International Journal of Computational and Theoretical Statistics
ISSN (2210-1519)
Int. J. Comp. Theo. Stat. 4, No. 2 (Nov-2017)

Additive Randomized Response Model with Known Sensitivity Level

Neeraj Tiwari[1] and Prachi Mehta[2]

[1,2]Department of Statistics, Kumaun University, S.S.J.Campus, Almora-263601, India

Received May 14, 2017, Revised July 6, 2017, Accepted August 1, 2017, Published November 1, 2017

Abstract : Randomized Response Technique (RRT) is an indirect survey method to collect sensitive information from the respondent without unveiling his/her true status. RRT has been first developed by Warner [12]. Tiwari & Mehta [11] proposed a model in which sensitivity level was taken as a known quantity under the assumption that if the respondent did not feel that the survey question was sensitive and can give an honest answer frankly, then it was not needed to conceal his/her identity and we can receive the true response from them without applying the randomization techniques. In this article, the idea of known sensitivity level proposed by Tiwari & Mehta [11] has been applied to One-Stage Optional Randomized Response Technique (ORRT) (Gupta et al. [5]), Two-Stage Optional Randomized Response Technique (ORRT) (Gupta et al. [6]) and Three-Stage Optional Randomized Response Technique (ORRT) (Mehta et al. [9]) based on quantitative data. The variances of the mean estimators of the models with sensitivity level considered as known, have been compared with the variances of the mean estimators of the models with unknown sensitivity level for all the above three models. It has been empirically established that the relative efficiency (RE) of the mean estimator for proposed model with known sensitivity level as compared to the RE of the mean estimator for the models with unknown sensitivity level was found to be greater than one for all the different cases considered by us. The proposed model is also simpler for use, as only one sample is required in it for collecting the sensitive information.

Keywords: Optional Randomized Response Model; Randomization Device; Sensitivity Level; Scrambling Variable.

1. INTRODUCTION

Exact information is highly applicable in any survey for estimating the parameter. However in the case of sensitive topics, it is difficult to get valid/exact information because most people underreport the sensitive behavior and/or hide the true information. For getting the reliable information on the sensitive issues and to maintain respondent's privacy, an efficient methodology known as Randomized Response Technique (RRT) was introduced by Warner [12]. RRT is an effective method which uses the concepts of probability theory to protect the respondent's privacy and has been practically used in many areas of research on sensitive issues.

Warner [12] used a randomization device in his model, by which each respondent chooses one of the two questions-

1) "Do you belong to A?"

and 2) "Do you not belong to A?"

with probabilities p and (1-p) respectively, where A denotes the sensitive group. After getting the 'yes' or 'no' answers, researcher estimates the value of required population proportion using the estimator suggested by Warner [12].

In the model suggested by Warner [12], both the questions were related to the same sensitive group. However it is required that the two questions should be unrelated to protect the respondent's privacy. Greenberg et al. [2] introduced an unrelated question technique. In this technique, the second question of the model given by Warner [12] was replaced by an unrelated question such as- 'Did you watch the 6:00 pm news yesterday?'

Mangat & Singh [8] developed 'Two-Stage Randomized Response Model' for getting more efficient estimator and respondent's cooperation. In this model two different randomization devices were used in two stages.

E-mail address: kumarn_amo@yahoo.com, prachimehta178@gmail.com

An Optional Randomized Response (ORR) Model was first introduced by Gupta [3]. In this optional model, the respondent gives a scrambled answer if he/she seems the survey question is sensitive and gives a true response if he/she feels the survey question is non-sensitive. So there is a choice or an option for the respondents to give their answer.

Sihm & Gupta [10] suggested Two-Stage Binary Optional Randomized Response Model which was based on ORR model given by Gupta [3]. Their method gives better results than Optional Randomized Response Technique (ORRT) given by Gupta [3].

Recently Tiwari & Mehta [11] proposed an improved methodology for RRT, in which the sensitivity level (W) was considered to be known. This makes the procedure simpler and more efficient compared to the procedures of Gupta [3] and Sihm & Gupta [10].

RRT is applicable for both the qualitative as well as quantitative data. In this discussion we restrict ourselves to the quantitative models only. Many researchers have worked on different aspects of quantitative RRT models.

Eichhorn & Hayre [1] proposed a multiplicative randomized response method for obtaining responses to sensitive questions when the answers were quantitative. In this method, the respondent multiply his/her answer by a random number from a known distribution and provide the product to the interviewer. The interviewer does not know the value of the random number and receives a scrambled response. This method ensures that the exact response of the respondent is not revealed to the interviewer.

To provide an option to the respondents who may not feel any need to scramble the response depending upon the nature of the question, Gupta et al. [4] suggested an Optional Randomized Response Technique (ORRT), in which a respondent chooses one of the two options-

1) 'Give the correct answer X'

Or 2) 'Give the scrambled answer SX'

Here S is the scrambling variable.

A problem with the multiplicative scrambling model was that it compromises anonymity because a non-zero reported response would mean that X could not have been zero, implying that the respondent has at least some level of sensitive behavior. One-stage additive ORRT model was introduced by Gupta et al. [5], in which additive scrambling was used in place of multiplicative scrambling.

For getting more efficient estimator and respondent's cooperation, two-stage additive ORRT model was proposed by Gupta et al. [6]. In this method, a known proportion of respondents (T) were asked to provide a true response of the sensitive question and rest of the respondents (1-T) give additive scrambling response.

A drawback of two-stage model was that a greater value of T might be required if the survey question was highly sensitive, which decreases the respondent's cooperation in the survey. To overcome this difficulty, three-stage ORRT was introduced by Mehta et al. [9]. Under this model, a known proportion (T) of the respondents were asked to tell the truth, again a known proportion (F) were asked to provide an additive scrambled response and rest of the respondents (1-T-F) follow usual ORRT with additive scrambling.

Huang [7] proposed an optional randomized response model using a linear combination scrambling which is a generalization of the multiplicative scrambling of Eichhorn and Hayre [1] and the additive scrambling of Gupta et al. ([5], [6]). The linear combination scrambling provided a higher level of anonymity as compared to additive scrambling but did not give a more efficient model in terms of estimator precision.

As discussed earlier, Tiwari & Mehta [11] considered the situation where a particular question may not be sensitive for some of the respondents, who provided a true response of that question. This provides us the sensitivity level of the question (W) as a known quantity. Taking inspiration from Tiwari & Mehta [11], this article proposes one-stage, two-stage and three-stage additive models for quantitative data that considers sensitivity level (W) as a known quantity. The proposed model considerably simplifies the collection of sensitive data due to the fact that only one sample is required in it. The expressions for the mean and variance of the estimate of the mean of the true response have been derived. It has been empirically established that the variance of the mean estimator of the proposed model is less than the variance of the mean estimator of the additive ORRT model considered by the earlier authors.

In Section 2, one-stage, two-stage and three-stage additive ORRT models proposed by earlier authors have been discussed. The proposed one-stage, two-stage and three-stage additive RRT models with known sensitivity level have been discussed in Section 3. The proposed models have been empirically compared with the existing models in Section 4. The findings of the paper have been discussed in Section 5.

2. ADDITIVE OPTIONAL RANDOMIZED RESPONSE TECHNIQUE (ORRT) MODELS

In this section, we have given brief description about one-stage additive ORRT model (Gupta et al. [5]), two-stage additive ORRT model (Gupta et al. [6]) and three-stage additive ORRT model (Mehta et al. [9]).

2.1 One-stage additive ORRT model

In this model, respondents give a true response X if they feel the question is non-sensitive and give a scrambled response (X+S) if they feel the question is sensitive, while maintaining anonymity. The scrambling variable S and the true response variable X are assumed to be mutually independent. Let the unknown mean and variance of X be μ_x and σ_x^2, respectively. Let w be the sensitivity level of the survey question. Since there are two parameters (μ_x and W) to estimate, Gupta et al. [5] used the split sample approach. The sample of size n was split into two subsamples of sizes n_1 and n_2, where each subsample used a different scrambling device. For two subsamples, let the known mean and variance of S_i (i=1,2) be θ_i and $\sigma_{s_i}^2$ respectively. Let Z_i be the reported response in the i^{th} subsample. Thus we have

$$Z_i = \begin{cases} X \\ X + S_i \end{cases} \qquad \text{i =1, 2}$$

with probability (1-W) and W respectively. For i=1, 2, the mean and the variance of Z_i are

$$E(Z_i) = \mu_x + \theta_i W \tag{1}$$

And
$$\sigma_{z_i}^2 = \sigma_x^2 + \sigma_{s_i}^2 W + \theta_i^2 W(1-W) \tag{2}$$

By solving the equation (1) for i=1, 2, we get

$$\mu_x = \frac{\theta_2 E(Z_1) - \theta_1 E(Z_2)}{\theta_2 - \theta_1}, \qquad \theta_1 \neq \theta_2$$

and
$$W = \frac{E(Z_2) - E(Z_1)}{\theta_2 - \theta_1}, \qquad \theta_1 \neq \theta_2 \tag{3}$$

The unbiased estimators of μ_x and W, can be obtained by estimating $E(Z_i)$ by $\overline{Z}_i$ (i=1,2). Thus the unbiased estimators $\hat{\mu}_x$ and $\hat{W}$ are

$$\hat{\mu}_x = \frac{\theta_2 \overline{Z}_1 - \theta_1 \overline{Z}_2}{\theta_2 - \theta_1}, \qquad \theta_1 \neq \theta_2$$

and
$$\hat{W} = \frac{\overline{Z}_2 - \overline{Z}_1}{\theta_2 - \theta_1}, \qquad \theta_1 \neq \theta_2 \tag{4}$$

For $\theta_1 \neq \theta_2$, $\hat{\mu}_x$ ~ AN(μ_x, V_1) , and $\hat{W}$ ~ AN(W, V_2), where

$$V_1 = \frac{1}{(\theta_2 - \theta_1)^2}\left[\theta_2^2 \frac{\sigma_{z_1}^2}{n_1} + \theta_1^2 \frac{\sigma_{z_2}^2}{n_2}\right], \qquad \theta_1 \neq \theta_2$$

and $$V_2 = \frac{1}{(\theta_2-\theta_1)^2}\left[\frac{\sigma_{z_1}^2}{n_1}+\frac{\sigma_{z_2}^2}{n_2}\right], \qquad \theta_1 \neq \theta_2 \tag{5}$$

2.2 Two-stage additive ORRT model

In the two-stage model of Gupta et al. [6], a known proportion of respondents (T) were asked to provide a true response X to the sensitive question in the first stage while maintaining anonymity. Rest of the respondent (1-T) follow the usual additive optional randomized response technique (ORRT) in the second stage. Let W be the sensitivity level of the survey question. There are two parameters μ_x and W, which are estimated by using two samples. Assume there are two independent samples of sizes n_1 and n_2 respectively (n_1+n_2=n). Let the unknown mean and variance of X be μ_x and σ_x^2, respectively. Let S_i be the scrambling variable used to scramble the responses in the i^{th} subsample (i=1, 2). The variables X, S_1, S_2 are mutually independent. Let the known mean and variance of S_i (i=1,2) be θ_i and $\sigma_{s_i}^2$ respectively. Let Z_i be the reported response in the i^{th} subsample. Thus we have

$$Z_i = \begin{cases} X \\ X+S_i \end{cases} \qquad i=1, 2$$

with probability {T+(1-W)(1-T)} and W(1-T) respectively.

The mean and the variance of Z_i are respectively given by

$$E(Z_i)=\mu_x+\theta_i W(1-T) \tag{6}$$

and $$\sigma_{z_i}^2 = \sigma_x^2+\sigma_{s_i}^2 W(1-T)+\theta_i^2[W(1-T)][1-\{W(1-T)]] \tag{7}$$

From equation (10), we have

$$\mu_x = \frac{\theta_2 E(Z_1)-\theta_1 E(Z_2)}{\theta_2-\theta_1}, \qquad \theta_1 \neq \theta_2$$

and $$W = \frac{E(Z_2)-E(Z_1)}{(\theta_2-\theta_1)(1-T)}, \qquad T\neq 1,\ \theta_1 \neq \theta_2 \tag{8}$$

The unbiased estimators of μ_x and W, can be obtained by estimating $E(Z_i)$ by $\overline{Z}_i$ (i=1,2). Thus the unbiased estimators $\hat{\mu}_x$ and $\hat{W}$ are

$$\hat{\mu}_x = \frac{\theta_2 \overline{Z}_1-\theta_1 \overline{Z}_2}{\theta_2-\theta_1}, \qquad \theta_1 \neq \theta_2$$

and $$\hat{W} = \frac{\overline{Z}_2-\overline{Z}_1}{(\theta_2-\theta_1)(1-T)}, \qquad T\neq 1,\ \theta_1 \neq \theta_2 \tag{9}$$

The estimators $\hat{\mu}_x$ and $\hat{W}$ are unbiased with variances given by

$$V(\hat{\mu}_x) = \frac{1}{(\theta_2-\theta_1)^2}\left[\theta_2^2\frac{\sigma_{z_1}^2}{n_1}+\theta_1^2\frac{\sigma_{z_2}^2}{n_2}\right], \qquad \theta_1 \neq \theta_2$$

and $$V(\hat{W}) = \frac{1}{(\theta_2-\theta_1)^2(1-T)^2}\left[\frac{\sigma_{z_1}^2}{n_1}+\frac{\sigma_{z_2}^2}{n_2}\right], \qquad \theta_1 \neq \theta_2,\ T\neq 1 \tag{10}$$

2.3 Three-stage additive ORRT model

In this model, the sample of size n is again split into two subsamples of sizes n_1 and n_2. In each sample, a fixed predetermined proportion of respondents (T) were asked to provide a true response X and a fixed predetermined proportion (F) of respondents was instructed to scramble their response additively. The remaining proportion (1-T-F) of respondents has an option to scramble their response additively, if they consider the survey question is sensitive or they can report their true response X. Let W be the sensitivity level of the survey question. Let the unknown mean and variance of X be μ_x and σ_x^2, respectively. Let S_i be the scrambling variable used to scramble the responses in the i^{th} subsample (i=1, 2). The variables X, S_1, S_2 are mutually independent. Let the known mean and variance of S_i (i=1,2) be θ_i and $\sigma_{s_i}^2$ respectively. Let Z_i be the reported response in the i^{th} subsample. Thus we have

$$Z_i = \begin{cases} X \\ X + S_i \end{cases} \qquad i = 1, 2$$

with probability {T+(1-T-F)(1-W)} and F+(1-T-F)W respectively.

The mean and the variance of Z_i are respectively given by

$$E(Z_i) = \mu_x + \theta_i [F+(1-T-F)W] \qquad (11)$$

and

$$\sigma_{z_i}^2 = \sigma_x^2 + \sigma_{s_i}^2 [F + (1-T-F)W] + \theta_i^2 [F + (1-T-F)W][1 - \{F + (1-T-F)W\}]$$

(12)

From equation (11), we have

$$\mu_x = \frac{\theta_2 E(Z_1) - \theta_1 E(Z_2)}{\theta_2 - \theta_1}, \qquad \theta_1 \neq \theta_2$$

and

$$W = \frac{1}{(1-T-F)} \left(\frac{E(Z_2) - E(Z_1)}{\theta_2 - \theta_1} - F \right), \qquad T+F \neq 1, \ \theta_1 \neq \theta_2 \qquad (13)$$

The unbiased estimators of μ_x and W, can be obtained by estimating $E(Z_i)$ by $\bar{Z}_i$ (i=1,2). Thus the unbiased estimators $\hat{\mu}_x$ and $\hat{W}$ are

$$\hat{\mu}_x = \frac{\theta_2 \bar{Z}_1 - \theta_1 \bar{Z}_2}{\theta_2 - \theta_1}, \qquad \theta_1 \neq \theta_2$$

and

$$\hat{W} = \frac{1}{(1-T-F)} \left(\frac{\bar{Z}_2 - \bar{Z}_1}{\theta_2 - \theta_1} - F \right), \qquad T+F \neq 1, \ \theta_1 \neq \theta_2 \qquad (14)$$

The estimators $\hat{\mu}_x$ and $\hat{W}$ are unbiased with variances given by

$$V(\hat{\mu}_x) = \frac{1}{(\theta_2 - \theta_1)^2} \left[\theta_2^2 \frac{\sigma_{z_1}^2}{n_1} + \theta_1^2 \frac{\sigma_{z_2}^2}{n_2} \right], \qquad \theta_1 \neq \theta_2$$

and

$$V(\hat{W}) = \frac{1}{(\theta_2 - \theta_1)^2 (1-T-F)^2} \left[\frac{\sigma_{z_1}^2}{n_1} + \frac{\sigma_{z_2}^2}{n_2} \right], \qquad \theta_1 \neq \theta_2, \ T+F \neq 1 \qquad (15)$$

3. THE PROPOSED ADDITIVE RANDOMIZED RESPONSE TECHNIQUE (RRT) MODELS WITH KNOWN SENSITIVITY LEVEL (W)

The proposed models are based on the fact that the respondent can give a direct response if he/she doesn't feel the particular question is sensitive. By this way interviewer can separate sensitive and non-sensitive group and get the value of the sensitivity level (W). Taking W as a known quantity only one sample is required for collecting the sensitive information, which makes the collection of sensitive information quite easy for the interviewer, with the protection of those respondent's privacy, who consider that the particular question is sensitive for them.

3.1 One-stage additive randomized response technique (RRT) model with known sensitivity level

In this model, respondents give a true response X if they feel the question is non-sensitive and give additive scrambled response (X+S) if they feel the question is sensitive, while maintaining anonymity. The scrambling variable S and the true response variable X are assumed to be mutually independent. Let the unknown mean and variance of X be μ_x and σ_x^2, respectively. Let the known mean and the known variance for S be θ and σ_s^2 respectively. W is the sensitivity level of the survey question. If Z be the reported response, then we have

$$Z = \begin{cases} X \\ X+S \end{cases}$$

with probability (1-W) and W respectively. The mean and the variance of Z are

$$E(Z) = \mu_x + \theta W \tag{16}$$

and

$$\sigma_z^2 = \sigma_x^2 + \sigma_s^2 W + \theta^2 W(1-W) \tag{17}$$

From equation (16),

$$\mu_x = E(Z) - \theta W \tag{18}$$

The unbiased estimator for μ_x is obtained by estimating E(Z) by $\overline{Z}$, thus

$$\hat{\mu}_x = \overline{Z} - \theta W$$

$$\hat{\mu}_x \sim AN(\mu_x, V_1),$$

where $V_1 = \frac{\sigma_z^2}{n}$ and σ_z^2 is given in the equation (17).

3.2 Two-stage additive randomized response technique (RRT) model with known sensitivity level (W)

In the two-stage model, a known proportion of respondents (T) are asked to provide a true response X to the sensitive question in the first stage while maintaining anonymity. In the second stage, rests of the respondents (1-T) have an option to give an additive scrambled response if they feel the survey question is sensitive, else provide a true response X. Let the unknown mean and variance of X be μ_x and σ_x^2, respectively. Let the known mean and the known variance for S be θ and σ_s^2 respectively. W is the sensitivity level of the survey question. If Z be the reported response, then we have

$$Z = \begin{cases} X \\ X+S \end{cases}$$

with probability {T+(1-W)(1-T)} and W(1-T) respectively.

The mean and the variance of Z are respectively given by

$$E(Z) = \mu_x + \theta\, W(1-T) \tag{19}$$

and

$$\sigma_z^2 = \sigma_x^2 + \sigma_s^2 W(1-T) + \theta^2 [W(1-T)][1-\{W(1-T)\}] \tag{20}$$

From equation (19),

$$\mu_x = E(Z)-\theta\, W(1-T) \tag{21}$$

The unbiased estimator for μ_x is obtained by estimating E(Z) by $\overline{Z}$, thus

$$\hat{\mu}_x = \overline{Z} - \theta\, W(1-T)$$

$$\hat{\mu}_x \sim AN(\mu_x, V_1),$$

where $V_1 = \frac{\sigma_z^2}{n}$ and σ_z^2 is given in the equation (20).

3.3 Three-stage additive randomized response technique (RRT) model with known sensitivity level (W)

In three-stage additive RRT model, a fixed predetermined proportion of respondents (T) are asked to provide a true response X and a fixed predetermined proportion (F) of the respondents are instructed to scramble their response additively. The remaining proportion (1-T-F) of respondents have an option to scramble their response additively, if they consider the survey question is sensitive or they can report their true response X. Let the unknown mean and variance of X be μ_x and σ_x^2, respectively. Let the known mean and the known variance for S be θ and σ_s^2 respectively. W is the sensitivity level of the survey question. If Z be the reported response, then we have

$$Z = \begin{cases} X \\ X + S \end{cases}$$

with probability {T+(1-T-F)(1-W)} and F+(1-T-F)W respectively.
The mean and the variance of Z are respectively given by

$$E(Z)= \mu_x + \theta\, [F+ (1-T-F)W] \tag{22}$$

and

$$\sigma_z^2 = \sigma_x^2 + \sigma_s^2 [F + (1-T-F)W] + \theta^2 [F + (1-T-F)W][1-\{F + (1-T-F)W\}]$$

(23)

From equation (22),

$$\mu_x = E(Z)-\theta\, [F+ (1-T-F)W] \tag{24}$$

The unbiased estimator for μ_x is obtained by estimating E(Z) by $\overline{Z}$, thus

$$\hat{\mu}_x = \overline{Z} - \theta\, [F+ (1-T-F)W]$$

$$\hat{\mu}_x \sim AN(\mu_x, V_1),$$

where $V_1 = \frac{\sigma_z^2}{n}$ and σ_z^2 is given in the equation (23).

4. EMPIRICAL COMPARISON

In this section, we have computed the relative efficiency of the proposed estimators with respect to the estimators suggested by Gupta et al. [5], Gupta et al. [6] and Mehta et al. [9], by using the formula of relative efficiency.

4.1 Empirical comparison for different one-stage additive randomized response technique (RRT) models:

We have obtained the values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Gupta et al. [5] by the formula

$$RE = \frac{V(\hat{\mu}_x)_1}{Var(\hat{\mu}_x)_{n_1}}$$

Here $V(\hat{\mu}_x)_1$ is the variance of the mean estimator for one stage optional randomized response technique (ORRT) model given by Gupta et al. [5] and $V(\hat{\mu}_x)_{n_1}$ is the variance of the mean estimator for the proposed one-stage RRT model.

Let us consider that $n_1 = n_2$ =500, n=1000, X~ poisson(7), S_1 ~ poisson(3), S_2 ~ poisson(4) and S~ poisson(7). The values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Gupta et al. [5] for various values of W are given in Table 1.

TABLE 1. RELATIVE EFFICIENCY (RE) OF THE PROPOSED ONE STAGE RRT MODEL RELATIVE TO ONE STAGE ORRT MODEL GIVEN BY GUPTA ET AL. [5]

W=0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
50	34.5950	29.3703	26.8762	25.7129	25.3508	25.7205	26.5135	28.2647	31.2429	37

Let us consider that $n_1 = n_2$ =500, n=1000, X~ poisson(5), S_1 ~ χ^2(4,8), S_2 ~ χ^2(5,10) and S~ χ^2(9,18). The values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Gupta et al. [5] for various values of W are given in table 2.

TABLE 2. RELATIVE EFFICIENCY (RE) OF THE PROPOSED ONE STAGE RRT MODEL RELATIVE TO ONE STAGE ORRT MODEL GIVEN BY GUPTA ET AL. [5]

W=0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
82	44.3971	37.5	35.1094	34.2405	34.1107	34.8295	36.1271	38.3333	42.1754	49.1304

From Table 1 and Table 2, it is clear that the values of relative efficiencies (RE) are greater than 1, for different values of W and different distributions for S_1, S_2 and S. Thus it may be concluded that the proposed one-stage additive RRT model performs better than the one-stage additive ORRT model proposed by Gupta et al. [5].

4.2 Empirical comparison for different two-stage RRT models:

We have obtained the values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Gupta et al. [6] by the formula

$$RE = \frac{V(\hat{\mu}_x)_2}{V(\hat{\mu}_x)_{n_2}}$$

Here $V(\hat{\mu}_x)_2$ is the variance of the mean estimator for two stage ORRT model given by Gupta et al. [6] and $V(\hat{\mu}_x)_{n_2}$ is the variance of the mean estimator for the proposed two-stage RRT model.

Let us consider that $n_1 = n_2$ =500, n=1000, X~ poisson(7), S_1 ~ poisson(3), S_2 ~ poisson(4) and S~ poisson(7). The values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Gupta et al. [6] for various values of W are given in Table 3.

TABLE 3. RELATIVE EFFICIENCY (RE) OF THE PROPOSED TWO STAGE RRT MODEL RELATIVE TO TWO STAGE ORRT MODEL GIVEN BY GUPTA ET AL. [6]

W	T=0	T=0.1	T=0.3	T=0.5	T=0.7	T=0.9
0	50	50	50	50	50	50
0.1	34.5950	35.5431	37.3084	39.76288	43.2325	47.0263
0.3	26.87628	27.5081	28.9638	31.37062	35.5431	43.2325
0.5	25.3508	24.7043	26.2038	27.9329	31.3706	39.7731
0.7	26.5135	25.8815	25.3876	26.20388	28.9638	37.3084
0.9	31.2429	28.4504	25.8815	25.4798	25.9027	35.5431

Let us consider that $n_1 = n_2$ =500, n=1000, X~ poisson(5), $S_1 \sim \chi^2(4,8)$, $S_2 \sim \chi^2(5,10)$ and S~ $\chi^2(9,18)$. The values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Gupta et al. [6] for various values of W are given in Table 4.

TABLE 4. RELATIVE EFFICIENCY (RE) OF THE PROPOSED TWO STAGE RRT MODEL RELATIVE TO TWO STAGE ORRT MODEL GIVEN BY GUPTA ET AL. [6]

W	T=0	T=0.1	T=0.3	T=0.5	T=0.7	T=0.9
0	82	82	82	82	82	82
0.1	44.3971	45.5488	49.0956	53.8144	61.3076	73.3898
0.3	35.0966	35.6511	37.2342	40.1111	45.7207	60.5316
0.5	34.1107	34.1389	34.5454	36.0323	40.1111	53.8144
0.7	36.1271	35.1306	34.0938	34.5454	37.2342	49.0956
0.9	42.1754	38.7312	35.1306	34.1389	53.7894	45.5488

From Table 3 and Table 4, it is clear that the values of relative efficiencies (RE) are greater than 1, for various combinations of W and T and different distributions for S_1, S_2 and S. Thus it may be concluded that the proposed two-stage additive RRT model performs better than the two-stage additive ORRT model proposed by Gupta et al. [6]. It is also clear from Table 3 and Table 4 that in two-stage model, for smaller values of W, relative efficiency decreases as T increases and for larger values of W, relative efficiency initially increases as T increases and then start decreasing with increase in T. Thus for low sensitive questions, we can take any value of T and when the questions are highly sensitive; greater value of T are needed.

4.3 Empirical comparison for different three-stage RRT models:

We have obtained the values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Mehta et al. [9] by the formula

$$RE = \frac{V(\hat{\mu}_x)_3}{V(\hat{\mu}_x)_{n_3}}$$

Here $V(\hat{\mu}_x)_3$ is the variance of the mean estimator for three-stage ORRT model given by Mehta et al. [9] and $V(\hat{\mu}_x)_{n_3}$ is the variance of the mean estimator for the proposed three-stage RRT model.

Let us consider $n_1 = n_2$ =500, n=1000, X~ poisson(7), S_1 ~ poisson(3), S_2 ~ poisson(4) and S~ poisson(7). The values of the relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Mehta et al. [9] for various values of F and T are given in Table 5.

TABLE 5. RELATIVE EFFICIENCY (RE) OF THE PROPOSED THREE STAGE RRT MODEL RELATIVE TO THREE STAGE ORRT MODEL GIVEN BY MEHTA ET AL. [9]

F	T=0	T=0.1	T=0.3	T=0.5	T=0.7	T=0.9
0	28.2647	26.8082	25.5895	25.7129	28.1477	36.2321
0.1	28.7839	27.0879	25.6637	25.6118	27.6589	_
0.3	29.8306	27.875	25.7598	25.4598	_	_
0.5	31.2429	35.1411	26.1106	_	_	_
0.7	33.1533	29.8306	_	_	_	_
0.9	35.5337	_	_	_	_	_

Let us consider that $n_1 = n_2$ =500, n=1000, X~ poisson(5), S_1 ~ χ^2(4,8), S_2 ~ χ^2(5,10) and S~ χ^2(9,18). The values of relative efficiencies (RE) of the proposed estimator with respect to the estimator suggested by Mehta et al. [9] for various values of F and T are given in Table 6.

TABLE 6. RELATIVE EFFICIENCY (RE) OF THE PROPOSED THREE STAGE RRT MODEL RELATIVE TO THREE STAGE ORRT MODEL GIVEN BY MEHTA ET AL. [9]

F	T=0	T=0.1	T=0.3	T=0.5	T=0.7	T=0.9
0	35.1094	35.6472	37.2342	40.1111	45.5488	60.5316
0.1	34.4032	34.6006	35.3863	36.9692	39.5828	_
0.3	34.2151	34.1686	34.1238	34.4784	_	_
0.5	35.3644	35.0369	34.4971	_	_	_
0.7	38.1656	37.2835	_	_	_	_
0.9	43.8407	_	_	_	_	_

From Table 5 and Table 6, it is seen that for W=0.8 (i.e. the survey question is highly sensitive) and for W=0.3 (i.e. the survey question is less sensitive), the values of relative efficiencies (RE) are greater than 1, for various combinations of F and T and different distributions for S_1, S_2 and S. This shows that the proposed model appears to be more efficient that the existing models for highly sensitive as well as less sensitive questions. The values of the relative efficiencies of the mean estimator for the proposed three-stage additive RRT model relative to the mean estimator for additive ORRT model given by Mehta et al. [9] were also calculated for different values of W lying between 0 and 1. It was found that the relative efficiencies of the mean estimator for the proposed three-stage additive RRT model relative to the mean estimator for additive ORRT model given by Mehta et al. [9] are always greater than 1 for all the values of W. The details of these calculations are omitted for brevity.

5. DISCUSSION

In this article, an attempt has been made to improve the one-stage, two-stage and three-stage additive ORRT models based on quantitative data proposed by Gupta et al. [5], Gupta et al. [6] and Mehta et al. [9] respectively. Taking the idea of known sensitivity level proposed by Tiwari and Mehta [11], we have proposed three improved additive RRT models. An empirical comparison of the proposed models with the existing additive ORRT models has been carried out using different values of sensitivity level and different combinations of distributions for scrambling variable. In all the cases, the proposed models appear to perform better than the existing models.

By taking the sensitivity level as a known quantity complexity of data collection and computation is minimized as only one sample is required in the proposed models as compared to two samples required for the existing models for collecting the sensitive information.

ACKNOWLEDGMENTS

The authors are grateful to the editor and referees for his/her constructive comments, which have led to considerable improvement in presentation of this manuscript.

REFERENCES

[1] B. H. Eichhorn, and L. S. Hayre, "Scrambled randomized response methods for obtaining sensitive quantitative data," Journal of Statistical Planning and Inference, vol. 7, pp. 307-316, 1983.

[2] B. G. Greenberg, A. L. A. Abul-Ela, W. R. Simmons, and D. G. Horvitz, "The unrelated question randomized response model- theoretical framework," *Journal of American Statistical Association*, vol. 64 (326), pp. 520-539, 1969.

[3] S. N. Gupta, "Qualifying the sensitivity level of binary response personal interview survey questions," Journal of Combinatorics, Information and System Sciences, vol. 26 (1-4), pp. 101-109, 2001.

[4] S. N. Gupta, B. C. Gupta, S. Singh, "Estimation of sensitivity level of personal interview survey questions," Journal of Statistical Planning and Inference, vol. 100, pp. 239-247, 2002.

[5] S. N. Gupta, B. Thornton, J. Shabbir, and S. Singhal, "A comparison of multiplicative and additive optional RRT models," Journal of Statistical theory and application, vol. 5(3), pp. 226-239, 2006.

[6] S. N. Gupta, J. Shabbir, and S. Sehra, "Mean and sensitivity estimation in optional randomized response models," Journal of Statistical Planning and Inference, vol. 140, pp. 2870-2874. 2010.

[7] K. C. Huang, "Unbiased estimators of mean, variance and sensitivity level for quantitative characteristics in finite population sampling," Metrika, vol. 71, pp. 341-352, 2010.

[8] N. S. Mangat, and R. Singh, "An alternative randomized response procedure," Biometrika, vol. 77 (2), pp. 439-442, 1990.

[9] S. Mehta, B. K. Dass, J. Shabbir, and S. Gupta, "A three-stage optional randomized response model," Journal of Statistical Theory and Practice, vol. 6(3), pp. 417-427, 2012.

[10] J. S. Sihm, and S. N. Gupta, "A two stage binary optional randomized response model," Communications in Statistics-Simulation and Communication, vol. 44 (9), pp. 2278-2296, 2014.

[11] N. Tiwari and P. Mehta, "An improved two stage optional RRT model," Journal of Indian Society of Agricultural Statistics, vol. 70(3), pp. 197-203, 2016.

[12] S. L. Warner, "Randomized response : a survey technique for eliminating evasive answer bias," Journal of the American Statistical Association, vol. 60 (309), pp. 63-69, 1965.