1980

A Necessary Condition for Admissibility

Lawrence D. Brown
University of Pennsylvania

Recommended Citation

Brown, L. D. (1980). A Necessary Condition for Admissibility. *The Annals of Statistics, 8* (3), 540-544. http://dx.doi.org/10.1214/aos/1176345007

A Necessary Condition for Admissibility

Abstract
The main theorem of this note is required in a paper of Brown. Briefly, the theorem shows that procedures which can be improved on in a neighborhood of infinity are either inadmissible or are generalized Bayes for a (possibly improper) prior whose rate of growth at infinity is of an appropriate order.

This theorem is applied here to show that the risk of the usual estimator of a two dimensional normal mean, θ, cannot be improved on near ∞ at order $\|\theta\|^{-2}$.

Keywords
admissibility, generalized Bayes procedures

Disciplines
Statistics and Probability

The Annals of Statistics
1980, Vol. 8, No. 3, 540–544

A NECESSARY CONDITION FOR ADMISSIBILITY[1]

BY LAWRENCE D. BROWN

Rutgers University and Cornell University

The main theorem of this note is required in a paper of Brown. Briefly, the theorem shows that procedures which can be improved on in a neighborhood of infinity are either inadmissible or are generalized Bayes for a (possibly improper) prior whose rate of growth at infinity is of an appropriate order.

This theorem is applied here to show that the risk of the usual estimator of a two dimensional normal mean, θ, cannot be improved on near ∞ at order $\|\theta\|^{-2}$.

Consider a statistical decision problem $\mathcal{P}$ with sample space X, $\mathcal{B}_X$; parameter space Θ, $\mathcal{B}_\Theta$; decision space $\mathcal{A}$, $\mathcal{B}_\mathcal{A}$; distributions $\{F_\theta: \theta \in \Theta\}$; and loss function $L: \Theta \times \mathcal{A} \to [0, \infty)$. Assume that $\{F_\theta\}$ is a dominated family and let ν be a σ-finite measure such that $\{F_\theta\} \approx \nu$ i.e., $(\nu(B) > 0 \Leftrightarrow \exists \theta \ni F_\theta(B) > 0)$. Let $f_\theta = dF_\theta/d\nu$. Assume that Θ, $\mathcal{A}$ are both locally compact second countable topological spaces and their σ fields are the respective Borel fields. Let g be a real valued function. If the space is not compact, the symbolism $\liminf_{\theta\to\infty} g(\theta)$ is defined by $\liminf_{\theta\to\infty} g(\theta) = \sup\{\inf\{g(\theta): \theta \notin S\}: S \subset \Theta, S \text{ is compact}\}$. The symbolism $\limsup_{\theta\to\infty} g(\theta)$ and $\lim_{\theta\to\infty} g(\theta)$ is similarly defined.

Assume $L(\theta,\cdot)$ is lower semicontinuous on $\mathcal{A}$ for each $\theta \in \Theta$. If $\mathcal{A}$ is not compact assume there exists a second countable compactification $\mathcal{A}^k$ of $\mathcal{A}$ and a measurable map $h: \mathcal{A}^k \to \mathcal{A}$ such that $L(\theta, h(a)) \leqq \liminf_{\alpha\to a:\, a\in\mathcal{A}} L(\theta, a)$. (If $\mathcal{A}$ is not compact and $\lim_{a\to\infty} L(\theta, a) = \sup_{a\in\mathcal{A}} L(\theta, a)$ this condition is easily satisfied, for then one may choose $\mathcal{A}^k = \mathcal{A} \cup \{\infty\}$ and $h(a) = a$ for $a \in \mathcal{A}$, and $h(\infty) = a_0 \in \mathcal{A}$ for any fixed $a_0 \in \mathcal{A}$.)

PROPOSITION 1. *Under the above assumptions the space of risk functions is "subcompact". That is, define* $\hat{\Gamma} = \{r: \Theta \to [0, \infty] | \exists$ *(measurable) procedure* $\delta \ni r(\theta) \geqq R(\theta, \delta) \forall \theta\}$. *Then* $\hat{\Gamma}$ *is compact in the topology of pointwise convergence.*

PROOF. See Le Cam (1955). □

Let G be any nonnegative σ-finite measure on Θ, $\mathcal{B}_\Theta$ giving finite measure to compact subsets of Θ. Define $B_G(a|x) = \int L(\theta, a) f_\theta(x) G(d\theta) / \int f_\theta(x) G(d\theta) = V_G(a|x)/W_G(x)$. G is called a generalized prior if $G(\Theta) \neq 0$. A procedure δ is called generalized Bayes for G if $\{x: \delta(A(x)|x) \neq 1\}$ has (outer) ν-measure zero where $A(x) = A_G(x) = \{a: V_G(a|x) = \inf_{\alpha\in\mathcal{A}} V_G(\alpha|x)\}$. If $S \in \mathcal{B}_X$ then δ is

Received May 1978; revised July 1978.
[1]This research was supported in part by NSF MCS 75-23343 A01.
AMS 1970 *subject classifications.* Primary: 62C15, 62C07; secondary. 62C10.
Key words and phrases. Admissibility, generalized Bayes procedures.

called generalized Bayes a.e. (ν) on S for G if the (outer) ν-measure of $S \cap \{x: \delta(A(x)|x) \neq 1\}$ is 0. [Some authors use a slightly different terminology. They would require $W_G(x) < \infty$, $x \in S$, before calling δ generalized Bayes on S for G.]

PROPOSITION 2. *Under the above assumptions, for any given generalized prior G there exists a (nonrandomized) generalized Bayes procedure. If, also X is locally compact second countable with Borel field $\mathfrak{B}_X$ then there is a (nonrandomized) generalized Bayes procedure with $\delta(A(x)|x) = 1$ for all $x \in X$.*

Note: this proposition is also valid if $L: \Theta \times \mathcal{A} \to (-\infty, \infty)$ with $L(\theta, a) \geqq \underline{L}(\theta)$ for all $a \in \mathcal{A}$ and $\int |\underline{L}(\theta)| G(d\theta) < \infty$.

PROOF. See Brown and Purves (1973). □

The following theorem involves two further continuity assumptions: Assume $L(\cdot, a): \Theta \to [0, \infty]$ is continuous for each $a \in \mathcal{A}$, and $f_{\cdot}(x): \Theta \to [0, \infty)$ is also continuous for each $x \in X$.

THEOREM. *Make the above assumptions. Let δ_0 be any admissible procedure. If Θ is compact let*

(1) $$S = \{x: x \in X, \sup_{\theta \in \Theta} R(\theta, \delta_0) f_\theta(x) < \infty\}.$$

Then δ_0 is generalized Bayes a.e. (ν) *on S for some prior G_0 with $G_0(\Theta) = 1$.*

If Θ is not compact assume that there is a continuous $h: \Theta \to (0, \infty)$ and a procedure δ' with $R(\theta, \delta') < \infty$ for all $\theta \in \Theta$ and

(2) $$\liminf_{\theta \to \infty} h(\theta)(R(\theta, \delta_0) - R(\theta, \delta')) = \lambda > 0.$$

Let

$$S = \{x: x \in X, \lim_{\theta \to \infty} h(\theta) L(\theta, a) f_\theta(x) = 0 \forall a \in \mathcal{A}, \lim_{\theta \to \infty} h(\theta) f_\theta(x) = 0,$$

and

$$\sup_{\theta \in \Theta} h(\theta) R(\theta, \delta_0) f_\theta(x) < \infty\}.$$

Then δ_0 is generalized Bayes a.e.(ν) *on S for some generalized prior G_0 satisfying*

(3) $$\int h^{-1}(\theta) G_0(d\theta) < \infty.$$

[Note that if $x \in S$ and (3) is satisfied then $W_{G_0}(x) = \int f_\theta(x) h(\theta) h^{-1}(\theta) G_0(d\theta) < \infty$ and $V_{G_0}(a|x) = \int h(\theta) L(\theta, a) f_\theta(x) h^{-1}(\theta) G_0(d\theta) < \infty$. Thus, except for certain trivial statistical situations, not all procedures will be generalized Bayes on S for G_0. In fact if $L(\theta, \cdot)$ is strictly convex than the generalized Bayes procedure for G_0 is (essentially) uniquely determined on S.]

PROOF. Consider the case where Θ is not compact. Let δ_0 be admissible and h, δ', S as in the statement of the theorem. Consider a modified problem, $\mathcal{P}^*$, with loss function $L^*(\theta, a) = (L(\theta, a) - R(\theta, \delta_0)) h(\theta)$. (The risk, etc., in problem $\mathcal{P}^*$ will be denoted by R^*, etc.) Note that in this problem the procedure δ_0 is admissible and has risk function $R^*(\theta, \delta_0) \equiv 0$. Let $r_i(\theta) \equiv -i^{-1}$. Then $r(\cdot) \notin \hat{\Gamma}^*$

since δ_0 is admissible in $\mathcal{P}^*$. Hence $r_i(\cdot)$ can be separated from the compact set $\hat{\Gamma}^*$ by some finite measure, G_i, say. It is easy to check that G_i may be taken to be a probability measure, and

$$(4)\qquad \begin{aligned} -i^{-1} &= \int r_i(\theta) G_i(d\theta) \leqq \inf_{\delta \in \mathfrak{D}} \int R^*(\theta, \delta)\, G_i(d\theta) \\ &= \int R^*(\theta, \delta_i) G_i(d\theta) \leqq \int R^*(\theta, \delta_0) G_i(d\theta) \\ &= 0 \end{aligned}$$

where δ_i denotes a Bayes procedure for G_i. Note that $\delta_i(A^*_{G_i}(x)|x) = 1$ a.e.(ν) by Proposition 2. [Equation (4) merely says that δ_0 is Bayes in the wide sense for problem $\mathcal{P}^*$.]

By taking subsequences, if necessary, assume that $G_i \to G$ "weakly" (in the sense that $\int c(\theta) G_i(d\theta) \to \int c(\theta) G(d\theta)$ for all continuous c such that $\lim_{\theta \to \infty} c(\theta) = 0$). Note that $B^*_{G_i}(a|x) \to B^*_G(a|x)$ for all $x \in S$, $a \in \mathcal{A}$.

Let $B = \sup R^*(\theta, \delta') < \infty$. Let $C \subset \Theta$ be a compact subset such that $R^*(\theta, \delta') < -\lambda/2$ for $\theta \notin C$. C exists by virtue of condition (2). Then $-i^{-1} \leqq \int R^*(\theta, \delta') G_i(d\theta) \leqq B G_i(C) - (\lambda/2)(1 - G_i(C))$. It follows that $\liminf_{i \to \infty} G_i(C) > 0$, and hence $G(C) > 0$.

Let

$$(5)\qquad S_i = \left\{x\colon x \in S, \int B^*_{G_i}(a|x)\delta_0(da|x) - \int B^*_{G_i}(a|x)\delta_i(da|x) \leqq i^{-\frac{1}{2}}\right\}.$$

Then

$$(6)\qquad \int_{x \in S - S_i} \int f_\theta(x) G_i(d\theta)\nu(dx) \leqq i^{-\frac{1}{2}}$$

by (4) since $\int (R^*(\theta, \delta_0) - R^*(\theta, \delta_i)) G_i(d\theta) \geqq \int_{x \in S - S_i} \int B^*_{G_i}(a|x)(\delta_0(da|x) - \delta_i(da|x)) f_\theta(x) G_i(d\theta)\nu(dx)$. Let $S' = \limsup S_i$. Then $\nu(S - S') = 0$ by (6).

Let $x \in S'$. Let $\{i'\} \supset \{i\}$ be a subsequence such that $x \in S_{i'}$ for $i' \in \{i'\}$. Then

$$(7)\qquad \begin{aligned} &\inf_a B^*_G(a|x) \\ &\qquad \geqq \lim_{i' \to \infty} \inf_a B^*_{G_{i'}}(a|x) \\ &\qquad \geqq \liminf_{i' \to \infty} \int B^*_{G_{i'}}(a|x)\delta_0(da|x) \\ &\qquad \geqq \int B^*_G(a|x)\delta_0(da|x). \end{aligned}$$

This proves that δ_0 is generalized Bayes a.e.(ν) on S for the generalized prior G in problem $\mathcal{P}^*$.

Let $G_0(d\theta) = h(\theta)G(d\theta)$. Then, for $x \in S$, $W_{G_0}(x) < \infty$ and

$$\begin{aligned} W_{G_0}(x) B_{G_0}(a|x) &= \int L(\theta, a) f_\theta(x) G_0(d\theta) \\ &= \int h(\theta) L(\theta, a) f_\theta(x) G(d\theta) \\ &= W_G(x) B^*_G(a|x) + \int h(\theta) R(\theta, \delta_0) f_\theta(x) G(d\theta). \end{aligned}$$

since $\int h(\theta) R(\theta, \delta_0) f_\theta(x) G(d\theta) \leqq \sup h(\theta) R(\theta, \delta_0) f_\theta(x) < \infty$. Hence $A_{G_0}(x) = A^*_G(x)$ and so δ_0 is also generalized Bayes in problem $\mathcal{P}$ for the generalized prior

G_0. And, $\int h^{-1}(\theta)G_0(d\theta) = \int G(d\theta) \leqq 1$. This proves the theorem when Θ is not compact.

When Θ is compact the proof is similar, but simpler. Let $L^* = L$ so that $\mathcal{P} = \mathcal{P}^*$. The sequence G_i is constructed as before, and $G_i \to G$ weakly, without loss of generality. (That is, $\int c(\theta)G_i(d\theta) \to \int c(\theta)G(d\theta)$ for all continuous c.) Follow the sequence of steps from (5) through (7) to show that δ_0 is generalized Bayes a.e. (ν) on S for problem $\mathcal{P} = \mathcal{P}^*$. Letting $G_0 = G$ completes the proof since $\int G(d\theta) = \lim_{i\to\infty} \int G_i(d\theta) = 1$. □

We conjecture that the theorem remains true if the continuity condition on f is replaced by the condition that $f\colon \Theta \to L_1(X, \mathcal{B}_X, \nu)$ be continuous.

This theorem has some interesting applications concerning the existence of prior distributions for which the given procedure is Bayes. These follow from the theorem together with the simple extension provided by the following proposition.

PROPOSITION 3. *Suppose δ_0 is generalized Bayes relative to G_0 on S, and $\nu(X - S) = 0$. Then $\int R(\theta, \delta_0)G_0(d\theta) = \inf_\delta \int R(\theta, \delta)G_0(d\theta) \leqslant \infty$.*

PROOF. Let δ_1 be any procedure. Then

$$\begin{aligned}\int R(\theta, \delta_0)G_0(d\theta) &= \int_\Theta \int_X \int_{\mathcal{A}} L(\theta, a)\delta_0(da|x)f_\theta(x)\nu(dx)G_0(d\theta)\\ &= \int_S \int_\Theta \int_{\mathcal{A}} L(\theta, a)\delta_0(da|x)f_\theta(x)G_0(d\theta)\nu(dx)\\ &\leqq \int_S \int_\Theta \int_{\mathcal{A}} L(\theta, a)\delta_1(da|x)f_\theta(x)G_0(d\theta)\nu(dx)\\ &\qquad\qquad (\text{since } \delta_0 \text{ is generalized Bayes})\\ &= \int_\Theta \int_X \int_{\mathcal{A}} L(\theta, a)\delta_1(da|x)f_\theta(x)\nu(dx)G_0(d\theta)\\ &= \int R(\theta, \delta_1)G_0(d\theta).\end{aligned}$$

□

The following corollary and application provide an example of the results attainable.

COROLLARY 1. *Suppose Θ is not compact but the problem has a finite minimax value, $m < \infty$. Suppose*

$$\lim_{\theta\to\infty} f_\theta(x) = 0, \ \lim_{\theta\to\infty} L(\theta, a)f_\theta(x) = 0 \tag{8}$$

for all $x \in X$, $a \in \mathcal{A}$. Let δ_0 be any admissible procedure with $R(\theta, \delta_0) < \infty$ for all $\theta \in \Theta$ such that

$$\liminf_{\theta\to\infty} R(\theta, \delta_0) - m > 0. \tag{9}$$

Then δ_0 is Bayes for some probability measure G_0, and the Bayes risk is finite.

PROOF. Apply the theorem with δ' a minimax procedure (the existence of δ' can be deduced from Proposition 1) and with $h(\theta) = (1 + R(\theta, \delta_0))^{-1}$. Then δ_0 is generalized Bayes for some nonnegative measure G_0 such that $\int(1 + R(\theta, \delta_0))G_0(d\theta) < \infty$. Hence $\int G_0(d\theta) < \infty$ and G_0 can be normalized to be a probability measure. Condition (8) implies that $S = X$ so that δ_0 is Bayes (as well as generalized Bayes) relative to G_0 and $\int R(\theta, \delta_0)G_0(d\theta) < \infty$, as claimed. □

EXAMPLE 1. Consider a location parameter problem with $X = \mathcal{Q} = \Theta = R^k f_\theta(x) = f_0(x - \theta)$ relative to Lebesque measure, and $L(\theta, a) = l(\theta - a)$. Suppose $\lim_{\|t\| \to \infty} l(t) = \infty$ and $\int l(t) f_0(t)\, dt < \infty$. (This implies $m < \infty$.) Consider the linear estimators defined by $\delta_c(\{cx\}|x) = 1, 0 < c < 1$. Any such estimator obviously has $\liminf_{\theta \to \infty} R(\theta, \delta_c) = \infty > m$. Hence, such a linear estimator can be admissible only if it is Bayes. □

It is also possible to use this theorem to derive results concerning the existence of least favorable distributions in testing problems. For example Theorems 3.1 and 4.1 of Lehmann (1952) are direct consequences of the Theorem and Proposition 3.

EXAMPLE 2. For a final application, consider the common problem of estimating a p-dimensional multivariate normal mean θ with squared error loss when the variance covariance matrix is known to be the identity. When the dimension is $p = 2$ then the usual estimator δ_0 (given by $\delta_0(\{x\}|x) = 1$ in the previous notation) is admissible and generalized Bayes for the uniform prior. It follows that no estimator can have smaller risk at ∞ of order $\|\theta\|^{-2}$. (To be precise, $\limsup_{\|\theta\| \to \infty} \|\theta\|^2 (R(\theta, \delta) - R(\theta, \delta_0)) \geqq 0$ for any procedure δ.)

In dimension $p = 1$ one gets only the weaker result that no estimator can have smaller risk than δ_0 of order $\|\theta\|^{-1}$. Here it is possible to have smaller risk at ∞ of order $\|\theta\|^{-2}$ than δ_0. In fact the generalized Bayes estimator for the prior with density $|\theta|\, d\theta$ does have smaller risk than δ_0 of order $\|\theta\|^{-2}$; and this latter estimator cannot be improved in risk at ∞ of order $\|\theta\|^{-2}$. This generalized Bayes estimator is discussed more fully in Brown (1971, page 897).

REFERENCES

BROWN, L. (1971). Admissible estimators, recurrent diffusions and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–903.

BROWN, L. (1979). A heuristic method for determining admissibility of estimators—with applications. *Ann. Statist.* **7** 960–994.

BROWN, L. and PURVES, R. (1973). Measurable selections of extrema *Ann. Statist.* **1** 902–912.

LE CAM., L. (1955). An extension of Wald's Theory of statistical decision functions, *Ann. Math. Statist.* **26** 69–81.

LEHMANN, E. (1952). On the existence of least favorable distributions *Ann. Math. Statist.* **23** 408–416.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
WHITE HALL
ITHACA, NY 14850