A New Multinomial Distribution Approach to Quantitative Randomized Response Technique

Jong-Min Kim[a] and William D. Warde[b]
[a] Division of Science and Math., University of Minnesota, Morris, MN , 56267, USA
[b] Department of Statistics, Oklahoma State University, Stillwater, OK , 74078, USA

Keywords: Multinomial Distribution; Randomized Response Design Matrix; Product-Moment Correlation.

Abstract

A new quantitative randomized response technique is presented in this paper. The proposed technique will use a Hopkins' randomizing device to derive a multinomial distribution for sensitive categories. After obtaining the observed estimates for sensitive category proportions which also include the random responses from the Hopkins' randomizing device, we derive the true estimates of the proportions for the sensitive categories in a situation where a model accounting for the respondent to lie is used. For contingency tables, we derive a Pearson product-moment correlation between two different sensitive questions.

Introduction and Literature Review

Since the introduction of the randomized response technique by Warner (1965), the theory and technique for randomized response (RR) technique have been considerably developed. Abul-Ela et al.(1967) extended Warner's dichotomous RR technique to a polychotomous RR technique but the Abul-Ela et al. RR technique had a drawback. The drawback is that the complexity of the estimation procedure increases as the number of categories in the polychotomy increases. There has been much research on enhancing RR techniques for polichotomies. In particular, Greenberg et al. (1971) adapted the unrelated question qualitative RR technique of Horvitz et al. (1967) to produce the unrelated question quantitative RR technique. A number of quantitative RR techniques have been proposed since Greenberg's quantitative RR technique. Bourke and Dalenius (1976) presented some new ideas in the realm of randomized response. They pointed out that Greenberg's quantitative RR technique leads to the loss of useful information on the sensitive trait because of the unrelated or nonsensitive question in the quantitative RR technique. To deal with the disadvantage of Greenberg's quantitative RR technique, Eriksson (1973) and Liu and Chow (1976) presented discrete quantitative RR techniques which modified the Greenberg quantitative RR technique. Kim and Flueck (1978) and Himmelfarb and Edgell (1980) developed the additive model approach to RR technique. Pollock and Bek (1976) and Eichhorn and Hayre (1983) introduced the multiplicative RR technique which is the method where a respondent multiplies his or her answer to the sensitive question by a random number from a known distribution. Therefore a validation check for RR technique has also been attempted by Abernathy et al. (1970), Bradburn and Sudman (1979), Tracy and Fox (1981), Danermark and Swensson (1987), Duffy and Waterton (1988), Han (1993) and Kerkvliet (1994). These researchers compared RR interviews and direct interviews based on a statistical measure of efficiency and respondents' protection.

Estimation of Proportions in a Multinomial Distribution

Our RR technique utilizes the Hopkins' device to estimate a multinomial distribution for a sensitive variable (A). Thus our new quantitative RR technique follows the same procedure as Liu and Chow's (1976) RR technique. There are two different colors of balls, red and green, in the device. Each of the green balls has a discrete number marked on it, $1,2,\cdots,k+1$. Suppose that all green balls consist of a set of non-sensitive categories, $B=\{B_1,B_2,\cdots,B_{k+1}\}$, such that all the values of sensitive categories $A=\{A_1,A_2,\cdots,A_{k+1}\}$ are included. With t different interviewees performing the Hopkins' device, each interviewee belongs to one of $k+1$ mutually exclusive and exhaustive categories $T=\{T_1,T_2,\cdots,T_{k+1}\}$ which consist of sensitive categories and non-sensitive categories. Let t_i denote the number of observations in a category T_i so that $t=\sum_{i=1}^{k+1} t_i$. We let a_i be the number of observations in a category A_i so that $a=\sum_{i=1}^{k+1} a_i$ and b_i be the number of observations in a category B_i so that $b=\sum_{i=1}^{k+1} b_i$. We assume that $T_i=t_i$ is the sum of $A_i=a_i$ and $B_i=b_i$. Thus we are attempting to estimate $P_{a1},P_{a2},\ldots,P_{a(k+1)}$ the proportions in the population who are in sensitive categories $A_1,A_2,\cdots,A_{k+1}$. Based on green balls with number in the Hopkins' device, we can derive the proportions in the population who are in categories $B_1,B_2,\cdots,B_{k+1}$ by $P_{bi}=g_i/g$. Let $P_{t1},P_{t2},\ldots,P_{t(k+1)}$ denote the proportions in the population who are in categories $T_1,T_2,\cdots,T_{k+1}$. When t different interviewees finish performing the Hopkins' device, we can derive b the total number of people who are in $B = \{B_1, B_2, \cdots, B_{k+1}\}$ by $b \geq tg/(r+g)$ where b is an integer. We can also derive $b_1, b_2, \ldots, b_k$ in the same way that $b_i \geq tg_i/(r+g_i)$ where b_i is an integer. Thus, we can get $b_{k+1} = b - (b_1 + b_2 + \ldots + b_k)$. Then we can define a multinomial distribution of T, A and B as follow:

$$T = (T_1, T_2, \ldots, T_k) \sim \text{MULT}\,(t, P_{t1,} P_{t2}, \ldots, P_{tk})$$
$$A = (A_1, A_2, \ldots, A_k) \sim \text{MULT}\,(a, P_{a1,} P_{a2}, \ldots, P_{ak})$$
$$B = (B_1, B_2, \ldots, B_k) \sim \text{MULT}\,(b, P_{b1,} P_{b2}, \ldots, P_{bk}) .$$

Suppose that $T = A + B$ and respondents give truthful answers to one of two different questions. From the moment generating functions of T, A and B or directly from the marginal probability mass functions, we can compute moments. So

$$E(T_h) = tP_{th}, \quad E(A_h) = aP_{ah} \text{ and } E(B_h) = bP_{bh}$$

where $h = 1,2,\ldots,k+1$.

For $T_h = A_h + B_h$,

$$E(A_h) = E(T_h) - E(B_h) = tP_{th} - bP_{bh} .$$

Since $E(A_h) = aP_{ah}$,

$$P_{ah} = \frac{tP_{th} - bP_{bh}}{a} = \frac{tP_{th} - bP_{bh}}{t-b} .$$

Let $\hat{P}_{ah}$ denote the estimate of P_{ah} and $\hat{P}_{th}$ denote the estimate of P_{th}. Since $\hat{P}_{bh} = g_h/g$,

$$\hat{P}_{ah} = \frac{t\hat{P}_{th} - b(g_h/g)}{t-b}$$

which is an unbiased estimator of P_{ah}.

The estimate of variance is

$$v\left(\hat{P}_{ah}\right) = \frac{t\hat{P}_{th}(1-\hat{P}_{th})}{(t-b)^2} .$$

The estimate of covariance is

$$C\hat{o}v\left(\hat{P}_{ah}, \hat{P}_{ai}\right) = -\frac{t\hat{P}_{th}\hat{P}_{ti}}{(t-b)^2} \qquad \text{where } h \neq i .$$

A Random Transformation to the True Estimate

In the previous section, we assumed that respondents report truthfully. But in a case of untruthful reporting, we need to derive an estimator for population proportion P_{ah} with the prior information when a respondent reports untruthfully. Let R_{ij} denote the probability that a person of category i announces himself or herself as one of category j. Suppose that respondents report truthfully when they have a non-sensitive question. Then we can apply the lying model of Mukhopadhyay (1980) to the sensitive question. Assume that there is a sensitive category A_i for $i = 1,2,3,4$ such that A_1 has no social stigma and that there is more social stigma as i increases. Intuitively, we can stipulate the following:

$$R_{12} = R_{13} = R_{14} = R_{23} = R_{24} = R_{34} = 0, \quad R_{11} = 1, R_{21} + R_{22} = 1,$$
$$R_{31} + R_{32} + R_{33} = 1 \quad \text{and} \quad R_{41} + R_{42} + R_{43} + R_{44} = 1.$$

Let π_i represent the true proportion of respondents who belong to a sensitive category i and P_{ai} represent the observed proportion of respondents who belong to a sensitive category i. Under the assumptions, we can derive the following:

$$P_{a1} = R_{11}\pi_1 + R_{21}\pi_2 + R_{31}\pi_3 + R_{41}\pi_4 = \pi_1 + R_{21}\pi_2 + R_{31}\pi_3 + R_{41}\pi_4$$
$$P_{a2} = R_{12}\pi_1 + R_{22}\pi_2 + R_{32}\pi_3 + R_{42}\pi_4 = R_{22}\pi_2 + R_{32}\pi_3 + R_{42}\pi_4$$
$$P_{a3} = R_{13}\pi_1 + R_{23}\pi_2 + R_{33}\pi_3 + R_{43}\pi_4 = R_{33}\pi_3 + R_{43}\pi_4$$
$$P_{a4} = R_{14}\pi_1 + R_{24}\pi_2 + R_{34}\pi_3 + R_{44}\pi_4 = R_{44}\pi_4 .$$

Then

$$P = \begin{pmatrix} P_{a1} \\ P_{a2} \\ P_{a3} \\ P_{a4} \end{pmatrix} = \begin{pmatrix} 1 & R_{21} & R_{31} & R_{41} \\ 0 & R_{22} & R_{32} & R_{42} \\ 0 & 0 & R_{33} & R_{43} \\ 0 & 0 & 0 & R_{44} \end{pmatrix} \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & R_{21} & R_{31} & R_{41} \\ 0 & 1-R_{21} & R_{32} & R_{42} \\ 0 & 0 & 1-R_{31}-R_{32} & R_{43} \\ 0 & 0 & 0 & 1-R_{41}-R_{42}-R_{43} \end{pmatrix} \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{pmatrix} .$$

We can extend the four-category sensitive case to the k-category sensitive case. Assume that there is a sensitive category A_i for $i = 1,2,\ldots,k$ such that A_1 has no social stigma and A_i is more social stigma as i increases. We can stipulate the following:

$$R_{ij} = 0 \quad \text{if} \quad i < j \quad \text{where } i, j = 1,2,\ldots,k$$
$$\sum_{j=1}^{k} R_{ij} = 1 \quad \text{for } i = 1,2,\ldots,k .$$

We can derive the following:

$$P_{aj} = \sum_{i=1}^{k} R_{ij}\pi_i \quad \text{for } j = 1,2,\ldots,k .$$

Then

$$P = \begin{pmatrix} P_{a1} & P_{a2} & \cdots & P_{a(k-1)} & P_{ak} \end{pmatrix}^T = R\pi$$

where $\pi = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_{k-1} & \pi_k \end{pmatrix}^T$ and

$$R = \begin{pmatrix} 1 & R_{21} & \cdots & R_{(k-1)1} & R_{k1} \\ 0 & 1-R_{21} & \cdots & R_{(k-1)2} & R_{k2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1-\sum_{j=1}^{k-2} R_{(k-1)j} & R_{k(k-1)} \\ 0 & 0 & 0 & 0 & 1-\sum_{j=1}^{k-1} R_{kj} \end{pmatrix}$$

If R is nonsingular, we can derive the true proportions for sensitive categories:

$$\pi = R^{-1}P .$$

The Maximum Likelihood estimator of $\pi = (\pi_1 \quad \pi_2 \quad \cdots \quad \pi_{k-1} \quad \pi_k)^T$ is given by

$$\hat{\pi} = R^{-1}\hat{P} .$$

where $\hat{P}$ is a estimate vector of P, provided that the vector $\hat{\pi}$ satisfies $\hat{\pi}_i \geq 0$ for all i and $\sum_{i=1}^{4} \hat{\pi}_i = 1$. The estimate of covariance of $\hat{\pi}$ is

$$C\hat{o}v(\hat{\pi}) = R^{-1}C\hat{o}v(\hat{P})(R^{-1})^T$$

where

$$C\hat{o}v(\hat{P}_{ai}, \hat{P}_{aj}) = \begin{cases} \dfrac{t\hat{P}_{ti}(1-\hat{P}_{ti})}{(t-b)^2} & if\ i = j \\ -\dfrac{t\hat{P}_{ti}\hat{P}_{tj}}{(t-b)^2} & if\ i \neq j \end{cases} .$$

Correlation between Two Different Sensitive Questions

Fox and Tracy (1984) considered estimating the correlation between two sensitive variables which are surveyed under the quantitative RR technique by Greenberg et al. (1971). In this paper, we will consider estimating the correlation between two sensitive variables which is based on a new quantitative RR technique. For an interview involving two sensitive questions, a researcher prepares two Hopkins' devices which have different ratios of red balls and green balls with designated numbers. An interviewee will face two devices so that she or he will use different device for each sensitive question independently. For each question, the respondent will shake the device and will get a ball. If the ball is red then the respondent should answer the sensitive question. Otherwise, if the ball is green with a designated number then the respondent will just say the number on the green ball. For two different questions, we are going to use the multivariate randomized response design of Bourke (1981). We denote θ_{ij} to be the probability that a respondent gives the i th category for the first question and the j th category for the second question. Let P_{ij} denote the true proportion of respondents who fall in the i th category for the first question and the j th category for the second question. Suppose that the first question has I categories and the second question has J categories. For the conditional probability $P[k\,l \mid i\ j]$ that a respondent of category i and category j announces himself or herself as one of category k and category l, we have

$$\theta_{ij} = \sum_{i=1}^{I}\sum_{j=1}^{J} P[k\,l \mid i\ j]\lambda_{ij}$$

where λ_{ij} is the true proportion that a respondent belongs in the i th category for the first question and belongs in the j th category for the second question. Since two devices are independently performed by a respondent, we can write $P[k\,l \mid i\ j] = P_1[k \mid i]P_2[l \mid j]$. It can be rewritten like this:

$$\theta_{ij} = \sum_{i=1}^{I}\sum_{j=1}^{J} P_1[k \mid i]P_2[l \mid j]\lambda_{ij} .$$

By Bourke (1981), we can express the vectors $\theta^{(2)}$ and $\lambda^{(2)}$ so that θ_{ij} is the r th element of the vector $\theta^{(2)}$ and λ_{ij} is the c th element of the vector $\lambda^{(2)}$, where

$$r = l + (k-1)I , \quad c = j + (i-1)J .$$

We can express that $P[k\,l \mid i\ j] = P[k \mid i]P[l \mid j]$ is in the (r, c) position of the matrix $M^{(2)}$. The $M^{(2)}$ is the Kronecker product of two matrixes M_1 and M_2 so that $P[k \mid i]$ is a element of M_1 and $P[l \mid j]$ is a element of M_2. Therefore the vector $\theta^{(2)}$ can be expressed as follows:

$$\theta^{(2)} = M^{(2)}\lambda^{(2)} = (M_1 \otimes M_2)\lambda^{(2)} .$$

If $M_1 \otimes M_2$ is nonsingular, we can derive $\lambda^{(2)}$ as follows:

$$\lambda^{(2)} = (M_1 \otimes M_2)^{-1}\theta^{(2)} .$$

If $\hat{\theta}^{(2)}$ is the asymptotic Maximum Likelihood estimator of $\theta^{(2)}$ then we can estimate

$$\hat{\lambda}^{(2)} = (M_1 \otimes M_2)^{-1}\hat{\theta}^{(2)} .$$

Using these cell proportions $\hat{\lambda}^{(2)}$, we can consider the product moment correlation between two sensitive variables ($A^{(1)}$ and $A^{(2)}$). From the interview, we can directly estimate the Pearson product-moment correlation between two different variables ($T^{(1)} = A^{(1)} + B^{(1)}$ and $T^{(2)} = A^{(2)} + B^{(2)}$). When $T^{(1)}$ is a row variable and $T^{(2)}$ is a column variable, we let $A(r_i)$ denote a value assigned to the i th row category, and $A(c_j)$ denote a value assigned to the j th column category. Suppose $A(r_1) \leq A(r_2) \leq \cdots \leq A(r_I)$ and $A(c_1) \leq A(c_2) \leq \cdots \leq A(c_J)$.

For $I \times J$ contingency table

$$\rho_T = \frac{Cov\left(T^{(1)}, T^{(2)}\right)}{\sqrt{Var(T^{(1)})}\sqrt{Var(T^{(2)})}}$$

$$= \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} \lambda_{ij}\left(A(r_i) - A(\bar{r})\right)\left(A(c_j) - A(\bar{c})\right)}{\sqrt{\left\{\sum_{i=1}^{I} \lambda_{i+}\left(A(r_i) - A(\bar{r})\right)^2\right\}\left\{\sum_{j=1}^{J} \lambda_{+j}\left(A(c_j) - A(\bar{c})\right)^2\right\}}}$$

where $\lambda_{i+} = \sum_{j=1}^{J} \lambda_{ij}$, and $\lambda_{+j} = \sum_{i=1}^{I} \lambda_{ij}$, and

$$A(\bar{r}) = \sum_{i=1}^{I} \lambda_{i+}A(r_i) \text{ and } A(\bar{c}) = \sum_{j=1}^{J} \lambda_{+j}A(c_j) .$$

The estimator is

$$r_T = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{\lambda}_{ij}\left(A(r_i) - A(\hat{\bar{r}})\right)\left(A(c_j) - A(\hat{\bar{c}})\right)}{\sqrt{\left\{\sum_{i=1}^{I}\hat{\lambda}_{i+}\left(A(r_i) - A(\hat{\bar{r}})\right)^2\right\}\left\{\sum_{j=1}^{J}\hat{\lambda}_{+j}\left(A(c_j) - A(\hat{\bar{c}})\right)^2\right\}}}$$

where $A(\hat{\bar{r}}) = \sum_{i=1}^{I}\left(\hat{\lambda}_{i+}/\hat{\lambda}_{++}\right)A(r_i)$ and $A(\hat{\bar{c}}) = \sum_{j=1}^{J}\left(\hat{\lambda}_{+j}/\hat{\lambda}_{++}\right)A(c_j)$.

Since $A^{(1)}$ and $B^{(1)}$ are independent, and $A^{(2)}$ and $B^{(2)}$ are independent. Then

$$Var(T^{(1)}) = Var(A^{(1)}) + Var(B^{(1)}) \quad \text{and}$$
$$Var(T^{(2)}) = Var(A^{(2)}) + Var(B^{(2)}).$$

Suppose $A^{(1)}$ and $B^{(2)}$, $A^{(2)}$ and $B^{(1)}$, and $B^{(1)}$ and $B^{(2)}$ are uncorrelated each other. Then the covariance of two variable $T^{(1)}$ and $T^{(2)}$ is

$$\begin{aligned} Cov(T^{(1)}, T^{(2)}) &= Cov(A^{(1)} + B^{(1)}, A^{(2)} + B^{(2)}) \\ &= Cov(A^{(1)}, A^{(2)}) + Cov(A^{(1)}, B^{(2)}) \\ &\quad + Cov(B^{(1)}, A^{(2)}) + Cov(B^{(1)}, B^{(2)}) \\ &= Cov(A^{(1)}, A^{(2)}). \end{aligned}$$

The product moment correlation between two sensitive variables $A^{(1)}$ and $A^{(2)}$ is

$$\rho_A = \frac{Cov\left(A^{(1)}, A^{(2)}\right)}{\sqrt{Var(A^{(1)})}\sqrt{Var(A^{(2)})}}$$

$$= \frac{Cov\left(T^{(1)}, T^{(2)}\right)}{\sqrt{Var(T^{(1)}) - Var(B^{(1)})}\sqrt{Var(T^{(1)}) - Var(B^{(1)})}}$$

$$= \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\lambda_{ij}\left(A(r_i) - A(\bar{r})\right)\left(A(c_j) - A(\bar{c})\right)}{\sqrt{\left\{\sum_{i=1}^{I}\lambda_{i+}\left(A(r_i) - A(\bar{r})\right)^2 - Var\left(B^{(1)}\right)\right\}\left\{\sum_{j=1}^{J}\lambda_{+j}\left(A(c_j) - A(\bar{c})\right)^2 - Var\left(B^{(2)}\right)\right\}}}$$

From Eriksson(1973), we can derive the mean and variance of a designated number i :

$$\mu_B = \sum_{i=1}^{k} i\,\frac{g_i}{1-P} \quad \text{and} \quad Var(B) = \sum_{i=1}^{k}(i - \mu_B)^2\,\frac{g_i}{1-P}$$

where the proportion of green balls with designated number i is g_i such that $1 - P = \sum_{i=1}^{m} g_i$. If a researcher uses two Hopkins' devices which have different ratios of red balls and green balls with a designated number then she or he can derive the variances of a designated number i, that is, $Var(B^{(1)})$ and $Var(B^{(2)})$.

The estimator of ρ_A is

$$r_A = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{\lambda}_{ij}\left(A(r_i) - A(\hat{\bar{r}})\right)\left(A(c_j) - A(\hat{\bar{c}})\right)}{\sqrt{\left\{\sum_{i=1}^{I}\hat{\lambda}_{i+}\left(A(r_i) - A(\hat{\bar{r}})\right)^2 - Var\left(B^{(1)}\right)\right\}\left\{\sum_{j=1}^{J}\hat{\lambda}_{+j}\left(A(c_j) - A(\hat{\bar{c}})\right)^2 - Var\left(B^{(2)}\right)\right\}}}$$

where $A(\hat{\bar{r}}) = \sum_{i=1}^{I}\left(\hat{\lambda}_{i+}/\hat{\lambda}_{++}\right)A(r_i)$ and $A(\hat{\bar{c}}) = \sum_{j=1}^{J}\left(\hat{\lambda}_{+j}/\hat{\lambda}_{++}\right)A(c_j)$.

If the value of r_A equals zero then it means that two sensitive variables $A^{(1)}$ and $A^{(2)}$ are independent. The farther the absolute value of r_A is from zero, the stronger the relationship between two sensitive variables $A^{(1)}$ and $A^{(2)}$ correlate with each other.

Discussion

Eriksson (1973) and Liu and Chow (1976) have presented a quantitative randomized response technique which is modified by Greenberg et al. (1971). But the result of their researches focused on estimating the proportions which are the observed estimates of sensitive category proportions. Furthermore they did not apply their randomized response models to the multivariate randomized response design for a sensitive variable. So the multinomial distribution approach to a new RR technique using a Hopkins' device was introduced. It is advantageous to treat ordinal data in a quantitative manner by assigning ordered scores to the categories. In a new quantitative RR technique, we derived the true proportion estimates of the sensitive categories based on the observed estimates of sensitive category proportions. A Pearson product-moment correlation between two sensitive variables was presented in this research. Since researchers often deal with categorical data of sensitive issues in a real life, the Pearson product-moment correlation is more appropriate than the correlation between two sensitive variables presented by Fox and Tracy (1984). Through the Pearson product-moment correlation presented in this research, researchers may discover the important fact that if researchers choose two sensitive issues highly correlated then they may obtain more useful information, for example, like the correlation between abortion and alcohol abuse, in addition to get a reliable data.

References

Abernathy, J.R., Greenberg, B.G., and Horvitz, D.G., 1970. Estimates of induced abortion in urban North Carolina.. Demography, 7, 19-29.

Abul-Ela, A.A., Greenberg, B.G., and Horvitz, D.G., 1967. A multiproportion randomized response model. Journal of American Statistical Association, 62, 990-1008.

Bourke, P.D., 1981. On the analysis of some multivariate randomized response designs for categorical data. Journal of Statistical Planning and Inference, 5, 165-170.

Bourke, P.D., 1982. Randomized response multivariate designs for categorical data. Communications in Statistics: Theory and Methods, 11, 2889-2901.

Bourke, P.D., and Dalenius, T., 1976. Some new ideas in the realm of randomized inquires. International Statistics Review, 44, 219-221.

Bradburn, N.M., and Sudman, S., 1979. Improving interview method and questionnaire design. San Francisco: Jossey-Bass.

Danermark, B., and Swensson, B., 1987. Measuring drug use among Swedish adolescents: randomized response versus anonymous questionnaires. Journal of Official Statistics, 3, 439-448.

Duffy, J.C., and Waterton, J.J., 1988. Randomized response vs. direct questioning: estimating the prevalence of alcohol related problems in a field survey. Australian Journal of Statistics, 30, 1-14.

Eichhorn, B.H., and Hayre, L.S., 1983. Scrambled randomized methods for obtaining sensitive quantitative data. Journal of Statistical Planning and Inference, 7, 307-316.

Eriksson, S.A., 1973. A new model for randomized response. International Statistics Review, 41, 101-113.

Fox, J.A., and Tracy, P.E., 1984. Measuring associations with randomized response. Social Science Research.13, 188-197.

Greenberg, B.G., Kuebler, R.R., Abernathy, J.R. and Horvitz, D.G., 1971. Applications of the RR technique in obtaining qualitative data. . Journal of American Statistical Association, 66, 243-250.

Han, G., 1993.Correaltion analysis for the randomized response models. Ph.D. Dissertation, Oklahoma State Univ.

Himmelfarb, S., and Edgell, S.E., 1980. Additive constants model: A randomized response technique for eliminating evasiveness to quantitative response questions. Psychological Bulletin, 87, 525-530.

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

Kerkvliet, J., 1994. Estimating a logit model with randomized data: the case of cocaine use. . Australian Journal of Statistics, 36, 9-20.

Kim, J-I., and Flueck, J.A., 1978. An additive randomized response model. Proceedings of Social Statistics Section, American Statistical Association, 351-355.

Liu, P.T., and Chow, L.P., 1976. A new discrete quantitative randomized response models for quantitative data. Journal of American Statistical Association, 71, 72-73.

Mukhopadhyay, P., 1980. On estimation of some confidential characters from survey data. Calcutta Statistical Association Bulletin , 29, 133-141.

Pollock, K.H., and Beck, Y., 1976. A comparison of three randomized response models for quantitative data. Journal of American Statistical Association 71, 884-886.

Tracy, P.E., and Fox, J.A., 1981. The validity of randomized response for sensitive measurements. American Sociological Review, 46, 187-200.

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