*The Annals of Statistics*
1985, Vol. 13, No. 2, 811–817

# SOME ADMISSIBLE NONPARAMETRIC AND RELATED FINITE POPULATION SAMPLING ESTIMATORS

BY GLEN MEEDEN[1], MALAY GHOSH[2] AND STEPHEN VARDEMAN[1]

*Iowa State University, University of Florida and Iowa State University*

Given a random sample from an unknown distribution $F$, which is assumed to belong to some nonparametric family of distributions, consider the problem of estimating $\gamma(F)$, some function of $F$. When the loss function is squared error, admissible estimators are exhibited for a large class of $\gamma$'s. A relationship between these estimators and similar ones in finite population sampling is demonstrated.

**1. Introduction.** Let $X_1, \cdots, X_n$ be a random sample from an unknown distribution $F$, which is assumed to belong to $\Theta$, some large nonparametric family of distribution functions on the set of real numbers. We wish to estimate $\gamma(F) = \int \phi(t)\, dF(t)$, where $\phi$ is some specific function, with squared error loss.

Cohen and Kuo (1983) recently demonstrated that the empirical distribution function is an admissible estimator of $F$ itself for a class of loss functions. They showed that to prove admissibility it was enough to just consider the subfamily of $\Theta$ consisting of all discrete distributions with at most a finite number of jumps. For this subfamily, admissibility was proved by adopting an argument of Alam (1979) for estimating multinomial probabilities. This type of argument, which was originally discussed in Johnson (1971), can be thought of as an example of proving admissibility by demonstrating that the estimator is stepwise Bayes against a sequence of priors (see Hsuan, 1979; and Brown, 1981).

Cohen and Kuo then modify their argument to prove that the empirical distribution function is an admissible estimator of the population distribution function in finite population sampling.

Recently the authors have proved various admissibility results for finite population sampling (see Meeden and Ghosh, 1983; and Vardeman and Meeden, 1983) using an argument based on a result in Meeden and Ghosh (1981) (which is closely related to Hsuan, 1979).

In this paper we review all these arguments and show that they are essentially the same. In addition, we demonstrate that there is a natural duality between admissible estimators in finite population sampling and admissible estimators in the nonparametric problem given above. This comes about because both problems are related to proving admissibility for a multinomial problem. This observation generalizes the work of Cohen and Kuo (1983) and for the first time proves the admissibility of many standard nonparametric estimators.

---

Received July 1983; revised September 1984.
[1] Research supported by NSF Grant Number MCS-8202116.
[2] Research supported by NSF Grant Number MCS-8218091.
*AMS* 1980 *subject classifications.* Primary 62C15, 62G05; secondary 62D05, 62C10.
*Key words and phrases.* Admissibility, nonparametric estimation, finite population sampling, stepwise Bayes.



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

For estimating $F$ itself the recent work of Brown (1984) uses similar techniques to give a detailed study of the admissibility of various invariant nonparametric estimators.

**2. Proving admissibility.** In what follows we will be proving admissibility for several closely related problems. The first is the nonparametric problem. Let $\phi$ be a specified Borel measurable function defined on the real numbers. Let $\Theta$ denote the class of all distribution functions $F$ for which

$$\gamma(F) = \int \phi(t)\, dF(t) \tag{2.1}$$

exists. Given $X_1, \cdots, X_n$, a random sample from $F$, we wish to estimate $\gamma(F)$ with squared error loss when $F$ is assumed to belong to $\Theta$. As we will see, one can find admissible estimators for this problem by finding admissible estimators for the following simpler problem.

Let $\alpha_1, \cdots, \alpha_r$ be $r$ distinct real numbers. Let $\Theta(\alpha_1, \cdots, \alpha_r)$ denote all distribution functions which concentrate all their mass on $\alpha_1, \cdots, \alpha_r$. We now consider the problem of estimating $\gamma(F)$ when $F$ is assumed to belong to $\Theta(\alpha_1, \cdots, \alpha_r)$. In this case, $X_1, \cdots, X_n$ is a random sample from a multinomial $(1; p_1, \cdots, p_r)$ population where $p_i = P(X_j = \alpha_i)$ for $i = 1, \cdots, r$ and $j = 1, \cdots, n$. If $x = (x_1, \cdots, x_n)$ denotes a possible set of outcomes for the random sample, then we let $w_j(x)$ be the number of the $x_i$'s equal to $\alpha_j$ for $j = 1, \cdots, r$. Note that $\Theta(\alpha_1, \cdots, \alpha_r)$ is equivalent to the $(r-1)$-dimensional simplex

$$T = \{p = (p_1, \cdots, p_r): p_i \geq 0 \text{ for } i = 1, \cdots, r \text{ and } \textstyle\sum_{i=1}^r p_i = 1\}.$$

Each $p \in T$ determines a unique $F$, say $F_p$, and we write $\gamma(p) = \gamma(F_p) = \sum_{i=1}^r \phi(\alpha_i)p_i$. As was noted by Cohen and Kuo (1983), to prove that an estimator is admissible for the nonparametric problem, it is enough to show that it is admissible for the multinomial problem with parameter space $\Theta(\alpha_1, \cdots, \alpha_r)$ for every choice of $\alpha_1, \cdots, \alpha_r$ for $r = 1, 2, \cdots$.

We will now show how these two estimation problems are closely related to two estimation problems in finite population sampling. Consider a finite population $U$ with units labeled $1, 2, \cdots, N$. Let $y_i$ be the value of a single characteristic attached to the unit $i$. The vector $y = (y_1, \cdots, y_N)$ is the unknown state of nature and is assumed to belong to the parameter space $\Gamma \subset R^N$. A subset $s$ of $\{1, 2, \cdots, N\}$ is called a sample. Let $n(s)$ denote the number of elements in $s$. Let $\Delta$ be a design which assigns positive mass only to sets of size $n$. Given $y \in \Gamma$ and $s = \{i_1, \cdots, i_n\}$ where $1 \leq i_1 < i_2 < \cdots < i_n \leq N$, let $y(s) = (y_{i_1}, \cdots, y_{i_n})$.

It is conventional to take $\Gamma = R^N$. However, other choices are often sensible as well. In particular, we will find it convenient to take $\Gamma$ equal to

$$\Gamma(\alpha_1, \cdots, \alpha_r) = \{y: y_i = \alpha_j \text{ for some } j = 1, \cdots, r \text{ for all } i = 1, \cdots, N\}$$

where $\{\alpha_1, \cdots, \alpha_r\}$ is a set of $r$ distinct real numbers. No matter how $\Gamma$ is defined, given a $y \in \Gamma$ we denote by $F_y$ the distribution function which assigns mass $1/N$ to each component $y_i$ of $y$. We consider the problem of estimating, with squared error loss,

$$\gamma(y) = \gamma(F_y) = \int \phi(t)\, dF_y(t) = \textstyle\sum_{i=1}^N \phi(y_i)/N. \tag{2.2}$$

For $y \in \Gamma(\alpha_1, \cdots, \alpha_r)$, let $w_j(y)$ be the number of $y_i$'s equal to $\alpha_j$, and $w_j(y(s))$ be the number of $y_i$'s with $i \in s$ equal to $\alpha_j$.

Let $e(s, y)$ denote an estimator of $\gamma(y)$. ($e(s, y)$ depends on $y$ only through $y(s)$.) If the design $\Delta$ is used in conjunction with the estimator $e$, then the risk function is $r(y; \Delta, e) = \sum_s [e(s, y) - \gamma(y)]^2\Delta(s)$.

It was demonstrated in Meeden and Ghosh (1983) that if an estimator is admissible for the design $\Delta$ when $\Gamma$ is taken to be $\Gamma(\alpha_1, \cdots, \alpha_r)$ for every choice of $\alpha_1, \cdots, \alpha_r$, for $r = 1, 2, \cdots$, then it is admissible for the design $\Delta$ when $\Gamma = R^N$ as well.

Note that this relationship is very similar to the one between the multinomial estimation problem and the nonparametric estimation problem discussed above. We will now indicate how these two sets of problems are related.

First consider the multinomial problem with parameter space $T = \Theta(\alpha_1, \cdots, \alpha_r)$. Let $G$ be a prior distribution over $T$. Given a sample $(X_1 = x_1, \cdots, X_n = x_n)$ the Bayes estimator of $p_j$ against $G$ is

$$E_G(p_j \mid x) = \frac{\int \cdots \int p_j \prod_{i=1}^r p_i^{w_i(x)}\, dG(p_1, \cdots, p_r)}{\int \cdots \int \prod_{i=1}^r p_i^{w_i(x)}\, dG(p_1, \cdots, p_r)} = P(\alpha_j \mid x, G) \tag{2.3}$$

where $x = (x_1, \cdots, x_n)$ and $P(\alpha_j \mid x, G)$ is the $G$ posterior probability that an additional observation takes the value $\alpha_j$. From this it follows that

$$E_G(\gamma \mid x) = \textstyle\sum_{i=1}^r \phi(\alpha_i) P(\alpha_i \mid x, G) \tag{2.4}$$

is the Bayes estimator of $\gamma$ against $G$ based on $x$.

Next we consider the finite population estimation problem with $\Gamma = \Gamma(\alpha_1, \cdots, \alpha_r)$. First note that the prior distribution $G$ on $T$ induces a prior distribution $G^*$ over $\Gamma(\alpha_1, \cdots, \alpha_r)$ by the relationship

$$G^*(y) = \int \cdots \int \textstyle\prod_{i=1}^r p_i^{w_i(y)}\, dG(p_1, \cdots, p_r)$$

for $y \in \Gamma(\alpha_1, \cdots, \alpha_r)$. If the design $\Delta$ is used to pick the sample $s$ which results in the observations $y(s)$, then the form of the Bayes estimator of $\gamma(y)$ and the fact that the estimator does not depend on the design $\Delta$ are well known (see Basu, 1969). That is,

$$E_{G^*}(\gamma \mid y(s)) = (n/N) \textstyle\sum_{i=1}^r \phi(\alpha_i) w_i(y(s))/n + ((N - n)/N) E_G(\gamma \mid y(s)) \tag{2.5}$$

is the Bayes estimator of $\gamma$ against $G^*$ based on the sample $y(s)$. (Note that here we are identifying $y(s)$, the observed values in our sample of size $n$, with $x$, the vector of observations of length $n$ in the multinomial problem.)

Note that there is an interesting relationship between the estimators in (2.4) and (2.5). Suppose that we are given a set of $n$ observations each of which belongs to the set $\{\alpha_1, \cdots, \alpha_r\}$. If we assume that the observations arose from the multinomial model with prior $G$, then the estimator in (2.4) is the proper one. If, on the other hand, we assume that the observations were generated by the finite population model with prior $G^*$, then the estimator in (2.5) is the proper one. This suggests that the term $\sum_{i=1}^{r} \phi(\alpha_i)w_i(y(s))/n$ in (2.5) can be interpreted as the finite population correction factor to the estimators in (2.4).

If we let $F_{y(s)}$ denote the distribution function which puts mass $1/n$ on each member of $y(s)$ and

$$\delta_{G^*}(y(s)) = E_{G^*}(\gamma \mid y(s)) \quad \text{and} \quad \delta_G(x) = E_G(\gamma \mid x),$$

then equation (2.5) can be written as

$$\delta_{G^*}(y(s)) = (n/N) \int \phi(t)\, dF_{y(s)}(t) + ((N - n)/N)\delta_G(y(s)). \tag{2.6}$$

If $G$ is such that $\delta_G$ is the unique Bayes estimator against $G$ for the multinomial problem with parameter space $T = \Theta(\alpha_1, \cdots, \alpha_r)$, then $\delta_G$ is admissible. In addition, $\delta_{G^*}$ is the unique Bayes estimator against $G^*$ for the finite population sampling problem with parameter space $\Gamma(\alpha_1, \cdots, \alpha_r)$ and hence $\delta_{G^*}$ is admissible. More generally, if $\delta$ is the unique stepwise Bayes estimator against a sequence of priors for the multinomial problem with parameter space $T = \Theta(\alpha_1, \cdots, \alpha_r)$, then $\delta$ is admissible. In addition, the estimator

$$\delta^*(y(s)) = (n/N) \int \phi(t)\, dF_{y(s)}(t) + ((N - n)/N)\delta(y(s)) \tag{2.7}$$

is a unique stepwise Bayes estimator for the finite population sampling problem with parameter space $\Gamma(\alpha_1, \cdots, \alpha_r)$ and hence is admissible.

Arguing as in Meeden and Ghosh (1983) or Cohen and Kuo (1983), it follows that if for every finite set of real numbers $(\alpha_1, \cdots, \alpha_r)$, $\delta$ is admissible for the multinomial problem with parameter space $T = \Theta(\alpha_1, \cdots, \alpha_r)$, then $\delta$ is admissible for the nonparametric estimation problem as well. In addition, $\delta^*$ is admissible for the finite population sampling problem with parameter space $\Gamma = R^N$. More generally, if for every finite set of real numbers $\{a_1, \cdots, a_k\}$ there exists a set $\{\alpha_1, \cdots, \alpha_r\} \supset \{a_1, \cdots, a_k\}$ such that $\delta$ is admissible for the multinomial problem with parameter space $T = \Theta(\alpha_1, \cdots, \alpha_r)$, then $\delta$ is admissible for the nonparametric estimation problem and $\delta^*$ is admissible for the finite population sampling problem with parameter space $\Gamma = R^N$.

The preceding argument is summarized in the following theorem.

THEOREM. *Let $\phi$ be a Borel measurable function with $\Theta = \{F: \int | \phi(t) |\, dF(t) < \infty\}$. Let $X_1, \cdots, X_n$ be iid $F$ where $F \in \Theta$. If $\delta$ is an estimator of $\gamma(F)$ under squared error loss such that for each finite set of real numbers $\{a_1, \cdots, a_k\}$ there exist a set of real numbers $\{\alpha_1, \cdots, \alpha_r\}$ such that $\{a_1, \cdots, a_k\} \subset \{\alpha_1, \cdots, \alpha_r\}$ and $\delta$ is a unique stepwise Bayes estimator of $\gamma(F)$ for the multinomial problem with parameter space $T = \Theta(\alpha_1, \cdots, \alpha_r)$, then $\delta$ is an admissible estimator of $\gamma(F)$ for the nonparametric problem when the parameter space is $\Theta$. In addition, for the* *finite population sampling problem with parameter space $R^N$ the estimator $\delta^*$, given in (2.7), is admissible for estimating $\gamma$, given in (2.2), for every design of fixed sample size n.*

**3. Some examples.** We now give some examples where admissibility follows from the above theorem and its obvious generalizations. The actual details are omitted but can be found in Meeden, Ghosh and Vardeman (1984).

EXAMPLE 1. $\sum_{i=1}^{n} \phi(x_i)/n$ is admissible for estimating $\gamma(F)$ for the nonparametric problem and $\sum_{i\in s} \phi(y_i)/n$ is admissible for estimating $\sum_{i=1}^{N} \phi(y_i)/N$ in the finite population sampling problem. The sequence of priors needed is the one used in Brown (1981) (see also Alam, 1979) in the special case where $\phi(t) \equiv t$ for the multinomial problem. This sequence was also used in Meeden and Ghosh (1983) for the special case where $\phi(t) \equiv t$ in the finite population problem.

EXAMPLE 2. Let $\phi$ be an arbitrary nonconstant Borel measurable function with $\underline{c} = \inf_t \phi(t) < \sup_t \phi(t) = \bar{c}$. Let $\mu^* \in (\underline{c}, \bar{c})$ and $M > 0$ be given. Then using a sequence of priors given in Vardeman and Meeden (1983) it follows from the theorem that $M\mu^*/(M + n) + \sum_{i=1}^{n} \phi(x_i)/(M + n)$ is admissible for estimating $\gamma(F)$ in the nonparametric problem. In the special case $\phi(t) = t$, this is the estimator given in equation (7) of Ferguson (1973) if one identifies $M$ with the parameter $\alpha(R)$ of the Dirichlet process.

EXAMPLE 3. The arguments of Section 2 can be generalized to obtain admissible estimators of parameters $\gamma(F)$ of types other than those satisfying equation (2.1). For example, let $\gamma(F)$ be an estimable parameter with kernel $\phi$ and degree $m > 1$, i.e.

$$\gamma(F) = \int \cdots \int \phi(x_1, \cdots, x_m)\, dF(x_1), \cdots, dF(x_m),$$

where, without loss of generality, it can be assumed that $\phi$ is symmetric in its arguments. One special case of interest for $m = 2$ is $\phi(u, v) = (u - v)^2/2$. Using the sequence of priors of Example 1, it follows that $(n + 1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is an admissible estimator of the population variance.

It is interesting to note that if one assumes the population to be normal then this estimator is the best invariant estimator of the population variance $\sigma^2$. However, as shown by Stein (1964), such an estimator is not then an admissible estimator of $\sigma^2$ under squared error loss.

If the sequence of priors of Example 2 is used then one obtains, as an admissible estimator of the population variance, the estimator given in equation (15) of Ferguson (1973). For the finite population analogues of these two estimators, see Ghosh and Meeden (1983) and Vardeman and Meeden (1983).

EXAMPLE 4. Another possible generalization of the theorem is to two sample problems. Suppose we have independent random samples from two different distributions, say $F$ and $G$. Consider the problem of estimation

$$\gamma(F, G) = \int F(x)\, dG(x)$$

the probability that an observation from $F$ is less than or equal to an observation from $G$ with squared error loss. The usual unbiased estimator is the Mann-Whitney $U$-statistic. Another estimator is given in part (f) of Section 5 of Ferguson (1973). It can be shown that both these estimators are admissible.

EXAMPLE 5. The theorem can be extended to nonparametric estimation problems when sampling from a bivariate or, more generally, a multivariate population.

For example, for a sample of size $n$, from a bivariate distribution, the Kendall tau statistic multiplied by $(n - 1)/(n + 1)$ is admissible for estimating the probability of concordance minus the probability of discordance nonparametrically.

Finally, suppose we have a random sample of size $n$ from a $k$-dimensional distribution and wish to estimate the $k$-dimensional vector of marginal means when the loss function is the sum of the squared error losses of each component. It then follows that the sample mean vector is an admissible estimator of the population mean vector for the nonparametric problem. This shows the absence of the Stein effect for the nonparametric problem. For related results see Gutmann (1982) and Joshi (1977, 1979).

REMARK. Several of the above estimators for which admissibility has been demonstrated are proper Bayes estimators with respect to Dirichlet process priors of Ferguson (1973). It is natural to wonder why this fact does not in itself guarantee their admissibility. In the usual theory, there are two standard arguments for showing the admissibility of Bayes estimators.

In the first, admissibility follows when all the risk functions are continuous and the support of the prior is the entire parameter space. In the present case, this fails because the Dirichlet prior is concentrated on the discrete distributions and there is no way to topologize $\Theta$ so that all risk functions are continuous and the closure of the discrete distributions, which is the support of the prior, is $\Theta$.

In the second approach, admissibility follows if a Bayes estimator is unique a.e. $F$ for all $F \in \Theta$. This fails here because the Bayes estimator is only uniquely determined for almost all $X = (X_1, \cdots, X_n)$ relative to the marginal distribution of $X$. For example, suppose the parameter defining the Dirichlet prior is absolutely continuous with respect to Lebesgue measure. The marginal distribution of $X$ is a mixture of distributions. One component is absolutely continuous with respect to Lebesgue measure on $R^n$. The others are absolutely continuous with respect to Lebesgue measure on lower dimensional hyperplanes defined by the equality of some subset of coordinates of $X$. Hence it is easy to find a subset of $R^n$ containing a finite number of elements which has zero probability under the marginal distribution of $X$ but which is assigned probability one when $X_1, \cdots, X_n$ are iid $F$ for all $F \in \Theta(\alpha_i, \cdots, \alpha_r)$ for some choice of $\alpha_1, \cdots, \alpha_r$.

Since $\Theta(\alpha_1, \cdots, \alpha_r) \subset \Theta$, an estimator which is Bayes against a Dirichlet prior is not unique a.e. $F$ for all $F \in \Theta$.

## REFERENCES

ALAM, K. (1979). Estimation of multinomial probabilities. *Ann. Statist.* **7** 282–283.

BASU, D. (1969). Role of the sufficiency and likelihood principles in sample survey. *Sankhyā Sec. B* **31** 441–454.

BROWN, L. (1981). A complete class theorem for statistical problems with finite sample spaces. *Ann. Statist.* **9** 1289–1300.

BROWN, L. (1984). Admissibility in discrete and continuous invariate nonparametric estimation problems, and their multinomial analogue. Submitted to *Ann. Statist.*

COHEN, M. and KUO, L. (1983). The admissibility of the empirical distribution function. *Ann. Statist.* **13** 262–271.

FERGUSON, T. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

GHOSH, M. and MEEDEN, G. (1983). Estimation of the variance in finite population sampling. *Sankhyā Sec. B* **45** 362–375.

GUTMANN, S. (1982). Stein's Paradox is impossible in problems with finite sample space. *Ann. Statist.* **10** 1017–1020.

HSUAN, F. C. Y. (1979). A stepwise Bayesian procedure. *Ann. Statist.* **7** 860–868.

JOHNSON, B. M. (1971). On the admissible estimators for certain fixed sample binomial problems. *Ann. Math. Statist.* **42** 1579–1587.

JOSHI, V. M. (1977). A note on estimators in finite populations. *Ann. Statist.* **5** 1051–1053.

JOSHI, V. M. (1979). Joint admissibility of the sample means as estimators of the means of finite populations. *Ann. Statist.* **7** 995–1002.

MEEDEN, G. and GHOSH, M. (1981). Admissibility in finite problems. *Ann. Statist.* **9** 846–852.

MEEDEN, G. and GHOSH, M. (1983). Choosing between experiments: Application to finite population sampling. *Ann. Statist.* **11** 296–305.

MEEDEN, G., GHOSH, M. and VARDEMAN, S. (1984). Some admissible nonparametric and related finite population sampling estimators; technical details. University of Florida Technical Report.

STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

VARDEMAN, S. and MEEDEN, G. (1983). Admissible estimators in finite population sampling employing various types of prior information. *J. Statist. Plann. Inference* **7** 329–341.

M. GHOSH
DEPARTMENT OF STATISTICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611

G. MEEDEN
S. VARDEMAN
DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011