ON A CLASS OF UNIFORMLY ADMISSIBLE ESTIMATORS FOR FINITE POPULATIONS

BY RM. SEKKAPPAN AND M. E. THOMPSON

University of Waterloo

Let C' be a class of sampling designs of fixed expected sample size n and fixed inclusion probabilities π_i and C be the subclass of C' consisting of designs of fixed size n and inclusion probabilities π_i . Then it is established that the pair (e^*, p^*) where $p^* \in C$ and $e^*(s, \mathbf{x}) = \sum_{i \in s} b_i x_i$, $b_i > 1$, and $\sum_{i=1}^{N} (b_i)^{-1} = E(n(s)) = n$, is strictly uniformly admissible among pairs (e_i, p_i) where $p_i \in C'$ and e_i is any measurable estimate.

1. Summary and introduction. Admissible estimation in relation to survey-sampling has been studied in great detail by Godambe and Joshi [4] and Joshi [6], [7]. Joshi [7], Godambe [3] and Ericson [1] have established the uniform admissibility of some classes of estimator-design pairs for a finite population total or for a finite population mean. In particular, Joshi [7] showed that the sample mean and a sampling design of fixed sample size n are uniformly admissible for the population mean, when the competing designs have expected sample size not less than n.

In this paper, we consider the pair (\bar{e}, p') , where $\bar{e} = \sum_{i \in s} b_i x_i$ is a linear estimator of the population total. The important Horvitz-Thompson estimator is a special case of \bar{e} . The estimator \bar{e} is shown by Joshi [6] to be admissible (with respect to a given sampling design) when the sampling design is of fixed size n and $\sum_{i=1}^{N} b_i^{-1} = n$, $b_i \geq 1$. It might be supposed that, as a generalization of Joshi's result concerning the sample mean, any pair (\bar{e}, p') would be uniformly admissible among estimator-design pairs with expected sample size n, provided that p' had fixed size n. We are able here to prove this, when the competing estimators are measurable and the competing designs have the same inclusion probabilities as the fixed size design p'. Moreover, we show by an example that the (expected) more general extension of Joshi's result does not hold; the pair (\bar{e}, p') need not be uniformly admissible if the competing designs are allowed to have inclusion probabilities differing from those of p'.

The notation and definitions follow those of [3], [4] and [6]. The population consists of N units denoted by the integers $i = 1, 2, \dots, N$. Let x_i be the real (unknown) value associated with the ith population unit; then $\mathbf{x} = (x_1, \dots, x_N)$ is a vector in the N-dimensional Euclidean space R^N . The population total $T(\mathbf{x})$ is $\sum_{i=1}^N x_i$. Any subset s of the integers $1, 2, \dots, N$ is called a sample. If S is a subset of the set of all possible samples s and p is any real function on S such

The Annals of Statistics. STOR

www.jstor.org

Received January 1973; revised May 1974.

AMS 1970 subject classifications. Primary 62; Secondary D05.

Key words and phrases. Uniform admissibility, finite populations, unequal probability sampling, Horvitz-Thompson estimator.

that $0 < p(s) \le 1$ for all $s \in S$ and $\sum_{S} p(s) = 1$ then (S, p) is called a sampling design. For a sample s, n(s) will denote the number of distinct elements in the sample. A fixed size design of size n is one for which $s \in S$ implies n(s) = n. Any real function $e(s, \mathbf{x})$ on the product space $S \times R^N$, such that e depends on \mathbf{x} only through those x_i for which $i \in s$, is called an estimator of $T(\mathbf{x})$.

DEFINITION 1.1. A pair (e', p') consisting of an estimator e' and a sampling design (S', p') is said to be uniformly superior to another pair (e, p) if for all $\mathbf{x} \in \mathbb{R}^N$

(1)
$$\sum_{s'} p'(s)[e'(s, \mathbf{x}) - T(\mathbf{x})]^2 \leq \sum_{s} p(s)[e(s, \mathbf{x}) - T(\mathbf{x})]^2,$$

strict inequality holding for at least one x; here S' and S denote the sets of all possible samples s for which $p'(s) \neq 0$ and $p(s) \neq 0$ respectively.

DEFINITION 1.2. With respect to a class C^* of sampling designs of fixed expected sample size, a class D of estimators and a squared error loss function, a pair (e', p') of an estimator $e' \in D$ and a sampling design $(S', p') \in C^*$ is said to be uniformly admissible if no other pair (e, p) such that $(S, p) \in C^*$ and $e \in D$ is uniformly superior to (e', p').

2. Uniform admissibility of the estimate. We shall prove the uniform admissibility for a general class of estimates of which the Horvitz-Thompson estimate (H.T. estimate for short) is a particular case. As an intermediate step in the development of the argument, we prove a weaker type of uniform admissibility.

DEFINITION 2.1. A pair (e, p) of an estimator $e \in D$ and a sampling design $(S, p) \in C^*$ is said to be weakly uniformly admissible if there exists no other pair (e', p'), where $e' \in D$ and $(S', p') \in C^*$, such that the inequality (1) is satisfied for almost all (Lebesgue measure) $\mathbf{x} \in R^N$, the strict inequality in (1) holding on a non-null subset of R^N .

To distinguish the uniform admissibility defined by Definition 1.2 from weak uniform admissibility the former will be called strict uniform admissibility. The argument is then completed by proving that weak uniform admissibility implies strict uniform admissibility.

THEOREM 2.1. Consider the class of sampling designs

(2)
$$C' = \begin{cases} p: & \text{(i)} \quad \sum_{s} n(s)p(s) = n, \\ & \text{(ii)} \quad \sum_{s \ni i} p(s) = \pi_i, \end{cases}$$

 π_i being the inclusion probability for the individual i, and let C be the subclass of C' consisting of fixed size designs of size n and inclusion probabilities π_i . Let $e^*(s, \mathbf{x})$ be an estimate given by

$$e^*(s, \mathbf{x}) = \sum_{i \in s} b_i x_i,$$

where the coefficients b_i satisfy

(4) (i)
$$b_i > 1$$
 $i = 1, 2, \dots, N$ and (ii) $\sum_{i=1}^{N} (b_i)^{-1} = E(n(s)) = n$.

Then the pair (e^*, p^*) where p^* belongs to C is weakly uniformly admissible among pairs (e_1, p_1) where $p_1 \in C'$ and e_1 is any measurable estimate.

PROOF. If possible, let (e_1, p_1) be a pair satisfying

(5)
$$\sum_{s_1} p_1(s) [e_1(s, \mathbf{x}) - T(\mathbf{x})]^2 \leq \sum_{s_1} p_1(s) [e_1(s, \mathbf{x}) - T(\mathbf{x})]^2$$

for almost all $x \in R^N$ (Lebesgue measure), with strict inequality holding for a non-null (Lebesgue) set in R^N . Putting

(6)
$$x_i = \frac{y_i}{b_i}, \qquad i = 1, 2, \dots, N,$$

(7)
$$f(s, \mathbf{y}) = e_1(s, \mathbf{y}) - \sum_{i \in s} \frac{y_i}{b_i},$$

(8)
$$A(s) = \sum_{i \in s} \frac{1}{b_i} = n - \sum_{i \in s} \frac{1}{b_i},$$

(9)
$$\bar{y}(s) = \left[\sum_{i \in s} \left(1 - \frac{1}{b_i} \right) y_i \right] / B(s) ,$$

where

$$B(s) = \sum_{i \in s} \left(1 - \frac{1}{b_i}\right),$$

in (5), we have

(11)
$$\sum_{s_1} p_i(s) \left[f(s, \mathbf{y}) - \sum_{i \in s} \frac{y_i}{b_i} \right]^2 \leq \sum_{s^*} p^*(s) \left[A(s) \bar{y}(s) - \sum_{i \in s} \frac{y_i}{b_i} \right]^2.$$

(We note that B(s) = A(s) in the R.H.S. of (11).)

We now take expectations of both sides of (11) with respect to a prior distribution on R^N , under which each variate y_i $i=1,2,\dots,N$ is distributed independently and normally with common mean θ and variance $\sigma_i^2 = K/(1-1/b_i)$, K>0; and we get

(12)
$$\sum_{s_1} p_1(s) E \left[f(s, \mathbf{y}) - A(s)\theta - \sum_{i \notin s} \frac{y_i - \theta}{b_i} \right]^2$$

$$\leq \sum_{s^*} p^*(s) E \left[A(s) \bar{y}(s) - A(s)\theta - \sum_{i \notin s} \frac{y_i - \theta}{b_i} \right]^2.$$

Simplifying (12) we have

(13)
$$\sum_{s_1} p_1(s) A^2(s) E[g(s, \mathbf{y}) - \theta]^2 \leq \sum_{s^*} p^*(s) A^2(s) E[\bar{y}(s) - \theta]^2,$$
 where

(14)
$$g(s, \mathbf{y}) = \frac{1}{A(s)} f(s, \mathbf{y}).$$

It is easily seen that for $\bar{y}(s)$ in (9)

(15)
$$E[\bar{y}(s) - \theta]^2 = \frac{K}{B(s)}.$$

Now inserting

(16)
$$g(s, \mathbf{y}) = \bar{y}(s) + h(s, \mathbf{y})$$

in (13), we get

(17)
$$\sum_{s_1} p_1(s) A^2(s) E[h^2(s, \mathbf{y})] + 2 \sum_{s_1} p_1(s) A^2(s) E[h(s, \mathbf{y})(\bar{y}(s) - \theta)] + \sum_{s_1} p_1(s) A^2(s) E[\bar{y}(s) - \theta]^2 \leq \sum_{s^*} p^*(s) A^2(s) E[\bar{y}(s) - \theta]^2.$$

By (8), (10) and (15) we have

(18)
$$\sum_{S_1} p_1(s) A^2(s) E[\bar{y}(s) - \theta]^2 = K \sum_{S_1} p_1(s) \left[n - \sum_{i \in s} \frac{1}{b_i} \right]^2 / B(s) = K \sum_{S_2} p_1(s) (n - n(s))^2 / B(s) + \sum_{S_3^*} p^*(s) A^2(s) E[\bar{y}(s) - \theta]^2.$$

Substituting (18) in (17) and noting that the first term in the R.H.S. of (18) is positive, we get by cancelling out the common term

(19)
$$\sum_{s_1} p_1(s) A^2(s) E[h^2(s, \mathbf{y})] + 2 \sum_{s_1} p_1(s) A^2(s) E[h(s, \mathbf{y})(\bar{y}(s) - \theta)] \leq 0.$$

Adding the term $\sum_{s_1} p_1(s) A^2(s) E[\bar{y}(s) - \theta]^2$ to both sides of (19), we get

(20)
$$\sum_{s_1} p_1(s) A^2(s) E[g(s, \mathbf{y}) - \theta]^2 \leq \sum_{s_1} p_1(s) A^2(s) E[\bar{y}(s) - \theta]^2.$$

We may now proceed using the Cramér-Rao inequality in the manner of Joshi [6], or apply the following argument. The above inequality simplifies to

(21)
$$\sum_{s_1} p_1(s) A^2(s) E[g(s, \mathbf{y}) - \bar{y}(s)]^2 \\ \leq 2 \sum_{s_2} p_1(s) A^2(s) E[(\bar{y}(s) - g(s, \mathbf{y}))(\bar{y}(s) - \theta)].$$

Using (20) and (15) we get, by Schwarz's inequality,

(22)
$$\sum_{s_1} p_1(s) A^2(s) E[|(\bar{y}(s) - g(s, \mathbf{y}))(\bar{y}(s) - \theta)|] \le 2K \sum_{s_1} p_1(s) A^2(s) / B(s).$$
 Clearly

(23)
$$T(\theta) \leq 4K \sum_{s_1} p_1(s) A^2(s) / B(s) ,$$

where

(24)
$$T(\theta) = \sum_{s_1} p_1(s) A^2(s) E[g(s, \mathbf{y}) - \bar{y}(s)]^2$$

and K is a constant. Thus $T(\theta)$ is a bounded function of θ . Let $-\infty < a < 0 < b < +\infty$. Integrating both sides of (21) with respect to θ , we get

(25)
$$\int_{a}^{b} T(\theta) d\theta \leq 2 \int_{a}^{b} \sum_{s_{1}} p_{1}(s) A^{2}(s) E[(\bar{y}(s) - g(s, \mathbf{y}))(\bar{y}(s) - \theta)] d\theta \\
= 2 \sum_{s_{1}} p_{1}(s) A^{2}(s) \int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} (\bar{y}(s) - g(s, \mathbf{y})) \\
\times \left[\int_{a}^{b} (\bar{y}(s) - \theta) \left[\prod_{i \in s} \frac{1}{(2\pi\sigma_{i}^{2})^{\frac{1}{2}}} \right] \\
\times \exp\left[-\frac{1}{2} \sum_{i \in s} \left[\frac{y_{i} - \theta}{\sigma_{i}} \right]^{2} \right] d\theta \right] \prod_{i \in s} dy_{i}.$$

After simplification, the R.H.S. of (25) works out to

(26)
$$2 \sum_{s_1} p_1(s) A^2(s) C(s) \prod_{i \in s} \frac{1}{(2\pi\sigma_i^2)^{\frac{1}{2}}} \int \cdots \int_{\mathbb{R}^n} (\bar{y}(s) - g(s, \mathbf{y})) \times \left[\exp\left[-\frac{1}{2} \sum_{i \in s} \left[\frac{y_i - \dot{b}}{\sigma_i}\right]^2\right] - \exp\left[-\frac{1}{2} \sum_{i \in s} \left[\frac{y_i - a}{\sigma_i}\right]^2\right] \right] \prod_{i \in s} dy_i$$

where $C(s) = 1/(\sum_{i \in s} 1/\sigma_i^2) \le M$ for all s.

Replacing C(s) by N and applying Schwarz's inequality to each term separately in (26) tells us that

(27)
$$\int_a^b T(\theta) d\theta \leq 2M(\sum_{s, p_1(s)} A^2(s))^{\frac{1}{2}} ((T(b))^{\frac{1}{2}} + (T(a))^{\frac{1}{2}}).$$

From (23) and (27), it follows that

(28)
$$\int_a^b T(\theta) d\theta \text{ is bounded.}$$

If for $\varepsilon > 0$ there is a b_0 such that for all $b \ge b_0$, $(T(b))^{\frac{1}{2}} \ge \varepsilon$, then

$$\int_a^b T(\theta) d\theta \ge \int_a^b T(\theta) d\theta \to \infty$$
 as $b \to \infty$,

which is impossible on account of (28). Hence $\liminf_{b\to\infty} T(b) = 0$. Similarly $\liminf_{a\to\infty} T(a) = 0$.

These two together imply that $\int_a^b T(\theta) d\theta = 0$ and hence

(29)
$$g(s, \mathbf{y}) = \bar{y}(s)$$
 a.e. for all $s \in S_1$.

Now it follows from (16) that

(30)
$$h(s, \mathbf{y}) = 0$$
 a.e. for all $s \in S_1$.

Combining (30), (17) and (18), we have that

$$n(s) = n$$
 for all $s \in S_1$.

Hence by (14), (7), (9) and (29)

$$e_1(s, \mathbf{x}) = e^*(s, \mathbf{x})$$
 a.e. in R^N for any $s \in S_1$.

This leads to

(31)
$$\sum_{s, p_1(s)} [e^*(s, \mathbf{x}) - T(\mathbf{x})]^2 \leq \sum_{s^*} p^*(s) [e^*(s, \mathbf{x}) - T(\mathbf{x})]^2$$
 a.e. in \mathbb{R}^N .

By (18) it follows that the strict inequality does not hold in (31) on a non-null set either. The theorem is thus proved.

3. Strict uniform admissibility. The method of this proof utilizes Theorem 5.2 of Joshi [6].

THEOREM 3.1. Weak uniform admissibility of the pair (e^*, p^*) in Theorem 2.1 implies its strict uniform admissibility.

PROOF. Suppose

(32)
$$\sum_{s, p_1(s)} [e_1(s, \mathbf{x}) - T(\mathbf{x})]^2 \leq \sum_{s, p} [e^*(s, \mathbf{x}) - T(\mathbf{x})]^2$$

for all $x \in R^N$ with strict inequality for at least one $x \in R^N$. By (31), we have

$$\sum_{s_1} p_1(s) [e^*(s, \mathbf{x}) - T(\mathbf{x})]^2 \le \sum_{s_1} p^*(s) [e^*(s, \mathbf{x}) - T(\mathbf{x})]^2$$

a.e. in \mathbb{R}^N . Using (2), (3) and (6), the computation of (31) leads to

(33)
$$2 \sum_{i \le j}^{N} y_i y_i (\pi_{ij} - \pi_{ij}^*) \le 0$$
 a.e. in R^N ,

where $\pi_{ij} = \sum_{s \ni i,j} p_i(s)$ and $\pi_{ij}^* = \sum_{s \ni i,j} p^*(s)$.

Moreover, from continuity considerations (33) must be true everywhere in \mathbb{R}^N . Consider the inequality (33) at points $\mathbf{y} \in \mathbb{R}^N$, at which only two particular coordinates y_i and y_j differ from zero; then we have

$$(34) y_i y_j \pi_{ij} \leq y_i y_j \pi_{ij}^* in R^N$$

Since (34) holds for both positive and negative values of the product $y_i y_j$, we get

(35)
$$\pi_{ij} = \pi_{ij}^* \quad \text{for each pair} \quad i, j.$$

Hence

(36)
$$\sum_{s_1} p_1(s) (e^*(s, \mathbf{x}) - T(\mathbf{x}))^2 = \sum_{s_1} p^*(s) (e^*(s, \mathbf{x}) - T(\mathbf{x}))^2$$

for all $\mathbf{x} \in \mathbb{R}^N$.

Using (36) in (32) we have

(37)
$$\sum_{s_1} p_1(s) (e_1(s, \mathbf{x}) - T(\mathbf{x}))^2 \leq \sum_{s_1} p_1(s) (e^*(s, \mathbf{x}) - (T(\mathbf{x}))^2$$

for all $x \in R^N$, with strict inequality holding for at least one $x \in R^N$. By Theorem 5.2 of Joshi [6], we have $e_1(s, x) = e^*(s, x)$ for all $x \in R^N$, $s \in S_1$. Thus the proof of the theorem is complete.

COROLLARY. The pair (\bar{e}, p^*) , where $p^* \in C$ and \bar{e} is the H.T. estimate, is strictly uniformly admissible in the class of all pairs (e, p) with e in the class of all measurable estimates and $p \in C'$.

PROOF. We may take $b_i = 1/\pi_i$ in (3) and (4). This gives $e^*(s, \mathbf{x}) = \sum_{i \in s} b_i x_i = \sum_{i \in s} x_i/\pi_i$, which is the H.T. estimator.

- 4. Possible extensions of the main result. Upon examination of the conditions of Theorem 2.1, several questions, such as the following, naturally arise.
 - Q.1: Need the design (S^*, p^*) have fixed sample size?
- Q.2: Is it necessary that the inclusion probabilities of competing designs should coincide with those of (S^*, p^*) ?
 - Q.3: Should the same results hold with fewer restrictions on the b_i ?
 - Q.4: Is the measurability restriction on competing estimators necessary?
- Q.1 has been answered affirmatively by Godambe and Joshi [4], who show that if e^* is the Horvitz-Thompson estimator corresponding to p^* , and if (S^*, p^*) has varying sample size, then except in trivial situations one can find an estimator which has smaller mean square error than e^* for all x, without altering the design.

Q.2 can also be answered affirmatively, in a somewhat weaker sense. For suppose that the population consists of just two units 1, 2. Let $S^* = \{\{1\}, \{2\}\}\}$, so that a sampling design defined on S^* may be written $p^* = \{p_1^*, p_2^*\}$. Then the mean square error of e^* with respect to p^* is

$$M(e^*, p^*) = p_1^* [b_1 x_1 - (x_1 + x_2)]^2 + p_2^* [b_2 x_2 - (x_1 + x_2)]^2.$$

Setting $x_i = y_i/b_i$ and noting that $1/b_1 + 1/b_2 = 1$, we may write

$$M(e^*, p^*) = \left[\frac{p_1^*}{b_2^2} + \frac{p_2^*}{b_1^2}\right] (y_1 - y_2)^2.$$

Case (i). Suppose $b_1 = b_2$. Then whatever p_1^* , p_2^* we choose, the mean square error is the same for all y in \mathbb{R}^2 , and in fact Joshi's result of [7] guarantees that (e^*, p^*) is uniformly admissible.

Case (ii). When the b_i 's are different and p_1^* , $p_2^* > 0$ we can choose another design $p = \{p_1, p_2\}$ so that $M(e^*, p) < M(e^*, p^*)$ for all $\mathbf{y} \in R^2$. For example, suppose $b_1 = 3$, $b_2 = \frac{3}{2}$, $p_1^* = p_2^* = \frac{1}{2}$, $p_1 = \frac{1}{4}$, $p_2 = \frac{3}{4}$. Then $M(e^*, p) < M(e^*, p^*)$. Thus for N = 2, $b_1 \neq b_2$, the pair (e^*, p^*) is not uniformly admissible if competing designs are allowed to have inclusion probabilities different from those of (S^*, p^*) . Of course, it may be conjectured that if the competing design is required to have inclusion probabilities not less than those of (S^*, p^*) , the pair (e^*, p^*) is still uniformly admissible.

With regard to Q.3, we may note that Godambe [3] has shown the following result. With respect to the class $C_n = \{p: \sum_S n(s)p(s) = n\}$ of sampling designs and all estimators the pair (e^*, p^*) is uniformly admissible, where e^* is the estimator given by

$$e^*(s, \mathbf{x}) = \sum_{i \in s} x_i + \sum_{i \notin s} \lambda_i$$

 $\lambda_1, \dots, \lambda_N$ being any fixed numbers. Moreover, Ericson [1] has proved that with respect to the class of designs $D_n = \{p : n(s) \neq n \Rightarrow p(s) = 0\}$ and all estimators the pair (e^*, p^*) is uniformly admissible where

$$e^*(s, \mathbf{x}) = \alpha_n \sum_{i \in s} x_i + \beta$$

 α_n and β being fixed values such that $1 < \alpha_n < N/n$. These results suggest that it might be possible to establish uniform admissibility for (e^*, p^*) if $b_i \ge 1$ for all i and $\sum_{i=1}^N 1/b_i \ge n$, and this uniform admissibility has in fact been established by Sekkappan [9] when $b_i > 1$ and $\sum_{i=1}^N 1/b_i > n$. The proof involves showing that e^* is a Bayes estimator with finite Bayes risk.

Finally, in connection with Q.4, it is easy to see that the uniform admissibility of the estimate e^* in Theorem 2.1 can be proved directly by an algebraic method when the sample size n is equal to one, so that no measurability restriction applies. It seems likely that the result may be true also when the sample size is greater than one.

Acknowledgment. The authors wish to thank Professor V. P. Godambe for helpful discussions during the preparation of this paper.

REFERENCES

- [1] Ericson, W. A. (1970). On a class of uniformly admissible estimators of a finite population total. *Ann. Math. Statist.* 41 1369-1372.
- [2] GODAMBE, V. P. (1960). An admissible estimate for any sampling design. Sankhyā 22 285-288.
- [3] GODAMBE, V. P. (1969). Admissibility and Bayes estimation in sampling finite populations— V. Ann. Math. Statist. 40 672-676.
- [4] GODAMBE, V. P. and JOSHI, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations—I. Ann. Math. Statist. 36 1707-1723.
- [5] HODGES, J. L. JR. and LEHMANN, E. L. (1951). Some applications of the Cramér-Rao inequality. Proc. Second Berkeley Symp. Math. Statist. Prob. 13-22. Univ. of California Press.
- [6] Joshi, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations— II, III. Ann. Math. Statist. 36 1723-1742.
- [7] Joshi, V. M. (1966). Admissibility and Bayes estimation in sampling finite populations— IV. Ann. Math. Statist. 37 1658-1670.
- [8] Joshi, V. M. (1969). Admissibility of estimates of the mean of a finite population. New Developments in Survey-Sampling. Wiley, New York. 188-212.
- [9] Sekkappan, Rm. (1973). Bayes estimation and uniform admissibility for sampling from finite populations. Thesis, Univ. of Waterloo.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO, CANADA