ON HORVITZ AND THOMPSON'S T_1 CLASS OF LINEAR ESTIMATORS[1]

By S. G. Prabhu Ajgaonkar

Deccan College, Poona

1. Introduction. The estimators usually employed to estimate a population mean or total from sample survey data are linear functions of the observations. Horvitz and Thompson (1952) among others have discussed various subclasses of estimators, for example.

(1) Class T_1—estimators of the form $\sum \alpha_r x_r$ where the weight α_r is a function of the order of drawing.

(2) Class T_2—estimators of the form $\sum \beta_i x_i$ where the weight β_i is a function of x_i, the unit drawn.

(3) Class T_3—estimators of the form $\nu_{s_n} \sum x_r$ where the weight ν_{s_n} is a function of the sample taking into account the order of the units.

Obviously, additional classes of estimators can be obtained by associating weights that depend on a combination of these possibilities, (a fact noted by Koop (1957) in his Ph.D. thesis).

Godambe (1955) showed that there does not exist a best (minimum variance unbiased) estimator for the class encompassing the T_2 and T_3 classes.

The present paper discusses the T_1 class estimators as defined in (1) above for the case of sampling with arbitrary probabilities of selection at each draw, and examines some particular sampling designs in the light of this discussion.

2. Notation and terminology. Let X_i ($i = 1, 2, \cdots, N$) be the value of ith unit under consideration. The problem of present concern is the estimation of

$$T = \sum_{i=1}^{N} X_i \tag{2.1}$$

from a sample of n units.

Let us suppose that the sample is an ordered set—the order being that of the order of draw. Let

$$p_{ir}, \quad i = 1, 2, \cdots, N; \quad r = 1, 2, \cdots, n, \tag{2.2}$$

denote the probability of selecting the ith unit of the population at the rth draw, and

$$P = (p_{ir}) \tag{2.3}$$

be the corresponding $N \times n$ probability matrix in which the rth column denotes

Received 13 May 1964; revised 15 December 1966.

[1] This paper is an incidental publication arising out of the author's research studies on "Some Aspects of Successive Sampling," which was submitted to Karnatak University, Dharwar, India, in partial fulfillment of the requirements for the Ph.D. degree in Statistics. This work was financially supported by the University Grants Commission, Government of India.

the probabilities of selecting units at the rth draw. It may be noted that the probability matrix employed here is quite general; and that the discussion of T_1 class of this and next sections pertains to this probability matrix. Let x_r $(r = 1, 2, \cdots, n)$ stand for the value of the X-characteristic observed on a unit selected at the rth draw.

The T_1 class of estimators is unbiased if

$$E(T_1) = \sum_{r=1}^{n} \alpha_r E(x_r) = \sum_{i=1}^{N} X_i \sum_{r=1}^{n} \alpha_r p_{ir} = T,$$

which will hold for all X_i $(i = 1, 2, \cdots, N)$ if and only if

$$\sum_{r=1}^{n} \alpha_r p_{ir} = 1, \qquad i = 1, 2, \cdots, N. \tag{2.4}$$

It is well known that the above set of equations admits a consistent solution if and only if the rank of P is the same as the rank of the augmented matrix P^+ where

$$P^+ = \begin{pmatrix} p_{11} & \cdots & p_{1n} & 1 \\ \vdots & & & \vdots \\ p_{N1} & \cdots & p_{Nn} & 1 \end{pmatrix}.$$

3. Empty and non-empty classes. We present some definitions which are used in the sequel:

DEFINITION 1. Empty class: If for a given sampling system there does not exist an unbiased estimator in the class, which is independent of the population values, that class is termed empty class for that sampling scheme.

DEFINITION 2. Non-empty class: For a given sampling system a class is designated as non-empty, whenever that class contains an unbiased estimator, independent of the population values.

That there are empty classes is easily seen by the following example.

Consider the sampling system for which

$$p_{ir} = p_i \quad \text{for} \quad i = 1, 2, \cdots, N; \quad r = 1, 2, \cdots, n,$$

where there exists at least one p_i which is different from the rest. It is easily noted that the rank of the matrix P is one, while the rank of the augmented matrix is two. Therefore for this sampling procedure, the T_1 class is an empty class.

4. The variance of T_1 class of estimators. In this section we derive an expression for the variance of T_1 class estimators when the probability system described in (2.2) is employed. We have

$$\operatorname{Var}(T_1) = \sum_{r=1}^{n} \alpha_r^2 \operatorname{Var}(x_r) + 2 \sum_{s>r}^{n} \alpha_r \alpha_s \operatorname{Cov}(x_r, x_s) \tag{4.1}$$

where

$$\operatorname{Var}(x_r) = \sum_{i=1}^{N} p_{ir} X_i^2 - \left(\sum_{i=1}^{N} p_{ir} X_i\right)^2; \tag{4.2}$$

and

$$\operatorname{Cov}(x_r, x_s) = \sum_{i=1}^{N} \sum_{j=1}^{N} X_i X_j \{p(j, s/i, r) - p_{js}\} p_{ir} \tag{4.3}$$

where $p(j, s/i, r)$ denotes the conditional probability of selecting the jth unit of the population on the sth draw when the ith unit of the population is already included in the sample on the rth draw.

5. Simple random sampling without replacement. In this sampling system,

$$p_{ir} = 1/N \qquad \text{for all} \quad i, r. \tag{5.1}$$

From this it is easily noted that for this sampling system the T_1 class is a non-empty class. Since the ranks of the matrix P and the augmented matrix P^+ are one, there is only one effective equation in the coefficients α's namely

$$\sum_{r=1}^{n} \alpha_r = N. \tag{5.2}$$

To determine uniquely values of the coefficients α's, we resort to the principle of least variance. We have

$$\begin{aligned}\mathrm{Var}\,(T_1) &= N(N-1)^{-1}\sigma^2 \sum_{r=1}^{n} \alpha_r^2 - N^2(N-1)^{-1}\sigma^2 \\ &= N(N-1)^{-1}\sigma^2\{\sum_{r=1}^{n} (\alpha_r - N/n)^2 + N^2/n - N\}\end{aligned}$$

where $\mathrm{Var}\,(x_r) = \sigma^2$ for $r = 1, 2, \cdots, n$. It may be noted that $\mathrm{Var}\,(T_1)$ is minimized for the variation of α_r's subject to the condition in (5.2), when

$$\alpha_r = N/n \quad \text{for all} \quad r,$$

and for this set of values of α_r we get the well-known results:

(1) the best estimator for the T_1 class is given by

$$t_1 = (N/n) \sum_{r=1}^{n} x_r\,,$$

and (2)

$$\mathrm{Var}\,(t_1) = N^2(N-n)\{n(N-1)\}^{-1}\sigma^2.$$

6. Midzuno's scheme of sampling. In this sampling system the number of draws is always greater than one, and the units are selected with varying probabilities at the first draw, but on the subsequent draws they are selected with equal probabilities without replacement. Let the probability of selecting the ith unit of the population on the first draw be p_i for $i = 1, 2, \cdots, N$ where there is at least a pair of the p_i's differing in values, which is assumed to be $p_1 \neq p_2$ without loss of generality. Thus

$$\begin{aligned}p_{ir} &= p_i\,, & r &= 1, \\ &= (1-p_i)(N-1)^{-1}, & r &= 2, \cdots, n,\end{aligned}$$

for $i = 1, 2, \cdots, N$. It is noted that the rank of the matrix P is two, which is also the rank of the augmented matrix P^+. Therefore the T_1 class is non-empty for this sampling scheme.

The effective equations in the coefficients α's are two, namely

$$\begin{aligned}\alpha_1 &= 1, \\ \alpha_2 + \alpha_3 + \cdots + \alpha_n &= N - 1\end{aligned} \tag{6.1}$$

In order to determine uniquely the values of coefficients α_r's $(r > 2)$, we make use of the principle of minimum variance. We have

(6.2) $$\text{Var}(T_1) = \sigma_1^2 + \sigma_2^2 \sum_{r=2}^n \alpha_r^2 + 2b_1 \sum_{r=2}^n \alpha_r + b_2\{(\sum_{r=2}^n \alpha_r)^2 - \sum_{r=2}^n \alpha_r^2\}$$

where

$$\text{Var}(x_1) = \sigma_1^2;$$

$$\text{Var}(x_r) = \sigma_2^2 \quad \text{for} \quad r = 2, \cdots, n;$$

$$\text{Cov}(x_1, x_r) = b_1 \quad \text{for} \quad r = 2, 3, \cdots, n;$$

$$\text{Cov}(x_r, x_s) = b_2 \quad \text{for} \quad r \neq s = 2, 3, \cdots, n.$$

With the help of Equation (6.1),

$$\text{Var}(T_1) = \sigma_1^2 + 2(N-1)b_1 + (N-1)^2 b_2 + (\sigma_2^2 - b_2) \sum_{r=2}^n \alpha_r^2.$$

From this it is noted that the minimization of Var (T_1) subject to the condition given in (6.1), leads to the minimization of $\sum_{r=2}^n \alpha_r^2$, provided $\sigma_2^2 > b_2$, which is so, is proved at the end of this section. Accordingly, we have

$$\alpha_r = (N-1)(n-1)^{-1} \quad \text{for} \quad r = 2, \cdots, n.$$

Consequently t_1, the best estimator for the T_1 class is given by

$$t_1 = x_1 + (N-1)(n-1)^{-1} \sum_{r=2}^n x_r,$$

when Midzuno's sampling procedure is employed. Further

$$\text{Var}(t_1) = (N-n)(n-1)^{-1}\{\sum_{r=1}^N (1-p_i)X_i^2 - \sum_{i \neq j=1}^N (1-p_i-p_j)(N-2)^{-1}X_iX_j\}.$$

The inequality $\sigma_2^2 > b_2$ is noted from the fact that

$$(\sigma_2^2 - b_2) = (n-1)[(N-1)(N-n)]^{-1} \text{Var}(t_1).$$

7. Sampling with probability proportion to estimated size. This sampling system is discussed by Horvitz and Thompson (1952) on page 679 of their paper. Moreover, this has already been referred to by Narain (1951), Yates and Grundy (1953). If p_i represents the probability of drawing the ith unit of the population at the first draw for $i = 1, 2, \cdots, N$; then on the second draw the jth unit of the population is selected with probability $p_j(1 - p_i)^{-1}$ for $j \neq i = 1, 2, \cdots, N$, from the remaining $N - 1$ units. Moreover the sample comprises only two draws. Accordingly, for this sampling procedure

$$p_{i1} = p_i \quad \text{for} \quad i = 1, 2, \cdots, N,$$

$$p_{i2} = \{S - p_i(1-p_i)^{-1}\}p_i \quad \text{for} \quad i = 1, 2, \cdots, N,$$

where

$$S = \sum_{i=1}^N p_i(1-p_i)^{-1}.$$

Further, let there be at least three different p_i's.

It is noted that the rank of the matrix P is two, however the rank of the augmented matrix P^+ is three. Therefore for this sampling system the T_1 class is an empty class.

8. Acknowledgment. The author is thankful to Dr. B. D. Tikkiwal, Head of the Department of Statistics, University of Rajasthan, Jaipur, for suggesting this problem for his research and the necessary guidance. Further he remains deeply thankful to Dr. J. C. Koop for sending his publications. He is also indebted to the referee and the Associate Editor, who have offered many useful suggestions. To Dr. S. M. Katre, Director, the author owes his thanks for providing many facilities to carry on the research at the Deccan College Post-graduate and Research Institute.

REFERENCES

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

[2] HORVITZ, D. G. and THOMPSON, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* **47** 663–685.

[3] KOOP, J. C. (1957). Contributions to the general theory of sampling finite populations without replacement and with unequal probabilities. Ph.D. thesis. N. Carolina State College Library, Raleigh.

[4] KOOP, J. C. (1963). On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory for finite populations, Part I, II and III. *Metrika* **7** 81–114 and 165–204.

[5] MIDZUNO, H. (1952). On the sampling system with probability proportionate to sum of sizes. *Ann. Inst. Statist. Math. Tokyo* **3** 99–107.

[6] NARAIN, R. D. (1951). On sampling without replacement with varying probabilities. *J. Indian. Soc. Agric. Statist.* **3** 169–174.

[7] PRABHU AJGAONKAR, S. G. (1962). Some aspects of successive sampling. Ph.D. thesis submitted to Karnatak University, Dharwar, India.

[8] SUKHATME, P. V. (1953). *Sampling Theory of Surveys with Applications.* The Indian Society of Agricultural Statistics, New Delhi, India and the Iowa State College Press.

[9] YATES, F. and GRUNDY, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *J. Roy. Statist. Soc. B* **15** 253–261.