Estadística Española • Volumen 59, número 193 / 2017, pp. 63-72

On Comparison Of Horvitz-Thompson And Murthy's Sampling Strategies For Estimating Sensitive Finite Population Totals Under Scrambled Randomized Response Plans

Samindranath Sengupta
Department of Statistics, University of Calcutta

Abstract

We consider the problem of unbiased estimation of a finite population total related to a sensitive quantitative variable under two scrambled randomized response plans and compare the relative efficiency of the unequal probability sampling strategies due to Horvitz–Thompson (1952) and Murthy (1957) under a super-population model depending on a parameter g. It is shown that for the linear plan the model expected variance is smaller for Murthy's (1957) strategy if $g \leq 1$, while for the multiplicative plan the model expected variance is smaller for the Horvitz-Thompson (1952) strategy if $g \geq 2$. We also address the problem of unbiased estimation of the variances of these two sampling strategies under the two randomized response plans and study the non-negative property of the variance estimators.

Key Words: Model expected variance, Population total, Randomized response, Sampling strategy, Super-population model, Unequal probability sampling, Variance estimation.

AMS Classification: 62D05

En la comparación de Horvitz-Thompson y Estrategias de muestreo de Murthy para la estimación Totales de población finita sensible debajo Planes de respuesta aleatorizados codificados

Resumen

Consideramos que el problema de la estimación no sesgada de un total finito de población está relacionado a una variable cuantitativa sensible bajo dos planes aleatorizados de respuesta aleatoria y comparar la eficiencia relativa de las estrategias desiguales de muestreo probabilístico debido a Horvitz-Thompson (1952) y Murthy (1957) bajo una superpoblación modelo dependiendo de un parámetro g. Se muestra que para el plan lineal, el modelo la varianza esperada es menor para la estrategia de Murthy (1957) si g ≤ 1, mientras que para el plan multiplicativo, la varianza esperada del modelo es menor para los Horvitz-Estrategia de Thompson (1952) si g ≥ 2. También abordamos el problema de la estimación de las varianzas de estas dos estrategias de muestreo bajo los dos planes de respuesta aleatoria y estudiar la propiedad no negativa de la varianza estimadores.

Palabras clave: Varianza esperada del modelo, Total de la población, Respuesta aleatorizada, Estrategia de muestreo, Modelo de superpoblación, Muestreo de probabilidad desigual, Estimación de varianza.

Clasificación AMS: 62D05

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.

In comparing the relative efficiency of unbiased sampling strategies for estimating a finite population parameter it is customary to compare their model expected variances under some super-population models. Rao (1966), Hanurav (1967), Rao (1967), Chaudhuri and Arnab (1979) and Sengupta (2016) had compared some unequal probability sampling strategies for estimating the population total of a quantitative variable in direct surveys under a certain super-population model depending on a parameter g. In this paper we make a similar comparison of the unequal probability sampling strategies due to Horvitz-Thompson (1952) and Murthy (1957) in terms of their model expected variances under such a super-population model when the character is a sensitive one and these strategies are based on data obtained through an RR survey employing either of the two scrambled RR plans mentioned in Warner (1971) and extensively studied in Pollock and Bek (1976) and Eichhorn and Hayre (1983). It is shown that for the linear plan the model expected variance is smaller for Murthy's (1957) strategy if $g \leq 1$, while for the multiplicative plan the model expected variance is smaller for the Horvitz-Thompson (1952) strategy if $g \geq 2$.

We also address the problem of unbiased estimation of the variances of these two sampling strategies under the two randomized response plans and following Arnab (1994) study the non-negative property of the variance estimators which are useful for determination of confidence intervals of the population total.

2. Notations and Preliminaries

Let $U = \{1, 2,..., i,\ldots, N\}$ be a finite population of N labeled units and Y be a quantitative variable with unknown value y_i for the population unit i, $1 \le i \le N$. The problem of interest is to estimate unbiasedly the unknown population total $\theta = \sum_{i=1}^{N} y_i$ on surveying a sample of units s selected from a set of samples S with a given probability $p(s)$ (> 0) i.e. according to a given sampling design p. A sampling design p together with an unbiased estimator e of θ is called an unbiased sampling strategy for estimating θ and is denoted by $H = (p, e)$.

We consider the following two unbiased strategies in an open set-up due to Horvitz-Thompson (1952) and Murthy (1957) based on known normed size measure w_i for unit i, $1 \le i \le N$, $w_i > 0\ \forall i$, $\sum_{i=1}^{N} w_i = 1$.

$H = (p,e)$: n units are selected without replacement (WOR) from U such that the inclusion probability π_i of unit i in a sample, defined as $\sum_{s:i\in s} p(s)$, is nw_i and $e = \frac{1}{n}\sum_{i\in s}\frac{y_i}{w_i}$.

$H' = (p',e')$: n units are selected with probability proportional to size without replacement and $e'(s) = \frac{\sum_{i\in s} y_i p'(s|i)}{p'(s)}$ where $p'(s|i)$ is the conditional probability of selecting the sample s given that the first unit selected is the population unit i.

The variances of these two strategies are given by

$$V(H) = \frac{1}{n}\sum_{i=1}^{N}\frac{y_i^2}{w_i} + \frac{1}{n^2}\sum_{i\neq}\sum_{j=1}^{N}\frac{y_i y_j}{w_i w_j}\pi_{ij} - \theta^2 \quad [2.1]$$

$$\mathrm{V(H')} = \sum_{s\in S}\frac{\left(\sum_{i\in s} y_i p'(s|i)\right)^2}{p'(\mathrm{s})} - \theta^2 \quad [2.2]$$

where π_{ij} is the joint inclusion probability of population units i and j in a sample under p defined as $\sum_{s:i,j\in s} p(s)$.

Sengupta (2016) had compared the model expected variances of these two sampling strategies under a super-population model M i.e. a class of prior distributions α of $\boldsymbol{y} = (y_1,..., y_N)$ under which

$$E_\alpha(y_i) = \beta w_{i,} V_\alpha(y_i) = \sigma^2 w_i^g, 1 \le 1 \le N, Cov_\alpha(y_i, y_j) = 0, 1 \le i \ne j \le N \quad [2.3]$$

where $\beta, \sigma^2 (> 0)$ are unknown parameters and g (≥ 0) is known or unknown and the suffix α on *E, V or Cov* is used to denote the expectation, variance or covariance with respect to the prior distribution α. The results obtained are summarized in the following theorem.

Theorem 2.1 For $n \ge 2$ and w_i's not all equal,
(i) $E_\alpha V(H') < E_\alpha V(H)\ \forall \alpha$ if $g \le 1$
(ii) $E_\alpha V(H) < E_\alpha V(H')\ \forall \alpha$ if $g \ge 2$

Suppose now the character to be sensitive and some RR device R be employed to produce a randomized response z_i on the population unit i when included in s. In what follows we shall consider two RR plans, to be denoted respectively as R_1 and R_2, mentioned in Warner (1971) and extensively studied in Pollock and Bek (1976) and Eichhorn and Hayre (1983) wherein the population unit i is asked to report $z_i = y_i + A$ or Ay_i, where A is a random variable with known probability distribution with $E(A) \ne 0$. This may be implemented e.g. by asking a sampled unit i to choose at random a number A_j out of a given set of numbers $A_1, ..., A_L, \sum_{j=1}^{L} A_j \ne 0$ and to report the value $z_i = y_i + A_j$ or $A_j y_i$ (see Chaudhuri and Christofides, 2013, Chapter 5). Writing $r_{i1} = z_i - E(A)$ and $r_{i2} = z_i / E(A)$ it follows that under R_t, r_{it}'s are independently distributed with

$$E_{R_t} = (r_{it}) = y_i, t = 1, 2, V_{R_1}(r_{i1}) = k_1, V_{R_2}(r_{i2}) = k_2 y_i^2, \quad 1 \le i \le N \quad [2.4]$$

with $k_1 = V(A), k_2 = V(A)/E^2(A)$, 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.

An RR strategy $H_R = (p, e_R)$ is said to be unbiased for estimating θ if $E_{pR}(e_R) = E_p E_R = (e_R) = \theta\ \forall y = (y_1, ..., y_N)$. It can be readily verified using [2.4] that under R_t, $t = 1, 2$, two unbiased RR strategies $H_R = (p, e_R)$ and $H'_R = (p', e'_R)$ can be derived from the strategies H and H' in the open set-up replacing y_i by r_{it} in e and e'. Also the variances of H_R and H'_R are given by

$$V(H_R) = V_p E_R(e_R) + E_p V_R(e_R) = V(H) + E_p V_R(e_R) \quad [2.5]$$

$$V(H'_R) = V_{p'} E_R(e'_R) + E_{p'} V_R(e'_R) = V(H') + E_{p'} V_R(e'_R) \quad [2.6]$$

3. Comparison under R_1

We first compare the relative efficiency of the sampling strategies under the super-population model M when these are based on data obtained from an RR survey employing the RR plan R_1. The comparison is based on the following lemma.

Lemma 3.1 For $n > 1$ and w_i's not all equal, $E_{p'}V_R(e'_R) < E_pV_R(e_R)$ under the RR plan R_1.

Proof. We note that under R_1

$$E_pV_R(e_R) = \frac{k_1}{n^2} E_p \sum_{i\in s} \frac{1}{w_i^2} = \frac{k_1}{n} \sum_{i\in 1}^{N} \frac{1}{w_i} \quad [3.1]$$

$$E_{p'}V_R(e'_R) = k_1 E_{p'}\left[\frac{\sum_{i\in s} p'^2(s|i)}{p'^2(s)}\right] = k_1 \sum_{s\in S} \frac{\sum_{i\in s} p'^2(s|i)}{p'(s)} = k_1 \sum_{i\in 1}^{N} \sum_{s:i\in s} \frac{p'^2(s|i)}{p'(s)} \quad [3.2]$$

Now as $p'(s) = \sum_{i\in s} w_i p'(s|i), \ \sum_{s:i\in s} p'(s|i) = 1,$

$$\sum_{i\in 1}^{N}\left[\frac{1}{nw_i} - \sum_{s:i\in s} \frac{p'^2(s|i)}{p'(s)}\right] = \frac{1}{n}\sum_{i=1}^{N} \frac{1}{w_i} \sum_{s:i\in s} \frac{p'(s|i)\{p'(s) - nw_i p'(s|i)\}}{p'(s)}$$

$$= \frac{1}{n}\sum_{i=1}^{N} \frac{1}{w_i} \sum_{s:i\in s} \frac{p'(s|i) \sum_{j(\neq i)\in s} \{w_j p'(s|j) - w_i p'(s|i)\}}{p'(s)}$$

$$= \frac{1}{n}\sum_{s\in S}^{N} \frac{1}{p'(s)} \sum_{i\neq} \sum_{j\in s} \frac{p'(s|i)\{w_j p'(s|j) - w_i p'(s|i)\}}{w_i}$$

$$= \frac{1}{n}\sum_{s\in S}^{N} \frac{1}{p'(s)} \sum_{i<} \sum_{j\in s} \left\{\frac{p'(s|i)}{w_i} - \frac{p'(s|i)}{w_j}\right\}\{w_j p'(s|j) - w_i p'(s|i)\}$$

$$= \frac{1}{n}\sum_{s\in S}^{N} \frac{1}{p'(s)} \sum_{i<} \sum_{j\in s} \left\{\frac{w_j p'(s|i) - w_i p'(s|i)}{w_i w_j}\right\}\{w_j p'(s|j) - w_i p'(s|i)\} =$$

$$\frac{1}{n}\sum_{s\in S}^{N} \frac{1}{p'(s)} \sum_{i<} \sum_{j\in s} \frac{\left[\{w_j p'(s|i) - w_i p'(s|i)\}^2 + (w_i + w_j)\{p'(s|i) - p'(s|j)\}\{w_j p'(s|j) - w_i p'(s|i)\}\right]}{w_i w_j}$$

by Lemma A.1 in the Appendix. Hence, follows the Lemma.

Combining Lemma 3.1 with Theorem 2.1 we immediately obtain the following theorem which shows that the relative efficiency under the super-population model M is greater for H'_R if g ≤ 1 under the RR plan R1.

Theorem 3.2 For $n > 1$ and w_i's not all equal, $E_\alpha V(H'_R) < E_\alpha V(H_R) \forall \alpha$ if $g \leq 1$ under the RR plan R_1.

4. Comparison under R_2

We now compare the relative efficiency of the sampling strategies under the super-population model M when these are based on data obtained from an RR survey employing the RR plan R_2. The comparison is again based on the following lemma.

Lemma 4.1 For $n > 1$ and w_i's not all equal, $E_\alpha E_p V_R(e_R) < E_\alpha E_{p'} V_R(e'_R) \forall \alpha$ if g ≥ 2 under the RR plan R2.

Proof. We note that under R_2

$$E_\alpha E_p V_R(e_R) = \frac{1}{n^2} E_p \sum_{i \in s} \frac{k_2\left(\beta^2 w_i^2 + o^2 w_i^g\right)}{w_i^2} = \frac{1}{n}\sum_{i=1}^{N} \frac{k_2\left(\beta^2 w_i^2 + o^2 w_i^g\right)}{w_i} = \frac{k_2}{n}\left(\beta^2 + o^2 \sum_{i=1}^{N} w_i^{g-1}\right) \quad [4.1]$$

$$E_\alpha E_{p'} V_R(e'_R) = E_{P'}\left[\frac{\sum_{i \in s} k_2\left(\beta^2 w_i^2 + o^2 w_i^g\right) p'^2(s|i)}{p'^2(s)}\right] = \sum_{s \in S} \frac{\sum_{i \in s} k_2\left(\beta^2 w_i^2 + o^2 w_i^g\right) p'^2(s|i)}{p'(s)}$$

$$= k_2\left(\beta^2 \sum_{i=1}^{N} w_i^2 \sum_{s:i \in s} \frac{p'^2(s|i)}{p'(s)} + o^2 \sum_{i=1}^{N} w_i^g \sum_{s:i \in s} \frac{p'^2(s|i)}{p'(s)}\right) \quad [4.2]$$

As in the proof of Lemma 3.1, we have

$$\frac{1}{n} - \sum_{i=1}^{N} w_i^2 \sum_{s:i \in s} \frac{p'^2(s|i)}{p'(s)} = \sum_{i=1}^{N} w_i \left[\frac{1}{n}\sum_{s:i \in s} p'(s|i) - w_i \sum_{s:i \in s} \frac{p'^2(s|i)}{p'(s)}\right]$$

$$= -\frac{1}{n}\sum_{s \in s} \frac{\sum\sum_{i<j \in s} \left\{w_i p'(s|i) - w_j p'(s|j)\right\}^2}{p'(s)} < 0 \quad [4.3]$$

and similarly

$$\frac{1}{n}\sum_{i=1}^{N} w_i^{g\text{-}1} - \sum_{i=1}^{N} w_i^{g} \sum_{s:i\in s} \frac{p'^2(s|i)}{p'(s)} = \sum_{i=1}^{N} w_i^{g\text{-}1}\left[\frac{1}{n}\sum_{s:i\in s} p'(s|i) - w_i \sum_{s:i\in s}\frac{p'^2(s|i)}{p'(s)}\right]$$

$$= -\frac{1}{n}\sum_{s\in S}\frac{\sum\sum_{i<j\in s}\left\{w_i p'(s|i) - w_j p'(s|j)\right\}\left\{w_i^{g-1}p'(s|i) - w_j^{g-1}p'(s|j)\right\}}{p'(s)} < 0 \qquad [4.4]$$

for $g \geq 2$ by Lemma A.1 in the Appendix. The proof now follows from (4.1) – (4.4).

Combining Lemma 4.1 with Theorem 2.1 we immediately obtain the following theorem which shows that the relative efficiency under the super-population model M is greater for H_R if $g \geq 2$ under the RR plan R_2.

Theorem 4.2 For $n > 1$ and w_i's not all equal, $E_\alpha V(H_R) < E_\alpha V(H'_R)\forall\alpha$ if $g \geq 2$ under the RR plan R_2.

Remark 4.3 For $g = 2$, the result follows as a special case of a more general result (see Arnab, 1995, 1998) that for g =2, the model expected variance of H_R under the super-population model M is smaller than that of any other linear unbiased strategy for the RR plan R_2.

5. Unbiased Variance Estimation

In this section, we consider the problem of unbiased estimation of the variances of the above two sampling strategies under the two randomized response plans R_1 and R_2 and study the non-negative property of the variance estimators which are useful for determination of confidence intervals of the population total.

For $n \geq 2$, it is well known that in an open set-up $V(H)$ and $V(H')$ can be unbiasedly estimated, respectively, by

$$v = \sum\sum_{i<j\in s}\left(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}}\right)\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2 \qquad [5.1]$$

and $$v' = \frac{1}{p'^2(s)}\sum\sum_{i<j\in s}\left[p'(s)p'(s|ij) - p'(s|i)p'(s|j)\right]w_iw_j\left(\frac{y_i}{w_i} - \frac{y_j}{w_j}\right)^2 \qquad [5.2]$$

where $p'(s|ij)$ is the conditional selection probability of s under p' given that the first two units selected are the population units i and j (see Yates and Grundy, 1953; Murthy, 1957). It may be noted that the estimator v' is uniformly nonnegative since $p'(s)p'(s|ij) - p'(s|i)p'(s|j) \geq 0 \;\; \forall 1 \leq i \neq j \leq N$ (see Pathak and Shukla, 1966; Andreatta and Kaufman, 1986). Also a sufficient condition for the estimator v to be uniformly nonnegative is $\pi_i\pi_j - \pi_{ij} \geq 0 \ \forall 1 \leq i \neq j \leq N$ which holds for many choices of the sampling design p (see Brewer and Hanif, 1983; Chaudhuri and Vos, 1988).

Since $V_{R_1}(r_{i1}) = k_1 \forall i$, it now follows from Arnab (1994) that under R_1 for $n \geq 2$, $V(H_R)$ and $V(H_R^{/})$ can be unbiasedly estimated, respectively, by

$$v_R = \sum_{i<}\sum_{j\in s}\left(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}}\right)\left(\frac{r_{i1}}{\pi_i} - \frac{r_{j1}}{\pi_j}\right)^2 + N k_1 \qquad [5.3]$$

and $v'_R = \frac{1}{p'^2(s)}\sum_{i<}\sum_{j\in s}\left[p'(s)p'(s|ij) - p'(s|i)p'(s|j)\right]w_iw_j\left(\frac{r_{i1}}{w_i} - \frac{r_{j1}}{w_j}\right)^2 + N k_1$ [5.4]

which are clearly uniformly nonnegative if the estimator v, defined in (5.1), is so in the open set-up.

It can also be readily verified that under p and $p^{/}$, $\sum_{i=1}^{N} V_{R2}(r_{i2}) = k_2\sum_{i=1}^{N} y_i^2 = \frac{k_2}{k_2+1}\sum_{i=1}^{N} E_{R2}(r_{i2}^2)$ can be unbiasedly estimated under R_2, respectively, by $\frac{k_2}{k_2+1}\sum_{i\in s}\frac{r_{i2}^2}{\pi_i}$ and $\frac{k_2}{k_2+1}\sum_{i\in s}\frac{r_{i2}^2 p'(s|i)}{p'(s)}$

Hence, it similarly follows from Arnab (1994) that under R_2 for $n \geq 2$, $V(H_R)$ and $V(H_R^{/})$ can be unbiasedly estimated, respectively, by

$$v_R = \sum_{i<}\sum_{j\in s}\left(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}}\right)\left(\frac{r_{i2}}{\pi_i} - \frac{r_{j2}}{\pi_j}\right)^2 + \frac{k_2}{k_2+1}\sum_{i\in s}\frac{r_{i2}^2}{\pi_i} \qquad [5.5]$$

and

$$v'_R = \frac{1}{p'^2(s)}\sum_{i<}\sum_{j\in s}\left[p'(s)p'(s|ij) - p'(s|i)p'(s|j)\right]w_iw_j\left(\frac{r_{i2}}{w_i} - \frac{r_{j2}}{w_j}\right)^2 + \frac{k_2}{k_2+1}\sum_{i\in s}\frac{r_{i2}^2 p'(s|i)}{p'(s)} \qquad [5.6]$$

which are again uniformly nonnegative if the estimator v, defined in (5.1), is so in the open set-up.

Appendix

Lemma A.1 If for $i \neq j \in s$, $w_i \geq w_j$ then (i) $w_i^k p'(s|i) \leq w_j^k p'(s|j)$ for k ≤ 0 and (ii) $w_i^k p'(s|i) \geq w_j^k p'(s|j)$ for k ≥ 1 with equalities if and only if $w_i = w_j$.

The proof of the lemma is given in Sengupta (2016).

References

ANDREATTA, G. AND KAUFMAN, G. M. (1986). «Estimation of finite population properties when sampling is without replacement and proportional to magnitude». J. Amer. Statist. Assoc., 81, 657 -666.

ARNAB, R. (1994). «Nonnegative variance estimation in randomized response surveys». Commun. Statist.- Theory Meth., 23, 1743-1752

ARNAB, R. (1995). «Optimal estimation of a finite population total under randomized response surveys». Statistics, 27, 175-180.

ARNAB, R. (1998). «Randomized response surveys: Optimum estimation of a finite population total». Statistical Papers, 39, 405-408.

BREWER, K. R. W. AND HANIF, M. (1983). *Sampling With Unequal Probabilities*. Lecture notes in Statistics Series, Springer – Verlag, 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. AND ARNAB, R. (1979). «On the relative efficiencies of sampling strategies under a super population model». Sankhya, Ser C, 41, 40-43.

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

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

CHAUDHURI, A. AND VOS, J. W. E. (1988). *Unified Theory and Srategies of Survey Sampling*. North Holland, Amsterdam.

EICHHORN, B. H. AND HAYRE, L. S. (1983). «Scrambled randomized response methods for obtaining sensitive quantitative data». Statist. Plann. Inference, 7, 307-316.

HANURAV, T. V. (1967). «Optimum utilization of auxiliary information: πps sampling of two units from a stratum». J. Roy. Statist. Soc., Ser B, 29, 374-390.

HORVITZ, D. G. AND THOMPSON, D. J. (1952). «A generalization of sampling without replacement». J. Amer. Statist. Assoc., 47, 663-685.

MURTHY, M. N. (1957). «Ordered and unordered estimators in sampling without replacement». Sankhya, 18, 379=390.

PATHAK, P. K. AND SHUKLA, N. D. (1966). «Non-negativity of a variance estimator». Sankhya, Ser A, 28, 41-46.

POLLOCK, K. H. AND BEK, Y. (1976). «A comparison of three randomized response models for quantitative data». J. Amer. Statist. Assoc., 71, 884-886

RAO, J. N. K. (1966). «On the relative efficiency of some estimators in pps sampling for multiple characteristics». Sankhya, Ser A, 28, 61-70.

RAO, T. J. (1967). «On the choice of a strategy for ratio method of estimation». J. Roy. Statist. Soc,. Ser B, 29, 392-397.

SENGUPTA, S. (2016). «On the relative efficiency of Murthy's sampling strategy under a super-population model». Model Assisted Statistics & Applications, 11, 231-234.

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

WARNER, S. L. (1971). «The linear randomized response model». J. Amer. Statist. Assoc., 66, 884-888.

YATES, F. AND GRUNDY, P. M. (1953). «Selection without replacement within strata with probabilities proportional to sizes». J. Roy. Statist. Soc,. Ser B, 15, 243-261.