Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 75(3) 2021 195–204

Randomized Response Trial using Geometric Distribution

Raghunath Arnab
University of KwaZulu-Natal, South Africa

Received 27 January 2020; Revised 07 January 2022; Accepted 18 January 2022

SUMMARY

Randomized response (RR) technique is used to collect data relating to sensitive issues. RR using the Geometric distribution was proposed by Singh and Grewal (2013) and showed empirically that their method performs better than the Kuk's (1990) model. In this paper, it is shown that the comparison proposed by Singh and Grewal was not fair. A more realistic comparison is proposed in this paper and it is found that the Kuk's model performs better than the Singh and Grewal's model in most situations.

Keywords: Estimation of proportion; Geometric distribution; Randomized response; Relative efficiency.

1. INTRODUCTION

In surveys relating to sensitive issues such as induced abortion, drug use, sexual behaviour, suffering from HIV/AIDS, tax evasion among others, the direct method of interview is not suitable because a high percentage of respondents report untrue responses or even refuse to respond because of social stigma and or fear (Arnab, 2017). Interviewers also feel embarrassed to ask such sensitive and personal information. In order to improve the quality of data, protect respondents' privacy and reduce nonresponse rate in surveys, Warner (1965) proposed an indirect method of interview procedure known as Randomized Response (RR) technique. The Warner (1965) RR technique was extended by various researchers including Horvitz *et al.* (1967), Greenberg *et al.* (1969), Kim (1978), Arnab (1990, 1996), Kuk (1990), Christofides (2003), Kim and Warde (2004), Arnab and Singh (2007, 2014), Chaudhuri and Dihidar (2014), Rueda *et al.* (2015), Arcos *et al.* (2015) for further improvement of respondents' co-operations and enhancing efficiencies of the estimates of the population characteristics of interest. Applications of the RR methods in real live surveys were provided by Van der Heijden *et al.* (1998) and Arnab and Mothupi (2015). A comprehensive review of the RR techniques is given by Chaudhuri and Mukerjee (1988) and Arnab (2017).

Recently, Singh and Grewal (2013) proposed the geometric distribution in RR surveys and claimed that their method performed better than the Kuk's (1990) model. In this paper, it is shown that their comparison is not fair enough as their method is based on larger number of trials than that of Warner's (1965) and Kuk's (1990) RR method.

In this paper a new method of comparing efficiencies of the different randomized response (RR) techniques are proposed. In the existing method, efficiencies of the RR techniques are compared by comparing magnitudes of the variances of the estimators of the parameter of interest for the different RR techniques. In the proposed method, the efficiencies of the RR techniques are compared by comparing the magnitude of the variances of the estimators keeping the expected number of trials γ fixed to a certain level.

For example, in the Warner's (1965) model, a respondent performs only one trial (draws only one card from a pack i. e $\gamma=1$). In Kuk's (1990) model, a respondent draws a fixed number $c\ (\geq 1)$ cards while in Sing and Grewal's (2013) model, the number of

E-mail address: arnab.raghunath@gmail.com

cards draws by a respondent is not fixed, it is a random variable that follows a geometric distribution.

Under the proposed method of comparison, Kuk's RR model is compared with Warner's model by comparing variance of the estimator of the Kuk's model based on γ number of trials with the variance of Warner's estimator assuming each respondent performs Warner's trial $\gamma(=c)$ number of times independently. For the comparison of the efficiency of the Warner's model with the Singh-Grewal's model the variance of Singh and Grewal's estimator with γ expected number of trials is compared with the variance of the Warner's estimator assuming that each respondent repeats the Warner's trial γ number of times.

Some of the RR techniques relevant to the present discussion are given in the sequel.

Consider a finite population $U=\{1,..,i,..,N\}$ of N identifiable units and let y_i be the value of the sensitive characteristic y for the i^{th} unit. We define $y_i=1$ if the i^{th} unit possesses the certain sensitive attribute A such as HIV + and $y_i=0$ if the i^{th} unit does not possesses the attribute i.e the unit belongs to $\overline{A}$, the complement of A. Our objective is to estimate the finite population proportion $\pi=\sum_{i\in U} y_i/N$ using a sample s of size n selected from the population U by using simple random sampling with replacement (SRSWR) method. Since y is a sensitive variable, no direct information of y is available from the respondent and, hence, indirect information relating to y is obtained from each of the selected respondents independently by using any of the following RR techniques.

1.1 Warner's Technique

In Warner's (1965) model, each of the selected respondents is asked to draw one card at random from a pack containing two types of cards. Card type-1, with proportion $P_w(\neq 1/2)$ bears statement "I belong to the sensitive group A" while card type-2, with proportion $1-P_w$ bears the statement "I do not belong to the sensitive group A". Respondents are asked to truthfully answer "Yes" if the statement of the drawn card matches with his/her own status and answer "No" otherwise. Let $z_i=1(0)$ if the i^{th} respondent provides a response "Yes" ("No"). Also, let $E_R(V_R)$ and $E_p(V_p)$ denote the expectation (variance) with respect to RR model and the sampling design p, respectively. Then we have

$$\begin{aligned} E_R(z_i) &= y_iP_w+(1-y_i)(1-P_w) \\ &= (2P_w-1)y_i+(1-P_w) \end{aligned} \tag{1.1.1}$$

and

$$\begin{aligned} V_R(z_i) &= y_iP_w+(1-y_i)(1-P_w)-\{y_iP_w+(1-y_i)(1-P_w)\}^2 \\ &= P_w(1-P_w) \end{aligned} \tag{1.1.2}$$

(noting $y_i=y_i^2=1$ or 0)

Now writing

$$r_i=\frac{z_i-(1-P_w)}{2P_w-1} \tag{1.1.3}$$

one finds that

$$E_R(r_i)=y_i,\ V_R(r_i)=\frac{P_w(1-P_w)}{(2P_w-1)^2}=\phi_w \text{ (say) and}$$

$$C_R(r_i,r_j)=0 \text{ for } i\neq j \tag{1.1.4}$$

where C_R is the co-variance with respect to the RR model.

Let $\overline{r}_w=\frac{1}{n}\sum_{i\in s} r_i=\frac{\overline{z}_w-(1-P_w)}{2P_w-1}$ and $\overline{z}_w=\frac{1}{n}\sum_{i\in s} z_i$ with $\sum_{i\in s}$ denoting the sum-over the units in the sample s with repetition. Then,

$$E(\overline{r}_w)=E_p\left[\frac{1}{n}\sum_{i\in s}E_R(r_i)\right]=\frac{1}{N}\sum_{i\in U}y_i=\pi \tag{1.1.5}$$

and

$$\begin{aligned} \text{Var}(\overline{r}_w) &= V_p[E_R(\overline{r}_w)]+E_p[V_R(\overline{r}_w)] \\ &= V_p\left[\frac{1}{n}\sum_{i\in s}y_i\right]+E_p\left[\frac{\phi_w}{n}\right] \\ &= \frac{\pi(1-\pi)}{n}+\frac{\phi_w}{n} \\ &= V_w \end{aligned} \tag{1.1.6}$$

From (1.1.5) and (1.1.6), we find the following theorem which was originally derived by Warner (1965).

Theorem 1.1.1.

(i) $\overline{r}_w$ is an unbiased estimator of π

(ii) $\text{Var}(\overline{r}_w)=\frac{\pi(1-\pi)}{n}+\frac{\phi_w}{n}$

where $\phi_w=\frac{P_w(1-P_w)}{(2P_w-1)^2}$

(iii) An unbiased estimator of $\text{Var}(\overline{r}_w)$ is

$$\widehat{\text{Var}}(\overline{r}_w)=\frac{\sum_{i\in s}(r_i-\overline{r}_w)^2}{n(n-1)}$$

1.2 Kuk's RR technique

The Kuk (1990) RR technique involves two pack of cards. Each of the respondents belonging to the sensitive group A were asked to draw $c(\geq 1)$ cards at random with replacement from a pack containing θ_1 proportion of red and $1-\theta_1$ proportion of black cards and asked to report the total number of red cards drawn. Similarly, the respondents belonging to the non-sensitive group $\bar{A}$ were asked to draw c cards with replacement from the other pack containing $\theta_2(\neq \theta_1)$ proportion of red cards and $1-\theta_2$ proportion of black cards and were further asked to report the total number of red cards drawn. Let $X_i(Q_i)$ be the number of red cards drawn if the i^{th} person belongs to the sensitive (non-sensitive) group $A(\bar{A})$. Thus, the total number of red cards drawn by the i^{th} respondent can be written as

$$z_i = y_i X_i + (1-y_i)Q_i \tag{1.2.1}$$

From Eq. (1.2.1), one finds that

$$\begin{aligned} E_R(z_i) &= y_i E_R(X_i) + (1-y_i)E_R(Q_i) \\ &= c\{y_i\theta_1 + (1-y_i)\theta_2\} \end{aligned} \tag{1.2.2}$$

and
$$\begin{aligned} V_R(z_i) &= y_i V_R(X_i) + (1-y_i)V_R(Q_i) \\ &= c\{y_i\theta_1(1-\theta_1) + (1-y_i)\theta_2(1-\theta_2)\} \end{aligned} \tag{1.2.3}$$

Now, let $r_i = \dfrac{\bar{z}_i - \theta_2}{\theta_1 - \theta_2}$ with $\bar{z}_i = z_i / c$. Then,

$$E_R(r_i) = y_i, V_R(r_i) = \frac{V_R(\bar{z}_i)}{(\theta_1-\theta_2)^2}$$

$$= \frac{y_i\theta_1(1-\theta_1) + (1-y_i)\theta_2(1-\theta_2)}{c(\theta_1-\theta_2)^2} = \phi_{ki}$$

and $C_R(r_i, r_j) = 0$ for $i \neq j$ (1.2.4)

Now using Theorem 1.1.1, one derives the results in Theorem 1.2.1 which were originally obtained by Kuk (1990).

Theorem 1.2.1

(i) $\bar{r}_k = \dfrac{1}{n}\sum_{i\in s} r_i = \dfrac{\bar{z}-\theta_2}{\theta_1-\theta_2}$ is an unbiased estimator of π

where $\bar{z} = \dfrac{1}{n}\sum_{i\in s}\bar{z}_i$

(ii) $$\text{Var}(\bar{r}_k) = \frac{\pi(1-\pi)}{n} + \frac{1}{nN}\sum_{i\in U}\phi_{ki} = \frac{\pi(1-\pi)}{n} + \frac{\phi_k}{n} = V_k \tag{1.2.5}$$

where $\phi_k = \dfrac{\pi\theta_1(1-\theta_1) + (1-\pi)\theta_2(1-\theta_2)}{c(\theta_1-\theta_2)^2}$

(iii) An unbiased estimator of $\text{Var}(\bar{r}_k)$ is

$$\hat{\text{Var}}(\bar{r}_k) = \frac{\sum_{i\in s}(r_i - \bar{r}_k)^2}{n(n-1)}$$

Remark 1.2.1

Kuk's (1990) RR technique reduces to Warner's (1965) technique if $c=1$, $\theta_1 = P_w$ and $\theta_2 = 1-P_w$.

1.3 Singh and Grewal's RR technique

In the Singh and Grewal's (2013) RR model, each of the selected respondents was given two decks of cards. Each of the deck comprises two types of cards as in Warner (1965) and Kuk (1990) model. Deck-I (Deck-II) comprises cards with proportions $\theta_g(\theta_g^*)$ written statement "I belong to the group A" and the remaining $1-\theta_g$ $(1-\theta_g^*)$ proportion of cards bearing statement "I belong to the group $\bar{A}$". Then, respondents were asked to choose Deck-I (Deck-II) if he/she belongs to the group $A(\bar{A})$ and draw cards one by one with replacement till he/she receives a card with a statement matches with his/ her own status the first time. The respondents were asked to report the total number of cards he/she has drawn as his/her randomized response. Let X_i (Q_i) be the number of cards drawn reported by the ith respondent if he/she belongs to the group $A(\bar{A})$. Then z_{gi}, the RR from the i^{th} respondent can be expressed as

$$z_{gi} = y_i X_i + (1-y_i)Q_i \tag{1.3.1}$$

Since X_i and Q_i follows the Geometric distribution, we have

$$\begin{aligned} E_R(z_{gi}) &= y_i E_R(X_i) + (1-y_i)E_R(Q_i) \\ &= y_i(\theta_g^* - \theta_g)/(\theta_g^*\theta_g) + 1/\theta_g^* \end{aligned} \tag{1.3.2}$$

and

$$\begin{aligned} V_R(z_{gi}) &= y_i V_R(X_i) + (1-y_i)V_R(Q_i) \\ &= y_i\frac{1-\theta_g}{\theta_g^2} + (1-y_i)\frac{1-\theta_g^*}{\theta_g^{*2}} \end{aligned} \tag{1.3.3}$$

Now, writing $r_{gi} = \dfrac{\theta_g^* z_{gi} - 1}{\theta_g^* - \theta_g}\theta_g$ and assuming $\theta_g \neq \theta_g^*$ one obtains

$$E(r_i) = y_i,\ V_R(r_{gi}) = \frac{\theta_g^2\theta_g^{*2}\left\{y_i\dfrac{1-\theta_g}{\theta_g^2} + (1-y_i)\dfrac{1-\theta_g^*}{\theta_g^{*2}}\right\}}{(\theta_g^*-\theta_g)^2} = \phi_{gi}$$

and $C_R(r_{gi}, r_{gj}) = 0$ for $i \neq j$ (1.3.4)

Finally, following Theorem 1.1.1, one can derive the following results (Theorem 1.3.1) obtained by Singh and Grewal (2013).

Theorem 1.3.1.

(i) $\bar{r}_{sg} = \frac{1}{n}\sum_{i\in s} r_{gi} = \frac{\theta_g^* \bar{z}_g - 1}{\theta_g^* - \theta_g}\theta_g$ is an unbiased estimator of π

where $\bar{z}_g = \frac{1}{n}\sum_{i\in s} \bar{z}_{gi}$

(ii) $\text{Var}(\bar{r}_{sg}) = \frac{\pi(1-\pi)}{n} + \frac{1}{nN}\sum_{i\in U}\phi_{gi}$

$$= \frac{\pi(1-\pi)}{n} + \frac{\phi_g}{n} = V_{sg} \quad (1.3.5)$$

where $\phi_g = \frac{\pi(1-\theta_g)\theta_g^{*2} + (1-\pi)(1-\theta_g^*)\theta_g^2}{(\theta_g^* - \theta_g)^2}$

(iii) An unbiased estimator of $\text{Var}(\bar{r}_{sg})$ is

$$\hat{\text{Var}}(\bar{r}_{sg}) = \frac{\sum_{i\in s}(r_{gi} - \bar{r}_g)^2}{n(n-1)}$$

2. EXISTING COMPARISON AMONG WARNER'S (1965), KUK'S (1990), AND SINGH AND GREWAL'S (2013) RR MODELS

2.1 Warner's (1965) vs. Kuk's (1990) model

The variance of the Kuk's (1990) estimator V_k=Var$(\bar{r}_k)$ decreases with the increase of the number of trials c and the absolute difference $|\theta_1 - \theta_2|$. However, to preserve the respondent's confidence, the difference $|\theta_1 - \theta_2|$ should be under a certain level. The percentage relative efficiency (RE) of Kuk's (1990) RR model with respect to Warner's (1965) model is given by

$$E_k = \frac{V_w}{V_k}\times 100 \quad (2.1.1)$$

2.2 Kuk's (1990) model vs. Singh and Grewal's (2013) model

The percentage relative efficiency of Singh and Grewal's (2013) model with respect to Kuk's (1990) model with $c = 1$ was proposed by Singh and Grewal (2013) is

$$E_{sg} = \frac{V_k}{V_{sg}}\times 100 \quad (2.1.2)$$

From Theorem 1.2.1 and Theorem 1.3.1, we note that for $c = 1$,

$V_{sg} \le V_k$

if

$$\frac{\pi(1-\theta_g)\theta_g^{*2} + (1-\pi)(1-\theta_g^*)\theta_g^2}{(\theta_g^* - \theta_g)^2} \le \frac{\pi\theta_1(1-\theta_1) + (1-\pi)\theta_2(1-\theta_2)}{(\theta_1 - \theta_2)^2} \quad (2.1.3)$$

Clearly, the expression (2.1.3) does not provide any simple comparison. Singh and Grewal (2013) computed relative efficiency E_{sg} numerically by keeping θ_1, θ_2 and π fixed to certain values then found the values of θ_g and θ_g^* for which relative exceed 100. For example, for θ_1 = 0.7, $\theta_2 = 0.2$ and $\pi = 0.1$, E_{sg} exceeds 100 for θ_g = 0.1 and θ_g^* = 0.3, 0.4,0.5,0.6, 0.7,0.8 and 0.9. However, the comparison is not fair because total number of RR trials required to draw a card bearing the respondent status for Singh and Grewal's (2013) model is a random variable which is at least one. It may be two, three and so on. Hence, the appropriate comparison should involve expected number of RR trials required to produce a card bearing respondent status the first time.

3. A NEW METHOD OF COMPARISON OF EFFICIENCIES

The total number of cards drawn by the i^{th} respondent in Singh and Grewal (2013) model is a random variable and it is given in Eq (1.3.1) as

$$z_{gi} = y_i X_i + (1 - y_i)Q_i$$

and the expected value of z_{gi} is given by

$$\gamma = E(z_{gi}) = E_p[y_i E_R(X_i) + (1-y_i)E_R(Q_i)]$$

$$= \pi\frac{1}{\theta_g} + (1-\pi)\frac{1}{\theta_g^*}$$

The values of γ for different values of π, θ_g and θ_g^* are given in Table 3.1. From Table 3.1 one finds that the expected number of trials γ for Singh and Grewal model ranges from 1.2 to 8.4.

3.1 Warner's model vs. Kuk's model

For the proposed new method of comparison of the relative efficiency of Kuk's model with $c(=\gamma)$ number of trials with respect to Warner's model, we assume each of the respondents selected in the sample for the Warner's model performs γ RR trials independently. Let $z_i(j)$ be the RR response obtained by the $i(\in s)$th respondent in the $j(=1,..,\gamma)$th trial. Then, an unbiased estimator π for the Warner's model becomes

Table 3.1. Expected number of trial under Singh and Grewal (2013) model

π	θ	θ*	γ	π	θ	θ*	γ	π	θ	θ*	γ	π	θ	θ*	γ
0.1	0.1	0.5	2.8	0.3	0.1	0.5	4.4	0.5	0.1	0.5	6.000	0.7	0.1	0.5	7.600
		0.6	2.5			0.6	4.167	0.5		0.6	5.833			0.6	7.500
		0.7	2.286			0.7	4	0.5		0.7	5.714			0.7	7.429
		0.8	2.125			0.8	3.875	0.5		0.8	5.625			0.8	7.375
		0.9	2			0.9	3.778	0.5		0.9	5.556			0.9	7.333
0.1	0.2	0.6	2	0.3	0.2	0.6	2.667	0.5	0.2	0.6	3.333	0.7	0.2	0.6	4.000
		0.7	1.786			0.7	2.5	0.5		0.7	3.214			0.7	3.929
		0.8	1.625			0.8	2.375	0.5		0.8	3.125			0.8	3.875
		0.9	1.500			0.9	2.278	0.5		0.9	3.056			0.9	3.833
0.1	0.3	0.7	1.619	0.3	0.3	0.7	2.000	0.5	0.3	0.7	2.381	0.7	0.3	0.7	2.762
		0.8	1.458			0.8	1.875	0.5		0.8	2.292			0.8	2.708
		0.9	1.333			0.9	1.778	0.5		0.9	2.222			0.9	2.667
0.1	0.4	0.8	1.375	0.3	0.4	0.8	1.625	0.5	0.4	0.8	1.875	0.7	0.4	0.8	2.125
		0.9	1.250			0.9	1.528	0.5		0.9	1.806			0.9	2.083
0.1	0.5	0.9	1.200	0.3	0.5	0.9	1.378	0.5	0.5	0.9	1.556	0.7	0.5	0.9	1.733
0.2	0.1	0.5	3.600	0.4	0.1	0.5	5.200	0.6	0.1	0.5	6.800	0.8	0.1	0.5	8.400
		0.6	3.333		0.1	0.6	5.000	0.6		0.6	6.667			0.6	8.333
		0.7	3.143		0.1	0.7	4.857	0.6		0.7	6.571			0.7	8.286
		0.8	3.000		0.1	0.8	4.750	0.6		0.8	6.500			0.8	8.250
		0.9	2.889		0.1	0.9	4.667	0.6		0.9	6.444		0.2	0.9	8.222
0.2	0.2	0.6	2.333	0.4	0.2	0.6	3.000	0.6	0.2	0.6	3.667			0.6	4.333
		0.7	2.143		0.2	0.7	2.857	0.6		0.7	3.571			0.7	4.286
		0.8	2.000		0.2	0.8	2.750	0.6		0.8	3.500			0.8	4.25
		0.9	1.889		0.2	0.9	2.667	0.6		0.9	3.444			0.9	4.222
0.2	0.3	0.7	1.810	0.4	0.3	0.7	2.190	0.6	0.3	0.7	2.571	0.8	0.3	0.7	2.952
		0.8	1.667		0.3	0.8	2.083	0.6		0.8	2.500			0.8	2.917
		0.9	1.556		0.3	0.9	2.000	0.6		0.9	2.444		0.4	0.9	2.889
0.2	0.4	0.8	1.500	0.4	0.4	0.8	1.750	0.6	0.4	0.8	2.000			0.8	2.250
		0.9	1.389		0.4	0.9	1.667	0.6		0.9	1.944		0.5	0.9	2.222
0.2	0.5	0.9	1.289	0.4	0.5	0.9	1.467	0.6	0.5	0.9	1.644			0.9	1.822

$$\bar{r}_w^* = \frac{1}{n}\sum_{i\in s}\bar{r}_i \qquad (3.1.1)$$

where $\bar{r}_i = \frac{1}{\gamma}\sum_{j=1}^{\gamma}\frac{z_i(j)-(1-P_w)}{2P_w-1}$.

and the variance of $\bar{r}_w^*$ is given by

$$V_w^* = \text{Var}(\bar{r}_w^*)$$
$$= V_p\left[E_R(\bar{r}_w^*)\right] + E_p\left[V_R(\bar{r}_w^*)\right]$$
$$= V_p\left[\frac{1}{n}\sum_{i\in s} y_i\right] + E_p\left[\frac{1}{n^2}\sum_{i\in s}\frac{\phi_w}{\gamma}\right]$$
$$= \frac{1}{n}\left[\pi(1-\pi)+\phi_w/\gamma\right] \qquad (3.1.2)$$

From Equations (1.2.5) and (3.1.2) one finds that the Kuk's model becomes more, equal and less efficient than the Warner model under the new method of comparison where each respondent performs γ trials independently is

$V_w^* \;>, = \text{or} \;< V_k^*$

(where V_k^* =Var($\overline{r}_k$) with $c = \gamma$)

i.e. if

$$\frac{P_w(1-P_w)}{(2P_w-1)^2} >, = < \frac{\pi\theta_1(1-\theta_1)+(1-\pi)\theta_2(1-\theta_2)}{(\theta_1-\theta_2)^2} \quad (3.1.3)$$

3.2 Warner's vs. Singh and Grewal model

From the Equations (3.1.2) and (1.3.5) we note that the Singh and Grewal model becomes more, equal and less efficient than the Warner's model under the new method of comparison if

$V_w^* \;>, = \text{or} \;< V_{sg}$

i.e.

$$\frac{P_w(1-P_w)}{\gamma(2P_w-1)^2} >, = \text{or} < \frac{\pi(1-\theta_g)\theta_g^{*2}+(1-\pi)(1-\theta_g^*)\theta_g^2}{(\theta_g^*-\theta_g)^2} \quad (3.2.1)$$

3.3 Kuk vs. Singh and Grewal model

For the new method of comparison Singh and Grewal model becomes more, equal and less efficient than the Kuk's model if

Table 3.2. Relative efficiencies of $\overline{r}_w^*$, $\overline{r}_k$, $\overline{r}_k^*$ and $\overline{r}_{sg}$

		π											
		0.1				**0.2**				**0.3**			
α	**β**	E_w^*	E_k	E_k^*	E_{sg}	E_w^*	E_k	E_k^*	E_{sg}	E_w^*	E_k	E_k^*	E_{sg}
0.1	0.4	325.00	100.99	328.20	909.80	400.00	106.50	426.16	593.96	475.00	111.40	529.21	478.54
0.1	0.5	280.00	91.25	255.50	604.60	360.00	97.42	350.71	405.04	440.00	102.30	450.06	342.10
0.1	0.6	250.00	86.67	216.70	416.70	333.30	93.45	311.49	297.07	416.70	98.19	409.11	266.39
0.1	0.7	228.60	86.03	196.70	295.50	314.30	92.92	292.02	230.11	400.00	97.06	388.25	220.19
0.1	0.8	212.50	89.66	190.50	213.70	300.00	95.15	285.44	185.81	387.50	97.86	379.21	189.91
0.1	0.9	200.00	100.00	200.00	156.20	288.90	100.00	288.89	154.95	377.80	100.00	377.78	168.95
0.2	0.5	230.00	94.80	218.00	533.90	260.00	98.33	255.66	376.34	290.00	101.60	294.70	303.89
0.2	0.6	200.00	91.56	183.10	414.60	233.30	95.07	221.82	280.97	266.70	98.04	261.43	228.35
0.2	0.7	178.60	92.86	165.80	314.00	214.30	95.63	204.91	213.45	250.00	97.70	244.25	178.57
0.2	0.8	162.50	100.00	162.50	234.20	200.00	100.00	200.00	165.38	237.50	100.00	237.5	144.61
0.2	0.9	150.00	118.69	178.00	172.10	188.90	109.30	206.50	130.49	227.80	104.80	238.66	120.59
0.3	0.6	183.30	96.23	176.40	368.90	200.00	97.53	195.05	273.82	216.70	98.71	213.88	224.27
0.3	0.7	161.90	100.00	161.90	319.60	181.00	100.00	180.95	221.35	200.00	100.00	200.00	178.38
0.3	0.8	145.80	111.37	162.40	263.90	166.70	107.40	178.99	176.5	187.50	104.60	196.07	143.33
0.3	0.9	133.30	139.94	186.00	209.10	155.60	122.60	190.70	140.07	177.80	113.10	200.99	116.83
0.4	0.7	153.60	106.43	163.50	297.30	164.30	104.80	172.17	225.05	175.0	103.40	180.90	185.20
0.4	0.8	137.50	122.36	168.30	281.60	150.00	115.90	173.79	194.06	162.50	110.90	180.21	154.56
0.4	0.9	125.00	161.97	202.50	252.20	138.90	138.50	192.29	162.04	152.80	124.40	190.05	127.78
0.4	0.1	925.00	168.08	155.50	218.70	850.00	144.40	1227.00	267.51	775.00	132.90	1029.70	307.80
0.5	0.8	132.50	132.55	175.60	279.90	140.00	125.00	174.97	209.17	147.50	118.90	175.30	170.37
0.5	0.9	120.00	184.36	221.20	293.60	128.90	156.80	202.07	191.94	137.80	139.10	191.61	148.49
0.5	0.1	920.00	143.61	1321.0	157.50	840.00	126.10	1059.30	205.73	760.00	118.30	899.11	241.62
0.5	0.2	470.00	128.25	602.80	139.70	440.00	120.50	530.24	163.04	410.00	115.30	472.78	184.95

0.6	0.9	116.70	207.03	241.50	325.00	122.20	178.10	217.71	226.94	127.80	158.10	202.04	178.10
0.6	0.2	466.70	117.86	550.00	100.70	433.30	111.90	484.68	123.82	400.00	108.20	432.66	143.50
0.6	0.1	916.70	126.51	1160.00	125.20	833.30	114.90	957.41	173.61	750.00	109.90	824.47	207.22
0.7	0.1	914.30	114.56	1047.00	106.20	828.60	107.80	893.03	155.05	742.90	104.90	779.59	187.33
0.7	0.2	464.30	108.35	503.10	77.55	428.60	105.1	450.54	100.82	392.9	103.3	405.74	119.14
0.7	0.3	314.30	100	314.30	78.35	295.20	100	295.24	93.51	276.2	100	276.19	107.45
0.7	0.4	239.3	96.29	230.40	88.97	228.60	97.56	223.01	100.47	217.9	98.71	215.06	112.37

		π											
		0.4				**0.5**				**0.6**			
α	**β**	E_w^*	E_k	E_k^*	E_{sg}	E_w^*	E_k	E_k^*	E_{sg}	E_w^*	E_k	E_k^*	E_{sg}
0.1	0.4	550.00	115.90	637.40	418.20	625.00	120.50	752.81	377.56	700.00	125.70	880.11	343.03
0.1	0.5	520.00	106.22	552.40	311.80	600.00	109.80	658.54	290.32	680.00	113.50	771.73	268.60
0.1	0.6	500.00	101.56	507.80	253.90	583.30	104.30	608.12	243.75	666.70	106.80	712.12	229.49
0.1	0.7	485.70	99.66	484.00	219.10	571.40	101.50	580.03	216.12	657.10	103.10	677.63	206.64
0.1	0.8	475.00	99.35	471.90	196.50	562.50	100.30	564.23	198.44	650.00	101.10	657.00	192.22
0.1	0.9	466.70	100.00	466.70	181.00	555.60	100.00	555.56	186.46	644.40	100.00	644.44	182.59
0.2	0.5	320.00	104.77	335.30	261.10	350.00	107.90	377.75	231.13	380.00	111.30	423.05	206.94
0.2	0.6	300.00	100.62	301.90	199.30	333.30	103.00	343.33	178.82	366.70	105.40	386.50	161.21
0.2	0.7	285.70	99.31	283.70	160.20	321.40	100.70	323.51	146.84	357.10	101.90	363.94	133.93
0.2	0.8	275.00	100.00	275.00	134.30	312.50	100.00	312.5	126.07	350.00	100.00	350.00	116.53
0.2	0.9	266.70	102.20	272.50	116.30	305.60	100.50	307.09	111.89	344.40	99.16	341.56	104.84
0.3	0.6	233.30	99.82	232.90	193.00	250.00	100.90	252.22	170.42	266.70	102.00	271.91	152.30
0.3	0.7	219.10	100.00	219.10	153.300	238.10	100.00	238.1	135.66	257.10	100.00	257.14	121.08
0.3	0.8	208.30	102.46	213.50	125.00	229.20	100.80	230.91	111.92	250.00	99.26	248.16	100.45
0.3	0.9	200.00	107.08	214.20	104.50	222.20	102.90	228.57	95.24	244.40	99.46	243.13	86.29
0.4	0.7	185.70	102.10	189.60	159.30	196.40	100.90	198.25	140.41	207.10	99.82	206.76	125.34
0.4	0.8	175.00	106.94	187.20	131.30	187.50	103.60	194.32	114.92	200.00	100.70	201.40	101.85
0.4	0.9	166.70	115.00	191.70	108.90	180.60	108.10	195.18	95.80	194.40	102.50	199.36	84.96
0.4	0.1	700.00	125.73	880.10	343.00	625.00	120.50	752.81	377.56	550.00	115.90	637.43	418.24
0.5	0.8	155.00	113.73	176.30	145.30	162.50	109.30	177.64	127.26	170.00	105.40	179.15	113.11
0.5	0.9	146.70	126.63	185.70	123.60	155.60	117.20	182.35	106.75	164.40	109.60	180.29	93.74
0.5	0.1	680.00	113.49	771.70	268.60	600.00	109.80	658.54	290.32	520.00	106.20	552.36	311.78
0.5	0.2	380.00	111.33	423.10	206.90	350.00	107.90	377.75	231.13	320.00	104.80	335.26	261.11
0.6	0.9	133.30	143.34	191.10	148.30	138.90	131.90	183.12	127.78	144.40	122.50	176.93	112.28
0.6	0.2	366.70	105.41	386.50	161.20	333.30	103.00	343.33	178.82	300.00	100.60	301.86	199.26
0.6	0.1	666.70	106.82	712.10	229.50	583.30	104.30	608.12	243.75	500.00	101.60	507.81	253.91

0.7	0.1	657.10	103.12	677.60	206.60	571.40	101.50	580.03	216.12	485.70	99.66	484.04	219.07
0.7	0.2	357.10	101.90	363.90	133.90	321.40	100.70	323.51	146.84	285.70	99.31	283.74	160.24
0.7	0.3	257.10	100.00	257.10	121.10	238.10	100.00	238.10	135.66	219.10	100.00	219.05	153.34
0.7	0.4	207.10	99.82	206.80	125.30	196.40	100.90	198.25	140.41	185.70	102.10	189.61	159.29

		π											
		0.7				0.8				0.9			
α	β	E_w^*	E_k	E_k^*	E_{sg}	E_w^*	E_k	E_k^*	E_{sg}	E_w^*	E_k	E_k^*	E_{sg}
0.1	0.4	775.00	132.86	1030.00	307.80	850.00	144.40	1227.00	267.51	925.00	168.10	1554.80	218.67
0.1	0.5	760.00	118.30	899.10	241.60	840.00	126.10	1059.30	205.73	920.00	143.60	1321.20	157.53
0.1	0.6	750.00	109.93	824.50	207.20	833.30	114.90	957.41	173.61	916.70	126.50	1159.60	125.2
0.1	0.7	742.90	104.94	779.60	187.30	828.60	107.80	893.03	155.05	914.30	114.60	1047.50	106.2
0.1	0.8	737.50	101.90	751.50	174.90	825.00	103.10	850.82	143.50	912.50	106.10	968.23	94.17
0.1	0.9	733.30	100.00	733.30	166.80	822.20	100.00	822.22	135.89	911.10	100.00	911.11	86.13
0.2	0.5	410.00	115.31	472.80	185.00	440.00	120.50	530.24	163.04	470.00	128.30	602.78	139.7
0.2	0.6	400.00	108.17	432.70	143.50	433.30	111.90	484.68	123.82	466.70	117.90	550.00	100.73
0.2	0.7	392.90	103.28	405.70	119.10	428.60	105.10	450.54	100.82	464.30	108.40	503.05	77.55
0.2	0.8	387.50	100.00	387.50	103.80	425.00	100.00	425.00	86.35	462.50	100.00	462.50	62.72
0.2	0.9	383.30	97.84	375.10	93.59	422.20	96.12	405.85	76.73	461.10	92.84	428.11	52.71
0.3	0.6	283.30	103.12	292.20	136.40	300.00	104.40	313.31	121.28	316.70	106.10	335.91	106.14
0.3	0.7	276.20	100.00	276.20	107.50	295.20	100.00	295.24	93.51	314.30	100.00	314.29	78.35
0.3	0.8	270.80	97.78	264.80	88.74	291.70	96.05	280.15	75.66	312.50	93.64	292.64	60.31
0.3	0.9	266.70	96.27	256.70	76.09	288.90	92.62	267.57	63.63	311.10	87.28	271.55	47.98
0.4	0.7	217.90	98.71	215.10	112.40	228.60	97.56	223.01	100.47	239.30	96.29	230.42	88.97
0.4	0.8	212.50	97.91	208.10	90.19	225.00	95.03	213.81	78.89	237.50	91.74	217.87	67.24
0.4	0.9	208.30	97.54	203.20	74.69	222.20	92.47	205.49	63.98	236.10	86.52	204.28	52.14
0.4	0.1	475.00	111.41	529.20	478.50	400.00	106.50	426.16	593.96	325.00	101.00	328.23	909.77
0.5	0.8	177.50	101.75	180.60	101.20	185.00	98.26	181.78	90.63	192.50	94.74	182.37	80.73
0.5	0.9	173.30	103.07	178.70	82.64	182.20	96.98	176.71	72.38	191.10	90.82	173.57	62.28
0.5	0.1	440.00	102.29	450.10	342.10	360.00	97.42	350.71	405.04	280.00	91.25	255.50	604.64
0.5	0.2	290.00	101.62	294.70	303.90	260.00	98.33	255.66	376.34	230.00	94.80	218.04	533.90
0.6	0.9	150.00	114.55	171.80	99.75	155.60	107.50	167.27	89.00	161.10	101.10	162.80	79.31
0.6	0.2	266.70	98.04	261.40	228.40	233.30	95.07	221.82	280.97	200.00	91.56	183.13	414.58
0.6	0.1	416.70	98.19	409.10	266.40	333.30	93.45	311.49	297.07	250.00	86.67	216.67	416.67
0.7	0.1	400.00	97.06	388.30	220.20	314.30	92.92	292.02	230.11	228.60	86.03	196.65	295.53
0.7	0.2	250.00	97.70	244.30	178.60	214.30	95.63	204.91	213.45	178.60	92.86	165.83	314.04
0.7	0.3	200.00	100.00	200.00	178.40	1810	100.00	180.95	221.35	161.90	100.00	161.90	319.61
0.7	0.4	175.00	103.37	180.90	185.20	164.30	104.80	172.17	225.05	153.60	106.40	163.45	297.32

$V_k^* \; >, = \text{or} < V_{sg}$

i.e. if

$$\frac{\pi\theta_1(1-\theta_1)+(1-\pi)\theta_2(1-\theta_2)}{\gamma(\theta_1-\theta_2)^2} <, = \text{or} > \frac{\pi(1-\theta_g)\theta_g^{*2}+(1-\pi)(1-\theta_g^*)\theta_g^2}{(\theta_g^*-\theta_g)^2} \quad (3.3.1)$$

3.4 Relative efficiencies

The percentage relative efficiencies of the Warner's with γ number of trials, Kuk with $c=1$, Kuk with $c=\gamma$ and Singh and Grewal's model with respect to the Warner's model with a single trial is given by

$$E_w^* = \frac{V_w}{V_w^*}\times 100 , \; E_k = \frac{V_w}{V_k}\times 100 , \; E_k^* = \frac{V_w}{V_k^*}\times 100 \text{ and}$$

$$E_{sg} = \frac{V_w}{V_{sg}}\times 100 \quad (3.4.1)$$

It is obvious to note that the Warner's model and the modified Kuk's model with γ number of trials are more efficient than the corresponding Warner's model with a single trial and Kuk's (1990) model with $c=1$ as $\gamma>1$.

The validity of the conditions (3.1.3), (3.2.1) and (3.3.1) stated above are very difficult to check as they depend on the values of the parameters π, P_w, θ_1, etc. Hence, we compare performances of the models numerically. Table 3.2 shows relative efficiencies E_w^*, E_k, E_k^* and E_{sg} for diffident values of π, α and β where $\theta_1=\theta_g=\alpha$, $\theta_2=\theta_g^*=\beta$ and $\theta_1-\theta_2=2P_w-1$. The estimator $\overline{r}_k$ is found to be least efficient in almost all situations. Among the estimators $\overline{r}_w^*$, $\overline{r}_k^*$, and $\overline{r}_{sg}$, none of them was found best in all situations. Efficiencies E_w^*, E_k^* and E_{sg} varies from 125 to 925, 161.9 to 1554.8 and 47.98 to 909.77 respectively. On the other hand, $\overline{r}_{sg}$ was found to be least efficient for some of the occasions. In most situations the modified Kuk's estimator was found to be most efficient.

4. DISCUSSIONS

The RR techniques are used to collect information on sensitive items. Warner (1965), Kuk (1990) and Singh and Grewal (2013) proposed the RR techniques for estimating π, the proportion of individuals belong to a certain sensitive group A.

In Warner's RR technique, respondents are asked to answer a sensitive question such as whether or not they belong to a certain sensitive group A. However, to answer such questions, the respondent may feel embarrassed or uncomfortable.

In Kuk's (1990) RR technique each respondent has to draw c cards at random with replacement from a pack containing certain number of black and white cards. The respondent is asked to report how many red cards he or she has drawn. So, in this method the respondent need not answer any sensitive or embarrassing question.

The Singh and Grewal (2013) technique is similar to Kuk (1990) technique. Here, each respondent is asked to draw cards from a pack of cards containing two types cards bearing statement "I belong to the group A" and "I belong to the group $\overline{A}$". Respondent draws cards at random one by one with replacement until the statement of drawn card matches his or her status the first time (for details see Section 1.3). The respondent is asked to report how many cards he/she has drawn. In this method, the number of cards drawn by a respondent is not fixed. It is a random variable and, hence, respondents may very often commit mistakes in drawing cards.

Singh and Grewal (2013) compared performance of their proposed estimator with the estimators proposed by Kuk (1990), with $c=1$, and Warner (1965) estimator where a respondent draws only one card. The comparison is not fair since the number of cards drawn by Singh and Grewal model is always greater than one. In fact, the expected number of cards drawn by a respondent in Singh and Grewal model γ ranges from 1.2 to 8.4. (see Table 3.1).

In the present paper, a more realistic comparison has been proposed where each of the respondents in the Warner (1965) and Kuk (1990) methods draws γ number of cards. The efficiencies of Warner, Modified Warner, Kuk, Modified Kuk and Singh and Grewal models are compared theoretically but no meaningful conclusion was reached because the comparisons involved many unknown parameters. The numerical comparisons indicate that none of the proposed model is best in all the situations. However, the modified Kuk's model performs fairly well in most situations and it may be preferred for its easy execution and that the fact that the respondent need not answer embarrassing questions.

ACKNOWLEDGEMENTS

The author wishes to thank the referee and Dr. Wiloson Moseki Thupeng for editing the manuscript.

REFERENCES

Arcos, A., Rueda, M., and Singh, S.A. (2015). Generalized approach to randomized response for quantitative variables, *Quality and Quantity*, **49**, 1239-1256.

Arnab, R. (1990). On commutativity of design and model expectations in randomized response surveys. *Communications in Statistics: Theory and Methods*, **19**, 3751-3757.

Arnab, R. (1996). Randomized response trials: a unified approach for qualitative data. *Communications in Statistic: Theory and Methods,* **25(6)**, 1173-1183.

Arnab, R. (2017). *Survey Sampling Theory and Applications.* Academic Press, U.K.

Arnab, R., and Singh, S. (2014). An alternative randomized response model using two deck of cards: A rejoinder, 2013, *Statistica*, anno LXXIII, 493-503.

Arnab, R., and Singh. S (2007): A note on unrelated question randomized response model, *Journal of the Indian Society of Agricultural Statistics,* **61(1)**, 77-79.

Arnab, R., and Mothupi, T. (2015). Randomized response techniques: A case study of the risky behaviors' of students of a certain University, *Model Assisted Statistics and Applications,* **10**, 421-430.

Chaudhuri, A. and Dihidar, K. (2014). Generating randomized response by inverse mechanism. *Model Assisted Statistics and Applications*, 9, 343-351.

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

Christofides, T.C. (2003). A generalized randomized response technique, *Metrika,* **57**, 195-200.

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

Horvitz, D.G., Shah, B.V., and Simmons, W.R., (1967). The unrelated question randomized response model. *Proceedings of Social Statistics Section*, *American Statistical Association*, 65-72.

Kim, J.I. (1978). *Randomized response technique for surveying human populations*. Ph.D. Dissertation, Temple University, Philadelphia, USA.

Kim, J.M. and Warde, W. D. (2004). Stratified Warner's randomized response model. *Journal of Statistical Planning and Inference*, **120**, 155-65.

Kuk, A.Y.C. (1990). Asking sensitive questions indirectly. *Biometrika,* **77**, 436-438.

Rueda, M., Cobo, B., Arcos, A., (2015). Package 'RRTCS': Randomized Response Techniques for Complex Surveys. *http://cran.r-project.org/web/packages/RRTCS/*.

Singh, S., and Grewal, I. (2013). Geometric distribution as randomization device: Implemented to the kuk's model. *International journal of Contemporary Mathematical Sciences*, **8**, 243-248.

Van der Heijden, P. G. M., Van Gils, G., Bouts, J., and Hox, J. J. (1998). A comparison of randomized response. CASAQ, and direct questioning; eliciting sensitive information in the context of social security fraud. *Kwantitatieve Methoden,* **59***,* 15-34.

Warner, S. L., (1965). Randomize response: a survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association,* **60**, 63-69.