Statistics and Applications {ISSN 2452-7395 (online)}
Volume 15 Nos. 1&2, 2017 (New Series), pp 27-36

A Generalized Mixture Estimator Of The Mean Of A Sensitive Variable In The Presence Of Non-Sensitive Auxiliary Information

Sat Gupta[1], Tanja Zatezalo[1] & Javid Shabbir[2]

[1]*Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro, North Carolina, USA*

[2] *Department of Statistics, Quaid-i-Azam University, Islambad, Pakistan*

Received: December 12, 2016; Revised: January 15, 2017; Accepted: January 25, 2017

Abstract

Gupta et al. (2012) proposed a generalized regression-cum-ratio estimator and Koyuncu et al. (2014) proposed a generalized exponential estimator for the mean of the sensitive variable utilizing a non sensitive auxiliary variable. We propose a new generalized mixture estimator for estimating the population mean of a sensitive study variable. The expressions for Bias and Mean Square Error are derived up to the first order of approximation. Numer-ical examples show that the proposed generalized mixture estimator performs better than many of the existing estimators.

Keywords: Generalized regression-cum-ratio estimator, Generalized exponential estimator, Generalized mixture Estimator, Population mean, Auxiliary information

1 Introduction

Randomized response technique (RRT) can be used to estimate the mean of a sensitive variable Y where direct observation on Y is subject to bias. We assume a non sensitive auxiliary variable X is available and can be observed directly. Sousa et al. (2010) introduced a ratio type estimator and Gupta et al. (2012) proposed a regression and generalized regression-cum-ratio estimators based on RRT models to deal with this situation. Following Bahl & Tuteja (1991), Koyuncu et al. (2014) also proposed a generalized exponential type estimator to improve the efficiency of the mean estimator based on RRT models.

In this paper we propose an ordinary exponential ratipo type estimator and two generalized mixture estimators where the RRT estimators of the mean of Y are further improved by using information from an auxiliary variable X. Expressions for the Bias and Mean Square Error are derived up to the first order of approximation. We will use the following notations.

Let Y be the sensitive study variable which cannot be observed directly. Let X be a non sensitive auxiliary variable which has a positive correlation with Y, and let S be a scrambling variable. Assume that S is 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 the true response for X. Let a random sample of size n be drawn without replacement from a finite population $U = (U_1, U_2, \ldots, U_N)$. For ith population dement, let y_i and x_i respectively be the values of the study variable Y and auxiliary variable X. Let $\overline{Y} = E(Y), \overline{X} = E(X)$ and $\overline{Z} = E(Z)$ be the population means for Y, X and Z respectively. We assume that the population mean $\overline{X}$ and the population variance S_x^2 of the auxiliary variable are known. Also, assume that population mean and the population variance for the scrambling variable S are known and given as $\overline{S} = E(S) = 0$

Corresponding author: Sat Gupta
E-mail: sngupta@uncg.edu

and S_S^2. Thus $E(Z) = E(Y)$ and $C_z^2 = C_y^2 + \left(S_s^2/\overline{Y}^2\right)$, where C_z and C_y are the coefficients of the variation of Z and Y respectively. We will use the same error terms as in Sukhatme and Sukhatme (1970), given as:

$e_z = \frac{\overline{z}-\overline{Z}}{\overline{Z}}$ and $e_x = \frac{\overline{x}-\overline{X}}{\overline{X}}$, where $E(e_z) = E(e_x) = 0$ and $E(e_z^2) = \lambda C_z^2$, $E(e_x^2) = \lambda C_x^2$, $E(e_z e_x) = \lambda C_{zx} = \lambda \rho_{zx} C_z C_x$, and $\lambda = \left(\frac{1}{n} - \frac{1}{N}\right)$.

2 Some Existing Estimators

In this section we will give some existing estimators with corresponding bias and mean square error.

2.1 RRT Sample mean

If information on X is ignored, then an unbiased estimator of $\overline{Y}$ is the ordinary RRT sample mean $(\overline{z})$ given by:

$$\hat{\mu}_Y = \overline{z}. \tag{1}$$

The MSE of $\hat{\mu}_Y$ is given by:

$$MSE(\hat{\mu}_Y) = \lambda \left(S_y^2 + S_s^2\right), \tag{2}$$

where $S_Y^2 = \frac{1}{N-1}\sum_{i=1}^N (y_i - \overline{Y})^2$, and $S_s^2 = \frac{1}{N-1}\sum_{i=1}^N \left(s_i - \overline{S}\right)^2$ are the population variances of the study sensitive variable (Y) and the scrambling variable (S).

2.2 RRT Ratio estimator

Sousa et al.(2010) proposed the ratio type estimator of the mean of a sensitive variable (Y) using a non sensitive auxiliary variable (X) given by:

$$\hat{\mu}_R = \overline{z}\frac{\overline{X}}{\overline{x}}. \tag{3}$$

The bias and the mean square error of this ratio estimator, up to the first order of approximation, are given by:

$$Bias(\hat{\mu}_R) \approx \lambda \overline{Y}\left(C_x^2 - \rho_{zx} C_z C_x\right), \tag{4}$$

$$MSE(\hat{\mu}_R) \approx \lambda \overline{Y}^2\left(C_x^2 - 2\rho_{zx} C_z C_x + C_z^2\right). \tag{5}$$

2.3 RRT Transformed ratio type estimator

Sousa et al.(2010) proposed the transformed ratio type estimator given by:

$$\hat{\mu}_{TR} = \overline{z}\left(\frac{c\overline{X} + d}{c\overline{x} + d}\right), \tag{6}$$

where c and d are the unit-free parameters, which may be quantities such as the coefficient of skewness $\beta_1(x)$ and coefficient of kurtosis $\beta_2(x)$ of the auxiliary variable (X). The bias and the mean square error of this estimator, up to the first order of approximation are given by:

$$Bias(\hat{\mu}_{TR}) \approx \lambda \overline{Y}\left(\eta^2 C_x^2 - \eta \rho_{zx} C_z C_x\right), \tag{7}$$

$$MSE(\hat{\mu}_{TR}) \approx \lambda \bar{Y}^2 \left(\eta^2 C_x^2 - 2\eta \rho_{zx} C_z C_x + C_z^2\right), \tag{8}$$

where $\eta = \frac{c\bar{X}}{c\bar{X}+d}$.

2.4 RRT Regression estimator

Gupta et al.(2012) proposed an ordinary regression type estimator of the population mean $\bar{Y}$ given by:

$$\hat{\mu}_{Reg} = \bar{z} + \hat{\beta}_{zx}(\bar{X} - \bar{x}), \tag{9}$$

where $\hat{\beta}_{zx} = \frac{s_{zx}}{s_x^2} = \frac{s_{yx}}{s_x^2}$ is the sample regression coefficient between Z and X. The bias of this regression estimator, up to the first order of approximation, is given as:

$$Bias(\hat{\mu}_{Reg}) \approx -\lambda \beta_{zx} \left(\frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}}\right), \tag{10}$$

where $\beta_{zx} = \frac{S_{zx}}{S_x^2} = \frac{S_{yx}}{S_x^2} = \rho_{yx}\frac{S_y}{S_x} = \beta_{yx}$ is the population regression coefficient and $\mu_{rs} = \sum_{i=1}^{n}(z_i - \bar{Z})^r(x_i - \bar{X})^s$. The mean square error of the regression estimator, up to the first order of approximation, is given as:

$$MSE(\hat{\mu}_{Reg}) \approx \lambda \bar{Y}^2 C_z^2\left(1-\rho_{zx}^2\right) = \lambda S_y^2 \left[\left(1 + \frac{S_s^2}{S_y^2}\right) - \rho_{yx}^2\right]. \tag{11}$$

2.5 Gupta et al. (2012) generalized RRT Regression-Cum-Ratio estimator

Gupta et al. (2012) proposed a generalized regression-cum-ratio estimator given as:

$$\hat{\mu}_{GRR} = \left[k_1\bar{z} + k_2(\bar{X} - \bar{x})\right]\left(\frac{\bar{X}}{\bar{x}}\right), \tag{12}$$

where k_1 and k_2 are suitably chosen constants. The bias of this estimator, up to the first order of approximation, is given by:

$$Bias(\hat{\mu}_{GRR}) = (k_1 - 1)\bar{Y} + k_1\bar{Y}\lambda(C_x^2 - \rho_{zx}C_zC_x) + k_2\bar{X}\lambda C_x^2. \tag{13}$$

The optimum values of k_1 and k_2 and corresponding mean square error, are given by

$$k_{1(opt)} = \frac{1-\lambda C_x^2}{1-\lambda\left[C_x^2 - C_z^2(1-\rho_{zx}^2)\right]}, \tag{14}$$

$$k_{2(opt)} = \frac{\bar{Y}}{\bar{X}}\left[1 + k_{1(opt)}\left(\frac{\rho_{zx}C_z}{C_x} - 2\right)\right], \tag{15}$$

and

$$MSE(\hat{\mu}_{GRR})_{min} \approx \bar{Y}^2 \frac{\lambda C_z^2\left[1-\rho_{zx}^2\right]\left[1-\lambda C_x^2\right]}{\lambda C_z^2\left[1-\rho_{zx}^2\right] + \left[1-\lambda C_x^2\right]}. \tag{16}$$

2.6 Koyuncu et al. (2014) generalized exponential estimator

Following Bahl & Tuteja (1991) and Gupta et al. (2012), Koyuncu et al. (2014) proposed a generalized exponential type estimator given by

$$\hat{\mu}_{GER} = \left[w_1\bar{z} + w_2(\bar{X} - \bar{x})\right] \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right). \tag{17}$$

The bias of this estimator, up to the first order of approximation, is given by

$$Bias(\hat{\mu}_{GER}) \approx (w_1 - 1)\bar{Y} + \lambda\left[\frac{1}{2}w_1\bar{Y}\left(\frac{3}{4}C_x^2 - C_{zx}\right) + \frac{1}{2}w_2\bar{X}C_x^2\right]. \tag{18}$$

The minimum mean square error at the optimum values of w_1 and w_2, are given by

$$w_{1(opt)} = \frac{1 - \frac{1}{8}\lambda C_x^2}{1 + \lambda C_z^2(1 - \rho_{zx}^2)}, \tag{19}$$

$$w_{2(opt)} = \frac{\bar{Y}}{\bar{X}}\left[\frac{1}{2} - w_{1(opt)}\left(1 - \rho_{zx}\frac{C_z}{C_x}\right)\right], \tag{20}$$

and

$$MSE_{min}(\hat{\mu}_{GER}) \approx \bar{Y}^2\left[\left(1 - \frac{1}{4}\lambda C_x^2\right) - \frac{(1 - \frac{1}{8}\lambda C_x^2)^2}{1 + C_z^2\left(1 - \rho_{zx}^2\right)}\right], \tag{21}$$

or

$$MSE_{min}(\hat{\mu}_{GER}) \approx \left\{\frac{MSE(\hat{\mu}_{Reg})}{\left[1 + \frac{MSE(\hat{\mu}_{Reg})}{\bar{Y}^2}\right]} - \frac{\lambda C_x^2\left[MSE(\hat{\mu}_{Reg}) + \lambda\frac{1}{16}C_x^2\bar{Y}^2\right]}{4\left[1 + \frac{MSE(\hat{\mu}_{Reg})}{\bar{Y}^2}\right]}\right\}. \tag{22}$$

3 Proposed Generalized Mixture RRT Estimator

Following Bahal & Tuteja we propose the exponential ratio type estimator for estimating the population mean of the sensitive variable using a non sensitive auxiliary variable. This estimator is given by:

$$\hat{\mu}_{ER} = \bar{z}\exp\left(\frac{\bar{X} - \bar{x}}{\bar{x} + \bar{X}}\right) \tag{23}$$

where $\bar{z}$ and $\bar{x}$ are the sample means of the reported responses and the auxiliary variable, respectively. Using to the first order of approximation, the estimator can be written as:

$$\hat{\mu}_{ER} - \bar{Z} \approx \bar{Z}\left(e_z - \frac{1}{2}e_x - \frac{1}{2}e_ze_x + \frac{3}{8}e_x^2\right) \tag{24}$$

Recognizing that $\bar{Z} = \bar{Y}$ in (24), the bias and mean square error of the exponential ratio type estimator are given by:

$$Bias(\hat{\mu}_{ER}) \approx \lambda\bar{Y}\frac{1}{2}\left(\frac{3}{4}C_x^2 - \rho_{zx}C_zC_x\right), \quad \text{and} \tag{25}$$

$$MSE(\hat{\mu}_{ER}) \approx \lambda \bar{Y}^2 \frac{1}{4}\left(4C_z^2 - 4\rho_{zx}C_zC_x + C_x^2\right). \tag{26}$$

It can be verified easily that:

a) $MSE(\hat{\mu}_{ER}) < MSE(\hat{\mu}_Y)$ if $\rho_{zx} > \frac{1}{4}\frac{C_x}{C_z}$

b) $MSE(\hat{\mu}_{ER}) < MSE(\hat{\mu}_R)$ if $\rho_{zx} < \frac{3}{4}\frac{C_x}{C_z}$

By combining the regression, ratio and exponential estimators we furher generalize the estimator (23) and propose a generalized mixture estimator given by:

$$\hat{\mu}_{GR} = \left\{ d_1\bar{z}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha} + d_2\left(\bar{X} - \bar{x}\right)\right\} \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{27}$$

where d_i $(i = 1, 2)$ and α are suitably chosen constants. We will consider two values for α ($\alpha = 1$ and $\alpha = 2$). To obtain the bias and mean square error, up to the first order of approximation, $\hat{\mu}_{GR}$ can be written in terms e_y and e_x as:

$$\hat{\mu}_{GR} = \left[d_1\bar{Z}(1+e_z)(1+e_x)^{-\alpha} - d_2\bar{X}e_x\right] \exp\left[\left(-\frac{e_x}{2}\right)\left(1 + \frac{e_x}{2}\right)^{-1}\right] \tag{28}$$

Note that,

$$\hat{\mu}_{GR} - \bar{Z} \approx (d_1 - 1)\bar{Z} + d_1\bar{Z}\left(e_z - Ae_x - Ae_ze_x + Be_x^2\right) - d_2\bar{X}\left(e_x - \frac{1}{2}e_x^2\right), \tag{29}$$

where

$$A = \alpha + \frac{1}{2} \qquad \text{and} \qquad B = \frac{1}{2}\alpha(\alpha+2) + \frac{3}{8}. \tag{30}$$

By taking expectation of (29) and recognizing that $\bar{Z} = \bar{Y}$, the bias of this estimator, up to the first order of approximation, is given by:

$$Bias(\hat{\mu}_{GMR}) \approx (d_1 - 1)\bar{Y} + \lambda d_1\bar{Y}\left(BC_x^2 - A\rho_{zx}C_zC_x\right) + \lambda d_2\bar{X}\,\frac{1}{2}C_x^2 \tag{31}$$

Squaring (29) and using first order of approximation, we get:

$$\begin{aligned}(\hat{\mu}_{GR} - \bar{Z})^2 \approx{}& (d_1-1)^2\bar{Z}^2 + d_1^2\bar{Z}^2\left[2e_z - 2Ae_x - 4Ae_ze_x + e_z^2 + \left(A^2 + 2B\right)e_x^2\right] \\ &+ d_2^2\bar{X}^2e_x^2 - 2d_1\bar{Z}^2\left[e_z - Ae_x - Ae_ze_x + Be_x^2\right] \\ &- 2d_1d_2\bar{X}\bar{Z}\left[e_x + e_ze_x - \left(A + \frac{1}{2}\right)e_x^2\right] + 2d_2\bar{X}\bar{Z}\left(e_x - \frac{1}{2}e_x^2\right).\end{aligned} \tag{32}$$

By taking expectation of (32) and recognizing that $\bar{Z} = \bar{Y}$, the mean square error of the proposed estimator, up to the first order of approximation, is given by:

$$\begin{aligned} MSE(\hat{\mu}_{GR}) \approx & (d_1-1)^2\bar{Y}^2+\lambda d_1^2\bar{Y}^2\left[\left(A^2+2B\right)C_x^2-4A\rho_{zx}C_zC_x+C_z^2\right] \\ & +\lambda d_2^2\bar{X}^2C_x^2-2\lambda d_1\bar{Y}^2\left[BC_x^2-A\rho_{zx}C_zC_x\right] \\ & -2\lambda d_1d_2\bar{X}\bar{Y}\left[\rho_{zx}C_zC_x-\left(A+\frac{1}{2}\right)C_x^2\right]-\lambda d_2\bar{X}\bar{Y}C_x^2. \end{aligned} \tag{33}$$

By taking partial derivatives of (33) with respect to d_1 and d_2, we get:

$$\begin{aligned} \frac{\partial MSE(\hat{\mu}_{GR})}{\partial d_1} = & 2(d_1-1)\bar{Y}^2+2\lambda d_1\bar{Y}^2\left[(A^2+2B)C_x^2-4A\rho_{zx}C_zC_x+C_z^2\right] \\ & -2\lambda\bar{Y}^2\left[BC_x^2-A\rho_{zx}C_zC_x\right]-2\lambda d_2\bar{X}\bar{Y}\left[\rho_{zx}C_zC_x-\left(A+\frac{1}{2}\right)C_x^2\right], \end{aligned} \tag{34}$$

and

$$\frac{\partial MSE(\hat{\mu}_{GR})}{\partial d_2}=2\lambda d_2\bar{X}^2C_x^2-2\lambda d_1\bar{X}\bar{Y}\left[\rho_{zx}C_zC_x-\left(A+\frac{1}{2}\right)\right]-\lambda\bar{X}\bar{Y}C_x^2. \tag{35}$$

$\frac{\partial MSE(\hat{\mu}_{GR})}{\partial d_i}=0$ $(i=1,2)$, the optimum value of d_1 and d_2 are given by:

$$d_{1(opt)}=\frac{1+\lambda\left[\left(B-\frac{1}{2}A-\frac{1}{4}\right)C_x^2+\left(\frac{1}{2}-A\right)\rho_{zx}C_zC_x\right]}{1+\lambda\left[\left(2B-A-\frac{1}{4}\right)C_x^2+(1-2A)\rho_{zx}C_zC_x+(1-\rho_{zx}^2)C_z^2\right]}, \quad \text{and} \tag{36}$$

$$d_{2(opt)}=\frac{\bar{Y}}{\bar{X}}\left\{\frac{1}{2}-d_{1(opt)}\left[\left(A+\frac{1}{2}\right)-\rho_{zx}\frac{C_z}{C_x}\right]\right\} \tag{37}$$

Substituting the optimum values of d_1 and d_2 in (33), the minimum mean square, up to the first order of approximation, is given by:

$$\begin{aligned} MSE_{min}(\hat{\mu}_{GR}) \approx \bar{Y}^2 & \left\{\left(1-\frac{1}{4}\lambda C_x^2\right)\right. \\ & \left. -\frac{\left[1+\lambda\left\{\left(B-\frac{1}{2}A-\frac{1}{4}\right)C_x^2+\left(\frac{1}{2}-A\right)\rho_{zx}C_zC_x\right\}\right]^2}{1+\lambda\left[\left(2B-A-\frac{1}{4}\right)C_x^2+(1-2A)\rho_{zx}C_zC_x+\left(1-\rho_{zx}^2\right)C_z^2\right]}\right\} \end{aligned} \tag{38}$$

For $\alpha=1$ the generalized mixture estimator is given by:

$$\hat{\mu}_{GR1}=\left[d_1\bar{z}\left(\frac{\bar{X}}{\bar{x}}\right)+d_2\left(\bar{X}-\bar{x}\right)\right]\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \tag{39}$$

The optimum values of d_1 and d_2 are given by:

$$d_{1GR1(opt)} = \frac{1+\left[\frac{7}{8}C_x^2 - \rho_{zx}C_zC_x\right]}{1+\lambda\left[2C_x^2 - 2\rho_{zx}C_zC_x + (1-\rho_{zx}^2)C_z^2\right]}$$

$$d_{2GR1(opt)} = \frac{\bar{Y}}{\bar{X}}\left[\frac{1}{2} - d_{1(opt)}\left(2-\rho_{zx}\frac{C_z}{C_x}\right)\right] \tag{40}$$

and the minimum mean square error is given by:

$$MSE_{min}(\hat{\mu}_{GMR1}) \approx \bar{Y}^2\left\{\left(1-\frac{1}{4}\lambda C_x^2\right) - \frac{\left[1+\lambda\left(\frac{7}{8}C_x^2 - \rho_{zx}C_zC_x\right)\right]^2}{1+\lambda\left[2C_x^2 - 2\rho_{zx}C_zC_x + \left(1-\rho_{zx}^2\right)C_z^2\right]}\right\}. \tag{41}$$

When $\alpha = 2$, the generalized mixture estimator is given by:

$$\hat{\mu}_{GR2} = \left[d_1\bar{z}\left(\frac{\bar{X}}{\bar{x}}\right)^2 + d_2\left(\bar{X}-\bar{x}\right)\right]\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right). \tag{42}$$

The optimum values of d_1 and d_2 are given by:

$$d_{1GR2(opt)} = \frac{1+\left[\frac{23}{8}C_x^2 - 2\rho_{zx}C_zC_x\right]}{1+\lambda\left[6C_x^2 - 4\rho_{zx}C_zC_x + (1-\rho_{zx}^2)C_z^2\right]}$$

$$d_{2GR2(opt)} = \frac{\bar{Y}}{\bar{X}}\left[\frac{1}{2} - d_{1(opt)}\left(3-\rho_{zx}\frac{C_z}{C_x}\right)\right]. \tag{43}$$

The minimum mean square error is given as:

$$MSE_{min}(\hat{\mu}_{GMR2}) \approx \bar{Y}^2\left\{\left(1-\frac{1}{4}\lambda C_x^2\right) - \frac{\left[1+\lambda\left(\frac{23}{8}C_x^2 - 2\rho_{zx}C_zC_x\right)\right]^2}{1+\lambda\left[-4\rho_{zx}C_zC_x + 6C_x^2 + \left(1-\rho_{zx}^2\right)C_z^2\right]}\right\}. \tag{44}$$

4 Efficiency Comparisons

In this section efficiency of the proposed estimator is compared with the some commonly used RRT estimators. Conditions under which the proposed estimator is more efficient are given below:

1. $MSE(\hat{\mu}_{GR}) < MSE(\mu_Y)$ if

$$\lambda C_z^2 - \left\{\left(1-\frac{1}{4}\lambda C_x^2\right)\right. \tag{45}$$

$$\left. - \frac{\left\{1+\lambda\left[\left(B-\frac{1}{2}A-\frac{1}{4}\right)C_x^2 + \left(\frac{1}{2}-A\right)\rho_{zx}C_zC_x\right]\right\}^2}{1+\lambda\left[\left(2B-A-\frac{1}{4}\right)C_x^2 + (1-2A)\rho_{zx}C_zC_x + \left(1-\rho_{zx}^2\right)C_y^2\right]}\right\} > 0$$

2. $MSE(\hat{\mu}_{GR}) < MSE(\hat{\mu}_R)$ if

$$\begin{aligned}&\lambda\,(C_x-\rho_{zx}C_z)^2+\lambda(1-\rho_{zx}^2)C_z^2 \\ &-\left\{\left(1-\frac{1}{4}\lambda C_x^2\right)-\frac{\left\{1+\lambda\left[\left(B-\frac{1}{2}A-\frac{1}{4}\right)C_x^2+\left(\frac{1}{2}-A\right)\rho_{zx}C_zC_x\right]\right\}^2}{1+\lambda\left[\left(2B-A-\frac{1}{4}\right)C_x^2+(1-2A)\rho_{zx}C_zC_x+\left(1-\rho_{zx}^2\right)C_z^2\right]}\right\}>0\end{aligned} \tag{46}$$

3. $MSE(\hat{\mu}_{GR}) < MSE(\hat{\mu}_{Reg})$ if

$$\begin{aligned}&\lambda\bar{Y}^2C_z^2\left(1-\rho_{zx}^2\right)-\left\{\left(1-\frac{1}{4}\lambda C_x^2\right)\right. \\ &\left.-\frac{\left\{1+\lambda\left[\left(B-\frac{1}{2}A-\frac{1}{4}\right)C_x^2+\left(\frac{1}{2}-A\right)\rho_{zx}C_zC_x\right]\right\}^2}{1+\lambda\left[\left(2B-A-\frac{1}{4}\right)C_x^2+(1-2A)\rho_{zx}C_zC_x+\left(1-\rho_{zx}^2\right)C_z^2\right]}\right\}>0\end{aligned} \tag{47}$$

4. $MSE(\hat{\mu}_{GR}) < MSE(\hat{\mu}_{ER})$ if

$$\begin{aligned}&\lambda\left(\frac{1}{2}C_x-\rho_{zx}C_z\right)^2+\lambda\left(1-\rho_{zx}^2\right)C_z^2 \\ &-\left\{\left(1-\frac{1}{4}\lambda C_x^2\right)-\frac{\left[1+\lambda\left\{\left(B-\frac{1}{2}A-\frac{1}{4}\right)C_x^2+\left(\frac{1}{2}-A\right)\rho_{zx}C_zC_x\right\}\right]^2}{1+\lambda\left[\left(2B-A+\frac{1}{4}\right)C_x^2+(1-2A)\rho_{zx}C_zC_x+\left(1-\rho_{zx}^2\right)C_z^2\right]}\right\}>0\end{aligned} \tag{48}$$

Numerical examples and simulation results show that these conditions are generally true, and hence the proposed estimator for $\alpha = 1$ and $\alpha = 2$ may be preferred over the existing estimators.

5 Numerical example

In this section, we compare the efficiency of proposed estimators with other existing RRT mean estimators considered in Section 2 using real data. The Population Statistics for the real data are given in Table 1. The scrambling variable S is taken to be a normal distribution with mean zero and standard deviation equal to two. The reported response is given by $Z = Y + S$. Table 2 gives Theoretical Percent Relative Efficiency (in bold) for various estimators based on the first order of approximation. The Theoretical Percent Relative Efficiency of the estimators as compared to the ordinary RRT sample mean are calculated from the following equation:

$$PRET(\hat{\mu}_i) = 100\times\frac{MSET(\hat{\mu}_y)}{MSET(\hat{\mu}_i)} \tag{49}$$

where $i = R, Reg, ER, GRR, GER, GR1$, and $GR2$.

Table 1: Population Statistics

Parameters	*Population 1*	*Population 2*	*Population 3*	*Population 4*
N	70	34	256	204
n	25	20	100	50
ρ_{yx}	0.7293	0.4491	0.887	0.71
ρ_{zx}	0.81079	0.44909	0.8867	0.7099
$\bar{Y}$	96.7	856.4118	56.47	966
$\bar{X}$	175.2671	208.8824	44.45	26441
S_x^2	19842.15	22650.18	3872.573	2061327175
S_y^2	3657.368	537544.3	6430.019	5711084
S_s^2	3.67395	3.67395	3.67395	3.67395
C_y	0.6254	0.8561	1.42	2.4739
C_x	0.8037	0.7205	1.40	1.7171
C_z	0.6257	0.8561	1.4204	2.4739
f	0.3571	0.5882	0.3906	0.2451

1. **Population 1 [Source: Singh and Chaudhary (1986), pp.108]**
2. **Population 2 [Source: Singh and Chaudhary(1986), pp. 177]**
3. **Population 3 [Source: Cochran (1977), pp. 196]**
4. **Population 4 [Source: Kadilar & Cingi (2005)]**

Table 2: The Theoretical Percent Relative Efficiency for the Mean Estimators

Estimators	PRET	*Population 1*	*Population 2*	*Population 3*	*Population 4*
$\hat{\mu}_Y$	**PRET**	**100**	**100**	**100**	**100**
$\hat{\mu}_R$	**PRET**	**176.3753**	**105.001**	**447.5094**	**201.5505**
$\hat{\mu}_{Reg}$	**PRET**	**291.8705**	**125.2645**	**467.9889**	**201.6534**
$\hat{\mu}_{ER}$	**PRET**	**269.5187**	**125.1390**	**271.1049**	**159.3275**
$\hat{\mu}_{GRR}$	**PRET**	**292.8943**	**126.7898**	**472.3173**	**211.3242**
$\hat{\mu}_{GER}$	**PRET**	**294.468**	**127.1320**	**478.3395**	**213.413**
$\hat{\mu}_{GR1}$	**PRET**	**303.6344**	**128.7935**	**485.3493**	**212.9479**
$\hat{\mu}_{GR2}$	**PRET**	**431.1358**	**137.8521**	**775.2617**	**242.964**

6 Conclusion

In this study, we proposed a generalized mixture estimator for the mean of a sensitive variable in simple random sampling without replacement by using information about a non sensitive auxiliary variable. The proposed generalized mixture estimator is a mixture of some of the commonly known RRT estimators. For the proposed estimators all the percent relative efficiencies are greater 100 indicating that all these estimators are better than the RRT ordinary mean estimator. We also note that both of the proposed generalized mixture estimators are more efficient than the other estimators considered here. Furthemore, the choice $\alpha = 2$ works better than $\alpha = 1$. We may note that at a theoretical level, one may be tempted to optimize α. Our goal though was to have a general family of estimators where many of the existing estimators become special cases of the proposed estimator with specific choice of α. For example, with $\alpha = 0$ our generalized mixture estimator II becomes combination of the regression and exponential ratio type estimators. For $\alpha = 1$, it involves the ratio term also. For $\alpha = -1$, it involves the product term.

References

Gupta S., Shabbir J. and Sehra S. (2010). Mean and sensitivity estimation in optional randomized response model. *Journal of Statistical Planning and Inference,* **140 (10)**, 2870-2874

Bahl,S. ,Tuteja, R.K. (1991). Ratio and product type exponential estimators. *Information and Optimization Sciences,* **12(1)**, 159-163

Gupta S., Shabbir J., Sousa R. and Corte-Real P. (2012). Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics - Theory and Methods,* **41 (13-14)**, 2394-2404

Koyuncu, N., Gupta S. and Sousa R. (2014). Exponential type estimators of the mean of a sensitive variable in the presence of non sensitive auxiliary information. *Communications in Statistics- Simulation and Computation*, **43**, 1583-1594

Sousa R., Shabbir J., Real P.C. and Gupta S. (2010). Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information. *Journal of Statistical Theory and Practice,* **4 (3)**, 495-507.

Sukhatme, P. , Sukhatme, B. and Sukhatme, S. (1970). *Sampling theory of surveys with applications.* Third edition, Iowa state university press