Jour. Ind. Soc. Ag. Statistics
55(2), 2002 : 209-219

A Modified Quantitative Randomized Response Model

N.S. Mangat and Sarjinder Singh[1]
Morrison Scientific Inc., 815-706, 7th Avenue SW, Calgary, Alberta, Canada
(Received : August, 2001)

SUMMARY

Greenberg *et al.* [2] proposed a randomized response (RR) model to collect numerical data on sensitive characters. Their model requires two randomized devices R_i to be used for each respondent in sample i, i = 1, 2. The variance of their estimator attains a minimum if probability p_2, representing the sensitive character in RR device R_2, is taken as zero. Although this choice provides an economical strategy, the response of the respondents selected in both the samples (with replacement) can be linked to their actual status. Thus the choice $p_2 = 0$ renders the RR model non-functional. In such a situation the obvious choice according to Greenberg *et al.* [2] is left to the usual strategy by choosing $p_1 + p_2 = 1$. Here we come up with an alternative strategy. This proposed strategy being more efficient can be used in almost all the practical situations instead of using $p_1 + p_2 = 1$ model for the situation when $p_2 = 0$ choice becomes non-functional.

Key words: Sensitive character, Quantitative model, Cost aspect, Randomized response, Efficiency.

1. Introduction

In sample surveys, it is difficult to obtain reliable information on stigmatized, criminal or socially unacceptable matters using the conventional direct question procedure. To overcome this problem, Warner [16] developed an interviewing procedure, popularly known as randomized response (RR) technique, designed to eliminate evasive answer bias. Among others, some modifications related to this model are due to Raghavarao [12], Mangat and Singh [9], Lakshmi and Raghavarao [6], Krishnamoorthy and Raghavarao [5] and Mangat [14].

In Warner's [16] model, both the questions were related to the sensitive issues. Horvitz *et al.* [3] felt that the confidence of the respondents in the anonymity provided by the RR technique might be further enhanced if one of

[1] Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

the two questions referred to a nonsensitive characteristic unrelated to the sensitive characteristic, and called it the unrelated question RR model. The theoretical framework for this model was developed by Greenberg *et al.* [1]. They considered both the situations: when π_y, the proportion of nonsensitive attribute in the population, is known and when it is not known. For the situation when π_y is known, some of the recent developments are due to Mangat *et al.* [10] and Singh *et al.* [9].

Greenberg *et al.* [2] observed that the RR procedure need not be restricted to qualitative traits only and may be applied to the case of quantitative characters as well. So, they developed an RR model to obtain quantitative data on sensitive issues on the lines of Greenberg *et al.* [1].

This model is reviewed briefly in the following section, and the problem arising on account of the optimal choice of probabilities representing the statements in RR device is discussed.

2. *Quantitative Randomized Response Model*

Analogous to Greenberg *et al.'s* [1] model, the RR device used for the quantitative model also comprised two statements: one about sensitive character A and the other for nonsensitive character Y, but requiring quantitative response. For example, according to Greenberg *et al.* [2], the question Q_A related to sensitive character and Q_Y regarding nonsensitive character in RR set R_i to be used for the respondents in sample i, i = 1, 2, could be

Q_A: About how many dollars did the head of this household earn last year?

Q_Y: About how many dollars do you think, on the average, the average head of a household of your size earns in a year?

represented with probabilities p_i and $(1 - p_i)$ respectively. Each respondent in the with replacement simple random sample of size $n_i, i = 1, 2,$ is asked to draw a statement from the RR device R_i at random and answer it without revealing the question selected. If μ_A, μ_Y and σ_A^2, σ_Y^2 be the means and variances of these characters in the population, the mean response for the respondents in sample i is

$$\mu_{z_i} = p_i\mu_A + (1 - p_i)\mu_Y, i = 1, 2$$

If $\bar{z}_1$ and $\bar{z}_2$ are the observed mean responses for the two samples, their estimator

$$\hat{\mu}_G = \frac{(1 - p_2)\bar{z}_1 - (1 - p_1)\bar{z}_2}{p_1 - p_2} \tag{2.1}$$

is unbiased for the population mean μ_A. The variance of $\bar{z}_i$ is

$$V(\bar{z}_i) = \frac{1}{n_i}\left[\sigma_Y^2 + p_i\left(\sigma_A^2 - \sigma_Y^2\right) + p_i\left(1 - p_i\right)\left(\mu_A - \mu_Y\right)^2\right], i = 1, 2$$

This yields

$$V\left(\hat{\mu}_G\right) = \frac{1}{\left(p_1 - p_2\right)^2}\left[\frac{\left(1 - p_2\right)^2 V\left(z_1\right)}{n_1} + \frac{\left(1 - p_1\right)^2 V\left(z_2\right)}{n_2}\right] \tag{2.2}$$

Under optimal choice of $\mu_A = \mu_Y$ and $\sigma_A^2 = \sigma_Y^2$, the variance expression of $\hat{\mu}_G$ is

$$V\left(\hat{\mu}_G\right) = \frac{\sigma_Y^2}{\left(p_1 - p_2\right)^2}\left[\frac{\left(1 - p_2\right)^2}{n_1} + \frac{\left(1 - p_1\right)^2}{n_2}\right] \tag{2.3}$$

Greenberg *et al.* [2] reported that $V\left(\hat{\mu}_G\right)$ will be minimum at $p_2 = 0$. According to them, "If such a procedure does not produce incompatible results in estimating μ_A it is obviously the method of choice. Unless there is evidence that the second sample can be used with $p_2 = 0$, the authors believe a conservative approach by having $p_1 + p_2 = 1$ is still worthwhile." Horvitz *et al.* [3] also reported, "As in the qualitative response case, the optimum procedure is to choose $p_2 = 0$, so that second sample is used to estimate μ_Y only. However what happens to the privacy of the respondents by following the recommendation of $p_2 = 0$ needs to be discussed".

According to this line, the investigation from the respondents included in one of the samples (say first sample) will be collected on both the characters by using RR device while from the respondents in other (second) sample, the information will be elicited only on the non-sensitive character Y by using direct question procedure. In order to examine this recommendation, we consider the following example.

Suppose the objective is to estimate the mean income of cab drivers in a certain city. As both the samples are drawn using replacement method, there is a possibility that some of the respondents may occur in both the samples. The two questions represented with probabilities p_i and $\left(1 - p_i\right)$ respectively are

Q_A: About how many dollars did you earn last year?

Q_Y: About how many dollars do you think a cab driver in this city can earn in a year if he works for the average amount of time?

Let the reply of the repeated respondent, while in the second sample, regarding question Q_Y be $40,000. If he selects the sensitive question Q_A while in the first sample, his response could be different from $40,000. So the later response, if different, could be linked to the sensitive character. The privacy of such a respondent is thus disclosed and Greenberg *et al.*'s [1] model with choice $p_2 = 0$ becomes non-functional. Obviously, in these situations, either we are to follow the choice $p_1 + p_2 = 1$ as pointed out by the second line of the quoted statement or we are to look for more efficient alternative strategy which makes the $p_2 = 0$ choice functional. In what follows, an attempt is made to make $p_2 = 0$ strategy functional and it is shown that this proposed strategy is better than the Greenberg *et al.'s* $p_1 + p_2 = 1$ choice in the situation discussed above. It is, therefore, necessary to look for an appropriate modification to make this optimal choice functional. Similar discussion on the choice of $p_2 = 0$ can be had from Singh [13] and Mahmood *et al.* [7]. Some researchers believe that the probability of repetition of a respondent is very small or negligible in case of WR sampling. It is interesting to note that Singh *et al.* [15] have shown that same problem occurs while using SRS without replacement sampling, which violates the assumption of low probability of repetition of a respondent in two samples.

In the present investigation, an attempt has been made with an interesting and meaningful model, which is more practical and depends upon the cost of survey, as discussed in the section below.

3. Proposed Strategy

In this procedure the selection criterion for the respondents remains the same as for the model discussed in Section 2. The difference is that each respondent in one (say first) of the two samples is asked to select randomly one of the two questions Q_A or Q_Y, represented with probabilities p_1 and $(1 - p_1)$ respectively, in the RR device. If Q_A is selected, the respondent is required to add a random variable S to the actual response and report the scrambled response (following the additive model of Pollock and Bek [12]). The random variable S has known mean μ_s and variance σ_s^2. The respondents in the second sample are used to collect information on the non sensitive character only. Let g(t) and h(t) be the probability (or probability density) functions associated with the scrambled response and the non-sensitive question respectively. Also, let t_{ij} be the response from individual j in sample i. Then the expectations

$$E_g(t) = \mu_A + \mu_s \text{ and } E_h(t) = \mu_Y \qquad (3.1)$$

where t is the response variable for the scrambled response and the unrelated character.

Further for each individual in a sample, we have

Sample 1. $\Psi_1(t_1) = p_1 g(t) + (1 - p_1) h(t)$ (3.2)

Sample 2. $\Psi_2(t_2) = h(t)$ (3.3)

where $\Psi_i(t_i)$ is the probability (or probability density) function of each individual in sample i. If T_i is the reported response of an individual in sample i, then

$$\mu_{t_1} = E(T_1) = p_1(\mu_A + \mu_s) + (1 - p_1)\mu_Y \tag{3.4}$$

$$\mu_{t_2} = E(T_2) = \mu_Y \tag{3.5}$$

We now consider the following estimator of μ_A

$$\hat{\mu}_A = \frac{\overline{T}_1 - (1 - p_1)\overline{T}_2}{p_1} - \mu_s \tag{3.6}$$

where $\overline{T}_1$ and $\overline{T}_2$ are the observed mean sample responses from the first and second sample respectively. Then we have the following theorems.

Theorem 3.1. The estimator $\hat{\mu}_A$ is unbiased for population mean μ_A.

Proof. We have

$$E(\hat{\mu}_A) = \frac{E(\overline{T}_1) - (1 - p_1)E(\overline{T}_2)}{p_1} - \mu_s$$

$$= \frac{p_1(\mu_A + \mu_s) + (1 - p_1)\mu_Y - (1 - p_1)\mu_Y}{p_1} - \mu_s$$

$$= \mu_A$$

Theorem 3.2. The variance of the proposed estimator $\hat{\mu}_A$ is

$$V(\hat{\mu}_A) = \frac{1}{p_1^2}\left[\frac{p_1\left\{(\sigma_A^2 + \sigma_s^2) - \sigma_Y^2\right\} + \sigma_Y^2 + p_1(1 - p_1)\{(\mu_A + \mu_s) - \mu_Y\}^2}{n_1} + \frac{(1 - p_1)^2\sigma_Y^2}{n_2}\right] \tag{3.7}$$

Proof. From (3.6) we have

$$V(\hat{\mu}_A) = \frac{1}{p_1^2}\left[\frac{V(T_1)}{n_1} + \frac{V(T_2)}{n_2}\right] \tag{3.8}$$

Now

$$V(T_1) = E(T_1^2) - \{E(T_1)\}^2$$

$$= p_1E_g(T_1^2) + (1 - p_1)E_h(T_1^2) - [p_1(\mu_A + \mu_s) + (1 - p_1)\mu_Y]^2$$

$$= p_1\left\{\sigma_A^2 + \sigma_s^2 + (\mu_A + \mu_s)^2\right\} + (1 - p_1)\left\{\sigma_Y^2 + \mu_Y^2\right\} - [p_1(\mu_A + \mu_s) + (1 - p_1)\mu_Y]^2$$

$$= p_1\left\{(\sigma_A^2 + \sigma_s^2) - \sigma_Y^2\right\} + \sigma_Y^2 + p_1(1 - p_1)\{(\mu_A + \mu_s) - \mu_Y\}^2$$

On substituting the value of $V(T_1)$ and $V(T_2) = \sigma_Y^2$ in (3.8), one gets (3.7).

The unbiased estimator of $V(\hat{\mu}_A)$ is immediate on replacing $V(T_1)$ and $V(T_2)$ in (3.8) by sample mean squares $s_1^2 = (n_1 - 1)^{-1}\left[\sum_{i=1}^{n_1} t_{i1}^2 - n_1 \overline{T}_1^2\right]$ and $s_2^2 = (n_2 - 1)^{-1}\left[\sum_{i=1}^{n_2} t_{i2}^2 - n_2 \overline{T}_2^2\right]$, respectively.

Next we are concerned with the optimal choice of p_1, non-sensitive character Y, random variable S and optimal allocation of n to n_1 and n_2. Differentiating $V(\hat{\mu}_A)$ in (3.7) with respect to p_1, one gets

$$\frac{\partial V(\hat{\mu}_A)}{\partial p_1} = -\frac{1}{p_1^2}\left[\frac{(\sigma_A^2 + \sigma_s^2)}{n_1} + \frac{\sigma_Y^2}{n_1}\left(\frac{2}{p_1} - 1\right) + \frac{\{(\mu_A + \mu_s) - \mu_Y\}^2}{n_1} + 2\left(\frac{1}{p_1} - 1\right)\frac{\sigma_Y^2}{n_2}\right] \tag{3.9}$$

which is always negative. The variance $V(\hat{\mu}_A)$ is, therefore, a decreasing function of p_1. The value of p_1 is, therefore, chosen close to 1 as much as possible without threatening the co-operation of the respondent. The value of p_1 close to 1 means somewhere in between $.8 \pm .1$ as recommended by Greenberg *et al.* [1]. The exact value of p_1 to be chosen depends on how high a value, the respondents are willing to tolerate. This is to be decided by the investigator keeping in view the respondents' attitude and co-operation.

So far as the choice of non-sensitive character and scrambling variable is concerned, one can easily choose Y and S in such a way so that $(\sigma_A^2 + \sigma_s^2 - \sigma_Y^2)$ and $(\mu_A + \mu_s - \mu_Y)$ are minimum. For this purpose good guess regarding $\sigma_A^2, \mu_A, \sigma_Y^2$ and μ_Y can be helpful. Minimizing σ_s^2 alone might help the interviewer to recognize the response as to which characteristic it could belong to. Therefore, one is to make a judicious choice of σ_s^2 keeping in view the choice of character Y. The optimal allocation of n_1 and n_2 can be obtained by equating $\frac{\partial V(\hat{\mu}_A)}{\partial n_1}$ to zero. This yields

$$\frac{n_1}{n_2} = \frac{\left[p_1\{(\sigma_A^2 + \sigma_s^2) - \sigma_Y^2\} + \sigma_Y^2 + p_1(1 - p_1)\{(\mu_A + \mu_s) - \mu_Y\}^2\right]^{1/2}}{(1 - p_1)\sigma_Y} \tag{3.10}$$

As already mentioned in the last paragraph of Section 2 that in case the Greenberg *et al.'s* [2] model with $p_2 = 0$ becomes non-functional, the investigator has to choose either the Greenberg *et al.'s* [2] model with $p_1 + p_2 = 1$ or the strategy proposed in the present paper. So in order to decide which of these two should be preferred, we compare the efficiency of the proposed estimator $\hat{\delta}_2$ with the usual one $\hat{\mu}_G$.

4. Efficiency Comparison

As compared to Greenberg *et al.'s* [2] procedure, the cost for collecting information from the respondents in the first sample by using the proposed strategy is expected to be little higher since the respondent is to make use of scrambling device in case the sensitive question is drawn from the RR device. However, the cost for collecting the information from the respondents in the second sample is quite small for the proposed procedure as the information is to be collected on the non-sensitive character only by using direct questions which can be done by mail/telephone. The efficiency comparison of both the procedures is, therefore, appropriate under the appropriate cost function.

Let the cost function be

$$C = C_0 + n_1\left[p_1 C_s + (1 - p_1) C_R\right] + n_2 C_D \tag{4.1}$$

where

$C =$ total budget at hand

$C_0 =$ the overhead cost

$C_s =$ the cost, per respondent, of obtaining scrambling response, by using RR device and scrambled variable

$C_R =$ the cost, per respondent, of obtaining the randomized response by using RR device and

$C_D =$ the cost, per respondent, of obtaining response by mail and/or telephone by using direct question

We obtain n_1 and n_2 such that the variance $V(\hat{\mu}_A)$ is minimum subject to cost constraint (4.1). If λ be the Lagrange multiplier, minimizing the function

$$\phi = V(\hat{\mu}_A) - \lambda\left[C - C_0 - n_1 C_1 - n_2 C_D\right]$$

with respect to n_1 and n_2, one gets

$$n_1 = \frac{\{C - C_0\}\sqrt{V(T_1)}}{K\sqrt{C_1}}$$

and $$n_2 = \frac{\{C - C_0\}(1 - p_1)\sqrt{V(T_2)}}{K\sqrt{C_D}}$$

where $$K = \sqrt{C_1}\sqrt{V(T_1)} + (1 - p_1)\sqrt{C_D}\sqrt{V(T_2)}$$

and $$C_1 = p_1 C_s + (1 - p_1) C_R$$

The variance in (3.8) in terms of n_1 and n_2 obtained above be written as

$$V_C(\hat{\mu}_A) = \frac{\left[\sqrt{C_1 V(T_1)} + (1 - p_1)\sqrt{C_D V(T_2)}\right]^2}{p_1^2 (C - C_0)} \tag{4.2}$$

Similarly the variance expression (2.2) in terms of cost can be written as

$$V_C(\hat{\mu}_G) = \frac{C_R\left[(1 - p_2)\sqrt{V(Z_1)} + (1 - p_1)\sqrt{V(Z_2)}\right]^2}{(p_1 - p_2)^2 (C - C_0)} \tag{4.3}$$

The proposed estimator $\hat{\mu}_A$ will be more efficient than the usual estimator $\hat{\mu}_G$ if

$$V_C(\hat{\mu}_A) < V_C(\hat{\mu}_G)$$

Under optimality conditions, $\sigma_A^2 + \sigma_S^2 = \sigma_Y^2$ and $\mu_A + \mu_S = \mu_Y$ for (4.2), and $\sigma_A^2 = \sigma_Y^2$ and $\mu_A = \mu_Y$ for (4.3) and letting $C_R = m_1 C_D$ and $C_S = m_2 C_R$, the above inequality reduces to

$$\sqrt{p_1 m_1 m_2 + (1 - p_1) m_1} + (1 - p_1) - \frac{p_1\sqrt{m_1}}{2p_1 - 1} < 0 \tag{4.4}$$

Thus we have

Theorem 4.1. The proposed strategy based estimator $\hat{\mu}_A$ will be more efficient than Greenberg *et al.* [2] with $p_1 + p_2 = 1$ strategy if the inequality (4.4) holds.

As the respondent in the first sample is to add random variable S to the randomized response obtained for sensitive character A, the cost C_S is expected to be slightly more than C_R. However the cost C_D for obtaining direct response on character Y from the respondents in the second sample is far less as compared to C_R. This means m_1 is much larger than m_2. So, the inequality (4.4) is expected to hold always in practice except few situations where p_1 is very high and m_1 is close to m_2. As this combination (p_1 very high and m_1 close to m_2) is expected to be encountered rarely, the proposed strategy will be preferred over the Greenberg *et al.* [2] model in almost all practical situations.

The magnitude of relative efficiencies obtained in Table 1 emphasize that the proposed strategy will yield much precise estimates than Greenberg *et al.* [2] model with $p_1 + p_2 = 1$. The strategy proposed is, therefore, advantageous in the situations where Greenberg *et al.* [2] with $p_2 = 0$ strategy becomes non-functional.

Table 1. Percent RE of the proposed estimator $\hat{\mu}_A$ with respect to the usual estimator $\hat{\mu}_G$

m_1	m_2	p_1				
		0.70	0.75	0.80	0.85	0.90
2	1.05	202.6	157.5	131.8	116.1	105.9
	1.10	197.1	152.8	127.5	111.9	101.8
	1.15	191.9	148.4	123.5	108.1	98.1
	1.20	187.0	144.2	119.7	104.5	94.5
	1.25	182.3	140.3	116.2	101.2	91.3
4	1.05	224.7	172.0	141.8	122.7	110.1
	1.10	218.3	166.7	137.0	118.3	105.7
	1.15	212.3	161.7	132.5	114.1	101.7
	1.20	206.5	156.9	128.3	110.2	98.0
	1.25	201.1	152.5	124.4	106.6	94.6
6	1.05	235.7	179.2	146.5	125.9	112.0
	1.10	228.8	173.5	141.5	121.2	107.5
	1.15	222.3	168.1	136.8	116.9	103.4
	1.20	216.2	163.1	132.4	112.9	99.6
	1.25	210.4	158.4	128.3	109.1	96.1
8	1.05	242.7	183.6	149.5	127.8	113.2
	1.10	235.5	177.7	144.3	123.0	108.6
	1.15	228.7	172.2	139.5	118.6	104.5
	1.20	222.3	167.0	135.0	114.5	100.6
	1.25	216.2	162.1	130.7	110.7	97.0
10	1.05	247.6	186.8	151.6	129.2	114.0
	1.10	240.1	180.7	146.3	124.3	109.4
	1.15	233.2	175.0	141.3	119.8	105.2
	1.20	226.6	169.7	136.7	115.7	101.3
	1.25	220.4	164.7	132.4	111.8	97.7

Remark 4.1. Generalization as Linear Model

As suggested by Warner [17] and later on Greenberg *et al.* [2] the proposed model can be treated as linear randomized model with some changes consistent with (3.1)-(3.5). One can write

$$Z = X\beta + U$$

where

$$Z' = \left[Z_{11}, Z_{12}, \ldots, Z_{1n_1}, Z_{21}, Z_{22}, \ldots, Z_{2n_2}\right]$$

$$X' = \begin{bmatrix} p_1, p_1, \ldots, p_1, 0, 0, \ldots, 0 \\ q_1, q_1, \ldots q_1, 1, 1, \ldots, 1 \end{bmatrix}$$

$$\beta' = \left[\mu_A^*, \mu_Y\right] \text{ and } U' = [U_1, U_2]$$

with

$$U_1 = \left[U_{11}, U_{12}, \ldots, U_{1n_1}\right]$$

$$U_2 = \left[U_{21}, U_{22}, \ldots, U_{2n_2}\right]$$

and

$$q_1 = 1 - p_1 \text{ and } \mu_A^* = \mu_A + \mu_S$$

As in the Greenberg *et al.'s* [2] model, E(U) is the null vector. The weighted least squares estimator of β is given by

$$\hat{\beta} = \begin{bmatrix} \hat{\mu}_A^* \\ \hat{\mu}_Y \end{bmatrix} = \left(X'V^{-1}X\right)^{-1} X'V^{-1}Z$$

where V is the diagonal matrix having $V(Z_1)$ for each first n_1 elements along the diagonal and $V(Z_2)$ for each of the succeeding n_2 elements along the diagonal. Obviously, $\hat{\mu}_A = \mu_A^* - \mu_S$ is the required estimator.

ACKNOWLEDGEMENTS

The authors' are thankful to the editor and referee for critically reviewing the original version of the manuscript. The opinions and results in this paper are of authors' and not necessary of any institute or organization.

REFERENCES

[1] Greenberg, B.G., Abul-Ela, A.L.A., Simmons, W.R. and Horvitz, D.G. (1969). The unrelated question randomized response model: theoretical framework. *Jour. Amer. Statist. Assoc.*, **64**, 520-539.

[2] Greenberg, B.G., Kuebler, R.R, Jr., Abernathy, J.R. and Horvitz, D.G. (1971). Application of the randomized response technique in obtaining quantitative data. *Jour. Amer. Statist. Assoc.*, **66**, 243-250.

[3] Horvitz, D.G., Greenberg, B.G. and Abernathy, J.R. (1976). Randomized response: a data gathering device for sensitive questions. *Int. Statist. Rev.*, **44**, 181-196.

[4] Horvitz, D.G., Shah, B.V. and Simmons, W.R. (1967). The unrelated question randomized response model. *Proc. Social Statist. Sec., Amer. Statist. Assoc.*, 65-72.

[5] Krishnamoorthy, K. and Raghavarao, D. (1993). Untruthful answering in repeated randomized response procedures. *Canadian. J. Statist.*, **21**, 233-236.

[6] .Lakshmi, D.V. and Raghavarao, D. (1992). A test for detecting untruthful answering in randomized response procedures. *J. Statist. Planning Infer.*, **31**, 387-390.

[7] Mahmood, M., Singh, S. and Horn, S. (1998). On the confidentiality guaranteed under randomized response sampling: A comparison with several new techniques. *Biom. J.*, **40(2)**, 237-242.

[8] Mangat, N.S. (1994). An improved randomized response strategy. *J. Roy. Statist. Soc.*, **B 56**, 93-95.

[9] Mangat, N.S. and Singh, R. (1990). An alternative randomized response procedure. *Biometrika*, **77**, 439-442.

[10] Mangat, N.S., Singh, R. and Singh, S. (1992). An improved unrelated question randomized response strategy. *Calcutta Statist. Assoc. Bull.*, **42**, 277-281.

[11] Pollock, K.H. and Bek, Y. (1976). A comparison of three randomized response models for quantitative data. *Jour. Amer. Statist. Assoc.*, **71**, 884-886.

[12] Raghavarao, D. (1978). On an estimation problem in Warner's randomized response technique. *Biometrics*, **34**, 87-90.

[13] Singh, S. (1999). An addendum to the confidentiality guaranteed in randomized response sampling by Mahmood, Singh and Horn. *Biom. J.*, **41(8)**, 955-966.

[14] Singh, S., Singh, R., Mangat, N.S. and Tracy, D.S. (1994). An alternative device for randomized responses. *Statistica*, **54**, 233-243.

[15] Singh, S., Singh, R. and Mangat, N.S. (2000). Some alternative strategies to Moor's model in randomized response sampling. *J. Statist. Plann. and Infer.*, **83**, 243-255.

[16] Warner, S.L. (1965). Randomized response: a survey technique for eliminating evasive answer bias. *Jour. Amer. Statist. Assoc.*, **60**, 63–69.

[17] Warner, S.L. (1971). The linear randomized response model. *Jour. Amer. Statist. Assoc.*, **66**, 884-888.