The Annals of Statistics
1975, Vol. 3, No. 3, 669-679

CHOICE OF AN OPTIMUM SAMPLING STRATEGY—I

BY M. K. RAMAKRISHNAN

University of Singapore

The problem of finding an optimum sampling strategy from the class of all linear unbiased strategies with a given expected sample size is considered. The paper deals with the properties of admissibility, completeness and strong admissibility of the Horvitz-Thompson strategies.

1. Introduction. One main feature of the theory of survey sampling which distinguishes it from other parts of statistical inference is that the randomness involved in survey sampling is deliberately injected by the statistician, and furthermore, the nature of this randomness is within certain limits at the disposal of the statistician. As a result, the problem in survey sampling is not just to choose an optimum estimator but to choose an optimum combination of sampling and estimation procedures.

In this series of papers we intend to study systematically the central problem of the choice of an optimum combination of sampling and estimation procedures. In this paper we discuss the problem of finding an optimum sampling strategy from the class $LH^*(\mu)$ of all linear unbiased strategies with a certain expected sample size μ. After proving the admissibility of any Horvitz-Thompson strategy (HT-strategy, for short) we show that the set of all such strategies is not complete in $LH^*(\mu)$ in situations of practical interest.

After noting that there does not exist a hyperadmissible strategy in $LH^*(\mu)$ we introduce a new criterion called "strong admissibility" which is stronger than admissibility but weaker than hyperadmissibility. We prove that the set of all HT-strategies in $LH^*(\mu)$ is complete in $LH^*(\mu)$ with respect to strong admissibility.

2. Notations and definitions. Let U denote the population consisting of N units, denoted by the integers $1, 2, \cdots, N$. If the variate value associated with the unit i is y_i, $i = 1, 2, \cdots, N$, then $\mathbf{y} = (y_1, y_2, \cdots, y_N)$ is a point in R_N, the N-dimensional Euclidean space, and the population total is a function on R_N given by

$$Y = Y(\mathbf{y}) = \textstyle\sum_1^N y_i . \tag{1}$$

A nonempty subset s of U is called a sample and S denotes the set of all possible samples s. A real function p on S, such that

$$p(s) \geqq 0 \quad \text{for all} \quad s, \qquad \sum p(s) = 1 \tag{2}$$

Received September 1972; revised November 1973.
AMS 1970 *subject classification.* 62D05.
Key words and phrases. Admissibility, HT-strategy, strong admissibility.

is called a (sampling) design. For a design p let

$$\pi_i(p) = \sum_{s \ni i} p(s) \qquad 1 \leqq i \leqq N \tag{3}$$

and

$$\pi_{ij}(p) = \sum_{s \ni i,j} p(s) \qquad 1 \leqq i \neq j \leqq N \tag{4}$$

where in (3) the sum is over all samples that contain the unit i and in (4) the sum is over all samples that contain the units i and j. They are called the first and second order inclusion probabilities respectively and they play an important role in the choice of optimum sampling strategies.

The conventional problem in survey sampling is to estimate the population total Y by observing the values of y_i for which $i \in s$, where s is a sample drawn according to a design p.

DEFINITION 2.1. As estimator e is a real-valued function on $S \times R_N$ which depends on y only through those y_i for which $i \in s$, that is $e(s, \mathbf{y}) = e(s, \mathbf{y}')$ for any two $\mathbf{y}, \mathbf{y}'$ such that $y_i = y_i'$ for all $i \in s$.

From practical considerations it is evident that the estimate $e(s, \mathbf{y})$ need not be defined for those samples for which $p(s) = 0$. An estimator is said to be linear if it is of the form

$$e(s, \mathbf{y}) = \sum_{i=1}^{N} b(s, i) y_i \tag{5}$$

where b is a function on $S \times U$ such that $b(s, i) = 0$ if $i \notin s$.

Let $M(e, \mathbf{y})$ denote the mean squared error (mse, for short) of an estimator e. If e is unbiased the mean square error is the same as the variance and is denoted by $V(e, \mathbf{y})$. Godambe (1955) has proved the celebrated result that for any design p there does not exist a uniformly minimum variance (umv, for short) estimator in the class $L^*(p)$ of all linear p-unbiased estimators of the population total. However, later Godambe (1965), Hege (1965) and Hanurav (1965) pointed out some exceptions to the theorem and gave nontrivial designs where best estimators exist. Such designs were called unicluster designs by Hanurav (1965).

DEFINITION 2.2. A design p is said to be a unicluster design if $s_1, s_2 \in \bar{S}$, $s_1 \neq s_2$ implies $s_1 \cap s_2 = \emptyset$ where $\bar{S}$ is the set of all samples for which $p(s) > 0$.

DEFINITION 2.3. With respect to a design p, an estimator e_0 belonging to a class D of estimators is said to be admissible in D if for no other estimator $e \in D$,

$$M(e, \mathbf{y}) \leqq M(e_0, \mathbf{y}) \tag{6}$$

for all $\mathbf{y} \in R_N$, strict inequality being true for at least one $\mathbf{y} \in R_N$.

DEFINITION 2.4. A design p together with an estimator e of Y is called a (sampling) strategy for the estimation of Y and is denoted by (p, e) or sometimes just by H.

A strategy (p, e) is said to be unbiased for Y if e is p-unbiased for Y. The expectation, variance or mean square error of a strategy are defined as the expectation, variance or mean square error of the corresponding estimator. Analogous to the Definition 2.3 of admissibility of an estimator we have

DEFINITION 2.5. A strategy H_0 belonging to a class $\mathscr{H}$ of strategies is said to be admissible in $\mathscr{H}$ if for no other strategy $H \in \mathscr{H}$,

$$M(H, \mathbf{y}) \leqq M(H_0, \mathbf{y}) \tag{7}$$

for all $\mathbf{y} \in R_N$, strict inequality being true for at least one $\mathbf{y} \in R_N$.

A principal hypersurface (phs, for short) of R_N is defined as a linear subspace of all points $\mathbf{y} = (y_1, y_2, \cdots, y_N)$ with $y_{i_1} = y_{i_2} = \cdots = y_{i_k} = 0$ where $0 \leqq k < N$ and $(i_1, i_2, \cdots, i_k)$ is a subset of $(1, 2, \cdots, N)$. Clearly the whole space R_N corresponds to the case $k = 0$ and there are, in all, $2^N - 1$ phs's of R_N. Let $\mathscr{H}$ be a class of strategies for the estimation of Y.

DEFINITION 2.6. $H_0 \in \mathscr{H}$ is hyperadmissible in $\mathscr{H}$, if it is admissible (Definition 2.5) in $\mathscr{H}$ when we restrict $\mathbf{y}$ to any of the $2^N - 1$ phs's of R_N.

DEFINITION 2.7. A subclass $\mathscr{H}_1$ of $\mathscr{H}$ is said to be complete in $\mathscr{H}$ if for any $H_0 \in \mathscr{H} - \mathscr{H}_1$, there exists an $H_1 \in \mathscr{H}_1$ such that

$$M(H_1, \mathbf{y}) \leqq M(H_0, \mathbf{y}) \tag{8}$$

for all $\mathbf{y} \in R_N$. Further, if every $H_1 \in \mathscr{H}_1$ is admissible in $\mathscr{H}$, $\mathscr{H}_1$ is called a minimal complete subclass of $\mathscr{H}$.

Evidently if one wants to search for an optimum sampling strategy from among members of $\mathscr{H}$ one may restrict one's attention to any complete subclass of $\mathscr{H}$.

3. Completeness of unbiased strategies. A strategy $H = (p, e)$ is said to be linear unbiased if $e \in L^*(p)$. The expected sample size of a strategy $H = (p, e)$ is defined as $\mu(H) = \sum_s n(s)p(s)$ where $n(s)$ is the number of units in s. Let $AH(\mu)$, $H^*(\mu)$ and $LH^*(\mu)$ respectively denote the classes of all, all unbiased and all linear unbiased strategies with expected sample size μ, a given number.

THEOREM 3.1. *The class $H^*(\mu)$ is complete in $AH(\mu)$ if and only if $\mu = N$.*

PROOF. Suppose $\mu < N$ and that $H^*(\mu)$ is complete in $AH(\mu)$. Let p_1 be a sampling design with expected sample size μ and $p_1(s_0) = (\mu - 1)/(N - 1)$ where s_0 denotes the sample consisting of all the units in the population. Consider the strategy $H_1 = (p_1, e_1)$ where

$$\begin{aligned} e_1(s, \mathbf{y}) &= Y \qquad \text{for} \quad s = s_0 \\ &= \delta \qquad \text{for} \quad s \neq s_0 \end{aligned}$$

and δ a nonzero real number. By hypothesis there exists an $H_0 \in H^*(\mu)$ such that

$$V(H_0, \mathbf{y}) \leqq M(H_1, \mathbf{y}) = (N - \mu)/(N - 1)(\delta - Y)^2 . \tag{9}$$

If $H_0 = (p, e)$ it follows from (9) that $e(s, \mathbf{y}) = \delta$ for all s and $\mathbf{y}$ except perhaps for the sample s_0. Since H_0 is unbiased it is clear that $p(s_0) > 0$ and $e(s_0, \mathbf{y}) = ((Y - \delta(1 - p(s_0))/p(s_0)$ so that

$$V(H_0, \mathbf{y}) = (1 - p(s_0))/p(s_0)(\delta - Y)^2 . \tag{10}$$

Comparing (9) and (10) we get $p(s_0) \geqq (N-1)/(2N - \mu - 1)$. The expected sample size of $H_0 \geqq Np(s_0) + 1(1 - p(s_0)) \geqq 1 + (N-1)^2/(2N - \mu - 1) > \mu$, a contradiction. If $\mu = N$, clearly $H^*(\mu)$ has a member H_0 with $V(H_0, \mathbf{y}) = 0$ for all $\mathbf{y} \in R_N$. The proof of the theorem is complete.

REMARK 3.1. From the above theorem we see that $H^*(\mu)$ will not be complete in $AH(\mu)$ except in the trivial case $\mu = N$, that is, when we have a complete census. In practice, $\mu < N$, and so one cannot exclude biased strategies from the point of view of mean square error criterion alone.

4. Horvitz–Thompson strategies.

A particular unbiased estimator suggested by Horvitz and Thompson (1952) has received much attention recently. It is called the Horvitz–Thompson estimator (HT-estimator, for short) and is defined by

$$\bar{e}(s, \mathbf{y}) = \sum_{i \in s} y_i/\pi_i \tag{11}$$

where π_i, $i = 1, 2, \cdots, N$ are the first order inclusion probabilities for the units. In the sequel the HT-estimator will always be denoted by $\bar{e}$. Any strategy $H = (p, \bar{e})$ where $\bar{e}$ is the HT-estimator for the design p is called a Horvitz–Thompson strategy (HT-strategy for short) for the estimation of Y and its variance is given by

$$V(H, \mathbf{y}) = \sum_{i=1}^N \left(\frac{1}{\pi_i(p)} - 1\right) y_i^2 + \sum_{i \neq j}^N \sum \left(\frac{\pi_{ij}(p)}{\pi_i(p)\pi_j(p)} - 1\right) y_i y_j . \tag{12}$$

THEOREM 4.1. *Any strategy $H_0 = (p_0, \bar{e}) \in H^*(\mu)$ is admissible in $H^*(\mu)$.*

PROOF. Given any strategy $H_1 = (p_1, e_1) \in H^*(\mu)$ with

$$\sum p_1(s) e_1^2(s, \mathbf{y}) \leqq \sum p_0(s) \bar{e}^2(s, \mathbf{y}) \tag{13}$$

for all $\mathbf{y}$ we shall show that the strict inequality in (13) cannot hold at any point $\mathbf{y}$. Applying (13) at the origin and remembering that e_1 is an estimator, we have, for every s with $p_1(s) > 0$

$$e_1(s, \mathbf{y}) = 0 \quad \text{if} \quad y_i = 0 \quad \text{for all} \quad i \in s . \tag{14}$$

Next consider a point $\mathbf{y}'$ which has only one non-vanishing coordinate, say y_i'. Unbiasedness of e_1 at $\mathbf{y}'$ together with (14) gives $\sum_{s \ni i} p_1(s) e_1(s, \mathbf{y}') = y_i'$, so that by the Cauchy inequality

$$\sum_{s \ni i} p_1(s) e_1^2(s, \mathbf{y}') \geqq y_i'^2/\pi_i(p_1) . \tag{15}$$

Applying (13) at $\mathbf{y}'$ and using (15), we get $\pi_i(p_1) \geqq \pi_i(p_0)$. Since i is arbitrary and $\mu(H_0) = \mu(H_1)$ it follows that

$$\pi_i(p_1) = \pi_i(p_0) = \pi_i \quad \text{(say)} , \qquad 1 \leqq i \leqq N . \tag{15a}$$

We also obtain that the sign of equality must hold in (15) and consequently

$$e_1(s, \mathbf{y}) = y_i/\pi_i \tag{15b}$$

for any s with $p_1(s) > 0$ and any point $\mathbf{y}$ such that the set of coordinates y_j, $j \in s$ include only one non-vanishing coordinate y_i. Next consider a point $\mathbf{y}''$ which has only two non-vanishing coordinates, say y_i'', y_j''. Unbiasedness of e_1 at $\mathbf{y}''$ together with (14) and (15b) gives $\sum_{s \ni i,j} p_1(s)e_1(s, \mathbf{y}'') = \pi_{ij}(p_1)(y_i''/\pi_i + y_j''/\pi_j)$, so that by the Cauchy inequality

$$\sum_{s \ni i,j} p_1(s)e_1^2(s, \mathbf{y}'') \geqq \pi_{ij}(p_1)\left(\frac{y_i''}{\pi_i} + \frac{y_j''}{\pi_j}\right)^2 . \tag{15c}$$

Applying (13) at $\mathbf{y}''$ and using (15a) we get

$$\sum_{s \ni i,j} p_1(s)e_1^2(s, \mathbf{y}'') \leqq \pi_{ij}(p_1)\left(\frac{y_i''^2}{\pi_i^2} + \frac{y_j''^2}{\pi_j^2}\right) + \frac{2\pi_{ij}(p_0)}{\pi_i \pi_j} y_i'' y_j'' . \tag{15d}$$

Comparison of (15c) and (15d) gives $\pi_{ij}(p_1)y_i''y_j'' \leqq \pi_{ij}(p_0)y_i''y_j''$, and since the sign of $y_i''y_j''$ may be positive or negative it follows that $\pi_{ij}(p_1) = \pi_{ij}(p_0)$. As i and j are arbitrary we see that the first and second order inclusion probabilities of p_1 and p_0 coincide. Let $\bar{e}_1$ be the HT-estimator for p_1. Then $\sum_s p_1(s)\bar{e}_1^2(s, \mathbf{y}) = \sum_s p_0(s)\bar{e}^2(s, \mathbf{y})$. Thus the strict inequality cannot hold in (13), for otherwise $\bar{e}_1$ would be inadmissible for the design p_1. Hence the theorem.

REMARK 4.1. From Theorem 4.1 it follows that there does not exist a best (in the sense of uniformly minimum variance) strategy in $H^*(\mu)$ for there are at least two, in fact infinitely many, HT-strategies belonging to $H^*(\mu)$.

REMARK 4.2. Neither of two HT-strategies with the same expected sample size is better than the other since both of them are admissible.

REMARK 4.3. Given an unbiased strategy H with $\mu(H) < \mu$, one can easily construct another strategy $H' \in H^*(\mu)$ such that $V(H', \mathbf{y}) \leqq V(H, \mathbf{y})$ for all $\mathbf{y} \in R_N$. Hence it follows that Theorem 4.1 remains valid with $H^*(\mu)$ replaced by $H^*(\leqq \mu)$ where $H^*(\leqq \mu)$ denotes the class of all unbiased strategies with expected sample size less than or equal to μ.

Let $\mathrm{HT}(\mu)$ denote the class of all HT-strategies with expected sample size μ. The following theorem shows that in situations of practical interest we cannot exclude strategies other than the HT-strategies using the criterion of minimum variance alone.

THEOREM 4.2. *The class* $\mathrm{HT}(\mu)$ *is complete in* $LH^*(\mu)$ *if and only if* $\mu = 1$ *or* N.

Before proving the theorem we digress a little to prove a generalization of a result due to Joshi (1966). Removing the condition of unbiasedness and considering squared error as loss function we establish the admissibility of a particular strategy in the class $AH(\mu)$ of all strategies with expected sample size μ.

In the sequel e^* will denote the estimator given by

$$e^*(s, \mathbf{y}) = \frac{N}{n(s)} \sum_{i \in s} y_i \tag{16}$$

where $n(s)$ is the number of units in s. Concerning this estimator Joshi (1966) has proved the following two theorems.

THEOREM 4.3. *For any design p, the estimator e^* is admissible for Y in the class $A(p)$ of all estimators.*

THEOREM 4.4. *For any design p of fixed size μ, the strategy (p, e^*) is admissible for Y in the class $AH(\mu)$ of all strategies H with expected sample size $\mu(H) = \mu$.*

We now generalize the latter of these two results to cover some cases when p is not of fixed size.

THEOREM 4.5. *Let p be a design of expected size μ, put $m = [\mu]$ (= the integral part of μ) and $f = \mu - m$ (= the fractional part of μ); suppose that*

$$V_p(n(s)) = f(1 - f)\,. \tag{17}$$

Then the strategy $H = (p, e^*)$ is admissible in $AH(\mu)$.

In order to prove the theorem we require the following

LEMMA 4.1. (Joshi (1965)). *If*

(a) *$y_1, y_2, \cdots, y_N$ are independently and identically distributed random variables,*
(b) *$\phi_n(y)$ is a real function of $y_1, y_2, \cdots, y_n$ for $n = 1, 2, \cdots, N$,*
(c) *$\bar{y}_n = (1/n) \sum_{i=1}^n y_i$ for $n = 1, 2, \cdots, N$,*
(d) *for every common finite discrete frequency function w of $y_1, y_2, \cdots, y_N$,*

$$\sum_{n=1}^N A_n^2 E_w(\phi_n(y) - \theta)^2 \leqq \sum_{n=1}^N A_n^2 E_w(\bar{y}_n - \theta)^2\,,$$

E_w denoting the expectation, θ the common mean of $y_1, y_2, \cdots, y_N$ and $A_1, A_2, \cdots, A_N$ being arbitrary constants, then for every $\mathbf{y} = (y_1, y_2, \cdots, y_N) \in R_N$ we have $\phi_n(\mathbf{y}) = \bar{y}_n$ for all n for which $A_n \neq 0$.

PROOF OF THEOREM 4.5. Suppose the theorem is not true. Then there exists a strategy $H' = (p', e')$ such that $\mu(H') = \mu$ and

$$M(H', \mathbf{y}) \leqq M(H, \mathbf{y}) \qquad \text{for all} \quad \mathbf{y} \in R_N \tag{18}$$

with strict inequality for at least one $\mathbf{y}$. Let $\bar{S}'$ be the set of samples s for which $p'(s) > 0$, and let $\bar{S}$ denote the corresponding set for the sampling design p. Evidently

$$V_p(n(s)) = f(1 - f) \Leftrightarrow \sum_{s:n(s)=k} p(s) = 1 - f \quad \text{for} \quad k = m$$
$$= f \quad \text{for} \quad k = m + 1\,. \tag{19}$$

Clearly

$$\sum n(s)p'(s) = \sum n(s)p(s) = m + f \tag{20}$$

and

$$\sum p'(s)(e'(s, \mathbf{y}) - Y)^2 \leqq \sum p(s)(e^*(s, \mathbf{y}) - Y)^2 \tag{21}$$

where strict inequality holds for at least one $\mathbf{y} \in R_N$. Taking expectations on both sides of (21) with respect to a prior distribution on R_N under which $y_1, y_2, \cdots, y_N$ are distributed independently and identically with mean θ and variance σ^2, we get

$$\sum p'(s)E(e'(s, \mathbf{y}) - Y)^2 \leqq \sum p(s)E(e^*(s, \mathbf{y}) - Y)^2 . \tag{22}$$

Defining

$$g'(s, \mathbf{y}) = [N - n(s)]^{-1}(e'(s, \mathbf{y}) - \textstyle\sum_{i \in s} y_i) \tag{23}$$

and making use of the fact that $y_1, y_2, \cdots, y_N$ are independently distributed, we get

$$E(e'(s, \mathbf{y}) - Y)^2 = (N - n(s))^2 E(g'(s, \mathbf{y}) - \theta)^2 + (N - n(s))\sigma^2 . \tag{24}$$

Similarly

$$E(e^*(s, \mathbf{y}) - Y)^2 = (N - n(s))^2 E(\bar{y}_s - \theta)^2 + (N - n(s))\sigma^2 \tag{25}$$

where $\bar{y}_s = (n(s))^{-1} \sum_{i \in s} y_i$. Now substituting (24) and (25) in (22) and cancelling out the common terms, (22) becomes

$$\sum p'(s)(N - n(s))^2 E(g'(s, \mathbf{y}) - \theta)^2 \leqq \sum p(s)(N - n(s))^2 E(\bar{y}_s - \theta)^2 . \tag{26}$$

Putting $g'(s, \mathbf{y}) = \bar{y}_s + h'(s, \mathbf{y})$ and noting that $E(\bar{y}_s - \theta)^2 = \sigma^2/n(s)$ we have from (26), after cancelling out the common term,

$$\begin{aligned} &\sum p'(s)(N - n(s))^2 E(h'^2(s, \mathbf{y})) \\ &\quad + 2 \sum p'(s)(N - n(s))^2 E(h'(s, \mathbf{y})(\bar{y}_s - \theta)) + \sigma^2 N^2 \sum \frac{p'(s)}{n(s)} \\ &\quad \leqq \sigma^2 N^2 \left(\frac{1 - f}{m} + \frac{f}{m + 1}\right) . \end{aligned} \tag{27}$$

Clearly

$$\sum \frac{p'(s)}{n(s)} = \sum_{i=1}^{N} \frac{p_i}{i} \tag{28}$$

where $p_i = \sum^{(i)} p'(s)$ where the summation is taken over all those samples s which contain exactly i units. One can easily check that

$$\sum_{i=1}^{N} \frac{p_i}{i} \geqq \frac{1 - f}{m} + \frac{f}{m + 1} . \tag{29}$$

Combining (27), (28) and (29), we have

$$\begin{aligned} &\sum p'(s)(N - n(s))^2 E(h'^2(s, \mathbf{y})) \\ &\quad + 2 \sum p'(s)(N - n(s))^2 E(h'(s, \mathbf{y})(\bar{y}_s - \theta)) \leqq 0 . \end{aligned} \tag{30}$$

The inequality (30) is equivalent to the inequality contained in clause (d) in

Lemma 4.1 and hence using the lemma, it follows that for all s such that $p'(s) > 0$ we have

$$h'(s, \mathbf{y}) = 0 \tag{31}$$

so that for such samples s we get $g'(s, \mathbf{y}) = \bar{y}_s$ and by (16) and (23)

$$e'(s, \mathbf{y}) = e^*(s, \mathbf{y}) \quad \text{if} \quad p'(s) > 0 . \tag{32}$$

Because of (31) the first two terms in the left-hand side of (27) vanish, and hence due to (29) the sign of equality must hold in both (27) and (29), so that the sampling design p' is such that

$$V_{p'}(n(s)) = f(1 - f) . \tag{33}$$

We shall now show that the strict inequality in (18) cannot hold. Since e^* is a linear estimator the mean squared errors of the strategies H' and H are quadratic forms in $y_1, y_2, \cdots, y_N$ given by

$$M(H', \mathbf{y}) = \sum \sum_{i,j=1}^N a'_{ij} y_i y_j$$
$$M(H, \mathbf{y}) = \sum \sum_{i,j=1}^N a_{ij} y_i y_j .$$

Using the first order inclusion probabilities one can easily show that

$$\sum_1^N (a'_{ii} - a_{ii}) = 0 . \tag{34}$$

From (18) and (34) it follows that $a'_{ij} = a_{ij}$ for all i and j. Hence $M(H', \mathbf{y}) = M(H, \mathbf{y})$ for all $\mathbf{y} \in R_N$, and thus the strict inequality in (18) cannot hold. The proof of the theorem is complete.

REMARK 4.4. We get Joshi's Theorem 4.4 by putting $f = 0$ in Theorem 4.5.

REMARK 4.5. Given a strategy H with $\mu(H) < \mu$, one can easily construct another strategy H' with $\mu(H') = \mu$ such that $M(H', \mathbf{y}) \leqq M(H, \mathbf{y})$ for all $\mathbf{y} \in R_N$. Hence it follows that Theorem 4.5 remains valid with $AH(\mu)$ replaced by $AH(\leqq \mu)$ where $AH(\leqq \mu)$ denotes the class of all strategies with expected sample size less than or equal to μ.

REMARK 4.6. If the strategy $H = (p, e^*)$ of Theorem 4.5 is such that $V_p(n(s)) > f(1 - f)$, then it may become inadmissible. We illustrate this with an example: Let p be the design obtained by simple random sampling with replacement with size $n \geqq 3$. Clearly $\pi_i(p) = 1 - ((N - 1)/N)^n$ for $i = 1, 2, \cdots, N$; hence

$$\mu(p) = N\left[1 - \left(\frac{N-1}{N}\right)^n\right] = m + f, \quad \text{say}$$

and

$$V_p(n(s)) > f(1 - f) .$$

Now we show that the strategy (p, e^*) where p is as defined above and e^* is given by (16) is inadmissible in $AH(\mu)$. Consider the strategy (p_1, e^*) where p_1 is the design obtained by simple random sampling without replacement with size m or

$m + 1$ with probabilities $1 - f$ and f respectively. From Ramakrishnan (1969) it follows that (p_1, e^*) is uniformly more efficient than (p, e^*). The author feels that the condition $V_p(n(s)) = f(1 - f)$ is also necessary for the validity of the theorem. For the sake of future use we denote the strategy (p_1, e^*), above, by

$$H_N = H(p_1, N\bar{y}, \mu, N) \tag{35}$$

where μ denotes the expected sample size and N, the size of the population. We are now in a position to prove Theorem 4.2.

PROOF OF THEOREM 4.2. If $\mu = 1$ or N, for any strategy $(p, e) \in LH^*(\mu)$, the design p will be unicluster and hence the HT-estimator $\bar{e}(s, \mathbf{y})$ is the umv estimator in $L^*(p)$. In either case it is seen that $\mathrm{HT}(\mu)$ is complete in $LH^*(\mu)$. This proves the "if" part of the theorem.

In case $1 < \mu < N$ and μ is not an integer it follows as a consequence of Theorem 4.5 that the strategy $H(p_1, N\bar{y}, \mu, N)$ in (35) is admissible in $LH^*(\mu)$. Since the above strategy does not belong to $\mathrm{HT}(\mu)$, it follows that $\mathrm{HT}(\mu)$ is not complete in $LH^*(\mu)$.

The only case left out is: μ an integer, $2 \leqq \mu \leqq N - 1$ and $N \geqq 3$. Choose a number δ such that $0 < \delta < 1$ and

$$1 < \mu_0 = \mu - \delta < N - 1 .$$

From the strategy $H_{N-1} = H(p_1, (N - 1)\bar{y}, \mu_0, N - 1)$ corresponding to the population consisting of the first $(N - 1)$ units, we construct a strategy $H' = (p', e')$ corresponding to the population of N units as follows: Any sample s for p_1 goes into two samples (s, N) and s for p' with probabilities $\delta p_1(s)$ and $(1 - \delta)p_1(s)$ respectively. The estimator e' is defined by

$$e'((s, N), \mathbf{y}) = (N - 1)\bar{y}_s + \frac{y_N}{\delta}$$

and

$$e'(s, y) = (N - 1)\bar{y}_s .$$

Since $\pi_N(p') = \delta$ and $(N - 1)\bar{y}_s$ is unbiased for $\sum_1^{N-1} y_i$ in the design p_1, it is clear that e' is unbiased for $Y (= \sum_1^N y_i)$ in p'. Also since $e' \neq \bar{e}$

$$H' = (p', e') \in LH^*(\mu) - \mathrm{HT}(\mu) .$$

Next we show that given any strategy $H = (p, \bar{e}) \in \mathrm{HT}(\mu)$, there exists a point $\mathbf{y}_0$ (which may depend on H and H') such that

$$V(H', \mathbf{y}_0) < V(H, \mathbf{y}_0)$$

which will show that $\mathrm{HT}(\mu)$ is not complete in $LH^*(\mu)$. If $\pi_N(p) < \delta$, one can easily check that

$$V(H', \mathbf{y}^{(N)}) < V(H, \mathbf{y}^{(N)})$$

where $\mathbf{y}^{(N)} = (0, 0, \cdots, 0, y_N)$ and y_N is any nonzero real number. If $\pi_N(p) \geqq \delta$, and p gives positive probability to the sample consisting of unit N alone, then also it is easy to check that

$$0 = V(H', \mathbf{y}_0) < V(H, \mathbf{y}_0)$$

where $\mathbf{y}_0 = (k, k, \cdots, k, 0)$ and k is any nonzero real number. If $\pi_N \geqq \delta$ and p gives zero probability to the sample consisting of unit N alone, then construct the strategy $H_2 = (p_2, \bar{e})$ corresponding to the population of the first $(N-1)$ units where p_2 is obtained from p by removing unit N from all those samples for p containing it, the probability structure remaining unchanged. Since $\pi_N(p) \geqq \delta$ and $\sum \pi_i(p) = \mu$,

$$\textstyle\sum_1^{N-1} \pi_i(p_2) = \sum_1^{N-1} \pi_i(p) \leqq \mu - \delta = \mu_0 \,.$$

Hence from Remark 4.5, there exists a point $\mathbf{y}_0' = (y'_{01}, \cdots, y'_{0N-1}) \in R_{N-1}$ such that

$$V(H_{N-1}, \mathbf{y}_0') < V(H_2, \mathbf{y}_0') \,. \tag{36}$$

Since

$$V(H_{N-1}, \mathbf{y}_0') = V(H', \mathbf{y}_0) \tag{37}$$

and

$$V(H_2, \mathbf{y}_0') = V(H, \mathbf{y}_0) \tag{38}$$

where $\mathbf{y}_0 = (y'_{01}, y'_{02}, \cdots, y'_{0N-1}, 0)$, we have, on comparing (36), (37) and (38)

$$V(H', \mathbf{y}_0) < V(H, \mathbf{y}_0) \,.$$

The proof of the theorem is complete.

5. Strong admissibility.

After having proved the nonexistence of a best strategy (in the sense of uniformly minimum variance) in $LH^*(\mu)$ and that the complete class of admissible strategies is wider than $\mathrm{HT}(\mu)$ in most of the situations, our next step is to impose further criteria which will give a narrow enough class of strategies. One can easily check that there does not exist a hyperadmissible strategy in $LH^*(\mu)$. In the following we weaken this criterion and characterize the class of all strategies in $LH^*(\mu)$ that satisfy the new criterion.

DEFINITION 5.1. In a class $\mathscr{H}$ of unbiased strategies for Y, a strategy $H_1 \in \mathscr{H}$ is said to be "strongly admissible" in $\mathscr{H}$ if it is admissible in $E_1, E_2, \cdots, E_N$ separately, where $E_r = \bigcup_{i=1}^{\binom{N}{r}} R_i^r$ and R_i^r is the ith phs of dimension r.

The definition of a strongly admissible estimator is straightforward. For the case of a single design, it can be noted that the criteria of strong admissibility and hyperadmissibility are effectively equivalent in arriving at an optimum estimator. While there exists no hyperadmissible strategy, there exist strongly admissible strategies in $LH^*(\mu)$ and we characterize the set of all strongly admissible strategies in $LH^*(\mu)$ in the following

THEOREM 5.1. $\mathrm{HT}(\mu)$ *is precisely the set of all strongly admissible strategies in* $LH^*(\mu)$. *In other words* $\mathrm{HT}(\mu)$ *is complete in* $LH^*(\mu)$ *with respect to strong admissibility.*

PROOF. For any design d, the unique strongly admissible estimator in $L^*(p)$ is given by $\bar{e}(s, \mathbf{y})$, which shows that the set of all strongly admissible strategies in $LH^*(\mu)$ is contained in $\mathrm{HT}(\mu)$. Further, it can be easily noted that, from the proof of Theorem 4.1, in fact follows the strong admissibility of any strategy $H \in \mathrm{HT}(\mu)$ in $LH^*(\mu)$ which proves the theorem.

REMARK 5.1. Let M^* ($\leqq \mu$) be the class of all unbiased strategies with expected sample size $\leqq \mu$, in which the estimator $e(s, \mathbf{y})$ is subject only to the restriction that for every s with $p(s) > 0$, $e(s, y)$ is continuous in the variates y_i at the origin. By the result of Joshi (1971), $\bar{e}$ is the unique strongly admissible estimator in the class of estimators continuous at the origin, from which it follows that Theorem 5.1 holds for M^* ($\leqq \mu$).

The criterion of strong admissibility has some practical implications. For example, in case of estimation of a domain of total (or mean) where the domain size is known (say r) but the domain frame is not available (a unit can be classified into that domain only after surveying it) it is easily seen that the parameter ($\mathbf{y} = (y_1, \cdots, y_N)$ where exactly $N - r$ co-ordinates have fixed zero values) space is given by E_r. So if we start with a strongly admissible estimator, such domain totals (or means) can be admissibly estimated with the same estimator.

Acknowledgment. The author wishes to thank Dr. J. Lanke for many useful suggestions. Thanks are also due to the referee for pointing out changes which considerably improved an earlier version of the paper.

REFERENCES

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

GODAMBE, V. P. (1965). A review of the contributions towards a unified theory of sampling from finite populations. *Rev. Inst. Internat. Statist.* **33** 242–258.

HANURAV, T. V. (1965). Optimum sampling strategies. Ph. D. thesis submitted to the Indian Statistical Institute.

HEGE, V. (1965). Sampling designs which admit uniformly minimum variance unbiased estimates. *Bull. Calcutta Statist. Assoc.* **14** 160–162.

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.

JOSHI, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations—II. *Ann. Math. Statist.* **36** 1723–1729.

JOSHI, V. M. (1966). Admissibility and Bayes estimation in sampling finite populations. *Ann. Math. Statist.* **37** 1658–1670.

JOSHI, V. M. (1971). Hyperadmissibility of estimators for finite populations. *Ann. Math. Statist.* **42** 680–690.

RAMAKRISHNAN, M. K. (1969). Some results on the comparison of sampling with and without replacement. *Sankhyā Ser. A* **31** 333–342.

DEPARTMENT OF STATISTICS AND ECONOMICS
UNIVERSITY OF SINGAPORE
SINGAPORE—10.