

SENSITIVITY ESTIMATION FOR PERSONAL INTERVIEW SURVEY QUESTIONS

S. Gupta, J. Shabbir

1. INTRODUCTION

Social desirability response bias (SDB) is a major problem in survey research involving sensitive questions (Edwards, 1957). Randomized response technique (RRT), pioneered by (Warner, 1965), is one of several methods to partially overcome SDB. Other methods involve use of bogus pipeline (BPL) (Jones and Sigall, 1971) and a SDB scale (Crowne and Marlowe, 1960). A comparison of BPL and RRT methods is shown by (Gupta and Thornton, 2003) using survey data. They show that a “Partial” RRT is at least as effective in circumventing SDB as BPL, while being more friendly and portable. Since SDB is directly correlated with the sensitivity level of the question, it is important to assess the sensitivity level since this allows one to assign better trained interviewers to collect information on more sensitive questions.

Researchers in this area have generally focused on the estimation of the mean of the sensitive variable. However, in a survey, different questions may have different sensitivity levels and it may be useful to quantify this sensitivity. (Gupta *et al.*, 2002) gave a method to estimate sensitivity levels using an optional RRT method. However, their method does not allow simultaneous estimation of the mean of the sensitive variable. In this paper we consider an optional RRT method that allows simultaneous estimation of both the average response and the sensitivity level of a sensitive question. The method follows the approach of (Greenberg *et al.*, 1969).

Two of the background models are discussed in section 2. The proposed method and its properties are discussed in section 3. In section 4 we present a numerical example to illustrate the estimation procedure and report some simulation results. We also discuss the performance of the proposed method in comparison with two other models – the “Full” RRT model of (Eichhorn and Hayre, 1983) and the “Partial” RRT model of (Mangat and Singh, 1990). Both of these models are considered in (Gupta and Thornton, 2003). We make some concluding remarks in section 5.

2. THE "FULL" AND "PARTIAL" RRT MODELS

In the "Full" RRT model of (Eichhorn and Hayre, 1983), each subject provides a scrambled response. This model works as follows. Let X be a sensitive quantitative variable of interest with an unknown mean of μ_x and an unknown variance of σ_x^2 . Let there be a deck of flash cards that follows a probability distribution S , independent of X , with a known mean of $\mu_s (= \theta)$ and a known variance of σ_s^2 . The respondent is asked to draw a card from the deck and is requested to report the scrambled response which is the product of the true response and the number on the card, and divided by the mean of the scrambling variable. Therefore, the reported response Y is given by

$$Y = \frac{XS}{\theta}. \quad (1)$$

The expected response, therefore, is given by $E(Y) = \mu_x$. This suggests estimating μ_x by $\hat{\mu}_x$, where $\hat{\mu}_x = \bar{Y}$. The variance of $\hat{\mu}_x$ is given by

$$\text{Var}(\hat{\mu}_x) = \text{Var}(\bar{Y}) = \frac{\text{Var}(Y)}{n} = \frac{\left[\sigma_x^2 + \frac{\sigma_s^2}{\theta^2} (\sigma_x^2 + \mu_x^2) \right]}{n}. \quad (2)$$

In the "Partial" RRT model, a predetermined proportion of randomly selected respondents are asked to provide a true response and the rest provide a scrambled response, just as in the "Full" RRT model. (Mangat and Singh, 1990) gave their "Partial" RRT model for the binary response (Yes/No) case, but it can be easily adapted for the quantitative response case also. If T is the proportion of respondents providing a true response, then the reported response is given by

$$Y = \begin{cases} X & \text{with probability } T \\ \frac{XS}{\theta} & \text{with probability } (1-T) \end{cases}$$

The expected response is given by

$$\begin{aligned} E(Y) &= \mu_x T + \frac{\mu_x \mu_s (1-T)}{\theta}, \\ &= \mu_x \text{ since } \mu_s = \theta. \end{aligned}$$

This suggests estimating μ_x by $\hat{\mu}_x = \bar{Y}$. Obviously $\hat{\mu}_x$ is an unbiased estimator of μ_x since \bar{Y} is an unbiased estimator of $E(Y)$. The variance of this estimator is given by

$$\text{Var}(\hat{\mu}_x) = \text{Var}(\bar{Y}) = \frac{\text{Var}(Y)}{n} = \frac{\left[\sigma_x^2 + (1-T) \frac{\sigma_s^2}{\theta^2} (\sigma_x^2 + \mu_x^2) \right]}{n}. \quad (3)$$

It can be checked easily that the variance in (3) is smaller than the variance of the estimator in (2).

3. PROPOSED METHOD

We propose here an ‘‘Optional’’ RRT model where a respondent is allowed to report a true response or a scrambled response depending on whether the respondent finds the question sensitive or not. Let X be the sensitive variable with mean μ_x and variance σ_x^2 . Let W ($0 \leq W \leq 1$) be the sensitivity level of the underlying sensitive question in the sense that a proportion W of the respondents in a survey consider the question sensitive enough not to feel comfortable answering the question in a face-to-face interview. Unlike the ‘‘Full’’ and the ‘‘Partial’’ RRT models, we would like to estimate both μ_x and W .

Analogous to (Greenberg *et al.*, 1969), we use two independent samples (with replacement) of respondents. Let the sample sizes be n_1 and n_2 . Respondents in i^{th} sample ($i = 1, 2$) are given a randomization device R_i ($i = 1, 2$), where R_i follows some probability distribution with a mean of θ_i and variance σ_i^2 . A respondent in sample i ($i = 1, 2$) is instructed to answer truthfully if he/she considers the question non-sensitive and to report a scrambled response using randomization device R_i if he/she considers the question sensitive. The interviewer will not know which way the question is answered. The reported response Z_i ($i = 1, 2$) from i^{th} sample will be

$$Z_i = \begin{cases} X & \text{with probability } (1-W) \\ S_i X & \text{with probability } W, \end{cases}$$

where X is the true response and S_i is scrambling variable value from device R_i . We assume X , S_1 and S_2 are mutually independent.

Note that for $i = 1, 2$, we have

$$\begin{aligned} E(Z_i) &= E(X)(1-W) + E(S_i X)W \\ &= \mu_x(1-W) + \theta_i \mu_x W \\ &= [\theta_i W + (1-W)] \mu_x. \end{aligned} \quad (4)$$

Eliminating W from the two equations in (4), we get

$$\mu_x = \frac{E(Z_1)(\theta_2 - 1) - E(Z_2)(\theta_1 - 1)}{(\theta_2 - \theta_1)}, \quad (5)$$

requiring that the two randomization devices be such that $\theta_1 \neq \theta_2$. This suggests estimating μ_x by

$$\hat{\mu}_x = \frac{\bar{Z}_1(\theta_2 - 1) - \bar{Z}_2(\theta_1 - 1)}{(\theta_2 - \theta_1)}, \quad (6)$$

where \bar{Z}_i ($i = 1, 2$) is the sample mean of reported responses from sample i .

Theorem 1. $\hat{\mu}_x$ is an unbiased estimator of μ_x

Proof: From (5), and from the fact that $E(\bar{Z}_i) = E(Z_i)$ ($i = 1, 2$), we get

$$E(\hat{\mu}_x) = E\left[\frac{\bar{Z}_1(\theta_2 - 1) - \bar{Z}_2(\theta_1 - 1)}{(\theta_2 - \theta_1)}\right] = \frac{E(Z_1)(\theta_2 - 1) - E(Z_2)(\theta_1 - 1)}{(\theta_2 - \theta_1)} = \mu_x. \quad (7)$$

Theorem 2. The minimum variance of $\hat{\mu}_x$ is

$$\text{Var}(\hat{\mu}_x)_{\min} = \frac{1}{n(\theta_2 - \theta_1)^2} [(\theta_2 - 1)\sigma_{\bar{z}_1} + (\theta_1 - 1)\sigma_{\bar{z}_2}]^2. \quad (8)$$

Proof: From (6), we have

$$\begin{aligned} \text{Var}(\hat{\mu}_x) &= \text{Var}\left[\frac{\bar{Z}_1(\theta_2 - 1) - \bar{Z}_2(\theta_1 - 1)}{(\theta_2 - \theta_1)}\right] \\ &= \frac{1}{(\theta_2 - \theta_1)^2} \left[(\theta_2 - 1)^2 \frac{\sigma_{\bar{z}_1}^2}{n_1} + (\theta_1 - 1)^2 \frac{\sigma_{\bar{z}_2}^2}{n_2} \right]. \end{aligned} \quad (9)$$

Differentiating (9) w. r. t n_1 and n_2 , we get

$$n_1 = \frac{n(\theta_2 - 1)\sigma_{\bar{z}_1}}{(\theta_2 - 1)\sigma_{\bar{z}_1} + (\theta_1 - 1)\sigma_{\bar{z}_2}}, \quad (10)$$

$$n_2 = \frac{n(\theta_1 - 1)\sigma_{\bar{z}_2}}{(\theta_2 - 1)\sigma_{\bar{z}_1} + (\theta_1 - 1)\sigma_{\bar{z}_2}}, \quad (11)$$

where σ_{Z_1} and σ_{Z_2} are standard deviations of Z_1 and Z_2 respectively and can be estimated through pilot study or past experience etc. The result follows on substituting (10) and (11) in (9). In order to estimate W , we eliminate μ_x from equation (4) and get

$$W = \frac{E(Z_2) - E(Z_1)}{E(Z_1)(\theta_2 - 1) - E(Z_2)(\theta_1 - 1)}. \quad (12)$$

This suggests estimating W by

$$\hat{W} = \frac{\bar{Z}_2 - \bar{Z}_1}{\bar{Z}_1(\theta_2 - 1) - \bar{Z}_2(\theta_1 - 1)}. \quad (13)$$

Theorem 3. The bias of \hat{W} is given by

$$\text{Bias}(\hat{W}) \approx \frac{1}{(\theta_2 - \theta_1)^2 \mu_x^2} \left[\frac{\sigma_{\bar{x}_1}^2}{n_1} (\theta_2 - 1) \{1 + W(\theta_2 - 1)\} + \frac{\sigma_{\bar{x}_2}^2}{n_2} (\theta_1 - 1) \{1 + W(\theta_1 - 1)\} \right].$$

Proof: Define

$$\hat{T}_1 = \bar{Z}_2 - \bar{Z}_1$$

and

$$\hat{T}_2 = \bar{Z}_1(\theta_2 - 1) - \bar{Z}_2(\theta_1 - 1),$$

then

$$\hat{W} = \hat{T}_1 / \hat{T}_2. \quad (14)$$

Now using second order Taylor's expansion around (T_1, T_2) (Mood *et al.*, 1974, p. 181), we have

$$E(\hat{W}) = E\left(\frac{\hat{T}_1}{\hat{T}_2}\right) \approx \frac{T_1}{T_2} - \frac{1}{T_2^2} \text{Cov}(\hat{T}_1, \hat{T}_2) + \frac{T_1}{T_2^3} \text{Var}(\hat{T}_2). \quad (15)$$

Note that

$$E(\hat{T}_1) = (\theta_2 - \theta_1)W\mu_x = T_1 \text{ (say)}, \quad (16)$$

$$E(\hat{T}_2) = (\theta_2 - \theta_1)\mu_x = T_2 \text{ (say)}, \quad (17)$$

$$\text{Var}(\hat{T}_1) = \frac{\sigma_{\tilde{x}_1}^2}{n_1} + \frac{\sigma_{\tilde{x}_2}^2}{n_2}, \quad (18)$$

$$\text{Var}(\hat{T}_2) = (\theta_2 - 1)^2 \frac{\sigma_{\tilde{x}_1}^2}{n_1} + (\theta_1 - 1)^2 \frac{\sigma_{\tilde{x}_2}^2}{n_2}, \quad (19)$$

$$\begin{aligned} \text{Cov}(\hat{T}_1, \hat{T}_2) &= \text{Cov}[(\bar{Z}_2 - \bar{Z}_1), \{\bar{Z}_1(\theta_2 - 1) - \bar{Z}_2(\theta_1 - 1)\}] \\ &= -(\theta_1 - 1) \frac{\sigma_{\tilde{x}_2}^2}{n_1} - (\theta_2 - 1) \frac{\sigma_{\tilde{x}_1}^2}{n_2}, \end{aligned} \quad (20)$$

where

$$\sigma_{\tilde{x}_1}^2 = (\sigma_x^2 + \mu_x^2) \{1 - W + W(\theta_1^2 + \sigma_1^2)\} - \{\theta_1 W + (1 - W)\}^2 \mu_x^2$$

and

$$\sigma_{\tilde{x}_2}^2 = (\sigma_x^2 + \mu_x^2) \{1 - W + W(\theta_2^2 + \sigma_2^2)\} - \{\theta_2 W + (1 - W)\}^2 \mu_x^2.$$

The result follows on substituting (19) and (20) in (15).

Theorem 4. The minimum variance of \hat{W} is given by

$$\text{Var}(\hat{W})_{\min} = \frac{1}{n(\theta_2 - \theta_1)^2 \mu_x^2} [\sigma_{\tilde{x}_1} \{1 + W(\theta_2 - 1)\} + \sigma_{\tilde{x}_2} \{1 + W(\theta_1 - 1)\}]^2. \quad (21)$$

Proof: From (Mood *et al.*, (1974, p. 181), we have

$$\text{Var}(\hat{W}) = \text{Var}\left(\frac{\hat{T}_1}{\hat{T}_2}\right) \approx \left(\frac{T_1}{T_2}\right)^2 \left[\frac{\text{Var}(\hat{T}_1)}{T_1^2} + \frac{\text{Var}(\hat{T}_2)}{T_2^2} - \frac{2\text{Cov}(\hat{T}_1, \hat{T}_2)}{T_1 T_2} \right]. \quad (22)$$

Substituting (18) through (20) in (22), we get

$$\text{Var}(\hat{W}) \approx \frac{1}{(\theta_2 - \theta_1)^2 \mu_x^2} \left[\frac{\sigma_{\tilde{x}_1}^2}{n_1} \{1 + W(\theta_2 - 1)\}^2 + \frac{\sigma_{\tilde{x}_2}^2}{n_2} \{1 + W(\theta_1 - 1)\}^2 \right]. \quad (23)$$

Differentiating (23) w. r. t n_1 and n_2 , we get

$$n_1 = \frac{n\sigma_{z_1} \{1 + W(\theta_2 - 1)\}}{\sigma_{z_1} \{1 + W(\theta_2 - 1)\} + \sigma_{z_2} \{1 + W(\theta_1 - 1)\}}, \quad (24)$$

$$n_2 = \frac{n\sigma_{z_2} \{1 + W(\theta_1 - 1)\}}{\sigma_{z_1} \{1 + W(\theta_2 - 1)\} + \sigma_{z_2} \{1 + W(\theta_1 - 1)\}}. \quad (25)$$

Substituting (24) and (25) in (23), we get minimum variance of \hat{W} as given in (21). We evaluate the performance of the proposed estimator in the next section.

4. A NUMERICAL EXAMPLE AND SIMULATIONS

In this section, we first explain the estimation process with the help of a numerical example, and then present a more extensive simulation study. The first step in the estimation process is to split the total sample size optimally in two sub-samples so as to minimize $\text{Var}(\hat{\mu}_x) + \text{Var}(\hat{W})$ instead of minimizing just $\text{Var}(\hat{\mu}_x)$ or just $\text{Var}(\hat{W})$. This is because we are estimating both the sensitive variable mean μ_x and its sensitivity level W . The optimum values of (n_1, n_2) for minimizing $\text{Var}(\hat{\mu}_x) + \text{Var}(\hat{W})$ can easily be calculated just as in equations (10), (11), (24) and (25). These optimum values are given by

$$n_1 = \frac{n\sqrt{[\mu_x^2(\theta_2 - 1)^2 + \{1 + W(\theta_2 - 1)\}^2]\sigma_{z_1}}}{\sqrt{[\mu_x^2(\theta_1 - 1)^2 + \{1 + W(\theta_1 - 1)\}^2]\sigma_{z_2}} + \sqrt{[\mu_x^2(\theta_2 - 1)^2 + \{1 + W(\theta_2 - 1)\}^2]\sigma_{z_1}}}, \quad (26)$$

$$n_2 = \frac{n\sqrt{[\mu_x^2(\theta_1 - 1)^2 + \{1 + W(\theta_1 - 1)\}^2]\sigma_{z_2}}}{\sqrt{[\mu_x^2(\theta_1 - 1)^2 + \{1 + W(\theta_1 - 1)\}^2]\sigma_{z_2}} + \sqrt{[\mu_x^2(\theta_2 - 1)^2 + \{1 + W(\theta_2 - 1)\}^2]\sigma_{z_1}}}. \quad (27)$$

The unknown parameters (μ_x, W) involved in these equations can be estimated from a pilot study.

For the numerical example, we assume that the variable of interest has a Poisson distribution with a mean of 4.0 and a sensitivity level of .30. We plan to use a total sample size of 500. For the pilot study, we take two independent samples of size 100 each according to the ‘‘Optional’’ RRT model proposed here. The scrambling variable for the first sample is taken to be a Poisson with a mean of $\theta_1 = 2$. The scrambling variable for the second sample is taken to be a Poisson with a mean of $\theta_2 = 5$. In keeping with the chosen sensitivity level of .30, ran-

domly selected 30% observations in each sample are scrambled and the rest are left as it is. The sample means and variances $(\bar{Z}_1, \bar{Z}_2, \hat{\sigma}_{Z_1}^2, \hat{\sigma}_{Z_2}^2)$ for these two samples are 5.50, 9.4520, 17.9776 and 165.6369 respectively. With these sample statistics, estimated values for μ_x and W (from (6) and (13)) are 3.8833 and .315 respectively. Using these preliminary estimates and $n = 500$ in (26) and (27), optimum values for the two sample sizes are $n_1 = 279$ and $n_2 = 221$.

We now select two independent samples of size 279 and 221 respectively from a Poisson distribution with a mean of 4.0. In both samples, 30% of the observations are scrambled using the scrambling variables chosen in the pilot study. The sample statistics $(\bar{Z}_1, \bar{Z}_2, \hat{\sigma}_{Z_1}^2, \hat{\sigma}_{Z_2}^2)$ for these two samples are 5.032, 8.167, 20.25 and 92.2176 respectively. Using these in (6) and (13), we get $\hat{\mu}_x = 3.987$ and $\hat{W} = .262$ as our final estimates. The estimated variances of these estimates, calculated from (9) and (23), are given by $\hat{V}ar(\hat{\mu}_x) = .1754$ and $\hat{V}ar(\hat{W}) = .0068$.

For the more extensive simulation study we assume that the sensitive variable has a Poisson distribution with $\mu_x = 4$. The scrambling variables are also assumed to have Poisson distributions. All simulation results are averaged over 1000 replications. Table 1 below gives estimated values for various parameters. Note that estimation of μ_x , W and $\text{Var}(\hat{\mu}_x)$ is very good even for small sample size but estimation of $\text{Var}(\hat{W})$ can be off for small samples when sensitivity level is high. This is because we have used unconstrained estimation of W in the sense that the estimated value is allowed to go outside the usual range $(0, 1)$. Some of the simulation runs do give such values when sample size is small and sensitivity level is high. Corresponding estimates are denoted as “**”. Also, $\text{Var}(\hat{W})$ in (23) is accurate only up to second order approximation. Again, this approximation may not be very good for small samples. Both of these problems disappear when sample size is large.

Results in Table 2 below compare variances of various estimators for various choices of W and optimum values of sample sizes (n_1, n_2) . Note that the variances for “Optional” RRT method introduced here are higher as compared to the “Full” and the “Partial” RRT methods referenced in Section 1. However, this is a bit misleading because we estimate both μ_x and W from the same sample unlike the “Partial” RRT method where W is assumed to be known, and unlike the “Full” RRT method where all subjects provide a scrambled response.

TABLE 1

Simulation results for $\mu_x = 4$, $\theta_1 = 2$, $\theta_2 = 5$ and various choices of W . Sample sizes (n_1, n_2) are chosen optimally according to (26) and (27)

n	W	\hat{W}	$\hat{\mu}_x$	$\hat{V}(\hat{\mu}_x)$	$V(\hat{\mu}_x)$	$\hat{V}(\hat{W})$	$V(\hat{W})$
100	0.1	0.118	3.997	0.446	0.439	0.021	0.014
	0.2	0.241	3.986	0.741	0.731	0.061	0.029
	0.3	0.378	3.979	1.022	0.998	0.157	0.050
	0.4	**	3.970	1.243	1.240	**	0.076
	0.5	**	3.977	1.428	1.457	**	0.106
	0.6	**	3.983	1.615	1.649	**	0.141
	0.7	**	3.971	1.841	1.816	**	0.180
	0.8	**	3.966	2.028	1.957	**	0.223
	0.9	**	3.960	2.042	2.071	**	0.269
	1.0	**	3.966	2.082	2.158	**	0.316
500	0.1	0.102	4.004	0.089	0.088	0.003	0.003
	0.2	0.204	4.005	0.144	0.146	0.007	0.006
	0.3	0.307	4.006	0.210	0.200	0.011	0.010
	0.4	0.412	4.010	0.272	0.248	0.018	0.015
	0.5	0.512	4.015	0.283	0.291	0.023	0.021
	0.6	0.615	4.024	0.329	0.330	0.032	0.028
	0.7	0.721	4.023	0.377	0.363	0.042	0.036
	0.8	0.827	4.017	0.394	0.391	0.054	0.045
	0.9	0.931	4.021	0.442	0.414	0.073	0.054
	1.0	1.036	4.015	0.448	0.432	0.082	0.063
1000	0.1	0.102	4.000	0.043	0.044	0.001	0.001
	0.2	0.203	4.000	0.078	0.073	0.003	0.003
	0.3	0.303	4.006	0.102	0.100	0.005	0.005
	0.4	0.408	4.005	0.135	0.124	0.009	0.008
	0.5	0.509	4.006	0.156	0.146	0.012	0.011
	0.6	0.609	4.011	0.158	0.165	0.014	0.014
	0.7	0.712	4.011	0.186	0.182	0.019	0.018
	0.8	0.815	4.008	0.203	0.196	0.025	0.022
	0.9	0.917	4.003	0.210	0.207	0.029	0.027
	1.0	1.018	3.998	0.201	0.216	0.031	0.032

** : Unreliable estimates since many simulation runs produced a \hat{W} outside the range (0, 1)

TABLE 2

Values of $\text{Var}(\hat{\mu}_x)$ for “Full”, “Partial” and “Optional” RRT models for $\mu_x = 4$, $\theta_1 = 2$, $\theta_2 = 5$ and various choices of W . Sample sizes (n_1, n_2) are chosen optimally according to (26) and (27)

n	“Full” RRT	W	“Partial” RRT	“Optional” RRT
100	0.140	0.1	0.050	0.439
		0.2	0.060	0.731
		0.3	0.070	0.998
		0.4	0.080	1.240
		0.5	0.090	1.457
		0.6	0.100	1.649
		0.7	0.110	1.816
		0.8	0.120	1.957
		0.9	0.130	2.071
		1.0	0.140	2.158
500	0.028	0.1	0.010	0.088
		0.2	0.012	0.146
		0.3	0.014	0.200
		0.4	0.016	0.248
		0.5	0.018	0.291
		0.6	0.020	0.330
		0.7	0.022	0.363
		0.8	0.024	0.391
		0.9	0.026	0.414
		1.0	0.028	0.432
1000	0.014	0.1	0.005	0.044
		0.2	0.006	0.073
		0.3	0.007	0.100
		0.4	0.008	0.124
		0.5	0.009	0.146
		0.6	0.010	0.165
		0.7	0.011	0.182
		0.8	0.012	0.192
		0.9	0.013	0.207
		1.0	0.014	0.216

5. CONCLUDING REMARKS

The proposed method has the advantage of being able to simultaneously estimate both the average response and the sensitivity level of sensitive survey questions. While comparing this method with other methods, one should keep in mind that the proposed estimator estimates two parameters (μ_x, W) unlike the “Full” and the “Partial” RRT methods that estimate only μ_x , and hence have smaller variance. Also, due to unconstrained estimation of W , some of the estimated values may fall outside the range $(0, 1)$, but this problem disappears for larger samples. Estimation of the sensitivity level allows one to assign interviewers with specific skills to conduct surveys involving more sensitive questions.

*Department of Mathematical Sciences,
University of North Carolina at Greensboro*

SAT GUPTA

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

JAVID SHABBIR

ACKNOWLEDGEMENTS

The authors are very thankful to the referees for helpful suggestions that helped improve considerably the presentation of the paper. The second author wishes to acknowledge with thanks the facilities made available by the University of Southern Maine, USA during his post-doctoral work in 2003.

REFERENCES

- D. P. CROWNE, D. MARLOWE (1960), *A new scale of social desirability independent of psychopathy*, “Journal of consulting psychology”, 24, pp. 349-354.
- A. L. EDWARDS (1957), *The social desirability variable in personality assessment and research*, New York: Dryden.
- B. H. EICHHORN, L. S. HAYRE (1983), *Scrambled randomized response methods for obtaining sensitive quantitative data*, “Journal of Statistical Planning and Inference”, 7, pp. 307-316.
- S. GUPTA, B. GUPTA, S. SINGH (2002), *Estimation of sensitivity level of personal interview survey question*, “Journal of Statistical Planning and Inference”, 100, pp. 239-247.
- S. GUPTA, B. THORNTON (2003), *Circumventing social desirability response bias in personal interview surveys*, “American Journal of Mathematical and Management Sciences”, 22, pp. 369-383.
- B. G. GREENBERG, A. L. A. ABUL-ELA, W. R. SIMMONS, D. G. HORVITZ, (1969), *The unrelated question randomized response model- theoretical framework*, “Journal of American Statistical Association”, 64, pp. 520-539.
- E. E. JONES, H. SIGALL (1971), *The bogus pipe line: a new paradigm for measuring affect and attitude*, “Psychological Bulletin”, 76, pp. 349-364.
- N. S. MANGAT, R. SINGH (1990), *An alternative randomized response procedure*, “Biometrika”, 77, pp. 439-442.

- A. M. MOOD, F. A. GRAYBILL, D. C. BOES (1974), *Introduction to theory of statistical mathematics*. McGraw-Hill.
- S. L. WARNER (1965), *Randomized response: a survey technique for eliminating evasive answer bias*, "Journal of American Statistical Association", 60, pp. 63-69.

RIASSUNTO

Stima sensitiva delle domande personali in un'intervista

In questo lavoro viene proposto un metodo RRT per stimare il livello di sensitività in domande personali a risposta quantitativa. Il metodo è simile a quello usato da Gupta *et al.* (2002) ma permette di stimare sia il livello medio della risposta sia il livello di sensitività della domanda. Il procedimento di stima viene valutato sia sulla base di una applicazione numerica sia attraverso uno studio di simulazione. Viene inoltre confrontata la performance dello stimatore con quella dei modelli a risposta randomizzata "completa" e "parziale".

SUMMARY

Sensitivity estimation for personal interview survey questions

This paper proposes an optional RRT method to estimate the sensitivity level of personal interview survey questions with quantitative response. The method is similar to the one used by Gupta *et al.* (2002) but estimates both the average response level and the sensitivity level of the question. A numerical example explains the estimation process and a simulation study assesses the effectiveness of the proposed method. We also compare the performance of the proposed estimator with the "Full" and "Partial" randomized response models.