Asian Journal of Academic Research (AJAR)
ISSN-e: 2790-9379
Vol. 4, No. 1, (2023, Spring), 97-111.

The Estimation of the Variance of a Sensitive Variable in the Presence of Auxiliary Variable Using Randomized Response Technique

Amjad Mahmood,[1] Muhammad Hanif,[2] & Attique Hussain[3]

Abstract:

This paper presents a universal class of estimators for broad-spectrum population factors of the study variables via auxiliary information in example evaluation. Estimators of diverse parameters of the study variables based on auxiliary information can be examined as an associate of the class of estimators. It also investigates the issue of variance estimate in the occurrence of sensitive variable, which will be measured, and suggest a comprehensive variance estimator for a susceptible variable by the application of a randomized tool. A conservative estimator $t_{(a,b)}$ is used to characterize the population parameter, coefficient of variation (C.V), the population mean, and variance. A second-order ballpark figure for the minimum mean square estimator (MSE) has been analyzed. The competence of the suggested estimator with other estimators has been evaluated by using numerical study and simulation inquiry. The outcomes showed that the anticipated universal variance estimator is more competent than the previously available estimators of variance and ratio variance.

Keywords: Auxiliary variable, variances, conventional estimators, mean square error (MSE), percent relative efficiency (PRE), numerical study

INTRODUCTION

During the conduction of the survey, on many occasions, it is very hard to observe the variable of interest directly. These situations commonly happened when phenomena under study are prohibited in society, and people don't want to discuss them at all. To cover these topics in surveys, various participants could be able to provide true information about them (Gupta et al., 2020). Due

[1] PhD Scholar, Department of Statistics, National College of Business Administration and Economics, Lahore/ Assistant Professor, Punjab College of Informational Technology, Lahore, Punjab, Pakistan. Email: amjadmahmood502@gmail.com

[2] Professor, Department of Statistics, National College of Business Administration and Economics, Lahore. Email: drmianhanif@gmail.com

[3] MPhil Scholar, Department of Statistics, National College of Business Administration and Economics, Lahore. Email: Email: attiquehussain777@gmail.com

to the issue of the non-response of participants, measurement error occurred which can be avoided during the sensitive survey, and researchers opined that the obtained data is free from any error, another significant fact in surveys is the issue of non-responsiveness of participants, that happened when selected sample unable to answer the question in case of his non-availability at home. In case of a sensitive variable of interest, the participants vacillate to give private information which leads to non-responsiveness. Numerous previous literature supported that like Singh et al., (2011); Khare et al., (2013); Shabbir et al. (2018); Hansen and Huwitz (1946); Cochran (1977); Rao (1986); Khare and Srivastava (2010); Warner resolved this issue by the introduction of Randomized Response Technique (RRT) in 1965.

Variance is a vastly investigated topic by numerous researchers, even though fewer the problem of variance estimation. The variance analysis helped in planning managing and organizing expenses in a project by planning the monitoring against tangible costs. Likewise, in climate study, diverse traits of seeds and land fertility are expensive in farming research. The discrepancy of estimation is the standard of the square of the divergence of the estimator from its mean.

Variance Estimator (VE) is an important factor to predict quality. It aids to create applicable deductions about people. It is even employed to calculate the confidence intervals. Even a precise conclusion was drawn with its help. VE must have fine properties for sampling design. It was Liu, the first man who measured a universal class of quadratic impartial estimators for the variance of the population (Liu, 1974).

For the very first time, Olkin (1958) integrated multi-auxiliary inconsistent to presume the boundary of predestined general public and comprehend the ratio estimate. Srivastava and Jhajj (1980) employed AUI to recommend a capable group of variance evaluation. Even AUI pertained to foreseeing an objective (Rueda *et al.* (1996); Isaki (1983) described the ratio estimate; a neutral variance estimator for the result of finite population disparity and the smaller mean square error (*MSE).*

The expression of CV was introduced by Edgeworth in 1892 (Pearson, 1920). According to Martin and Gray (1971), CV is a consistent, non-dimensional determination of dispersion relative to an average set of data. It helps to assess multifaceted sets of data (Stevens, 1946), with diverse items of measurements (Soong, 2004), Karl Pearson was expected to be the first researcher, who proposed this measure (Pearson, 1896). According to Martin and Gray (1971), CV is frequently articulated in the form of percentage defined as the ratio of the standard deviation σ to the mean μ (or its absolute value, $|\mu|$). So, on similar lines, CV would be measured for statistics on a ratio scale (Michell, 1986), CV only takes nonnegative data. It may not take into view that data is based on an interval scale (Velleman & Wilkinson, 1993), or that set of data with discrete variables with zero points and equal intervals like the Likert scale. Usually, in these circumstances, facts did not contain any property that originally compels the use of CV; there is great variation in big phenomena but less in the case of small things (Snedecor & Cochran, 1989; Cattopadhyay & Kelley, 2016). In the case of proportions or ratios, which often showed heterogeneous variance, CV is difficult to interpret directly.

LITERATURE REVIEW

Shahzad *et al.* (2019); Saleem et al. (2019) proposed a group of ratio-type variance estimators by applying the quartiles of the AV to guess the inhabitant's inconsistency in the presence of non-response on the same lines by Taylor. Muneer *et al.* (2018) practiced the auxiliary information (AUI) to recommend the finer exponential ratio-product type estimate for anonymous population discrepancy for the study variable. Adichwal *et al.* (2016) proposed an estimator for the equation of bias and the mean square error (*MSE)* that was achieved by applying the Taylor series and the act of the projected estimator. A group of better comprehensive variance estimators using AUI under *SRSWOR* units was also suggested.

Mushtaq and Ul-Amin (2019) recommended a universal estimator of variance for a vulnerable variable by applying the randomized tool. One of the researchers anticipated the general estimate for a variance for the prediction of the bunch population mean under the adaptive cluster sampling technique (Qureshi et al., 2019). Das and Tripathi (1978) developed a chain of estimators to predict the variance of a finite population of study variables with the application of (AUI). Another piece of literature described the multivariate ratio and regression estimate to draw inferences about the variance of the finite population (Isaki, 1983). In another research, Singh and Singh (1991) projected a division of estimations for the variance of the population with two-stage sampling, which was in part cleared for the sole auxiliary variable. An impartial estimate for the variance of the population by employing AUI was proposed by Swain and Mishra (1992). Swain and Mishra (1994 a,b) found the limit sharing of the ratio estimate for the discrepancy in finite population and also designed the estimator for the variance of the population with the utilization of imbalanced chance sampling.

Another empirical research designed 3 groups of linked ration-kind estimations for a variance of people with well-defined auxiliary variables (AV) (Gupta et al., 1992); on the other hand, in 2007 Grover anticipated a wide and more competent group of estimations as compared to the one developed by Jhajj et al. (2005) as the population means and discrepancy with other known AV and proceed to estimate variance for the population. Kadilar and Cingi (2006) developed the estimator for populace variance in simple random sampling with AV.

The coefficient of variation signifies the ratio of the standard deviation to the mean, and it is a beneficial measurement for making the comparison of the degree of variation from one set of data to another data series, even if the means are considerably unlike each other. The CV is a relative degree of variability exhibited in the standard deviation size parallel to the mean. It is one of the standards that permit the researcher to do a comparison between the inconsistency of unequal group and their features. CV is also called relative standard deviation (RSD).

For the prediction of people in the areas, Kadilar and Cingi (2006) applied the CV as the AV to predict inconsistency in the restricted populace. On the same lines adopted by Chand (1975), one of the older studies anticipated the ratio series kind to predict the mean of the population with the use of two AV in sole and two-phase sampling by taking into account the linear association among the study and the AVs (Samiud din & Hanif, 2007). The variance estimation in simple random sampling (SRS) was used to prove this.

Many previous studies found an exponential estimate with the use of variance and mean in the populace of AV to lower assumed variance in finite populace under sole and two-typed samplings (Asghar et al., 2014; Sanaullah et al., 2019). Another exponential ratio prediction kind for a variance of people was given by Yadav and Kadilar (2013-a, b) and showed to be better than the estimator proposed by Isaki (1983). Kadilar and Cingi (2007) suggested the variance estimator in a simple random sampling technique in unsure conditions and established that the opted estimate possessed the least mean square error than the estimate of ratio and regression in confirmed given stated as developed by Isaki (1983).

To propose the estimator for the variance of a finite population, Shahzad, Perri, and Hanif (2019) found the suggestion of kernel matrix model. A better variety of variance estimators were also achieved with the help of, unlike choices of the general constants. While Muili *et al.* (2019) on the other hand, proposed the grouping of linear ratio-cum-product estimator to guess the inhabitant's variance for the study variable.

The variance estimate was introduced by Ahmed et al. (2013) with the application of using multi-auxiliary variables in multistage sampling. In the case of stratification in second-phase sampling, Neyman (1938) and Rao (1973) utilized multiphase sampling. Sen–Yates–Grundy (SYG) discrepancy was given by Hidiroglou et al. (2009) for estimation in multiphase sampling. They assumed the first-phase sample size as permanent and chose the second-phase sample size that was supported on the predetermined first-stage sample, and applied the stratification on both stages of sampling.

Numerous previous kinds of literature supported the variance judgment and recommended better estimations to predict discrepancies in the population (Singh, Singh & Solanki, 2014); Singh & Solanki, (2013). It was Asghar et al. in 2017 anticipated the regression-cum-exponential estimate to predict the discrepancy of population conducted with multistage sampling with the utilization of multi-auxiliary variables. It was given by Sanaullah et al. (2019) those two AVs in multistage sampling help to simplify the exponential estimator for inconsistency, and the state of affairs under which the estimate was assumed to work superior to available estimators was discussed.

Various empirical data were conceded to validate the show of the projected estimations. Yadav and Kadilar (2013-b) introduced an efficient exponential ratio-type estimate to predict the discrepancy in the populace and did a difference with commonly practiced discrepant estimators. A class of variance estimates to test the qualitative AV means to predict the people variance was introduced by Singh and Malik (2014) as variance ratio-type estimators with the use of twin conversion as previously completed by Sharma and Singh (2014).

To estimate the finite populace variance for the variable of attention, Bhat *et al.* (2018) initiated the linear blend of arithmetic mean, deciles, and kurtosis of the AVs to obtain the competent group of variance estimators. For the judgment of the variance of finite inhabitants in the presence of non-response, Shahzad *et al.* (2018) suggested a set of discrepancy estimators by the application of auxiliary attributes (AA). The mathematical outcomes clarified the excellent demonstration of the projected set of inconsistency estimators over the disparate estimator. To examine the discrepancy judgment of infrequent, obscured, merely jumble, and hard-to-reach units of people under used cluster sampling design, Qureshi *et al.* (2018, 2019) utilized the AI to predict general ratio exponential and general estimate of ratio. The estimated adopted estimators were harmonized with the non-adaptive estimators under (SRS) design. The outcomes recommended that adopted estimates will be obtained in a better situation than the non-adopted estimates for infrequent and discrepancy of the grouped population.

In the field of economics, environmental science, and natural and farming science, the coefficient of variation (CV) has been in use. The CV is vital in a situation where the center of attention is guessing the relative measure of dispersion as it is unit free. A little work has been done to investigate CV in the recent past. It was Das and Tripathi (1992-1993) found the estimator for CV using the SRSWOR method. Tripathi et al. (2002) planned a set of estimators that comprises ratio and regression estimators for the CV of the population by applying the sample information of the CV. Patel and Shah (2009) investigated the miniature sample MSE for numerous estimators under SRS with the application of a few estimates linked with hybrid class. Rajyaguru and Gupta (2006) projected three modules of estimates for the CV of the population under stratified sampling. Ahmed (2002) anticipated dwindling opening analysis for the estimate of a CV of the normal allocation when the extra sample is available to elevate the meticulousness of the estimator.

Many other empirical researchers like Singh et al., (2016); Chaudhary and Singh, (2014); Singh and Sharma, (2015); Tabasum and Khan, (2006); Sodipo and Obisesan, (2007), Singh and Kumar, (2008); studied the matter to predict mean in the presence of non-response or ME Mahmoundvand and Hassani (2007) anticipated a neutral estimator for the CV of population under the normal distribution. A neutral estimator for CVs was introduced by Breunig (2001) where a bias-corrected estimator is offered for the CVs. Patel and Shah (2009) on the pattern of Monte Carlo assessment proposed numerous estimators of finite population CV under simple random sampling without replacement (SRSWOR).

A lot of work has been done on the study variables but on the particular aspect of sensitive variables with the application of variance, no work has been conducted. To fill in the research gap this particular investigation has been practiced with the primary aim to explore the phenomena to estimate the general parameters for a sensitive inconstant by the application of variances for a finite population.

TERMINOLOGY

A finite population with N Units U= (U_1, U_2... U_N) out of which a sample size n was drawn with the use of simple random sampling without replacement was considered. Let Y be the study variable, a sensitive variable that cannot be observed directly due to respondent bias. Let X is the non-sensitive auxiliary variable that is correlated with Y. Let S will be the scrambling variable independent of Y and X.

The respondent is asked to report a scrambled response for Y given by Z=Y+S but is asked to provide a true response for X. Let ($\bar{y}$,$\bar{x}$) be the sample means corresponding to ($\bar{Y}$,$\bar{X}$) the population means of Y and X, respectively.

$$\mu_{rs} = \frac{1}{N}\sum^{N}\left(Z - \bar{Z}\right)^{r}\left(X - \bar{X}\right)^{s}$$

$\delta_{rs} = \dfrac{\mu_{rs}}{{\mu_Z}^{r/2}{\mu_Z}^{s/2}}$ And (r,s) is non-negative integers

$$\mu_{20} = {S_Z}^2, \quad \mu_{02} = {S_X}^2, \quad \mu_{11} = S_{XZ}, \quad {C_Z}^2 = \frac{{S_Z}^2}{\bar{Z}^2} = \frac{\mu_{20}}{\bar{Z}^2}$$

$${C_X}^2 = \frac{{S_X}^2}{\bar{X}^2} = \frac{\mu_{11}}{\bar{X}^2} \quad \text{And} \quad \rho_{XZ} = \frac{S_{XZ}}{S_X S_Z} = \frac{\mu_{11}}{\sqrt{\mu_{02}}\sqrt{\mu_{20}}}$$

Where,

$$\in_0 = \frac{\bar{z}-\bar{Z}}{\bar{Z}}, \quad \in_1 = \frac{s_z^2-\sigma_z^2}{\sigma_z^2}, \quad \in_2 = \frac{\bar{x}-\bar{X}}{\bar{X}}, \quad \in_3 = \frac{s_x^2-\sigma_x^2}{\sigma_x^2},$$

$$\mu_{20} = S_z^2, \mu_{02} = S_x^2, \mu_{11} = S_{xz}, C_z^2 = \frac{S_z^2}{\bar{Z}} = \frac{\mu_{20}}{\bar{Z}}, C_x^2 = \frac{S_x^2}{\bar{X}} = \frac{\mu_{11}}{\bar{X}},$$

$$\rho_{XZ} = \frac{S_{xz}}{S_x S_z} = \frac{\mu_{11}}{\sqrt{\mu_{02}}\sqrt{\mu_{20}}},$$

$$E(\in_0) = E(\in_1) = E(\in_2) = E(\in_3) = 0,$$

$$E(\in_\circ) = E(\in_1) = E(\in_2) = E(\in_3) = 0$$

$$E({\in_\circ}^2) = n^{-1}{C_Z}^2 \quad E({\in_1}^2) = n^{-1}(\delta_{40}-1) \quad E({\in_2}^2) = n^{-1}{C_X}^2 \quad E({\in_3}^2) = n^{-1}(\delta_{04}-1)$$

$$E(\in_\circ \in_1) = n^{-1}\delta_{30}C_Z \quad E(\in_\circ \in_2) = n^{-1}\rho_{ZX}C_X C_Z \quad E(\in_\circ \in_3) = n^{-1}\delta_{12}C_Z$$

$$E(\in_1 \in_2) = n^{-1}\delta_{21}C_X \quad E(\in_1 \in_3) = n^{-1}(\delta_{22}-1) \quad E(\in_2 \in_3) = n^{-1}\delta_{03}C_X$$

Notations used in Variance estimation

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2, s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2,$$

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i-\bar{Y})^2, \quad \sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})^2,$$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N}X_i, \ \bar{Y} = \frac{1}{N}\sum_{i=1}^{N}Y_i, \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n}y_i,$$

$$\delta_{rs} = \frac{\mu_{rs}}{\mu_{20}^{\frac{r}{2}}\mu_{02}^{\frac{s}{2}}}, \quad \mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(Z_i-\bar{Z})^r(X_i-\bar{X})^s,$$

$$s_z^2 = \sigma_z^2(1+\in_1), \quad \bar{z} = \bar{Z}(1+\in_0),$$

$$s_x^2 = \sigma_x^2(1+\in_2),$$

Conventional Estimator

The general parameter $t_{(a,b)}$ conventional estimator is given by

$$t'_{(a,b)} = \bar{z}^a\left(s_z^2-\sigma_s^2\right)^{\frac{b}{2}},$$

$$t'_{(a,b)} = \bar{Z}^a(1+\in_0)^a\left[\sigma_z^2-\sigma_s^2+\sigma_z^2\in_1\right]^{\frac{b}{2}},$$

$$t'_{(a,b)} = \bar{Z}^a(1+\in_0)^a\left[\sigma_y^2+\sigma_z^2\in_1\right]^{\frac{b}{2}},$$

$$t'_{(a,b)} = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right), \text{.......I}$$

$$MSE\left(t'_{(a,b)}\right) = \frac{t_{(a,b)}^2}{n}\left(a^2C_z^2 + ab\delta_{30}R_{zy} + \frac{b^2}{4}(\delta_{40}-1)R_{zy}^2\right)$$

$$MSE\left(t'_{(a,b)}\right) = \frac{t_{(a,b)}^2}{n} f_1(a,b)$$

where

$$f_1(a,b) = \left(a^2C_z^2 + ab\delta_{30}R_{zy} + \frac{b^2}{4}(\delta_{40}-1)R_{zy}^2\right)$$

PROPOSED ESTIMATOR

We propose the following class of ratio, product type; exponential ratio type, exponential product type estimators are given in this section.

Ratio Estimator

$$t_r^* = t'_{(a,b)}\left(\frac{S_x^2}{s_x^2}\right)$$

Using equation No. I

$$t_r^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\left(\frac{S_x^2}{S_x^2(1+\in_3)}\right),$$

$$t_r^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)(1+\in_3)^{-1}$$

Squaring both sides and neglecting higher order power, i.e. 3 or more

$$\left(t_r^* - t_{(a,b)}\right)^2 = t_{(a,b)}^2\left(a^2\in_0^2 + \frac{b^2}{4}\in_1^2 R_{zy}^2 + \in_3^2 + 2\frac{ab}{2}\in_0\in_1 R_{zy} - 2a\in_0\in_3 - 2\frac{b}{2}\in_1\in_3 R_{zy}\right)$$

Taking expectations

$$MSE\left(t_r^*\right) = t_{(a,b)}^2\left(a^2C_z^2 + \frac{b^2}{4}(\delta_{40}-1)R_{zy}^2 + ab\delta_{30}C_zR_{zy} + (\delta_{40}-1) - 2a\delta_{12}C_z - b(\delta_{22}-1)R_{zy}\right)$$

$$MSE\left(t_r^*\right) = t_{(a,b)}^2\left(f_1(a,b) - \left(2a\delta_{12}C_z + b(\delta_{22}-1)R_{zy}\right) + (\delta_{40}-1)\right)$$

$$MSE\left(t_r^*\right) = t_{(a,b)}^2\left(f_1(a,b) - f_2(a,b) + (\delta_{40}-1)\right)$$

where

$$f_2(a,b) = 2a\delta_{12}C_z + b(\delta_{22}-1)R_{zy}$$

Product Estimator

$$t_p^* = t'_{(a,b)}\left(\frac{s_x^2}{S_x^2}\right)$$

Using equation No. I

$$t_p^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\left(\frac{S_x^2(1+\in_3)}{S_x^2}\right),$$

$$t_p^* - t_{(a,b)} = t_{(a,b)}\left(a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy} + \in_3 + a\in_0\in_3 + \frac{b}{2}\in_1\in_3 R_{zy}\right)$$

Squaring expectations and neglecting higher order power, i.e. 3 or more

$$MSE(t_p^*) = t_{(a,b)}^2\left(a^2C_z^2 + \frac{b^2}{4}(\delta_{40}-1)R_{zy}^2 + ab\delta_{30}C_zR_{zy} + (\delta_{40}-1) + 2a\delta_{12}C_z + b(\delta_{22}-1)R_{zy}\right)$$

$$MSE(t_p^*) = t_{(a,b)}^2\left(f_1(a,b) + \left(2a\delta_{12}C_z + b(\delta_{22}-1)R_{zy}\right) + (\delta_{40}-1)\right)$$

$$MSE(t_p^*) = t_{(a,b)}^2\left(f_1(a,b) + f_2(a,b) + (\delta_{40}-1)\right)$$

Exponential Ratio Estimator

$$t_{er}^* = t'_{(a,b)}\exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)$$

Using equation No. I

$$t_{er}^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\exp\left(\frac{S_x^2 - S_x^2 - S_x^2\in_3}{S_x^2 + S_x^2 + S_x^2\in_3}\right),$$

$$t_{er}^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\exp\left(-\frac{\in_3}{2} + \frac{\in_3^2}{4} - \frac{\in_3^3}{8}\right)$$

$$t_{er}^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\left(1 - \frac{\in_3}{2} + \frac{\in_3^2}{4} - \frac{\in_3^3}{4} - \frac{\in_3^2}{8}\right)$$

$$t_{er}^* - t_{(a,b)} = t_{(a,b)}\left(a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy} - \frac{\in_3}{2} + \frac{a}{2}\in_0\in_3 - \frac{b\in_1\frac{\in_3}{2}}{4}R_{zy} + \frac{3\in_3^2}{8} - \frac{\in_3^3}{8}\right)$$

Squaring expectations and ignoring higher order power, i.e. 3 or more

$$MSE(t_{er}^*) = t_{(a,b)}^2\left(a^2C_z^2 + \frac{b^2}{4}(\delta_{40}-1)R_{zy}^2 + ab\delta_{30}C_zR_{zy} + \frac{(\delta_{40}-1)}{4} - a\delta_{12}C_z - \frac{b}{2}(\delta_{22}-1)R_{zy}\right)$$

$$MSE(t_{er}^*) = t_{(a,b)}^2\left(f_1(a,\mathrm{b}) - \frac{1}{2}f_2(a,b) + \frac{(\delta_{40}-1)}{4}\right).$$

Exponential Product Estimator

$$t_{ep}^* = t'_{(a,b)}\exp\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right)$$

Using equation No. I

$$t_{ep}^* = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\exp\left(\frac{S_x^2 + S_x^2\in_3 - S_x^2}{S_x^2 + S_x^2 + S_x^2\in_3}\right),$$

$$t^*_{ep} = t_{(a,b)}\left(1 + a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy}\right)\left(1 + \frac{\in_3}{2} - \frac{\in_3^2}{4} + \frac{\in_3^3}{4} + \frac{\in_3^2}{8}\right)$$

$$t^*_{ep} - t_{(a,b)} = t_{(a,b)}\left(a\in_0 + \frac{a(a-1)}{2}\in_0^2 + \frac{b}{2}\in_1 R_{zy} + \frac{b(b-2)}{8}R_{zy}^2\in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy} + \frac{\in_3}{2} + \frac{a}{2}\in_0\in_{33} + \frac{b\in_1\in_3}{4}R_{zy} - \frac{\in_3^3}{8}\right)$$

Squaring expectations by ignoring higher order power, i.e. 3 or more

$$MSE\left(t^*_{ep}\right) = t^2_{(a,b)}\left(a^2C_z^2 + \frac{b^2}{4}(\delta_{40}-1)R_{zy}^2 + ab\delta_{30}C_zR_{zy} + \frac{(\delta_{40}-1)}{4} + a\delta_{12}C_z + \frac{b}{2}(\delta_{22}-1)R_{zy}\right)$$

$$MSE\left(t^*_{ep}\right) = t^2_{(a,b)}\left(f_1(a,\mathrm{b}) + \frac{1}{2}f_2(a,b) + \frac{(\delta_{40}-1)}{4}\right)$$

SIMULATION STUDY

In this section, we have conducted numerical study to evaluate the performance of derived estimators when contrasted with each other. The formula Z = Y + S provides the reported response. From two bivariate normal populations following parameters are determined.

Population size = N = 1000

Population I $\mu = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 9 & 1.9 \\ 1.9 & 4 \end{bmatrix}$ ρ_{xy} = 0.3209

Population II $\mu = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix}$ ρ_{xy} = 0.8684

These 1000 values are considered as finite population. We consider three sample sizes which are 200, 300 and 500. The scrambling variable is considered to follow normal distribution with E(S) = 0 and the results are obtained using three different values of Var (S) = 0.2, 0.5, 1.

We have compared our derived estimators on the basis of percentage relative efficiency (PRE). This is given by as

$$\text{PRE} = \frac{MSE(t)}{MSE(t_\beta)} \times 100 \qquad \text{Where } \beta = p, r, ep, er$$

R Language is used to process the data. Table 1 and Table 2 give the PRE's comparison of the derived estimators on the basis of population I and Population II respectively.

Table 1

Var (S)	**n**	**PRE's** t^* **parison**	t^*_r	t^*_p	t^*_{er}	t^*_{ep}
	200	100.000	282.91	134.542	801.1	312.77
0.2	**300**	100.000	282.915	134.5420	801.11	312.8
	500	100.000	282.9149	134.5	801.115	312.768
0.5	**200**	100.000	291.174	136.3817	799.3779	312.503

	300	100.000	291.17	136.38	799.38	312.5030
	500	100.000	291.1743	136.3817	799.38	312.50
	200	100.000	304.4	139.207	793.607	311.6171
1	**300**	100.000	304.3611	139.2067	793.61	311.62
	500	100.000	304.39	139.21	793.608	311.6

Table 2

Var(S)	**n**	**PRE's** t^*	**parison** t_r^*	t_p^*	t_{er}^*	t_{ep}^*
	200	100.000	190.4639	90.11176	533.795	208.47
0.2	**300**	100.000	190.46	90.1118	533.7951	208.469
	500	100.000	190.464	90.112	533.80	208.4691
	200	100.000	198.574	91.8873	531.274	208.083
0.5	**300**	100.000	198.574	91.89	531.2735	208.0834
	500	100.000	198.6	91.887	531.27	208.1
	200	100.000	211.1828	94.50	523.21	206.8354
1	**300**	100.000	211.2	94.4981	523.213	206.84
	500	100.000	211.183	94.498	523.2128	206.835

REAL DATA APPLICATION

In this section we have carried out numerical study and compared our derived estimators based on real data obtained from Adichwal et al. (2022) paper "Estimation of general parameters using auxiliary in information SRSWOR"

The description of the data set is as follow

$$n=10,\quad \bar{X}=58.8,\quad \bar{Y}=101.1,\quad C_x=0.12381,\quad C_Y=0.1450,$$
$$\rho_{XY}=0.6500,\quad \delta_{12}=0.5714,\quad \delta_{21}=0.4537,\quad \delta_{03}=0.4861,$$
$$\delta_{30}=0.3248,\quad \delta_{04}=2.2387,\quad \delta_{40}=2.3523,\quad \delta_{13}=1.5041,$$
$$\delta_{31}=1.6923,\quad \delta_{22}=1.5432.$$

Following Table 3 showed PRE's Comparison based on above mentioned data.

Table 3

Var(S)	n	PRE t^* mparisor	t_r^*	t_p^*	t_{er}^*	t_{ep}^*
0.2	**200**	100.000	259.83	124.81470	750.3991	292.9
	300	100.000	259.8313	124.815	750.40	292.8912
	500	100.000	259.8	124.81	749.3991	292.89
0.5	**200**	100.000	260.17	124.892	750.40	292.891
	300	100.000	260.166	124.9	750.3961	292.8908
	500	100.000	260.2	124.8918	749.40	292.891
1	**200**	100.000	260.7224	125.02	750.3853	292.89
	300	100.000	260.72	124.88	750.4	292.889
	500	100.000	260.722	125.0	749.39	292.8891

As expected the numerical findings from the simulation study and real data illustrations support each other. The result presented in table 1 to 3 showed that t_{er}^* exponential ratio) estimator is better than all other derived estimators. Thus on the result of real data and simulation study we concluded that an estimator who is having greater efficiency value is efficient to deal with nonresponsive in certain type of situations. On the basis of numerical results, the exponential ratio estimator found to be more superior than all the estimators used in this research at different sample sizes.

CONCLUSION

There are numerous issues and problems faced by researchers in everyday life when we have to deal with sensitive nature of the variable under investigation. Respondents usually refused and sometimes provided wrong information to such types of variables or give fabricated answer to the questions that caused scrambled responses. In current paper the researchers have proposed ratio, product, exponential ratio and exponential product variance estimators for scrambled responses.

It was tried to derive mean square expressions for all the proposed estimators. A broad numerical comparison was also the part of this research to have more efficient estimators for under consideration problem.

From the numerical findings, it was concluded that exponential ratio estimator performs outstanding results for both real and artificial population on different sample sizes when it was compared with other derived estimators. So with the help of numerical results presented in table 1, table 2, table 3 illustrated above, we have concluded that exponential ratio estimator is better and more efficient than rest of the derived estimators in the manuscript.

References:

Adichwal, N. K., Ahmadini, A. A. H., Raghav, Y. S., Singh, R., & Ali, I. (2022). Estimation of general parameters using auxiliary information in simple random sampling without replacement. *Journal of King Saud University-Science*, *34*, 101754.

Adichwal, N. K., Kumar, J., & Singh, R. (2018). An improved generalized class of estimators for population variance using auxiliary variables. *Cogent Mathematics & Statistics*, *5*(1), 1-8.

Ahmed, A., Ahmad S. P., & Reshi, J. A. (2013). Bayesian analysis of Rayleigh distribution. *International Journal of Scientific and Research Publications*, *3*, 1-9.

Ahmed, S. E. (2002). Simultaneous estimation of co-efficient of variation. *Journal of Statistical Planning and Inference*, *104*, *31-51*.

Asghar, A., Sanaullah, A., & Hanif, M. (2014). Generalized exponential type estimator for population variance in survey sampling. *Revista Colombiana de Estadística*, *38*(2), 385-97.

Bhat, M. A., Maqbool, S., & Subzar., M. (2018). An improvement in variance estimator for the estimation of population variance, using known values of auxiliary information. *International Journal of Pure & Applied Bioscience, 6*(5), 135-38.

Breunig, R. (2001). An almost unbiased estimator of the coefficient of variation. *Economics Letters, Elsevier, 70*(1), 15-19.

Chand, L. (1975). *Some ratio-type estimators based on two or more auxiliary variables* (Ph.D. thesis Iowa State University, Ames, Iowa).

Chattopadhyay, B., & Kelley, K. (2016). Estimation of the coefficient of variation with minimum risk: A sequential method for minimizing sampling error and study cost. *Multivariate Behavioral Research, 51*(5), 627-48. https://doi.org/10.1080/00273171.2016.1203279

Chaudhary, M. K., & Singh, V. (2014). Estimating the population mean in two-stage sampling with equal size clusters under non-response using auxiliary characteristic. *Mathematical Journal of Interdisciplinary Sciences*, 2(1), 43-55.

Cochran, W. G. (1977). *Sampling Techniques*. John Wiley & Sons.

Das, A. K., & Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya C*, *40*, 139-48

Das, A.K., & Tripathi, T. P. (1992/1993). Use of auxiliary information in estimating the coefficient of variation. *Aligarh J. Stat., 12*(13), 51-58.

Grover, L. K. (2007). A wide and efficient class of estimators of population variance under sub-sampling scheme. *Model Assisted Statistics and Applications*, *2*(1), 41-51.

Gupta, R. K., Singh, S., & Mangat, N. S. (1992). Some chain ratio-type estimators for estimating finite population variance. *Aligarh Journal of Statistics*, *12,* 65-69.

Hansen, M. H., & Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association, 41*(236), 517–29.

Hidiroglou, M. A., & Lavallée, P. (2009). Sampling and estimation in business surveys. Handbook of statistics. *Elsevier, 29*, 441-70.

Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of American Statistical Association*, *381*, 117-23.

Jhajj, H. S., Sharma, M. K., & Grover, L.K. (2005). An efficient class of chain estimators of population variance under sub-sampling scheme. *Journal of Japan Statistical Society*, *35*(2), 273-86.

Kadilar, C., & Cingi, H. (2006). Improvement in variance estimation using auxiliary information. *Hacettepe Journal of Mathematics and Statistics, 35*(1), 111-15.

Kadilar, C., & Cingi, H. (2007). Sampling theory improvement in variance estimation in simple random sampling. *Communication in Statistics-Theory and methods*, *36*(11), 2075-81.

Khare, B., & Srivastava, S. (2010). Generalized two phase sampling estimators for the population mean in the presence of non-response. *Aligarh Journal of Statistics*, *30*, 39–54.

Khare, B., Pandey, S., & Kumar, A. (2013). Estimation of Population Mean in Sample Surveys Using Auxiliary Character, Method of Call Backs and sub-sampling from Non-respondents. *Proceedings of the National Academy of Sciences, India: Physical Sciences*, 83(1), 49–54.

Liu, T. P. (1974). A general unbiased estimator for the variance of a finite population. *Sankhya C*, *36*(1), 23-32.

Mahmoudvand, R., & Hassani, H. (2007). Two new confidence intervals for the C.V in a normal distribution. *Journal of Applied Statistics*, 36(4), *429-42*.

Martin, J. D., & Gray, L. N. (1971). Measurement of relative variation: Sociological Examples. *American Sociological Review, 36*, 496.

Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin, 100*(3), 398-407. https://doi.org/10.1037/0033-2909.100.3.398

Muili, J. O., Audu, A., Singh, R. V. K., Odeyale, A. B., & Adebiyi, A. (2019). Improved variance estimator using linear combination of tri-mean and quartile average. *Annals. Computer Science Series*, *17*(1), 142-47

Muneer, S., Khalil, A., Shabbir, J., and Narjis, G. (2018). A new improved ratio-product type exponential estimator of finite population variance using auxiliary information. *Journal of Statistical Computation and Simulation, 88*(16), 3179-92.

Mushtaq, N., & Noor-Ul-Amin, M. (2019). Generalized variance estimators using randomized device in the presence of auxiliary information. *Journal of Statistics and Management Systems*, *22*(8), 1417-24.

Neyman, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, 97(4), 558-625.

Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, *33*, 101-16.

Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika*, *45*, 154-65.

Patel, P. A., & Shah, R. (2009). A Monte Carlo comparison of some suggested estimators of co-efficient of variation infinite population. *Journal of Statistics sciences*, *1*(2), 137–47.

Pearson, F. B. (1920). Wisdom. *Journal of Education*, *91*(23), 635. https://doi.org/10.1177/002205742009102313

Pearson, K. (1896). Mathematical contributions to the theory of evolution-III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society A*, *187*, 253-318.

Qureshi, M. N., Khalil, S., & Hanif, M. (2018). Generalized semi exponential type estimator under systematic sampling. *Journal of Statistical Theory and Applications*, *17*(2), 283-90.

Qureshi, M. N., Riaz, N., & Hanif, M. (2019). Generalized estimator for the estimation of rare and clustered population variance in adaptive cluster sampling. *Journal of Statistical Computation and Simulation, 89*(11), 2084-101.

Rajyaguru, A., & Gupta, P. (2006). On the estimation of the coefficient of variation from finite population-II. *Model Assisted Statistics and application*, *1*(1**),** *57-66.*

Rao P. (1986). Ratio estimation with subsampling the non-respondents. *Survey Methodology, 12*(2), 217-30.

Rao, J. N. K. (1973). On double sampling for stratification and analytic surveys. *Biometrika*, *6*, 125-33.

Rueda, M., & Cebrian, A. A. (1996). Repeated substitution method the ratio estimator for the population variance. *Metrika, 43*, 101-05.

Saleem, I., Sanaullah, A., & Hanif, M. (2019). Double-sampling regression-cum-exponential estimator of the mean of a sensitive variable. *Mathematical Population Studies*, *26*(3), 163-82.

Samiuddin M., & Hanif, M. (2007). Estimation of population mean in single- and two-phase sampling with or without additional information. *Pak. J. Statist., 23*(2), 99-118.

Sanaullah, A., Saleem, I., & Shabbir, J. (2019). Use of scrambled response for estimating mean of the sensitivity variable. *Communications in Statistics-Theory and Methods*, 1-14.

Shabbir, J., Gupta, S., & Ahmed, S. (2018). A generalized class of estimators under two-phase stratified sampling for non-response. *Communications in Statistics-Theory and Methods*, 1–17.

Shahzad, U., Hanif, M., Koyuncu, N., & Sanaullah, A. (2018). On the estimation of population variance using auxiliary attribute in absence and presence of non-response. *Electronic Journal of Applied Statistical Analysis*, *11*(2), 608-21.

Shahzad, U., Perri, P. F., & Hanif, M. (2019). A new class of ratio-type estimators for improving mean estimation of nonsensitive and sensitive variables by using supplementary information. *Communications in Statistics - Simulation and Computation*, *48*(9), 2566-85

Sharma, P., & Singh, R. (2014). Improved dual to variance ratio type estimators for population variance. *Chil. J. Stat*, *5*(2), 45-54.

Singh, B. K., Chanu, W. W., & Tato, Y. (2016). A class of ratio cum dual to ratio estimator in presence of non-response using two phase sampling scheme. *GU J Sc.i, 29*(1), 213-21.

Singh, H. P., & Kumar, S. (2008b). Estimation of mean in presence of non-response using two phase sampling scheme. *Statistical Papers.* DOI 10.1007/s00362-008-040-5.

Singh, H. P., & Kumar, S. (2011). Combination of regression and ratio estimate in presence of nonresponse. *Brazilian Journal of Probability and Statistics*, *25*(2), 205-17.

Singh, H. P., Singh, A. K., & Solanki, R. S. (2014). Estimation of finite population variance using auxiliary information in sample surveys. *Statistica*, *74*, 99–116.

Singh, R. S., & Sharma, P. (2015). Method of estimation in the presence of nonresponse and measurement errors simultaneously. *Journal of Modern Applied Statistical Methods*, *14*(1), 12.

Singh, V. K., & Singh, G. N. (1991). Chain-type regression estimators with two auxiliary variables under double sampling scheme. *Metron International Journal of Statistics*, *49*, 279-89.

Snedecor, G. W., & Cochran, W. G. (1989) *Statistical Methods*. 8th Ed. Iowa State University Press.

Sodipo, A. A., & Obisesan, K. O. (2007). Estimation of the population mean using difference cum ratio estimator with full response on the auxiliary character. *Res. J. App. Sci*., *2*, 769-72.

Solanki, R. S., & Singh, H. P. (2013): An improved class of estimators for the population variance. *Model Assisted Statistics and Applications*, *8*(3), 229-38.

Srivastava, S. K., & Jhajj, H. S. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhya* C, *4*, 87-96.

Srivastava, S. K., & Jhajj, H. S. J. B (1981). A class of estimators of the population mean in survey sampling using auxiliary information. *Biometrika, 68*(1), 341-43.

Swain, A. K. P. C., & Mishra, G. (1992). Unbiased estimators of Infinite population variance using auxiliary informations. *Metron*, 201-15.

Swain, A. K. P. C., & Mishra, G. (1994 a). Limiting distribution of the ratio estimator of finite population variance, *Sankhya, B, 56*, 11-17.

Swain, A. K. P. C., & Mishra, G. (1994 b). Estimation of finite population variance under unequal probability sampling. *Sankhya, B*, 56, 374-88.

Tabasum, R., & Khan, I. A. (2006). Double sampling ratio estimator for the population mean in presence of non-response. *Assam Statistical Review*, 20, 73-83.

Tripathi, T.P., Singh, H.P., Upadhyaya, L.N. (2002). A general method of estimation and its application to the estimation of the coefficient of variation. *Stat. Transition, 5*(6), 887-09.

Velleman, P. F., & Wilkinson, L. (1993) Nominal, ordinal, interval and ration typology are misleading. *The American Statistician*, *47,* 65-72.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, *60*(309), 63-69.

Yadav, S. K., & Kadilar, C. (2013). A Class of ratio-cum-dual to ratio estimator of population variance. *Journal of Reliability and Statistical Studies, 6*(1), 29–34.

Yadav, S. K., & Kadilar, C. (2013). Improved exponential type ratio estimator of population variance. *RevistaColombiana de Estadistica*, *36*(1), 145-152.

Date of Publication	May 05, 2023