The Annals of Mathematical Statistics
1969, Vol. 40, No. 2, 672–676

ADMISSIBILITY AND BAYES ESTIMATION IN SAMPLING FINITE POPULATIONS—V

BY V. P. GODAMBE

University of Waterloo

1. Summary and introduction. Joshi (1965), (1966) studied in great detail admissible estimation, in relation to survey-sampling. He (1966) also established a property more demanding than admissibility namely *uniform admissibility* (previously called *global admissibility* by Godambe (1966)) for the conventional sample mean while estimating the population total. In this paper we establish uniform admissibility of a class of Bayes estimators.

Using the notation similar to that of Godambe and Joshi (1965) we denote the population units by integers 1, 2, $\cdots$, N. Any subset s of the integers 1, $\cdots$, N is called a sample. If S denotes the set of all possible samples, $(s \,\varepsilon\, S)$, any real function p on S such that $\sum_S p(s) = 1$ and $1 \geqq p(s) \geqq 0$, for all $s \,\varepsilon\, S$ is called a *sampling design*. Next we denote by x_i the real value associated with the unit i ($i = 1, \cdots, N$) of the population. $\mathbf{x} = (x_1, \cdots, x_i, \cdots, x_N)$ is a vector in the N-dimentional Euclidean space R_N. Any real function $e(\mathbf{x}, s)$ on the product space $R_N \times S$, such that e depends on $\mathbf{x}$ only through those x_i for which $i \,\varepsilon\, s$, is called an estimator. Since in this paper we would be concerned with estimation of the population total $T(\mathbf{x}) = \sum_1^N x_i$, the terms such as estimator, admissibility, uniform admissibility etc. used subsequently are to be understood *in relation to estimation of* T. Now to distinguish 'admissibility' from 'uniform admissibility' we introduce the following four definitions.

DEFINITION 1.1. For a *given* sampling design p, an estimator e' is said to be *superior* to the estimator e if for all $\mathbf{x} \,\varepsilon\, R_N$,

$$\sum_S p(s)[e'(s, \mathbf{x}) - T(\mathbf{x})]^2 \leqq \sum_S p(s)[e(s, \mathbf{x}) - T(\mathbf{x})]^2$$

strict inequality being true for at least one $\mathbf{x}$.

DEFINITION 1.2. For a *given* sampling design p, an estimator e is said to be *admissible* if no estimator e' is superior (Definition 1.1) to e.

DEFINITION 1.3. A pair (e', p') of an estimator e' and a sampling design p' is said to be *uniformly superior* to another pair (e, p) if for all $\mathbf{x} \,\varepsilon\, R_N$,

$$\sum_S p'(s)[e'(s, \mathbf{x}) - T(\mathbf{x})]^2 \leqq \sum_S p(s)[e(s, \mathbf{x}) - T(\mathbf{x})]^2$$

strict inequality holding for at least one $\mathbf{x}$.

DEFINITION 1.4. With respect to a class C of sampling designs, a pair (e, p) of an estimator e and a sampling design p is said to be *uniformly admissible* if no other pair (e', p') such that $p' \,\varepsilon\, C$, is uniformly superior to (e, p), (Definition 1.3).

For the discussion of the practical significance of the notion of uniform admissibility, especially if in Definition 1.4, the class $C = C_n$, where

Received 12 February 1968; revised 26 September 1968.

$$C_n = \{p: \textstyle\sum_s p(s)\cdot n(s) = \text{const.} = n\} \tag{1.1}$$

$n(s)$ being the number of units i such that $i \varepsilon s$, we refer to Joshi ((1966), Section 7). Obviously C_n above is the class of all sampling designs having a fixed 'average sample size.' The main result of this paper is the following

THEOREM 1.1 *With respect to the class C_n in (1.1) of sampling designs the pair (e^*, p^*), where e^* is the estimator given by,*

$$e^*(s, \mathbf{x}) = \textstyle\sum_{i\varepsilon s} x_i + \sum_{i\not\varepsilon s} \lambda_i , \tag{1.2}$$

$\lambda_1, \cdots, \lambda_i, \cdots, \lambda_N$ *being* any *arbitrarily fixed numbers and p^* is* any *sampling design belonging to the class C_n in (1.1), is uniformly admissible.* (Definition 1.4).

2. A Bayes estimator. Following the usual terminology we have

DEFINITION 2.1. For a given sampling design p an estimator e^* is said to be a *Bayes estimator*, with respect to the prior distribution ζ on R_N, for the population total $T(\mathbf{x})$, if for every other estimator e,

$$\int_{R_N} [\textstyle\sum_s p(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2]\, d\zeta \leqq \int_{R_N} [\sum_s p(s)(e(s, \mathbf{x}) - T(\mathbf{x}))^2]\, d\zeta. \tag{2.1}$$

For a discrete probability distribution ζ, the inequality (2.1) can be written as

$$\textstyle\sum_{R_N} \zeta(\mathbf{x})[\sum_s p(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2] \leqq \sum_{R_N} \zeta(\mathbf{x})[\sum_s p(s)(e(s, \mathbf{x}) - T(\mathbf{x}))^2]. \tag{2.2}$$

Now by changing the order of integration and summation in (2.1) and (2.2) and minimising the integrand for each fixed s, we obtain a Bayes estimator e^* as

$$e^*(s, \mathbf{x}) = \textstyle\sum_{i\varepsilon s} x_i + E_\zeta(\sum_{i\not\varepsilon s} x_i \mid s, x_i\colon i\, \varepsilon\, s), \tag{2.3}$$

where $E_\zeta(\cdot\mid\cdot)$ denotes conditional expectation, when ζ is the prior distribution·

NOTE. *The Bayes estimator (2.3) is independent of the sampling design p.*

Further we have from (2.3) the

THEOREM 2.1. *For any specified numbers $\lambda_1, \cdots, \lambda_i, \cdots, \lambda_N$ the estimator e^* given by (1.2) is a Bayes estimator for any prior distribution ζ such that, when distributed as ζ, $x_1, \cdots, x_N$ are probabilistically independent and $E_\zeta(x_i) = \lambda_i$, $i = 1, \cdots, N$.*

Now for proving the Theorems 1.1 and 3.1 to follow, we would need two lemmas, in the next section, based on the following classes Ω and Ω' of prior distributions on R_N.

$$\Omega = \left\{\begin{array}{l}\text{A class of discrete prior distributions } \zeta \text{ on } R_N,\\ \text{such that for any point, say } \mathbf{x}_0, \text{ in } R_N,\\ \text{there exists a } \zeta_0\, \varepsilon\, \Omega \text{ such that } \zeta_0(\mathbf{x}_0) > 0.\end{array}\right\}, \tag{2.4}$$

$$\Omega' = \left\{\xi\colon \begin{array}{l}\text{(a) } \xi \text{ is discrete}\\ \text{(b) } x_1, \cdots, x_N \text{ when distributed as } \xi \text{ are probabilistically independent,}\\ \text{(c) } E_\xi(x_i) = \lambda_i,\ i = 1, \cdots, N,\ \lambda_1, \cdots, \lambda_N \text{ being any specified numbers,}\\ \text{(d) } x_1, \cdots, x_N \text{ have a } \textit{common} \text{ but } \textit{unspecified} \text{ variance.}\end{array}\right\}. \tag{2.5}$$

3. Two lemmas. Noting, the facts: (1) although an estimator is defined as being Bayes relative to a particular prior distribution, typically a Bayes estimator depends only on certain characteristics of the prior and hence is Bayes relative to a class of prior distributions, and (2) the Bayes estimator (2.3) is independent of the sampling design, we have

LEMMA 3.1. *If* e^* *is the Bayes estimator* (2.3) *for all prior distributions* $\zeta \varepsilon \Omega$ *in* (2.4), *then for any sampling design* p, e^* *is admissible* (Definition 1.2).

The proof of the above lemma is quite straight forward. If e^* is not admissible then for some estimator e', for all $\mathbf{x} \varepsilon R_N$,

$$(3.1) \qquad \sum_s p(s)[e'(s, \mathbf{x}) - T(\mathbf{x})]^2 \leqq \sum_s p(s)[e^*(s, \mathbf{x}) - T(\mathbf{x})]^2,$$

with strict inequality for some $\mathbf{x}$ say $\mathbf{x}_0$. Now let ζ_0 be a prior distribution in Ω in (2.4), such that $\zeta_0(\mathbf{x}_0) > 0$. Multiplying both sides of (3.1) by $\zeta_0(\mathbf{x})$ and summing it over R_N, we have,

$$(3.2) \qquad \sum_{R_N} \zeta_0(\mathbf{x}) \sum_s p(s)[e'(s, \mathbf{x}) - T(\mathbf{x})]^2 < \sum_{R_N} \zeta_0(\mathbf{x}) \sum_s p(s)[e^*(s, \mathbf{x}) - T(\mathbf{x})]^2.$$

But (3.2) clearly contradicts the fact that e^* is a Bayes estimator (Definition 2.1) for the prior distribution ξ_0, proving the Lemma 3.1.

LEMMA 3.2. *For any point, say* $\mathbf{x}_0$, *in* R_N *there exists a* ξ, *say* $\xi_0 \varepsilon \Omega'$ *in* (2.5) *such that* $\xi_0(\mathbf{x}_0) > 0$.

To construct $\xi_0 \varepsilon \Omega'$ for which $\xi_0(\mathbf{x}_0) > 0$ we proceed as follows: Let the ith co-ordinate of the point $\mathbf{x}_0$ be x_{i0}, i.e. $\mathbf{x}_0 = (x_{10}, \cdots, x_{i0}, \cdots, x_{N0})$. Next we choose two other points $\mathbf{x}_1 = (x_{11}, \cdots, x_{i1}, \cdots, x_{N1})$ and $\mathbf{x}_2 = (x_{12}, \cdots, x_{i2}, \cdots, x_{N2})$ so as to satisfy the following equation (3.3) and for some *arbitrarily* chosen σ^2 equation (3.4). For $i = 1, \cdots, N$,

$$(3.3) \qquad x_{i1} + x_{i2} = 3\lambda_i - x_{i0},$$

$$(3.4) \qquad x_{i1}^2 + x_{i2}^2 = 3(\sigma^2 + \lambda_i^2) - x_{i0}^2.$$

Note that whatever x_{i0} and λ_i $(i = 1, \cdots, N)$, there would exist some x_{i1} and x_{i2} $(i = 1, \cdots, N)$, which would satisfy (3.3) and (3.4) provided σ^2 is sufficiently large. Now it is easy to see that the prior distribution ξ_0 on R_N which attaches equal probability $(=\frac{1}{3}^N)$ to all the 3^N points obtained by giving x_i (the ith co-ordinate of a generic $\mathbf{x}$) three values namely x_{i0}, x_{i1} and x_{i2}, belongs to Ω' in (2.5) and $\xi_0(\mathbf{x}_0) = \frac{1}{3}^N > 0$. This completes the proof of Lemma 3.2.

An immediate consequence of the Lemmas 3.1 and 3.2 is the following:

THEOREM 3.1. *For every sampling design* p, *the estimator* e^* *given by* (1.2), *for arbitrarily chosen* $\lambda_1, \cdots, \lambda_N$, *is admissible* (Definition 1.2.).

4. Proof of Theorem 1.1. Noting again the fact that the Bayes estimator (2.3) is independent of the sampling design, we first prove the following:

THEOREM 4.1. *If* e^* *is the Bayes estimator* (2.3), *with respect to every prior* *distribution* $\zeta \varepsilon \Omega$ *in* (2.4) *and if* C *is a class of sampling designs* p, *such that the Bayes risk*

$$\sum_{R_N} \zeta(\mathbf{x})[\sum p(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2] = \text{const. } (\zeta), \tag{4.1}$$

for all sampling designs $p \varepsilon C$ *and all* $\zeta \varepsilon \Omega$, *then with respect to* C, *the pair* (e^*, p^*) *where* p^* *is* any *sampling design in* C, *is* uniformly admissible. (Definition 1.4).

To prove the above theorem we note that if (e^*, p^*) is not uniformly admissible, there must be a pair (e', p'), $p' \varepsilon C$, which is uniformly superior (Definition 1.3) to (e^*, p^*), i.e.

$$\sum_s p'(s)(e'(s, \mathbf{x}) - T(\mathbf{x}))^2 \leqq \sum_s p^*(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2 \tag{4.2}$$

for all $\mathbf{x} \varepsilon R_N$ with strict inequality for some $\mathbf{x} = \mathbf{x}_0$ say. Further let $\zeta_0 \varepsilon \Omega$ (in (2.4)) be such that the probability $\zeta_0(\mathbf{x}_0) > 0$. Now multiplying both sides of (4.2) by $\zeta_0(\mathbf{x})$ and summing over all $\mathbf{x} \varepsilon R_N$ we get

$$\sum_{R_N} \zeta_0(\mathbf{x}) \sum_s p'(s)(e'(s, \mathbf{x}) - T(\mathbf{x}))^2 < \sum_{R_N} \zeta_0(\mathbf{x}) \sum_s p^*(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2. \tag{4.3}$$

Further because of (4.1) the right hand side of the inequality (4.3),

$$\sum_{R_N} \zeta_0(\mathbf{x}) \sum_s p^*(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2 = \sum_{R_N} \zeta_0(\mathbf{x}) \sum_s p'(s)(e^*(s, \mathbf{x}) - T(\mathbf{x}))^2. \tag{4.4}$$

Now (4.3) and (4.4) together contradict (2.2), i.e. the fact that e^* is a Bayes estimator (2.3) with respect to all $\zeta \varepsilon \Omega$ in (2.4) especially ζ_0. Hence (e', p') satisfying (4.2) cannot exist. This proves the Theorem 4.1.

Now consider the class C_n of sampling designs p, defined in (1.1) and for some specified $\lambda_1, \cdots, \lambda_N$ the class of prior distributions Ω' in (2.5), the estimator e^* in (1.2). Substituting (1.2) for e^* in the left hand side of (4.1) and interchanging the order of summation we can write the Bayes risk for e^* in (1.2) as

$$\sum_s p(s) \sum_{R_N} (\sum_{i \notin s} (\lambda_i - x_i))^2 \xi(\mathbf{x}) \tag{4.5}$$

where $\mathbf{x} = (x_1, \cdots, x_i, \cdots, x_N)$. If $\xi \varepsilon \Omega'$ in (2.5), the Bayes risk in (4.5) equals

$$\sum_s p(s)[N - n(s)]\sigma_\xi^2 = \sigma_\xi^2 N - \sigma_\xi^2 \sum_s n(s)p(s) \tag{4.6}$$

where $n(s)$ is the number of individuals $i \varepsilon s$ and σ_ξ^2 is the common variance referred to in (d) of (2.5). Now $\sigma_\xi^2 \sum_s n(s)p(s)$ is constant, depending on ξ only, for all sampling designs $p \varepsilon C_n$ in (1.1). Thus Theorem 2.1, Lemma 3.2 and Theorem 4.1 together imply Theorem 1.1.

It is easy to see from the above that Theorem 1.1 would hold even if the class of sampling designs C_n is replaced by any subset of C_n. Taking the subset to be just one sampling design we get Theorem 2.1 as a special case of Theorem 1.1.

Acknowledgment. The author wishes to thank Mrs. D. Bos and Mrs. I. Suerich for typing the manuscript. He also is indebted to the referee and the associate editor for suggestions which have improved the presentation.

REFERENCES

GODAMBE, V. P. and JOSHI, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations—I. *Ann. Math. Statist.* **36** 1707–1723.

GODAMBE, V. P. (1966). Bayes and empirical Bayes estimation in sampling finite populations (Abstract). *Ann. Math. Statist.* **37** 552.

JOSHI, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations—II, III. *Ann. Math. Statist.* **36** 1723–1742.

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