A BIVARIATE PROBABILITY MODEL TO IDENTIFY "HONESTY" VERSUS "CHEATING" IN ECONOMIC SURVEYS: XENOPHOBIA IS ILLUSTRATED

Ramalingam Shanmugam

School of Health Administration, Texas State University, San Marcos, Texas 78666, USA

ABSTRACT

Making successful policies for manufacturing or globalized business sectors require "honest" answers by the respondents in the surveys. When some respondents cheat in their answers to the questions in a survey for a variety of reasons, the survey results become useless. Currently, there is no appropriate probability model to address this phenomenon "cheating" in a survey. An efficient methodology is needed to estimate the proportion of "cheaters" in a survey and assess the statistical significance of the estimate. Such a methodology does not exist now in the literature. This article fulfils the need. In a pioneering manner, this article has formulated a bivariate probability model to explain the "cheating" in surveys. Several statistical properties of this new model are identified and explained. For illustration of the new probability model, the responses from 259 less and 217 more educated Germans about xenophobia are considered. The xenophobia is a stumbling block to a successful business operation in this globalized economy. Two psychometric reasons: "Insecurity level" and "social pressure" behind the xenophobia are captured with the help of the probability model and they are explained. The model of this article helps to predict "honesty" versus "cheating" levels in a survey. A statistical testing procedure is prepared to check whether an estimated "social pressure" level is significant.

1. MOTIVATION

When the questions in a survey are about a sensitive topic or one's illegal practice, the respondents are rightfully unwilling to give an answer. Conducting sensitive as much non-sensitive survey is common in social, health, epidemiologic and economic surveys. See Yahya and Adebayo (2013) for a health survey about the breast feeding by mothers. To circumvent such practical difficulties in a survey, Warner (1965) came up with an ingenious and pioneering methodology and named it Randomized Response Technique (RRT). In a RRT, every respondent uses a random device (such as rolling a die) in a private room and selects the sensitive questionnaire-I to answer if the outcome is such and such of a prescription or selects the questionnaire-II to answer if the outcome is outside

the prescription. The respondent does not have to reveal the outcome of the random device to anyone and hence, no one will ever know which questionnaire was answered by a particular respondent. This approach offers confidentiality to every respondent and hence, the RRT based survey increases the likelihood of obtaining truthful answers from the respondents.

Since Warner (1965), many articles and books have been written by statisticians, sociologists, psychologists, economists, epidemiologists and marketing researchers among other professionals. Most important articles and books are worth mentioning here in a chronological order. Greenberg *et al*. (1969) adapted unrelated questions in RRT. Campbell and Joiner (1973) suggested about how to get the answer without being sure you've asked the question. Campbell (1987) popularized among all scientists. Goodstadt and Gruson (1975) used the

RRT to test the efficacy of a drug. Maceli (1978) demonstrated about how to ask sensitive questions without getting punched in the nose. See Chaudhuri and Mukerjee (1989) for an excellent comprehension of the RRT. Clark and Desharnais (1998) introduced an approach to obtain honest answers to embarrassing questions and to detect cheating in an RRT. Fox (2005) applied RRT in item response theory models of educational studies. Fan and Chaloner (2006) designed a trinomial RRT response to configure an optimal dose level in clinical trials. Guerriero and Sandri (2007) compared randomized response procedures. Lynn (2008) advised on how to deal with the non-response in RRT. Musch *et al*. (2001) promoted RRT in survey research on the World Wide Web. Ostapczuk *et al*. (2009a) assessed sensitive attributes using the RRT and the evidence for the importance of response symmetry. Using the wonderful RRT, (Ostapczuk *et al*., 2009b; Musch *et al*., 2001) pondered over whether the education has a negative effect in resentful attitudes towards foreigners and this phenomenon is called xenophobia. Interestingly, they first classified the survey outcomes in terms of "honest yes", "honest no" and "cheating" and later provided reasons for xenophobia and questioned whether those with low education exhibited xenophobia more than those with high education among the respondents. Their article was somewhat incomplete only in a sense that an appropriate theoretical framework was missing to perform the statistical significance of the survey based estimate of the proportion of "honest yes", of "honest no" and of "cheating". To add supplementary concepts and tools, this article develops a bivariate probability model and utilizes it to address the "honesty" versus "cheating" in any survey. Using xenophobia data in Ostapczuk *et al*. (2009b), the contents of this article are explained and interpreted. A few comments and directions are mentioned in the end for future research work to improve survey methodology to capture truthful responses.

2. A BIVARIATE MODEL FOR "HONESTY" VERSUS "CHEATING" IN SURVEYS

In this section, we develop a bivariate probability model from the basics of the psychometric characters of the respondents in a survey about the xenophobia. A respondent (among a random sample of size n) might be xenophobic because of personal insecurity or social pressure. Recently, Ankudinov and Lebedev (2014) studied the role of insecurity among employees to engage in professional education. These two factors are

psychometric characters and they constitute dominant reasons for a respondent to be honest or cheating type in a survey. Neither factor is measurable directly and hence, are treated as parameters in this article.

To be specific, let $0 < \phi < 1$ be an unknown probability for a respondent to be "insecure". Independently, a respondent might yield to "social pressure" with a probability $0 < \rho < 1$. It is worth noticing that bothe parameters are open ended in the interval (0, 1) implying that the framework is meaningless if all respondents in the survey are "insecure" or yielding to "social pressure". There are four distinct and mutually exclusive possibilities for any respondent to be. That is, some respondents in a survey might be "insecure cheaters" with a probability $0 < \phi\rho < 1$, might be "insecure honesters" with a probability $0 < \phi(1-\rho) < 1$ to answer "yes", might be "secure cheaters" with a probability $0 < (1-\phi)\rho < 1$, or might be "secure honesters" to answer "no" with a probability $0 < (1-\phi)(1-\rho) < 1$. Every respondent ought to fall in any one of the four mutually exclusive possibilities.

Let X, Y and n-X-Y denote respectively the number of "honest yes", "honest no" and "cheaters" in a survey answered by n respondents. Then, the bona-fide bivariate model for X and Y is Equation 1:

$$\Pr[X = x, Y = y] =$$
$$\frac{n!}{x!y!(n-x-y)!} \rho^n [\phi(\frac{1}{\rho}-1)]^x [(1-\phi)(\frac{1}{\rho}-1)]^y;$$
$$0 < \phi < 1; 0 < \rho < 1;$$
$$x = 0,1,2,...,n; y = 0,1,2,...,n.$$
(1)

The number, *X* of "honest yes" and the number, *Y* of "honest no" marginally follow respectively a probability pattern Equation 2 and 3:

$$\Pr[X = x] = \sum_{y=0}^{n} \Pr[X = x, Y = y]$$
$$= \frac{n!}{x!(n-x)!} \rho^n [\phi(\frac{1}{\rho}-1)]^x [[\frac{1}{\rho} - \phi(\frac{1}{\rho}-1)]^{n-x}$$
$$0 < \phi < 1; 0 < \rho < 1; x = 0,1,2,...,n;$$
(2)

And:

$$\Pr[Y = y] = \sum_{x=0}^{n} \Pr[X = x, Y = y]$$
$$= \frac{n!}{y!(n-y)!} \rho^n [(1-\phi)(\frac{1}{\rho}-1)]^y [1 + \phi(\frac{1}{\rho}-1)]^{n-y}$$
$$0 < \phi < 1; 0 < \rho < 1; y = 0,1,2,...,n.$$
(3)

The expected value and variance of X, the number of "honest yes" are respectively nonlinear functions:

$$E[X = x] = n\phi(1 - \rho)$$

And:

$$Var[X = x] = [1 - \phi(1 - \rho)]E[X = x]$$

Likewise, the expected value and variance of Y, the number of "honest no" are respectively nonlinear functions:

$$E[Y = y] = n(1 - \phi)(1 - \rho)$$

And:

$$Var[Y = y] = [1 - (1 - \phi)(1 - \rho)]E[Y = y]$$

Furthermore, there exists an intrinsic relation between X and Y. That is:

$$\frac{Var(X)}{E(X)} + \frac{Var(Y)}{E(Y)} = 1 + \rho$$

Prompting that X and Y might be correlated. What is their correlation? The statistical dependence among the observed variables remains of vital interest to data analysts see (Nasser, 2007, for details). The correlations have been the basis of the connection between two measurable factors. Recently, Olatayo (2011) used the correlation to establish the similarities and differences between two minerals. Using the Probability Mass Function (PMF) in (1), the correlation between X, the number of "honest yes" and Y, the number of "honest no" is found and it is Equation 4:

$$corr(X, Y)$$
$$= -\sqrt{\frac{\phi(1 - \phi)}{[1 - \phi(1 - \rho)][1 - (1 - \phi)(1 - \rho)]}} \tag{4}$$

which is asymptotically near minus one when there is a negligible level of social pressure (that is, $\rho \rightarrow 0$) and it increases monotonically to $-\sqrt{\phi(1 - \phi)}$ when the social pressure is an increasing to its full level (that is, $\rho \rightarrow 1$).

Hence, the predictability of "honest yes" based on a known number of "honest no" and vice versa are quite possible. For that, their conditional PMFs are required. First, the conditional PMF of Y, the number of "honest no" answers given $X = x$, the number of "honest yes" answers is found and it is Equation 5:

$$Pr[Y = y | X = x] = \frac{Pr[x, y]}{P[x]}$$
$$= \frac{(n - x)!}{y!(n - x - y)!} \frac{[(1 - \phi)(\frac{1}{\rho} - 1)]^y}{[\frac{1}{\rho} - \phi(\frac{1}{\rho} - 1)]^{n-x}}; \tag{5}$$
$$0 < \phi < 1; 0 < \rho < 1; y = 0, 1, 2, ..., n - x.$$

The conditional mean, $E[Y = y | X = x]$ is really the regression of y for a given x with variance, $Var[Y = y | X = X]$. They are derived from the PMF (5) and are expressed as:

$$E[Y = y | X = x] = \beta_0 - \beta_1 x,$$

And:

$$Var[Y = y | X = x] = (\beta_0 - \beta_1 x)[1 - \frac{(\beta_0 - \beta_1 x)}{(n - x)}]$$

The intercept is $\dfrac{n(1 - \phi)(\frac{1}{\rho} - 1)}{[\frac{1}{\rho} - \phi(\frac{1}{\rho} - 1)]}$ which increases with

the number of participants, n in the survey and the slope

is $\dfrac{(1 - \phi)(\frac{1}{\rho} - 1)}{[\frac{1}{\rho} - \phi(\frac{1}{\rho} - 1)]}$ which increases as the intercept increases.

The conditional variance is recognized as heterogeneity level in the statistics. The randomized response sample survey is therefore quite heterogeneous. The heterogeneity plays a significant role in scientific enquires. Chiarella and He (2005) utilized heterogeneity to comprehend the dynamics among the producers.

Likewise, the conditional probability mass function of X, the number of "honest yes" answers given $Y = y$, the number of "honest no" is Equation 6:

$$Pr[X = x | Y = y] = \frac{Pr[x, y]}{P[y]}$$
$$= \frac{(n - y)!}{x!(n - x - y)!} \frac{[\phi(\frac{1}{\rho} - 1)]^x}{[1 + \phi(\frac{1}{\rho} - 1)]^{n-y}}; \tag{6}$$
$$0 < \phi < 1; 0 < \rho < 1; x = 0, 1, 2, ..., n - y.$$

The conditional mean, $E[X = x | Y = y]$ and variance, $Var[X = x | Y = y]$ of the PMF (5) are:

$$E[X = x | Y = y] = \gamma_0 - \gamma_1 y,$$

And:

$$Var[X = x | Y = y] = (\gamma_0 - \gamma_1 x)[1 - \frac{(\gamma_0 - \gamma_1 y)}{(n - y)}]$$

The conditional mean is really the regression of x for a given y with the intercept is $\dfrac{n\phi(\frac{1}{\rho} - 1)}{[1 + \phi(\frac{1}{\rho} - 1)]}$ which increases with the number of participants, n in the survey and the slope is $\dfrac{\phi(\frac{1}{\rho} - 1)}{[1 + \phi(\frac{1}{\rho} - 1)]}$ which increases as the intercept increases.

Most importantly, of interest to those who conducts the survey is the number of "cheaters" in the survey. Recall that X and Y denote respectively the number of respondents who give "honest yes" and "honest no" answers by the n respondents. That means Z = n-X-Y denotes the number of cheaters in the survey. Then, the PMF of Z is Equation 7:

$$\Pr[Z = z] = \sum_{x}^{n} \sum_{y}^{n} \Pr[X = x, Y = y]$$
$$= \frac{n!}{z!(n-z)!} \rho^z (1-\rho)^{n-z}; \qquad (7)$$
$$0 < \rho < 1; z = 0, 1, 2, ..., n.$$

The expected number and variance of the "cheaters" are respectively $\mu_Z = n\rho$ and $\sigma_z^2 = \mu_z(1 - \frac{\mu_z}{n})$. In other words, the survey becomes homogeneous when the number of expected cheaters is higher or lower. The predictability of "honesty" is quite connected to the level of "cheating" in a survey. That is, the conditional PMF of X, the number of "honest yes" and Y, the number of "honest no" for a given level of "cheating" in a survey is Equation 8:

$$\Pr[X = x, Y = y | z]$$
$$= \frac{(n-z)!}{x!y!} \rho^{(n-z)} [\phi(\frac{1}{\rho} - 1)]^x [(1-\phi)(\frac{1}{\rho} - 1)]^y; \qquad (8)$$
$$0 < \phi < 1; 0 < \rho < 1; x, y = 0, 1, 2, ..., (n-z).$$

How great is the survey? A survey is great if the number of "honest yes" answers is more than the number

of "cheaters". For this purpose, the conditional PMF of the total, $T = X+Y$ number of "honest" answers in a survey for a given level of the number, $Z = z$ of "cheaters" is needed and it is Equation 9:

$$\Pr[X + Y = t | Z = z] = \sum_{i=0}^{t} \Pr[X = i, Y = t - i | z]$$
$$= \frac{(n-z)!}{t!(n-z-t)!} \rho^{n-z} (\frac{1}{\rho} - 1)^t; \qquad (9)$$
$$0 < \rho < 1; t = 0, 1, ..., (n-z).$$

The conditional mean of the PMF (9) is $\mu_{t/z} = E[X + Y = t | Z = z] = n(1-\rho) - (1-\rho)z$ with the conditional variance:

$$\sigma_{t/z}^2 = Var[X + Y = t | Z = z] = \rho \mu_{t/z}$$

And they mean the following. For a given z, the number of "cheaters" among the n respondents in a survey, the expected total number, t of the "honest" answers increases more heterogeneously when the probability, (1-ρ) for any respondent not yielding to "social pressure" increases. The probability, (1-ρ) for any respondent to have not yielded to social pressure to xenophobia is the slope of the downward regression line of $X + Y = t$ on $Z = z$.

How is its converse? The conditional PMF of Z, the number of "cheating" answers in a survey for a given level of the total number, $X + Y = 1$ of "honest" answers is Equation 10:

$$\Pr[Z = z | X + Y = t]$$
$$= \frac{(n-t)!}{z!(n-t-z)!} (1-\rho)^{n-t} (\frac{\rho}{1-\rho})^z; \qquad (10)$$
$$0 < \rho < 1; z = 0, 1, ..., (n-t).$$

The conditional mean of the PMF (9) is $\mu_{z/t} = E[Z = z | X + Y = t] = n\rho - \rho t$ with the conditional variance:

$$\sigma_{z/t}^2 = Var[Z = z | X + Y = t] = \rho(1-\rho)\mu_{z/t}$$

They mean the following. For a given total number X + Y = t, of "honest" answers by the n respondents in a survey, the expected number, z of the "cheating" answers increases more heterogeneously when the odds, ρ/(1-ρ) for any respondent to yield to "social pressure" increases. The conditionally predictable number of "cheaters" in a survey proportionally decreases at the rate of probability, ρ for any responded to yield to the "social pressure".

Next, we need to estimate the parameters: ϕ, ρ to be useful in practice. The Maximum Likelihood Estimators

(MLE) are preferable over others as the MLE are most efficient and optimal see (Kendall *et al.*, 1994). For this purpose of finding the MLE, consider a random sample (x_i, y_i, z_i), I = 1,2,…, nof size n≥2 as a draw from the PMF (9). Then, the log likelihood is Equation 11:

$$\ln L(\phi, \rho)$$
$$= \ln[\frac{n!}{x! \, y! (n-x-y)!}] + n \ln \rho \tag{11}$$
$$+ x \ln \phi + y \ln(1-\phi) + (x+y) \ln(\frac{1}{\rho} - 1)$$

Differentiating separately with respect to ϕ and ρ equating to zero and solving them, their Maximum Likelihood Estimator (MLE) are obtained. They are Equation 12:

$$\hat{\phi}_{MLE} = \frac{x}{(x+y)} \tag{12}$$

Which is the ratio of the number of "honest yes" to the total "honest answers" and Equation 13:

$$\hat{\rho}_{MLE} = 1 - \frac{(x+y)}{n} \tag{13}$$

Where n denotes the number of "respondents" in the survey. The surveyors often wonder whether a sample estimate is statistically significant. To answer the question, a hypothesis testing procedure needs to be developed. The likelihood ratio test is most powerful and invariant (that is, the MLE of a function is simply the function of the MLE). Hence, the likelihood ratio is adapted here. Kendall *et al.* (1994) for details about the likelihood ratio concept and tools.

To be specific, suppose that the surveyors wonders whether an estimate $\hat{\rho}$ (13) in a data is significant or negligible? If it is negligible, its p-value ought to be large. For the purpose of finding p-value, the covariance of the estimates (12) and (13) is zero. The variance-covariance matrix of the MLE (12) and (13) is the inverse of the Fisher's information matrix:

$$I = E \begin{bmatrix} -\partial^2_{\hat{\phi}\hat{\phi}} \ln L & -\partial^2_{\hat{\phi}\hat{\rho}} \ln L \\ -\partial^2_{\hat{\rho}\hat{\phi}} \ln L & -\partial^2_{\hat{\rho}\hat{\rho}} \ln L \end{bmatrix}$$
$$= \begin{bmatrix} \dfrac{\mu_x}{\phi^2} + \dfrac{\mu_y}{(1-\phi)^2} & 0 \\ 0 & \dfrac{n}{\rho^2}\{1 + \dfrac{(\mu_x + \mu_y)(1-2\rho)}{(1-\rho)^2}\} \end{bmatrix}$$

The information matrix is a diagonal matrix. The variance-covariance matrix $\begin{bmatrix} \text{var}(\hat{\phi}) & \text{cov}(\hat{\phi}, \hat{\rho}) \\ \text{cov}(\hat{\phi}, \hat{\rho}) & \text{var}(\hat{\rho}) \end{bmatrix}$ is the inverse matrix SI^{-1} 1. Because the information matrix is diagonal, the covariance matrix is also diagonal with inverted elements. That means we can use the statistic $\left| \dfrac{\hat{\rho}}{\sqrt{\text{vâr}(\hat{\rho})}} \right| = \left| \sqrt{n[1 + \dfrac{\hat{\rho}(1-2\hat{\rho})}{(1-\hat{\rho})^2}]} \right|$ which follows the standard normal distribution. Hence, the p-value of $\hat{\rho}$ is:

$$2\Pr\{Z > \left| \sqrt{n[1 + \frac{\hat{\rho}(1-2\hat{\rho})}{(1-\hat{\rho})^2}]} \right|\} \tag{14}$$

The statistical power accepting the alternative hypothesis $H_1 : \rho = \rho_1$ where $\rho_1 \neq \hat{\rho}$ is:

$$\Pr\{-\frac{(\hat{\rho} - \rho_1)\sqrt{n[1 + \frac{\hat{\rho}(1-2\hat{\rho})}{(1-\hat{\rho})^2}]}}{\hat{\rho}} < Z <$$
$$\frac{(\hat{\rho} - \rho_1)\sqrt{n[1 + \frac{\hat{\rho}(1-2\hat{\rho})}{(1-\hat{\rho})^2}]}}{\hat{\rho}}\} \tag{15}$$

3. ILLUSTRATION

In this section, the results in the previous section are illustrated using the survey data about xenophobia in (Ostapczuk *et al.*, 2009a; Musch *et al.*, 2001). There were two groups. In the group-1 with low education, a random sample of $n_1 = 259$ Germans and in the group-2 with high education, a random sample of $n_1 = 217$ Germans were asked whether they hated foreigners. The numbers for X, Y and Z are displayed in **Table 1**. Using the MLE (12) and (13) respectively, the estimate of insecurity level and the level of yielding to the social pressure are found for each group and displayed in the **Table 1** Other results are also displayed in **Table 1** and are interpreted below.

Notice that the estimated probability of being insecure is 0.61 in group-1 with low education and 0.36 in group-2 with high education. The estimated proportion yielding to social pressure is 0.37 in group-1 with low education while it is only 0.17 in the group-2 with high education. Indeed, the education reduces xenophobia.

Table 1. Survey results in Ostapczuk *et al*. (2009b)

Response type	Low education ($n_1 = 259$)	High education ($n_2 = 217$)
X, honest yes	99.00000	65.000000
Y, honest no	62.00000	114.000000
Z, cheaters	98.00000	38.000000
Proportion with insecurity, $\hat{\phi} =$	0.61000	0.360000
Proportion yielding to social pressure, $\hat{\rho} =$	0.37000	0.170000
P-value of $\hat{\rho}$	0.00001	0.000010
Power with $\rho_1 = 0.40$	0.94600	0.899000
Expected honest yes	99.53370	64.839600
Expected honest no	63.63630	115.270400
Correlation (x, y)	-0.71572	-0.837190
Beta 1	12.28500	1.130213
Gamma 1	0.63000	0.830000
Expected cheaters	95.83000	36.890000
Expected decrease of honesty for an extra cheater	0.63000	0.830000
Expected decrease of cheater for an extra honesty	0.37000	0.170000

The estimated probability of yielding to "social pressure" is significant in both groups. The statistical power of accepting the true statement that the proportion of yielding to social pressure is 0.40 is 0.946 in group-1 and 0.899 in group-2. Such high power confirms that the methodology of this article is superior.

The expected number of "honest yes" "honest no" and "cheaters" are close to the observed counter-parts. The correlation between the number of "honest yes" and the number of "honest no" is estimated to be -0.71 in group-1 with low education and -0.83 in group-2 with high education. The education has impact.

For an increase in "honest yes", the expected decrease in the "honest no" is 12.28 in group-1 with low education but is only 1.13 in group-2 with high education. Again, it confirms that the education makes a difference in the reduction of "honest no".

The converse is more robust. That is, for an increase in "honest no", the expected decrease in the "honest yes" is 0.63 in group-1 with low education but is only 0.83 in group-2 with high education. The education makes a moderate difference in the reduction of "honest yes". Furthermore, the expected decrease in the total number of "honest" persons is 0.63 in group-1 but 0.83 in group-2, when the number of "cheaters" increase by one. Finally, for an increase of one more total "honest" persons, the expected decrease in the number of cheaters is 0.37 in group-1 with low education but is only 0.17 in group-2 with high education and it confirms that there is an impact of education.

4. FUTURE RESEARCH DIRECTIONS

This article is first of its kind to suggest the underlying probability structure foe the "cheating" responses in a survey. It is customary for the non-measurable characteristics of the respondents of a survey are treated as parameters. Accordingly, the non-measurable reasons: "Insecurity level" and "yielding to social pressure" in a survey about xenophobia are treated as parameters. In Bayesian approach, the dynamics of the parameters are tracked and explained with the application of loss function and optimal criteria. The prior and posterior distributions of the parameters need to be worked out. The Bayesian estimates are done differently from the classical (that is, frequentist) approach. This article has explored only the frequentist approach but not the Bayesian approach of the "cheating" versus "honesty" in the answers of a survey. A future research work is needed and hence, is recommended to construct the Bayesian approach in this topic.

5. LIMITATIONS OF OUR METHODOLOGY

The bivariate probability model, estimators of the model's parameters and the validity of the hypothesis testing are limited to the availability of a random sample. In surveys, commonly practiced data collection methods are systematic sampling, cluster sampling, or snow sampling. The contents of this article are unsuitable for non-random sample which are collected using systematic sampling, cluster sampling, or snow sampling in a survey.

6. CONCLUSION

The education, as a covariate, is noticed to make a significant difference not only with respect to xenophobia but also the number of "honest yes", "honest no" and "cheating" in the survey. Then, there could be many other covariates. Such covariates might not be orthogonal to each other but might be collinear.

Currently, there is no methodology to sort them out. There is a need to develop such needed methodology. There are plenty of health and medical data out there to benefit from those yet to be developed methodologies. These methodologies have scope beyond the health in engineering, commerce and economics among many others.

7. ACKNOELEDGEMENT

8. REFERENCES

Ankudinov, A.B. and D.V. Lebedev, 2014. Job insecurity and employees motivation to engage in professional education. Am. J. Applied Sci., 11: 860-865. DOI: 10.3844/ajassp.2014.860.865

Campbell, A.A., 1987. Randomized response technique. Science, 236: 1049-1049.

Campbell, C. and B. Joiner, 1973. How to get the answer without being sure you've asked the question. Am. Stat., 27: 229-231. DOI: 10.1080/00031305.1973.10479043

Chaudhuri, A. and R. Mukerjee, 1989. Randomized Response: Theory and Techniques. Marcel Dekker, New York, ISBN: 0824777859.

Chiarella, C. and X.Z. He, 2005. Learning dynamics in the cobweb model with heterogeneous producers. Am. J. Applied Sci., 2: 45-56. DOI: 10.3844/ajassp.2005.45.56

Clark, S.J. and R.A. Desharnais, 1998. Honest answers to embarrassing questions: Detecting cheating in the randomized response model. Psychol., Meth., 3: 160-168. DOI:10.1037/1082-989X.3.2.160

Fan, S.K. and K. Chaloner, 2006. Design for a Trinomial Response to Dose. In: Response Surface Methodology and Related Topics, Khuri, A.I., (Ed.)., World Scientific Publications Singapore, ISBN: 9812774734, pp: 225-250.

Fox, J.P., 2005. Randomized item response theory models. J. Edu. Behav. Stat, 30: 189-212. DOI: 10.3102/10769986030002189

Goodstadt, M. and V. Gruson, 1975. The randomized response technique: A test on drug use. JASA, 70: 814-818. DOI: 10.1080/01621459.1975.10480307

Greenberg, B.G., A.L.A. Abul-Ela, W.R. Simmons and D.G. Horvitz, 1969. The unrelated question randomized response model: Theoretical framework. J. Am. Stat. Assoc., 64: 520-39. DOI: 10.1080/01621459.1969.10500991

Guerriero, M. and M.F. Sandri, 2007. A note on the comparison of some randomized response procedures. J. Stat. Plann. Inference, 173: 2184-2190. DOI: 10.1016/j.jspi.2006.07.004

Kendall, M.G., A. Stuart and J. Ord, 1994. Kendall's Advanced Theory of Statistics. 1st Edn., Edward Arnold, London, ISBN-10: 047023380X, pp: 676.

Lynn, P., 2008. The Problem of non-Response. In: International Handbook of Survey Methodology, Leeuw, E.D.D., J. Hox and D. Dillman, (Eds.)., Erlbaum, ISBN-10: 0805857532, pp: 35-55.

Maceli, J., 1978. How to Ask Sensitive Questions without Getting Punched in the Nose. In: Modules in Applied Mathematics, Lucas, W.F. (Ed.)., Springer-Verlag, New York, ISBN: 978-1-4612-5432-4, pp: 169-182.

Musch, J., A. Broder and K. Klauer, 2001. Improving Survey Research on the World-Wide Web Using the Randomized Response Technique. In: Dimensions of Internet Science, Reips, U.D. and M. Bosnjak (Eds.), Lengerich, Germany, Pabst, pp: 179-192.

Nasser, R., 2007. Deciding whether there is statistical independence or not? J. Math. Stat., 3: 151-162. DOI: 10.3844/jmssp.2007.151.162

Olatayo, T.O., 2011. Bootstrap approach to correlation analysis of two mineral components. Am. J. Geosci., 2: 10-12. DOI: 10.3844/ajgsp.2011.10.12

Ostapczuk, M., J. Musch and M. Moshagen, 2009a. A randomized-response investigation of the education effect in attitudes towards foreigners. Eur. J. Soc. Psychol., 39: 920-31. DOI: 10.1002/ejsp.588

Ostapczuk, M., M. Moshagen, Z. Zhao and J. Musch, 2009b. Assessing sensitive attributes using the randomized-response-technique: Evidence for the importance of response symmetry. J. Edu.. Behav. Stat., 34: 267-87. DOI: 10.3102/1076998609332747

Warner, S.L. 1965. Randomized response: A survey technique for eliminating evasive answer bias. J. Am. Stat., Assoc., 60: 63-69. PMID: 12261830

Yahya, W.B. and S.B. Adebayo, 2013. Multilevel ordinal response modeling of trend of breastfeeding initiation. Am. J. Biostat., 3: 1-10. DOI: 10.3844/amjbsp.2013.1.10