Journal of Statistical Theory and Applications
Vol. **18(4)**, December 2019, *pp.* 351–360
DOI: https://doi.org/10.2991/jsta.d.191112.001; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

Improved Randomized Response in Optional Scrambling Models

Zawar Hussain*, Muhammad Imran Shahid

Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan

ARTICLE INFO

Article History
Received 07 Jul 2015
Accepted 05 Jun 2019

Keywords
Randomized response
Percent relative efficiency
Optimum sample size
Optional randomized response model
Sensitive quantitative variable

2000 Mathematics Subject Classification: 62D05

ABSTRACT

In the present study, we discuss the issue of increasing the respondents cooperation in sensitive surveys. When the question is highly sensitive then the cooperation from the respondents is decreased. We propose two optional randomized response models (ORRMs) to increase the respondents cooperation. A comparison of proposed ORRMs with [1] two- and three-stage models have been made. It is found that, in estimating mean and sensitivity level, the proposed strategies perform better than [1] two- and three-stage models. A comparison is also made among the [1] two- and three-stage models and proposed ORRMs to identify the best one. Numerical illustration are also given in favor of the algebraic results.

1. INTRODUCTION

Through usual methods (direct questioning) of surveys, obtaining the honest response on all the questions in a questionnaire is difficult due to stigmatizing nature of questions. Generally, when the questions are inoffensive, the true response on them are procured. In contrast, interviewees try to conceal their real response about stigmatizing behavior(s), possibly, due to security problems, religious faith, social desirability, etc. Thus, misreporting and non-response is expected in such situations. Consequently, a bias creeps into the estimates and statistical inferences are rendered invalid. Such type of bias is common in face to face surveys where the interviewees are inquired through direct questioning about the possession of the stigmatizing characteristics. To overcome such type of problems, Warner [2] suggests a model that reduces the response bias and increases cooperation from the respondents. The model is known as randomized response model (RRM). This model generates randomized response (RR). The name RR recommends that a reported response cannot be traced back to the actual response of the respondents. In addition, response of a given respondent may be different if he/she is asked for a repeated response. The RRM by Warner [2] consists of presenting two complementary questions to the respondents with known probabilities. A respondent has to answer one of them depending upon the random outcome of a Bernoulli experiment. In this way, the question actually answered by a respondent remains unknown to the interviewer and privacy of the respondents remains intact. This privacy protection and confidentiality are, perhaps, the main reasons of respondent's trust on RRM. The main aim of RRM is to estimate proportion of individuals possessing stigmatizing characteristic. For example, RRM may be of great help in estimating the proportion of smokers in the university, proportion of class bunkers in the college, proportion of people having extra marital affairs and average number of bottles of alcohol used per week by a person.

Greenberg *et al.* [3] presented the idea of quantitative response by using two quantitative questions. Of these, one is sensitive in nature and other is unrelated non-sensitive question. Since then, several authors, including, [4–6] and many others have worked on quantitative RRMs. Gupta and Thornton [7] propose a model where a predetermined proportion of the respondents are asked to provide true responses and the remaining are told to scramble their responses. Gupta and Shabbir [8] discussed that in Gupta *et al.* [6] model, some approximation is needed in the estimation of variances of the estimates. Gjestvang and Singh [9] present a forced quantitative randomized RRM. Huang [10] present his model to estimate mean and sensitivity level of stigmatizing variables. Gjestvang and Singh [11] present a different type of additive model in which selected respondents scramble their responses for both assertions. Other authors like [1,12–16] models also contributed toward the estimation of mean and sensitivity level of sensitive quantitative variables.

*Corresponding author. Email: zhlangah@yahoo.com

Our emphasis in this inquiry is on optional randomized response models (ORRMs) only. In such type of models, respondents which are selected in a random sample have a choice in providing the responses. If the interviewee feels the selected question sensitive, he/she has the option to scramble his/her response, otherwise, a true response is provided by him/her. In Gupta *et al.* [14], a large value of truth parameter (T) is needed, if the question is highly sensitive. Mehta *et al.* [1] present two- and three-stage models by introducing a forced scrambling parameter (F). Their models perform better in estimating mean but they did not discuss the performance of sensitivity estimator. Motivated by Mehta *et al.* [1] we propose two ORRMs which are more efficient than the [1], two- and three-stage models. Before presenting the proposed models in the next section, we briefly discuss the [1] two- and three-stage procedures.

2. SOME RECENT QUANTITATIVE ORRMs

The following sub-sections give a short description of [1] two- and three-stages procedures.

2.1. [1] Two-Stage Model

Mehta *et al.* [1] present a two-stage model when the question of interest is highly sensitive. A procedure similar to that of Gupta *et al.* [14] is adapted for collecting two responses with a difference at the first stage. Two independent sub-samples of size n_1 and n_2 are drawn from the population such that $n_1 + n_2 = n$. In the ith sub-sample, at stage-1, a known proportion of the respondents F is directed to scramble their responses and the remaining proportion $(1 - F)$ of the respondents provide the response as report the true response X if you feel the question insensitive, otherwise, report scrambled response $(X + Y_i)$. In the ith sub-sample, the distribution of reported response U_i may be written as

$$U_i = \begin{cases} X & \text{with probability } (1-F)(1-W) \\ (X+Y_i) & \text{with probability } F + (1-F)\,W. \end{cases} \tag{1}$$

The optional RR U_i in the ith sub-sample is given by

$$U_i = \beta\,(X+Y_i) + (1-\beta)\{\alpha\,(X+Y_i) + (1-\alpha)\,X\}, \quad i = 1, 2. \tag{2}$$

where, α and β are the Bernoulli random variables with mean W and F, respectively. The unbiased estimators of μ_X and W are given by

$$\hat{\mu}_{XM1} = \frac{\mu_1 \overline{U}_2 - \overline{U}_1 \mu_2}{\mu_1 - \mu_2}, \quad \mu_1 \neq \mu_2 \tag{3}$$

$$\hat{W}_{M1} = \frac{1}{(1-F)}\left\{\frac{\overline{U}_2 - \overline{U}_1}{\mu_1 - \mu_2} - F\right\}, \quad \mu_1 \neq \mu_2, \quad F \neq 1. \tag{4}$$

The variances of the estimators $\hat{\mu}_{XM1}$ and $\hat{W}_{M1}$ are given by

$$Var\,(\hat{\mu}_{XM1}) = \frac{1}{(\mu_1-\mu_2)^2}\left[(\mu_2-1)^2\left(\frac{\sigma^2_{U_1}}{n_1}\right) + (\mu_1-1)^2\left(\frac{\sigma^2_{U_2}}{n_2}\right)\right], \quad \mu_1 \neq \mu_2, \tag{5}$$

$$Var\left(\hat{W}_{M1}\right) = \frac{1}{(1-F)^2(\mu_1-\mu_2)^2}\left[\left(\frac{\sigma^2_{U_1}}{n_1}\right) + \left(\frac{\sigma^2_{U_2}}{n_2}\right)\right], \quad \mu_1 \neq \mu_2, \quad F \neq 1. \tag{6}$$

$$\text{where,} \quad \sigma^2_{U_i} = \sigma^2_X + (F + (1-F)\,W)\,\delta^2_1 + \mu^2_1\,(F + (1-F)\,W)\,[(1 - (F + (1-F)\,W))], \quad i = 1, 2. \tag{7}$$

2.2. [1] Three-Stage Model

In addition to two-stage model, Mehta *et al.* [1] also propose a three-stage model to estimate the mean and sensitivity level of stigmatizing variable. According to their strategy, in each sub-sample, a fixed predetermined proportion (T) of respondents are directed to give the true response (X) and a fixed predetermined proportion (F) of the respondents are instructed to provide a scrambled response $(X + Y_i)$. The remaining proportion $(1 - T - F)$ of respondents are directed to use the ORRM. In the ith sample, the distribution of reported response is given as

$$U_i^* = \begin{cases} X & \text{with probability } T + (1-T-F)\,(1-W) \\ (X+Y_i) & \text{with probability } F + (1-T-F)\,W. \end{cases} \tag{8}$$

The optional RR U_i^* in the ith sub-sample is given by

$$U_i^* = X\eta + \beta (X + Y_i) + (1 - \eta - \beta) \{\alpha (X + Y_i) + (1 - \alpha) X\}. \quad i = 1, 2. \tag{9}$$

The unbiased estimators of mean μ_X and sensitivity level W are given by

$$\hat{\mu}_{XM2} = \frac{\mu_1 \overline{U}_2^* - \overline{U}_1^* \mu_2}{\mu_1 - \mu_2}, \quad \mu_1 \neq \mu_2 \tag{10}$$

$$\hat{W}_{M2} = \frac{1}{(1 - T - F)} \left\{ \frac{\overline{U}_2^* - \overline{U}_1^*}{\mu_1 - \mu_2} - F \right\}, \quad \mu_1 \neq \mu_2, \quad T + F \neq 1. \tag{11}$$

The variances of the estimators $\hat{\mu}_{XM2}$ and $\hat{W}_{M2}$ are given by

$$Var(\hat{\mu}_{XM2}) = \frac{1}{(\mu_1 - \mu_2)^2} \left[(\mu_2)^2 \left(\frac{\sigma^2_{U_1^*}}{n_1} \right) + (\mu_1)^2 \left(\frac{\sigma^2_{U_2^*}}{n_2} \right) \right], \quad \mu_1 \neq \mu_2 \tag{12}$$

$$Var\left(\hat{W}_{M2}\right) = \frac{1}{(1 - T - F)^2 (\mu_1 - \mu_2)^2} \left[\left(\frac{\sigma^2_{U_1^*}}{n_1} \right) + \left(\frac{\sigma^2_{U_2^*}}{n_2} \right) \right], \quad \mu_1 \neq \mu_2, \quad T + F \neq 1. \tag{13}$$

$$\text{where,} \quad \sigma^2_{U_i^*} = \sigma_X^2 + \delta_i^2 \left[(F + (1 - T - F) W) \right] + \mu_i^2 (F + (1 - T - F) W) \{1 - (F + (1 - T - F) W)\}, \quad i = 1, 2. \tag{14}$$

Mehta *et al.* [1] two- and three-stage procedures deal with the problem of increasing the respondents cooperation in case of highly sensitive question Motivated by Mehta *et al.* [1] two- and three-stages procedures, we proposed two ORRMs. Through proposed models, we plan to improve [1] two- and three-stage models for estimating the mean and sensitivity level.

3. PROPOSE STRATEGIES

In this section, we present the proposed strategies and expressions for unbiased estimators and their minimum variances.

3.1. Proposed Two-stage Additive and Subtractive ORRM

A drawback of [1] two-stage model is that when the respondents scramble their responses, either at first or second stage, the resulting response may be in large magnitude. A typical respondent would hesitate to report a large response because it is thought to be associated with the sensitive variable. To avoid misreporting in such case, we use two scrambling variables in each sub-sample. In the ith sample, at stage 1, a known proportion F of the respondents are directed to scramble their responses as add (subtract) the scrambling variable $Y_i(Z_i)$ to (from) the actual response X. It is anticipated that under the suggested scrambling having a smaller response is more likely. Also it helps fulfilling the social desirability of the respondents. And at stage 2, remaining proportion $(1 - F)$ of respondents are provided the ORRM as

- Report the true response X, if you feel the question insensitive
- Report scrambled response $X + Y_i - Z_i$, if you feel the question sensitive

In the ith sub-sample the distribution of reported response V_i is given as

$$V_i = \begin{cases} X & \text{with probability } (1 - F)(1 - W) \\ (X + Y_i - Z_i) & \text{with probability } F + (1 - F) W. \end{cases} \tag{15}$$

The reported response in the ith sub-sample is written as

$$V_i = \beta (X + Y_i - Z_i) + (1 - \beta) \{\alpha (X + Y_i - Z_i) + (1 - \alpha) X\}. \tag{16}$$

The expected responses in first and second sub-samples are given as

$$E(V_1) = \mu_X + (\mu_1 - 1) \{F + (1 - F) W\} \tag{17}$$

$$E(V_2) = \mu_X + (\mu_2 - 1)\{F + (1-F)W\}. \tag{18}$$

On solving (17) and (18), we have

$$\begin{aligned} \mu_X &= \frac{E(V_2)(\mu_1 - 1) - E(V_1)(\mu_2 - 1)}{\mu_1 - \mu_2}, \quad \mu_1 \neq \mu_2 \\ W &= \frac{1}{(1-F)}\left\{\frac{E(V_1) - E(V_2)}{\mu_1 - \mu_2} - F\right\}, \quad \mu_1 \neq \mu_2, \quad F \neq 1. \end{aligned} \tag{19}$$

The unbiased estimators of mean μ_X and sensitivity level W are given by

$$\begin{aligned} \hat{\mu}_{XMI} &= \frac{\overline{V}_2(\mu_1 - 1) - \overline{V}_1(\mu_2 - 1)}{\mu_1 - \mu_2}, \quad \mu_1 \neq \mu_2 \\ \hat{W}_{MI} &= \frac{1}{(1-F)}\left\{\frac{\overline{V}_2 - \overline{V}_1}{\mu_1 - \mu_2} - F\right\}, \quad \mu_1 \neq \mu_2, \quad F \neq 1. \end{aligned} \tag{20}$$

The variances of the estimators $\hat{\mu}_{XMI}$ and $\hat{W}_{MI}$ are given as

$$Var(\hat{\mu}_{XMI}) = \frac{1}{(\mu_1 - \mu_2)^2}\left[(\mu_2 - 1)^2\left(\frac{\sigma^2_{V_1}}{n_1}\right) + (\mu_1 - 1)^2\left(\frac{\sigma^2_{V_2}}{n_2}\right)\right], \qquad \mu_1 \neq \mu_2 \tag{21}$$

$$Var\left(\hat{W}_{MI}\right) = \frac{1}{(1-F)^2(\mu_1 - \mu_2)^2}\left[\left(\frac{\sigma^2_{V_1}}{n_1}\right) + \left(\frac{\sigma^2_{V_2}}{n_2}\right)\right], \quad \mu_1 \neq \mu_2, \quad F \neq 1, \tag{22}$$

where,
$$\begin{aligned} \sigma^2_{V_i} &= \sigma^2_X + (F + (1-F)W)\left(\delta_i^2 + \gamma_i^2\right) \\ &\quad + (\mu_i - 1)^2 (F + (1-F)W)\left[1 - (F + (1-F)W)\right], \qquad i = 1, 2. \end{aligned} \tag{23}$$

To calculate optimum sub-sample sizes, we define a linear combination of $Var(\hat{\mu}_{XMI})$ and $Var\left(\hat{W}_{MI}\right)$ and minimize it subject to the restriction that $n_2 + n_2 = n$. Now consider,

$$\begin{aligned} Var\left(\hat{\mu}_{XMI}, \hat{W}_{MI}\right) &= \left[Var(\hat{\mu}_{XMI}) + Var\left(\hat{W}_{MI}\right)\right] - \lambda(n_1 + n_2 - n) \\ &= \frac{1}{(\mu_1 - \mu_2)^2(1-F)^2}\left[\frac{\left[(\mu_2 - 1)^2(1-F)^2 + 1\right]\sigma^2_{V_1}}{n_1}\right] \\ &\quad + \frac{1}{(\mu_1 - \mu_2)^2(1-F)^2}\left[\frac{\left[(\mu_1 - 1)^2(1-F)^2 + 1\right]\sigma^2_{V_2}}{n_2}\right] \\ &\quad - \lambda(n_1 + n_2 - n). \end{aligned} \tag{24}$$

Solving $\frac{\partial(\hat{\mu}_{XMI}, \hat{W}_{MI})}{\partial(n_i)} = 0$, we get

$$n_1 = \frac{n\sigma_{V_1}\sqrt{(\mu_2 - 1)^2(1-F)^2 + 1}}{\sigma_{V_2}\sqrt{(\mu_1 - 1)^2(1-F)^2 + 1} + \sigma_{V_1}\sqrt{(\mu_2 - 1)^2(1-F)^2 + 1}}, \tag{25}$$

and

$$n_2 = \frac{n\sigma_{V_2}\sqrt{(\mu_1 - 1)^2(1-F)^2 + 1}}{\sigma_{V_2}\sqrt{(\mu_1 - 1)^2(1-F)^2 + 1} + \sigma_{V_1}\sqrt{(\mu_2 - 1)^2(1-F)^2 + 1}}. \tag{26}$$

Thus, expressions for the minimum variances of $\hat{\mu}_{XMI}$ and $\hat{W}_{MI}$ are given by

$$\begin{aligned} Var(\hat{\mu}_{XMI})_{min} &= \frac{\left(\sigma_{V_2}\sqrt{(\mu_1 - 1)^2(1-F)^2 + 1} + \sigma_{V_1}\sqrt{(\mu_2 - 1)^2(1-F)^2 + 1}\right)}{n(\mu_1 - \mu_2)^2} \\ &\quad \left\{\frac{\sigma_{V_2}(\mu_1 - 1)^2}{\sqrt{(\mu_1 - 1)^2(1-F)^2 + 1}} + \frac{\sigma_{V_1}(\mu_2 - 1)^2}{\sqrt{(\mu_2 - 1)^2(1-F)^2 + 1}}\right\}, \end{aligned} \tag{27}$$

and

$$Var(\hat{W}_{MI})_{min} = \frac{\left(\sigma_{V_2}\sqrt{(\mu_1-1)^2(1-F)^2+1}+\sigma_{V_1}\sqrt{(\mu_2-1)^2(1-F)^2+1}\right)}{n(\mu_1-\mu_2)^2} \left\{\frac{\sigma_{V_2}}{\sqrt{(\mu_1-1)^2(1-F)^2+1}}+\frac{\sigma_{V_1}}{\sqrt{(\mu_2-1)^2(1-F)^2+1}}\right\}. \quad (28)$$

3.2. Proposed Three-stage Additive and Subtractive ORRM

The proposed model is an extension of [1] three-stage model. The main drawback of [1] three-stage model is that respondent's cooperation is low due to the use of additive scrambling variable in second and third stages. Therefore, it is desirable to use a new scrambling variable which is subtracted from the scrambled response of [1] three-stage model, which in turn, helps increasing the respondents' cooperation. Two independent sub-samples of size n_1 and n_2 are drawn from the population such that $n_1+n_2=n$. In each sub-sample, a fixed predetermined proportion T of respondents is instructed to tell the truth and a fixed predetermined proportion F of respondents is instructed to scramble their response. The remaining proportion $(1-T-F)$ of interviewees are directed to use ORRM as

- Report the true response X, if you feel the question insensitive
- Report scrambled response $X+Y_i-Z_i$, if you feel the question sensitive

In the ith sub-sample the distribution of reported response V_i^* is given by

$$V_i^* = \begin{cases} X & \text{with probability } T+(1-T-F)(1-W) \\ (X+Y_i-Z_i) & \text{with probability } F+(1-T-F)W. \end{cases} \quad (29)$$

The reported response in the ith sub-sample may also be written as

$$V_i^* = X\eta + \beta(X+Y_i-Z_i) + (1-\eta-\beta)\{\alpha(X+Y_i-Z_i)+(1-\alpha)X\}. \quad (30)$$

The expected responses from the first and second sub-samples are given by

$$E\left(V_1^*\right) = \mu_X + (\mu_1-1)\{F+(1-T-F)W\} \quad (31)$$

$$E\left(V_2^*\right) = \mu_X + (\mu_2-1)\{F+(1-T-F)W\}. \quad (32)$$

On simplifying (31) and (32), we have

$$\mu_X = \frac{E\left(V_2^*\right)(\mu_1-1) - E\left(V_1^*\right)(\mu_2-1)}{\mu_1-\mu_2}, \quad \mu_1 \neq \mu_2 \quad (33)$$

$$W = \frac{1}{(1-T-F)}\left\{\frac{E\left(V_1^*\right)-E\left(V_2^*\right)}{\mu_1-\mu_2} - F\right\}, \quad \mu_1 \neq \mu_2, \quad T+F \neq 1. \quad (34)$$

The unbiased estimators of mean and sensitivity level μ_X and W are given by

$$\hat{\mu}_{XMII} = \frac{\overline{V}_2^*(\mu_1-1) - \overline{V}_1^*(\mu_2-1)}{\mu_1-\mu_2}, \quad \mu_1 \neq \mu_2 \quad (35)$$

$$\hat{\mu}_{MII} = \frac{1}{(1-T-F)}\left\{\frac{\overline{V}_1^* - \overline{V}_2^*}{\mu_1-\mu_2} - F\right\}, \quad \mu_1 \neq \mu_2, \quad T+F \neq 1. \quad (36)$$

The variances of the estimators in (35) and (36) are given by

$$Var(\hat{\mu}_{XMII}) = \frac{1}{(\mu_1-\mu_2)^2}\left[(\mu_2-1)^2\left(\frac{\sigma^2_{V_1^*}}{n_1}\right) + (\mu_1-1)^2\left(\frac{\sigma^2_{V_2^*}}{n_2}\right)\right], \quad \mu_1 \neq \mu_2, \quad (37)$$

$$Var\left(\hat{W}_{MII}\right) = \frac{1}{(1-T-F)^2(\mu_1-\mu_2)^2}\left[\left(\frac{\sigma^2_{V_1^*}}{n_1}\right)+\left(\frac{\sigma^2_{V_2^*}}{n_2}\right)\right], \quad \mu_1 \neq \mu_2, \quad T+F \neq 1. \tag{38}$$

where,
$$\sigma^2_{V_i^*} = \sigma^2_X + (F+(1-T-F+2TF)\,W)\left(\delta_i^2+\gamma_i^2\right) + (\mu_i-1)^2\left[(F+(1-T-F)\,W)(1-(F+(1-T-F)\,W))+2WTF\right], \quad i=1,2. \tag{39}$$

We define a linear combination of $Var\left(\hat{\mu}_{XMII}\right)$ and $Var\left(\hat{\mu}_{XMII}\right)$ in order to find the optimum allocation of sample sizes. Consider,

$$\begin{aligned} Var\left(\hat{\mu}_{XMII}, \hat{W}_{MII}\right) &= \left[Var\left(\hat{\mu}_{XMII}\right) + Var\left(\hat{W}_{MII}\right)\right] - \lambda\left(n_1+n_2-n\right) \\ &= \frac{1}{(\mu_1-\mu_2)^2(1-T-F)^2}\left[\frac{\left[(\mu_2-1)^2(1-T-F)^2+1\right]\sigma^2_{V_1^*}}{n_1}\right] \\ &\quad + \frac{1}{(\mu_1-\mu_2)^2(1-T-F)^2}\left[\frac{\left[(\mu_1-1)^2(1-T-F)^2+1\right]\sigma^2_{V_2^*}}{n_2}\right] \\ &\quad - \lambda\left(n_1+n_2-n\right). \end{aligned} \tag{40}$$

Solving $\frac{\partial(\hat{\mu}_{XMII}, \hat{W}_{MII})}{\partial(n_i)} = 0$, we obtain

$$n_1 = \frac{n\sigma_{V_1^*}\sqrt{(\mu_2-1)^2(1-T-F)^2+1}}{\sigma_{V_2^*}\sqrt{(\mu_1-1)^2(1-T-F)^2+1}+\sigma_{V_1^*}\sqrt{(\mu_2-1)^2(1-T-F)^2+1}}, \tag{41}$$

and

$$n_2 = \frac{n\sigma_{V_2^*}\sqrt{(\mu_1-1)^2(1-T-F)^2+1}}{\sigma_{V_2^*}\sqrt{(\mu_1-1)^2(1-T-F)^2+1}+\sigma_{V_1^*}\sqrt{(\mu_2-1)^2(1-T-F)^2+1}}. \tag{42}$$

Thus, expressions for the minimum variances of $\hat{\mu}_{XMII}$ and $\hat{W}_{MII}$ are given by

$$Var(\hat{\mu}_{XMII})_{min} = \frac{\left(\sigma_{V_2^*}\sqrt{(\mu_1-1)^2(1-T-F)^2+1}+\sigma_{V_1^*}\sqrt{(\mu_2-1)^2(1-T-F)^2+1}\right)}{n(\mu_1-\mu_2)^2}\left\{\frac{\sigma_{V_2^*}(\mu_1-1)^2}{\sqrt{(\mu_1-1)^2(1-T-F)^2+1}}+\frac{\sigma_{V_1^*}(\mu_2-1)^2}{\sqrt{(\mu_2-1)^2(1-T-F)^2+1}}\right\}, \tag{43}$$

and

$$Var(\hat{W}_{MII})_{min} = \frac{\left(\sigma_{V_2^*}\sqrt{(\mu_1-1)^2(1-T-F)^2+1}+\sigma_{V_1^*}\sqrt{(\mu_2-1)^2(1-T-F)^2+1}\right)}{n(\mu_1-\mu_2)^2}\left\{\frac{\sigma_{V_2^*}}{\sqrt{(\mu_1-1)^2(1-T-F)^2+1}}+\frac{\sigma_{V_1^*}}{\sqrt{(\mu_2-1)^2(1-T-F)^2+1}}\right\}. \tag{44}$$

4. SIMULATION STUDY

In this section, we discuss the simulated results of suggested models and compared the proposed strategies with [1] two- and three-stage models. As it clear from (5), (6), (12), (13), (21), (22), (37) and (38) that algebraic comparison of variances deriving efficiency conditions is difficult. Therefore, to know the relative performance of proposed estimators we compare them with [1] two- and three-stage models, numerically. We assume that $Y_1(Y_2)$ and $Z_1(Z_2)$ follow Poisson distribution with parameters 2(5) and 1(1). That is, $\mu_1 = 2$, $\mu_2 = 5$ and $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$. The sensitive variable X is assumed to be normal with mean $\mu_X = 4$ and variance $\sigma^2_X = 4$. The $PRE_{\mu_{MI}}$ and $PRE_{W_{MI}}$ are the percentage relative efficiencies of mean and sensitivity level estimators relative to mean and sensitivity level estimators by Mehta *et al.* [1] two-stage model defined as $PRE_{\mu_{MI}} = \frac{Var(\hat{\mu}_{XM1})}{Var(\hat{\mu}_{XMI})} \times 100$ and $PRE_{W_{MI}} = \frac{Var(\hat{W}_{M1})}{Var(\hat{W}_{MI})} \times 100$. The simulated estimates and Percent Relative Efficiency (PRE)s of proposed two-stage ORRM are arranged in Tables 1 and 2. From the Tables 1 and 2, it is observed that when the level of W and F is increased the $PRE_{\mu_{MI}}$ and $PRE_{W_{MI}}$ increase up to a certain level then gradually decrease. The Table 2 shows the improved performance of the proposed two-stage model for different values of the parameters. The $PRE_{\mu_{MII}}$ and $PRE_{W_{MII}}$ are the percentage relative efficiencies of mean and sensitivity level estimators relative to mean and sensitivity level estimators by Mehta *et al.* [1] three-stage model, defined as $PRE_{\mu_{MII}} = \frac{Var(\hat{\mu}_{XM2})}{Var(\hat{\mu}_{XMII})} \times 100$ and $PRE_{W_{MII}} = \frac{Var(\hat{W}_{M2})}{Var(\hat{W}_{MII})} \times 100$. The simulated estimates and PREs of proposed three-stage ORRM are arranged in Tables 3 and 4. In three-stage ORRM, as the level of W, T are F is increased the $PRE_{\mu_{MII}}$ and $PRE_{W_{MII}}$ increase up to a certain level then gradually decrease. The results in Table 4 show the improved performance of the proposed strategy. A comparison is also made among the [1] two-, three-stage and proposed ORRMs, to find the best estimator. From the results in Tables 5 and 6, it is observed that proposed two-stage ORRM performed better than the other three models ([1] two- and three-stage models and proposed three-stage ORRM), in estimating both the mean and the sensitivity level. From the Table 6, it is noticed that when the values of parameters are changed the performance of proposed two-stage ORRM is improved further, in estimating both the mean and sensitivity level.

Table 1 Simulated estimates and PREs of the estimators $\hat{\mu}_{XMI}$ and $\hat{W}_{MI}$ relative to $\hat{\mu}_{XM1}$ and $\hat{W}_{M1}$ for $\mu_X = 4$, $\sigma_X^2 = 4$, $\mu_1 = 2$, $\mu_2 = 5$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 2$, $\delta_2^2 = 5$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$.

n	F	W	$\hat{\mu}_{XM1}$	$\hat{W}_{M1}$	$\hat{\mu}_{XMI}$	$\hat{W}_{MI}$	PRE_{μ_I}	PRE_{W_I}
	0.15	0.20	3.9929	0.3026	3.9944	0.3026	177.4862	104.2773
		0.30	3.9880	0.3031	3.9940	0.3024	177.5879	104.8565
100	0.25	0.40	3.9936	0.4045	3.9954	0.4058	177.9600	104.2666
		0.50	3.9878	0.5044	3.9894	0.5046	177.6040	104.1755
	0.35	0.70	4.0156	0.6921	4.0100	0.6922	170.0834	101.9307
		0.90	4.0148	0.8937	4.0096	0.8935	166.7103	100.3575
	0.15	0.20	4.0063	0.1978	4.0031	0.1981	178.6253	104.9768
		0.30	4.0024	0.2987	4.0007	0.2991	179.0435	105.4029
500	0.25	0.40	3.9988	0.3998	3.9964	0.4010	179.6554	104.4616
		0.50	3.9962	0.5007	3.9954	0.5012	175.1750	103.2359
	0.35	0.70	3.9907	0.7015	3.9923	0.7015	172.4169	101.4381
		0.90	3.9927	0.9027	3.9935	0.9031	167.2681	100.5399
	0.15	0.20	4.0014	0.1989	4.0002	0.1991	178.7889	104.9573
		0.30	3.9972	0.3006	3.9973	0.3006	178.5654	105.1008
1000	0.25	0.40	4.0013	0.3995	4.0006	0.3997	177.8660	104.9161
		0.50	4.0048	0.4976	4.0042	0.4970	179.1386	104.5887
	0.35	0.70	4.0004	0.7001	4.0006	0.6999	171.9624	100.1390
		0.90	3.9981	0.8995	3.9985	0.8991	167.6696	100.0602

Table 2 Simulated estimates and PREs of the estimators $\hat{\mu}_{XMI}$ and $\hat{W}_{MI}$ relative to $\hat{\mu}_{XM1}$ and $\hat{W}_{M1}$ for $\mu_X = 6$, $\sigma_X^2 = 3$, $\mu_1 = 3$, $\mu_2 = 6$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 3$, $\delta_2^2 = 6$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$.

n	F	W	$\hat{\mu}_{XM1}$	$\hat{W}_{M1}$	$\hat{\mu}_{XMI}$	$\hat{W}_{MI}$	PRE_{μ_I}	PRE_{W_I}
	0.15	0.20	6.0103	0.1969	6.0071	0.1961	172.8143	110.0909
		0.30	6.0148	0.2971	6.0194	0.2954	173.5834	110.6993
100	0.25	0.40	6.0107	0.3973	6.0087	0.3968	172.0078	108.7325
		0.50	6.0030	0.5009	6.0006	0.5019	169.6038	107.4458
	0.35	0.70	5.9934	0.7016	5.9975	0.7002	159.5388	104.0838
		0.90	5.9988	0.9029	6.0034	0.9029	151.7586	100.7551
	0.15	0.20	5.9975	0.2001	5.9970	0.2004	173.3019	109.9578
		0.30	5.9993	0.3001	5.9985	0.3003	172.6932	110.1628
500	0.25	0.40	5.9969	0.4002	5.9959	0.4003	171.9284	108.8302
		0.50	6.0116	0.4965	6.0091	0.4966	168.8598	107.7519
	0.35	0.70	5.9985	0.7004	5.9997	0.6999	160.3861	102.1692
		0.90	6.0003	0.8996	5.9985	0.9005	152.3747	100.7961
	0.15	0.20	5.9973	0.2006	5.9979	0.2008	173.4654	110.5120
		0.30	6.0033	0.2990	6.0015	0.2994	172.9812	109.9680
1000	0.25	0.40	5.9968	0.4022	5.9991	0.4011	173.7487	110.1246
		0.50	5.9960	0.5008	5.9963	0.5009	171.0512	108.7892
	0.35	0.70	6.0007	0.6996	6.0004	0.6996	160.4707	103.0604
		0.90	6.0007	0.8994	6.0010	0.8991	150.4549	101.6957

Table 3 Simulated estimates and PREs of the estimators $\hat{\mu}_{XMII}$ and $\hat{W}_{MII}$ relative to $\hat{\mu}_{XM2}$ and $\hat{W}_{M2}$ for $\mu_X = 4$, $\sigma_X^2 = 4$, $\mu_1 = 2$, $\mu_2 = 5$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 2$, $\delta_2^2 = 5$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$.

n	T	F	W	$\hat{\mu}_{XM2}$	$\hat{W}_{M2}$	$\hat{\mu}_{XMII}$	$\hat{W}_{MII}$	$PRE_{\mu_{II}}$	$PRE_{W_{II}}$
	0.20	0.15	0.20	3.9919	0.3038	3.9946	0.3033	178.4960	105.0356
			0.30	3.9982	0.3008	3.9983	0.3012	178.4229	105.3644
100	0.30	0.25	0.40	3.9936	0.4056	3.9968	0.4067	177.9101	104.3087
			0.50	3.9864	0.5079	3.9911	0.5066	179.0506	104.9941
	0.40	0.35	0.70	4.0184	0.6819	4.0114	0.6843	177.5183	104.7576
			0.90	4.0185	0.8788	4.0113	0.8810	176.8830	104.0436
	0.20	0.15	0.20	4.0081	0.1960	4.0050	0.1961	177.8733	104.5428
			0.30	4.0008	0.2993	3.9999	0.2994	178.4259	105.0207
500	0.30	0.25	0.40	3.9940	0.4035	3.9963	0.4033	180.2129	104.9297
			0.50	3.9969	0.5009	3.9960	0.5015	178.2884	105.0409
	0.40	0.35	0.70	3.9873	0.7096	3.9893	0.7111	180.2989	105.7806
			0.90	3.9894	0.9112	3.9919	0.9118	178.1306	104.7578
	0.20	0.15	0.20	4.0006	0.1995	3.9999	0.1997	178.8017	104.8441
			0.30	3.9966	0.3009	3.9972	0.3009	180.7398	105.5225
1000	0.30	0.25	0.40	4.0008	0.3992	4.0005	0.3993	178.9036	105.3841
			0.50	4.0062	0.4957	4.0040	0.4956	180.0386	105.4599
	0.40	0.30	0.70	3.9996	0.6997	3.9996	0.6996	177.2731	104.0512
			0.90	4.0001	0.8984	3.9994	0.8974	179.6923	105.1635

Table 4 Simulated estimates and PREs of the estimators $\hat{\mu}_{XMII}$ and $\hat{W}_{MII}$ relative to $\hat{\mu}_{XM2}$ and $\hat{W}_{M2}$ for $\mu_X = 6$, $\sigma_X^2 = 3$, $\mu_1 = 3$, $\mu_2 = 6$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 3$, $\delta_2^2 = 6$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$.

n	T	F	W	$\hat{\mu}_{XM2}$	$\hat{W}_{M2}$	$\hat{\mu}_{XMII}$	$\hat{W}_{MII}$	$PRE_{\mu_{II}}$	$PRE_{W_{II}}$
	0.20	0.15	0.20	6.0018	0.2004	6.0011	0.2008	172.4898	109.7678
			0.30	6.0227	0.2932	6.0183	0.2934	172.1270	109.5257
100	0.30	0.25	0.40	6.0028	0.3975	6.0005	0.3979	173.3882	109.8600
			0.50	5.9823	0.7167	5.9881	0.7133	172.0590	109.5446
	0.40	0.35	0.70	6.0094	0.6887	6.0058	0.6898	172.5855	109.6670
			0.90	6.0107	0.8945	6.0086	0.8963	170.8559	108.1835
	0.20	0.15	0.20	5.9993	0.1993	5.9971	0.2003	171.7968	109.5483
			0.30	6.0010	0.2991	5.9995	0.2996	173.1462	110.0219
500	0.30	0.25	0.40	5.9932	0.4013	5.9946	0.4010	173.9630	110.4000
			0.50	5.9999	0.5005	6.0020	0.4993	173.2602	110.0961
	0.40	0.35	0.70	5.9953	0.7042	5.9970	0.7035	172.5629	109.0104
			0.90	5.9999	0.9008	5.9993	0.9017	172.6907	109.7657
	0.20	0.15	0.20	5.9991	0.2006	5.9996	0.2005	172.2561	109.9642
			0.30	6.0016	0.2994	6.0014	0.2994	173.1791	109.8318
1000	0.30	0.25	0.40	5.9991	0.4003	5.9988	0.4005	173.5312	110.1113
			0.50	5.9987	0.4995	5.9984	0.4997	173.3324	110.0153
	0.40	0.35	0.70	6.0020	0.6962	6.0005	0.6970	172.4775	109.8343
			0.90	6.0000	0.8981	5.9995	0.8987	171.3794	108.7687

Table 5 Simulated PREs of estimators $\hat{\mu}_{MI}$ and $\hat{W}_{MI}$ relative to $\hat{\mu}_{M1}$, $\hat{\mu}_{M2}$, $\hat{\mu}_{MII}$ and $\hat{W}_{M1}$, $\hat{\mu}_{M2}$ and $\hat{W}_{MII}$ for $\mu_X = 4$, $\sigma_X^2 = 4$, $\mu_1 = 2$, $\mu_2 = 5$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 2$, $\delta_2^2 = 5$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$.

n	T	F	W	$PRE_{\mu_{I,1}}$	$PRE_{\mu_{I,2}}$	$PRE_{\mu_{I,II}}$	$PRE_{W_{I,1}}$	$PRE_{W_{II,2}}$	$PRE_{W_{I,II}}$
	0.20	0.15	0.20	177.4862	177.4990	99.4414	104.2773	178.5148	169.9564
			0.30	177.5879	176.8762	99.1331	104.8566	176.9275	167.9195
100	0.30	0.25	0.40	177.9600	173.2325	97.3707	104.2666	281.0522	269.4426
			0.50	177.6049	173.6579	96.9888	104.1755	278.3475	265.1077
	0.40	0.35	0.70	170.0834	172.4773	97.1602	99.93070	694.4239	662.8861
			0.90	166.7103	168.5491	95.2884	97.35750	696.9869	669.8984

(*continued*)

Table 5 Simulated PREs of estimators $\hat{\mu}_{MI}$ and $\hat{W}_{MI}$ relative to $\hat{\mu}_{M1}$, $\hat{\mu}_{M2}$, $\hat{\mu}_{MII}$ and $\hat{W}_{M1}$, $\hat{\mu}_{M2}$ and $\hat{W}_{MII}$ for $\mu_X = 4$, $\sigma_X^2 = 4$, $\mu_1 = 2$, $\mu_2 = 5$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 2$, $\delta_2^2 = 5$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$. *(Continued)*

n	T	F	W	$PRE_{\mu_{I,1}}$	$PRE_{\mu_{I,2}}$	$PRE_{\mu_{I,II}}$	$PRE_{W_{I,1}}$	$PRE_{W_{II,2}}$	$PRE_{W_{I,II}}$
	0.20	0.15	0.20	178.6253	176.6777	99.3278	104.9768	176.4362	168.7692
			0.30	179.0435	176.0226	98.6530	105.4029	177.1385	168.6700
500	0.30	0.25	0.40	179.6554	174.9141	97.0598	104.4616	279.2598	266.1397
			0.50	175.1750	175.7303	98.5651	103.2359	288.8495	274.9876
	0.40	0.35	0.70	172.4169	173.4596	96.2066	101.4381	705.4966	666.9432
			0.90	167.2681	172.4673	96.8201	97.53990	722.5723	689.7549
	0.20	0.15	0.20	178.7889	176.9410	98.9593	104.9573	178.4408	170.1952
			0.30	178.5654	178.0594	98.5169	105.1008	178.7083	169.3556
1000	0.30	0.25	0.40	177.8660	176.5509	98.6849	104.9161	288.8816	274.1207
			0.50	179.1386	177.4479	98.5610	104.5887	288.2034	273.2823
	0.40	0.35	0.70	171.9624	168.9254	95.2909	100.1390	680.8597	654.3446
			0.90	167.6698	175.0242	97.4021	97.36021	720.6047	685.2230

Table 6 Simulated PREs of estimators $\hat{\mu}_{MI}$ and $\hat{W}_{MI}$ relative to $\hat{\mu}_{M1}$, $\hat{\mu}_{M2}$, $\hat{\mu}_{MII}$ and $\hat{W}_{M1}$, $\hat{\mu}_{M2}$ and $\hat{W}_{MII}$ for $\mu_X = 6$, $\sigma_X^2 = 3$, $\mu_1 = 3$, $\mu_2 = 6$, $\mu_{Z_1} = 1$, $\mu_{Z_2} = 1$, $\delta_1^2 = 3$, $\delta_2^2 = 6$, $\gamma_1^2 = 1$ and $\gamma_2^2 = 1$.

n	T	F	W	$PRE_{\mu_{I,1}}$	$PRE_{\mu_{II,2}}$	$PRE_{\mu_{I,II}}$	$PRE_{W_{I,1}}$	$PRE_{W_{II,2}}$	$PRE_{W_{I,II}}$
	0.20	0.15	0.20	172.8143	170.0507	98.5859	110.0909	183.9822	167.6103
			0.30	173.5834	162.7850	94.5725	110.6993	177.1305	161.7249
100	0.30	0.25	0.40	172.0078	165.1365	95.2409	108.7325	292.5501	266.2934
			0.50	169.6038	163.5106	94.6306	107.4458	286.3443	262.0446
	0.40	0.35	0.70	159.5388	166.6039	96.8294	102.0838	731.7065	667.9530
			0.90	151.7586	167.6940	98.1494	96.75551	772.0319	713.6313
	0.20	0.15	0.20	173.3019	169.2371	98.5100	109.9578	181.6347	165.8033
			0.30	172.6932	170.7100	98.5931	110.1628	186.6265	169.6266
500	0.30	0.25	0.40	171.9284	168.0147	96.5807	108.8302	293.8508	266.1692
			0.50	168.8598	169.5892	97.8812	107.7519	299.1032	271.6746
	0.40	0.35	0.70	160.3861	164.9760	95.6033	102.1692	721.8251	662.1617
			0.90	152.3747	168.3917	97.5105	97.79614	775.7312	706.7153
	0.20	0.15	0.20	173.4654	166.5723	96.7003	110.5120	181.4366	164.9953
			0.30	173.1793	168.8176	96.9067	109.8318	184.1052	166.5418
1000	0.30	0.25	0.40	172.2482	166.6364	96.0267	109.3727	291.0022	264.2805
			0.50	171.0512	165.3644	95.4030	108.7892	291.8981	265.3249
	0.40	0.35	0.70	160.4707	167.7059	97.2334	103.0605	753.0921	685.6615
			0.90	150.4549	168.1901	98.1390	96.69570	779.4595	716.6210

5. CONCLUSION

In the present study, the proposed two- and three-stage ORRMs is found to be more efficient than the [1] two- and three-stage models. The estimators of mean and sensitivity level obtained from the suggested models are unbiased. The proposed ORRMs provide better estimators of mean in terms of percentage relative efficiency. The proposed two-stage ORRM are compared with [1] two- and three-stage models and proposed three-stage ORRM to obtain the most efficient model. From simulation study, we observed that two-stage ORRM performs better than the others in terms of percentage relative efficiency. Conclusively, we can say that proposed two-stage ORRM model is superior to the others. For the future work, the proposed models can be extended for other sampling schemes such as stratified random sampling, probability proportional to size and without replacement sampling, and multi-stage sampling, etc.

CONFLICT OF INTEREST

There is no conflict of interest between the authors.

AUTHORS' CONTRIBUTIONS

The idea of this work was conceived by Zawar Hussain. He also performed numerical analysis. The initial draft of the paper was written by Muhammad Imran Shahid and final draft was prepared by Zawar Hussain.

REFERENCES

1. S. Mehta, B. Dass, J. Shabbir, S. Gupta, J. Stat. Theory Pract. 6 (2012), 417–427.
2. S.L. Warner, J. Am. Stat. Assoc. 60 (1965), 63–69.
3. B.G. Greenberg, R.R. Kuebler Jr, J.R. Abernathy, D.G. Horvitz, J. Am. Stat. Assoc. 66 (1971), 243–250.
4. K. Pollock, Y. Bek, J. Am. Stat. Assoc. 71 (1976), 884–886.
5. B.H. Eichhorn, L.S. Hayre, J. Stat. Plan. Inference. 7 (1983), 307–316.
6. S. Gupta, B. Gupta, S. Singh, J. Stat. Plan. inference. 100 (2002), 239–247.
7. S. Gupta, B. Thornton, Am. J. Math. Manage. Sci. 22 (2002), 369–383.
8. S. Gupta, J. Shabbir, Statistica. 64 (2004), 643–653.
9. C.R. Gjestvang, S. Singh, J. R. Stat. Soc. Ser. B Stat. Methodol. 68 (2006), 523–530.
10. K.-C. Huang, Qual. Quant. 42 (2008), 679–686.
11. C.R. Gjestvang, S. Singh, J. Appl. Stat. 36 (2009), 1361–1367.
12. O. Odumade, S. Singh, Commun. Stat. Simul. Comput. 38 (2009), 473–502.
13. K.-C. Huang, Metrika. 71 (2010), 341–352.
14. S. Gupta, J. Shabbir, S. Sehra, J. Stat. Plan. Inference. 140 (2010), 2870–2874.
15. Z. Hussain, K. Khan, World Appl. Sci. J. 23 (2013), 1201–1206.
16. S. Gupta, S. Mehta, J. Shabbir, B. Dass, Commun. Stat. Theory Methods. 42 (2013), 4034–4042.