Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 28 (2025), No. 2, pp. 369–380
DOI : 10.47974/JIM-1290

Variance estimation using an optional randomized response technique

Badr Aloraini ‡
Department of Mathematics and Statistics
University of North Carolina
Greensboro 27412
U.S.A.

Sadia Khalil †
Department of Statistics
Lahore College for Women University
Lahore 54000
Pakistan

Muhammad Nouman Qureshi §
Department of Statistics
National College of Business Administration and Economics
Lahore 54660
Pakistan

Sat Gupta *
Department of Mathematics and Statistics
University of North Carolina
Greensboro 27412
U.S.A.

Abstract

Gupta et al. [2] used a Randomized Response Technique (RRT) to propose several estimators for finite population variance of a sensitive study variable using auxiliary information. In this paper, we modify those estimators by using the more efficient form of RRT, namely, Optional Randomized Response Technique (ORRT). Performance of the

‡ *E-mail:* **boalorai@uncg.edu**
† *E-mail:* **sadia_khalil@hotmail.com**
§ *E-mail:* **nqureshi633@gmail.com**
* *E-mail:* **sngupta@uncg.edu (Corresponding Author)**

proposed estimators is evaluated using a unified measure of respondent privacy and estimator efficiency. We also estimate the sensitivity level of the sensitive study variable.

Subject Classification: *62D05.*

Keywords: *Auxiliary information, Mean squared error, Optional randomized response technique, Respondent privacy, Variance estimation.*

1. Introduction

The problem of variance estimation is very well-studied in the context of survey sampling. Many authors including [6] have used auxiliary information to find improved estimators of the population variance when the study variable is directly observed. Others who have contributed in the area of variance estimation include Sanaullah et al., [9], Masood and Shabbir, [7], Asghar et al., [1] and Sanaullah et al., [10]. However, variance estimation has received much less attention when the study variable is sensitive in nature and direct observation is not possible. Gupta et al. [2] have estimated the population variance of a sensitive variable using the Randomized Response Technique (RRT) when the auxiliary information is available.

RRT was introduced by Warner [11]. The technique provides respondents enough privacy in order to reduce response bias. The main focus of this paper is on improving the [2] variance estimators by using the more efficient Optional Randomized Response Technique (ORRT) introduced by Gupta et al. [3]. The ORRT model is based on the assumption that a question may not be sensitive for all respondents. Thus, there is no need to force every respondent to give a scrambled response. Respondents can provide a scrambled response or a truthful response depending on whether they consider the question sensitive or not. With that motivation, we will re-examine Gupta et al. [2] in the context of the ORRT model.

Let Y be the study variable with mean $\overline{Y}$ and variance σ_y^2. Let X be the auxiliary variable with mean $\overline{X}$ and variance σ_x^2. Let W be the proportion of respondents who consider the question sensitive. It is assumed that Y and X are strongly positively correlated. Let $(\overline{y}, \overline{x})$ and (s_y^2, s_x^2) be the sample means and variances of the study and the auxiliary variable respectively. We use the additive RRT model $Z = Y + S$ studied by Pollock and Bek [8]. Since we will estimate both the population variance σ_y^2 and the sensitivity level W, we will use a split sample approach where the sample of size n is divided into two sub-samples of sizes n_1 and n_2 $(n_1$

$+ n_2 = n$). We use a scrambling variable S_i in the ith sub-sample $i = 1, 2$. We assume $E(S_i) = 0$ for $i = 1, 2$. We also assume that S_1, S_2 and Y are mutually uncorrelated.

2. The Proposed Estimators

Based on the ORRT model, the respondents are allowed to answer in two different ways depending on whether they feel the question is sensitive or not. Let

$$Z_i = \begin{cases} Y & \text{with probability } 1-W \\ Y+S_i & \text{with probability } W \end{cases} \quad i=1,2.$$

The mean and variance of the scrambled response Z_i are given by:

$$E(Z_i) = E(Y)(1-W) + E(Y+S_i)W = E(Y), \tag{1}$$

and

$$\sigma_{z_i}^2 = (1-W)\sigma_y^2 + W(\sigma_y^2 + \sigma_{s_i}^2). \tag{2}$$

Following Gupta et al. [5] and Zhang et al. [13] solving Equation (2) for the two unknown parameters σ_y^2 and W, we have

$$\sigma_y^2 = \left[\frac{\sigma_{z2}^2\sigma_{s1}^2 - \sigma_{z1}^2\sigma_{s2}^2}{\sigma_{s1}^2 - \sigma_{s2}^2}\right], \sigma_{s1}^2 \neq \sigma_{s2}^2$$

$$W = \left[\frac{\sigma_{z1}^2 - \sigma_{z2}^2}{\sigma_{s1}^2 - \sigma_{s2}^2}\right], \sigma_{s1}^2 \neq \sigma_{s2}^2$$

Estimating $\sigma_{z_i}^2$ and by its unbiased estimator $s_{z_i}^2$, we get the following estimators of σ_y^2 and W:

$$\hat{\sigma}_y^2 = \left[\frac{s_{z2}^2\sigma_{s1}^2 - s_{z1}^2\sigma_{s2}^2}{\sigma_{s1}^2 - \sigma_{s2}^2}\right], \sigma_{s1}^2 \neq \sigma_{s2}^2$$

$$\hat{W} = \left[\frac{s_{z1}^2 - s_{z2}^2}{\sigma_{s1}^2 - \sigma_{s2}^2}\right], \sigma_{s1}^2 \neq \sigma_{s2}^2$$

To obtain the Bias and MSE of the proposed variance estimator, we define the following error terms:

$$\delta_{z1}=\frac{s_{z1}^2-\sigma_{z1}^2}{\sigma_{z1}^2},\delta_{z2}=\frac{s_{z2}^2-\sigma_{z2}^2}{\sigma_{z2}^2},\delta_{x1}=\frac{s_{x1}^2-\sigma_{x}^2}{\sigma_{x}^2},\delta_{x2}=\frac{s_{x2}^2-\sigma_{x}^2}{\sigma_{x}^2}$$

where $E(\delta_{z1})=E(\delta_{z2})=E(\delta_{x1})=E(\delta_{x2})=E(\delta_{z1}\delta_{z2})=E(\delta_{z1}\delta_{x2})=E(\delta_{z2}\delta_{x1})=0$

$$E(\delta_{z1}^2)=\theta_1\lambda_{401}^*,E(\delta_{z2}^2)=\theta_2\lambda_{402}^*,E(\delta_{x1}^2)=\theta_1\lambda_{041}^*,E(\delta_{x2}^2)=\theta_2\lambda_{042}^*,$$

$$E(\delta_{z1}\delta_{x1})=\theta_1\lambda_{221}^*,E(\delta_{z2}\delta_{x2})=\theta_2\lambda_{222}^*,$$

where $\lambda_{rsi}=\dfrac{\mu_{rsi}}{\mu_{20i}^{\frac{r}{2}}\mu_{02i}^{\frac{s}{2}}},\mu_{rsi}=\dfrac{1}{N-1}\sum_{j=1}^{N}(Z_j-\overline{Z}_i)^r(X_j-\overline{X})^s,\theta_i=\dfrac{1}{n_i}-\dfrac{1}{N}$ and

$$\lambda_{rsi}^*=(\lambda_{rsi}-1)$$

With this background and notations, we introduce several variance estimators as discussed below:

2.1 *A Basic Variance Estimator Under the ORRT*

We have our first proposed estimator for the population variance given by

$$t_0(A)=\left[\frac{s_{z2}^2\sigma_{s1}^2-s_{z1}^2\sigma_{s2}^2}{\sigma_{s1}^2-\sigma_{s2}^2}\right],\sigma_{s1}^2\neq\sigma_{s2}^2 \tag{3a}$$

The estimator $t_0(A)$ can be expressed in terms of the error terms as shown below

$$t_0(A)=\frac{\sigma_{z2}^2(1+\delta_{z2})\sigma_{s1}^2-\sigma_{z1}^2(1+\delta_{z1})\sigma_{s2}^2}{\sigma_{s1}^2-\sigma_{s2}^2} \tag{3b}$$

Subtracting σ_y^2 on both sides, we obtain

$$(t_0(A)-\sigma_y^2)=\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2-\sigma_{s2}^2} \tag{3c}$$

To obtain the expression for the bias, up to a first-order approximation, we take the expectation on both sides of (3c), and we have

$$Bias(t_0(A))\approx E\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2-\sigma_{s2}^2}\right)=0 \tag{3d}$$

Subtracting σ_y^2, squaring and taking the expectation on both sides of (3b), the MSE, up to a first-order approximation, is obtained as

$$MSE(t_0(A)) = E\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2} - \sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2 - \sigma_{s2}^2}\right)^2$$

$$= \frac{1}{(\sigma_{s1}^2 - \sigma_{s2}^2)^2} E(\sigma_{z2}^4\sigma_{s1}^4\delta_{z2}^2 + \sigma_{z1}^4\sigma_{s2}^4\delta_{z1}^2 - 2\sigma_{z1}^2\sigma_{z2}^2\sigma_{s1}^2\sigma_{s2}^2\delta_{z1}\delta_{z2})$$

$$MSE(t_0(A)) \approx \frac{1}{(\sigma_{s1}^2 - \sigma_{s2}^2)^2}(\theta_2\sigma_{z2}^4\sigma_{s1}^4\lambda_{402}^* + \theta_1\sigma_{z1}^4\sigma_{s2}^4\lambda_{401}^*)$$

$$MSE(t_0(A)) \approx \frac{B_1}{ds^2}; \text{ where } B_1 = \theta_2\sigma_{z2}^4\sigma_{s1}^4\lambda_{402}^* + \theta_1\sigma_{z1}^4\sigma_{s2}^4\lambda_{401}^* \text{ and } ds = \sigma_{s1}^2 - \sigma_{s2}^2 \tag{3e}$$

2.2 *The Ratio Estimator Under the ORRT*

Isaki [6] proposed the traditional ratio estimator of population variance using auxiliary information as

$$t_1 = s_y^2\left(\frac{\sigma_x^2}{s_x^2}\right)$$

The ORRT version of t_1 is

$$t_1(A) = \left[\frac{s_{z2}^2\sigma_{s1}^2 - s_{z1}^2\sigma_{s2}^2}{\sigma_{s1}^2 - \sigma_{s2}^2}\right]\left(\frac{\sigma_x^2}{s_{x1}^2} + \frac{\sigma_x^2}{s_{x2}^2}\right)\left(\frac{1}{2}\right), \tag{4a}$$

where s_{x1}^2 and s_{x2}^2 respectively are sample variances of x in the two sub-samples.

$t_1(A)$ can be written as

$$t_1(A) = \left(\frac{1}{2}\right)\left[2\sigma_y^2 + 2\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2} - \sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2 - \sigma_{s2}^2}\right) - \sigma_y^2\delta_{x1}\right.$$
$$\left. - \delta_{x1}\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2} - \sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2 - \sigma_{s2}^2}\right) - \sigma_y^2\delta_{x2} - \delta_{x2}\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2} - \sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2 - \sigma_{s2}^2}\right)\right]$$

Subtracting σ_y^2 on both sides, we obtain,

$$(t_1(A)-\sigma_y^2)=\left[\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2-\sigma_{s2}^2}\right)-\frac{\sigma_y^2(\delta_{x1}+\delta_{x2})}{2}-\right.$$
$$\left.(\delta_{x1}+\delta_{x2})\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{2(\sigma_{s1}^2-\sigma_{s2}^2)}\right)\right]$$

By taking the expectation on both sides, the Bias of $t_1(A)$, up to a first-order approximation, is obtained as,

$$Bias(t_1(A))\approx\left[\frac{\theta_1\sigma_{z1}^2\sigma_{s2}^2\lambda_{221}^*-\theta_2\sigma_{z2}^2\sigma_{s1}^2\lambda_{222}^*}{2(\sigma_{s1}^2-\sigma_{s2}^2)}\right]$$

$$Bias(t_1(A))\approx-\frac{B_2}{2ds};\text{ where } B_2=\theta_2\sigma_{z2}^2\sigma_{s1}^2\lambda_{222}^*-\theta_1\sigma_{z1}^2\sigma_{s2}^2\lambda_{221}^* \tag{4b}$$

To obtain the MSE of $t_1(A)$, up to a first-order approximation, we write

$$(t_1(A)-\sigma_y^2)^2=\left[\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2-\sigma_{s2}^2}\right)-\frac{\sigma_y^2(\delta_{x1}+\delta_{x2})}{2}-\right.$$
$$\left.(\delta_{x1}+\delta_{x2})\left(\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{2(\sigma_{s1}^2-\sigma_{s2}^2)}\right)\right]^2$$

Expanding and simplifying, the MSE of $t_1(A)$ is obtained as

$$MSE(t_1(A))\approx\left[\frac{\theta_2\sigma_{z2}^4\sigma_{s1}^4\lambda_{402}^*+\theta_1\sigma_{z1}^4\sigma_{s2}^4\lambda_{401}^*}{(\sigma_{s1}^2-\sigma_{s2}^2)^2}+\frac{\sigma_y^4(\theta_1\lambda_{041}^*+\theta_2\lambda_{042}^*)}{4}\right.$$
$$\left.-\sigma_y^2\left(\frac{\theta_2\sigma_{z2}^2\sigma_{s1}^2\lambda_{222}^*-\theta_1\sigma_{z1}^2\sigma_{s2}^2\lambda_{221}^*}{\sigma_{s1}^2-\sigma_{s2}^2}\right)\right]$$

$$MSE(t_1(A))\approx\frac{1}{4ds^2}[4B_1+d_s\sigma_y^2(d_s\sigma_y^2B_3-4B_2)] \tag{4c}$$

where $B_3=\theta_1\lambda_{041}^*+\theta_2\lambda_{042}^*$

2.3 *A Generalized Variance Estimator Under the ORRT*

A generalized population variance estimator under the ORRT is given by

$$t_p(A)=\left[\frac{s_{z2}^2\sigma_{s1}^2-s_{z1}^2\sigma_{s2}^2}{\sigma_{s1}^2-\sigma_{s2}^2}+((\sigma_x^2-s_{x1}^2)+(\sigma_x^2-s_{x2}^2))\left(\frac{1}{2}\right)\right]$$

$$*\left(\frac{(\alpha\sigma_x^2+\beta)}{\omega[(\alpha s_{x1}^2+\beta)+(\alpha s_{x2}^2+\beta)]\left(\frac{1}{2}\right)+(1-\omega)(\alpha\sigma_x^2+\beta)}\right)^g \tag{5a}$$

In this estimator, α and β are suitably chosen constants associated with the auxiliary variable X. With $g=1$, we obtain various ratio estimators, and with $g=-1$ we obtain various product estimators. ω is an unknown constant, and its optimal value will be obtained later.

After some simplifications, we have

$$t_p(A)=\left[\sigma_y^2+\frac{\sigma_{z2}^2\sigma_{s1}^2\delta_{z2}-\sigma_{z1}^2\sigma_{s2}^2\delta_{z1}}{\sigma_{s1}^2-\sigma_{s2}^2}-\sigma_x^2\left(\frac{\delta_{x1}+\delta_{x2}}{2}\right)\right]$$

$$(1-g\omega\psi_i(\delta_{x1}+\delta_{x2}));\psi_i=\frac{\alpha\sigma_x^2}{2(\alpha\sigma_x^2+\beta)}$$

Using Taylor's approximation, up to a first-order approximation, we obtain the bias of the generalized estimator $t_p(A)$ as

$$Bias(t_p(A))\approx Q\left[\frac{\sigma_x^2}{2}(\theta_1\lambda_{041}^*+\theta_2\lambda_{042}^*)-\left(\frac{\theta_2\sigma_{z2}^2\sigma_{s1}^2\lambda_{222}^*-\theta_1\sigma_{z1}^2\sigma_{s2}^2\lambda_{221}^*}{\sigma_{s1}^2-\sigma_{s2}^2}\right)\right];$$

where $Q=g\omega\psi_i$

$$Bias(t_p(A))\approx\frac{Q}{ds}\left[\left(\frac{\sigma_x^2 d_s}{2}\right)B_3-B_2\right] \tag{5b}$$

The MSE is given by

$$MSE(t_p(A))\approx\left[\frac{\theta_2\sigma_{z4}^2\sigma_{s1}^4\lambda_{402}^*+\theta_1\sigma_{z1}^4\sigma_{s2}^4\lambda_{401}^*}{(\sigma_{s1}^2-\sigma_{s2}^2)^2}+\left(\frac{R_{xy}}{2}+Q\right)^2\sigma_y^4(\theta_1\lambda_{041}^*+\theta_2\lambda_{042}^*)\right.$$

$$\left.-2\sigma_y^2\left(\frac{\theta_2\sigma_{z2}^2\sigma_{s1}^2\lambda_{222}^*-\theta_1\sigma_{z1}^2\sigma_{s2}^2\lambda_{221}^*}{\sigma_{s1}^2-\sigma_{s2}^2}\right)\left(\frac{R_{xy}}{2}+Q\right)\right] \tag{5c}$$

where $R_{xy}=\frac{\sigma_x^2}{\sigma_y^2}$

Differentiating equation (5c) with respect to Q and setting the equation to zero, the optimum value of Q is given by

$$Q_{opt}=\frac{\theta_2\sigma_{z2}^2\sigma_{s1}^2\lambda_{222}^*-\theta_1\sigma_{z1}^2\sigma_{s2}^2\lambda_{221}^*}{\sigma_y^2(\sigma_{s1}^2-\sigma_{s2}^2)(\theta_1\lambda_{041}^*+\theta_2\lambda_{042}^*)}-\frac{R_{xy}}{2}$$

$$Q_{opt} = \frac{B_3}{\sigma_y^2 ds B_2} - \frac{R_{xy}}{2}$$

Putting Q_{opt} in (5c), the minimum MSE of $t_p(A)$ is

$$MSE(t_p(A)) \approx \left[\frac{\theta_2 \sigma_{z4}^2 \sigma_{s1}^4 \lambda_{402}^* + \theta_1 \sigma_{z1}^4 \sigma_{s2}^4 \lambda_{401}^*}{(\sigma_{s1}^2 - \sigma_{s2}^2)^2} - \frac{\theta_2 \sigma_{z2}^2 \sigma_{s1}^2 \lambda_{222}^* - \theta_1 \sigma_{z1}^2 \sigma_{s2}^2 \lambda_{221}^*}{(\sigma_{s1}^2 - \sigma_{s2}^2)^2 (\theta_1 \lambda_{041}^* + \theta_2 \lambda_{042}^*)} \right]$$

$$MSE(t_p(A)) \approx \frac{B_1}{ds^2} \left[1 - \frac{B_2^2}{B_1 B_3} \right]$$

$$MSE(t_p(A)) \approx \frac{B_1}{ds^2} [1 - \Gamma^2]; \text{ where } \Gamma = \frac{B_2}{\sqrt{B_1 B_3}} \tag{5d}$$

3. Simulation Study

We take a sample of size N = 1000 from a bivariate normal population and consider this as our finite population. The bivariate normal population has parameters given by

$$\mu_{y,x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \Sigma_{y,x} = \begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix}, \quad \rho_{yx} = 0.85$$

For the 1000 observations that have been made, the mean, variance and the correlation have the following values:

$$\mu_x = 1.98549, \mu_y = 1.97337, \sigma_x^2 = 1.99136, \sigma_y^2 = 5.98102, \rho_{yx} = 0.85086$$

These will be our finite population parameters. A simple random sample of size 500 without replacement is taken. We consider the equal split where $n_1 = n_2 = 250$. S_1 and S_2 are the scrambling variables with normal distribution with $E(S_1) = E(S_2) = 0$ and various values of Var(S_1) and Var(S_2). We use four different values (0.3, 0.5, 0.8, 1) for the sensitivity level W. In the generalized variance estimator, we choose $\alpha = 1$, $g = 1$ and $\beta = 0$. Our simulation results show that other choices of α and β have minimal impact.

The MSE is affected when varying the choices of Var(S_1) and Var (S_2). As the difference between the Var(S_1) and Var(S_2) increases, the MSE values tend to decrease. This makes sense because in the proposed estimators, the difference in variances (ds) shows up in the denominator, thereby reducing the MSE as the differences increase.

The Percent Relative Efficiency (PRE) of $t_i(A)$ is calculated as

$$PRE = \frac{MSE(t_0(A))}{MSE(t_i(A))} \times 100, \text{ where } i = 0,1 \text{ and } p$$

Since we are developing the proposed estimators based on randomized data, it is also important to consider the privacy level of the RRT model. A unified measure of estimator quality (δ) was introduced by Gupta et al. [4], and it is given by

$$\delta = \frac{\text{Estimator } MSE}{\Delta_{DP}}, \text{ where } \Delta_{DP} = \frac{\Sigma_{i=1}^{2}\sigma_{si}^{2}}{2}$$

is the average privacy level for the model $Z_i = Y + S_i$ with equal sub-sample sizes, as per Yan et al. [12]. We prefer smaller value of (δ) .

Table 1

Theoretical (bold) and empirical MSEs and PREs of the estimators with σ_y^2 = 5.98102 and ρ_{yx} = 0.85086.

σ_{S1}^2	σ_{S2}^2	W	$\hat{W}$	Estimator	Mean($\hat{\sigma}_y^2$)	MSE	PRE	δ
1.5	0.5	0.3	0.3128	$t_0(A)$	5.9957	**0.7504** 0.7535	**100** 100	0.7625
				$t_1(A)$	6.0326	**0.7386** 0.7401	**101.5893** 101.8118	0.7506
				$t_p(A)$	6.0003	**0.7032** 0.7128	**106.7106** 105.7090	0.7146
		0.5	0.4914	$t_0(A)$	6.0029	**0.7805** 0.7756	**100** 100	0.7931
				$t_1(A)$	6.0391	**0.7671** 0.7642	**101.7443** 101.4914	0.7795
				$t_p(A)$	6.0074	**0.7415** 0.7405	**105.2476** 104.7443	0.7535
		0.8	0.7959	$t_0(A)$	6.0072	**0.8134** 0.8227	**100** 100	0.8265
				$t_1(A)$	6.0427	**0.8046** 0.8008	**101.0887** 102.7421	0.8176
				$t_p(A)$	6.0112	**0.7795** 0.7844	**104.3392** 104.8790	0.7921
		1	0.9854	$t_0(A)$	6.0343	**0.8774** 0.8806	**100** 100	0.8916
				$t_1(A)$	6.0097	**0.8636** 0.8580	**101.5948** 102.6419	0.8776
				$t_p(A)$	6.0181	**0.8235** 0.8326	**106.5439** 105.7757	0.8368

Table 2

Theoretical (bold) and empirical MSEs and PREs of the estimators with σ_y^2 = 5.98102 and ρ_{yx} = 0.85086.

σ_{S1}^2	σ_{S2}^2	W	$\hat{W}$	Estimator	Mean($\hat{\sigma_y^2}$)	MSE	PRE	δ
3	1	0.3	0.2994	$t_0(A)$	6.0005	**0.8154** 0.8094	**100** 100	0.4122
				$t_1(A)$	6.0374	**0.7942** 0.7980	**102.6781** 101.4262	0.4014
				$t_p(A)$	6.0051	**0.7681** 0.7703	**106.1590** 105.0707	0.3883
		0.5	0.4841	$t_0(A)$	6.0109	**0.8722** 0.8680	**100** 100	0.4409
				$t_1(A)$	6.0469	**0.8478** 0.8537	**102.8882** 101.6722	0.4285
				$t_p(A)$	6.0152	**0.8213** 0.8309	**106.2062** 104.4656	0.4151
		0.8	0.8042	$t_0(A)$	6.0148	**0.9506** 0.9597	**100** 100	0.4805
				$t_1(A)$	6.0496	**0.9332** 0.9295	**101.8734** 103.2500	0.4717
				$t_p(A)$	6.0185	**0.9091** 0.9187	**104.5642** 104.4696	0.4596
		1	0.9953	$t_0(A)$	6.0053	**1.0546** 1.0411	**100** 100	0.5331
				$t_1(A)$	6.0101	**1.0170** 1.0114	**103.6969** 102.9433	0.5141
				$t_p(A)$	6.0388	**0.9969** 1.0008	**105.7888** 104.0309	0.5039

In Tables 1 and 2, the theoretical and empirical MSEs and PREs are given based on the ORRT proposed variance estimators. The results show that the generalized estimator is more efficient than the basic and the ratio estimators. By comparing corresponding cells in Tables 1 and 2, one can note that the MSEs increase as the variances of Si increase, which is on expected lines due to extra noise in the data. However, this loss in efficiency is off-set by the gain in privacy as shown by the δ-column. For example, the MSE of the generalized variance estimator is 0.7415 when W=0.5, Var(S_1)=1.5, Var(S_2)=0.5 in Table 1, and it is 0.8213 when W=0.5, Var(S_1)=3, Var(S_2)=1 in Table 2. However, δ decreases from 0.7535 to 0.4151. It is also observed that the MSEs of the proposed estimators increase asWincreases (going towards non-optional model). For instance, when Var(S_1)=3 and Var(S_2)=1, the MSE of the generalized variance estimator increased from 0.8213 to 0.9969 asWincreased from 0.5 to 1. This indicates that using ORRT models improves the efficiency of the estimators. Taking unified measure δ into account, we can note that the MSE increases as Var(S_1) and Var(S_2) increase. However, this increase is off-set by the corresponding gain in privacy level. For example, the MSE of the generalized variance estimator is 0.7795 when W=0.8, Var(S_1)=1.5, Var(S2)=0.5 in Table 1, and it is 0.9091 when W=0.8, Var(S1)=3, Var(S2)=1 in Table 2. However, δ decreases from 0.7921 to 0.4596.

4. Concluding Remarks

In conclusion, this study uses the ORRT model to estimate the population variance of a sensitive study variable. Several variance estimators are proposed using the additive ORRT model. Through a simulation study, it is seen that the generalized estimator is more efficient than the other estimators. Additionally, it is observed that the MSEs increases as the sensitivity levelWincreases, which indicates that the ORRT estimators are more efficient than those based on non-optional RRT. Furthermore, we examine the estimators' performance with respect to the unified measure of the estimator efficiency and the respondent privacy. We observe that an increase in the noise level results in an increase in the MSE . However, the unified measure decreases as the noise level increases. As such, it is crucial to examine both the MSE and respondent privacy.

References

[1] A. Asghar, A. Sanaullah, and M. Hanif, "A multivariate regression-cum-exponential estimator for population variance vector in two phase sampling," *Journal of King Saud University-Science*, vol. 30, no. 2, pp. 223-228 (2018).

[2] S. Gupta, B. Aloraini, M. Qureshi, and S. Khalil, "Variance estimation using randomized response technique," *REVSTAT*, vol. 18, no. 2, pp. 165-176 (2020).

[3] S. N. Gupta, B. C. Gupta, and S. Singh, "Estimation of sensitivity level of personal interview survey questions," *Journal of Statistical Planning and Inference*, vol. 2, pp. 239-247 (2002).

[4] S. Gupta, M. Samridhi, J. Shabbir, and S. Khalil, "A class of estimators for quantitative sensitive data," *Statistical Papers*, vol. 52, no. 3, pp. 633-650 (2018).

[5] S. Gupta, J. Shabbir, and S. Sehra, "Mean and sensitivity estimation in optional randomized response models," *Journal of Statistical Planning and Inference*, vol. 140, no. 10, pp. 2870-2874 (2010).

[6] C. T. Isaki, "Variance estimation using auxiliary information," *Journal of the American Statistical Association*, vol. 78, no. 381, pp. 117-123 (1983).

[7] S. Masood and J. Shabbir, "Generalized multi-phase regression-type estimators under the effect of measurement error to estimate the population variance," *Hacettepe Journal of Mathematics and Statistics*, vol. 45, no. 4, pp. 1297-1306 (2016).

[8] K. Pollock and Y. Bek, "A comparison of three randomized response models for quantitative data," *Journal of the American Statistical Association*, vol. 71, no. 356, pp. 884-886 (1976).

[9] A. Sanaullah, M. Hanif, and A. Asghar, "Generalized exponential estimators for population variance under two-phase sampling," *International Journal of Applied and Computational Mathematics*, vol. 2, no. 1, pp. 75-84 (2016).

[10] A. Sanaullah, I. Niaz, J. Shabbir, and I. Ehsan, "A class of hybrid type estimators for variance of a finite population in simple random sampling," *Communications in Statistics-Simulation and Computation*, pp. 1-11 (2020).

[11] 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).

[12] Z. Yan, J. Wang, and J. Lai, "An efficiency and protection degree-based comparison among the quantitative randomized response strategies," *Communications in Statistics-Theory and Methods*, vol. 38, pp. 404-408 (2009).

[13] Q. Zhang, S. Gupta, G. Kalucha, and S. Khalil, "Ratio estimation of the mean under RRT models," *Journal of Statistics and Management Systems*, vol. 22, no. 1, pp. 97-113 (2019).

Received April, 2021