Sankhyā : The Indian Journal of Statistics
2015, Volume 77-B, Part 1, pp. 75-83

Estimation Of Finite Population Proportion In Randomized Response Surveys Using Multiple Responses

S. Sengupta
University of Calcutta, Kolkata, India

Abstract

We consider the problem of unbiased estimation of a finite population proportion related to a sensitive attribute under a randomized response model when independent responses are obtained from each sampled individual as many times as he/she is selected in the sample. We identify a minimal sufficient statistic for the problem and obtain complete classes of unbiased and linear unbiased estimators. We also prove the admissibility of two linear unbiased estimators and the non-existence of a best unbiased or a best linear unbiased estimator.

AMS (2000) *subject classification*. MSC 62D05, 62C15.
Keywords and phrases. Admissibility, Best unbiased estimator, Complete class, Minimal sufficient statistic, Multiple responses, Population proportion, Randomized response

1 Introduction

Consider a finite population of labeled units and suppose that the problem is to estimate certain population parameters on surveying a random sample of units. In an open set-up it is assumed that an exact response can be obtained from each sampled unit through a direct survey. However, if the character of interest is sensitive or stigmatizing such as drinking alcohol or gambling habit, drug addiction, tax evasion, history of induced abortions etc., a direct survey is likely to yield unreliable responses and an alternative technique, introduced by Warner (1965), is to obtain responses through a *randomized response* (RR) survey wherein every sampled unit is asked to give a response through an RR device as per instructions from the investigator. We refer to Chaudhuri and Mukerjee (1988), Chaudhuri (2011) and Chaudhuri and Christofides (2013) for a comprehensive review of such RR procedures.

Sengupta and Kundu (1989) considered the problem of unbiased estimation of a finite population mean of a sensitive quantitative variable (or the proportion bearing a sensitive attribute) under an RR model and had obtained certain non-existence, admissibility and optimality results paralleling those based on direct responses assuming tacitly that a single randomized response is obtained from every sampled unit even if selected more than once in the sample. However, for with replacement sampling, a population unit may be selected more than once and independent randomized responses may be obtained from it as many times as it is selected in the sample. In fact, for several RR devices suggested in the literature including that of Warner (1965), an infinite population set-up (or equivalently simple random with replacement sampling) had been assumed and the proposed estimators of population parameters were based on independent repeated responses from the sampled units. Estimators based on independent repeated responses in the finite population set-up had also been discussed e.g. in Arnab (1999) and Chaudhuri et al. (2011, a, b)

In this paper we consider the problem of estimation of the population proportion related to a sensitive attribute based on such independent multiple randomized responses under an RR model. In Section 3 we identify a minimal sufficient statistic for the problem and obtain complete classes of unbiased and linear unbiased estimators. We prove the admissibility of two linear unbiased estimators in Section 4 and the non-existence of a best unbiased or a best linear unbiased estimator in Section 5.

2 Notations and Preliminaries

Let $U = \{1, 2, \ldots, i, \ldots, N\}$ be a finite population of N labeled units and Y be an indicator variable with unknown value y_i for the population unit $i, 1 \leq i \leq N$, where y_i is 1 or 0 according as the unit i does or does not possess a certain attribute. A sequence of units of U with or without repetitions is considered as a sample s which is selected with a known probability $p(s)(> 0)$ from a set of possible samples. It is assumed that $p(s)$ is independent of $\mathbf{y} = (y_1, y_2, \ldots, y_N)$. The collection of all such $p(s)$ is called a sampling design p. The problem of interest is to estimate unbiasedly the unknown population proportion $\theta = \frac{1}{N}\sum_{i=1}^{N} y_i$ on surveying a sample of units selected according to a given sampling design p. It is assumed that for p, the inclusion probability π_i of unit i in a sample, defined by $\sum_{s \supset i} p(s)$, is positive for every $i = 1, \ldots, N$. Any p is said to be a without replacement (WOR) sampling design if $f_{si} = 1$ for each i $\in$ s and for each s and a with replacement (WR) sampling design otherwise, where f_{si} is the number of times the population unit i isselected in s.

We consider the attribute to be sensitive and suppose that some RR device R is employed to produce a randomized response z_i on the population unit i when included in a sample. We assume that under R, z_i's are independently distributed with

$$Pr[z_i = 1] = a\ y_i + b = \phi(y_i), \text{say}, Pr[z_i = 0] = 1 - \phi(y_i), 1 \le i \le N \quad (2.1)$$

for some known constants $a(\neq 0)$ and b independent of $i, 0 < a + b, b < 1$. The RR model (2.1) holds for many RR devices e.g. for that due to Warner (1965) with $a = 2q - 1, b = 1 - q$ which consists of asking a population unit i when included in a sample to report y_i or $1 - y_i$ with probability q or $1-q(0 < q < 1, q \neq 1/2)$ which are known to the investigator. The RR device due to Eriksson (1973) is also a special case of (2.1) with $a = q_0, b = q_1$ in which a population unit i is asked to report y_i, 1 or 0 with given probabilities q_0, q_1 or $1 - q_0 - q_1(q_0, q_1 > 0, q_0 + q_1 < 1)$. Writing $r_i = (z_i - b)/a$ it follows that under (2.1)

$$E_R(r_i) = y_i, V_R(r_i) = c_1\ y_i + c_2, 1 \le i \le N \quad (2.2)$$

for some known constants c_1 and $c_2(c_1 + c_2, c_2 > 0)$, where the suffixes p, R and both on E (or V) are used to denote the expectations (or variances) with respect to p, R and both. As for example, for Warner (1965) device, $c_1 = 0, c_2 = q(1 - q)/(2q - 1)^2$, while for Eriksson (1973) device, $c_1 = (1 - q_0 - 2q_1)/q_0, c_2 = q_1(1 - q_1)/q_0^2$.

If a single response is obtained from each distinct population unit in s, one obtains the data d $= \{z_i : i \in s\}$, where $\{i : i \in s\}$ is the set of distinct population units in s. Let $\boldsymbol{A}_u$ be the class of all unbiased estimators e (d) based on d satisfying $E_{pR}(e(d)) = E_p E_R(e(d)) = \theta\ \forall\ \boldsymbol{y}$ and $\boldsymbol{L}_u$ be the sub-class of all linear unbiased estimators of the form

$$e(d) = \alpha_s + \sum_{i\in s} b_{si} r_i, \sum_s \alpha_s p(s) = 0, \sum_{s\supset i} b_{si} p(s) = 1/N\ \forall\ i. \quad (2.3)$$

The sub-class of estimators of the form (2.3) with $\alpha_s = 0$ will be denoted by $\boldsymbol{L}_{0u}$.

If, however, independent randomized responses are obtained from each population unit as many times as it is selected in the sample, one obtains the data $d^* = \{z_{ij},\ j = 1, \dots, f_{si} : i \in s\}$, where z_{ij} is the response from unit i in its jth selection. Let $\boldsymbol{A}_u^*$ be the class of all unbiased estimators e(d^*) based on d^* satisfying $E_{pR}(e(d^*)) = E_p\, E_R(e(d^*)) = \theta\ \forall\ \boldsymbol{y}$ and $\boldsymbol{L}_u^*$ be the sub-class of all linear unbiased estimators of the form

$$e(d^*) = \alpha_s + \sum_{i\in s}\sum_{j=1}^{f_{si}} b_{sij} r_{ij}, \sum_s \alpha_s p(s) = 0, \sum_{s\supset i} b'_{si} p(s) = 1/N\ \forall\ i \quad (2.4)$$

where $r_{ij} = (z_{ij} - b)/a$ and $b'_{si} = \sum_{j=1}^{f_{si}} b_{sij}$. The sub-class of estimators of the form (2.4) with $\alpha_s = 0$ will be denoted by $\boldsymbol{L}^*_{0u}$. The classes $\boldsymbol{A}_u, \boldsymbol{L}_u$ and $\boldsymbol{L}_{0u}$ are, respectively, sub-classes of $\boldsymbol{A}^*_u, \boldsymbol{L}^*_u$ and $\boldsymbol{L}^*_{0u}$ since d is equivalent to $\{z_{i1} : i \in s\}$ Writing $d^*_0 = \{z'_i ; i \in s\}$ with $z'_i = \sum_{j=1}^{f_{si}} z_{ij}$, we also denote by $\boldsymbol{A}^{**}_u$ the sub- class of $\boldsymbol{A}^*_u$ consisting of all unbiased estimators e(d^*_0) based on d^*_0 and by $\boldsymbol{L}^{**}_u$ the sub-class of $\boldsymbol{L}^*_u$ consisting of estimators of the form

$$e(d^*_0) = \alpha_s + \sum_{i \in s} b'_{si} r'_i, \sum_s \alpha_s p(s) = 0, \sum_{s \supset i} b'_{si} f_{si} p(s) = 1/N \ \forall \ i \tag{2.5}$$

where $r'_i = \sum_{j=1}^{f_{si}} r_{ij}$. The sub-class of $\mathbf{L}^*_{0u}$ consisting of estimators of the form (2.5) with $\alpha_s = 0$ will be denoted by $\mathbf{L}^{**}_{0u}$. For a WOR sampling design, the classes $\mathbf{A}_u, \mathbf{A}^*_u$ and $\mathbf{A}^{**}_u$ (or $\mathbf{L}_u, \mathbf{L}^*_u$ and $\mathbf{L}^{**}_u$ or $\mathbf{L}_{0u}, \mathbf{L}^*_{0u}$ and $\mathbf{L}^{**}_{0u}$) are all identical.

An unbiased estimator e = e (d^*) is said to be better than an unbiased estimator $e' = e'(d^*)$ if

$$V_{pR}(e) = V_p E_R(e) + E_p V_R(e) \leq V_{pR}(e') = V_p E_R(e') + E_p V_R(e') \, \forall \, \mathbf{y}$$

with strict inequality for at least one $\mathbf{y}$. An estimator e $\in$ a sub-class C of $\boldsymbol{A}^*_u$ is said to be admissible in C if there does not exist any estimator $e' \in$ C better than e and e is said to be uniformly best in C if $V_{pR}(e) \leq V_{pR}(e') \, \forall \, \mathbf{y} \, \forall \, e' \in$ C.

3 Minimal Sufficient Statistic and Complete Classes of Estimators

We first identify a minimal sufficient statistic for the problem in the followingtheorem.

Theorem 3.1. *For any given p,d^*_0 is a minimal sufficient statistic.*

Proof. Given the selection of s, $z_{ij} i \in s, j = 1, \ldots, f_{si}$ are independently distributed with

$$Pr_y[z_{ij} = 1|s] = \phi(y_i), Pr_y[z_{ij} = 0|s] = 1 - \phi(y_i).$$

Hence, the conditional likelihood function given the selection of s is

$$P_y(d^*|s) = \prod_{i \varepsilon s} \left\{ [\phi(y_i)]^{z'_i} \, [1 - \phi(y_i)]^{f_{si} - z'_i} \right\} = p_y(d^*_0), \text{ say}$$

and the unconditional likelihood is p(s)$p_y(d_0^*)$ whence the sufficiency of d_0^* follows by Factorization Theorem. Following the proof of Theorem 2.3 of Cassel et al. (1977), it can be shown that d_0^* is, in fact, minimal sufficient.

It now follows, by Rao - Blackwell Theorem, that given any estimator e$(d^*) \in \boldsymbol{A}_u^*$ but $\notin \boldsymbol{A}_u^{**}$ $\exists$ an estimator $e^*(d_0^*) = E(e(d^*)|d_0^*) \in \boldsymbol{A}_u^{**}$ which is better than e (d^*). For e(d^*) given by (2.4),

$$e^*(d_0^*) = E(e(d^*) \mid d_0^*) = \alpha_s + \sum_{i \in s} \frac{b'_{si}}{f_{si}} r'_i, b'_{si} = \sum_{j=1}^{f_{si}} b_{sij}$$

since $Pr[z_{ij} = 1|z'_i, s] = z'_i/f_{si}, Pr[z_{ij} = 0|z'_i, s] = 1 - z'_i/f_{si}$ implying that $E[z_{ij}|d_0^*] = z'_i/f_{si}, i \in s, j = 1, \ldots, f_{si}$.

We thus obtain the following theorem.

Theorem 3.1. *For any given p, the class of estimators* $\boldsymbol{A}_u^{**}(or \boldsymbol{L}_u^{**} or \boldsymbol{L}_{0u}^{**})$ *is complete in* $\boldsymbol{A}_u^*$ *(or in* $\boldsymbol{L}_u^*$ *or in* $\boldsymbol{L}_{0u}^*$*).*

Remark 3.1. *Arnab (1999) had shown that for a WR sampling design, any estimator of the form (2.3) can be improved upon by an estimator given by (2.5) with* $b'_{si} = b_{si}/f_{si}$*. This now follows as a corollary to Theorem 3.2 and more generally, for a WR sampling design, any unbiased estimator based on d can be improved upon by an unbiased estimator based on* d_0^**.*

4 Admissible Linear Unbiased Estimators

In this section we prove the admissibility of two linear unbiased estimators viz.

$$e_1 = \sum_{i \in s} b'_i r'_i, b'_i = \frac{1}{NE_p(f_{si})} = \frac{1}{N \sum_{s \ni i} f_{si} p(s)} \text{ and}$$

$$e_2 = \sum_{i \in s} b^*_{si} r'_i, b^*_{si} = \frac{1}{N f_{si}} \frac{\left(1 + \frac{c_1 + N c_2}{f_{si}}\right)^{-1}}{E_p \left(1 + \frac{c_1 + N c_2}{f_{si}}\right)^{-1}} = \frac{1}{N f_{si}} \frac{\left(1 + \frac{c_1 + N c_2}{f_{si}}\right)^{-1}}{\sum_{s \ni i} p(s) \left(1 + \frac{c_1 + N c_2}{f_{si}}\right)^{-1}}.$$

The results are given in the following theorem.

Theorem 4.1. *For a given p,*

(i) e_1 *is admissible in* $\boldsymbol{L}_u^{**}$ *(and hence in* $\boldsymbol{L}_u^*$*).*

(ii) e_2 *is admissible in* $\boldsymbol{L}_{0u}^{**}$ *(and hence in* $\boldsymbol{L}_{0u}^*$*).*

PROOF. (i) Consider an estimator e of the form (2.5). Then for $\theta = 0$ i.e. $\mathbf{y} = (0, \ldots 0)$,

$$\begin{aligned} V_{pR}(e) &= V_p E_R(e) + E_p V_R(e) \geq E_p V_R(e) \\ &= c_2 \sum_{i=1}^{N} \sum_{s \supset i} p(s) b_{si}'^2 f_{si} \geq \frac{c_2}{N^2} \sum_{i=1}^{N} \frac{1}{E_p(f_{si})} \end{aligned}$$

on using Cauchy - Schwartrz inequality and the condition of unbiasdness with equality holding if and only if $\alpha_s = 0 \,\forall\, s$ and $b_{si}' = b_i' \,\forall\, s \supset i \,\forall\, i$. Thus for $\theta = 0, V_{pR}(e)$ is uniquely minimized for $e = e_1$ and, hence, there can not exist any e $\in \boldsymbol{L}_u^{**}$ better than e_1.

(ii) Consider an e of the form (2.5) with $\alpha_s = 0 \,\forall\, s$ and let $V_i(e)$ be $V_{pR}(e)$ for any with $y_i = 1$ and $y_j = 0 \,\forall\, j (\neq i)$. Then, as in the proof of (i),

$$\begin{aligned} \sum_{i=1}^{N} V_i(e) &= \sum_{i=1}^{N} \sum_{s \supset i} p(s) b_{si}'^2 f_{si}^2 \left(1 + \frac{c_1 + N c_2}{f_{si}}\right) \\ &- \frac{1}{N} \geq \frac{1}{N^2} \sum_{i=1}^{N} \frac{1}{E_p \left(1 + \frac{c_1 + N c_2}{f_{si}}\right)^{-1}} - \frac{1}{N} \end{aligned}$$

and equality holds if and only if $b_{si}' = b_{si}^* \,\forall\, s \supset i \,\forall\, i$. Thus $\sum_{i=1}^{N} V_i(e)$ is uniquely minimized for e $= e_2$ and, hence, there can not exist any e $\in \boldsymbol{L}_{0u}^{**}$ better than e_2.

For a WOR sampling design, both the estimators e_1 and e_2 reduce to Horvitz and Thompson (1952) type estimator $e_{HT} = \frac{1}{N} \sum_{i \in s} \frac{r_i}{\pi_i}$. However, for a WR sampling design, the two estimators are different unless for each $i = 1, \ldots, N, f_{si} = f_i \,\forall\, s \supset i$.

For simple random sampling with replacement (SRSWR) involving n draws, the estimator e_1 reduces to Warner (1965) type estimator

$$e_W = \frac{1}{n} \sum_{i \in s} r_i'$$

For SRSWR involving n draws, Chaudhuri et al. (2011, a) had considered two other linear unbiased estimators

$$e_3 = \frac{1}{\nu} \sum_{i \in s} \frac{r_i'}{f_{si}} \text{ and } e_4 = \frac{1}{N\pi} \sum_{i \in s} \frac{r_i'}{f_{si}}$$

where $\nu = \nu(s)$ is the number of distinct population units in s and $\pi = 1 - (1 - 1/N)^n$ is the inclusion probability of each unit in a sample. These two estimators may, however, be inadmissible in $\boldsymbol{L}_{0u}^{**}$ which may be demonstrated through the following example.

Example 4.1. *Consider Warner (1965) RR device with q = 0.4. By the results in Chaudhuri et al. (2011, a)*

$$V_{pR}(e_W) - V_{pR}(e_3) = k_1\theta(1-\theta) - 6k_2 \leq \frac{k_1}{4} - 6k_2 \,\forall\, \boldsymbol{y} \tag{4.1}$$

and

$$V_{pR}(e_W) - V_{pR}(e_4) = \frac{k_3\theta - k_4\theta^2}{Nn\pi^2} - 6k_5 \leq \frac{k_3^2}{4Nn\pi^2 k_4} - 6k_5 \,\forall\, \boldsymbol{y} \tag{4.2}$$

with $k_1 = \frac{1}{n} - \frac{N}{N-1}\left\{E_p\left(\frac{1}{\nu}\right) - \frac{1}{N}\right\}, k_2 = E_p\left(\frac{1}{\nu^2}\sum_{i\in s}\frac{1}{f_{si}}\right) - \frac{1}{n}, k_3 = N\pi^2 - n(\pi - \pi^*), k_4 = N\{\pi^2 - n(\pi^2 - \pi^*)\}$ *and* $k_5 = \frac{1}{N^2\pi^2}E_p\left(\sum_{i\in s}\frac{1}{f_{si}}\right) - \frac{1}{n}$, *where* $\pi^* = 1 - 2(1-1/N)^n + (1-2/N)^n$ *is the joint inclusion probability of each pair of population units in a sample.*

Computations give $k_1 = 1/24, k_2 = 3/128$ and the RHS of (4.1) < 0 for $N = 4, n = 3$ and $k_3 = 7/27, k_4 = 11/27, k_5 = 1/25$ and the RHS of (4.2) < 0 for $N = 3, n = 2$. Thus e_W is better than e_3 and e_4 for these two sets of values of N and n, respectively.

Chaudhuri et al. (2011, b) had discussed another WR sampling scheme (see also Raj and Khamis 1958), known as inverse SRSWR , in which units are selected by SRSWR until a fixed number of distinct population units, ν say, are included in s. It can be readily verified that for this sampling scheme, $E_p(f_{si}) = N^{-1}E_p(n(s))$ for each i so that e_1 takes the form

$$e_{1l} = \frac{1}{E_p(n(s))}\sum_{i\in s} r_i'$$

where n(s) is the total number of units (including repetitions) in s. For inverse SRSWR, Chaudhuri et al. (2011, b) considered two other linear unbiased estimators

$$e_{3l} = \frac{1}{\nu}\sum_{i\in s}\frac{r_i'}{f_{si}} \text{ and } e_{4l} = \frac{1}{n(s)}\sum_{i\in s} r_i'.$$

From the results of Chaudhuri et al. (2011, b), it follows that, for Warner (1965) RR device, $V_{pR}(e_{4l}) - V_{pR}(e_{3l}) = k_1\theta(1-\theta) - c_2k_2$ for some $k_1(> 0)$ and $k_2 = \frac{1}{\nu^2}E_p\left(\sum_{i\in s}\frac{1}{f_{si}}\right) - E_p\left(\frac{1}{n(s)}\right)$. It may similarly be seen that $V_{pR}(e_{1l}) -$

$V_{pR}(e_{4l}) = k_3\theta - k_4\theta^2 - c_2k_5$ for some k_3, k_4 and $k_5 = E_p\left(\frac{1}{n(s)}\right) - \frac{1}{E_p(n(s))}$. Since $k_2, k_5 > 0$, examples can be obtained as in Example 4.1 to show that $V_{pR}(e_{1l}) < V_{pR}(e_{4l}) < V_{pR}(e_{3l}) \,\forall\, \mathbf{y}$ for certain values of q, N and ν. Thus both e_{3l} and e_{4l} may be inadmissible in $\boldsymbol{L}_{0u}^{**}$.

5 Non-existence of Best Estimators

Sengupta and Kundu (1989) had proved that for a given p, there does not exist a uniformly best estimator in $\boldsymbol{A}_u$ or in $\boldsymbol{L}_u$ unless p is a census i.e. every s includes all the population units and there exists a uniformly best estimator in $\boldsymbol{L}_{0u}$ if and only if p is a *unicluster sapling design* (UCSD) i. e. any two samples either include the same set of distinct population units or do not include any common population unit. Following the proof of Theorem 2.1 of Sengupta and Kundu (1989), it can be easily shown that there does not exist a uniformly best estimator in $\boldsymbol{A}_u^{**}$ or in $\boldsymbol{L}_u^{**}$ (and hence in $\boldsymbol{A}_u^{*}$ or in $\boldsymbol{L}_u^{*}$) for any non-census p. In the following theorem we prove that there also generally does not exist a uniformly best estimator in $\boldsymbol{L}_{0u}^{**}$ (and hence in $\boldsymbol{L}_{0u}^{*}$).

Theorem 5.1. *For a given sampling design p, there exists a uniformly best estimator in $\boldsymbol{L}_{0u}^{**}$ if and only if p is a UCSD with $f_{si} = f_i \,\forall\, s \supset i \,\forall\, i$.*

Proof. Let p be such that for at least one unit i and for two samples s_1 and s_2 containing i, $f_{s_1i} \neq f_{s_2i}$. Then e_1 and e_2, defined in the earlier section, are two different estimators both of which are admissible in $\boldsymbol{L}_{0u}^{**}$. Hence, there can not exist a uniformly best estimator in $\boldsymbol{L}_{0u}^{**}$.

Consider now a sampling design p for which $f_{si} = f_i \,\forall\, s \supset i \,\forall\, i$. Since $e_1 = \frac{1}{N}\sum_{i\in s}\frac{r_i'}{f_i\pi_i}$ is admissible in $\boldsymbol{L}_{0u}^{**}$, a uniformly best estimator in $\boldsymbol{L}_{0u}^{**}$, if exists, must be e_1 and, hence, by Rao (1952) Theorem

$$E_{pR}(e_1h) = 0 \,\forall\, \mathbf{y} \tag{5.1}$$

for any $h = h(d_0^*) = \sum_{i\in s} c_{si}r_i'$, $\sum_{s\supset i} c_{si}p(s) = 0 \,\forall\, i$. It is easy to verify that (5.1) holds if and only if

$$E_p[E_R(e_1)h^*] = E_p\left[\frac{1}{N}\sum_{i\in s}\frac{y_i}{\pi_i}h^*\right] = 0 \,\forall\, \mathbf{y}$$

for any $h^* = \sum_{i\in s} c_{si}y_i$, $\sum_{s\supset i} c_{si}p(s) = 0 \,\forall\, i$ i.e. the Horvitz and Thompson (1952) estimator $\frac{1}{N}\sum_{i\in s}\frac{y_i}{\pi_i}$ is the best homogeneous linear unbiased estimator of θ

based on direct responses which can be true if and only if p is a UCSD see (Godambe, 1955; Hanurav, 1966). Hence, there exists a uniformly best estimator in $\boldsymbol{L}_{0u}^{**}$ if and only if p is a UCSD.

References

ARNAB, R. (1999). On use of distinct respondents in RR surveys. *Biometrical Journal* **41**, 507–513.

CASSEL, C. M., SARNDAL, C. E. and WRETMAN, J. H. (1977). *Foundations of Inference in Survey Sampling*. John Wiley, New York.

CHAUDHURI, A. (2011). *Randomized Response and Indirect Questioning Techniques in Surveys. CRC Press, Chapman and Hall, Taylor & Francis Group*. Boca Raton, Florida,USA.

CHAUDHURI, A., BOSE, M. and DIHIDAR, K. (2011, a). Estimating sensitive proportions by Warner's randomized response technique using multiple responses from distinct persons. *Statistical Papers* **52**, 111–124.

CHAUDHURI, A., BOSE, M. and DIHIDAR, K. (2011, b). Estimation of a sensitive proportion by Warner's randomized response data through inverse sampling. *Statistical Papers* **52**, 343-354.

CHAUDHURI, A. and CHRISTOFIDES, T. C. (2013). *Indirect Questioning in Sample Surveys.* Springer-Verlag, Heidelberg, Germany.

CHAUDHURI, A. and MUKERJEE, R. (1988). *Randomized Response: Theory and Techniques.* Marcel Dekker, New York.

ERIKSSON, S. A. (1973). A new model for randomized response. *Interenat. Stat. Rev.* **41**, 101–113.

GODAMBE, V. P. (1955). A unified theory of sampling from finite populations. *J. Roy. Statist. Soc., Ser B* **17**, 269–278.

HANURAV, T. V. (1966). Some aspects of unified sampling theory. *Sankhya, Ser A* **28**, 175–204.

HORVITZ, D. G. and THOMPSON, D. J. (1952). A generalization of sampling without replacement. *J. Amer. Statist. Assoc.* **47**, 663–685.

RAJ, D. and KHAMIS, S. H. (1958). Some remarks on sampling with replacement. *Ann. Math. Statist* **39**, 550–557.

RAO, C. R. (1952). Some theorems on minimum variance estimation. *Sankhya* **12**, 27–42.

SENGUPTA, S. and KUNDU, D. (1989). Estimation of finite population mean in randomized response surveys. *J. Statist. Plann. Inference* **23**, 117–125.

WARNER, S. L. (1965). Randomized response - A survey technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.* **60**, 63–69.

S. SENGUPTA
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALCUTTA
35 BALLYGUNGE CIRCULAR ROAD
KOLKATA 700019 INDIA
E-mail: samindras@yahoo.co.in

Paper received: 15 January 2014.