ISSN 0971-0388

Aligarh Journal of Statistics
Vol. 41 (2021), 47-62

Generalization of Two-Stage Randomized Response with an Extension in Optional Randomized Response

Dipika Patra[1] and Sanghamitra Pal[2]
[Received on October 2019. Accepted on March 2021]

ABSTRACT

In the era of survey sampling, there is an increasing interests in sensitive features of the population that people generally prefer to hide from people. A novel technique called the Randomized Response Technique (RRT)is used to gather reliable data, protectthe confidentiality of respondents, and estimate the population proportion, bearing the sensitive attribute. But the perception of sensitivity is notthe same for every person. It has been observed that a group of people are willing to disclose their true nature rather than the compulsory Randomized Response (RR). Considering this fact, the concept of optional RRT was developed.In the present work, we reformulate an estimation method of sensitive population proportion addressed by Singh, Singh, Mangat and Tracy s'two-stage RRT. The procedure is more generalized here when respondents are chosen exclusively byan unequal probabilitysampling scheme and an unbiased estimator is derived along with its unbiased variance estimator. Further, we develop an Optional Randomized Response (ORR) method based on our proposed work. A simulation study has been carried out to find out the efficacy of the proposed RR and ORR procedures.

1. Introduction

In socio-economic surveys, often the study relates to personal features such as drinking habits, drug usage, abortion, dopingwhich people wish to keep secret. Sometimes the survey question itself is sensitive; an interviewer may hesitate to ask such a question. In such cases, the respondents may simply refuse to answer such questions. Thus, the direct survey fails to produce reliable estimates. Therefore, it is impossible to derive an unbiased estimator of the unknown population parameter θ_A, bearing the sensitivity attribute *A*.

✉ : Sanghamitra Pal
Email: mitrapal2013@gmail.com

Extended author information available after reference list of the article.

Warner (1965) introduced the Randomized Response Technique (RRT) inestimating the proportion of the population with a sensitive attribute where the sample was drawn by simple random sampling with replacement (SRSWR). In this method, a Randomized Response (RR) device is provided to each sampled person instead of direct questioning. The device may be a box, containing a sufficient number of identical cards marked bytwo types ofthe questions - "Do you belong to *A*?" and "Do you belong to A^c?" with proportions $p \neq \frac{1}{2}$ and $(1-p)$ respectively. Respondents are requested to draw a card, randomly, and report the answer "yes" or "no" to the question mentioned on the card regardinghis/her characteristics without divulging the card type to the interviewer. The techniquefurther was refined by Horvitz *et al.* (1967), Greenberg *et al.* (1969), Kuk (1990), Mangat and Singh (1990), Mangat (1992), and many others. In each case, a sample is drawn by SRSWR.Chaudhuri (2001) proposed a method to address the above problem where the sample was drawn by unequal probability sampling. Singh *et al.* (1995) suggested a two-stage RR procedure withthe simple random sampling frame which is an improved version of Mangat and Singh(1990), Tracy and Osahan (1999). Tracy and Singh (2018) extended the work of Singh *et al.* (1995) in stratified sampling. In India, National Sample Survey Organization (NSSO) draws samples byunequal probability sampling schemes. In this paper, the objective is to develop Singh *et al.* (1995)'s two-stage RR procedure in unequal probability sampling, taking the cue of Chaudhuri (2001) and toconstruct an optional RRT, using the proposed model. Chaudhuri and Mukerjee (1985) introduced the concept of the Optional Randomized Response (ORR) technique, recognizing the fact that the subject of the inquiry may not be sensitive enough toa group of respondents and they may prefer to respond directly. In this technique,a direct response option is offered along with the specified RR device to respondents. Respondents are requested to report their response directly if they do not feel the survey question is sensitive, otherwise, they have the option to answer withthe RR device, without divulging the option so exercised. A large number of developments were proposed by Gupta (2001), Gupta *et al.* (2002), Pal (2008), and many others, following this approach. Recently, Patra and Pal (2019) and Pal *et al.* (2020) extended a few well-known RRTs in ORR techniques in estimating the degree of privacy protection. An alternative ORR approach was illustrated by Chaudhuri and Mukerjee (1988), later Chaudhuri and Saha (2005) forgeneral sampling scheme. Subsequent developments of the related work are narrated in Arnab (2004), Arnab and Rueda (2016), and many others. In this paper, we have modified Singh *et al.* (1995)'s RRT for unequal probability sampling design in Section 3.

Section 4 is designed for an Optional Randomized ResponseTechnique (ORRT) following Chaudhuri and Mukerjee (1985)'s approach for the proposed generalized model.A simulation study is carried out to show the efficacy of our result in Section 5.

2. Singh *et al.* (1995) Model

Mangat and Singh (1990) considered two RR devices to ask an individual about his/her group membership related to the sensitive attribute which people usuallyprefer to hide from others. The first RR device R_1 consists of two statements say, "I belong to the sensitive group" and "Go to the second randomized device R_2" with probabilities T (0<T <1) and $(1-T)$ respectively. The device R_2 is the same as Warner's (1965) RR device, if and only if it is directed from the device R_1.

Following them, Tracy and Osahan (1999) modified the device R_2 of this procedure. According to them, R_2 is formed with three statements say, "I belong to the sensitive group", "yes" and "no" with proportion $p:(1-p)/2$ and $(1-p)/2$.

Another modification was suggested by Singh *et al.* (1995). Their approach isthe same of the above. Singh *et al.* modified the second device R_2 with the same three statements but in proportions $p,(1-p)w$ and $(1-p)(1-w)$ where $w\in[0,1]$ and $0<p<1,\ p\neq\frac{1}{2}$.

In Section 3, we try to modify their work for general sampling design.

3. Proposed RRT Model

Let $U=(1,2,.....,N)$ be a finite population on which the variables y and z are defined. The variable y is of our principal interest and sensitive in nature, and z is positive integer-valuedvariable which is highly correlated with the study variable y. Their y and z values are y_i and z_i with totals Y and Z respectively. The variable y relates to a sensitiveattribute A. A sample s is chosen from U bygeneralsampling design,using the size measure variable z ,and individuals are directed to respond truthfully with the following RR device.

Let for the person labelled i, we define.

$$y_i = \begin{cases} 1 & \text{if } i^{th} \text{ individual possess sensitive attribute } A \\ 0 & \text{otherwise,} \end{cases} \quad i = 1, 2, ..., N \,. \tag{3.1}$$

The population proportion is defined as

$$\theta_A = \frac{1}{N}\sum_{i=1}^{N} y_i \text{ which is unknown.} \tag{3.2}$$

Two boxes with different specificationsare given to the respondent. The first box, box I (say) contains two types of identical cards which are identical in length and width, and marked as "Do you possess *A*?" and "Go to the box II" with the probabilities T(0<T<1) and (1-T) respectively.

Box II consists of three types of cards marked as "I possess A", "yes" and "no" with proportions p:(1-p)w:(1-p)(1-w) where 0<p <1 and wÎ(0,1) ,are known to the interviewer and are specified before the investigation. He/she is requested to draw a card from the first box, box I, without divulging the card type. The respondent will draw a card from box II only if it is directed from box I. Finally, his/her responseis recorded without knowing any information about the card type.

Denoting I_i be the response received from i^{th} individual, we may write

$$I_i = \begin{cases} 1 & \text{if response is "yes"} \\ 0 & \text{if response is "no"} \end{cases} \tag{3.3}$$

Then, one may get

$$P(I_i = 1) = Ty_i + (1-T)\{py_i + (1-p)w\} \tag{3.4}$$

and $P(I_i = 0) = T(1-y_i) + (1-T)\{p(1-y_i) + (1-p)(1-w)\}$. (3.5)

Writing,

$E_R(I_i) = Ty_i + (1-T)\{py_i + (1-p)w\}$, denoting E_R as expectation due to the RR device.

Writing, $(1-T)(1-p) = \phi$, we get $E_R(\frac{I_i - \phi w}{1-\phi}) = y_i$. An unbiased estimator of y_i is

$$r_i = \frac{I_i - \phi w}{1-\phi} \quad \phi \neq 1. \tag{3.6}$$

The variance of the estimator (3.6) is formulated as follows:

$$V_R(r_i) = \frac{V_R(I_i)}{(1-\phi)^2} = \frac{E_R(I_i)(1-E_R(I_i))}{(1-\phi)^2} = \frac{\phi(1-\phi)(y_i - w)^2 + \phi w(1-w)}{(1-\phi)^2}, \tag{3.7}$$

where V_R stands for the variance for RR device and $\phi \neq 1$.

Clearly, $V_R(r_i)$, a function of y_i, is unknown as the variance depends on the value of y_i which is estimated through r_i.

It follows that,

$$V_R(r_i) = \frac{\emptyset w(1-\emptyset w)}{(1-\emptyset)^2} + \frac{\emptyset(1-w)(1-\emptyset+\emptyset w)-\emptyset w(1-\emptyset w)}{(1-\emptyset)^2} y_i \quad ; \emptyset \neq 1. \tag{3.8}$$

An unbiased variance estimator of $V_R(r_i)$ is denoted by $v_R(r_i)$,

$$v_R(r_i) = \frac{\phi\, w(1-\phi+\phi\, w)+\phi(1-2w)I_i}{(1-\phi)^2}\,; \phi \neq 1. \tag{3.9}$$

Now,employing Horvitz Thompson's (1952) unbiased estimator for θ_A, the final estimator becomes

$$t_1 = \frac{1}{N}\sum_{i\in s}\frac{r_i}{\pi_i}. \tag{3.10}$$

where π_i is the first order inclusion probability of the i^{th} unit in the sample.

Denoting the overall expectation and variance operator as E and V, we may write,

$E(t_1) = \theta_A$ and

$V(t_1) = E_P V_R(t_1) + V_P E_R(t_1)$

where E_P and V_P stands for expectation and variance for the sampling design P respectively.

Hence, the overall variance of the estimatordefined in (3.10) can be written as follows,

$$V(t_1) = \sum_{i}^{N}\sum_{<j=1}^{N}(\pi_i\pi_j - \pi_{ij})(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j})^2 + \sum_{i=1}^{N}\frac{V_R(r_i)}{\pi_i}. \tag{3.11}$$

It may be estimated by(Chaudhuri (2011, pp. 39))

$$v(t_1) = \frac{1}{N^2}[\sum_{i}^{n}\sum_{<j=1}^{n}(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}})(\frac{r_i}{\pi_i} - \frac{r_j}{\pi_j})^2 + \sum_{i=1}^{n}\frac{v_R(r_i)}{\pi_i}] \tag{3.12}$$

where $v_R(r_i)$ (see the equation (3.9)) is the unbiased variance estimator of $V_R(r_i)$.

3.1 Generalization of Proposed RRT Model

In this subsection, we generalize the above-mentionedprocedure,modifying the second RR device, box II.The feature of box I remainsthe same. Box I contains cards marked as "Do you possess A?" and "Go to the box II" with probabilities $T(0<T<1)$ and $(1-T)$ respectively. Box IIcontains those three types of cards as mentioned in Section 3. But the distribution of cards in box II will depend on the respondents' choices. The three types of cards marked as "I possess A","yes" and "no", are kept in three compartments in sufficient numbers. To build box II, respondent will insert $m_1(>0)$ number of "I possess A" cards in box II. He/she will take $m_2(>0)$ number of "yes" marked cards and $m_3(>0)$ number of "no" marked cards according to their own desire. The numbers m_2 and m_3 depend on the respondents and these number varies from person to person. Denoting the total number of cards as $m = m_1 + m_2 + m_3$, it is noted that m_1 and m are fixed for each respondent. Now, inbox II, for i^{th} respondent,these three types of cards are in proportion $p:(1-p)w_i^*:(1-p)(1-w_i^*)$ $(i=1,2,...,n)$. Here, $p=\frac{m_1}{m}$ is known and fixed. But $w_i^* \in (0,1)'s$ is unknown to the interviewer as m_2 depends on the respondents' choice.The remaining procedure of RRT is thesameas the previously proposed RRT model.

For instance, Respondents are requested to build box II with $m=12$ cards in which $m_1=4$ cards are strictly marked as "I possess A". The other $m=m_1=8$ cards are either "yes" or "no" type. The "yes"and "no" written cards are kept in a compartment in sufficient numbers.Then,the respondents are requested tofollow the above mentioned proposed generalized RRT procedure.

Let Respondent 1 puts $m_2=6$ cardsmarked as "yes" and the $m_3=2$ cards marked as "no". So in box II, the proportion of cards "I possess A": "yes": "no" is

$\frac{1}{3}:\frac{1}{2}:\frac{1}{6} \equiv \frac{1}{3}:(\frac{2}{3}\times\frac{3}{4}):(\frac{2}{3}\times\frac{1}{4})$. Here $p=\frac{1}{3},\ w_i^*=\frac{3}{4}$.

Similarly, consider the other respondent, says Respondent 2, put $m_2 = 5$ cards marked as "yes" and $m_3 = 3$ cards marked as "no". In that case, the proportion of cards in box II becomes $\frac{1}{3}:\frac{5}{12}:\frac{1}{4} \equiv \frac{1}{3}:(\frac{2}{3}\times\frac{5}{8}):(\frac{2}{3}\times\frac{3}{8})$.

Here w_i^* is changed to $\frac{5}{8}$ but p is fixed at $\frac{1}{3}$.

The whole process is repeated one more time, independently but with different box II while box I remainsthe same. Let J_i and J_i' are two independent responses for i^{th} respondent. Respondents are requested to build box II. For the first response of i^{th} individual, box II is built with the cards in proportion $p_1:(1-p_1)w_i^*:(1-p_1)(1-w_i^*)$ (say). For the second response, box IIis built with the cards in proportion $p_2:(1-p_2)w_i^*:(1-p_2)(1-w_i^*)$ by changing only the number of cards marked as "I possess *A*". Here $0 < p_j < 1,\ j = 1,2,\ p_1 \neq p_2$.

It can be written as

$$J_i = \begin{cases} 1 & \text{if response is "yes"} \\ 0 & \text{if response is "no"} \end{cases} \tag{3.13}$$

Then, one may get

$$P(I_i = 1) = Ty_i + (1-T)\{p_1 y_i + (1-p_1)w_i^*\} \tag{3.14}$$

and $P(I_i = 0) = T(1-y_i) + (1-T)\{p_1(1-y_i) + (1-p_1)(1-w_i^*)\}$. (3.15)

Taking the expectation due to RR device we get,

$$E_R(J_i) = Ty_i + (1-T)\{p_1 y_i + (1-p_1)w_i^*\}$$

For the second response J_i' we may write,

$$E_R(J_i') = Ty_i + (1-T)\{p_2 y_i + (1-p_2)w_i^*\}.$$

Then,

$$E_R((1-p_2)J_i - (1-p_1)J_i') = (p_1 - p_2)y_i$$

i.e. $E_R\left(\frac{(1-p_2)J_i - (1-p_1)J_i'}{(p_1 - p_2)}\right) = y_i,\ \ p_1 \neq p_2$.

So, an unbiased estimator of y_i takes the form,

$$r_i' = \frac{(1-p_2)J_i - (1-p_1)J_i'}{(p_1 - p_2)}; \;\; p_1 \neq p_2 \tag{3.16}$$

and hasa variance

$$V_R(r_i') = \frac{(1-p_2)^2 V_R(J_i) + (1-p_1)^2 V_R(J_i')}{(p_1 - p_2)^2}. \tag{3.17}$$

In the above expression, the variance $V_R(J_i)$ may be written as,

$$V_R(J_i) = \emptyset_1 w_i^*(1-\emptyset_1 w_i^*) + \emptyset_1\{(1-w_i^*)(1-\emptyset_1+\emptyset_1 w_i^*) - w_i^*(1-\emptyset_1 w_i^*)\}y_i, \text{ where } \emptyset_1 = (1-T)(1-p_1). \tag{3.18}$$

and $V_R(J_i')$ can be defined similarly.

Following Chaudhuri (2011), $V_R(r_i')$ may be unbiasedly estimated by

$$v_R(r_i') = r_i'(r_i'-1) = \frac{(1-p_1)(1-p_2)}{(p_1-p_2)^2}(J_i - J_i')^2. \tag{3.19}$$

Now,the Horvitz- Thompson (1952) unbiased estimator of θ_A can be written here as,

$$t_2 = \frac{1}{N}\sum_{i\in s}\frac{r_i'}{\pi_i}. \tag{3.20}$$

Thus, the variance of the above estimatorcan be written as follows,

$$V(t_2) = \sum_{i}^{N}\sum_{<j=1}^{N}(\pi_i\pi_j - \pi_{ij})(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j})^2 + \sum_{i=1}^{N}\frac{V_R(r_i')}{\pi_i}. \tag{3.21}$$

An unbiased variance estimator of t_2 is written as,

$$v(t_2) = \frac{1}{N^2}[\sum_{i}^{n}\sum_{<j=1}^{n}(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}})(\frac{r_i'}{\pi_i} - \frac{r_j'}{\pi_j})^2 + \sum_{i=1}^{n}\frac{v_R(r_i')}{\pi_i}]. \tag{3.22}$$

4. Proposed ORRT Model

Chaudhuri and Mukerjee (1985) developed the Optional Randomized Response (ORR) model to allow the option of direct response (DR) to those who do not consider this particular attribute is stigmatizing enough and prefer to divulge the truth. The respondents may opt for RRT also. In this section, we propose a new ORR model using the proposed generalized RR device mentioned in subsection 3.1. Chaudhuri (2011), and Chaudhuri, Christofides and Rao (2016) in their books provide extensive developments in RRT, also in ORRT. In the proposed ORRT model, we provide the choice of direct response with the proposed generalized RR device sothat the respondent may respond directly with his/her probability $c_i (0 \le c_i \le 1)$ and by the RR device with probability $(1-c_i)$.

The response of i^{th} individual under the ORRT model may be written as,

$$Z_i = \{ \begin{array}{ll} y_i & \text{with probability } c_i \\ I'_i & \text{with probability } (1-c_i) \end{array} \qquad (4.1)$$

Here I'_i be the randomized response using the proposed model.

Thus,

$$E_R(Z_i) = c_i y_i + (1-c_i)\{Ty_i + (1-T)\{p_1 y_i + (1-p_1)w_i^*\}\} .$$

Two independent responses are required here to get an unbiased estimator of y_i.

Another response is recorded as Z'_i by changing the proportion of cards as suggested in the proposedRR device in the previous section.

So,

$$E_R(Z'_i) = c_i y_i + (1-c_i)\{Ty_i + (1-T)\{p_2 y_i + (1-p_2)w_i^*\}\} \quad ; \quad 0 < p_j < 1,$$

$p_1 \ne p_2$ $(j = 1, 2)$.

It follows that,

$$E_R((1-p_2)Z_i - (1-p_1)Z'_i) = (p_1 - p_2)\{c_i y_i + (1-c_i)Ty_i + (1-c_i)(1-T)y_i\}$$

$$i.e.\ r_i^* = \frac{(1-p_2)Z_i - (1-p_1)Z'_i}{p_1 - p_2} \quad \text{where } p_1 \ne p_2 . \qquad (4.2)$$

It easily follows that, $E_R(r_i^*) = y_i$ and the estimator (4.2) has variance as

$$V_R(r_i^*) = \frac{(1-p_2)^2 V_R(Z_i) + (1-p_1)^2 V_R(Z'_i)}{(p_1 - p_2)^2} \qquad (4.3)$$

where

$$V_R(Z_i) = \phi_1^* w_i^* (1-\phi_1^* w_i^*) + \phi_1^* (1-w_i^*)(1-\phi_1^* + \phi_1^* w_i^*) - w_i^* (1-\phi_1^* w_i^*) y_i$$

$$\text{taking } \phi_1^* = (1-c_i)(1-T)(1-p_1). \qquad (4.4)$$

Similarly, $V_R(Z'_i)$ can be formulated.

The variance $V_R(r_i^*)$ can be estimated through the unbiased variance estimator say,

$$v_R(r_i^*) = r_i^*(r_i^* - 1) = \frac{(1-p_1)(1-p_2)}{(p_1 - p_2)^2}(Z_i - Z'_i)^2 . \qquad (4.5)$$

So, an unbiased estimator of θ_A under ORRT model becomes

$$t_3 = \frac{1}{N}\sum_{i\in s}\frac{r_i^*}{\pi_i}. \qquad (4.6)$$

Its variance (Chaudhuri (2011))is

$$V(t_3) = \frac{1}{N^2}[\sum_{i}^{N}\sum_{<j=1}^{N}(\pi_i\pi_j - \pi_{ij})(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j})^2 + \sum_{i=1}^{N}\frac{V_R(r_i^*)}{\pi_i}]. \qquad (4.7)$$

An unbiased estimator of $V(t_3)$ (Chaudhuri (2011)) may bederived as

$$v(t_3) = \frac{1}{N^2}[\sum_{i}^{n}\sum_{<j=1}^{n}(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}})(\frac{r_i^*}{\pi_i} - \frac{r_j^*}{\pi_j})^2 + \sum_{i=1}^{n}\frac{v_R(r_i^*)}{\pi_i}]. \qquad (4.8)$$

5. Simulation Study

In order to demonstrate how the proposed method works,we consider afictitiouspopulationof size 117. The values of y_i 's and size measures z_i 's are taken from Chaudhuri *et al.* (2009) in which

$y_i = 1$ if ith person evades income tax

0, otherwise

The size measure variable z refers to the monthly per capita consumption. The study variable and the size measure variable are highly correlated (0.8325).

A sample of size $n = 45$ is drawn by Lahiri(1951)-Midzuno(1952)-Sen (1953) sampling scheme. In our data $\theta_A = 0.803$.

To judge the efficacy of the proposed models, a large number of samples (say 1000) are drawn. Table 1shows the results of the 527th simulation for all methods. Table 2 compares all methods with their average coverage probability (ACP in %), average coefficient of variance (ACV in %), and average length (AL) of confidence interval. The point estimator will be judged well if the ACV, the average over 1000 replications of estimated coefficient of variations, has a small magnitude, preferably less than10% or at most 30%. The probability that theconfidence interval $CI_{0.95} = [LCL, UCL]$ contains θ_A is referredas coverage probability (CP) and ACP is the average of all CP's. Different methods are compared by Box plots for estimated proportions (Figure 1.1) and estimated variances (Figure 1.2). Confidence limitsfrom 200th to 500th simulationsofeach method are shown diagrammatically in Figures 2,3, and 4. Judging by range, method 2 (proposed generalized RRT) and method 3 (proposed ORRT) have more variability than method 1 (proposed RRT). It is noteworthy here that the interquartile range (IQR) (from Figure 1.1 and Figure 1.2) is more spreadableformethod 2. In figures 2, 3, and 4, we have plotted the respective

[LCL, UCL] by dotted lines with an estimate of θ_A. The three methods named "Proposed RRT", "Proposed Generalized RRT" and "Proposed ORRT" are described in Sections 3, 3.1, and 4, respectively, and these methodsare compared in this study.

Table 1:
Comparison of different methods (527th simulation).

(T,p_1, p_2)	**Proposed Method**	**Estimated Proportion**	**Estimated Variance**	**Coefficient of Variance**	**Average Length**
(0.67,0.34, 0.56)	**1. RRT**	0.7115	0.0053	10.2122	0.2848
	2.Generalized RRT	0.7818	0.0205	18.3354	0.5619
	3.ORRT	0.7595	0.0142	15.6776	0.4668

Table 2:
Comparison by ACV, ACP, AL.

Proposed Method	**ACP**	**ACV**	**AL**
1. RRT	52.8	11.3829	0.2893
2. Generalized RRT	85.2	26.1468	0.6165
3. ORRT	94.1	22.7253	0.6011

6. Concluding Remarks

In this paper, Singh *et al.* (1995)'s two-stage RRT has been generalized forunequal probability sampling scheme and modified by generating different proportions of "yes" and "no" cards which arecreated by respondents, ensuring more protection to them. The proposed RRT model is alsoextended for the Optional Randomized Response (ORR) model.The proposed methods provide satisfactory results in terms of the comparison criteria ACP, ACV, and AL. In comparing boxplots of different estimates and their variance estimates, we may conclude that the proposed method performs well in ORRT. The proposed ORRT method is quite satisfactory in terms of the above criteria also.

Acknowledgment

The authors are grateful to the editor and the referees for their valuable comments and suggestions.

References

Arnab, R. (2004): Optional randomized response techniques for complex survey designs. *Biometrical Journal*, **46(1)**, 114-124.

Arnab, R. and Rueda, M. (2016): Optional randomized response: A critical review. Handbook of Statistics, Elsevier, **34**, 253-271.

Chaudhuri, A. (2001): Using randomized response from a complex survey to estimate a sensitive proportion in a dichotomous finite population. *Journal of Statistical Planning and Inference*, **94**, 37-42.

Chaudhuri, A. (2011): Randomized response and indirect questioning techniques in surveys. CRC Press.

Chaudhuri, A., Christofides, T.C. and Rao, C.R. (2016): Handbook of statistics 34, Data Gathering, Analysis and Protection of Privacy through Randomized Response Techniques: Qualitative and Quantitative Human Traits. Elsevier, NL.

Chaudhuri, A., Christofides, T.C. and Saha, A. (2009): Protection of privacy in efficient application of randomized response techniques. *Statistical Methods and Applications*, **18**, 389–418.

Chaudhuri, A. and Mukerjee, R. (1985): Optionally randomized responses techniques. *Calcutta Statistical Association Bulletin*, **34**, 225–229.

Chaudhuri, A. and Mukerjee, R. (1988): Randomized Responses: Theory and Techniques. Marcel Dekker, New York, NY.

Chaudhuri, A. and Saha, A. (2005): Optional versus compulsory randomized response techniques in complex surveys. *Journal of Statistical Planning and Inference*, **135(2)**, 516-527.

Greenberg, B.G., Abul-Ela, A.-L., Simmons, W.R. and Horvitz, D.G. (1969): The unrelated question RR model: Theoretical framework. *Journal of American Statistical Association*,**64**, 520–539.

Gupta, S. (2001): Qualifying the sensitivity level of binary response personal interview survey questions. *Journal of Combinatorics, Information & System Sciences* **26 (1- 4),**101-109.

Gupta, S., Gupta, B. and Singh, S. (2002): Estimation of sensitivity level of personal interview survey question. *Journal of Statistical Planning and Inference*, **100**, 239-247.

Horvitz, D.G., Shah, B.V. and Simmons, W.R. (1967): The unrelated question randomized response model. *Proceeding of the Social Stat. Sec. Amer. Stat. Assoc.*, 65-72.

Horvitz, D.G. and Thompson, D.J. (1952): A generalization of sampling without replacement from a finite universe. *Journal of American Statistical Association*, **47**, 663–685.

Kuk, A.Y.C. (1990): Asking sensitive questions indirectly. *Biometrika,* **77(2)**, 436–438.
Lahiri, D. B. (1951): A method of sample selection providing unbiased ratio estimates. *Bulletin of the International Statistical Institute,* **3**, 133-140.
Mangat, N.S. (1992): Two stage randomized response sampling procedure using unrelated question. *Journal of Indian Society of Agricultural Statistics,* **44(1)**, 87-88.
Mangat, N.S. and Singh, R. (1990): An alternative randomized response procedure. *Biometrika,* **77**, 439-442.
Midzuno, H.(1952): On the sampling system with probability proportional to the sum of the sizes, *Annals of the Institute of Statistical Mathematics,* **3**, 99-107.
Pal, S. (2008): Unbiasedly estimating the total of a stigmatizing variable from a complex survey on permitting options for direct or randomized responses. *Statistical Papers,* **49**, 157–164.
Pal, S., Chaudhuri, A., and Patra, D. (2020): How Privacy may be Protected in Optional Randomized Response Surveys. *Statistics in Transition,* **21(2)**, 61-87.
Patra, D., and Pal, S. (2019): Privacy Protection in Optional Randomized Response Surveys for Quantitative Characteristics. *Statistics and Applications,* **17(2)**, 47-54.
Sen, A. R. (1953): On the estimator of the variance in sampling with varying probabilities. *Journal of Indian Society of Agricultural Statistics,* **5**, 119-127.
Singh, R., Singh, S., Mangat, N.S. and Tracy, D. S. (1995): An improved two stage randomized response strategy.*Statistical Papers,* **36**, 265-271.
Tracy, D.S. and Osahan, S.S. (1999): An improved randomized response technique. *Pakistan Journal of Statistics,* **15(1)**, 1-6.
Tracy, D.S. and Singh, H.P.(2018): A proficient two-stage stratified randomized response strategy. *Journal of Modern Applied Statistical Methods,* **17(1)**, 1-11.
Warner, S. L. (1965): Randomized response: a survey technique for eliminating evasive answer bias. *Journal of American Statistical Association,* **60**, 63-69.

Authors and Affilliations

Dipika Patra[1] and Sanghamitra Pal[2]

Dipika Patra
dipika.patra1988@gmail.com

[1,2]Department of Statistics, West Bengal State University, India.

Box plot for estimated proportions (proposed RRT, proposed generalized RRT, proposed ORRT)

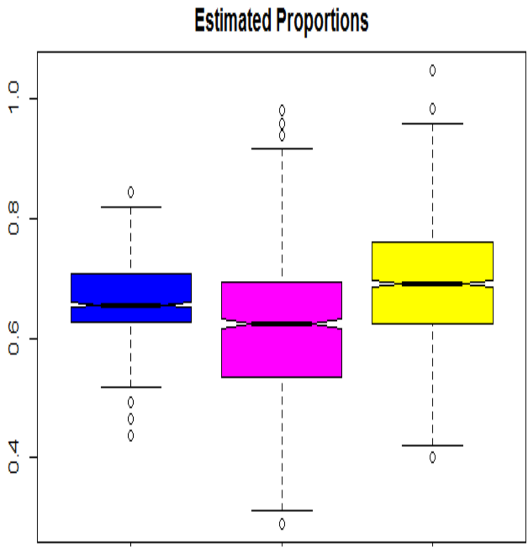

Fig. 1.1

Box plot for estimated variances (proposed RRT, proposed generalized RRT, proposed ORRT)

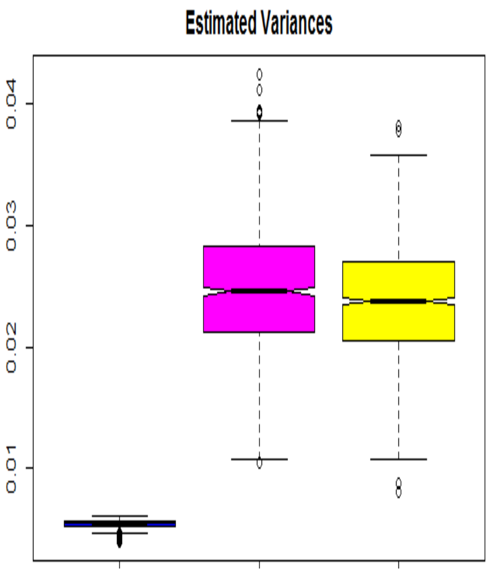

Fig. 1.2

[LCL, UCL] for proposed RRT model

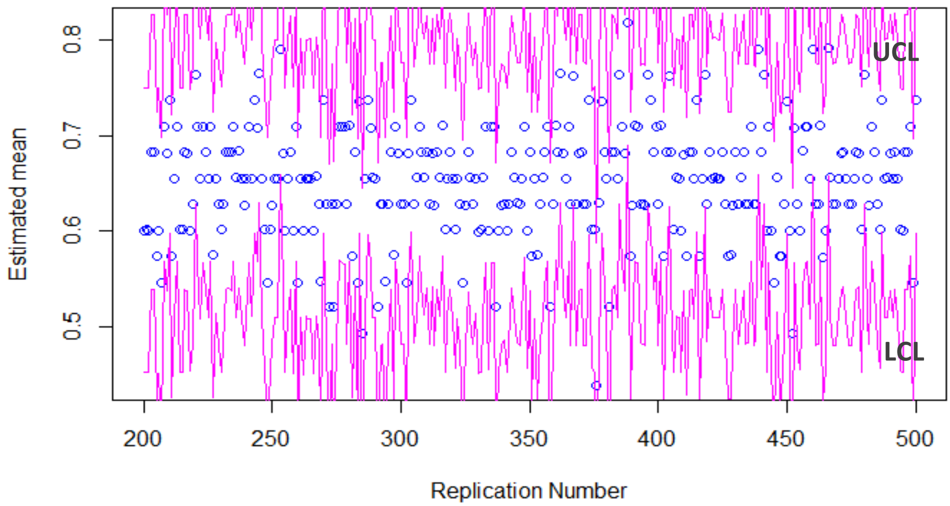

Fig. 2

[LCL, UCL] for proposed generalized RRT model

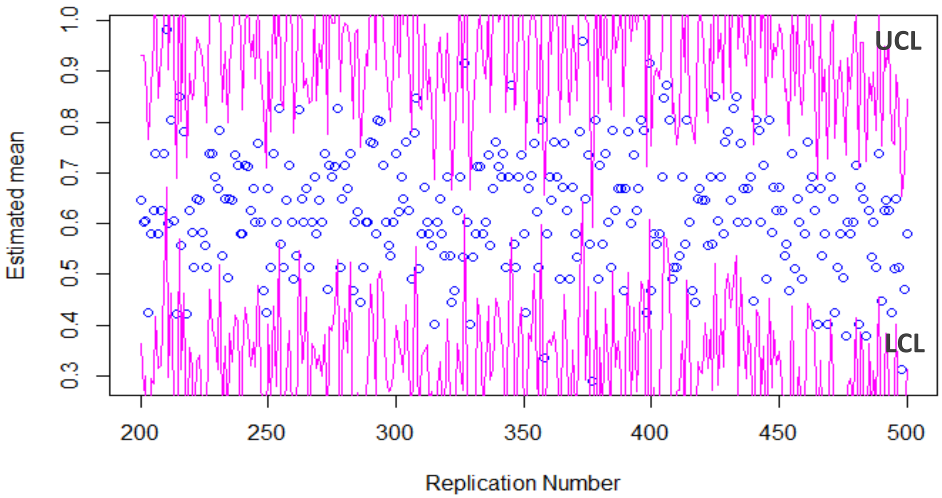

Fig. 3

***[LCL, UCL]* for proposed ORRT model**

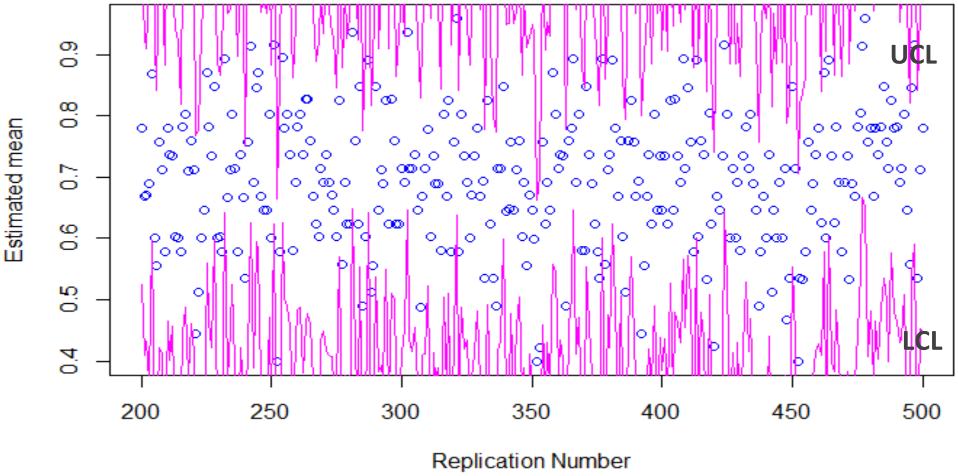

Fig. 4