The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 839–841

ADMISSIBILITY OF ARBITRARY ESTIMATES AS UNBIASED ESTIMATES OF THEIR EXPECTATIONS IN A FINITE POPULATION

BY V. M. JOSHI

Secretary, Maharashtra Government, Bombay

0. Summary. A conjecture of Hanurav (1968), that subject only to a mild restriction, any arbitrary estimate is admissible in the class of all unbiased estimates of its expectation (which is a population function) is shown to be false.

1. Preliminary. $U = \{u_1, u_2, \cdots, u_N\}$ denotes a finite population. A sample s means any non-empty subset of U. S denotes the set of all possible samples s. A sampling design d is determined by defining on S a probability p;

$$p(s) \geqq 0 \qquad \text{for all} \quad s \in S, \quad \text{and} \tag{1}$$

$$\textstyle\sum_{s \in S} p(s) = 1.$$

With each unit u_i, $i = 1, 2, \cdots, N$, is associated a variate value x_i. $\mathbf{x} = (x_1, x_2, \cdots, x_N)$ denotes a point in the N-dimensional space R_N. Estimates and their admissibility in the unbiased class are defined as follows.

DEFINITION 1.1. An estimate e is a function defined on SXR_N, such that for any $s \in S$, e depends on $\mathbf{x}$ through only those x_i for which $u_i \in s$.

DEFINITION 1.2. For a given sampling design d, an estimate $e = \{e(s, \mathbf{x})\}$ is an unbiased estimate of a population function $V(\mathbf{x})$, if

$$\textstyle\sum_{s \in S} p(s) \cdot e(s, \mathbf{x}) = V(\mathbf{x}), \qquad \text{for all} \quad \mathbf{x} \in R_N. \tag{2}$$

DEFINITION 1.3. For a given sampling design d, an unbiased estimate e of a function $V(\mathbf{x})$ is admissible in the class of all unbiased estimates of $V(\mathbf{x})$, if there does not exist any other unbiased estimate $e_1(s, \mathbf{x})$ of $V(\mathbf{x})$, such that

$$\textstyle\sum_{s \in S} p(s)[e_1(s, \mathbf{x}) - V(\mathbf{x})]^2 \leqq \sum_{s \in S} p(s)[e(s, \mathbf{x}) - V(\mathbf{x})]^2 \tag{3}$$

for all $\mathbf{x} \in R_N$ and the strict inequality holds in (3) for at least one $\mathbf{x} \in R_N$.

2. A conjecture regarding admissibility in the unbiased class. In the following we take the sampling design d as fixed. For convenience, let $\bar{S}$ denote the subset of S consisting of all those samples s for which $p(s) > 0$. Let $g = g(s, \mathbf{x})$ be any arbitrary estimate (Definition 1.1). Then g is an unbiased estimate of the population function

$$G(\mathbf{x}) = \textstyle\sum_{s \in \bar{S}} p(s) g(s, \mathbf{x}). \tag{4}$$

Received February 11, 1969; revised February 11, 1970.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics*. www.jstor.org

Hanurav (1968) has expressed the conjecture, that if for each $s \in \bar{S}$, $g(s, \mathbf{x})$ depends on $(s, \mathbf{x})$ fully, i.e. $g(s, \mathbf{x})$ is not independent of any x_i, $i \in s$, then g is admissible in the class of all unbiased estimates of $G(\mathbf{x})$. We shall show this conjecture to be false by constructing a counter-example.

Let $h = h(s, \mathbf{x})$ be an arbitrary estimate (Definition 1.1) such that

$$\sum_{s \in \bar{S}} p(s)h(s, \mathbf{x}) = 0 \qquad \text{for all} \quad \mathbf{x} \in R_N, \text{ and} \tag{5}$$

$$h(s, \mathbf{x}) \neq 0 \qquad \text{for at least one } s \in \bar{S}, \text{ for at least one } \mathbf{x} \in R_N. \tag{6}$$

Let $K = \{K(s, \mathbf{x})\}$ be any arbitrary nonnegative estimate, i.e.

$$K(s, \mathbf{x}) \geqq 0 \qquad \text{for every } s \in \bar{S}, \text{ and every } \mathbf{x} \in R_N. \tag{7}$$

Put

$$g(s, \mathbf{x}) = K(s, \mathbf{x}) \cdot h(s, \mathbf{x}), \qquad \text{and} \tag{8}$$

$$e(s, \mathbf{x}) = g(s, \mathbf{x}) + \alpha h(s, \mathbf{x}), \quad \text{where } \alpha > 0 \text{ is a constant.} \tag{9}$$

Then both $g(s, \mathbf{x})$ and $e(s, \mathbf{x})$ are unbiased estimates of $G(\mathbf{x})$ defined by (4).

The estimate $e = \{e(s, \mathbf{x})\}$ is inadmissible for $G(\mathbf{x})$ because by (9),

$$\begin{aligned}
&\sum_{s \in \bar{S}} p(s)[e(s, \mathbf{x}) - G(\mathbf{x})]^2 - \sum_{s \in \bar{S}} p(s)[g(s, \mathbf{x}) - G(\mathbf{x})]^2 \\
&\quad = \alpha^2 \sum_{s \in \bar{S}} p(s)h^2(s, \mathbf{x}) + 2\alpha \sum_{s \in \bar{S}} p(s)h(s, \mathbf{x})[g(s, \mathbf{x}) - G(\mathbf{x})] \\
&\quad = \alpha^2 \sum_{s \in \bar{S}} p(s)h^2(s, \mathbf{x}) + 2\alpha \sum_{s \in \bar{S}} p(s)h(s, \mathbf{x})g(s, \mathbf{x}), && \text{by (5)} \\
&\quad = \alpha^2 \sum_{s \in \bar{S}} p(s)h^2(s, \mathbf{x}) + 2\alpha \sum_{s \in \bar{S}} p(s)h^2(s, \mathbf{x})K(s, \mathbf{x}) && \text{by (8)} \\
&\quad \geqq \alpha^2 \sum_{s \in \bar{S}} p(s)h^2(s, \mathbf{x}) && \text{by (7),} \\
&\quad \geqq 0;
\end{aligned} \tag{10}$$

and by (6) the strict inequality holds in the extreme right-hand side of (10) for at least one $\mathbf{x} \in R_N$.

The estimate $e(s, \mathbf{x})$ in (9) is therefore inadmissible.

Since $h(s, \mathbf{x})$ and $K(s, \mathbf{x})$ are arbitrary subject only to the restrictions in (5), (6) and (7), they can be so chosen that for each s, $e(s, \mathbf{x})$ depends on all x_i, $i \in s$, and further so that $G(\mathbf{x})$ in (4) is not a mere constant but involves all the x_i, $i = 1, 2, \cdots, N$. The following is a simple illustration.

The population consists of three units, $U = \{u_1, u_2, u_3\}$; the sampling design assigns positive probabilities to only three samples $s_1 = (u_1, u_2)$; $s_2 = (u_2, u_3)$ and $s_3 = (u_3, u_1)$; $p(s_1) = p(s_2) = p(s_3) = \frac{1}{3}$.

Let $h(s_1, \mathbf{x}) = x_1 - x_2$; $h(s_2, \mathbf{x}) = x_2 - x_3$; and $h(s_3, \mathbf{x}) = x_3 - x_1$, so that $\sum_{s \in \bar{S}} p(s)h(s, \mathbf{x}) \equiv 0$ for all $\mathbf{x} \in R_3$.

Putting in (8),

$$\begin{aligned} K(s, \mathbf{x}) &= 1 && \text{if} \quad h(s, \mathbf{x}) > 0 \\ &= 0 && \text{if} \quad h(s, \mathbf{x}) \leqq 0, \end{aligned}$$

we have

$$\begin{aligned} g(s_1, \mathbf{x}) &= x_1 - x_2 && \text{if} \quad x_1 > x_2, \\ &= 0 && \text{if} \quad x_1 \leqq x_2; \\ g(s_2, \mathbf{x}) &= x_2 - x_3 && \text{if} \quad x_2 > x_3, \\ &= 0 && \text{if} \quad x_2 \leqq x_3; \\ g(s_3, \mathbf{x}) &= x_3 - x_1 && \text{if} \quad x_3 > x_1, \\ &= 0 && \text{if} \quad x_3 \leqq x_1. \end{aligned}$$

Then

$$\begin{aligned} G(\mathbf{x}) &= \tfrac{1}{3}[g(s_1, \mathbf{x}) + g(s_2, \mathbf{x}) + g(s_3, \mathbf{x})] \\ &= \tfrac{1}{3}[\max(x_1, x_2, x_3) - \min(x_1, x_2, x_3)]. \end{aligned}$$

Thus $G(\mathbf{x})$ depends on each of x_1, x_2 and x_3. Next taking $\alpha = 1$ in (9), we have

$$\begin{aligned} e(s_1, \mathbf{x}) &= 2(x_1 - x_2) && \text{if} \quad x_1 > x_2, \\ &= x_1 - x_2 && \text{if} \quad x_1 \leqq x_2; \\ e(s_2, \mathbf{x}) &= 2(x_2 - x_3) && \text{if} \quad x_2 > x_3, \\ &= x_2 - x_3 && \text{if} \quad x_2 \leqq x_3; \end{aligned}$$

and

$$\begin{aligned} e(s_3, \mathbf{x}) &= 2(x_3 - x_1) && \text{if} \quad x_3 > x_1, \\ &= x_3 - x_1 && \text{if} \quad x_3 \leqq x_1. \end{aligned}$$

Thus for each s, $e(s, \mathbf{x})$ depends on all the x_i, $i \in s$, but the estimate $e = \{e(s, \mathbf{x})\}$ is inadmissible being inferior in variance to $g = \{g(s, \mathbf{x})\}$.

REFERENCE

1] HANURAV, T. V. (1968). Hyperadmissibility and optimum estimators for sampling finite populations. *Ann. Math. Statist.* **39** 621–642, (particularly 637).