The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 680–690

HYPERADMISSIBILITY OF ESTIMATORS FOR FINITE POPULATIONS

BY V. M. JOSHI

Secretary, Maharashtra Government, Bombay

1. Introduction. This note deals with a point arising from the recently published paper of Hanurav (1968). Defining the notion of hyperadmissibility for estimators for finite populations, Hanurav proves that (i) if the sampling design is a non-unicluster design, then the Horvitz Thomson estimator (H-T estimator for short) for the population total is the unique unbiased and hyperadmissible estimator, in the class of all polynomial estimators; he further claims to prove that (ii) if the sampling design is a unicluster design there is always a class of unbiased hyperadmissible estimators. Hanurav has also expressed the conjecture that his result (i) is probably true for the entire class of unbiased estimators of the population total.

We show that (ii) is false; for any unicluster design which has three or more clusters, the H–T estimator is the unique hyperadmissible estimator. Thus for obtaining a unique hyperadmissible estimator, the restraint on the sampling design of non-uniclusterness is not the correct one. A revised condition if formulated, and it is shown that if the sampling design satisfies this revised condition, then the H–T estimator is (as conjectured by Hanurav), the unique hyperadmissible estimator in the entire class of all unbiased estimators of the population total.

The revised restraint on the sampling design is a mild one, which would be satisfied for most designs—whether unicluster or non-unicluster—met with in practical work. For the remaining cases of non-unicluster designs, which do not satisfy the revised condition, Hanurav's result (i) continues to apply, but even in these cases, the restriction to polynomial estimators is unnecessary, and the result remains valid if the class of estimators is restricted only by requiring that the estimators should be continuous functions of the variate values at the single point at which all the variate values vanish.

2. Notation. For convenience we use the same notation and definitions as Hanurav. For ready reference the notation and the relevant definitions are briefly reproduced here. The population $\mathscr{U}$ consists of distinct units $U_1, U_2, \cdots, U_N$. A sample s is a finite, ordered, sequence of units, not necessarily distinct, drawn from $\mathscr{U}$. S is the set of all possible samples s. A sampling design P (or more briefly a design) is determined by defining a probability P on S. P_s denotes the probability of the sample s when the sampling design is P. $\mathscr{Y}$ is a real variable defined on $\mathscr{U}$ which takes the value Y_i on U_i, $i = 1, 2, \cdots, N$. Y denotes the population total of the $\mathscr{Y}$-values, i.e.

(1) $$Y = \sum_{i=1}^{i=N} Y_i$$

Received April 22, 1969; revised October 17, 1970.

$\mathbf{Y} = (Y_1, Y_2, \cdots, Y_N)$ is a vector in the N-dimensional space R_N. An estimator T is a real function defined on $S \times R_N$, such that for each $s \in S$, the value of T depends on $\mathbf{Y}$ through only those Y_i, for which the unit U_i occurs in the sample (sequence) s.

$i \in s$ means that the unit U_i occurs at least once in the sample s. $\sum_{i \in s}$ denotes that the sum is taken over the distinct units U_i which occur in s, i.e. each unit is taken only once whether it occurs once or more often in the sample (sequence) s. Similarly $\sum_{s \ni i}$ denotes a sum taken over the samples s, in which the unit U_i occurs, the sample s being taken once only irrespective of the number of times, the unit U_i occurs in s. With this notation the inclusion probability π_i of the unit U_i, is given by

$$\pi_i = \sum_{s \ni i} P_s. \tag{2}$$

$\mathscr{M}^*(P)$ denotes the class of all unbiased estimators of the population total Y, which for each $s \in S$, are polynomials in Y_i.

3. Unicluster designs. A pair of samples s_1 and s_2 are said to be disjoint if the set of distinct units which occur in s_1 is disjoint from the set of distinct units which occur in s_2. A pair of samples s_1 and s_2 are said to be effectively equivalent, in symbols $s_1 \sim s_2$, if the set of distinct units which occur in s_1, is identical with the set of distinct units which occur in s_2. For a given sampling design P, we denote by $\bar{S}$ the subset of S consisting of all those samples s for which $P_s > 0$. A sampling design P is said to be a unicluster design if for every pair of samples (sequences) $s_1, s_2 \in \bar{S}$, s_1 and s_2 are either disjoint or effectively equivalent. For such sampling designs, Hanurav's theorem is as follows:

THEOREM 5.1 OF [1]. *If the unicluster design is such that*

$$0 < \pi_i, \qquad i = 1, 2, \cdots, N, \tag{3}$$

then any estimator $T \in \mathscr{M}^*(P)$ *is admissible*, iff, *T is of the form*

$$T = \{T_s(\mathbf{Y}), s \in S, \mathbf{Y} \in R_N\}, \tag{4i}$$

and for $s \in \bar{S}$,

$$T_s(\mathbf{Y}) = K_s + \sum_{i \in s} Y_i/\pi_i \tag{4ii}$$

where K_s are constants (i.e. *independent of* $\mathbf{Y}$) *satisfying*,

$$K_{s_1} = K_{s_2}, \qquad \text{if} \quad s_1 \sim s_2, \quad \text{and} \tag{5i}$$

$$\sum_{s \in \bar{S}} P_s K_s = 0. \tag{5ii}$$

Further every T satisfying (4) *and* (5) *is h-admissible* (short for hyperadmissible).

(Note. (3) is obviously a necessary condition in order that unbiased estimation of Y should be possible at all.)

It is the further part of this theorem which is not valid as is shown in the Remark below Example 5.1 in this paper. This is an immediate consequence of the main Theorem 4.1 proved in the next section.

4. Revised restriction on the sampling design. The H–T estimator is given by

$$\hat{Y}_{HT} = \sum_{i \in s} Y_i/\pi_i, \tag{6}$$

the sum in (6) being taken over the distinct units which occur in s. We shall show that the uniqueness of $\hat{Y}_{HT}$ does not depend on whether the design is unicluster or not but the uniqueness is secured instead by a different condition on P. For each i, $i = 1, 2, \cdots, N$, let $\bar{S}_i$ denote the subset of $\bar{S}$ consisting of all those samples s, in which the unit U_i occurs and $\bar{S}_i{}^*$ the subset of $\bar{S}$ consisting of all those samples s in which U_i does not occur. Clearly $\bar{S} = \bar{S}_i + \bar{S}_i{}^*$

Then,

CONDITION 4.1. The sampling design P should be such that it is possible to determine an ordered series of integers, $i_1, i_2, \cdots, i_k$, such that

$$\bar{S} = \bigcup_{r=1}^{k} \bar{S}_{i_r}^* \tag{7}$$

and for each j, $j = 2, 3, \cdots, k$, the set $\bar{S}_{i_j}^*$ has at least one sample in common with the set $\bigcup_{r=1}^{j-1} \bar{S}_{i_r}^*$.

Though Condition 4.1 appears complicated it will be seen to be satisfied for most designs considered in practical work. We shall now prove the following

THEOREM 4.1. *If the sampling design P satisfies Condition* 4.1 *and also the condition in* (3) *then the Horvitz–Thomson estimator given by* (6) *is the unique hyperadmissible estimator in the entire class of all unbiased estimators of the population total.*

PROOF. Let

$$T = \{T_s(\mathbf{Y}), \quad s \in S, \quad \mathbf{Y} \in R_N\} \tag{8}$$

be a hyperadmissible estimator. Hyperadmissibility as defined by Hanurav means that T is an unbiased estimator of the population total Y defined by (1) and further that T is admissible in the class of unbiased estimators of Y in every subspace $R(i_1, i_2, \cdots, i_m)$ of R_N where $[i_1, i_2, \cdots, i_m]$ is any set of distinct integers such that, $1 \leqq m \leqq N$, $1 \leqq i_j \leqq N$, for $j = 1, 2, \cdots, m$, and the subspace $R(i_1, i_2, \cdots, i_m)$ is defined by

$$\mathbf{Y} \in R(i_1, i_2, \cdots, i_m) \qquad \text{iff,}$$

$$Y_{i_j} \neq 0, \qquad j = 1, 2, \cdots, m \tag{9i}$$

$$Y_k = 0 \qquad \text{for every } k \notin [i_1, i_2, \cdots, i_m]. \tag{9ii}$$

Hanurav defines the admissibility of an estimator in the usual way with the squared error as loss function. We shall however take a mere general loss function

$W(d)$ where W is a nonnegative, non-decreasing and strictly convex function of the absolute value of the difference d between the estimate and the parametric function under estimation, i.e.

$$d = |T_s(\mathbf{Y}) - Y|. \tag{10}$$

This loss function includes the squared error as a special case.

Now consider the unbiasedness and admissibility of the estimator T in (8) in the subspace R (i), for some fixed integer i, $1 \leqq i \leqq N$. In this subspace, Y_i being the only variable coordinate, by the definition of an estimator for $s \in \bar{S}$ $T_s(\mathbf{Y})$ is some function $f_s(Y_i)$ of Y_i alone, and for $s \in \bar{S}_i{}^*$, $T_s(\mathbf{Y})$ is equal to some constant K_s.

Thus

$$T_s(\mathbf{Y}) = f_s(Y_i) \qquad \text{for } s \in \bar{S}_i,\ \mathbf{Y} \in R_{(i)}; \tag{11i}$$

$$T_s(\mathbf{Y}) = K_s \qquad \text{for } s \in \bar{S}_i{}^*,\ \mathbf{Y} \in R_{(i)}. \tag{11ii}$$

Note: By the definition of an estimator, (11ii) must hold also at the origin, i.e. at the point $Y_i = 0$, $i = 1, 2, \cdots, N$, for all $s \in \bar{S}$.

Put,

$$\begin{aligned} \bar{K}_i &= (1-\pi_i)^{-1} \textstyle\sum_{s \in \bar{S}_i{}^*} P_s K_s, && \text{if } \pi_i < 1, \\ &= 0 && \text{if } \pi_i = 1, \end{aligned} \tag{12}$$

and

$$\bar{f}(Y_i) = (\pi_i)^{-1} \textstyle\sum_{s \in \bar{S}_i} P_s f_s(Y_i). \tag{13}$$

Note that in (13), $\pi_i > 0$ by (3).

We now define a new unbiased estimator $\bar{T} = \{\bar{T}_s(\mathbf{Y})\}$ by

$$\bar{T}_s(\mathbf{Y}) = \bar{f}(Y_i) + \textstyle\sum_{j \in s,\, j \neq i} Y_j/\pi_j, \qquad \text{for } s \in \bar{S}_i,\ \mathbf{Y} \in R_N; \text{ and} \tag{14i}$$

$$\bar{T}_s(\mathbf{Y}) = \bar{K}_i + \textstyle\sum_{j \in s} Y_j/\pi_j \qquad \text{for for } s \in \bar{S}_i{}^*,\ \mathbf{Y} \in R_N. \tag{14ii}$$

Note that the set $\bar{S}_i{}^*$ is empty if $\pi_i = 1$.

Now for $\mathbf{Y} \in R_{(i)}$

$$\text{the population total} \quad Y = Y_i \tag{15}$$

and hence by (11), (12) and (13), the unbiasedness of T implies that

$$\pi_i \bar{f}(Y_i) + (1-\pi_i)\bar{K}_i = Y_i. \tag{16}$$

Using (16) it is easily verified that $\bar{T}$ in (14) is an unbiased estimator of Y for all $\mathbf{Y} \in R_N$.

Next consider the admissibility of T for $\mathbf{Y} \in R(\mathrm{i})$. Since $Y_j = 0$ for $j \neq i$,

(17) Risk of the estimator T

$$= \textstyle\sum_{s \in \bar{S}_i} P_s W(|f_s(Y_i) - Y_i|) + \sum_{s \in \bar{S}_i^*} P_s W(|K_s - Y_i|) \qquad \text{by (8) and (11),}$$
$$\geqq \pi_i \textstyle\sum_{s \in \bar{S}_i} W(|\bar{f}(Y_i) - Y_i|) + (1-\pi_i) W(|\bar{K}_i - Y_i|) \qquad \text{by (12) and (13),}$$
$$= \text{risk of the estimator } \bar{T} \qquad \text{by (14).}$$

In (17) we have used the well-known Jensen's inequality. Now in (17) the sign of strict inquality holds for some $\mathbf{Y} \in R(i)$, unless

(18i) $$f_s(Y_i) = \bar{f}(Y_i) \qquad \text{for all } s \in \bar{S}_i \text{ and all } Y_i \neq 0, \text{ and}$$

(18ii) $$K_s = \bar{K}_i \qquad \text{for all } s \in \bar{S}_i^*.$$

Since by assumption, T is admissible in $R(i)$, both the equalities in (18) must hold.

Using condition 4.1 we next show that $\bar{K}_i = 0$. For let $s \in \bar{S}_{i_1}^* \cap \bar{S}_{i_2}^*$. Then using the definition of an estimator, we obtain that

(19) $$\bar{K}_{i_1} = \bar{K}_{i_2}$$

We now obtain the result by induction. Suppose that

(20) $$\bar{K}_{i_r} = K, \qquad \text{for } r = 1, 2, \cdots, j,\ 1 < j \leqq k-1 \text{ where } k \text{ is the constant in Condition 4.1.}$$

Next consider,

$$s \in \bar{S}_{i_{j+1}}^* \cap \{\textstyle\bigcup_{r=1}^{j} \bar{S}_{i_r}^*\}.$$

Then by the definition of an estimator, (20) implies that

(21) $$\bar{K}_{i_{j+1}} = K.$$

Hence by induction

(22) $$\bar{K}_{i_r} = K \qquad \text{for } r = 1, 2, \cdots, k.$$

Since by Condition 4.1 $\bar{S} = \bigcup_{r=1}^{k} \bar{S}_{i_r}$, (22) and (18ii) imply that

(23) $$K_s = K \qquad \text{for all } s \in \bar{S}.$$

Now consider the unbiasedness of T at the origin, i.e. at the point $Y_i = 0$, $i = 1, 2, \cdots, N$. We obtain from the note below (11ii), together with (18ii) and (23) that

(24i) $$K = 0,$$

(24ii) $$K_s = 0 \qquad \text{for all } s \in \bar{S}, \text{ and}$$

(24iii) $$\bar{K}_i = 0 \qquad i = 1, 2, \cdots, N.$$

Hence by (16), (18) and (11) for $\mathbf{Y} \in R(i)$,

(25i) $$T_s(\mathbf{Y}) = Y_i/\pi_i \qquad \text{if } s \in \bar{S}_i,$$

(25ii) $$T_s(\mathbf{Y}) = 0 \qquad \text{if } s \in \bar{S}_i^*.$$

We now complete the proof by induction. We make the following inductive assertion:

ASSERTION A_m: Let m be a given integer $\leqq N$. For given m, let $[i_1, i_2, \cdots, i_m]$ be any set of distinct integers, each $\leqq N$. Then $R(i_1, i_2, \cdots, i_m)$ being the subspace of R_N defined by (9),

$$(26) \qquad T_s(\mathbf{Y}) = \textstyle\sum_{j \in s} Y_j/\pi_j, \qquad \text{for} \quad \mathbf{Y} \in R(i_1, i_2, \cdots, i_m) \quad \text{and all} \quad s \in \bar{S}.$$

(25) means that the Assertion A_m is true for $m = 1$. Now suppose the assertion is true for all $m \leqq h-1$ where h is some integer $\leqq N$. We shall show that the assertion must then hold also for $m = h$.

Let $\bar{S}(i_1, i_2, \cdots, i_h)$ denote the subset of $\bar{S}$ consisting of all those samples $s \in \bar{S}$ which contain each of the units $U_{i_1}, U_{i_2}, \cdots, U_{i_h}$, i.e.

$$(27) \qquad \begin{aligned} &s \in \bar{S}(i_1, i_2, \cdots, i_h) \quad \text{iff,} \\ &s \in \bar{S},\ U_{i_r} \in s, \qquad \text{for} \quad r = 1, 2, \cdots, h. \end{aligned}$$

Let

$$(28) \qquad \bar{S}^*(i_1, i_2, \cdots i_h) = \bar{S} - \bar{S}(i_1, i_2, \cdots, i_h).$$

Consider a particular sample $s' \in \bar{S}^*(i_1, i_2, \cdots, i_h)$. Let $U_{j_1}, U_{j_2}, \cdots, U_{j_f}$ be the distinct units contained in s' such that for

$$(29) \qquad \begin{aligned} &\mathbf{Y} \in R(i_1, i_2, \cdots, i_h) \\ &\mathbf{Y}_{j_r} \neq 0, \qquad r = 1, 2, \cdots, f. \end{aligned}$$

It is seen that (28) implies that

$$(30) \qquad f \leqq (h-1)$$

and that the set of integers $[j_1, j_2, \cdots, j_f]$ is a proper subset of the set $[i_1, i_2, \cdots, i_h]$. Now by the definition of an estimator,

$$(31) \qquad \begin{aligned} &T_{s'}(\mathbf{Y}) && \text{for} \quad \mathbf{Y} \in R(i_1, i_2, \cdots, i_h) \\ &\quad = T_{s'}\mathbf{Y}) && \text{for} \quad \mathbf{Y} \in R(j_1, j_2, \cdots, j_f) \\ &\quad = \textstyle\sum_{j \in s'} Y_j/\pi_j && \text{by the} \end{aligned}$$

inductive assertion which is assumed to hold for all $f \leqq (h-1)$.
Hence

$$(32) \quad T_s(\mathbf{Y}) = \textstyle\sum_{j \in s} Y_j/\pi_j, \quad \text{for all} \quad s \in \bar{S}^*(i_1, i_2, \cdots, i_h) \quad \text{and} \quad \mathbf{Y} \in R(i_1, i_2, \cdots, i_h).$$

Next consider the samples $s \in \bar{S}(i_1, i_2, \cdots, i_h)$. Now two alternatives are possible viz. that (A) the set $\bar{S}\,(i_1, i_2, \cdots, i_h)$ in (27) is empty or (B) it is non-empty. Suppose (A) is true. Then $\bar{S}^*(i_1, i_2, \cdots, i_h) = \bar{S}$ so that (32) holds for all $s \in \bar{S}$.

Next suppose alternative (B) is true.

Put

$$(33) \qquad \gamma = \textstyle\sum_{s \in \bar{S}(i_1, i_2, \ldots, i_h)} P_s.$$

Since $\bar{S}(i_1, \cdots, i_h)$ is non-empty and $P_s > 0$ for every $s \in \bar{S} \supset \bar{S}(i_1, i_2, \cdots, i_h)$

$$\gamma > 0. \tag{34}$$

For any $\mathbf{Y} \in R_N$, let $\mathbf{z}_h$ be the projection of $\mathbf{Y}$ on the subspace $R(i_1, i_2, \cdots, i_h)$. Then by the definition of an estimator, for all $s \in \bar{S}(i_1, i_2, \cdots, i_h)$, and $\mathbf{Y} \in R(i_1, i_2, \cdots, i_h)$ $T_s(\mathbf{Y})$ is some function of the vector $\mathbf{z}_h$ alone, i.e.

$$T_s(\mathbf{Y}) = f_s(\mathbf{z}_h), \qquad \text{for} \quad s \in \bar{S}(i_1, \cdots, i_h) \quad \mathbf{Y} \in R(i_1, \cdots, i_h). \tag{35}$$

Now put

$$\bar{f}(\mathbf{z}_h) = (\gamma)^{-1} \textstyle\sum_{s \in S(i_1, i_2, \dots, i_h)} P_s f_s(\mathbf{z}_h) \tag{36}$$

and define a new estimator $\bar{T} = \{\bar{T}_s(\mathbf{Y})\}$ by

$$\bar{T}_s(\mathbf{Y}) = T_s(\mathbf{Y}) \qquad \text{for} \quad \mathbf{Y} \in R_N \quad \text{and} \quad s \in \bar{S}^*(i_1, i_2, \dots, i_h), \tag{37i}$$

$$\bar{T}_s \mathbf{Y}) = \bar{f}(\mathbf{z}_h) + \textstyle\sum_{j \notin [i_1, i_2, \dots, i_h], j \in s}, Y_j/\pi_j, \quad \text{for} \quad s \in \bar{S}(i_1, i_2, \cdots, i_h)\ \mathbf{Y} \in R_N. \tag{37ii}$$

Now since the estimator T is unbiased, by taking a point $\mathbf{Y} \in R(i_1, i_2, \cdots, i_h)$, we get

$$\textstyle\sum_{r=1}^h Y_{i_r} = \sum_{s \in \bar{S}^*(i_1, i_2, \dots, i_h)} P_s T_s(\mathbf{Y}) + \sum_{s \in \bar{S}(i_1, i_2, \dots, i_h)} P_s f_s(\mathbf{z}_h). \tag{38}$$

Substituting in the first term in the right-hand side of (38) by (32) and in the second term by (36), we obtain after some algebraic simplification that

$$\bar{f}(\mathbf{z}_h) = \textstyle\sum_{r=1}^h Y_{i_r}/\pi_{i_r}. \tag{39}$$

Combining (39) with (37ii) and (37i) with (32) we obtain that

$$\bar{T}_s(\mathbf{Y}) = \textstyle\sum_{j \in s} Y_j/\pi_j, \qquad \text{for all} \quad s \in \bar{S}, \mathbf{Y} \in R_N. \tag{40}$$

The estimator $\bar{T}$ defined by (40), is the same as the H–T estimator which is unbiased. Further considering the admissibility of the estimator T in the subspace $R(i_1, i_2, \cdots, i_h)$ and using Jensen's inequality, we obtain as before,

$$T_s(\mathbf{Y}) = \bar{T}_s(\mathbf{Y}) \qquad \text{for} \quad s \in \bar{S}(i_1, i_2, \cdots, i_h). \tag{41}$$

Combining (41), (40) and (32), we obtain that

$$T_s(\mathbf{Y}) = \textstyle\sum_{j \in s} Y_j/\pi_j \qquad \text{for all} \quad s \in \bar{S} \quad \text{and} \quad \mathbf{Y} \in R(i_1, i_2, \cdots, i_h). \tag{42}$$

Thus whether the alternative (A) holds or the alternative (B), the relation (42) is valid. Hence if the inductive assertion A_m is true for $m \leqq h-1$, it is true for $m = h$. By (25), the assertion is true for $m = 1$. Hence it is true for all $m \leqq N$. From this it follows that the equality in (42) holds for all $\mathbf{Y} \in R_N$ except perhaps at the origin, i.e., at the point $Y_i = 0$, $i = 1, 2, \cdots, N$. But by the note below (11ii) together with (24ii) the equality in (42) holds at the origin also.
Hence we finally get

$$T_s(\mathbf{Y}) = \textstyle\sum_{j \in s} Y_j/\pi_j \qquad \text{for all} \quad s \in \bar{S}, \quad \text{and} \quad \mathbf{Y} \in R_N. \tag{43}$$

This completes the proof of the theorem.

5. General remarks. As observed before, Condition 4.1 will be satisfied by most sampling designs considered in practice. To see the necessity of imposing this condition, we give the following example.

EXAMPLE 5.1. The population consists of six units $U_1, U_2, \cdots, U_6$; the sampling design assigns positive probabilities to only two samples $s_1 = (U_1, U_2, U_3, U_2)$ and $s_2 = (U_4, U_4, U_5, U_6, U_5)$; let the probabilities of s_1, s_2 be respectively $p_1, p_2, p_1, p_2 > 0$ and $p_1 + p_2 = 1$. Then the estimator given by

$$(44)\qquad \begin{aligned} T_{s_1}(\mathbf{Y}) &= (p_1)^{-1}(Y_1 + Y_2 + Y_3) + K_1 \\ T_{s_2}(\mathbf{Y}) &= (p_2)^{-1}(Y_4 + Y_5 + Y_6) + K_2 \end{aligned}$$

where $p_1K_1 + p_2K_2 = 0$ is seen to be an unbiased hyperadmissible estimator.

REMARK. The sampling design in Example 5.1 is a unicluster one, where $\bar{S}$ contains two non-equivalent samples. It is easily seen that any non-unicluster design which assigns positive probability to not less than three mutually non-equivalent samples, necessarily satisfies Condition 4.1 of this paper so that the H–T estimator is the unique hyperadmissible estimator. Hence the further part of Theorem 5.1. of Hanurav is valid only for those unicluster designs which have only two or one non-equivalent samples with positive probability.

To show the necessity of imposing Condition 4.1 for a non-unicluster design, we give the following example.

EXAMPLE 5.2. The population consists of three units U_1, U_2, U_3; $\bar{S}$ consists of three samples, $s_1 = (U_1, U_2, U_1)$; $s_2 = (U_1, U_3, U_1)$ and $s_3 = (U_2, U_3, U_2, U_3)$ with probabilities, $p_1, p_2, p_3, p_1, p_2, p_3 > 0$, $p_1 + p_2 + p_3 = 1$. Consider the estimator

$$T = \{T_s(\mathbf{Y})\},$$

$$(45)\qquad \begin{aligned} T_{s_1}(\mathbf{Y}) &= Y_1/\pi_1 + Y_2/\pi_2 - (p_2a_1 + p_3a_2)/p_1, && \text{if } Y_1 \neq 0,\ Y_2 \neq 0, \\ &= Y_1/\pi_1 + a_1, && Y_1 \neq 0,\ Y_2 = 0, \\ &= Y_2/\pi_2 + a_2, && Y_1 = 0,\ Y_2 \neq 0, \\ &= -\pi_3a_3/p_1, && Y_1 = 0,\ Y_2 = 0, \\ T_{s_2}(\mathbf{Y}) &= Y_1/\pi_1 + Y_3/\pi_3 - (p_1a_1 + p_3a_3)/p_2, && Y_1 \neq 0,\ Y_3 \neq 0, \\ &= Y_1/\pi_1 + a_1, && Y_1 \neq 0,\ Y_3 = 0, \\ &= Y_3/\pi_3 + a_3, && Y_1 = 0,\ Y_3 \neq 0, \\ &= -\pi_2a_2/p_2, && Y_1 = 0,\ Y_3 = 0, \\ T_{s_3}(\mathbf{Y}) &= Y_2/\pi_2 + Y_3/\pi_3 - (p_1a_2 + p_2a_3)/p_3, && Y_2 \neq 0,\ Y_3 \neq 0, \\ &= Y_2/\pi_2 + a_2, && Y_2 \neq 0,\ Y_3 = 0, \\ &= Y_3/\pi_3 + a_3, && Y_2 = 0,\ Y_3 \neq 0, \\ &= -\ \pi_1a_1/p_3, && Y_2 = 0,\ Y_3 = 0. \end{aligned}$$

where a_1, a_2, a_3 are arbitrary subject to the restriction that

$$\pi_1 a_1 + \pi_2 a_2 + \pi_3 a_3 = 0.$$

In (45), π_1, π_2, π_3 have the usual meanings i.e. $\pi_1 = p_1 + p_2$, etc.

It is easily verified that the estimator in (45) is unbiased for all $\mathbf{Y} \in R_N$ and is hyperadmissible. Thus there is no unique hyperadmissible estimator. The estimator in (45) however reduces to the H–T estimator if we impose the additional restriction that the estimator should be continuous in Y_i at the point at which all $Y_i, i \in s$ vanish. This suggests that when the sampling design is a unicluster design which does not satisfy Condition 4.1, the imposition of the additional restriction of the estimators of continuity of the origin will secure uniqueness of the H–T estimator. That this is so, is shown in the next session.

6. Non-unicluster designs. We consider a non-unicluster design, for which Condition 4.1 is not satisfied. According to Theorem 3.1 of [1], in this case, the H–T estimator is the unique hyperadmissible estimator in the class of unbiased, polynomial estimators. We shall show that this uniqueness holds for the wider class of unbiased estimators subject only to the restriction that for each $s \in \bar{S}$, $T_s(\mathbf{Y})$ is continuous in Y_i at the point $Y_i = 0$ for all $i \in s$. The proof requires only minor modifications in the proof of Theorem 3.1 in [1].

PROOF. As the sampling design is a non-unicluster one, there exists at least one pair i, j, satisfying

$$0 < \pi_{ij} < \pi_i. \tag{46}$$

Let $T = \{T_s(\mathbf{Y})\}$ be a hyperadmissible estimator which satisfies the restriction of continuity at the point $Y_i = 0$, $i \in s$. Then by considering the admissibility of T in $R(i)$, we obtain from (16) and (18), on putting in (16) $K_2(i) = \bar{K}_i$ and $K_1(i) = -(\pi_i)^{-1}(1-\pi_i)\bar{K}_i$,

for $\mathbf{Y} \in R(i)$,

$$T_s(\mathbf{Y}) = Y_i/\pi_i + K_1(i) \qquad \text{for all} \quad s \in \bar{S}_i, \tag{47i}$$

$$T_s(\mathbf{Y}) = K_2(i) \qquad \text{for all} \quad s \in \bar{S}_i^*. \tag{47ii}$$

Similarly, for $\mathbf{Y} \in R(j)$,

$$T_s(\mathbf{Y}) = Y_j/\pi_j + K_1(j) \qquad \text{for all} \quad s \in \bar{S}_j, \tag{48i}$$

$$T_s(\mathbf{Y}) = K_2(j) \qquad \text{for all} \quad s \in \bar{S}_j^*. \tag{48ii}$$

Since $\pi_{ij} > 0$, there is one sample s_0 say, which belongs to both $\bar{S}_i$ and $\bar{S}_j$. Hence putting $s = s_0$ in (47i) and (48i) and taking the limit as $Y_i \to 0$ in (47i) and as $Y_j \cdots 0$ in (48i), and using the hypothesis of continuity of $T_{s_0}(\mathbf{Y})$ at $Y_i = 0$, $Y_j = 0$, we get

(49) $$K_1(i) = K_1(j).$$

Also since $\pi_i > \pi_{ij}$ in (46), there is at least one sample s_1 which belongs to $\bar{S}_i$ and $\bar{S}_j^*$. Hence putting $s = s_1$ in (47i) and taking the limit of $T_{s_1}(\mathbf{Y})$ when $Y_i \to 0$, and similarly putting $s = s_1$ in (48ii), we get

(50) $$K_1(i) = K_2(j);$$

combining (49) and (50) we get that in (48), $K_1(j) = K_2(j)$ and hence using the condition of unbiasedness of T for $\mathbf{Y} \in R(j)$, we get that in (48)

(51) $$K_1(j) = K_2(j) = 0.$$

Next consider any integer r, $r \neq j$. r may assume the value i, which occurs in (46). By considering the admissibility of T in $R(r)$, we have as in (47), for $\mathbf{Y} \in R(r)$,

(52i) $$T_s(\mathbf{Y}) = Y_r/\pi_r + K_1(r) \qquad \text{if} \quad r \in \bar{S}_r,$$

(52ii) $$T_s(\mathbf{Y}) = K_2(r), \qquad \text{if} \quad r \in \bar{S}_r^*.$$

If $\bar{S}_r^*$ is empty, then $K_1(r) = 0$ in (52) by the unbiasedness of T for $\mathbf{Y} \in R(r)$. Suppose $\bar{S}_r^*$ is not empty. Then it has a non-empty intersection with at least one of the sets $\bar{S}_j$ and $\bar{S}_j^*$. Suppose that its intersection with $\bar{S}_j$ is non-empty, so that there exists a sample s_2 which belongs to $\bar{S}_j$ and also to $\bar{S}_r^*$. Now put $s = s_2$ in (48i) and in (52ii) and take the limit in (48i) when $Y_j \to 0$. We thus get

(53) $$K_2(r) = K_1(j) = 0, \qquad \text{by (51).}$$

Similarly if $\bar{S}_r^*$ has a non-empty intersection with $\bar{S}_j^*$ we get $K_2(r) = K_2(j) = 0$.

Thus always $K_2(r) = 0$ and hence by the unbiasedness of T for $\mathbf{Y} \in R(r)$, in (51) $K_1(r)$ also $= 0$.

It thus follows from (51), (52) and (53), that for any r, $1 \leqq r \leqq N$, for $\mathbf{Y} \in R(r)$

(54) $$T_s(\mathbf{Y}) = Y_r/\pi_r \qquad \text{if} \quad s \in \bar{S}_r,$$
$$= 0 \qquad \text{if} \quad s \in \bar{S}_r^*.$$

(54) is the same as the assertion A_1 in (26). Hence the further proof is identical with that from (26) onward of Theorem 4.1 in this paper.

Remark 6.1. The above proof and Example 5.2 show, that when the sampling design does not satisfy Condition 4.1, the H–T estimator may not be the unique hyperadmissible estimator. Uniqueness is secured only by imposing on the estimators the additional restriction of being continuous at the origin. But the requirement of continuity does not seem to be a natural one in the context of the notion of hyperadmissibility.

Acknowledgment. I am grateful to Professor J. N. K. Rao for suggesting this investigation and to the referee for his valuable comments which have greatly improved the presentation in this paper. I have also to thank Mr. V. J. A. Mudaliar for doing the repeated typing work with neatness and patience.

REFERENCE

[1] HANURAV, T. V. (1968). Hyper-admissibility and optimum estimator for sampling finite populations. *Ann. Math. Statist.* **39** 621–641.