Jour. Ind. Soc. Ag. Statistics
Vol. XL, No. 3 (1988), pp. 173-177

TRANSFORMATION ON THE AUXILIARY VARIATE FOR MIDZUNO-SEN SAMPLING SCHEME

T. J. RAO
Indian Statistical Institute, Calcutta
(Received : June, 1983)

SUMMARY

Prasad and Srivenkataramana [5] have used a transformation on the study variate to obtain an estimator of the population total better than the Horvitz-Thompson estimator under the Midzuno-Sen sampling scheme. We observe that this transformation and the resulting estimator depend on an unknown parameter and suggest an alternative transformation on the *auxiliary* variate and obtain a simple estimator which is empirically found to be almost as efficient as the improved estimator.

Keywords : Midzuno-Sen sampling scheme, Horvitz-Thompson estimator, Transformation on the auxiliary variate.

Introduction

Consider a finite population $(U_1, U_2, \ldots, U_N)$ of size N. Let $\mathcal{Y}$ be the study variate taking value Y_i on the unit U_i, $i = 1, 2, \ldots, N$. Auxiliary information on a variable $\mathcal{X}$ related to $\mathcal{Y}$ taking positive value X_i on U_i is available for all the units. Under the Midzuno-Sen (Midzuno [4], Sen [11]) sampling scheme, consider the Horvitz-Thompson [2] estimator given by

$$\hat{Y}_{HT} = \sum_{i=1}^{n} (y_i/\pi_i). \tag{1.1}$$

based on a sample of size n for estimating the population total $Y = \sum_{i=1}^{N} Y_i$. Here the inclusion probability of the ith unit is

$$\pi_i = \alpha_1 + \alpha_2 P_i \tag{1.2}$$

where $\alpha_1 = (n-1)/(N-1)$, $\alpha_2 = (N-n)/(N-1)$, $P_i = X_i/X$ and

$X = \sum_{1}^{N} X_i$.

Prasad and Srivenkataramana (P-S [5]) observed that Y_i/π_i will not be a constant even when $Y_i = \beta X_i$ for all i, while $(Y_i + \beta\, \alpha_1\alpha_2^{-1} X)/\pi_i$ will be so. This motivated them to consider a transformation on the study variate y, given by

$$Z_i = Y_i + \alpha_1\alpha_2^{-1} b, \qquad i = 1, 2, \ldots, N, \tag{1.3}$$

where b is a scalar. Their estimator for Y is

$$\hat{Y}_1 = \sum_{i=1}^{n} (z_i/\pi_i) - N\alpha_1\alpha_2^{-1} b. \tag{1.4}$$

It is seen that $V(\hat{Y}_1)$ is minimised by the choice

$$b_{\text{opt}} = -\Delta_2/\Delta_1, \tag{1.5}$$

where

$$\Delta_1 = \alpha_1^2\alpha_1^{-2} \sum\sum_{i>j} (\pi_i\pi_j - \pi_{ij})((1/\pi_i) - (1/\pi_j))^2$$

$$= \alpha_1^2\alpha_2^{-2}\, V\left(\sum_{i=1}^{n} 1/\pi_i\right),$$

$$\Delta_2 = \alpha_1\alpha_2^{-1} \sum\sum_{i>j} (\pi_i\pi_j - \pi_{ij})((Y_i/\pi_i) - (Y_j/\pi_j))((1/\pi_i) - (1/\pi_j))$$

$$= \alpha_1\alpha_2^{-1} \operatorname{Cov}\left(\sum_{i=1}^{n} (y_i/\pi_i),\ \sum_{i=1}^{n} (1/\pi_i)\right)$$

with

$$\pi_{ij} = \alpha_1\{(N-n)(P_i + P_j) + (n-2)\}/(N-2)$$

and

$$V_{\min}(\hat{Y}_1) = V(\hat{Y}_{HT}) - (\Delta_2^2/\Delta_1). \tag{1.6}$$

P-S's transformation involves an unknown parameter and moreover, it is on the study variate. Their transformation is guided by the observation that Y_i/π_i will not be constant even when $Y_i = \beta X_i$. Instead, we now consider a transformation on the auxiliary variate, values of which are readily available. We shall use a transformation of the type

$$X_i' = X_i + d\bar{X}$$

(cf. Reddy and Rao [8]) and determine d for which the corresponding inclusion probability π_i' is proportional to X_i. This $d = -N(n-1)/n(N-1)$.

Thus, when $Y_i = \beta X_i$, Y_i/π` is a constant for all $i = 1, 2, \ldots, N$, a property which is required by P-S [5] and is satisfied easily. Motivated by this we consider the transformation given in (1.7) with $d = -N(n-1)/n(N-1$. Now, if we take a sample of size n by Midzuno-Sen sampling scheme or by Lahiri's [3] method, the probability of inclusion of the ith unit is

$$\pi_i' = \alpha_1 + \alpha_2 P_i' = nP_i \tag{1.8}$$

where

$$P_i' = X_i'/X', \; X' = \sum_{i=1}^{N} X_i' .$$

Then, one can consider the estimator

$$\hat{Y}_{HT}' = \sum_{i=1}^{n} (y_i/\pi_i') \tag{1.9}$$

which is simpler than $\hat{Y}_1$ of (1.4) suggested by P-S which involves the choice of an unknown parameter. Sankaranarayanan [10] considered the same estimator (1.9) while constructing an Inclusion Probability Proportional to Size sampling scheme using Lahiri's method of selection. As mentioned by him, there is a minor restriction for the above procedure to be applicable in practice, viz.,

$$P'_i > 0 \;\forall\; i \tag{1.10}$$

or equivalently,

$$P_i > (n - 1)/n(N - 1). \tag{1.11}$$

However, when one uses Lahiri's method of selection, it is sufficient to have

$$\sum_{i=1}^{n} P_i > \alpha_1 = (n - 1)/(N - 1). \tag{1.12}$$

2. Comparison of Strategies

We now have the following alternative strategies :

H_1 = (Midzuno Sen sampling scheme (Lahiri's scheme), $\hat{Y}_{HT}$),

H_2 = (P-S transformation on the study variate, Midzuno-Sen (Lahiri) scheme, $\hat{Y}_1^{opt}$), and

H_3 = (Reddy-Rao type transformation on the auxiliary variate, Midzuno-Sen (Lahiri) scheme, $\hat{Y}'_{HT}$).

Empirical calculations show that for Cochran's ([1], p. 203), Rao's ([6], p. 64) and Sampford's ([9], p. 72) data sets the percent relative efficiency of H_3 compared to H_2 was 99.71, 105.04, 99.19, 98.19 and 93.71 for the values (N, n) = (10, 2), (9, 2), (12, 4) and (11, 3) respectively. Thus the suggested strategy H_3 is almost as efficient as H_2 used by P-S with optimum choice of b. For a complete description of the data sets and other details we refer to Rao [7].

REFERENCES

[1] Cochran, W. G. (1977) : *Sampling Techniques*, 3rd edition, Wiley, New York.

[2] Horvitz, D. G. and Thompson, D. J. (1952) : A generalisation of sampling without replacement from a finite universe, *J. Am. Statist. Assoc.* 47 : 663-85.

[3] Lahiri, D. B. (1951) : A method of sample selection providing unbiased ratio estimates, *Bull. Int. Statist. Inst.* 33(2) : 133-40.

[4] Midzuno, H. (1952) : On the sampling system with probability proportional to sum of sizes, *Ann. Inst. Stat. Math.* 3 : 99-107.

[5] Prasad, N. G. N. and Srivenkataramana, T. (1980) : A modification to the Horvitz-Thompson estimator under the Midzuno sampling scheme, *Biometrika* 67 : 709-11.

[6] Rao, T. J. (1967) : Contributions to the theory of sampling strategies, *Ph.D. thesis* submitted to the Indian Statistical Institute, Calcutta.

[7] Rao, T. J. (1983) : Transformation on the auxiliary variate for Midzuno-Sen sampling scheme, Tech. Report No. 8/83, I. S. I.

[8] Reddy, V. N. and Rao, T. J. (1977) : Modified PPS method of estimation, *Sankhya, C* 39 : 185-97.

[9] Samford, M. R. (1962) : *An Introduction to Sampling Theory*, Oliver and Boyd, London.

[10] Sankaranarayanan, K. (1969) : An IPPS sampling scheme using Lahiri's method of selection, *J. Ind. Soc. Agr. Stat.* **21**(2) : 58-66.

[11] Sen, A. R. (1952) : Present status of probability sampling and its use in estimation of farm characteristics (abstract), *Econometrica* **20** : 103.