.47

populations are [4, p. 23] and [7, p. 433]. The description and parameters of the populations are given in Tables 2 and 3.

Table 4 indicates that the performance of the generalized exponential estimator is better than usual unbiased estimator (t_0), modified double version of [11] ratio estimator and [21] generalized exponential estimator (t_{SE}) for both populations. Table 4 indicates that the performance of the generalized exponential estimator is better than usual unbiased estimator (t_0), modified double version of [11] ratio estimator and [21] generalized exponential estimator (t_{SE}) for both populations.

Conclusion

We have suggested the generalized exponential estimator for estimating population variance. From the results given in “Empirical Results and Discussions” section we infer that the proposed estimator is better than the existing estimators mentioned in literature in the sense of having less mean square error. Hence, the proposed estimator is recommended for its practical use.

References

1. Ahmed, M.S., Raman, M.S., Hossain, M.I.: Some competitive estimators of finite population variance multivariate auxiliary information. *Inf. Manag. Sci.* **11**(1), 49–54 (2000)
2. Asghar, A., Sanaullah, A., Hanif, M.: Generalized exponential- type estimator for population variance in survey sampling. *Revis. Colomb. Estad.* **37**, 211–222 (2014)
3. Bahl, S., Tuteja, R.K.: Ratio and product type exponential estimator. *Inf. Optim. Sci.* **12**, 159–163 (1991)
4. Cochran, W.G.: *Sampling Techniques*. John Wiley, New York (1977)
5. Chaudhuri, A., Roy, D.: Model assisted survey sampling strategy in two phases. *Metrika* **41**, 355–362 (1994)
6. Dash, P.R., Mishra, G.: An improved class of estimators in two-phase sampling using two auxiliary variables. *Commun. Stat. Theory Methods* **40**, 4347–4352 (2011)
7. Gujarati, D.N.: *Basic Econometrics*, 4th edn. The McGraw-Hill Companies, New York (2004)
8. Gupta, S., Shabbir, J.: Variance estimation in simple random sampling using auxiliary information. *Hacet. J. Math. Stat.* **37**(1), 57–67 (2008)
9. Hanif, M., Hamad, N., Shahbaz, M.Q.: Some new regression type estimators in two phase sampling. *World Appl. Sci. J.* **8**(7), 799–803 (2010)
10. Hammad, N., Hanif, M., Haider, N.: A regression type estimator with two auxiliary variables for two-phase sampling. *Open J Stat.* **3**, 74–78 (2013)
11. Isaki, C.: Variance estimation using auxiliary information. *J. Am. Stat. Assoc.* **78**, 117–123 (1983)

12. Khan, H., Shouket, S., Sanaullah, A.: A class of improved estimators for estimating population mean regarding partial information in double sampling. *Glob. J. Sci. Front. Res. Math. Decis. Sci.* **12**(14), 2249–24626 (2012)
13. Murthy, M.: *Sampling Theory and Methods*. Calcutta Statistical Publishing Society, Kolkata (1967)
14. Neyman, J.: Contributions to the theory of sampling human populations. *J. Am. Stat. Assoc.* **33**, 101–116 (1938)
15. Prasad, B., Singh, H.P.: Some improved ratio type estimators of finite population variance in sample surveys. *Commun. Stat. Theory Methods* **19**, 1127–1139 (1990)
16. Rao, J.N.K.: On double sampling for stratification and analytic surveys. *Biometrika* **6**, 125–133 (1973)
17. Samiuddin, M., Hanif, M.: Estimation of population mean in single phase and two phase sampling with or without additional information. *Pak. J. Stat.* **23**(2), 99–118 (2007)
18. Sanaullah, A., Ali, A.H., Noor-ul-Amin, M., Hanif, M.: Generalized exponential chain ratio estimators under stratified two-phase random sampling. *Appl. Math. Comput.* **226**, 541–547 (2014)
19. Sanaullah, A., Khan, H., Ali, A.H., Singh, R.: Improved ratio-type estimators in survey sampling. *J. Reliab. Stat. Stud.* **5**(2), 119–132 (2012)
20. Sharma, P., Singh, R.: A generalized class of estimators for finite population variance in presence of measurement errors. *J. Mod. Appl. Stat. Methods* **12**(2), 231–241 (2013)
21. Singh, R., Chauhan, P., Sawan, N., Smarandache, F.: Improved exponential estimator for population variance using two auxiliary variables. *Ital. J. Pure Appl. Math.* **28**(N), 101–108 (2011)
22. Singh, H., Karpe, N.: Estimation of mean, ratio and product using auxiliary information in the presence of measurement errors in sample surveys. *J. Stat. Theory Pract.* **4**(1), 111–136 (2010)
23. Singh, R., Kumar, M.: A note on transformations on auxiliary variable in survey sampling. *Model Assist. Stat. Appl.* **6**(1), 17–19 (2011)
24. Singh, H.P., Ramkrishna, S.: Improved estimation of finite population variance using auxiliary information. *Commun. Stat. Theory Methods* **42**, 2718–2730 (2013)
25. Singh, H.P., Solanki, R.S.: A new procedure for variance estimation in simple random sampling using auxiliary information. *Stat. Pap.* **54**(2), 479–497 (2013)
26. Singh, H.P., Vishwakarma, K.: Some families of estimators of variance of stratified random sample mean using auxiliary information. *J. State Theory Prac.* **2**(1), 21–43 (2008)
27. Solanki, R.S., Singh, H.P.: An improved class of estimators for the population variance. *Model Assist. Stat. Appl.* **8**(3), 229–238 (2013)
28. Subramani, J., Kumarapandiyan, G.: Variance estimation using quartiles and their functions of an auxiliary variable. *Int. J. Stat. Appl.* **2**(5), 67–72 (2012)
29. Subramani, J., Kumarapandiyan, G.: Estimation of variance using known coefficient of variation and median of an auxiliary variable. *J. Mod. Appl. Stat. Methods* **12**(1), 58–64 (2013)
30. Upadhyaya, L.N., Singh, H.P.: Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical* **41**, 627–636 (1999)
31. Yadav, S.K., Kadilar, C.: Improved exponential type ratio estimator of population variance. *Rev. Colomb. Estad.* **36**(1), 145–152 (2013)