BULLETIN 13 (3), 1970

INADMISSIBILITY OF THE MAXIMUM LIKELIHOOD ESTIMATOR IN THE PRESENCE OF PRIOR INFORMATION

BY
B. K. KALE

0. Lehmann [1] in his lecture notes on estimation shows that for estimating the unknown mean of a normal distribution, $N(\theta, 1)$, the usual estimator $\bar{x}$ is neither minimax nor admissible if it is known that θ belongs to a finite closed interval $[a, b]$ and the loss function is squared error. It is shown that $\hat{\theta}(x)$, the maximum likelihood estimator (MLE) of θ, has uniformly smaller mean squared error (MSE) than that of $\bar{x}$. It is natural to ask the question whether the MLE $\hat{\theta}(x)$ of θ in $N(\theta, 1)$ is admissible or not if it is known that $\theta \in [a, b]$. The answer turns out to be negative and the purpose of this note is to present this result in a slightly generalized form.

1. Let the r.v. X have a p.d.f. belonging to exponential class of densities. Thus $dP_\theta(x) = \beta(\theta)\, e^{x\theta}\, d\mu(x)$, where μ is a regular σ-finite measure on R_1, $\theta \in [a, b]$, a proper subset of the natural range of the parameter space. For the basic properties of exponential family we refer to Lehmann [2]. Further, without any loss of generality, we suppose only one observation on X, because since the densities belong to the exponential class the same result will hold for a random sample of any size.

We assume that we are estimating $\eta(\theta) = E(x \mid \theta)$, and the loss function is squared error. Now

$$\begin{aligned} \hat{\theta}(x) &= \eta(a) && x < \eta(a) \\ &= x && \eta(a) \le x \le \eta(b) \\ &= \eta(b) && x > \eta(b). \end{aligned}$$

One can easily prove that MSE $[\hat{\theta}(x)] <$ MSE $[x]$, noting that $\eta(\theta)$ is strictly increasing and x takes values outside $[\eta(a), \eta(b)]$ with positive probabilities under each θ. Thus x is inadmissible.

2. We first study the structure of any admissible estimator of $\eta(\theta)$. The loss function is $W(\delta, \theta) = [\delta - \eta(\theta)]^2$, $\delta \in R_1$, $\theta \in [a, b]$ and is continuous in both variables. Noting that the parameter space is a finite closed interval, the standard results on complete class theorems [3, Ch. 5] imply that every admissible estimator of $\eta(\theta)$ must be Bayes with respect to a proper prior distribution $\Lambda(\theta)$ on $[a, b]$. Let $\xi_\Lambda(x)$ be Bayes estimator of $\eta(\theta)$ relative to $\Lambda(\theta)$. Then

$$\begin{aligned} \xi_\Lambda(x) &= E[\eta(\theta) \mid x] \\ &= \frac{\int_a^b \eta(\theta)\, e^{\theta x} \beta(\theta)\, d\Lambda(\theta)}{\int_a^b e^{\theta x} \beta(\theta)\, d\Lambda(\theta)} \end{aligned} \tag{3}$$

As $\xi_\Lambda(x)$ is a conditional expectation given x, $\xi_\Lambda(x)$ is unique up to a set of probability measure zero under each $\theta \in [a, b]$, and any other estimator $T(x)=\xi_\Lambda(x)$, a.e. $\{P_\theta, \theta \in [a, b]\}$, will be Bayes and thus admissible. Let

$$\psi_1(x) = \int_a^b e^{\theta x}\beta(\theta)\, d\Lambda(\theta)$$

and

$$\psi_2(x) = \int_a^b e^{\theta x}\beta(\theta)\eta(\theta)\, d\Lambda(\theta).$$

Note that $\eta(\theta)$ being bounded on $[a, b]$, both $\psi_1(x)$ and $\psi_2(x)$ are well defined for each $x \in R_1$. Next considering $x=\xi+i\eta$, where $(\xi, \eta) \in R_2$, we show that $\psi_1(x)$ and $\psi_2(x)$ are analytic functions. Consider $\psi_1(x)$, then

$$\left|\frac{\psi_1(x+h)-\psi_1(x)}{h}\right| \le \int_a^b e^{\theta\xi}\beta(\theta)\left|\frac{e^{\theta h}-1}{h}\right| d\Lambda(\theta) \tag{4}$$

Now

$$\left|\frac{e^{\theta h}-1}{h}\right| \le \frac{\exp(\delta|\theta|)}{\delta} \quad \text{for } |h| \le \delta$$

and therefore

$$e^{\theta\xi}\left|\frac{e^{\theta h}-1}{h}\right|\beta(\theta) \le \frac{1}{\delta}\{\exp[(\delta+\xi)\theta]+\exp[(\xi-\delta)\theta]\}\beta(\theta). \tag{5}$$

Since the R.H.S. of (5) is integrable for any $\delta>0$ and $\xi \in R_1$, we have by the Lebesgue dominated convergence theorem, $\psi_1(x)$ is differentiable at each x and thus $\psi_1(x)$ is analytic. A similar argument shows that $\psi_2(x)$ is also analytic and therefore $\xi_\Lambda(x)$, being the ratio of two analytic functions, is itself analytic at each point x.

Therefore if $T(x)$ is an admissible estimator of $\eta(\theta)$, then either (A) $T(x)$ is analytic at each point $x \in R_1$ or (B) there exists $T_1(x)$ analytic at each point $x \in R_1$ such that $T(x)=T_1(x)$, a.e. $\{P_\theta, \theta \in [a, b]\}$.

We consider first the case where μ is absolutely continuous with respect to Lebesgue measure so that $dP_\theta(x)=\beta(\theta)\, e^{\theta x}h(x)\, dx$. In this case we have $T(x)$ is admissible for $\eta(\theta)$ provided either $T(x)$ is analytic at each x or else there exists an analytic function $T_1(x)$ such that $T(x)=T_1(x)$ a.e. [Lebesgue]. In such a case it is now obvious that $\hat{\theta}(x)$ is not admissible, being neither analytic at $x=\eta(a)$ and $x=\eta(b)$, nor does there exist an estimator $T_1(x)$ such that $T_1(x)$ is analytic everywhere and $T_1(x)=\hat{\theta}(x)$ a.e. [Lebesgue]. In particular for $N(\theta, 1)$, it follows that the MLE of θ, when $\theta \in [a, b]$, is not admissible.

When μ is not absolutely continuous with respect to Lebesgue measure the above criterion is of no use. The difficulty lies in the fact that the Bayes estimator need be defined only a.e. $[\mu]$. We illustrate this by binomial density $b(1, p)$ where $p \in [\frac{1}{4}, \frac{3}{4}]$. Let $\Lambda(p)$ be uniform over $[\frac{1}{4}, \frac{3}{4}]$. Then $\xi_\Lambda(x)$ is given by $\xi_\Lambda(0)=\frac{11}{24}$ and

$\xi_\Lambda(1)=\frac{13}{24}$, and $\xi_\Lambda(x)$ can be defined arbitrarily at any other point. One can construct several nonanalytic functions, $T(x)$ such that $T(0)=\frac{11}{24}$ and $T(1)=\frac{13}{24}$ and $T(x)$ such that $T(x)=\xi_\Lambda(x)$ a.e. under $p \in [\frac{1}{4}, \frac{3}{4}]$.

References

1. E. L. Lehmann, *Notes on theory of estimation*, Associated Students' Store. Univ. of California, 1949.
2. ———, *Testing statistical hypotheses*, Wiley, New York, 1959.
3. A. Wald, *Statistical decision functions*, Wiley, New York, 1950.

University of Manitoba,
Winnipeg, Manitoba

https://doi.org/10.4153/CMB-1970-076-x Published online by Cambridge University Press