The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 446–456

SUFFICIENT CONDITIONS FOR THE ADMISSIBILITY UNDER SQUARED ERRORS LOSS OF FORMAL BAYES ESTIMATORS[1]

BY J. V. ZIDEK

The University of British Columbia

1. Introduction and summary. This paper is concerned with the problem of finding reasonably explicit sufficient conditions for the almost admissibility of formal Bayes estimators (for definitions, see Section 2), where the underlying distribution is assumed known up to a single real parameter, and a real function of this parameter is being estimated with squared error as loss. The parameter space is assumed to be a possibly unbounded interval. These conditions are derived in Section 3. They are similar, in appearance, to results obtained by Karlin [3] when the underlying distribution is a member of the family of one parameter exponential distributions and the mean of this distribution is being estimated. The results of Section 3 should be viewed as a refinement of a heuristic argument given by Stein ([6] and [7] pages 233–240).

In Section 2, some preliminary results and definitions are given. The results of Section 3 are applied in Section 4 to problems involving either the one dimensional exponential family or the estimation of a function of a single location parameter. Some of the results are known, in at least a similar form, while others are new.

A counterexample based on a one dimensional location parameter problem is given in Section 5. It suggests that conditions of the type obtained here may even be necessary.

2. Definitions and preliminary results. Let $(\mathscr{X}, \mathscr{B})$ denote a measurable space and X a random variable taking its values in $\mathscr{X}$. Assume X is distributed according to an unknown but unique member of a family of probability distributions indexed by a set Θ, a subinterval of the real line with upper and lower endpoints θ_u and θ_l, respectively. After observing X, a real-valued function $g: \Theta \to R$ is to be estimated with squared error as loss, that is, if the estimate is t and $\theta \in \Theta$ is the "true state of nature" a loss, $L(t, \theta) = (t-g(\theta))^2$, is incurred.

Suppose μ is a σ-finite measure on $\mathscr{B}$ which dominates the family of underlying probability distributions. Let $p(\cdot \mid \theta)$, $\theta \in \Theta$, denote the density of the probability distribution corresponding to θ. We assume $p(\cdot \mid \cdot)$ is jointly measurable in its arguments.

Suppose given a probability measure Π on the Borel subsets of Θ. The Bayes procedure with respect to Π, ϕ_Π, is given by

$$\phi_\Pi(x) = \int_\Theta g(\theta)p(x \mid \theta)\, d\Pi(\theta) / \int_\Theta p(x \mid \theta)\, d\Pi(\theta), \tag{2.1}$$

provided its Bayes risk is finite.

Received May 31, 1968; revised September 26, 1969.

[1] This research comprised a portion of the author's Ph.D. dissertation completed at Stanford University. It was supported by the Office of Naval Research Contract number Nonr 225(72). Reproduction in whole or in part is permitted for any purpose of the United States Government.

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

ϕ_Π depends on Π only through the posterior probability distribution, P_Π, defined by

$$P_\Pi(B \mid X = x) = \int_B p(x \mid \theta)\, d\Pi(\theta), \qquad \text{a.e. } [\mu], \tag{2.2}$$

for all Borel subsets, B, of Θ.

From some points of view it is reasonable to allow Π to be a σ-finite measure. Provided $P_\Pi(\Theta \mid X = x) < \infty$, a.e. $[\mu]$, Π is called a *prior measure* (*improper* if $\Pi(\Theta) = \infty$). We can define the *formal posterior distribution of* θ using Equation (2.2). A *formal Bayes estimator* of $g(\theta)$ is defined as any measurable function on $\mathscr{X}$ which, evaluated at x, minimizes and makes finite

$$\int (t - g(\theta))^2 P_\Pi(d\theta \mid X = x) \tag{2.3}$$

a.e. $[\mu]$ as a function of t. If such a procedure exists, it is unique and given by (2.1), except, possibly, on a set B for which

$$\int_B d\mu(x) \int_\Theta p(x \mid \theta)\, d\Pi(\theta) = 0. \tag{2.4}$$

The condition

$$\int (1 + g^2(\theta)) p(x \mid \theta)\, d\Pi(\theta) < \infty, \qquad \text{a.e. } [\mu], \tag{2.5}$$

is sufficient to insure that ϕ_Π is the formal Bayes estimator with respect to Π.

Let ϕ denote any estimator of $g(\theta)$. Its risk function will be denoted by $r(\phi, \theta)$, $\theta \in \Theta$. ϕ is called *almost admissible* with respect to Π, if for any other estimator ϕ^*, satisfying $r(\phi^*, \theta) \leqq r(\phi, \theta)$, $\theta \in \Theta$, $r(\phi^*, \theta) = r(\phi, \theta)$ a.e. $[\Pi]$. Obviously, any Bayes procedure is almost admissible with respect to the prior from which it is constructed.

The results of this paper will be concerned with almost admissibility rather than admissibility. A useful theorem giving conditions under which admissibility follows from almost admissibility is the following (see, for example, Stein [8]).

THEOREM 2.1. *Suppose for every element $\theta_0 \in \Theta$ and every set $B \in \mathscr{B}$ for which $\int_B p(x \mid \theta_0)\, d\mu(x) > 0$, $\Pi\{\theta : \int_B p(x \mid \theta)\, d\mu(x) > 0\} > 0$. Then if ϕ_Π is almost admissible, it is admissible.*

Assume Π is absolutely continuous with respect to Lebesgue measure. Denote its density by π. We can regard (X, θ) as an "improper" random variable with a joint distribution whose density with respect to $\mu \times m$ (m denoting Lebesgue measure) is $p(x \mid \theta)\pi(\theta)$.

The formal posterior distribution of θ, given $X = x$, has a density with respect to m. It will be denoted by $p_\Pi(\cdot \mid X = x)$ and is obtained from equation (2.2). It is useful to think of $\int p(x \mid \theta)\pi(\theta)\, d\theta$ as the marginal density of X even though its integral with respect to μ may be infinite.

By adopting this point of view we achieve a simplicity of notation and, in addition, a more intuitive conception of the nature of many otherwise unintuitive quantities which appear in the sequel. We can write, for example, $\phi_\Pi(x) = E_\Pi(g \mid X = x)$ and, letting $\rho(\theta) = r(\phi_\Pi, \theta)$, $\theta \in \Theta$,

$$\int f(\theta)\rho(\theta)\pi(\theta)\, d\theta = E_\Pi\{f(\theta) E[(g(\theta) - E_\Pi(g \mid X))^2 \mid \theta]\}. \tag{2.6}$$

Denote $E(\cdot\,|\,\theta)$ and $E(\cdot\,|\,X)$ by $E^{\theta}(\cdot)$ and $E_{\Pi}{}^{X}(\cdot)$, respectively (similar notation will be used for conditional covariances and variances). The results of this paper will be consequences of a theorem due to Stein [6] (earlier forms of this result are given in [2] and [8]). In stating this result, the following notation will be useful. Let $J \subset \Theta$ be any compact subinterval of Θ. Denote by F_J, the class of all non-negative functions, f, on Θ satisfying $f(\theta) \geqq 1$, $\theta \in J$, $Ef(\theta)\rho(\theta) < \infty$.

THEOREM 2.2. *Suppose Π is a positive prior measure which assigns finite measure to every compact subinterval of Θ which does not contain either endpoint of Θ. If for every such compact subinterval, J, of Θ and $\mathscr{E} > 0$, there exists $f \in F_J$ such that*

$$E\{\{\mathrm{Cov}_{\Pi}{}^{X}(f, g)\}^2 / E_{\Pi}{}^{X}(f)\} < \mathscr{E}, \tag{2.7}$$

ϕ_{Π} is almost admissible, with respect to Π, as an estimator of $g(\theta)$, under squared error loss.

PROOF. Suppose ϕ_{Π} is not almost admissible with respect to Π. Then there exists ϕ^* such that $r(\phi^*, \cdot) \leqq r(\phi_{\Pi}, \cdot)$ and $r(\phi^*, \theta) < r(\phi, \theta)$ for all θ in a set of positive Π measure. It follows that there exists $\delta > 0$, a set S_{δ}, and a compact subinterval J, not containing the endpoints of Θ such that $\Pi(S_{\delta} \cap J) > 0$ and $r(\phi_{\Pi}, \theta) - r(\phi^*, \theta) \geqq \delta$, $\theta \in S_{\delta}$. Choose $\mathscr{E} < \delta\Pi(S_{\delta} \cap J)$ and $f \in F_J$ satisfying (2.7). Then

$$\begin{aligned}\delta\Pi(J \cap S_{\delta}) &\leqq \textstyle\int_{J \cap S_{\delta}} f(\theta)\, d\Pi(\theta)\{r(\phi_{\Pi}, \theta) - r(\phi^*, \theta)\} \\ &\leqq E_{\Pi}(f\rho) - \inf_{\phi} E_{\Pi}\{f(\theta)E^{\theta}(g(\theta) - \phi(x))^2\}\end{aligned}$$

which is less than $\mathscr{E}$, since the last quantity is the left-hand side of (2.7). To see this observe that the infimum in this quantity is

$$E_{\Pi} f(\theta)\{g(\theta) - E_{\Pi}{}^{X}(fg)/E_{\Pi}{}^{X}(f)\}^2 = E_{\Pi}(f\rho) - E\{[\mathrm{Cov}_{\Pi}{}^{X}(f, g)]^2 / E_{\Pi}{}^{X}(f)\}.$$

From the contradiction we have obtained, the conclusion of the theorem follows.

3. A sufficient condition for almost admissibility. In this section, Theorem 2.2. is applied to obtain a sufficient condition for almost admissibility. This condition and Theorem 3.1, in which it is obtained, involve a function $M: \mathscr{X} \times \Theta \to (-\infty, \infty)$ defined by

$$\begin{aligned} M(x, \theta) &= \textstyle\int_{\theta}^{\theta_u} (g(t) - \phi(x))p(t\,|\,X = x)\, dt / p(\theta\,|\,X = x), &\quad p(\theta\,|\,X = x) > 0 \\ &= 0, &\quad p(\theta\,|\,X = x) = 0. \end{aligned} \tag{3.1}$$

In (3.1) we have omitted the subscript Π. As there is little danger of confusion, we shall continue to do so throughout the remainder of this paper. In particular, ϕ will always represent ϕ_{Π}.

Let

$$h(t) = E(M^2(X, \theta)\,|\,\theta = t). \tag{3.2}$$

Assume:

(I) $\pi(t)h(t)$ is bounded away from zero on compact subintervals of Θ

(II) $\{\theta : p(x\,|\,\theta) > 0\}$ is an interval a.e. $[\mu]$.

THEOREM 3.1. *Under Assumptions* I *and* II, ϕ_Π *is almost admissible with respect to* Π *as an estimator of* g *if, when* (i) $\int_c^{\theta_u} \pi(t)\rho(t)\,dt = \infty$, (ii) $\int_c^{\theta_u} dt/[\pi(t)h(t)] = \infty$ *and when* (i)′ $\int_{\theta_l}^c \pi(t)\rho(t)\,dt = \infty$, (ii)′ $\int_{\theta_l}^c dt/[\pi(t)h(t)] = \infty$, *where* $c \in (\theta_l, \theta_u)$.

PROOF. In the notation of Theorem 2.2, suppose $J = [c_1, c_2]$ is a compact subinterval of Θ, where c_i is not an endpoint of Θ, $i = 1, 2$. Let f be any nonnegative function which is absolutely continuous, vanishes outside a compact subinterval of Θ, has bounded derivative, and is identically 1 on J. Then

$$f(\theta) = \int_{\theta_l}^{\theta} f'(t)\,dt + f(\theta_l), \qquad \theta \in \Theta. \tag{3.3}$$

It follows that

$$\mathrm{Cov}^X(f, g) = \int_{\theta}^{\theta_u} f'(t) E^X[(g(\theta) - E^X(g))\psi(\theta, t)]\,dt \tag{3.4}$$

where $\psi(\theta, t)$ is 1 or 0 according as $\theta > t$ or $\theta \leqq t$. For convenience, let $f = r^2$ so that $f' = 2rr'$. Then

$$[\mathrm{Cov}^X(f, g)]^2 = 4\{\textstyle\int r'(t) r(t)(E^X[(g(\theta) - E^X(g))\psi(\theta, t)]/p(t \mid X)) p(t \mid X)\,dt\}^2, \tag{3.5}$$

where, according to a convention that will be adopted here, $(p(t \mid X))^{\frac{1}{2}}/p(t \mid X) = 0$ when $p(t \mid X) = 0$. Equality holds in equation (3.5) because of Assumption II, for it implies $E^X[(g(\theta) - E^X(g))\psi(\theta, t)]$ vanishes when $p(t \mid X)$ vanishes. After applying Schwarz's inequality on the right-hand side of equation (3.5) we obtain

$$E\{[\mathrm{Cov}_X(f, g)]^2 / E^X(r^2(\theta))\} \leqq 4 \textstyle\int (r'(t))^2 \pi(t) h(t)\,dt, \tag{3.6}$$

where h is defined by equation (3.2).

The remainder of the proof is suggested by an argument due to Stein ([7] pages 235–236). Define monotone transformations ψ_i, $i = 1, 2$, by

$$\psi_2(t) = \int_{c_2}^{t} ds/[\pi(s)h(s)], \qquad t \in \Theta \cap (c_2, \theta_u] \quad \text{and} \tag{3.7}$$

$$\psi_1(t) = \int_{t}^{c_1} ds/[\pi(s)h(s)], \qquad t \in \Theta \cap [\theta_l, c_1). \tag{3.8}$$

By Assumption I, ψ_2 (ψ_1) is one-to-one and onto if condition (ii) ((ii)′) holds.

Let A be any positive constant. Let $r(t) = 1$ for $t \in [c_1, c_2]$. If condition (i) fails to hold let $r(t) = 1$, $t \in \Theta \cap (c_2, \theta_u]$. If condition (i) holds, let $r(t) = 1 - A^{-1}\psi_2(t)$, $c_2 < t \leqq \psi_2^{-1}(A)$ and $r(t) = 0$, $\theta_u \geqq t > \psi_2^{-1}A$. In a similar way using ψ_1 and depending on whether (i)′ does or does not hold, $r(t)$ is defined on the remainder of Θ. For definiteness, assume both (i) and (i)′ hold (the remaining cases are treated in a similar way). Then

$$\begin{aligned} r'(t) &= 0, && c_1 < t < c_2 \\ &= -1/[A\pi(t)h(t)], && c_2 < t < \psi_2^{-1}(A) \\ &= 1/[A\pi(t)h(t)], && \psi_1^{-1}(A) < t < c_1 \\ &= 0, && \theta_l < t < \psi_1^{-1}(A), \quad \psi_2^{-1}(A) < t < \theta_u \end{aligned} \tag{3.9}$$

and, consequently,

$$\int (r'(t))^2 \pi(t) h(t)\, dt = A^{-2}\psi_1(\psi_1^{-1}(A)) + A^{-2}\psi_2(\psi_2^{-1}(A)). \tag{3.10}$$

The last quantity is just $2/A$. If A is chosen sufficiently large, condition (2.7) is satisfied. The conclusion then follows from Theorem 2.2.

4. Applications. Theorem 3.1 is applied, in this section, to several examples involving either the one dimensional exponential family or the estimation of a single location parameter.

Consider first the case of the exponential family. The following specializations are involved: $\mathscr{X} \subset (-\infty, \infty)$, $p(x \mid \theta) = \beta(\theta)\exp(x\theta)$,

$$\Theta = \{\theta : 1/\beta(\theta) = \int e^{x\theta}\, d\mu(x) < \infty\}. \tag{4.1}$$

Θ is, of course, an interval.

Suppose we are interested in estimating an arbitrary, piecewise continuous function, g. Define a function π by

$$\pi(\theta) = \exp[-\int_c^\theta g(\omega)\, d\omega]/\beta(\theta), \tag{4.2}$$

where $c \in (\theta_l, \theta_u)$. Then

$$(x - g(\theta))\beta(\theta)\, e^{x\theta}\, \pi(\theta) = (d/d\theta)\beta(\theta)\, e^{x\theta}\, \pi(\theta). \tag{4.3}$$

If

$$\int \beta(\theta)\, e^{x\theta}\, \pi(\theta)\, d\theta = \int \exp[x\theta - \int_c^\theta g(\omega)\, d\omega]\, d\theta < \infty, \tag{4.4}$$

π is the density of a prior measure Π. Furthermore, if

$$\exp[x\theta - \int_c^\theta g(\omega)\, d\omega] \to 0 \tag{4.5}$$

as $\theta \to \theta_u$ or $\theta \to \theta_l$, the formal Bayes or Bayes estimator of g, say ϕ, is easily shown to be $\phi(x) = x$. The function M which appears in equation (3.1) is just 1.

THEOREM 4.1. *Under the conditions given in* (4.4) *and* (4.5), *X is an admissible estimator of g if*

$$\int_c^{\theta_u} \beta(\theta)\exp[\int_c^\theta g(\omega)\, d\omega]\, d\theta = \int_{\theta_l}^c \beta(\theta)\exp[\int_c^\theta g(\omega)\, d\omega]\, d\theta = \infty. \tag{4.6}$$

PROOF. With π defined as in equation (4.2), X is a Bayes or formal Bayes estimator of g. Theorem 3.1 and Theorem 2.1 together imply X is admissible when the hypotheses are satisfied.

A special case of this theorem which has been treated by Cheng Ping [1], is that concerned with the estimation of the function $g(\theta) = \alpha + \gamma E^\theta(X)$, where α and γ are constants. Here

$$\int_c^\theta g(\omega)\, d\omega = \alpha\theta - \alpha c - \gamma \ln \beta(\theta) + \gamma \ln \beta(c). \tag{4.7}$$

We conclude X is an admissible estimator of g provided

$$e^{(x-\alpha)\theta}\, \beta^\gamma(\theta) \to 0 \quad \text{a.e.} \quad [\mu] \tag{4.8}$$

as $\theta \to \theta_l$ or $\theta \to \theta_u$, and

$$\int_c^{\theta_u} \beta^{1-\gamma}(\theta)\, e^{\alpha\theta}\, d\theta = \int_{\theta_l}^{c} \beta^{1-\gamma}(\theta)\, e^{\alpha\theta}\, d\theta = \infty. \tag{4.9}$$

It is not difficult to show that condition (4.4) is fulfilled. Of these conditions, only (4.9) is necessary for Cheng Ping's result since he, like Karlin [3] (who considers the case where $\alpha = 0$) does not require that the estimator in question be a Bayes or formal Bayes estimator.

We turn now to the problem of estimating a single location parameter, with a single observation drawn from the underlying distribution. In this example, the following specializations occur: $\Theta = \mathscr{X} = (-\infty,\infty)$, $\mu(dx) = dx$, $p(x \mid \theta) = p(x-\theta)$, $E^\theta(X) = \theta$. In order that Assumption II hold it is necessary to assume I$'$: $\{x : p(x) > 0\}$ is an interval. We also impose condition II$'$: π is the density of a prior measure, is continuous and satisfies

$$\pi(t)/\pi(\theta) \leqq a_1 + a_2 \left|t-\theta\right|^\alpha, \qquad -\infty < t, \quad \theta < \infty, \tag{4.10}$$

where α and the a_i, $i = 1, 2$ are nonnegative constants.

Without real loss of generality we take, in Assumption II$'$, $a_1 = a_2 = 1$. The formal Bayes or Bayes estimator of θ, ϕ is given by

$$\phi(x) = \int \theta p(x-\theta)\pi(\theta)\, d\theta / \int p(x-\theta)\pi(\theta)\pi(\theta)\, d\theta.$$

And the following theorem gives conditions for its almost admissibility. Its proof, which is straightforward, is omitted. We remark that hypothesis (ii) is used to obtain a constant which is a uniform bound for h.

THEOREM 4.2. *Subject to Assumptions* I$'$ *and* II$'$, ϕ *is almost admissible as an estimator of* θ *if*

(i) $\int_0^\infty d\theta/\pi(\theta) = \int_{-\infty}^0 d\theta/\pi(\theta) = \infty$ *and*

(ii) $\int p(x)(1+|x|^{2\alpha})\{[p(x)]^{-1}\int I(x,t)(1+|t|^{\alpha+1})p(t)\,dt\}^2\, dx < \infty$, *where*

$$I(x,t) = 1, \qquad t > x, \quad x > 0$$
$$= 0, \qquad t \leqq x, \quad x > 0$$
$$= 1 - I(|x|, t), \qquad x \leqq 0.$$

Theorem 4.2 covers the case of estimating the mean of a normal distribution where, without loss of generality, it is assumed that one observation is taken. From it, we conclude that any formal Bayes estimator with respect to a prior measure having a density, π, which satisfies Assumption II$'$ and hypothesis (i) of the theorem, is almost admissible.

Before considering a generalization of the single observation location parameter problem, we prove a lemma which leads to a simplification of the conditions of Theorem 3.1 at the expense of imposing a greater number of assumptions.

Assume I$''$: $\pi(\theta)h(\theta)$ is bounded away from zero on compact subsets of $(-\infty, \infty)$ and II$''$: $g(\theta_u) = -g(\theta_l) = \infty$; g is continuously differentiable and $g' > 0$.

Define A_θ by $A_\theta = \{x : x \in \mathscr{X},\ \phi(x) < g(\theta)\}$. Suppose *given* positive functions, c and d, each with domain $\mathscr{X} \times \Theta$. Define a function $H: \Theta \times \mathscr{X} \times \Theta \to (0, \infty)$ by

$$
\begin{aligned}
H_{m'}(t; x, \theta) &= \sup_{t' \geqq t} m(t'; x, \theta) p(t' \mid x)/g'(t), && x \in A_\theta \\
&= \sup_{t' \leqq t} m(t'; x, \theta) p(t' \mid x)/g'(t), && x \notin A_\theta \\
&= 0, && \text{otherwise,}
\end{aligned}
\tag{4.11}
$$

where

$$
m(t; x, \theta) = [1 + ((g(t) - \phi(x))/c(x, \theta))^2]^{m'}, \qquad x \in \mathscr{X},\quad \theta \in \Theta,\quad t \in \Theta, \tag{4.12}
$$

m' is any real number, with $m' > 1$.

Assume III″: $H_{m'}(\theta; x, \theta) g'(\theta)/[p(\theta \mid x) m(\theta; x, \theta)]$ is bounded by $d^2(x, \theta)$, for some $m' > 1$.

LEMMA 4.1. *Under Assumptions* I″–III″, ϕ *is an almost admissible estimator of* $g(\theta)$ *provided*

(i) $\int_c^{\theta_u} (g'(\theta))^2 / [\pi(\theta)(\mu(\theta) + \mu_4(\theta))]\, d\theta = \infty$ *when* $\int_c^{\theta_u} \pi(\theta)\rho(\theta)\, d\theta = \infty$,

and

(ii) $\int_{\theta_l}^{c} (g'(\theta))^2 / [\pi(\theta)(\mu(\theta) + \mu_4(\theta))]\, d\theta = \infty$ *when* $\int_{\theta_l}^{c} \pi(\theta)\rho(\theta)\, d\theta = \infty$,

where $c \in (\theta_l, \theta_u)$,

$$
\begin{aligned}
\mu(t) &= E\{c^4(X, \theta)\, d^4(X, \theta) \mid \theta = t\}, \\
\mu_4(t) &= E\{d^4(X, \theta)(g(\theta) - \phi(X))^4 \mid \theta = t\}.
\end{aligned}
\tag{4.13}
$$

PROOF. Suppose $x \in A_\theta$, $t \in (\theta_l, \theta_u)$. Then

$$
\begin{aligned}
&-\frac{1}{2}\frac{\partial}{\partial t}\left\{H_{m'}(t; x, \theta) c^2(x, \theta)\left[1 + \left(\frac{g(t) - \phi(x)}{c(x, \theta)}\right)^2\right]^{-(m'-1)} \Big/ (m'-1)\right\} \\
&\quad = \frac{1}{2}\left(-\frac{\partial}{\partial t} H_{m'}(t; x, \theta)\right) c^2(x, \theta)\left[1 + \left(\frac{g(t) - \phi(x)}{c(x, \theta)}\right)^2\right]^{-(m'-1)} \Big/ (m'-1) \\
&\qquad + H_{m'}(t; x, \theta) g'(t)[g(t) - \phi(x)]/m(t; x, \theta) \\
&\quad \geqq H_{m'}(t; x, \theta) g'(t)[g(t) - \phi(x)]/m(t; x, \theta).
\end{aligned}
$$

Similarly, for $x \notin A_\theta$, $\theta_l < t < \theta_u$,

$$
\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial t}\left\{H_{m'}(t; x, \theta) c^2(x, \theta)\left[1 + \left(\frac{g(t) - \phi(x)}{c(x, \theta)}\right)^2\right]^{-(m'-1)}\right\} \Big/ (m'-1) \\
&\qquad \geqq H_{m'}(t; x, \theta) g'(t)\, |g(t) - \phi(x)| / m(t; x, \theta).
\end{aligned}
$$

Recall that if f is a monotone increasing function on $[a, b]$, f' exists almost everywhere and

$$
\int_a^b f'(x)\, dx \leqq f(b) - f(a)
$$

(see, for example, Royden [4] page 82). Using this fact and the first of the two inequalities obtained above, with $x \in A_\theta$,

$$M(x,\theta) \leqq \int_\theta^{\theta_u} g'(t)[g(t)-\phi(x)]H_{m'}(t;x,\theta)/p(\theta \mid x)m(t;x,\theta)\,dt$$
$$\leqq d^2(x,\theta)c^2(x,\theta)\left[1+\left(\frac{g(\theta)-\phi(x)}{c(x,\theta)}\right)^2\right]\Big/\{2(m'-1)g'(\theta)\}.$$

Observe that in evaluating the integral which bounds the integral in the first of these last two inequalities at its upper limit we obtain

$$-\lim_{t\to\theta_u} \tfrac{1}{2}H_{m'}(t;x,\theta)c^2(x,\theta)\left[1+\left(\frac{g(t)-\phi(x)}{c(x,\theta)}\right)^2\right]^{-(m'-1)} = 0$$

since $H_{m'}(t;x,\theta) \leqq H_{m'}(\theta;x,\theta)$, $t \geqq \theta$ while $g(t) \to \infty$ as $t \to \theta_u$.

The same inequality holds when $x \notin A_\theta$. This is proved with the help of the identity

$$\int_\theta^{\theta_u} (g(t)-\phi(x))p(t \mid X = x)\,dt = \int_{\theta_l}^{\theta} |g(t)-\phi(x)|\, p(t \mid X = x)\,dt, \tag{4.14}$$

for $x \notin A_\theta$. Thus, there exists a constant M such that

$$M^2(x,\theta) \leqq Md^4(x,\theta)\{c^4(x,\theta)+(g(\theta)-\phi(x))^4\}/(g'(\theta))^2. \tag{4.15}$$

Using the hypotheses of this lemma and Theorem 3.1, together with inequality (4.15), the desired conclusion follows.

While there do not appear to be natural choices for the functions, c, and d, in the general problem, we believe that in applications they may be suggested by the structure of the problem.

Suppose we observe a random variable (X, Y) with X real valued and Y taking its values in a space $\mathscr{Y}$ and having a marginal distribution ν. Furthermore, given $Y = y$ we assume X has a conditional distribution with density given by $p^*(x-\theta \mid y)$ where $\theta \in \Theta = (-\infty, \infty)$ is unknown and p^* satisfies

$$\int p^*(x \mid y)\,dx = 1, \qquad \int xp^*(x \mid y)\,dx = 0.$$

In this context, the function h, defined in equation (3.2) can be written as

$$\int d\nu(y) \int dx \left(\int_\theta^\infty [g(t)-\phi(x,y)]p^*(x-t \mid y)\pi(t)\,dt/[p^*(x-\theta \mid y)\pi(\theta)]\right)^2 \times p^*(x-\theta \mid y), \tag{4.16}$$

where $\phi(x, y)$ denotes the formal Bayes estimator with respect to the prior measure with density π.

Reasonable choices, in this problem, of the functions c and d, which were supposed given in Lemma 4.1, seem to be $c \equiv d$ with $d^2(x, y, \theta) = 1+\sigma^2(x-\theta, y)$, where $\sigma^2(x, y) = E[(g(\theta)-\phi(x, y))^2 \mid X = x, Y = y]$.

Assume I$'''$: π is continuous, is the density of a prior measure, and satisfies

$$\pi(t)/\pi(\theta) \leqq 1+|\theta-t|^\alpha, \tag{4.17}$$

II$'''$: $g(\infty) = -g(-\infty) = \infty$. g is continuously differentiable and $g' > 0$, and

III$'''$: $g'(\theta)H^*_{m'}(\theta;x,y,\theta) \div \{p^*(x-\theta \mid y)[1+(g(\theta)-\phi(x,y))^2/(1+\sigma^2(x-\theta,y))]^{m'}\}$

is uniformly bounded by $1+\sigma^2(x-\theta, y)$ for some $m' > 1$, where if

$$(4.18)\quad G(t; x, y, \theta) = (1+|t-\theta|^{\alpha})[1+(g(t)-\phi(x, y))^2/(1+\sigma^2(x-\theta, y))]^{m'} \times p^*(x-t \mid y)/g'(t),$$

we define $H^*_{m'}$ by

$$(4.19)\quad \begin{aligned} H^*_{m'}(\theta; x, y, \theta) &= \sup_{t \geqq \theta} G(t; x, y, \theta), \qquad \theta > \phi(x, y) \\ &= \sup_{t \leqq \theta} G(t; x, y, \theta), \qquad \theta \leqq \phi(x, y). \end{aligned}$$

THEOREM 4.3. *Under Assumptions* I‴–III‴, *above*, ϕ *is an almost admissible estimator of* g *provided*

(i) $\int_0^\infty (g'(\theta))^2\, d\theta/[\pi(\theta)(1+\mu_4(\theta))] = \infty$ *when* $\int_0^\infty \pi(\theta)\rho(\theta)\, d\theta = \infty$;

(ii) $\int_{-\infty}^0 (g'(\theta))^2\, d\theta/[\pi(\theta)(1+\mu_4(\theta))] = \infty$ *when* $\int_0^\infty \pi(\theta)\rho(\theta)\, d\theta = \infty$,
with μ_4 *as defined in equation* (4.13);

(iii) $\int d\nu(y) \int dx\, p^*(x \mid y)\{E[(g(\theta)-E^{X,Y}(g))^4 \mid X = x, Y = y]\}^{\frac{1}{2}} < \infty$.

PROOF. Observe that $\mu(\theta)$ (see equations (4.13)) is uniformly bounded by the quantity whose finiteness is asserted in hypothesis (iii). Thus the conclusion of this theorem is an immediate consequence of Lemma 4.1.

In [5], Stein treats the case where $g(\theta) = \theta$ and $\pi = 1$. He concludes X is an admissible estimator of θ if

$$(4.20)\quad \int d\nu(y)[\int x^2 p^*(x \mid y)\, dx]^{\frac{1}{2}} < \infty.$$

Subject to Assumptions I‴–III‴, Theorem 4.3 yields the same conclusion (with the help of Theorem 2.1) if

$$(4.21)\quad \int d\nu(y)[\int x^4 p^*(x \mid y)\, dx]^{\frac{1}{2}} < \infty.$$

5. Counter example. In order to make hypothesis (i) of Theorem 4.2 seem more natural, a counter example will now be given, which suggests that this condition may even be necessary. Suppose a single observation is drawn from a uniform distribution on $[\theta-\frac{1}{2}, \theta+\frac{1}{2}]$. Let $\pi(\theta) = 1+|\theta|^{\alpha}$, $\alpha > 1$. Then the formal Bayes estimator of θ, with respect to the prior measure determined by π, say ϕ_α, is

$$(5.1)\quad \phi_\alpha(x) = [x+(|x+\tfrac{1}{2}|^{2+\alpha}-|x-\tfrac{1}{2}|^{2+\alpha})/(2+\alpha)] \times \{1+(\operatorname{sgn}(x+\tfrac{1}{2})\,|x+\tfrac{1}{2}|^{1+\alpha}-\operatorname{sgn}(x-\tfrac{1}{2})\,|x-\tfrac{1}{2}|^{1+\alpha})/(1+\alpha)\}^{-1}.$$

Assume $x > \frac{1}{2}$. Then the quantity $(1-1/(x+\frac{1}{2}))^a$, $a = \alpha+1$, $\alpha+2$, can be expanded in powers of $z^{-1} = (x+\frac{1}{2})^{-1}$. After doing so we obtain

$$(5.2)\quad \phi_\alpha(x) = x+z^{-1}\textstyle\sum_{r=0}^\infty (a_r/z^r)[1+z^{-\alpha}+z^{-1}\sum_{r=0}^\infty b_r/z^r]^{-1},$$

where $a_r = (r+1)/[2(r+2)(r+3)]\binom{\alpha}{r+1}(-1)^r$, $b_r = \binom{\alpha+1}{r+2}(-1)^{r+1}(\alpha+1)$, $r = 0, 1, 2, \cdots$. Assume x is sufficiently large, say $x > M_1 > \frac{1}{2}$ so that $|z^{-\alpha}+z^{-1}\sum_{r=0}^\infty b_r/z^r| < 1$. Then, after expanding $[1+z^{-\alpha}+(z^{-1}\sum_{r=0}^\infty b_r/z^r)]^{-1}$, in powers of $z^{-\alpha}+z^{-1}\sum_{r=0}^\infty b_r/z^r$

and collecting terms in the resulting expression for ϕ_α, we obtain $\phi_\alpha(x) = x+k(z)/z$, where k is of the form

$$(5.3)\qquad k(z) = \frac{\alpha}{12}+\frac{\alpha}{24}(z^{-1})+\frac{\alpha(8\alpha^2-16\alpha+9)}{360}z^{-2}-\frac{\alpha}{12}z^{-\alpha}-\frac{\alpha}{12}z^{-2\alpha} -\frac{\alpha(\alpha+1)}{24}z^{-(\alpha+1)}+o(z^{-2\alpha})+o(z^{-2}).$$

The risk of ϕ_α is, for $\theta > M_1+\frac{1}{2}$,

$$(5.4)\qquad r(\theta,\phi_\alpha) = \rho+2E^\theta([Z-\theta-\tfrac{1}{2}]k(Z)/Z)+E^\theta(k(Z)/Z)^2,$$

where $\rho = E^\theta(X-\theta)^2$ and $Z = X+\frac{1}{2}$. After considerable computation, this reduces to

$$(5.5)\qquad r(\theta,\phi_\alpha) = \rho+\frac{\alpha^2-2\alpha}{144}\theta^{-2}+o(\theta^{-2})+O(\theta^{-(1+\alpha)}).$$

Observe that $\alpha^2-2\alpha$ is increasing for $\alpha > 1$. Choose β subject to $1 < \beta < \alpha$ and $M_2(>M_1+\frac{1}{2})$ large enough so that if $\theta > M_2$, $r(\cdot, \phi_\alpha) > r(\cdot, \phi_\beta)$. Furthermore, choose $M_3 > M_2$ so that on $[M_3, \infty)$, $\phi_\alpha-\phi_\beta > 0$ and $\phi_\alpha+\phi_\beta$ is strictly increasing. Define an estimator, ϕ^*, by

$$(5.6)\qquad \begin{aligned}\phi^*(x) &= \phi_\alpha(x), \qquad x < M_3+1,\\ &= \phi_\beta(x), \qquad x \geqq M_3+1.\end{aligned}$$

As we shall see, $r(\theta, \phi_\alpha) \geqq r(\theta, \phi^*)$ with strict inequality when $\theta-\frac{1}{2} > M_3+1$.

Let P^θ denote the (uniform) probability distribution of X. Then, for $\theta+\frac{1}{2} \leqq M_3+1$, $P^\theta(\phi^*(x) = \phi_\alpha(x)) = P^\theta(X < M_3+1) = 1$. Consequently, $r(\theta, \phi_\alpha) = r(\theta, \phi^*)$ for this range of θ. Now suppose $\theta-\frac{1}{2} > M_3+1$. Then $P^\theta(\phi^*(x) = \phi_\beta(x)) = P^\theta(X \geqq M_3+1) = 1$, and $r(\theta, \phi_\beta) < r(\theta, \phi_\alpha)$ because of equation (5.5) and $\theta > M_2$. It remains to consider the range where $\theta+\frac{1}{2} > M_3+1 \geqq \theta-\frac{1}{2}$. There

$$r(\theta,\phi_\alpha)-r(\theta,\phi^*) = \int_{M_3+1}^{\theta+\frac{1}{2}} (\phi_\alpha(x)-\theta)^2-(\phi_\beta(x)-\theta)^2\,dx.$$

For convenience let $T(x;\theta) = (\phi_\alpha(x)-\theta)^2-(\phi_\beta(x)-\theta)^2$. Then $T(x;\theta) = (\phi_\alpha(x)-\phi_\beta(x))(\phi_\alpha(x)+\phi_\beta(x)-2\theta)$. Also $\phi_\alpha(x)-\phi_\beta(x) > 0$, for $x > M_3$ and $\phi_\alpha(x)+\phi_\beta(x)$ is strictly increasing for $x > M_3$. Thus either $T(x;\theta) > 0$ for all x in $[M_3+1, \theta+\frac{1}{2}]$ in which case $r(\theta, \phi_\alpha)-r(\theta, \phi^*) > 0$ and the proof is complete or $T(x;\theta) = 0$ at a unique point $x = x_0 \in [M_3+1, \theta+\frac{1}{2}]$. This last result is a consequence of the strictly increasing character of $(\phi_\alpha(x)+\phi_\beta(x)-2\theta)$ and

$$(5.7)\qquad 0 < r(\theta,\phi_\alpha)-r(\theta,\phi_\beta) = \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} T(x;\theta)\,dx$$

which imply $T(\theta+\frac{1}{2};\theta) > 0$. It follows that $T(x;\theta) < 0$ for $\theta-\frac{1}{2} \leqq x < M_3+1$. Thus, using (5.7), $\int_{M_3+1}^{\theta+\frac{1}{2}} T(x;\theta)\,dx > 0$ which completes the argument. Thus ϕ_α is inadmissible.

Acknowledgment. I am indebted to Professor Charles Stein for suggesting the problem to which the work of this paper is directed and for guiding me in the research reported in it. I wish to thank the referee for suggestions which led to an improvement in the presentation of these results.

REFERENCES

[1] CHENG, PING (1964). Minimax estimates of parameters of distributions belonging to the exponential family. *Chinese Math.—Acta.* **5** 277–299.

[2] JAMES, W. and STEIN, CHARLES (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press, 361–379.

[3] KARLIN, S. (1968). Admissibility for estimation with quadratic loss. *Ann. Math. Statist.* **29** 411–415.

[4] ROYDEN, H. L. (1963). *Real Analysis.* Macmillan, New York.

[5] STEIN, C. (1959). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30** 970–979.

[6] STFIN, C. (1965). Address given at the Fifth Berkeley Symposium. Math. Statist. Prob. (unpublished).

[7] STEIN, C. (1965a). Approximation of improper prior measures by prior probability measures, in *Bernoulli, Bayes, Laplace, Anniversary Volume.* Springer Verlag, New York.

[8] STEIN, C. (1965b). Unpublished notes on decision theory.