The Annals of Statistics
1973, Vol. 1, No. 3, 577-579

AN ALTERNATIVE PROOF OF THE ADMISSIBILITY OF THE HORVITZ-THOMPSON ESTIMATOR

BY M. K. RAMAKRISHNAN[1]

University of Sheffield

Godambe and Joshi (1965) have shown that for any design the Horvitz-Thompson estimator is admissible in the class of all unbiased estimators of a finite population total. In this note we give a simple alternative proof of this result using induction on the size of the population.

Let $U(N)$ denote the population consisting of the N units $1, 2, \cdots, N$. A subset of integers from $U(N)$ is called a sample and a nonnegative function p on the set S of all possible samples such that $\sum_s p(s) = 1$ is called a (sampling) design on $U(N)$. 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 the N-dimensional Euclidean space R_N which is assumed to be the parameter space. The conventional problem in survey sampling is to estimate the population total $Y(N) = \sum y_i$ by observing the values of those y_i for which $i \in s$ where s is a sample drawn according to a design p.

An estimator e is a real-valued function on $S \times R_N$ such that $e(s, \mathbf{y})$ depends on $\mathbf{y}$ 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 $i \in s$. For any design p for which $\pi_i = \sum_{s \ni i} p(s) > 0$, $i = 1, 2, \cdots, N$, the Horvitz-Thompson estimator of $Y(N)$ is defined by

$$\bar{e}(s, \mathbf{y}) = \sum_{i \in s} y_i/\pi_i . \tag{1}$$

In a class D of unbiased estimators of $Y(N)$, an estimator $e_0 \in D$ is said to be admissible in D if for no other estimator $e \in D$

$$V(e, \mathbf{y}) \leqq V(e_0, \mathbf{y}) \tag{2}$$

for all $\mathbf{y} \in R_N$. We denote by $A_N(p)$ the class of all unbiased estimators of $Y(N)$ for the design p. Godambe and Joshi (1965) have shown that the estimator $\bar{e}(s, \mathbf{y})$ in (1) is admissible in $A_N(p)$ for any design p for which it is defined. We give an alternative proof of this result.

THEOREM. *For any design p for which $\pi_i > 0$, $i = 1, 2, \cdots, N$ the estimator $\bar{e}(s, \mathbf{y})$ is admissible in $A_N(p)$.*

PROOF. We propose to prove the theorem using induction on N. For $N = 1$ the theorem is evident since $0 = V(\bar{e}, \mathbf{0}) < V(e, \mathbf{0})$ for any other $e \in A_1(p)$. We now assume that the theorem is true for N, that is, for any design p on $U(N)$ for which $\pi_i > 0$, $i = 1, 2, \cdots, N$ the estimator $\bar{e}(s, \mathbf{y})$ is admissible in $A_N(p)$. To show that the theorem holds for $N + 1$, we will show that for any given design p^* on $U(N + 1)$ for which $\pi_i^* = \pi_i(p^*) > 0$, $i = 1, 2, \cdots, N + 1$,

$$e^*(s^*, \mathbf{y}^*) \in A_{N+1}(p^*)$$

Received February 1972; revised August 1972.

[1] Present address: University of Singapore, Singapore 10.

and

(3) $$V(e^*, \mathbf{y}^*) \leqq V(\bar{e}^*, \mathbf{y}^*) \qquad \text{for all} \quad \mathbf{y}^* \in R_{N+1}$$

imply

$$e^*(s^*, \mathbf{y}^*) = \bar{e}^*(s^*, \mathbf{y}^*) \qquad \text{for all} \quad s^* \in S^*$$

such that $p^*(s^*) > 0$ and $\mathbf{y}^* \in R_{N+1}$ (in this note * will always correspond to something concerning the population $U(N+1)$). Consider the population $U(N)$ obtained by removing unit $N+1$ from $U(N+1)$. From p^* we construct a design p on $U(N)$ as follows:

(4) $$p(s) = p^*(s) + p^*((s, N+1))$$

where $(s, N+1)$ denotes the sample consisting of the units belonging to s and the unit $N+1$. Define the estimator

(5) $$e(s, \mathbf{y}) = (p^*(s)e^*(s, \mathbf{y}^*) + p^*((s, N+1))e^*((s, N+1), \mathbf{y}^*))/p(s)$$

where

(6) $$\mathbf{y} = (y_1, y_2, \cdots, y_N), \qquad \mathbf{y}^* = (y_1, y_2, \cdots, y_N, 0).$$

Since $e^* \in A_{N+1}(p^*)$ it is clear that $e \in A_N(p)$. Further

(7) $$V(\bar{e}, \mathbf{y}) = V(\bar{e}^*, \mathbf{y}^*)$$

and

(8) $$V(e, \mathbf{y}) \leqq V(e^*, \mathbf{y}^*)$$

where $\mathbf{y}$ and $\mathbf{y}^*$ are as in (6), (7) is obvious. To prove (8) it is enough to show

$$\sum e^{*2}(s^*, \mathbf{y}^*)p^*(s^*) \geqq \sum e^2(s, \mathbf{y})p(s)$$

that is

$$\sum (e^{*2}(s, \mathbf{y}^*)p^*(s) + e^{*2}((s, N+1), \mathbf{y}^*)p^*((s, N+1))) \geqq \sum e^2(s, \mathbf{y})p(s).$$

One can easily check using Cauchy's inequality and the definition of e in (5) that

$$e^{*2}(s, \mathbf{y}^*)p^*(s) + e^{*2}((s, N+1), \mathbf{y}^*)p^*((s, N+1)) \geqq e^2(s, \mathbf{y})p(s)$$

which will prove (8).

From (3), (7) and (8) we have

$$V(e, \mathbf{y}) \leqq V(\bar{e}, \mathbf{y})$$

for all $\mathbf{y} \in R_N$ with strict inequality for some $\mathbf{y}$ unless

$$e^*(s, \mathbf{y}^*) = e^*((s, N+1), \mathbf{y}^*)$$

for all $\mathbf{y}^*$ of the form (6) and for all s with $p^*(s) > 0$, $p^*(s, N+1)) > 0$. The induction hypothesis now gives

(9) $$e(s, \mathbf{y}) = \bar{e}(s, \mathbf{y})$$

for all s such that $p(s) > 0$ and $\mathbf{y}$, and

(10) $$e^*(s, \mathbf{y}^*) = e^*((s, N+1), \mathbf{y}^*)$$

for all $\mathbf{y}^*$ of the form (6) and for all s with $p^*(s) > 0$, $p^*((s, N+1)) > 0$. Clearly (9) and (10) imply

$$e^*(s^*, \mathbf{y}^*) = \bar{e}^*(s^*, \mathbf{y}^*)$$

for every sample s^* with $p^*(s^*) > 0$ not containing the unit $N + 1$ and every $\mathbf{y}^* \in R_{N+1}$. Also note that if $p^*(s, N+1) = 0$ then $e^*(s, \mathbf{y}^*) = e(s, \mathbf{y}) = \bar{e}(s, \mathbf{y})$. Since we could have removed any unit j instead of $N + 1$, it follows that

$$e^*(s^*, \mathbf{y}^*) = \bar{e}^*(s^*, \mathbf{y}^*) \tag{11}$$

for all $s^* \in S^*$ and $\mathbf{y}^* \in R_{N+1}$ except perhaps for the sample $s_0^* = (1, 2, \cdots, N+1)$. From (11) and noting that e^* and $\bar{e}^*$ have the same expectation, it follows that $e^*(s_0^*, \mathbf{y}^*) = \bar{e}^*(s_0^*, \mathbf{y}^*)$. The proof of the theorem is complete.

The author feels that most of the difficult proofs due to Joshi (e.g. 1965 and 1966) on the admissibility of various estimators in survey sampling can be simplified using the above approach.

Acknowledgment. The author wishes to thank Dr. J. Lanke for his useful suggestions and the referee for pointing out an error in the original version of the paper.

REFERENCES

GODAMBE, V. P. and JOSHI, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations I. *Ann. Math. Statist.* **36** 1707–1722.

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 IV. *Ann. Math. Statist.* **37** 1658–1670.

DEPARTMENT OF PROBABILITY AND STATISTICS
THE UNIVERSITY
SHEFFIELD S3 7RH
ENGLAND