Research Article

Generalized Difference-Cum Exponential Class of Estimators for Estimating Population Parameters of the Sensitive Variable

Gajendra K. Vishwakarma 1, Subhash Kumar Yadav 2, and Tarushree Bari 2

1Department of Mathematics & Computing, Department Indian Institute of Technology Dhanbad, Dhanbad-826004, India
2Department of Statistics, School of Physical & Decision Sciences, Babasaheb Bhimrao Ambedkar University, Lucknow-226025, India

Correspondence should be addressed to Subhash Kumar Yadav; drskystats@gmail.com and Tarushree Bari; taru9494@gmail.com

Received 21 January 2022; Accepted 28 February 2022; Published 2 May 2022

Academic Editor: Gajendra Vishwakarma

Copyright © 2022 Gajendra K. Vishwakarma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a generalized class of estimators to estimate general population parameters of susceptible research variables using additional auxiliary information in SRSWOR. The population constants such as the coefficient of variation, the population mean, the standard deviation, and the population mean square are defined using a conventional estimator. The expression for the mean square errors of the proposed class is derived up to the first order of approximation. To compare the effectiveness of the proposed class of estimators, an empirical study is conducted utilizing real and simulated data sets. Theoretical and empirical research demonstrates that the suggested generalized class of estimators outperforms other current estimators.

1. Introduction

The collection of data on sensitive topics such as induced abortions, drug misuse, and family income through personal interviews and surveys is a severe problem. Some questions, for example, are delicate: (a) On your 2009 tax return, how much did you underreport your income? (c) Have you undergone any abortions? (b) Have you molested any children? (e) Do you use illicit substances? One approach to get individuals to answer honestly is to use randomised response techniques. Horvitz et al. [1] and Greenberg et al. [2] have expanded Warner’s [3] model to include numeric responses to the sensitive question rather than simple “yes” or “no” responses. The respondent chooses one of two questions using a randomization device: one is the sensitive question, and the other is unrelated. Direct true responses on the study variable might become difficult to obtain in survey sampling, especially when the variable is sensitive. When the study variable is sensitive in nature, many survey statisticians have estimated the population mean of sensitive variables, such as Eichhron and Hayre [4], Gupta et al. [5], Saha [6], and Diana and Perri [7]. Different ratio, regression, and exponential estimators for estimating population parameters of sensitive variables based on scrambled responses were reported by Sousa et al. [8], Koyuncu et al. [9], Kalucha et al. [10], and Gupta et al. [11, 12]. Using a generalized quantitative optional randomised response model, Noor-Ul-Amin et al. [13] developed estimators based on generalized ratio and regression types, in which the nature of the auxiliary variable is nonsensitive. In the presence of measurement error, Khalil et al. [14] and Zahid and Shabbir [15] developed improved mean estimators of a sensitive research variable. Sanaullah et al. [16] introduced a new family of difference-cum-exponential-type estimators of the finite population mean of the susceptible research variable by using a single nonsensitive secondary information. In the case of nonresponse using RRT under two-phase sampling, Sanaullah et al. [17] proposed a generalized family of estimators for the finite population mean of a susceptible research variable. Quantitative scrambled randomised response models are considerably improved by Saleem et al. [18] by assessing the performance of the population mean estimator. Waseem et al. [19] presented a generalized exponential type estimator for the mean of a sensitive variable using information from a nonsensitive auxiliary variable. Singh et al. [20] provided some alternative additive...
randomised response models by offering a blank card option for predicting the population mean of a quantitative sensitive variable. Vishwakarma and Singh [21] developed the estimation procedure to estimate the effect of measurement errors under additive scramble response for sensitive variables.

There have been a few extreme values in numerous populations, and estimating the unknown population parameters without considering this knowledge is extremely delicate. Otherwise, the results will be understated or exaggerated, respectively. In order to resolve this problem, it is necessary to incorporate this knowledge into the estimation of population parameters. Several researchers, including Isaki [22], Bahl and Tuteja [23], Upadhyaya and Singh [24], Kadilar and Cingi [25], Dubey and Sharma [26], Singh and Vishwakarm [27], Shabbir and Gupta [28], and Yadav et al. [29], have proposed some broader classes of estimators for estimating finite population variance. Gupta et al. [30] and Kadilar and Cingi [25], Dubey and Sharma [26], Singh and Vishwakarm [27], Shabbir and Gupta [28], and Yadav et al. [29], have proposed some broader classes of estimators for estimating finite population variance. Gupta et al. [30] and Daraz and Khan [31] presented several variance estimators for sensitive variables that make use of additional information in SRSWOR.

Motivated by Adichwai et al. [32], we suggest the general difference-cum exponential class of estimators for estimating general population parameters of the sensitive research variable.

Let $Y$ represent the research variable, which contains sensitive elements that may not be accessible precisely as a result of the respondent’s response. Let $X$ be a nonsensitive secondary variable with a positive correlation to $Y$. Let $S$ be an uncorrelated scrambled variate with a predetermined distribution. The variable $S$ is used to scramble the sensitive variable $Y$. Each respondent is given the task of selecting a random number from the $S$ distribution, say $s$, and adding it to the actual value of $Y$ to assess $Z$. Let $Z$ stand for the reported scrambled answer to $Y$, which was first proposed by Warner [3] and then developed by Pollock and Bek [33].

The general population parameter used in the study are discussed as follows:

$$
\mu_{rs} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{Y}) (x_i - \bar{X})^t. \tag{1}
$$

$$
\delta_{rs} = \mu_{rs}/(\mu_{r0}^2 \mu_{s0}^2)^{1/2}
$$

where $r$ and $s$ are nonnegative integers.

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

The error terms used in the study are defined as follows:

$$
e_0 = \frac{\pi}{\bar{Z}} - 1,
e_1 = \frac{\bar{Z}^2}{\mu_{20}} - 1,
e_2 = \frac{x}{\bar{X}} - 1,
e_3 = \frac{x^2}{\mu_{02}} - 1. \tag{3}
$$

Under the condition that

$$
E(e_i) = 0;
i = 0, 1, 2, 3,.
E(e_i^2) = n^{-1} C_z^2,
E(e_i^2) = n^{-1} (\delta_{40} - 1),
E(e_i^2) = n^{-1} C_x^2,
E(e_i^2) = n^{-1} (\delta_{04} - 1), \tag{4}
$$

$$
E(e_i e_j) = n^{-1} \delta_{03} C_z,
E(e_i e_j) = n^{-1} \rho_{xz} C_z C_x,
E(e_i e_j) = n^{-1} \delta_{12} C_z,
E(e_i e_j) = n^{-1} \delta_{21} C_x,
E(e_i e_j) = n^{-1} (\delta_{22} - 1),
E(e_i e_j) = n^{-1} \delta_{03} C_x.
$$

2. Conventional Estimators

The population parameter of a sensitive research variable under discussion can be stated in its most general form as

$$
z_{(a,b)} = \bar{Z} a S_b^3, \tag{5}
$$

where $a$ and $b$ are suitably chosen scalars that take different values in order to get different parameters of a population. The parameters obtained after substituting a different combination of the scalars are given by

(i) For $(a = 1, b = 0)$, $z_{(1,0)}$ reduces to $\bar{Z}$.
(ii) For $(a = 0, b = 1)$, $z_{(0,1)}$ reduces to $S_z$.
(iii) For $(a = -1, b = 1)$, $z_{(-1,1)}$ reduces to $C_z$.
(iv) For $(a = 0, b = 2)$, $z_{(0,2)}$ reduces to $S_x^2$.

For estimating the general population parameters discussed previously, the general traditional estimator $\tilde{z}_{(a,b)}$ is given by

$$
\tilde{z}_{(a,b)} = \bar{Z}^a S_b^3. \tag{6}
$$

On expressing (6) in terms of errors, we have
where \( k_1 \) and \( k_2 \) are considered as scalar constants which take the values \((0, -1, 1)\), and \( \alpha \), \( \beta \), and \( \beta \) are suitably chosen constants. Expressing (11) in terms of the errors, we get

\[
\begin{align*}
\hat{z}_d &= \left[ \hat{z}_{(a,b)} \right] + wX_e \left[ \frac{k_1(\hat{X} - \hat{X})}{\hat{X} + (\alpha - 1)\hat{X}} \right] \\
& \cdot \left[ \frac{k_3(\hat{S}_e^2 - \hat{S}^2_2)}{\hat{S}_e^2 + (\beta - 1)\hat{S}^2_e} \right] \\
&= \left[ \hat{z}_{(a,b)} \right] + wX_e \left[ \frac{1 + a\hat{e}_0 + b\hat{e}_1}{2} + \frac{(\alpha - 1)}{2} \hat{e}_0 \right] \\
& \cdot \left[ \frac{k_3(\hat{S}_e^2 - \hat{S}^2_2)}{\hat{S}_e^2 + (\beta - 1)\hat{S}^2_e} \right].
\end{align*}
\]

where \( k_1 \), \( k_2 \), and \( k_3 \) are considered as scalar constants which take the values \((0, -1, 1)\), and \( \alpha \), \( \beta \), and \( \beta \) are suitably chosen constants. Expressing (11) in terms of the errors, we get

\[
\begin{align*}
\hat{z}_d &= \left[ \hat{z}_{(a,b)} \right] + wX_e \left[ \frac{1 + a\hat{e}_0 + b\hat{e}_1}{2} + \frac{(\alpha - 1)}{2} \hat{e}_0 \right] \\
& \cdot \left[ \frac{k_3(\hat{S}_e^2 - \hat{S}^2_2)}{\hat{S}_e^2 + (\beta - 1)\hat{S}^2_e} \right].
\end{align*}
\]
Squaring (14) on both the sides, and neglecting error terms of power greater than two, we get

\[(z_d - z_{(a,b)})^2 = [z_{(a,b)}a e_0 + z_{(a,b)} \left( \frac{b}{2} \right) e_1 - \left( z_{(a,b)} \left( \frac{k_1}{\alpha} \right) \right) e_2 - z_{(a,b)} \left( \frac{k_2}{\beta} \right) e_3]^2 \]

\[= \left[ z_{(a,b)} \left\{ a e_0 + \left( \frac{b}{2} \right) e_1 - \left( \frac{k_1}{\alpha} \right) e_2 - \left( \frac{k_2}{\beta} \right) e_3 \right\} - wX e_2 \right]^2. \tag{15} \]

Equivalently,

\[(z_d - z_{(a,b)})^2 = z_{(a,b)}^2 \left[ a^2 e_0^2 + \left( \frac{b^2}{4} \right) e_1^2 + \left( \frac{k_1}{\alpha} \right)^2 e_2^2 + \left( \frac{k_2}{\beta} \right)^2 e_3^2 + 2 \left( \frac{ab}{2} \right) e_0 e_1 - 2a \left( \frac{k_1}{\alpha} \right) e_0 e_2 - 2a \left( \frac{k_2}{\beta} \right) e_0 e_3 \right. \]

\[- 2 \left( \frac{b}{2} \right) \left( \frac{k_1}{\alpha} \right) e_1 e_2 - \left( \frac{b}{2} \right) \left( \frac{k_2}{\beta} \right) e_1 e_3 + 2 \left( \frac{k_1}{\alpha} \right) \left( \frac{k_2}{\beta} \right) e_2 e_3 \]

\[- 2z_{(a,b)} a wX e_0 e_2 - 2z_{(a,b)} \left( \frac{b}{2} \right) wX e_1 e_2 + 2z_{(a,b)} \left( \frac{k_1}{\alpha} \right) wX e_2^2 + 2z_{(a,b)} \left( \frac{k_2}{\beta} \right) wX e_3 + w^2 X e_2. \tag{16} \]

Taking expectations on both sides of the abovementioned equation, we get MSE of \(z_d\), given as

\[\text{MSE}(z_d) = \text{MSE}(z_{(a,b)}) + \frac{z_{(a,b)}^2}{n} \left[ \left( \frac{k_1}{\alpha} \right)^2 C_x^2 + \left( \frac{k_2}{\beta} \right)^2 (\delta_{04} - 1) - 2 \left( \frac{k_1}{\alpha} \right) \left( a_{p_{xX}} C_x + \left( \frac{b}{2} \right) \delta_{21} \right) C_x \right. \]

\[\left. - 2 \left( \frac{k_2}{\beta} \right) \left( a \delta_{12} C_x + \left( \frac{b}{2} \right) (\delta_{22} - 1) \right) + 2 \left( \frac{k_1}{\alpha} \right) \left( \frac{k_2}{\beta} \right) \delta_{03} C_x \right] \tag{17} \]

or

\[\text{MSE}(z_d) = \text{MSE}(z_{(a,b)}) + \frac{z_{(a,b)}^2}{n} \left[ \left( p^2 C_x^2 + Q^2 (\delta_{04} - 1) - 2pF_z (a,b) C_x - 2QF_y (a,b) + + 2PQ \delta_{03} C_x \right) \right. \]

\[\left. + 2 \frac{z_{(a,b)}^2}{n} wX \left[ p_{C_x^2} + Q \delta_{03} C_x \right] - 2 \frac{z_{(a,b)}^2}{n} \left( wX F_z (a,b) C_x + \frac{w^2 X^2}{n} C_x \right) \right]. \tag{18} \]

where

\[P = \frac{k_1}{\alpha} \]

\[Q = \frac{k_2}{\beta} \]

\[F_z (a,b) = \left\{ a_{p_{xX}} C_x + \left( \frac{b}{2} \right) \delta_{21} \right\}, \tag{19} \]

\[F_y (a,b) = \left\{ a_{p_{xX}} C_x + \left( \frac{b}{2} \right) (\delta_{22} - 1) \right\}. \]

Differentiating (17) partially w.r.t P and Q, and equating to zero, we get

\[\begin{bmatrix} z_{(a,b)} C_x & z_{(a,b)} \delta_{03} \end{bmatrix} \begin{bmatrix} P \\ z_{(a,b)} \delta_{03} C_x \end{bmatrix} = \begin{bmatrix} z_{(a,b)} \delta_{03} C_x \end{bmatrix} \]

\[\begin{bmatrix} z_{(a,b)} C_x \delta_{04} - 1 \end{bmatrix} \begin{bmatrix} Q \\ z_{(a,b)} (\delta_{04} - 1) \end{bmatrix} \tag{20} \]

\[= \begin{bmatrix} z_{(a,b)} C_x + \left( \frac{b}{2} \right) \delta_{21} - wX C_x \end{bmatrix}. \]

\[\begin{bmatrix} z_{(a,b)} \delta_{03} C_x \end{bmatrix} \]

\[= \begin{bmatrix} z_{(a,b)} \delta_{03} C_x \end{bmatrix} \]

\[\begin{bmatrix} z_{(a,b)} \delta_{03} C_x \end{bmatrix} \]
After substituting the optimal values of \( P \) and \( Q \), we get the minimum MSE (\( z_d \)) as

\[
MSE(z_d) = MSE(\tilde{z}_{(a,b)}) - \frac{z_{(a,b)}^2}{n} \left[ \frac{[F_3(a,b)]^2 - 2F_2(a,b)F_3(a,b)\delta_{03} + (\delta_{04} - 1)[F_2(a,b)]^2}{(\delta_{04} - \delta_{03}^2 - 1)} \right]
\]

\[
= MSE(\tilde{z}_{(a,b)}) - \frac{z_{(a,b)}^2}{n} \left[ F_2(a,b)\delta_{03} - F_3(a,b) \right]^2 \left[ (\delta_{04} - \delta_{03}^2 - 1) \right]
\]

\[
= MSE(\tilde{z}_{(a,b)}) - \frac{z_{(a,b)}^2}{n} \left[ F_3(a,b) \right]^2 \left[ (\delta_{04} - \delta_{03}^2 - 1) \right]
\]

(21)

Using known population mean \( \bar{X} \) and substituting \( k_2 = 0 \), in the general difference-cum exponential class of estimators (11), we get the following general estimator for estimating population parameters of the sensitive research variable as

\[
z_{d1} = \frac{\tilde{z}_{(a,b)} + w(\bar{X} - \bar{X})}{a + (\alpha - 1)\bar{X}} \exp \left[ \frac{k_1(\bar{X} - \bar{X})}{\bar{X} + (\alpha - 1)\bar{X}} \right],
\]

(22)

where \( k_1 \) takes real values (-1, 0, 1) and \( w \) and \( \alpha \) are suitably chosen constants. Following the abovementioned procedure, the MSE of \( z_{d1} \) is given by

\[
MSE(z_{d1}) = MSE(\tilde{z}_{(a,b)}) + \frac{z_{(a,b)}^2}{n} \left[ \frac{(k_1/\alpha)^2}{\bar{X}} - \frac{\left(\frac{k_1}{\alpha}\right)F_2(a,b)C_x}{2} \right].
\]

The minimum value of MSE (\( z_{d1} \)) is obtained, after minimizing the above equation for \( (k_1/\alpha)_{opt} = F_2(a,b)/C_x \), as

\[
MSE(z_{d1})_{min} = MSE(\tilde{z}_{(a,b)}) - \frac{z_{(a,b)}^2}{n} \left[ F_2(a,b) \right]^2.
\]

(24)

Table 1 illustrates the existing estimators derived from the family of \( z_{d1} \) estimators by selecting appropriate values for \( a, b, w, k_1, \alpha \) and applying them appropriately. Using known population variance \( S_x^2 \) and substituting \( k_1 = 0 \) in the general difference-cum exponential class of estimators (11), we get another general estimator for estimating population parameters of the sensitive research variable as

\[
z_{d2} = \frac{\tilde{z}_{(a,b)} + w(\bar{X} - \bar{X})}{a + (\alpha - 1)\bar{X}} \exp \left[ \frac{k_3(s_x^2 - s_x^2)}{(s_x^2 + (\beta - 1)(s_x^2)} \right],
\]

where \( k_3 \) takes real values (-1, 0, 1) and \( w \) and \( \beta \) are suitably chosen constants. Following the abovementioned procedure, the MSE of \( z_{d2} \) is given by

\[
MSE(z_{d2}) = MSE(\tilde{z}_{(a,b)}) + \frac{z_{(a,b)}^2}{n} \left[ \frac{(k_3/\beta)^2}{\bar{X}} - \frac{\left(\frac{k_3}{\beta}\right)F_3(a,b)}{2} \right].
\]

The minimum value of MSE (\( z_{d2} \)) is obtained, after minimizing the above equation for \( (k_2/\beta)_{opt} = F_3(a,b)/(\delta_{04} - 1) \), as

\[
MSE(z_{d2})_{min} = MSE(\tilde{z}_{(a,b)}) - \frac{z_{(a,b)}^2}{n} \left[ \frac{F_3(a,b)}{\delta_{04} - 1} \right]^2.
\]

Table 2 demonstrates the estimators derived from the family of \( z_{d2} \) estimators by selecting appropriate values for \( a, b, k_2, \alpha \) and applying them appropriately.

3.1. Efficiency Comparison of \( z_d \):

\[
MSE(\tilde{z}_{(a,b)}) - MSE(z_d) = \frac{z_{(a,b)}^2}{n} \left[ \frac{[F_3(a,b)]^2 - 2F_2(a,b)F_3(a,b)\delta_{03} + (\delta_{04} - 1)[F_2(a,b)]^2}{(\delta_{04} - \delta_{03}^2 - 1)} \right] \geq 0,
\]

(25)

\[
MSE(z_{d1}) - MSE(z_d) = \left( \frac{z_{(a,b)}^2}{n} \right) \left( \frac{F_2(a,b)\delta_{03} - F_3(a,b)}{\delta_{04} - \delta_{03}^2 - 1} \right) \geq 0,
\]

(26)

\[
MSE(z_{d2}) - MSE(z_d) = \left( \frac{z_{(a,b)}^2}{n} \right) \left( \frac{(\delta_{04} - 1)F_3(a,b) - \delta_{03}F_3(a,b)}{\delta_{04} - \delta_{03}^2 - 1} \right) \geq 0,
\]

(27)
Table 1: Several members of the $z_d$ family of estimator.

| S.No. | Estimators | $a$ | $b$ | $w$ | $k_1$ | $\alpha$ |
|-------|------------|-----|-----|-----|-------|---------|
| 1     | $z_{d1}(1) = z_{(a,b)}\exp\{(\bar{X} - \bar{x})/\bar{x} + (\alpha - 1)\bar{x}\}$ | a   | b   | 0   | 1     | $\alpha$ |
| 2     | $z_{d1}(2) = z_{(a,b)}\exp\{(S_x^2 - s_x^2)/(S_x^2 + (\beta - 1)(s_x^2)\}$ | 1   | 0   | 0   | 1     | $\alpha$ |
| 3     | $z_{d1}(3) = z_{(a,b)}\exp\{(\bar{X} - \bar{x})/\bar{x} + x\}$ | 1   | 0   | 0   | 1     | 2       |
| 4     | $z_{d1}(4) = [z + w(\bar{X} - \bar{x})]$ | 1   | 0   | 0   | $w$   | $\alpha$ |

Table 2: Several members of the $z_d$ family of estimator.

| S.No. | Estimators | $a$ | $b$ | $w$ | $k_2$ | $\beta$ |
|-------|------------|-----|-----|-----|-------|---------|
| 1     | $z_{d1}(i) = z_{(a,b)}\exp\{(S_x^2 - s_x^2)/(S_x^2 + (\beta - 1)(s_x^2)\}$ | a   | b   | 0   | 1     | $\beta$ |
| 2     | $z_{d1}(2) = z_{(a,b)}\exp\{(S_x^2 - s_x^2)/(S_x^2 + (\beta - 1)(s_x^2)\}$ | 0   | 2   | 0   | 1     | $\beta$ |
| 3     | $z_{d1}(3) = z_{(a,b)}\exp\{(S_x^2 - s_x^2)/(S_x^2 + (\beta - 1)(s_x^2)\}$ | 0   | 2   | 0   | 1     | 2       |

provided that $(\delta_{04} - \delta_{03}^2) \geq 0$, the proposed class of estimator $z_d$ outperforms other competing estimators.

4. Estimation of Population Mean of the Research Variable

Using known population parameters and substituting different values of $a, b, k_1, \alpha, \beta$ as $(1, 0, w, k_1, \alpha, \beta)$ in the general difference-cum-exponential class of estimators (equation 4), we obtain a class of estimators for estimating the population mean of the sensitive research variable as

$$z'_d = \frac{k_1}{\alpha}(\bar{X} - \bar{x})\exp\left[\frac{k_2(S_x^2 - s_x^2)}{\bar{X} + (\alpha - 1)\bar{x}}\right].$$

(29)

The expression of MSE of $z'_d$ up to O(n-1) is given by

$$\text{MSE}(z'_d) = \text{MSE}(\bar{X}) + \frac{\bar{X}^2}{n} \left[p^2C_x^2 + Q^2(\delta_{04} - 1)\right] - 2P\rho_{xz}C_xC_z - 2Q\delta_{12}C_x + 2PQ\delta_{03}C_x]$$

(30)

where

$$P = \frac{k_1}{\alpha},$$

$$Q = \frac{k_2}{\beta}$$

MSE($z'_d$) is minimized for the optimal values of $P$ and $Q$, given by

$$P = \frac{[(\delta_{04} - 1)\rho_{xz} - \delta_{03}\delta_{12}]C_x}{(\delta_{04} - \delta_{03}^2 - 1)C_x} - \frac{w\bar{X}}{Z},$$

(32)

$$Q = \frac{(\delta_{12} - \delta_{03}\rho_{xz})C_x}{(\delta_{04} - \delta_{03}^2 - 1)}$$

On substituting the optimal values of $P$ and $Q$ in (30), the least value of MSE($\bar{X}$') takes the form, given by

$$\text{MSE}(\bar{X}') = \text{Var}(\bar{X}_0) - \frac{Z^2(\delta_{12}^2 - 2\rho_{xz}\delta_{03}\delta_{12} + (\delta_{04} - 1)\rho_{xz}^2)C_x^2}{n(\delta_{04} - \delta_{03}^2 - 1)}.$$  

(33)

Table 3 shows the members of the proposed generalized class of estimators $z_d$, and their values of different constant used in MSES.

5. Relative Efficiency

In the current section, we analyze the theoretical efficiency of the proposed class of estimators and the criteria is discussed under which the estimator performs better than other existing estimators.

(i) $\text{Var}(\bar{X}_0) - \text{MSE}(\bar{X}_0) = \frac{Z^2[\delta_{12}^2 - 2\rho_{xz}\delta_{03}\delta_{12} + (\delta_{04} - 1)\rho_{xz}^2]C_x^2}{(\delta_{04} - \delta_{03}^2 - 1)} \geq 0$

(ii) $\text{MSE}(z_{d1}(2)) - \text{MSE}(\bar{X}_0) = \frac{Z^2[\delta_{12}^2 - 2\rho_{xz}\delta_{03}\delta_{12} + (\delta_{04} - 1)\rho_{xz}^2]C_x^2}{(\delta_{04} - \delta_{03}^2 - 1)} - Z^2/n\rho_{xz}C_x \geq 0$

(iii) $\text{MSE}(z_{d1}(3)) - \text{MSE}(\bar{X}_0) = \frac{Z^2[\delta_{12}^2 - 2\rho_{xz}\delta_{03}\delta_{12} + (\delta_{04} - 1)\rho_{xz}^2]C_x^2}{(\delta_{04} - \delta_{03}^2 - 1)} - Z^2/n[C_x^2/4 - \rho_{xz}C_x] \geq 0$

(iv) $\text{MSE}(z_{d1}(4)) - \text{MSE}(\bar{X}_0) = \frac{Z^2[\delta_{12}^2 - 2\rho_{xz}\delta_{03}\delta_{12} + (\delta_{04} - 1)\rho_{xz}^2]C_x^2}{(\delta_{04} - \delta_{03}^2 - 1)} - Z^2/n[C_x^2/4 - \rho_{xz}C_x] \geq 0$

6. Empirical Study

In a view to analysing the accuracy of the theoretical results, we use an empirical investigation to assess the effectiveness of the suggested class of estimators. The real population included in the study is Cochran (134), page no. 325), where $X$ denotes the number of rooms per block and $Y$ denotes the number of person per block. The descriptive statistics of the real data population are given as
Table 3: Several members of the $\mathcal{F}_d^c$ family of estimator.

| Estimators | a | b | c | d | MSE |
|------------|---|---|---|---|-----|
| $\dot{z}_{d1}(2) = \frac{p}{\lambda} \exp \left[ -\frac{(X - \lambda)}{\lambda} \right]$ | 0 | 1 | 0 | $\alpha$ | $\text{Var}(\dot{z}) \sim \frac{\lambda^2}{n^2} C^2_{d1}$ |
| $\dot{z}_{d1}(3) = \frac{p}{\lambda} \exp \left( -\frac{(X - \lambda)}{\lambda} \right)$ (Gupta et al. [12]) | 0 | 1 | 0 | 2 | $\text{Var}(\dot{z}) \sim \frac{\lambda^2}{n^2} C^2_{d1}$ |
| $\dot{z}_{d1}(4) = [\dot{z} + w(x - \lambda)]$ (Difference estimator) | 1 | 0 | 0 | $\alpha$ | $\text{Var}(\dot{z}) \sim \frac{\lambda^2}{n^2} C^2_{d1}$ |

Table 4: MSE and PRE of recommended and competing estimators for empirical data.

| S. No. | Estimators | MSE | PRE |
|--------|------------|-----|-----|
| 1      | $\bar{z}_0$ | 23.91043 | 100.00 |
| 2      | $\dot{z}_{d1}(2)$ | 13.75848 | 173.78 |
| 3      | $\dot{z}_{d1}(3)$ | 33.01291 | 72.43 |
| 4      | $\dot{z}_{d1}(4)$ | 13.75848 | 173.79 |
| 5      | $\bar{z}_j$ | 11.72704 | 203.89 |

Table 5: PRE of proposed and competing estimators for Normal population.

| $X^* \sim N(50, 4)$, $Y^* \sim N(40, 3)$, $S \sim N(0, 1)$ | Estimators | $\rho = 0.6$ | $\rho = 0.7$ | $\rho = 0.8$ | $\rho = 0.9$ |
|---------------|------------|---------------|---------------|---------------|---------------|
| $N = 40$      | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 157.39       | 188.02        | 253.31        | 437.25        |
|               | $\dot{z}_{d1}(3)$ | 73.92        | 67.38         | 62.50         | 57.17         |
|               | $\dot{z}_{d1}(4)$ | 157.39       | 189.02        | 253.31        | 437.25        |
|               | $\bar{z}_j$ | 160.83        | 194.10        | 257.55        | 454.78        |
|               | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 146.11       | 177.43        | 234.58        | 361.75        |
| $N = 60$      | $\dot{z}_{d1}(3)$ | 77.57        | 69.85         | 63.60         | 58.94         |
|               | $\dot{z}_{d1}(4)$ | 146.11       | 177.43        | 234.58        | 361.75        |
|               | $\bar{z}_j$ | 164.15        | 199.46        | 263.62        | 404.02        |
|               | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 122.61       | 143.31        | 180.72        | 257.52        |
| $n = 80$      | $\dot{z}_{d1}(3)$ | 94.95        | 80.56         | 69.70         | 62.19         |
|               | $\dot{z}_{d1}(4)$ | 122.61       | 143.31        | 180.72        | 257.52        |
|               | $\bar{z}_j$ | 126.90        | 148.24        | 186.71        | 265.23        |
|               | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 150.99       | 183.45        | 242.34        | 374.37        |
| $n = 100$     | $\dot{z}_{d1}(3)$ | 75.67        | 68.81         | 63.14         | 58.79         |
|               | $\dot{z}_{d1}(4)$ | 150.99       | 183.45        | 242.34        | 374.37        |
|               | $\bar{z}_j$ | 154.06        | 187.08        | 246.87        | 380.30        |

Table 6: PRE of proposed and competing estimators for beta population.

| $X^* \sim B(15, 3)$, $Y^* \sim B(3, 5)$, $S \sim B(2, 3)$ | Estimators | $\rho = 0.6$ | $\rho = 0.7$ | $\rho = 0.8$ | $\rho = 0.9$ |
|---------------|------------|---------------|---------------|---------------|---------------|
| $N = 40$      | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 100.87       | 103.30        | 108.50        | 119.73        |
|               | $\dot{z}_{d1}(3)$ | 94.92        | 92.63         | 88.83         | 86.00         |
|               | $\dot{z}_{d1}(4)$ | 100.87       | 103.30        | 108.50        | 119.73        |
|               | $\bar{z}_j$ | 101.57        | 104.15        | 109.63        | 121.49        |
|               | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 118.84       | 125.00        | 132.26        | 140.43        |
| $N = 60$      | $\dot{z}_{d1}(3)$ | 86.59        | 84.14         | 81.88         | 79.87         |
|               | $\dot{z}_{d1}(4)$ | 118.84       | 125.00        | 132.26        | 140.43        |
|               | $\bar{z}_j$ | 122.48        | 128.72        | 135.99        | 144.00        |
|               | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 117.23       | 123.74        | 132.22        | 143.66        |
| $n = 80$      | $\dot{z}_{d1}(3)$ | 87.70        | 85.16         | 82.66         | 80.12         |
|               | $\dot{z}_{d1}(4)$ | 117.23       | 123.74        | 132.22        | 143.66        |
|               | $\bar{z}_j$ | 117.95        | 124.56        | 133.20        | 144.89        |
|               | $\bar{z}_0$ | 100.00        | 100.00        | 100.00        | 100.00        |
|               | $\dot{z}_{d1}(2)$ | 107.58       | 112.60        | 120.22        | 132.72        |
| $n = 100$     | $\dot{z}_{d1}(3)$ | 93.57        | 89.97         | 86.18         | 82.04         |
|               | $\dot{z}_{d1}(4)$ | 107.58       | 112.60        | 120.22        | 132.72        |
\[
\text{PRE} = \frac{\text{MSE}(\hat{\theta}) - \text{MSE}(\bar{\theta})}{\text{MSE}(\bar{\theta})} \times 100.
\]  

7. Simulation Study

A simulation model is constructed to examine the dynamics of the suggested class of estimators, incorporating the additive scrambling model given by Pollock and Bek [33]. The mean square errors of the proposed class of estimators and the competing estimators are compared using six artificial populations, namely, normal, beta,
gamma, Poisson, uniform, and log-normal with differing population parameters. Following the procedure adopted by Singh and Horn [35], we generated study and ancillary variables as \( Y = 10 + \sqrt{(1 - \rho_{xy}^2)}Y^* + \rho_{xy}(S_y/S_x)X^* \) and \( X = 5 + X^* \), where \( X^* \) and \( Y^* \) are independent variates of pertaining family of distributions. The sensitive variable \( Z \) is derived by scrambling the variable \( Y \) with \( S \) where \( S \) follows the corresponding parent distribution with predefined parameters. With various selections of correlation coefficient, i.e., 0.6, 0.7, 0.8, and 0.9, samples of different sizes, i.e., 40, 60, 80, and 100, are collected from each bivariate distribution.

### 8. Results and Discussion

The findings of the study for real data set are presented in Table 4 which describes the MSE and PRE of suggested class of estimators \( \hat{Z}_d \) along with the competing estimators, i.e., RRT sample mean \((\bar{X}_0)\), \((\bar{Z}_d)\), Gupta et al. [12]’s estimator \((\hat{Z}_d)\), and difference estimator \((\bar{Z}_d)\). Table 5–10 summarise Table 6 PRE Table 7 of known Table 8 and proposed Table 9 estimators for various sample sizes, i.e., 40, 60, 80, and 100) and correlation coefficients, i.e., (0.6, 0.7, 0.8, and 0.9). Under simulation investigation, the proposed estimate \( \hat{Z}_d \) has the lowest MSE for all available regression coefficient

#### Table 9: PRE of proposed and competing estimators for uniform population.

| \( X^* \sim U(50,100) \) \( Y^* \sim U(40, 80) \) \( S \sim U(5,10) \) | \( \hat{Z}_0 \) | \( \hat{Z}_{d1} \) | \( \hat{Z}_{d2} \) | \( \hat{Z}_{d3} \) | \( \hat{Z}_{d4} \) | \( \hat{Z}_{d5} \) |
|---|---|---|---|---|---|---|
| \( \hat{Z}_0 \) | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| \( \hat{Z}_{d1} \) | 175.69 | 244.64 | 438.10 | 320.25 | 75.85 | 51.70 |
| \( \hat{Z}_{d2} \) | 72.12 | 63.22 | 370.25 | 370.25 | 75.85 | 75.85 |
| \( \hat{Z}_{d3} \) | 175.69 | 244.64 | 438.10 | 320.25 | 75.85 | 75.85 |
| \( \hat{Z}_{d4} \) | 181.03 | 253.28 | 459.54 | 3959.40 | 75.85 | 75.85 |
| \( \hat{Z}_{d5} \) | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| \( \hat{Z}_{d6} \) | 167.76 | 221.54 | 355.87 | 1266.64 | 75.85 | 75.85 |

#### Table 10: PRE of proposed and competing estimators for uniform population.

| \( X^* \sim LN(10,2) \) \( Y^* \sim LN(10,3) \) \( S \sim LN(5,3) \) | \( \hat{Z}_0 \) | \( \hat{Z}_{d1} \) | \( \hat{Z}_{d2} \) | \( \hat{Z}_{d3} \) |
|---|---|---|---|---|
| \( \hat{Z}_0 \) | 100.00 | 100.00 | 100.00 | 100.00 |
| \( \hat{Z}_{d1} \) | 174.89 | 220.82 | 313.07 | 591.36 |
| \( \hat{Z}_{d2} \) | 72.12 | 63.22 | 370.25 | 370.25 |
| \( \hat{Z}_{d3} \) | 174.89 | 220.82 | 313.07 | 591.36 |
| \( \hat{Z}_{d4} \) | 152.22 | 225.03 | 319.08 | 602.90 |
| \( \hat{Z}_{d5} \) | 100.00 | 100.00 | 100.00 | 100.00 |
| \( \hat{Z}_{d6} \) | 165.25 | 202.36 | 267.22 | 406.01 |

#### Table 11: PRE of proposed and competing estimators for uniform population.

| \( X^* \sim LN(10,2) \) \( Y^* \sim LN(10,3) \) \( S \sim LN(5,3) \) | \( \hat{Z}_0 \) | \( \hat{Z}_{d1} \) | \( \hat{Z}_{d2} \) | \( \hat{Z}_{d3} \) |
|---|---|---|---|---|
| \( \hat{Z}_0 \) | 100.00 | 100.00 | 100.00 | 100.00 |
| \( \hat{Z}_{d1} \) | 174.89 | 220.82 | 313.07 | 591.36 |
| \( \hat{Z}_{d2} \) | 72.12 | 63.22 | 370.25 | 370.25 |
| \( \hat{Z}_{d3} \) | 174.89 | 220.82 | 313.07 | 591.36 |
| \( \hat{Z}_{d4} \) | 152.22 | 225.03 | 319.08 | 602.90 |
| \( \hat{Z}_{d5} \) | 100.00 | 100.00 | 100.00 | 100.00 |
| \( \hat{Z}_{d6} \) | 165.25 | 202.36 | 267.22 | 406.01 |
values, indicating that it is more efficient than the existing estimator in the literature reveal.

9. Conclusion

(i) A new difference-cum exponential ratio type estimator in SRWOR is proposed in this study, which takes advantage of the additional information available to provide a more accurate general parameter estimate. Adopting the estimator $z_d$ provided in Section 3, it is feasible to estimate a variety of population parameters. Tables 1 and 2 illustrate specific applications of the class of estimators for estimating the mean and variance of the sensitive variable using various choices of arbitrary constants.

(ii) In Section 4, the specific situations of the generalized estimator are presented in order to estimate the mean of the sensitive research variable, and the relative efficiency of the class of estimators is discussed.

(iii) We tested the effectiveness of the suggested estimator on real and simulated data sets for the case of population mean estimation solely. The findings of the numerical illustration are demonstrated in Table 4 which indicates the supremacy of the recommended class of estimator over other estimators discussed under Section 4.

(iv) We can infer from the results of the simulation study conducted on six artificial populations that for all the different values of correlation coefficient, the PRE of the recommended class of estimators is greater than that of the other estimators. Hence, the suggested class of estimators is more efficient than other estimators.

(v) The suggested general class of estimator can be used to estimate various population parameters of the sensitive research variable. Also, the suggested approach can be used in different sampling techniques.

Abbreviations:

MSE: Mean square error
PRE: Percent relative efficiency
RRT: Randomized response technique.

Data Availability

The data present in this study are within this manuscript.

Ethical Approval

Ethical approval is not required for this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors’ Contributions

GKV conceptualized the study. SKY developed the data analysis methodology. TB analyzed the data and drafted this manuscript. SKY reviewed the drafted manuscript and provided technical inputs to GKV to finalize the manuscript. All the authors read and approved the final manuscript.

Acknowledgments

The authors are deeply indebted to the editor Professor Sat N. Gupta and learned referees for their constructive suggestions leading to improving the quality of the contents and presentation of the original manuscript.

References

[1] D. G. Horvitz, B. V. Shah, and W. R. Simmons, “The unrelated question RR model,” in Proceedings of the Social Statistics Section of the American Statistical Association, vol. 65-72, ASA, Alexandria, VA, June 1967.

[2] B. G. Greenberg, R. R. Kuebler, J. R. Abernathy, and D. G. Horvitz, “Application of the randomized response technique in obtaining quantitative data,” Journal of the American Statistical Association, vol. 66, no. 334, pp. 243–250, 1971.

[3] S. L. Warner, “Randomized response: a survey technique for eliminating evasive answer bias,” Journal of the American Statistical Association, vol. 60, no. 309, pp. 63–69, 1965.

[4] B. H. Eichhron and L. S. Hayre, “Scrambled randomized response methods for obtaining sensitive quantitative data,” Journal of Statistical Planning and Inference, vol. 7, no. 4, pp. 307–316, 1983.

[5] S. Gupta, B. Gupta, and S. Singh, “Estimation of sensitivity level of personal interview survey questions,” Journal of Statistical Planning and Inference, vol. 100, no. 2, pp. 239–247, 2002.

[6] A. Saha, “A randomized response technique for quantitative data under unequal probability sampling,” Journal of Statistical Theory and Practice, vol. 2, no. 4, pp. 589–596, 2008.

[7] G. Diana and P. F. Perri, “New scrambled response models for estimating the mean of a sensitive quantitative character,” Journal of Applied Statistics, vol. 37, no. 11, pp. 1875–1890, 2010.

[8] R. Sousa, J. Shabbir, P. C. Real, and S. Gupta, “Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information,” Journal of Statistical Theory and Practice, vol. 4, no. 3, pp. 495–507, 2010.

[9] N. Koyuncu, S. Gupta, and R. Sousa, “Exponential-type estimators of the mean of a sensitive variable in the presence of non sensitive auxiliary information,” Communications in Statistics - Simulation and Computation, vol. 43, no. 7, pp. 1583–1594, 2014.

[10] G. Kalucha, S. Gupta, and B. K. Dass, “Ratio estimation of finite population mean using optional randomized response models,” Journal of Statistical Theory and Practices, vol. 9, no. 3, pp. 633–645, 2015.

[11] S. Gupta, J. Shabbir, R. Sousa, and P. C. Real, “Estimation of the mean of a sensitive variable in the presence of auxiliary information,” Communications in Statistics - Theory and Methods, vol. 41, no. 13-14, pp. 2394–2404, 2012.

[12] S. Gupta, T. Zatezalo, and J. Shabbir, “A Generalized mixture estimator of the mean of a sensitive variable in the presence of
non-sensitive auxiliary information,” *Statistics and Applications*, vol. 15, no. 1-2, pp. 27–36, 2017.

[13] M. Noor-Ul-Amin, N. Mushtaq, and M. Hanif, “Estimation of mean using generalized optional scrambled responses in the presence of non-sensitive auxiliary variable,” *Journal of Statistics & Management Systems*, vol. 21, no. 2, pp. 287–304, 2018.

[14] S. Khalil, Q. Zhang, and S. Gupta, “Mean estimation of sensitive variables under measurement errors using optional RRT models,” *Communications in Statistics - Simulation and Computation*, vol. 50, no. 5, pp. 1417–1426, 2019.

[15] E. Zahid and J. Shabbir, “Estimation of finite population mean for a sensitive variable using dual auxiliary information in the presence of measurement errors,” *PLOS ONE*, vol. 14, no. 2, pp. 1–17, 2019.

[16] A. Sanaullah, I. Saleem, S. Gupta, and M. Hanif, “Mean estimation with generalized scrambling using two-phase sampling,” *Communications in Statistics - Simulation and Computation*, 2020b.

[17] A. Sanaullah, I. Saleem, and J. Shabbir, “Use of scrambled response for estimating mean of the sensitivity variable,” *Communications in Statistics - Theory and Methods*, vol. 49, no. 11, pp. 2634–2647, 2020a.

[18] I. Saleem, A. Sanaullah, and M. Hanif, “Double-sampling regression-cum-exponential estimator of the mean of a sensitive variable,” *Mathematical Population Studies*, vol. 26, no. 3, pp. 163–182, 2019.

[19] Z. Waseem, H. Khan, and J. Shabbir, “Generalized exponential type estimator for the mean of sensitive variable in the presence of non-sensitive auxiliary variable,” *Communications in Statistics - Theory and Methods*, vol. 50, no. 14, pp. 3477–3488, 2020.

[20] G. N. Singh, A. Kumar, and G. K. Vishwakarma, “Some alternative additive randomized response models for estimation of population mean of quantitative variable in the presence of scramble variable,” *Communications in Statistics - Simulation and Computation*, vol. 49, no. 11, pp. 2785–2807, 2020.

[21] G. K. Vishwakarma and N. Singh, “Computing the effect of measurement errors under additive scramble response for sensitive variable,” *Journal of Computational and Applied Mathematics*, vol. 395, Article ID 113593, 2021.

[22] C. T. Isaki, “Variance estimation using auxiliary information,” *Journal of the American Statistical Association*, vol. 78, no. 381, pp. 117–123, 1983.

[23] S. Bahl and R. K. Tuteja, “Ratio and product type exponential estimators,” *Journal of Information and Optimization Sciences*, vol. 12, no. 1, pp. 159–164, 1991.

[24] L. Upadhyaya and H. Singh, “An estimator for population variance that utilizes the kurtosis of an auxiliary variable in sample surveys,” *Vikram Mathematical Journal*, vol. 19, no. 1, pp. 14–17, 1999.

[25] C. Kadilar and H. Cingi, “Ratio estimators for the population variance in simple and stratified random sampling,” *Applied Mathematics and Computation*, vol. 173, no. 2, pp. 1047–1059, 2006.

[26] V. Dubey and H. Sharma, “On estimating population variance using auxiliary information,” *Statistics in Transition New Series*, vol. 9, no. 11, pp. 7–18, 2008.

[27] H. P. Singh and G. K. Vishwakarma, “Modified exponential ratio and product estimators for finite population mean in double sampling,” *Austrian Journal of Statistics*, vol. 36, no. 3, pp. 217–225, 2007.