5-1-2010

Ranked Set Sampling Using Auxiliary Variables of a Randomized Response Procedure for Estimating the Mean of a Sensitive Quantitative Character

Carlos N. Bouza

Universidad de la Habana, Havana, Cuba, bouza@matcom.uh.cu

Recommended Citation

Ranked Set Sampling Using Auxiliary Variables of a Randomized Response Procedure for Estimating the Mean of a Sensitive Quantitative Character

Carlos N. Bouza
Universidad de la Habana
Havana, Cuba

The analysis of the behavior of estimators of the mean of a sensitive variable is considered when a randomized response procedure is used. The results deal with the inference based on simple random sampling with replacement study design. A study of the behavior of the procedures for a ranked set sampling design is developed. A gain in accuracy is generally associated with the proposed alternative model.

Key words: Order statistics, expected model variance, unbiasedness.

Introduction

Let Y be a sensitive variable evaluated in a finite population $U = \{u_1, \ldots, u_N\}$. The individual u_i with a value of Y that carries a stigma will tend to give incorrect information or to refuse to answer. It is well known that, when dealing with sensitive questions, researchers face the need to reduce response refusals as well as response bias. A possibility is to replace a direct response to a sensitive question by using a random response query.

The seminal work on this issue was conducted by Warner (1965), who dealt with a qualitative question with possible responses of yes or no when one question is identified to have the stigma. The goal of the surveyor is to estimate the probability of having the stigma. It is expected that a large percent of the persons bearing it will either lie or refuse to answer. Warner's method consists of placing the question associated with the stigma together with some insensitive questions. The respondent randomly chooses a question and answers it without revealing which was selected. When dealing with a quantitative character, similar reasoning can be used. Chaudhuri-Stenger (1992) proposed the use of a randomized responses (RR) model and they developed an approach for designing a general RR procedure when the stigma is related with a quantitative character. See, for example, Singh-Singh (1993) and Zou (1997), who analyzed the behavior of various estimators of the mean when simple random sampling is used for selecting the sample.

Ranked set sampling (RSS) is an alternative sample design that provides gains in accuracy with respect to simple random sampling with replacement (SRSWR). It was proposed by McIntyre (1952) to estimate mean pasture yield, and it was found to be more efficient than selecting the sample using a simple random sampling (SRS) design.

The units can be ranked by means of a cheap procedure and selecting an order statistic from each of the independent samples selected using SRS with replacement (SRSWR). Results indicated that the use of ranked set sampling is highly beneficial and leads to estimators that are more precise than the usual sample mean per unit. The method is now referred to as the ranked set sampling (RSS) method. (See Patil (2002) and Patil, et al. (1994, 1999) for detailed discussions.) This research developed a study of the use of alternative RR procedures when RSS was used instead of SRS with replacement (SRSWR) under a model.

Carlos N. Bouza is a Professor in the Faculty of Mathematics and Computation. Email him at: bouza@matcom.uh.cu.

A Model Based Randomized Responses Procedure Under SRSSWR

The RR procedure proposed in Chaudhuri-Stenger (1992) begins with an individual $u_i \in U$ and the sets of known variables $A = \{A_1, ..., A_T\}$ and $B = \{B_1, ..., B_S\}$. After these are fixed the following are calculated:

$$\mu_A = \sum_{t=1}^{T} A_t / T \neq 0,$$

$$\sigma_A^2 = \sum_{t=1}^{T} (A_t - \mu_A)^2 / T,$$

$$\mu_B = \sum_{s=1}^{S} B_s / S,$$

and

$$\sigma_B^2 = \sum_{s=1}^{S} (B_s - \mu_B)^2 / S.$$

After u_i is selected from the population $U = \{1, ..., N\}$ the respondent will not report the value of Y_i directly. Instead a random experiment is performed and its result is the independent selection of $a \in A$ and a $b \in B$, (A_i, B_i). The report made by the interviewee is: $Z_i = A_i Y_i + B_i$.

The first procedure for deriving information on Y_i is to use the report to compute

$$R_i = \frac{Z_i - \mu_B}{\mu_A}.$$

The model expectation is

$$E_R(R_i) = Y_i$$

and the corresponding variance is

$$V_R(R_i) = \frac{Y_i^2 \sigma_A^2 + \sigma_B^2}{\mu_A^2} = V_i.$$

The selection of a sample of size n using a sampling design generates the reports $R_1, ..., R_n$. The sample means of the computed calculated variables are used for estimating the mean of the sensitive variable.

The RR procedure generates the data $D(R) = \{(u_i, Y_i, A_i, B_i) | u_i \in s, A_i \in A, B_i \in B\}$. The selection made in A and B produces random variables that are not related with the stigmatized character. The estimator of the mean of the variable R is

$$\bar{R} = \frac{\sum_{I=1}^{n} R_i}{n} \qquad (2.1)$$

Because R_i is model unbiased for Y_i, the model expectation of (2.1) is the sample mean of Y. It is a design unbiased estimator of the population mean. Therefore,

$$E_d E_R(\bar{R}) = E_d(\bar{y}) = \mu_Y.$$

The independence of the selection, provided by the use of SRSWR, supports that the model variance of (2.1) is:

$$V_R(\bar{R}) = \frac{\sum_{i=1}^{n} V_i^2}{n^2}$$

and the design-expectation of the variance is

$$E_d V_R(\bar{R}) = \frac{\sum_{i=1}^{n} E_d \left[\dfrac{Y_i^2 \sigma_A^2 + \sigma_B^2}{\mu_A^2} \right]}{n^2} \qquad (2.2)$$

$$= \left[\frac{\mu_Y^2 + \sigma_Y^2}{n \mu_A^2} \right] \sigma_A^2 + \frac{\sigma_B^2}{n \mu_A^2}$$

Ranked Set Sampling for the Randomized Responses Procedures

To implement RSS, m independent samples of size n are selected using SRSWR. The units in each sample are ranked without knowing the value of Y. Either personal judgment or the evaluation of a covariate X that is correlated with Y may be used to rank Y and the individual with rank i in the ranked sample $s_{(i)}$ is interviewed. The procedure is repeated independently r times (cycles) and n = mr. David-Levine (1972) studied the effect of judgmental errors and established that it does not affect the properties of RSS.

Suppose that the ranking is made on Y. The sampler asks interviewee u_i to randomly select A_i and B_i. The report of the ith ranked sample in the t^h cycle is: $Z_{(i)t} = A_i Y_{(i)t} + B_i$. Using this report, the following can be computed for each u_i

$$R_{(i)t} = \frac{Z_{(i)t} - \mu_B}{\mu_A}$$

and, if its expectation is $E_R(R_{(i)t}) = Y_{(i)t}$, then

$$\bar{R}_t = \frac{\sum_{i=1}^{m} R_{\{i\}t}}{m},$$

and its unbiasedness may be derived using:

$$E_d E_R(\bar{R}_t) = \frac{\sum_{i=1}^{m} \mu_{Y(i)}}{m} = \mu_Y.$$

From these results an estimator was derived that uses the information provided in the r cycles and is unbiased. Here it is proposed that

$$\bar{R}_{(rss)} = \frac{\sum_{t=1}^{r}\sum_{i=1}^{m} R_{(i)t}}{rm} \qquad (3.1)$$

be used. The fact that

$$V_R(R_{(i)t}) = V_R\left(\frac{Z_{(i)t} - \mu_B}{\mu_A}\right)$$

$$= \frac{Y_{(i)t}^2 \sigma_A^2 + \sigma_B^2}{\mu_A^2}$$

along with the independence of the involved variables are used for deriving the expected error of the RSS-estimator, which is:

$$E_d V_R(\bar{R}_{(rss)}) = \frac{\sum_{i=1}^{m}\left(\sigma_{(i)}^2 + \mu_{Y8(i)}^2\right)\sigma_A^2 + \sigma_B^2}{rm^2\mu_A^2}$$

$$(3.2)$$

given that $E_d(Y_{(i)t}^2) = \sigma_{(i)}^2 + \mu_{Y(i)}^2$.

The relation between $\sigma^2_{Y(i)}$, the variance of the ith os, and the variance of the distribution σ^2_Y permits (see Dell-Clutter, 1972): $\sum_{i=1}^{m} \sigma^2_{Y(i)}$

$= m\sigma^2_Y - \sum_{i=1}^{m}(\mu_{Y(i)} - \mu_Y)^2 = m\sigma^2_Y - \sum_{i=1}^{m} \Delta^2_{Y(i)}$.

Thus, (3.1) can be rewritten as:

$$E_d V_R(\bar{R}_{(rss)}) =$$

$$\frac{\sigma_A^2}{\mu_A^2}\left[\sigma_Y^2 - \frac{\sum_{i=1}^{m}\Delta^2_{Y(i)}}{mn}\right] + \frac{\sigma_A^2}{n\mu_A^2}\left[\frac{\sum_{i=1}^{m}\mu^2_{Y(i)}}{m}\right] + \frac{\sigma_B^2}{n\mu_A^2}$$

and leads to the proposition 3.1.

Proposition 3.1
Consider the use of RRS for selecting a sample of n = rm individuals and the ranking with respect to Y of the reports $Z_{(i)t}$, i = 1, m, and t = 1, …, r. $G(0, 1)$ represents a gain in accuracy measured by

$$G(0,1) = \frac{\sigma_A^2}{n\mu_A^2}\left[\sum_{i=1}^{m}\frac{\mu_Y^2 - \mu^2_{Y(i)}}{m} + \sum_{i=1}^{m}\frac{\Delta^2_{Y(i)}}{m}\right],$$

$\Delta_i^2 = (\mu_{Y(i)} - \mu_Y)^2$, for estimating the mean using (3.1) instead of the SRSWR and the sample mean. This result is a natural extension of the classic RSS procedure (note that RSS is not necessarily more accurate than SRSWR).

The ranking may be implemented using the information provided by the selection of the auxiliary variables. The persons included in each sample randomly select the corresponding insensitive variables A and B; they communicate their values to one another for establishing their ranks. The person in position j in the sample j gives the report. The procedure and the m independent samples in each cycle are evaluated. The report of an individual u_i is:

$$Z(i) = \begin{cases} Z_{A(i)} = A_{(i)}Y_i + B_i \\ \qquad if\ A\ det\,er\,min\,es\ the\ order \\ Z_{B(i)} = A_i Y_i + B_{(i)} \\ \qquad if\ B\ det\,er\,min\,es\ the\ order \end{cases}$$

If the ranks are made on A, then the tth cycle report is $Z_{A(i)t} = A_{(i)t}Y_i + B_i$, and the response variable computed is

$$R_{A(i)t} = \frac{Z_{A(i)t} - \mu_B}{\mu_{A(i)}}.$$

The model expectation is the value of the sensitive variable, $E_R\left(R_{A(i)t}\right) = Y_i$, therefore, to average the reports generates an unbiased estimation of the mean of Y. The corresponding results are fixed in the Proposition 3.2.

Proposition 3.2

The use of RRS for selecting a sample of n = rm individuals and the ranking, with respect to A, of the reports $Z_{(i)t}$, i = 1., …, m and t = 1, …,.r, represents a gain in accuracy when

$$\sum_{i=1}^{m} \frac{1}{m\mu_{A(i)}^2\mu_A^2} + \sum_{i=1}^{m} \frac{\Delta_{A(i)}^2}{nm\mu_{A(i)}^2\mu_A^2} > \sum_{i=1}^{m} \frac{\mu_{A(i)}^2}{m\mu_{A(i)}^2\mu_A^2}$$

with the estimator given by:

$$\bar{R}_{A(rss)} = \frac{\sum_{t=1}^{r}\sum_{i=1}^{m} R_{A(i)t}}{rm} \tag{3.3}$$

which is unbiased. The proof for this is as follows.

The unbiasedness of (3.3) follows from the fact that the reports are model unbiased for the corresponding Y_i and the arithmetic mean is design unbiased. Its model variance for the ith os in the cycle t is:

$$V_R\left(R_{A(i)t}\right) = \frac{Y_i^2\sigma_{A(i)}^2 + \sigma_B^2}{\mu_{A(i)}^2}$$

where $\sigma_{A(i)}^2$ and $\mu_{A(i)}$ are the variance and mean of $A_{(i)}$. The design expectation of the model error for the ith os is:

$$E_dV_R\left(R_{A(i)t}\right) = \frac{\left(\sigma_Y^2 + \mu_Y^2\right)\sigma_{A(i)}^2 + \sigma_B^2}{\mu_{A(i)}^2}$$

and

$$E_dV_R\left(\bar{R}_{A(rss)}\right) =$$

$$\left(\sigma_Y^2 + \mu_Y^2\right)\sum_{i=1}^{m} \frac{\sigma_{A(i)}^2}{nm\mu_{A(i)}^2} + \sum_{i=1}^{m} \frac{\sigma_B^2}{nm\mu_{A(i)}^2}$$

is the expected error of the estimator.

The relation between the variance of an os and the population variance is used again for rewriting the design expectation of the model variance of the estimator. The result derived is:

$$E_dV_R\left(\bar{R}_{A(rss)}\right) =$$

$$\left(\sigma_Y^2 + \mu_Y^2\right)\left[\frac{\sigma_A^2}{n}\sum_{i=1}^{m}\frac{1}{m\mu_{A(i)}^2} - \sum_{i=1}^{m}\frac{\Delta_{A(i)}^2}{nm\mu_{A(i)}^2}\right]$$

$$+ \sum_{i=1}^{m} \frac{\sigma_B^2}{nm\mu_{A(i)}^2}$$

$$\tag{3.4}$$

where $\Delta_{A(i)} = \mu_{A(i)} - \mu_A$.

Comparing (3.4) with (2.2) the gain is

$$G(0,2) =$$

$$\frac{\left(\sigma_Y^2 + \mu_Y^2\right)\sigma_A^2 + \sigma_B^2}{n}\left[\sum_{i=1}^{m}\frac{\mu_A^2 - \mu_{A(i)}^2}{m\mu_{A(i)}^2\mu_A^2} + \sum_{i=1}^{m}\frac{\Delta_{A(i)}^2}{nm\mu_{A(i)}^2\mu_A^2}\right]$$

which is positive only if the relation stated holds. Note that the conditions stated for granting use of the strategy characterized in the propositions are better than the SRSWR strategy when

$$\sum_{i\neq i'} \frac{\mu_{C(i)}\mu_{Ci'}}{m\mu_{C(i)}^2\mu_C^2} > 0, \ \ C = Y, A$$

is satisfied. The designer of the inquiry is able to fix the possible values of A and calculate the expectation of the different order statistics. Thus, it is possible to have a previous evaluation of the gain in accuracy when the model is based on ranking the auxiliary variable A.

When the ranking is based on the results for B the report is $Z_{B(i)t} = A_{it}Y_i + B_{(i)t}$ and the following is computed

$$R_{B(i)t} = \frac{Z_{B(i)t} - \mu_{B(i)}}{\mu_A}.$$

Once again the model expectation provides the true value of Y, that is $E_R(R_{B(i)t}) = Y_i$ and an unbiased estimation of μ_Y is

$$\overline{R}_{B(rss)} = \frac{\sum_{t=1}^{r}\sum_{i=1}^{m} R_{B(i)t}}{rm} \qquad (3.5)$$

because $E(Z_{B(i)}) = \mu_A Y_i + \mu_{B(i)}$. Denoting by $\sigma^2_{B(i)}$ and $\mu_{B(i)}$ the variance and mean of the ith os of B gives the design expectation of the error of the proposed estimator as:

$$E_d V_R\left(\overline{R}_{B(rss)}\right)$$

$$= \left(\sigma_Y^2 + \mu_Y^2\right)\left(\frac{\sigma_A^2}{n\mu_A^2}\right) + \sum_{i=1}^{m}\frac{\sigma_{B(i)}^2}{nm\mu_A^2}$$

$$= \left(\sigma_Y^2 + \mu_Y^2\right)\left(\frac{\sigma_A^2}{n\mu_A^2}\right) + \frac{\sigma_B^2}{n\mu_A^2} - \sum_{i=1}^{m}\frac{\Delta_{B(i)}^2}{nm\mu_A^2}$$

Therefore, to rank in B provides a gain in accuracy that does not depend on the differences between the mean of the variable used for ranking and expectation of the involved os's. Proposition 3.3 states this result.

Proposition 3.3

The use of RRS for selecting a sample of n = rm individuals and the ranking, with respect to B, of the reports $Z_{(i)t}$, i = 1, …, m and t = 1.,,.r, and of (3.5) for estimating the mean of Y is more accurate than the SRSWR strategy and the a gain in accuracy is measured by $\sum_{I=1}^{m}$ $\Delta^2_{B(i)}/m$, where $\Delta_{B(i)} = \mu_{B(i)} - \mu_B$. Therefore the surveyor is able to increase the gain in accuracy by taking into account that the smaller the mean of A, the larger the gain.

Comparison of the Different Alternatives

Different strategies for estimating the mean of a sensitive character are available because respondents may be ranked using Y, A or B. It is known that when Y is obtained by a direct response the RSS estimator of the mean is generally more accurate than SRSWR (see Patil, et al., 1999 for a clarifying discussion of this fact). The comparison of the accuracy of the estimators will be made analyzing the design expected model variances.

Comparing the SRSWR's estimator and its RSS counterpart when the rank is made using the sampler's judgment on Y. Taking

$$D_{rss,srs} = E_d V_R(\overline{R}) - E_d V_R(\overline{R}_{(rss)})$$

results in

$$D_{rss,srs} = -\frac{\sum_{i=1}^{m}\Delta_{Y(i)}^2}{mn\mu_A^2} + \frac{\sigma_A^2}{n\mu_A^2}\left[\frac{\sum_{i=1}^{m}\mu_{Y(i)}^2}{m} - \mu_Y^2\right]$$

This difference depends heavily on the parameters of A. Note that a large variance will provide a large value of the second term in $D_{srs,rss}$. It may be argued that this fact allows the design of a RR procedure that diminishes the sample error by determining an adequate set $\{A_1, …, A_T\}$.

In practice, ranking Y when it is a sensitive variable should be difficult. A more practical approach is that the sampled persons rank the selected A_i's and provides the pair (A_i, rank(A_i)) to the sampler. Comparing the expected errors of (2.1) and (3.2) the following are obtained

$$D_{A(rss),srs} = A* + A** - A***,$$

where

$$A* = \frac{\left(\sigma_Y^2 + \mu_Y^2\right)\sigma_A^2}{nm\mu_A^2}\left[\sum_{i=1}^{m}\left(\frac{\mu_{A(i)}^2 - \mu_A^2}{\mu_{A(i)}^2}\right)\right]$$

$$A** = \frac{\sigma_B^2}{nm\mu_A^2}\left[\sum_{i=1}^{m}\left(\frac{\mu_{A(i)}^2 - \mu_A^2}{\mu_{A(i)}^2}\right)\right]$$

$$A*** = \frac{\left(\sigma_Y^2 + \mu_Y^2\right)}{nm}\left[\sum_{i=1}^{m}\left(\frac{\Delta_{A(i)}^2}{\mu_{A(i)}^2}\right)\right]$$

and

$$\Delta_{A(i)} = \mu_{A(i)} - \mu_A$$

The gain in accuracy of RSS thus depends of the magnitude of differences between the expectation of the m os and the population mean of the ranked variable A.

The ranking of B produces the expression

$$D_{srs,B(rss)} = -\sum_{i=1}^{m} \left(\frac{\Delta_{B(i)}^2}{mn\mu_A^2} \right)$$

with $\qquad \Delta_{B(i)} = \mu_{B(i)} - \mu_B$

A comparison between two RSS alternatives may be developed by comparing their gains in accuracy. The ratio is proposed

$$D_{p,q} = \frac{D_{p,srs}}{D_{q,srs}} \qquad (4.1)$$

Methodology

To evaluate the performance of each estimator Monte Carlo experiments were developed. The first set of experiments consisted in considering $\{Y_1, \ldots, Y_N\}$, $\{A_1, \ldots, A_T\}$ and $\{B_1, \ldots, B_S\}$ as independently distributed Uniform random variables in [0, 1]. After these sets were determined (4.1) was computed, H = 1,000 experiments were conducted and $G_{p,q} = \sum_{h=1}^{1,000} \Delta_{p,q,h}/1,000$ was calculated for the different combinations of p, q = RSS, A(RSS), B(RSS) and m = 2,.,5, r = 1,..,5; results are shown in Tables 4.1 - 4.3.

Results

Table 4.1 suggests that ranking on Y provides considerable gains in accuracy with respect to ranking on A. This fact should be generated by the role of A in the reports and the absence of errors in the ranking of Y. The gains seem to be increased when m is small with respect to r for a fixed sample size n=mr.

Table 4.1: Values of $G_{rss,A(rss)}$ for the U[0, 1] Distribution

.r	.m=2	.m=3	.m=4	.m=5
1	0.725	0.087	0.047	0.033
2	0.500	0.085	0.043	0.035
3	0.483	0.083	0.055	0.035
4	0.576	0.080	0.045	0.034
5	0.435	0.088	0.055	0.029

Table 4.2 shows similar behavior of the ranking using B but the gains in accuracy with respect to the use of ranks in Y are considerably smaller. Table 4.3 establishes the preference to rank using B as opposed to A because generally it provides more accurate estimations. To increase m provides larger gains.

Table 4.2: Values of $G_{rss,B(rss)}$ for the U[0, 1] Distribution

.r	.m=2	.m=3	.m=4	.m=5
1	0.530	0.260	0.179	0.135
2	0.537	0.274	0.167	0.141
3	0.500	0.266	0.275	0.186
4	0.524	0.258	0.175	0.139
5	0.588	0.282	0.200	0.118

Table 4.3: Values of $G_{B(rss),A(rss)}$ for the U[0,1] Distribution

.r	.m=2	.m=3	.m=4	.m=5
1	1.377	0.310	0.264	0.254
2	0.934	0.310	0.264	0.253
3	0.971	0.310	0.258	0.257
4	1.111	0.310	0.203	0.251
5	0.746	0.310	0.252	0.246

Another set of experiments was developed considering an exponential with $\lambda = 1$ for generating the involved variables; results are shown in Tables 4.4 - 4.6. Table 4.4 suggests a

253

preference for ranking on Y. The gains in accuracy are larger than when the distribution is uniform.

Table 4.4: Values of $G_{rss,A(rss)}$ for the Exp[1] Distribution

.r	.m=2	.m=3	.m=4	.m=5
1	0.072	0.012	0.017	0.017
2	0.062	0.021	0.023	0.019
3	0.075	0.014	0.017	0.014
4	0.077	0.016	0.016	0.012
5	0.093	0.012	0.015	0.016

The results shown in Table 4.5 establish that the ranking using B may be better than to rank using Y for m > 3; Table 4.6 indicates the preference for using B instead of A.

Table 4.5: Values of $G_{rss,B(rss)}$ for the Exp[1] Distribution

.r	.m=2	.m=3	.m=4	.m=5
1	0.688	0.342	1.234	1.413
2	0.672	0.543	1.340	1.238
3	0.700	0.320	1.467	1.350
4	0.683	0.345	1.650	1.379
5	0.780	0.300	1.350	1.626

Table 4.6: Values of $G_{B(rss),A(rss)}$ for the Exp[1] Distribution

.r	.m=2	.m=3	.m=4	.m=5
1	0.105	0.035	0.013	0.013
2	0.092	0.039	0.017	0.015
3	0.107	0.043	0.012	0.011
4	0.113	0.043	0.010	0.015
5	0.119	0.034	0.011	0.009

Conclusions

The results of these experiments indicate that the best strategy is to have the sampled persons communicate only the value of B and to rank its values accordingly. This indication is sustained by the usual difficulty for obtaining a perfect ranking of Y.

References

Barabasi, L., & Pisani, C. (2002). Ranked set sampling for replicated sampling designs. *Biometrics*, *58*, 586-592.

Bouza, C. N. (2002a). Estimation of the mean in ranked set sampling with non responses. *Metrika*, *56*, 171-179

Bouza, C. N. (2002b). Ranked set subsampling the non response strata for estimating the difference of means. *Biometrical Journal*, *44*, 903-915.

Chaudhuri, A., Maiti, T., & Roy, R. (1996). A note on competing variance estimators in randomized response surveys. *Australia & New Zealand Journal of Statistics*, *38*, 35-42.

Chaudhury, A., & Stenger, H. (1992). *Sampling survey*. New York: M. Dekker.

David, H. A., & Levine, D. W. (1972). Ranked set sampling in the presence of judgment error. *Biometrics*, *28*, 553-555.

Patil, G. P., Sinha, A. K., & Taillie, C. (1994). Ranked set sampling. In *Handbook of statistics*, *Vol. 12*, G. P. Patil & C. R. Rao, Eds. Amsterdam: North Holland.

Patil, G. P., Sinha, A. K., & Taillie, C. (1999). Ranked set sampling: a bibliography. *Environmental & Ecological Statistics*, *6*, 91-98.

Patil, G. P (2002). Ranked set sampling. In *Encyclopedia of Enviromentrics*, *Vol. 3*, A. H. El-Shaarawi & W. W. Pieegoshed, Eds., 1684-1690. Chichester: Wiley.

Singh S., Horn, S. & Chowdhury, H. (1998). Estimation of stigmatized characteristics of a hidden gang in a finite population. *Australia & New Zealand Journal of Statistics*, *40*, 291-297.

Warner, S. (1965). Randomized responses: A survey technique for eliminating evasive survey bias. *Journal of the American Statistical Association*, *60*, 63-69.

Zou, G. (1997). Two-stage randomized response technique as a single stage procedure. *Australia & New Zealand Journal of Statistics*, *39*, 235-236.