Aligarh Journal of Statistics
Vol. 1, (1981) No. 1, 52–58

SOME T_2—CLASS OF ESTIMATORS BETTER THAN H–T ESTIMATOR

PULAKESH MAITI AND T.P. TRIPATHI

The T_2-class of linear estimator [Horvitz and Thompson (1952), Koop (1963)] for Y, the population total of a character y, in case of general sampling designs, is revisited and a subclass of biased estimators from T_2 better than H–T estimator $\hat{Y}_{H-T}$ is identified. It is found that, in case of a class of sampling designs, we may always generate estimators better (in the sense of having smaller mean square error) than $\hat{Y}_{H-T}$. We also study another biased subclass of estimator $T_2^* = \lambda \sum_{i\varepsilon s} y_i/p_i$ where $p_i = x_i/\sum_1^N x_i$, i= 1,2, ...N, x being an auxiliary character, and λ is a suitably chosen constant. Some members from T_2^* are shown to be better than $\hat{Y}_{H-T}$, under a super-population model.

1. Introduction—Let $U = \{1,2 \ldots ,N\}$ be a finite population of N (given) units and y be a variate under study which takes value y_i for the i-th unit of the population (i=1,2, . . . ,N). Let

$$\bar{Y} = \sum_1^N y_i/N, \quad \sigma_y^2 = \sum_{i=1}^N (y_i - \bar{Y})^2/N \text{ and } C_y = \sigma_y/\bar{Y} \text{ be the mean,}$$

the variance and the coefficient of variation respectively of the character y in the population.

The T_2–class of linear estimators [Horvitz and Thompson (1952); Koop (1963)] for population total $Y = N\bar{Y}$ based on any sampling design (not necessarily of fixed sample size) is defined by

$$T_2 = \sum_{i\varepsilon s} \beta_i y_i \tag{1.1}$$

where β_i (i=1,2, . . . ,N) is the coefficient attached with a specified unit i of the population. It may be shown that its mean square error is given by

$$M(T_2)=Y^2[\beta'A\beta-2\beta'd+1] \tag{1.2}$$

where,

$$\beta=(\beta_1, \beta_2, \ldots, \beta_N)', \quad A=(a_{ij})_{N\times N}, \quad d=(d_1, d_2, \ldots, d_N)'$$
$$a_{ij}=(y_i/Y)(y_j/Y)\,\pi_{ij}, \quad d_i=(y_i/Y)\,\pi_i, \quad i, j=1,2, \ldots, N$$

and π_{ij} for $j=i$ is interpreted as π_i.

It is known, for general sampling designs, that the Horvitz-Thompson (H-T) estimator,

$$\hat{Y}_{H\text{-}T}=\sum_{i\varepsilon s} y_i/\pi_i \tag{1.3}$$

with

$$V(\hat{Y}_{H\text{-}T})=\sum_{i=1}^{N}\frac{(1-\pi_i)}{\pi_i}y_i^2+\sum_{i\neq j=1}^{N}\sum\frac{(\pi_{ij}-\pi_i\pi_j)}{\pi_i\pi_j}y_iy_j \tag{1.4}$$

is the best (in the sense of having smallest variance) estimator in the unbiased sub-class of T_2 where π_i and π_{ij} are the first and second order inclusion probabilities respectively.

Here in this paper, an effort is made to search for some estimators in T_2 better (in the sense of having smaller mean square error) than H-T estimator.

2. Existence of Estimators in T_2 better than $\hat{Y}_{H\text{-}T}$—From (1.2), it may be shown that optimum choice β_0 of β which minimises $M(T_2)$ is a solution of

$$A\beta_0=d \tag{2.1}$$

and (the resulting) optimum mean square error is found to be

$$M_0(T_2)=Y^2[1-d'A^{-1}d] \tag{2.2}$$

Let $\nu(s)$ be the number of distinct units in a sample s (effective sample size) and $\nu=E\nu(s)=\sum_{s\varepsilon S}\nu(s)P_s$ be the average effective sample size.

It may be shown that a particular solution of (2.1) for a design p, for which $s\varepsilon S$, $p(s)>0 \Rightarrow \nu(s)=\nu$ is $\beta_0=(\beta_{01}, \beta_{02}, \ldots, \beta_{0N})$ where

$$\beta_{0i}=(Y/\nu\, y_i), \quad i=1, \ldots, N \tag{2.3}$$

Clearly this optimum choice of β_i reduces T_2 to Y itself. We note that the UMMSE (uniformly minimum mean square error) estimator for Y does not exist at all in T_2–class of linear estimators.

REMARK : It may be shown that for simple random sampling without replacement (SRSWOR), optimum choice of β_i is unique and is found to be

$$\beta_{oi} = (Y/ny_i) \tag{2.4}$$

where n is the sample size.

Since the best (UMMSE) estimator in T_2–class does not exist at all we look for some other estimators which may be better than $\hat{Y}_{H\text{-}T}$.

Let the weights β_i in T_2 be chosen such that $\beta_i = \alpha/\pi_i$ which reduces T_2 into

$$T_2' = \alpha \sum_{i \in S} y_i/\pi_i \tag{2.5}$$

where α is a constant.

Let

$$D = \sum_{i=1}^{N} \sum_{j=1}^{N} (\pi_{ij}/\pi_i \pi_j)(y_i/y)(y_j/y)$$

$$a_1' = \min_{1 \leqslant i \leqslant N} \left\{ 1/\pi_i \right\} \tag{2.6}$$

$$\text{and } a_2' = \min_{1 \leqslant i \neq j \leqslant N} \left\{ \pi_{ij}/\pi_i \pi_j \right\}$$

It may be noted that

$$D = \left[V\left(\hat{Y}_{H\text{-}T}\right)/Y^2 \right] + 1 > 1 \tag{2.7}$$

Let

$$D_1 = \left[\left(1 + C_1^2\right)/N\right]\left(a_1' - a_2'\right) + a_2'$$

where C_1^2 is any known quantity such that

$$\left[N\left(1 - a_2'\right)/\left(a_1' - a_2'\right)\right] - 1 < C_1^2 < C_y^2 \tag{2.8}$$

where C_y^2 is the square of coefficient of variation of y.

Now we prove the following

THEOREM 1 : The estimators in T_2', with α satisfying $((2/D_1)-1) \leqslant \alpha < 1$, will always be better than $\hat{Y}_{H\text{-}T}$.

PROOF : From (1.2) and (1.4), it may be shown that

$$M(T_2) \leqslant V(\hat{Y}_{H\text{-}T})$$

$$\text{iff} \quad \sum_{i=1}^{N}\sum_{j=1}^{N} (y_i/Y)(y_j/Y)\,\pi_{ij}\left(\beta_i\beta_j - \frac{1}{\pi_i\pi_j}\right) \leqslant 2\sum_{i=1}^{N}(y_i/Y)\left(\beta_i - \frac{1}{\pi_i}\right)\pi_i. \tag{2.9}$$

Hence from (2.5) and (1.4)

$$M(T_2') \leqslant V(\hat{Y}_{H\text{-}T})$$

$$\text{iff} \qquad (1-\alpha)\,[2-(1+\alpha)\,D] \leqslant 0 \tag{2.10}$$

Noting that $D > 1$, (2.10) can never be satisfied for $\alpha > 1$. Further observing that $1 < D_1 < D$, a sufficient condition that (2.10) is satisfied would be

$$[\,(2/D_1) - 1] < \alpha < 1.$$

This proves the theorem.

REMARKS : (i) From (2.10), it may also be shown that T_2' will be better than $\hat{Y}_{H\text{-}T}$ if α lies between $(2/D)-1$ and 1, the best choice of α being $1/D$. Further none of the estimators in T_2' would be better than $\hat{Y}_{H\text{-}T}$ if $D=1$.

(ii) It may be shown that in case of SRSWOR, the best choice of α in T_2' would be $\alpha_0=1/[1+KC_y^2]$, where $K=(N-n)/n(N-1)$. Thus, in this case, if C_y^2 is known exactly, the estimator due to Searls (1964) would be the best in T_2'. As shown by Maiti and Tripathi [(1979), (1980)], estimators better than $\hat{Y}=N\bar{y}$ for Y, $\bar{y}$ being the sample mean, may be generated even if C_y is not known exactly.

3. Some other biased Estimators in T_2 better than H—T Estimator— Let x be an auxiliary character (which may be some suitable real valued function of some other variate, say z), closely associated with the main character y. Getting motivation from (2.3), we set up $\beta_i = \lambda / p_i, p_i = x_i / X$ $X = \sum_{1}^{N} x_i$ giving a subclass of T_2 as

$$T_2^* = \lambda \sum_{i \varepsilon s} y_i / p_i \qquad (3.1)$$

where λ is a suitably chosen constant.

We shall study the performance of T_2^* compared to $\hat{Y}_{H\text{-}T}$. for fixed sample size designs (of size n) and under a super population model [Hanurav, (1966)] $\mathcal{E}$, specified by

$$\mathcal{E}(y_i) = a x_i$$

$$V(y_i) = \sigma^2 x_i^g$$

$$\mathrm{Cov}(y_i, y_j) = 0 \qquad (3.2)$$

where $\mathcal{E}$, V and Cov denote the expected value, variance and covariance respectively for a given vector value of the auxiliary variable x, g is the super-population parameter and a and σ^2 are constants. It is pointed out that under (3.2), for g=2, and $\Sigma \pi_i p_i \leq 1$

$$\mathcal{E} M(T_2^*) \leq X^2 (\sigma^2 + n a^2) \left(n\lambda^2 - 2\lambda \sum_{i=1}^{N} \pi_i p_i \right) + X^2 \left(\sigma^2 \sum_{1}^{N} p_i^2 + a^2 \right) \qquad (3.3)$$

The value of λ which minimises $\mathcal{E} M(T_2^*)$ is given by

$$0 < \lambda_0 = \sum_{i=1}^{N} \pi_i p_i / n + \frac{(1 - \Sigma \pi_i p_i)}{n + \sigma^2 / a^2} < 1 \qquad (3.4)$$

and the (resulting) optimum expected mean square error would be

$$\mathcal{E} M_0 (T_2^*) \leq X^2 \left[-(\sigma^2 + n a^2) \frac{1}{n} \left(\sum_{1}^{N} p_i \pi_i \right)^2 + \sigma^2 \sum_{1}^{N} p_i^2 + a^2 \right] \qquad (3.5)$$

Further under (3.2), for g=2,

$$\mathcal{E} V(\hat{Y}_{H\text{-}T}) = X^2 \left[a^2 . t + \sigma^2 \sum_{1}^{N} p_i^2 (1 - \pi_i) / \pi_i \right] \qquad (3.6)$$

where $t = \overset{*}{V}(\hat{X}_{H-T})/X^2$, $\overset{*}{V}(\hat{X}_{H-T})$ being given by (1.4) with y_i's replaced by x_i's.

Next, we prove the following:

THEOREM 2. Under the model (3.2) with $g=2$, and $\sum \pi_i p_i \le 1$ a sufficient condition for T_2^* to be $\mathcal{E}$—better (in the sense of having smaller expected mean square error) than $\hat{Y}_{H-T}$ is $\pi_i \leqslant 1/2$ and $t \geqslant 1-\left(\sum_1^N \pi_i p_i\right)^2$.

PROOF: From (3.5) and (3.6).

$$\mathcal{E}M(T_2^*) - V(\hat{Y}_{H-T}) \leq X^2\left[-(\sigma^2+na^2)\frac{1}{n}\left(\sum_1^N \pi_i p_i\right)^2 - \sigma^2\sum_1^N p_i^2 (1-2\pi_i)/\pi_i + a^2(1-t)\right]$$

Thus $\mathcal{E}_\sigma M(T_2^*) \leqslant \mathcal{E}V(\hat{Y}_{H-T})$

$$\text{iff } (\sigma^2/a^2) \geqslant \left[1-t-\left(\sum_1^N \pi_i p_i\right)^2\right] \Bigg/ \left[\left(\sum_1^N p_i \pi_i\right)^2\Big/ n + \sum_1^N p_i^2 (1-2\pi_i)/\pi_i\right] \qquad (3.8)$$

and the result follows.

REMARKS: (i) In case of SRSWOR, (3.8) reduces to

$$\sigma^2/a^2 \geqslant \left[1-KC_x^2 - f^2\right] \Bigg/ \left[f^2/n + \frac{(1-2f)}{f}\sum_1^N p_i^2\right]$$

where $f = n/N$ is the sampling fraction and $K = (N-n)/n(N-1)$. Obviously for $f \leqslant \frac{1}{2}$ and $C_x^2 \geqslant (N-1)f(1+f)$, T_2^* would be $\mathcal{E}$—better than $\hat{Y} = N\bar{y}$.

(ii) In case of π—P—S scheme where $\pi_i = np_i$, condition (3.8) reduces to

$$\sigma^2/a^2 \geqslant \left[n\left(1+n \sum_{i=1}^{N} p_i^2\right)\left(1-n \sum_{i=1}^{N} p_i^2\right)\right] \Big/ \left(n \sum_{i=1}^{N} p_i^2 - 1\right)^2$$

which always holds in case

$$n > 1/\sum_{i=1}^{N} p_i^2 .$$

REFERENCES

Hanurav, T.V. (1966). Some aspects of unified sampling theory, *Sankhyā, Ser. A*, 175-204.

Horvitz, D.G. and Thompson D J. (1952). A generalisation of sampling without replacement from a finite universe, *J. Amer. Statist Assoc.* 47, 663-685.

Koop, J.C. (1963). On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory for finite universes-I and II, *Metrika*, 7, 81-114, 165-204.

Maiti, P. and Tripathi, T.P. (1979), Use of prior information on some parameters in estimating population mean. *Tech Report No. 23/79, Stat-Math. Division, I.S.I., Calcutta.*

Maiti, P. and Tripathi, T.P. (1980) Some results on T1-class of linear estimators. *Tech. Report No. 25/80, Stat-Math. Division, I.S.I., Calcutta.*

Searls, D.T. (1964). The utilisation of a known co-efficient of variation in the estimation procedure. *J. Amer. Statist. Assoc.* 59, 1225-1226.