ON T_2 AND T_3 – CLASSES OF ESTIMATORS IN SAMPLING WITH VARYING PROBABILITIES AND WITHOUT REPLACEMENT

BY B.D. TIKKIWAL
University of Rajasthan, Jaipur

0. INTRODUCTION AND SUMMARY

Horvitz and Thompson (1952), while considering sampling with varying probabilities and without replacement, to estimate the population total T, gave three classes of linear estimators and mentioned that these are not exhaustive classes of possible linear estimators.

Prabhu Ajgaonkar and Tikkiwal (1961) and Prabhu Ajgaonkar (1962) examined the theory of sampling with varying probabilities and with or without replacement further and noted that there are in all seven classes of linear estimators. This point was also noted independently by Koop (1961, 63). Prabhu Ajgaonkar and Tikkiwal showed that an unbiased estimator dependent on population values always exists in a given class. There is need, however, to find either an unbiased estimator independent of population values or an unbiased estimator dependent on such population values which are either known in advance or which can be easily estimated.

The two authors approach the problem in the following way. While considering a given T-class; in its sub-class of unbiased estimators they examine whether the equations, obtained through condition of unbiasedness restricting the estimators to be independent of population values, are consistent. If the equations are not consistent then the class of unbiased estimators independent of population values is an empty class to start with. But if the equations are consistent, then the authors proceed further to find a minimum-variance-linear-unbiased estimator in the class of unbiased estimators. Such a minimum-variance-linear-unbiased estimator always exists in a given class. But the estimator either depends on the population values or it does not. When it depends on the population values, it is called a dependent estimator otherwise it is called an independent estimator.

The two authors made a detailed study of Horvitz and Thompson's three classes of linear estimators and the linear estimator in T_4-class where the weight associated to an element of the population in the sample depends upon the element itself and also on the draw at which this element occurs in the sample. It was shown by them that T_4-class, in general, has linear unbiased estimators independent of population values and that, when the probability system satisfies a certain condition, the independent minimum-variance-linear-unbiased estimator in T_4-class lies in Horvitz and Thompson's T_2-class. For some well known sampling schemes, the minimum-variance-linear-unbiased estimators or simply the linear unbiased estimators in T_1-class either did not exist or were obviously inefficient as compared to the corresponding estimators in T_2-class. The authors gave estimators in T_2-class but did not compare their efficiency with estimators in T_2-class.

This paper presents inter and intra class comparisons for known estimators in T_2 and T_3-classes for estimating different population totals with different sampling schemes and different probability sets. It is noted that the relative efficiency of an estimator in T_3-class as compared to that of an estimator in T_2-class for the same sampling scheme depends upon the population and the probability set used. It is further noted that an estimator based on simple random sampling can be more efficient than an estimator based on sampling with varying probabilities.) This points to the need of exercising necessary caution in the use of sampling with varying probabilities. Further, when we do use sampling with varying probabilities, we should not always use T_2-class estimators as per normal practice. It is shown that Stuart's method (1954) for obtaining optimal sampling results does not always give optimal results.)

1. The Estimators in T_2 and T_3-classes

Let there be a sample of size n drawn with varying probabilities and without replacement out of N units in the population to estimate the population total $T = \sum_{i=1}^{N} x_1$. Let a linear estimator in T_3-class be denoted by

$$\hat{T}_3 = \gamma_{sn} \left(\sum_{r=1}^{n} x_r \right)_{sn} \tag{1·1}$$

where γ_{s_n} is the weight associated with the total of s_n th sample for $s_n = 1, 2, \ldots, NC_n$ and x_r denotes the out-come at the rth draw for $r = 1, 2, \ldots n$.

It may be noted that we do not take into account the ordering of units in a given sample of n units as the minimum-variance-linear unbiased estimator is independent of the order [Murthy (1957), Prabhu Ajgaonkar and Tikkiwal (1961,[62)].

It is shown by Prabhu Ajgaonkar and Tikkiwal (1961, 62) that there is no minimum-variance-linear-unbiased estimator in T_3-class which is independent of population values even for simple random sampling with or without replacement. For the latter scheme, they further showed that the classical estimator $\hat{T}_c = N(\sum_{r=1}^{n} x_r)/n$ is the minimum-variance-linear-unbiased estimator in a sub-class of T-$_3$ class where the weight associated with a sample depends upon the outcome at the first draw. It is known that $\hat{T}_c$ is minimum-variance-linear-unbiased estimator in T_1 and T_2 classes.

It is of interest to get some idea of the loss of efficiency in using $\hat{T}_c$ instead of the minimum-variance-linear-unbiased estimator in T_3-class. For this we present for $N=4$, $n=2$ the results given by Prabhu Ajgaonkar and Tikkiwal [1961, 62].

Let w_{ij} be the weight to be associated with a sample containing the units x_i and x_j. Prabhu Ajgaonkar has shown that the weights for the minimum-variance-linear-unbiased estimator $(\hat{T}_3)_{min}$ in T_3-class assume the following values :

$$W_{12}=W_{34}=\frac{\lambda}{(x_1+x_2)^2+(x_3+x_4)^2};$$

(1·2) $$W_{13}=W_{24}=\frac{\lambda}{(x_1+x_3)^2+(x_2+x_4)^2};$$

and

$$W_{14}=W_{23}=\frac{\lambda}{(x_1+x_4)^2+(x_2+x_3)^2};$$

where

$$\lambda=6\left[\frac{1}{(x_1+x_2)^2+(x_s+x_4)^2}+\frac{1}{(x_1+x_3)^2+(x_2+x_4)}+\frac{1}{(x_1+x_4)^2+(x_2+x_3)^2}\right]^{-1}.$$

The variance of this estimator is given by

(1·3) $$(\lambda-T^2).$$

This is clear from the above discussion that this estimator is at least as efficient as the classical estimator $\hat{T}_c$. This is also so if

$$Q=\frac{\lambda}{3}\left[3-2\left(\frac{1}{c_1}+\frac{1}{c_2}+\frac{1}{c_3}\right)\right]\leqslant 0\,;$$

where $c_1+c_2+c_3=6$. The above inequality is seen to be true by noting that Q assumes maximum value for $c_1=c_2=c_3$.

Let a sample of size $n=2$ be drawn from each of following pseudo four populations of size 4 with simple random sampling without replacement

TABLE 1

	Pop. I	*Pop. II*	*Pop. III*	*Pop. IV*
x_1	−5	−10	−400	5
x_2	−8	− 8	−300	8
x_3	7	14	299	5
x_4	4	8	399	4

The variances of the estimators $(\hat{T}_3)_{min}$ and $\hat{T}_c$ for the above four populations are given below in order to show the degree of efficiency of the former estimator over the latter.

TABLE 2

	Pop. I	*Pop. II*	*Pop. III*	*Pop. IV*
Var $(\hat{T}_3)_{min}$	6·84	51·70	7·99	11·93
Var $(\hat{T}_c)$	204·00	560·00	664801.33	12·00

We now present an estimator in T_3-class for the general situation. In order that $\hat{T}_3$ in (1·1) be an unbiased estimator of T, we have

$$(1{\cdot}4)\qquad \sum_{sn\,(i)}^{N-1\,C_{n-1}} \gamma_{sn}(i)\, P_{S_n}(i)=1;\ i=1,2,\ldots,N$$

where $s_n^{(i)}$ denote the sample in which the i_{th} unit occurs in the sample and $P_{s_n}^{(i)}$ denote the probability of occurrence of the sample $s_n^{(i)}$. Let

$$\sum_{s_n} \Psi (P_{s_n}) F(x_1, x_2, \ldots, x_n) = k$$

where k is some constant quantity.

Let

$$(1{\cdot}5) \qquad \phi = \left[\text{Var} (\hat{T}_3) + T^2 \right] \times \sum_{s_n} \Psi (P_{s_n}) F(x_1, x_2, \ldots, x_n).$$

By Cauchy inequality,

$$(1{\cdot}6) \qquad \phi \geqslant \left[\sum_{s_n} \gamma s_n (\Sigma x_r) s_n (P_{s_n})^{1/2} . \{\Psi (P_{s_n}) F(x_1, x_2, \ldots, x_n)\}^{1/2} \right]^2.$$

The equality is achieved, if

$$(1{\cdot}7) \qquad \frac{\gamma s_n (P_{s_n})^{1/2} (\Sigma x_r) s_n}{[\Psi (P_{s_n}) F(x_1, x_2, \ldots x_n)]^{1/2}} = \lambda$$

where λ is constant quantity. It may be noted that the values of γs_n obtained by Equation (1·7) do not give the minimum value of ϕ as the right hand side of Equation (1·6) is not independent of γs_n as was true of all the cases considered by Stuart (1954) where he used Cauchy inequality to obtain optimum sampling results. Thus, we note that the use of Cauchy inequality does not always give optimum sampling results. Further, the Cauchy inequality cannot be used to obtain optimum sampling results, when the variance function is to be minimised subject to more than one restriction is clear from the fact that Lagrange method of undetermined multiples is analogous to the method using Cauchy inequality when there is only one restriction. This is clear from the following proof showing that the two methods are analogus when there is one restriction. Let $f(x_1, x_2, \ldots, x_n)$, the function of n variables be minimised subject to the restriction $g(x_1, x_2, \ldots, x_n) = k$. Then, from Lagrange method, we have to minimise the function

$$(1{\cdot}8) \qquad \phi = f(x_1, x_2, \ldots, x_n) + c/g(x_1, x_2, \ldots, x_n)$$

where c is some constant. This is same as minimising.

$$(1{\cdot}9) \qquad g(x_1, x_2, \ldots, x_n)\ \phi - c = f(x_1, x_2, \ldots x_n) . g(x_1, x_2, \ldots, x_n).$$

Iu order that Equations (1·4) and (1·7) hold for all probability systems, it is suffice to have

(1·10)
$$\begin{cases} \Psi(P_{s_n}) = 1/P_{s_n} \\ \text{and} \\ F(x_1, x_2, \ldots, x_n) = (\sum_r x_r)^2 s_n. \end{cases}$$

For simple random sampling without replacement it will mean that

(1·11)
$$\sum_{s_n} (\sum_r x_r)^2 s_n = k'$$

where k' is some other constant. By (1·6) and (1·7)

$$V_{s_n}(P_{s_n}) = \lambda.$$

Therefore, from Equation (1·4)

(1·12)
$$\gamma_{s_n} = \frac{1}{\binom{N-1}{n-1} P_{s_n}}, \quad s_n = 1, 2, \ldots, \binom{N}{n}.$$

Therefore,

(1·13)
$$\hat{T}_3' = \left[\left(\sum_{r=1}^{n} x_r\right) s_n\right] \Big/ \binom{N-1}{n-1} P_{s_n}$$

and

(1·14)
$$\text{Var}(\hat{T}'_3) = \left[1 \Big/ \binom{N-1}{n-1}\right]^2 \left[\sum_{s_n} (\sum x_r)^2 s_n / P_{s_n}\right] - T^2.$$

The above estimator and its variance was given by Des Raj (1954, Lemma 2, pp. 131) while obtaining an unbiased ratio estimator. An unbiased estimator of the variance can be obtained by his Lemma. 1. Des Raj stated in his Lemma 2 that if from a finite population of size N, a sample s_n of size n be selected with probability P_{s_n}, the only unbiased estimator, of the total of the population, of the form in (1·1) is $\hat{T}'_3$.

Since Lemma 2 is given without proof it is not clear how Des Raj obtained the estimator and how above statement was made. But it will be seen that the statement is not correct in view of the estimator given in (1·15) which is also an unbiased estimator of the form given in (1·1).

When the unit at the first draw is drawn with varying probabilities and at the subsequent draws it is drawn with equal probabilities without replacement as in Midzuno's scheme of sampling, the minimum-variance-linear-unbiased estimator of T, as given by Prabhu Ajgaonkar and Tikkiwal [1961, 62], in a sub-class of T_3-class where the weight associated with a sample depends upon the out come at the first draw is

$$(1{\cdot}15) \qquad \hat{T}''_3 = \frac{1}{np_{i_1}} \left(\sum_{r=1}^{n} x_{i_r} \right);$$

where x_{i_r} denote the value of i_rth unit drawn at the rth draw. It may be noted that $\hat{T}''_3$ also reduce to $\hat{T}_c$ for simple random sampling without replacement. This was the basis of Prabhu Ajgaonkar and Tikkiwal's statement that $\hat{T}_c$ is minimum-variance-linear-unbiased estimator in a sub-class of T_3-class as stated in the beginning of this section.

The variance of this estimator as derived by Parabhu Ajgaonkar and Tikkiwal is

$$(1{\cdot}16) \qquad \sum_{i=1}^{N} \frac{n-1}{n^2(N-1)} \left\{ s + \frac{N-n}{(n-1)\,p_i} \right\} x_i^2 \Big/ \frac{N-n}{(n-1)\,p_i}$$

$$+ \sum_{i \neq j=1}^{N} \frac{n1}{n^2(N-1)} \left\{ \frac{n-2}{N-2} s + \frac{N-n}{N-2} \left(\frac{1}{p_i} + \frac{1}{p_j} \right) \right\} x_i\, x_j - T^2$$

where $\qquad s = \sum_{i=1}^{N} (1/p_i).$

The variances of Horvitz and Thomson's estimators in T_2-class, $(\hat{T}_2)_{HT}$ and $(\hat{T}_2)_{M}$; corresponding to the sampling scheme 2 of these two authors and corresponding to Midzuno's sampling scheme given by these authors as samling scheme 1, can be easily worked out. In the following table, we present the variances of $(\hat{T}_2)_{HT}$, $(\hat{T}_2)_{M}$, $\hat{T}_3''$ and of the estimators $(\hat{T}_3')_{HT}$ and $(\hat{T}_3')_{M}$ obtained from Equation (1·14) for above two sampling schemes for the first three populations with probability sets (.3, .3, .2, .2), (.1, .2, .3, .4) and (.15, .25, .40, .20) for $n=2$.

TABLE III

	Pop. I	*Pop. II*	*Pop. III*
Var $(\hat{T}_2)_{HT}$	208·54	747·02	836609 08
Var $(\hat{T}_2)_M$	201·18	566·82	721208·56
Var $(\hat{T}_3')_{HT}$	218·68	943·36	825547·31
Var $\hat{T}_3')_M$	205·39	614·48	692442·14
Var $(\hat{T}_3'')$	206·00	671·22	755197·00

From tables II and III we note that Prabhu Ajgaonkar and Tikkiwal's minimum-variance-linear-unbiased estimate in T_3-class for simple random sampling with weights as in (1·2) is the most efficient estimator among all estimators in T_2 and T_3 classes given in the two tables for different sampling schemes. However, the estimator suffers from the fact that the weights depend on population values and therefore have to be estimated from the sample.

If we know about the population only through a sample of size two, we may assume $(x_i+x_j)^2$ to be the same for all (i, j) pairs. Such an assumption in (1·2) will make W_{ij} to be the same for all i, j and the estimator will then reduce to $\hat{T}_c$. In case the information on an auxiliary variable is available, the weights can be estimated with the help of the auxiliary information but the procedure would appear to be complicated for large N. This matter needs further examination.

We note that $\hat{T}_c$ can be obtained from $(\hat{T}_3')$ in (1·13) by putting $P_s=1/NC_n$ for simple random sampling without replacement. Since $\hat{T}_3'$ is obtained by using Cauchy inequality as in Stuart method for obtaining optimal sampling results, $\hat{T}_c$ can be said to be obtained by Stuart method. But $\hat{T}_c$ is not a minimum-variance-linear-unbiased estimator in T_3-class is clear from Table II and from discussions preceding Table I. This shows that Stuart's method for obtaining optimal sampling results does not always give optimal results as noted earlier.

When there is sampling with varying probabilities and without replacement, we normally use T_2-class estimator. In the above empirical study, the first three populations are arranged in increasing order of their ranges of variation. The T_2-class estimators for given sampling schemes are more efficient than the estimators in T_3-class for first two populations but less for the third population with the highest range of variation. Further, the relative efficiency of T_2-class estimators over T_3-class estimators is more for population II than for population I. Thus, the relative efficiency depends on population-range and therefore on population as the probability set used.

In Tables II and III we have presented estimators with three different sampling schemes, (*i*) Simple random sampling, (*ii*) Midzuno's scheme of sampling, (*iii*) Hor vitz and Thompson's seheme of sampling. The three populations have correlations —·97, ·83 and. ·49 with their respective probability sets for Table III. As judged by the variances of all the estimators excluding those of $(\hat{T}_3)_{min}$ and of $\hat{T}_3''$; the first sampling scheme is better than the second, which in turn, is better than the third for estimating totals of populations II and III in T_2 and T_3-classes and for estimating total of population one in T_3-class. For cstimating total of population 1 in T_2-class, the second sampling scheme is better than the first, which in turn, is better than the third. Even for estimating total of population I in T_2-class, there is hardly any difference in the efficiency of the two estimators with the first two sampling schemes.

In the literature on varying probabilities beginning from the work of Hansen and Hurvitz (1943) to up-to-date the author has not come across anywhere in any empirical study where the estimator based on simple random sampling is more efficient than the estimators based on varying probabilities. The present study indicates of such a possibility. Hence there is need of exercising necessary caution in the use of varying probabilities.

It may be noted that the estimator $(\hat{T}')_M$ is always more efficient than Prabhu Ajgaonkar and Tikkiwal's estimator $\hat{T}_3''$. This result is true in general is seen by noting that $(\hat{T}_3')_M$ is an unordered estimator of $(\hat{T}_3'')$ as shown by Murthi (1957, Section 4, pp. 385). We present here a direct proof of this result.

Since $\hat{T}_3'$ and $\hat{T}_3''$ are unbiased estimator of the same quantity T, it is enough for our purpose if we compare only $E(\hat{T}_3')^2$ and $E(\hat{T}_3'')^2$. Let $(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$ denote a sample s_n for $s_n=1,\ldots 2,\ldots, \binom{N}{n}$. Then the contribution of this sample to $E(\hat{T}_3'')^2$ is

$$\frac{1}{\binom{N-1}{n-1}} \sum_{r=1}^{n} \left\{ \frac{1}{n^2 p_{i_r}} \left\{ \sum_{r=1}^{n} x_{i_r} \right\}^2 \right\}.$$

Its contribution to $E(\hat{T}_3')^2$ is

$$\frac{1}{\binom{N-1}{n-1}} \frac{1}{\sum_{r=1}^{n} p_{i_r}} \left\{ \sum_{r=1}^{n} x_{i_r} \right\}^2.$$

We will have $E(\hat{T}_3'')^2 \geqslant E(\hat{T}_3')^2$, if

$$(\Sigma\, 1/p_{i_r}).\,(\Sigma\, p_{i_r}) \geqslant n^2,$$

which is true by Cauchy's inequality.

Acknowledgements

As stated in Section I, Dr. S.G. Prabhu Ajgaonkar and the author have shown that there is no minimum-variance-linear-unbiased estimator in T_3-class which is independent of population values even for simple random sampling with or without replacement. This result is in contrast with the following statement made by Horvitz and Thomson on pp. 669 of their paper (1952) :

'In connection with this remark, although it can be easily shown, when sampling with equal probabilities of selection for each draw, that the α's, β's and γ's are all equal to N/*n* for the best linear unbiased estimator of *T* for each of the three sub-classes...'

Though no proof of this statement relevant to T_3-class is given by the authors in their paper, one of the authors, Dr. D.G. Horvitz at present at the Research Triangle Institute, North Carolina, was kind enough to provide the proof of the statement for T_3-class on reference of this matter to him. Dr. Horvitz's proof is on the same lines as that for obtaining the estimator $\hat{T}_3'$ in this paper if we replace k by k' given in (1·11) of this paper. Thus, his method is based on Stuart's method. The use of function k for the general investigation was suggested by Dr. G.R. Seth of the Institute of Agricultural Research Statistics, New Delhi. The various calculations in this paper were done and checked by the Research Scholars of the Department.

REFERENCES

(1) Des Raj (1954) Ratio estimation in sampling with equal and unequal probabilities. Jour. Ind. Soc. Agric. Stat., Vol. **6**, 127-38.

(2) Hansen, M.H. and Hurvitz, W.N. (1943). On the theory of sampling from finite populations. Ann. Math. Statist. Vol. **15**, 333-362.

(3) Horvitz, D.G. and Thomson, D.J. (1952). A generalisation of sampling without replacement from a finite population Jour. Amer. Statist. Assoc. Vol. **47**, 663-685.

(4) Koop, J.C. (1961) Contributions to the general theory of sampling finite populations without replacement and with unequal probabilities. Memeo. Series No. 296, Institute of Statistics North Carolina, U.S A., Ph. D. thesis 1957, North Carolina State College.

(5) ————(1961) On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory for finite populations. Ann. Math. Statistics, Vol. **32**, 1349 (Abstract).

(6) ————(1963) *Ibid.* Metrika 7 (2 and 3) : 8-114, 165-204.

(7) Murthy, M.N. (1957) Ordered and unordered estimators in sampling without replacement. Sankhya, Vol. **18**, 379-390.

(8) Prabhu Ajgaonkar, S.G. (1962) Some aspects of the class of linear estimators with special reference to successive sampling. Ph. D. thesis, Karnatak University, Dharwar, India.

(9) Prabhu Ajgaonkar, S.G. and Tikkiwal, B.D. (1961) Horvitz and Thomson's T-class estimators. Ann. Math. Statist. Vol. **32**, 923 (Abstract).

(10) Stuart, A. (1954) A simple presentation of optimum, sampling results, Jour. Roy. Statist. Society, Series B, Vol. **16**, 239-241.