GENERALIZED BAYES SOLUTIONS IN ESTIMATION PROBLEMS[1]

BY JEROME SACKS

Northwestern University

0. Summary. In estimation problems where the parameter space is not compact the class of Bayes solutions ($\mathcal{B}$) is usually not a complete class and it is necessary to take the closure (in a suitable sense) of $\mathcal{B}$ to obtain a complete class. When the parameter to be estimated is that of an exponential density the limits of Bayes solutions can be characterized as generalized Bayes solutions in the sense that they minimize a posteriori risk where the a priori distribution may have infinite variation (theorem and corollaries in Section 2). The extent to which exponential densities are necessary for this characterization and some consequences of this characterization are contained in a series of remarks at the end of Section 2. In Section 1 we motivate the ideas by obtaining the above characterization for the problem of estimating the mean of a normal distribution with the parameter space restricted to $[0, \infty)$ and with squared-error loss function.

1. Introduction. In estimating the mean of a normal distribution with known variance and squared-error loss function it has long been observed that the usual estimator (the sample mean, $\bar{x}$) while not a Bayes estimator is the limit of a sequence of Bayes estimators and, moreover, can be thought of as a Bayes estimator if the notion of a priori distribution is enlarged to include Lebesgue measure on the real line. In particular $\bar{x}$ can be written as the a posteriori expected value of the unknown mean where the "a priori distribution" is Lebesgue measure. One of the main consequences of the Wald decision theory is that the class of Bayes estimators and their limits (in an appropriate sense) form a complete class. In analogy with $\bar{x}$ it is natural then to inquire whether every limit of Bayes estimators bears a similar representation, namely, if τ is a limit of Bayes estimators is there a measure F on the real line such that τ can be written as the a posteriori expected value of the mean when F is the "a priori distribution"? If τ can be so represented we will call it a Generalized Bayes Solution (GBS) with respect to F. It is a consequence of the Theorem of Section 2 that, for this normal example, every limit of Bayes solutions is a GBS, i.e., the question raised above has an affirmative answer.

With the normal example of the last paragraph as motivation let us consider the more general estimation problem with the space of states of nature $\Omega = \{\omega\}$ a subset of the reals, decision space $T = \{t\} = \Omega$, loss function W, a fixed number of observations n, and $x = (x_1, \cdots, x_n)$ the observation vector having density $p(x, \omega)$ with respect to some σ-finite measure μ when ω is "true." For such a

Received May 28, 1962.

[1] Research partly sponsored by Air Force Contract AF49(638)-877 and NSF Grant G24500.

problem we will call a decision function δ a GBS with respect to a measure F on Ω which gives finite measure to bounded subsets of Ω if, for almost all $x(\mu)$, δ selects (perhaps in a randomized way) a decision among those t's which minimize the (generalized) a posteriori loss

$$(1.1) \qquad \int_\Omega W(\omega, t)p(x, \omega)F(d\omega) \bigg/ \int_\Omega p(x, \omega)F(d\omega).$$

When F is a probability measure or a finite measure (as would be the case if Ω were bounded) δ is a Bayes solution so that the generalization is reached only when Ω is unbounded and $F(\Omega) = +\infty$. We are only concerned with F's such that (1.1) is finite for some t and a.e. (x) as is the case in the context of the normal example when F is Lebesgue measure. The question that this article deals with is: what is the relation of the closure in the sense of regular convergence (see below and Section 2) of the class of Bayes solutions (call the closure $\mathcal{B}'$) to the class of GBS's ($\mathcal{B}^*$)? For the appropriate decision theoretic definitions and results we refer to LeCam's [3] generalization of the Wald theory.

Let $\mathcal{B}_0$ denote that part of $\mathcal{B}'$ which excludes those δ's with infinite risk (hence inadmissible). $\mathcal{B}_0$ is obviously the pertinent part of $\mathcal{B}'$ and is a complete class because $\mathcal{B}'$ is. The Theorem of Section 2, which is stated for exponential densities and $\Omega = (-\infty, +\infty)$, shows that, under the assumptions made there, if $\delta \in \mathcal{B}_0$ then there is an F such that (1.1) is satisfied except, possibly, for certain x's. When these exceptional x's have measure 0 then we can conclude (Corollary 1) that $\mathcal{B}_0 \subset \mathcal{B}^*$ and therefore $\mathcal{B}^*$ is a complete class. Corollary 2 shows that if all members of $\mathcal{B}_0$ are nonrandomized then $\mathcal{B}_0 = \mathcal{B}^*$. Remark 3 at the end of Section 2 indicates how the theorem and corollaries may be generalized to include other than exponential densities but it is not clear how far matters may be extended nor to what extent the assumptions we make are essential.

As indicated below, the arguments of Section 2 work when, for example, $\Omega = [0, \infty)$; indeed the proofs are somewhat simpler in this case. We choose to prove the results in the case $\Omega = (-\infty, +\infty)$ because the argument for the full-line requires certain considerations which do not appear in the half-line case and which will also appear when generalizing to higher dimensions (see Remark 4). The main ideas, however, are already present in the half-line case. In fact, to exhibit the main ideas in Section 2 let us consider the example of estimating the mean ω of a normal distribution with variance 1 where ω is known to be non-negative, the loss function is squared-error, and with one observation. Let $\{\xi_n\}$ be a sequence of a priori distributions and let $\{\delta_n\}$ be the corresponding Bayes solutions. Since the loss is squared error, δ_n is non-randomized and let us denote by $\tau_n(x)$ the point (in $[0, \infty)$) which δ_n selects when x is observed. Suppose $\tau_n(x) \to \tau(x) < \infty$ a.e. (Lebesgue measure). We want to show that there is a non-decreasing function F on $[0, \infty)$ such that

$$(1.2) \qquad \tau(x) = \int_0^\infty \omega p(x, \omega)\, dF \bigg/ \int_0^\infty p(x, \omega)\, dF$$

where $p(x, \omega) = e^{-\frac{1}{2}\omega^2} e^{x\omega}$. Note that since δ_n is Bayes with respect to ξ_n,

$$\tau_n(x) = \int_0^\infty \omega p(x, \omega)\, d\xi_n \Big/ \int_0^\infty p(x, \omega)\, d\xi_n$$

We show first that there is an $a > 0$ such that

$$\limsup_{n\to\infty} \xi_n(v)/\xi_n(a) < \infty \tag{1.3}$$

for any $v > 0$ (compare Lemma 1 below). If (1.3) is not the case then for each a there is a $v > a$ and a sequence $\{n_j\}$ such that $\lim_{j\to\infty} \xi_{n_j}(v)/\xi_{n_j}(a) = +\infty$ (define $\xi_n(v)/\xi_n(a) = +\infty$ if $v > a$ and $\xi_n(a) = 0$). Let

$$\gamma(a) = \inf \{p(x, \omega) \mid \omega \in [a, v]\}.$$

Then, for any x,

$$\begin{aligned} \lim_{j\to\infty} \int_0^\infty p(x, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)} &\geqq \lim_{j\to\infty} \int_a^v p(x, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)} \\ &\geqq \gamma(a) \lim_{j\to\infty} \frac{\xi_{n_j}(v) - \xi_{n_j}(a)}{\xi_{n_j}(a)} = +\infty. \end{aligned} \tag{1.4}$$

Hence

$$\begin{aligned} \tau(x) = \lim_{j\to\infty} \tau_{n_j}(x) &= \lim_{j\to\infty} \frac{\int_0^a \omega p(x, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)} + \int_a^\infty \omega p(x, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)}}{\int_0^a p(x, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)} + \int_a^\infty p(x, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)}} \\ &= \lim_{j\to\infty} \int_a^\infty \omega p(x, \omega)\, d\xi_{n_j} \Big/ \int_a^\infty p(x, \omega)\, d\xi_{n_j} \geqq a. \end{aligned} \tag{1.5}$$

But if (1.5) is true for any x and all $a > 0$ we must have $\tau(x) = +\infty$ for all x, which violates our assumption that τ is finite valued. Thus (1.3) is established.

Put $F_n = \xi_n/\xi_n(a)$. As a consequence of (1.3) we can find a subsequence F_{n_k} and a non-decreasing function F such that $F_{n_k} \to F$ and $F(\omega) < \infty$ for each ω (of course $F(+\infty)$ may be $+\infty$). We might as well suppose that $\{F_{n_k}\}$ is the sequence $\{F_n\}$ so that $F_n \to F$.

The next step is to show that

$$\limsup_{n\to\infty} \int_0^\infty p(x, \omega)\, dF_n < \infty \quad \text{a.e.} \tag{1.6}$$

(compare Lemma 2 below). If (1.6) fails for some x then there is a subsequence $\{n_j\}$ such that

$$\lim_{j\to\infty} \int_0^\infty p(y, \omega) \frac{d\xi_{n_j}}{\xi_{n_j}(a)} = +\infty \tag{1.7}$$

for all $y \geqq x$. From (1.3) and (1.7) we conclude that

$$\lim_{j\to\infty} \int_0^\infty p(y, \omega)\, d\xi_{n_j}/\xi_{n_j}(v) = \infty$$

for all $v > 0$ so that, by (1.4) and (1.5) (replacing a there by v), we obtain $\tau(y) = +\infty$ for all $y \geqq x$. This contradiction establishes (1.6).

The final step is to verify

$$\lim_{n\to\infty} \int_0^\infty \omega p(x, \omega)\, dF_n = \int_0^\infty \omega p(x, \omega)\, dF \quad \text{a.e.} \tag{1.8}$$

Since $\omega p(x, \omega)$ is continuous in ω for each x, we need only show uniform integrability, i.e.,

$$\lim_{A\to\infty} \limsup_{n\to\infty} \int_A^\infty \omega p(x, \omega)\, dF_n = 0. \tag{1.9}$$

For $y > x$ observe that

$$\int_A^\infty \omega p(x, \omega)\, dF_n = \int_A^\infty \omega e^{(x-y)\omega} p(y, \omega)\, dF_n \leqq R(A) \int_A^\infty p(y, \omega)\, dF_n$$

where $R(A) = \sup_{\omega \geqq A} \omega e^{(x-y)\omega} \to 0$ as $A \to \infty$. Thus, if (1.9) is false then, for all $y > x$, we must have $\limsup_{n\to\infty} \int_A^\infty p(y, \omega)\, dF_n = +\infty$ which contradicts (1.6). The same argument shows that $\lim_{n\to\infty} \int_0^\infty p(x, \omega)\, dF_n = \int_0^\infty p(x, \omega)\, dF$.

Note that the first part of the argument ((1.3)) results from the observation: if $\{\xi_n\}$ sends mass to infinity "too rapidly" then τ_n could not converge to a finite τ. The other parts of the argument ((1.6) and (1.9)) show that we can take the appropriate limits under the integral sign. It will be seen in the example in Remark 3 that, even though $\{\xi_n\}$ behaves properly, the non-interchangeability of limit with integration can cause the theorem to fail when the assumptions below are not satisfied.

2. Results and proofs. Let μ be a σ-finite measure on the real line and let the space of states of nature $\Omega = \{\omega \mid \int e^{x\omega}\mu(dx) < \infty\}$. For $\omega \in \Omega$ let $p(x, \omega) = \rho(\omega)e^{x\omega}$ where $\rho^{-1}(\omega) = \int e^{x\omega}\mu(dx)$. We will assume that μ is such that $\Omega = (-\infty, +\infty)$. It will be easy to see that there is nothing lost by this apparent lack of generality in assuming Ω to be the full real line rather than some interval. Let the decision space T be the full real line. We suppose one observation whose density with respect to μ is $p(x, \omega)$ when $\omega \in \Omega$ is the state of nature; because the densities are exponential the same results will hold for any (fixed) number of observations.

For our purposes we define a decision function δ by: for fixed x, δ is a probability measure on the Borel sets of the extended real line, and for a fixed Borel set A, δ is μ-measurable in x. We represent the value δ assigns to the pair A, x by $\delta(A \mid x)$. A sequence $\{\delta_n\}$ is said to converge regularly to δ if (LeCam [3]), for every $f \varepsilon L_1(\mu)$ and every bounded continuous real valued function u on the extended real line

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x)u(t)\delta_n(dt \mid x)\mu(dx) \to \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x)u(t)\delta(dt \mid x)\mu(dx).$$

A consequence of $\delta_n \to \delta$ regularly is:

$$\begin{aligned}&\lim_{n\to\infty}\delta_n(C \mid x) = 0 \text{ for all bounded sets } C \text{ and all } x \varepsilon E \text{ with}\\ &\mu(E) > 0 \text{ implies } \delta(C \mid x) = 0 \text{ for all bounded } C \text{ a.e. } (\mu) \text{ on } E.\end{aligned} \tag{2.1}$$

A further consequence is:

$$\begin{aligned}&\lim_{n\to\infty}\delta_n(C \mid x) = 1 \text{ for } x \varepsilon E \text{ with } \mu(E) > 0 \text{ and}\\ &\qquad C \text{ compact implies } \delta(C \mid x) = 1 \text{ a.e. } (\mu) \text{ on } E.\end{aligned} \tag{2.2}$$

Denote by $W(\omega, t)$ the loss when ω is the state of nature and t is the decision. We shall assume (Assumption 1 below) that $W(\omega, t) \to +\infty$ as $t \to -\infty$ or $+\infty$. Consequently, any decision function δ which gives positive measure to $+\infty$ or $-\infty$ for some set of x which has positive μ-measure will have infinite risk at every point ω. Such a δ can safely be ignored. In any case, if we denote by $\mathcal{B}$ the class of all decision functions which are Bayes solutions and by $\mathcal{B}'$ the closure of $\mathcal{B}$ in the topology of regular convergence then ([3]) $\mathcal{B}'$ is complete. Let $\mathcal{B}_0$ be that subset of $\mathcal{B}'$ which *excludes* all δ's in $\mathcal{B}'$ which give positive measure to $+\infty$ or $-\infty$ on a set of x of positive μ-measure. The remarks at the beginning of this paragraph imply that $\mathcal{B}_0$ is complete. Our goal is to obtain a representation for the members of $\mathcal{B}_0$. We make the following assumptions.

ASSUMPTION 1. W is non-negative, finite, and continuous in both variables. $W(\omega, t) \to +\infty$ as $t \to +\infty$ or $-\infty$ uniformly on compact sets of ω.

ASSUMPTION 2. If $\omega > t > s$ then $W(\omega, t) < W(\omega, s)$ and

$$\inf_{\omega \geqq t}[W(\omega, s) - W(\omega, t)] > 0.$$

If $\omega < t < s$ then $W(\omega, t) < W(\omega, s)$ and $\inf_{\omega \leqq t}[W(\omega, s) - W(\omega, t)] > 0$.

ASSUMPTION 3. For each t and each $\epsilon > 0$,

$$\sup_{\omega \leqq 0} W(\omega, t)e^{\epsilon\omega} < \infty, \qquad \sup_{\omega \geqq 0} W(\omega, t)e^{-\epsilon\omega} < \infty,$$

and $\lim_{A\to\infty}\sup_{\omega \geqq A} W(\omega, t)e^{-\epsilon\omega} = \lim_{A\to\infty}\sup_{\omega \leqq -A} W(\omega, t)e^{\epsilon\omega} = 0$.

Examples of W's which satisfy these three assumptions are easy to give. In particular, $W(\omega, t) = |\omega - t|^\alpha$, $\alpha > 0$ is one such family of W's. Recall that we are assuming that Ω is the entire real line. Parts of each of the three assumptions are not needed when, for example, $\Omega = [0, \infty)$. The necessary modifications in any of the cases where Ω is not the entire real line are easy to make.

Let $\{\delta_n, n \geqq 1\} \subset \mathcal{B}$ and let $\{\xi_n\}$ be the sequence of a priori distributions such that δ_n is Bayes with respect to ξ_n. Let $\delta \varepsilon \mathcal{B}_0$ and suppose that δ is the regular limit of δ_n. In what follows K_1, K_2, etc., will denote "constants" appropriately chosen to suit the context in which they appear. The dependence of these "constants" on certain auxiliary quantities will be displayed unless there is no issue. We begin by collecting some notation and facts before proceeding with the main argument.

Let

$$\alpha(x, n) = \int_{-\infty}^{0} p(x, \omega)\xi_n(d\omega);$$

(2.3)

$$\beta(x, n) = \int_{0}^{\infty} p(x, \omega)\xi_n(d\omega); \qquad \pi(x, n) = \alpha(x, n)/\beta(x, n).$$

Some modification in the limits of integration is necessary if ξ_n has positive measure at 0 but we will not worry about it. When $\beta(x, n) = 0$, i.e., when $\xi_n(0, \infty) = 0$ ($\xi_n(a, b)$ is the measure ξ_n assigns to the interval (a, b)), we take $\pi(x, n) = +\infty$. The arguments all become simpler when all the ξ_n's are concentrated on either $[0, \infty)$ or $(-\infty, 0]$ but we shall proceed as if the support of ξ_n lies on both sides of 0 (the support of a measure is the smallest closed set whose complement has measure 0). Observe that, for fixed n,

(2.4) $\alpha(x, n)$, $1/\beta(x, n)$, and $\pi(x, n)$ decrease in x.

Let

(2.5) $$X_0 = \{x \mid \liminf_{n\to\infty} \pi(x, n) \leqq 1\}.$$

Whenever X_0 is not empty it is an infinite (to the right) interval by virtue of (2.4). Let $x^* = \inf \{x \mid x \,\varepsilon\, X_0\}$. When X_0 is empty put $x^* = +\infty$. It may or may not be that $x^* \,\varepsilon\, X_0$ and this ambiguity will require some trivial circumlocutions.

When x is observed, t is the decision made, and ξ_n is the a priori distribution, we denote the a posteriori expected loss by

(2.6) $$B(t, x, n) = \int_{-\infty}^{\infty} W(\omega, t)p(x, \omega)\xi_n(d\omega) \Big/ [\alpha(x, n) + \beta(x, n)].$$

We will only be concerned with the situation when, for each x, n, t, $\alpha(x, n)$, $\beta(x, n)$, and $B(t, x, n)$ are all finite. The more general situation can be handled with simple modifications. Since δ_n is Bayes with respect to ξ_n we know that for each x, $\delta_n(\cdot \mid x)$ assigns measure 1 to $\{t \mid B(t, x, n)$ is minimized$\}$ (Assumption 1 guarantees that the minimization actually takes place).

Assumption 2 implies that, if $\gamma > 0$ and $t > \gamma$, then

(2.7)
$$\inf_{c\leqq\gamma} \inf_{\omega\geqq t} [W(\omega, c) - W(\omega, t)] = K_1(\gamma, t) > 0,$$
$$\inf_{c\geqq-\gamma} \inf_{\omega\leqq-t} [W(\omega, c) - W(\omega, -t)] = K_2(\gamma, t) > 0.$$

Thus, if $t > \gamma > 0$, we have ((2.6) and (2.7)), for all $c \leqq \gamma$,

$$
\begin{aligned}
[\alpha(x, n) + \beta(x, n)]B(c, x, n) &\geqq \int_t^\infty W(\omega, c)p(x, \omega)\xi_n(d\omega) \\
&= \int_t^\infty W(\omega, t)p(x, \omega)\xi_n(d\omega) \\
&\quad + \int_t^\infty [W(\omega, c) - W(\omega, t)]p(x, \omega)\xi_n(d\omega) \\
&\geqq \int_t^\infty W(\omega, t)p(x, \omega)\xi_n(d\omega) \\
&\quad + K_1(\gamma, t)\int_t^\infty p(x, \omega)\xi_n(d\omega)
\end{aligned} \tag{2.8}
$$

and, for all $c \geqq -\gamma$,

$$
\begin{aligned}
[\alpha(x, n) + \beta(x, n)]B(c, x, n) &\geqq \int_{-\infty}^{-t} W(\omega, -t)p(x, \omega)\xi_n(d\omega) \\
&\quad + K_2(\gamma, t)\int_{-\infty}^{-t} p(x, \omega)\xi_n(d\omega).
\end{aligned} \tag{2.9}
$$

Since $p(x, \omega)$ is continuous and positive for all x and ω we have

$$
\begin{aligned}
\int_t^\infty p(x, \omega)\xi_n(d\omega) &= \beta(x, n) - \int_0^t p(x, \omega)\xi_n(d\omega) \\
&\geqq \beta(x, n) - K_3(x, t)\xi_n(0, t) \\
\int_{-\infty}^{-t} p(x, \omega)\xi_n(d\omega) &\geqq \alpha(x, n) - K_4(x, t)\xi_n(-t, 0).
\end{aligned} \tag{2.10}
$$

Assumption 1 implies $\int_{-t}^{t} W(\omega, -t)p(x, \omega)\xi_n(d\omega) \leqq K_5(x, t)\xi_n(-t, t)$ and $\int_{-t}^{t} W(\omega, -t)p(x, \omega)\xi_n(d\omega) \leqq K_6(x, t)\xi_n(-t, t)$ so that

$$
\begin{aligned}
[\alpha(x, n) + \beta(x, n)]B(t, x, n) &\leqq \int_t^\infty W(\omega, t)p(x, \omega)\xi_n(d\omega) \\
&\quad + K_5(x, t)\xi_n(-t, t) + \int_{-\infty}^{-t} W(\omega, t)p(x, \omega)\xi_n(d\omega)
\end{aligned} \tag{2.11}
$$

and

$$
\begin{aligned}
[\alpha(x, n) + \beta(x, n)]B(-t, x, n) &\leqq \int_{-\infty}^{-t} W(\omega, -t)p(x, \omega)\xi_n(d\omega) \\
&\quad + K_6(x, t)\xi_n(-t, t) + \int_t^\infty W(\omega, -t)p(x, \omega)\xi_n(d\omega).
\end{aligned} \tag{2.12}
$$

If $\epsilon > 0$ and $y + \epsilon < x$ we have, by Assumption 3,

$$
\begin{aligned}
&\int_{-\infty}^{-t} W(\omega, t)p(x, \omega)\xi_n(d\omega) \\
&\quad = \int_{-\infty}^{-t} W(\omega, t)e^{(x-y)\omega}p(y, \omega)\xi_n(d\omega) \leqq K_7(\epsilon, t)\alpha(y, n)
\end{aligned} \tag{2.13}
$$

and, similarly, if $y > x + \epsilon$

$$\int_t^\infty W(\omega, -t)p(x, \omega)\xi_n(d\omega) \leqq K_8(\epsilon, t)\beta(y, n) \tag{2.14}$$

where $K_7(\epsilon, t) \to 0$ as $t \to +\infty$ and similarly for K_8. Returning to (2.7) note that $K_1(\gamma, t)$ and $K_2(\gamma, t)$ are both increasing with t, so that, for fixed $\epsilon > 0$, we can find t_0 (which will depend on γ and ϵ) such that, for all $t \geqq \max(t_0, \gamma)$,

$$\begin{aligned} &K_1(\gamma, t)/3 - 2K_7(\epsilon, t) \geqq K_9(\text{say}) > 0; \\ &\qquad K_2(\gamma, t)/3 - 2K_8(\epsilon, t) \geqq K_{10} > 0. \end{aligned} \tag{2.15}$$

One further deduction from the nature of $p(x, \omega)$ is that, for any $u > 0$,

$$\alpha(x, n) + \beta(x, n) \geqq \int_{-u}^{u} p(x, \omega)\xi_n(d\omega) \geqq K_{11}(x, u)\xi_n(-u, u). \tag{2.16}$$

Let $\{m_j\}$ be a sequence of integers (all different) such that $\pi(x^* + \epsilon, m_j) \leqq 2$ for all j where ϵ is some fixed positive number. That we can find such a sequence $\{m_j\}$ follows from (2.5) at least if X_0 is not empty; if X_0 is empty then the remarks in this paragraph are vacuous. γ is fixed and ϵ is fixed so we can find t_0 as described above (2.15) and let us take $t > \max(t_0, \gamma)$. The choice of $\{m_j\}$, (2.3), and (2.10) used in (2.8) yields, for all $x \geqq x^* + \epsilon$,

$$\begin{aligned} \inf_{c \leqq \gamma} B(c, x, m_j) \geqq \int_t^\infty \frac{W(\omega, t)p(x, \omega)\xi_{m_j}(d\omega)}{\alpha(x, m_j) + \beta(x, m_j)} \\ + \frac{K_1(\gamma, t)}{3} - \frac{K_1(\gamma, t)K_3(x, t)\xi_{m_j}(0, t)}{\alpha(x, m_j) + \beta(x, m_j)}. \end{aligned} \tag{2.17}$$

Now take $y = x^* + \epsilon$ in (2.13) and conclude that for all $x \geqq x^* + 2\epsilon$

$$\begin{aligned} \int_{-\infty}^{-t} W(\omega, t)p(x, \omega)\xi_{m_j}(d\omega) &\leqq K_7(\epsilon, t)\alpha(x^* + \epsilon, m_j) \\ &= K_7(\epsilon, t)\pi(x^* + \epsilon, m_j)\beta(x^* + \epsilon, m_j) \\ &\leqq 2K_7(\epsilon, t)\beta(x^* + \epsilon, m_j) \\ &\leqq 2K_7(\epsilon, t)\beta(x, m_j). \end{aligned} \tag{2.18}$$

Using (2.18) in (2.11) we have for all $x \geqq x^* + 2\epsilon$

$$B(t, x, m_j) \leqq \frac{\int_t^\infty W(\omega, t)p(x, \omega)\xi_{m_j}(d\omega) + K_5(x, t)\xi_{m_j}(-t, t)}{\alpha(x, m_j) + \beta(x, m_j)} + 2K_7(\epsilon, t). \tag{2.19}$$

Using (2.15) we obtain from (2.19) and (2.17) (lumping some constants together and observing that $\xi_{m_j}(0, t) \leqq \xi_{m_j}(-t, t)$)

$$(2.20)\quad \inf_{c \leqq \gamma} B(c, x, m_j) \geqq B(t, x, m_j) + K_9 - \frac{K_{12}(x, t, \gamma)\xi_{m_j}(-t, t)}{\alpha(x, m_j) + \beta(x, m_j)}$$

for all $x \geqq x^* + 2\epsilon$.

Since $x^* - \epsilon$ is not in X_0 we have $\liminf_{n\to\infty} \pi(x^* - \epsilon, n) > 1$ and, therefore, for all but a finite number of n, $\pi(x^* - \epsilon, n) \geqq \frac{1}{2}$. An argument like that in the preceding paragraph enables us to conclude that, for all but a finite number of n,

$$(2.21)\quad \inf_{c \geqq -\gamma} B(c, x, n) \geqq B(-t, x, n) + K_{10} - \frac{K_{13}(x, t, \gamma)\xi_n(-t, t)}{\alpha(x, n) + \beta(x, n)}$$

for all $x \leqq x^* - 2\epsilon$.

Lemma 1. *There exists a subsequence* $\{n_k\}$ *and a number* a *such that, for all* $b < \infty$, $\limsup_{k\to\infty} \xi_{n_k}(-b, b)/\xi_{n_k}(-a, a) < \infty$.

Proof. We will argue by showing that the negation of the lemma implies the existence of a set E of positive μ-measure such that $\delta(C \mid x) = 0$ on all bounded sets C and all $x \varepsilon E$. Thus δ could not be in $\mathcal{B}_0$ and this is contrary to our assumption.

Let X_0 be as defined in (2.5) and suppose (x^*, ∞) has positive μ-measure. If this is not the case then $\mu(-\infty, x^*)$ has positive μ-measure and a parallel argument will work. If (x^*, ∞) has positive μ-measure then, for some $\epsilon > 0$, $(x^* + 2\epsilon, \infty)$ has positive μ-measure. Let $\{m_j\}$ be the sequence described above (2.17). Consider the interval $[-\gamma, \gamma]$. Let $t > \max(t_0, \gamma)$ and suppose that the lemma is false. Then there is a number $b > t$ and a subsequence $\{n_k\}$ of $\{m_j\}$ such that $\lim_{k\to\infty} \xi_{n_k}(-b, b)/\xi_{n_k}(-t, t) = +\infty$. To avoid circumlocutions define $\xi_n(-u, u)/\xi_n(-v, v) = +\infty$ if $u > v$ and $\xi_n(-v, v) = 0$ and in the same circumstances define $\xi_n(-v, v)/\xi_n(-u, u) = 0$. From (2.16) and (2.20) we obtain

$$(2.22)\quad \inf_{c \leqq \gamma} B(c, x, n_k) \geqq B(t, x, n_k) + K_9 - K_{12}(x, t, \gamma)K_{11}(x, b)\,\frac{\xi_{n_k}(-t, t)}{\xi_{n_k}(-b, b)}$$

for all $x \geqq x^* + 2\epsilon$. Thus, by (2.22) for each $x \geqq x^* + 2\epsilon$ there is a k_0 sufficiently large so that for all $k \geqq k_0$

$$(2.23)\quad \inf_{c \leqq \gamma} B(c, x, n_k) > B(t, x, n_k).$$

Let $X_{r,\gamma} = \{x \mid x \geqq x^* + 2\epsilon, \text{(2.23) holds for all } k \geqq r\}$. Let $I^* = (x^* + 2\epsilon, \infty)$. Note that $\lim_{r\to\infty} X_{r,\gamma} = I^*$. Choose r_γ such that

$$(2.24)\quad \mu(X_{r_\gamma,\gamma}) \geqq \mu(I^*)(1 - 1/2\gamma).$$

Since $\delta_{n_k}([-\gamma, \gamma] \mid x) = 0$ for all $k \geqq r_\gamma$ and all $x \varepsilon X_{r_\gamma,\gamma}$ we have by (2.2)

$$\delta([-\gamma, \gamma] \mid x) = 0$$

a.e. on $X_{r_\gamma,\gamma}$. We will let γ assume only positive integers as values. For each such γ we can find a subsequence $\{n_k\}$ such that (2.22) holds for all $x \geqq x^* + 2\epsilon$

and, consequently, we can define $X_{r_\gamma,\gamma}$ such that (2.24) holds and such that $\delta([-\gamma, \gamma] \mid x) = 0$ a.e. on $X_{r_\gamma,\gamma}$. $X_{r,\gamma} \subset I^*$ for each r and γ and from (2.24) we have

$$\mu\left(\bigcap_{\gamma=1}^{\infty} X_{r_\gamma,\gamma}\right) \geqq \tfrac{1}{2}\mu(I^*).$$

Take $E = \bigcap_{\gamma=1}^{\infty} X_{r_\gamma,\gamma}$ and we conclude that, for every bounded set C and every $x \varepsilon E$, $\delta(C \mid x) = 0$. But this implies that $\delta \not\varepsilon \mathcal{B}_0$ which is a contradiction.

If (x^*, ∞) has μ-measure 0 then $(-\infty, x^*)$ has positive μ-measure and we can use (2.21) instead of (2.20) to yield the same conclusion. This finishes the proof of Lemma 1.

Let $\{n_k\}$ and a be as in Lemma 1. Define measures $F_{n_k} = \xi_{n_k}/\xi_{n_k}(-a, a)$. Because of Lemma 1 a subsequence $\{r_i\}$ can be extracted so that F_{r_i} converges (weakly) to a measure F on the real line with the property that F is finite on bounded intervals and is not everywhere 0 (in particular, $F(-a, a) = 1$). This being so we might as well assume that $\{\xi_{r_i}\}$ is the original sequence $\{\xi_n\}$ so that, in addition to $\delta_n \to \delta$, we now have $F_n = \xi_n/\xi_n(-a, a) \to F$. F, of course, may or may not assign finite measure to the whole line.

LEMMA 2. *For almost all* $x(\mu)$, $\limsup_{n\to\infty} \int_{-\infty}^{+\infty} p(x, \omega)F_n(d\omega) < \infty$.

PROOF. Let

$$X_1 = \{x \mid \limsup_{n\to\infty} \beta(x, n)/\xi_n(-a, a) = +\infty\}$$

$$X_2 = \{x \mid \limsup_{n\to\infty} \alpha(x, n)/\xi_n(-a, a) = +\infty\}.$$

We wish to show that both $\mu(X_1)$ and $\mu(X_2)$ are 0. Considering X_1 first let $x_1 = \inf\{x \mid x \varepsilon X_1\}$. Since X_1 is an infinite (to the right) interval ($\beta(x, n)$ increases in x) we have two situations to be concerned with, namely, either $\mu(x_1, \infty) > 0$ or $\mu(\{x_1\}) > 0$ and $\mu(x_1, \infty) = 0$.

CASE 1. $\mu(x_1, \infty) > 0$: In this case there is a positive number ϵ_1 such that $\mu(x_1 + 2\epsilon, \infty) > 0$. Let $\{n_k\}$ be a sequence of integers such that

$$\lim_{k\to\infty} \beta(x_1 + \epsilon, n_k)/\xi_{n_k}(-a, a) = +\infty. \tag{2.25}$$

Let $X_0^1 = \{x \mid \liminf_{k\to\infty} \pi(x, n_k) \leqq 1\}$ and let $x_1^* = \inf\{x \mid x \varepsilon X_0^1\}$. If $x_1^* \leqq x_1 + 2\epsilon_1$ then we can proceed as in Lemma 1 as follows: let $\{m_j\}$ be a subsequence of $\{n_k\}$ such that $\pi(x_1 + 2\epsilon_1, m_j) \leqq 2$ for all j and observe by use of (2.20) that, for $\gamma > 0$ and $t > \max(t_0, \gamma)$ (see the proof of Lemma 1),

$$\begin{aligned}\inf_{c \leqq \gamma} B(c, x, m_j) &\geqq B(t, x, m_j) + K_9 - K_{12}(x, t, \gamma)\frac{\xi_{m_j}(-t, t)}{\alpha(x, m_j) + \beta(x, m_j)} \\ &\geqq B(t, x, m_j) + K_9 - K_{12}(x, t, \gamma)\frac{\xi_{m_j}(-t, t)}{\beta(x, m_j)}\end{aligned}$$

for all $x > x_1 + 2\epsilon_1$. Since $\{m_j\}$ is a subsequence of $\{n_k\}$ we obtain from (2.25) and the fact that $F_{m_j}(-t, t) \to F(-t, t) < \infty$

$$\frac{\xi_{m_j}(-t, t)}{\beta(x, m_j)} \leqq \frac{\xi_{m_j}(-t, t)}{\beta(x_1 + \epsilon_1, m_j)} = \frac{\xi_{m_j}(-t, t)}{\xi_{m_j}(-a, a)} \frac{\xi_{m_j}(-a, a)}{\beta(x_1 + \epsilon_1, m_j)} \to 0$$

as $j \to \infty$. We can now argue as in Lemma 1 to obtain a subset E of

$$(x_1 + 2\epsilon_1, \infty)$$

with $\mu(E) > 0$ and $\delta(C \mid x) = 0$ for all $x \varepsilon E$ and all bounded sets C.

If $x_1^* > x_1 + 2\epsilon_1$ and $\mu(x_1^*, \infty) > 0$ then we can argue as above to obtain a subset E of (x_1^*, ∞) etc. If $x_1^* > x_1 + 2\epsilon_1$ and $\mu(x^*, \infty) = 0$ then there is a number $\eta_1 > 0$ such that $x_1 + 2\epsilon_1 < x_1^* - \eta_1$ and $\mu(x_1 + 2\epsilon_1, x_1^* - \eta_1) > 0$. Since $x_1^* - \eta_1 \varepsilon X_1$ and since $\liminf_{k\to\infty} \pi(x_1^* - \eta_1, n_k) > 1$ we have $x_1^* - \eta_1 \varepsilon X_2$ and, in fact $\lim_{k\to\infty} \alpha(x, n_k)/\xi_{n_k}(-a, a) = +\infty$ for all $x \leqq x_1^* - \eta_1$. Thus $\mu(X_2) > 0$ and we can now argue as in the first part of this Case 1 using (2.21) instead of (2.20).

Case 2. $\mu(\{x_1\}) > 0$ and $\mu(x_1, \infty) = 0$: Here we can find a sequence of integers $\{n_k\}$ with $\lim_{k\to\infty} \beta(x_1, n_k)/\xi_{n_k}(-a, a) = +\infty$. If x_1^* (as defined in Case 1) is smaller than x_1 we can proceed as above to show that $\delta(C \mid x_1) = 0$ for all bounded sets C. If $x_1^* = x_1$ and $x_1^* \varepsilon X_0^1$ the same argument holds. If $x_1^* = x_1$ and $x_1^* \not\varepsilon X_0^1$ or if $x_1^* > x_1$ then $x_1 \varepsilon X_2$ and again we refer to the previous argument.

We have now shown that $\mu(X_1) = 0$. A parallel argument shows that $\mu(X_2) = 0$, and this concludes the proof of Lemma 2.

Let $z_0 = \inf\{x \mid x \varepsilon \text{ support of } \mu\}$ and let $z_1 = \sup\{x \mid x \varepsilon \text{ support of } \mu\}$. z_0 may be $-\infty$ and z_1 may be $+\infty$.

Lemma 3. *For almost all $x(\mu) \varepsilon (z_0, z_1)$*

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} p(x, \omega)F_n(d\omega) = \int_{-\infty}^{+\infty} p(x, \omega)F(d\omega) < \infty \tag{2.26}$$

and

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} W(\omega, t)p(x, \omega)F_n(d\omega) = \int_{-\infty}^{+\infty} W(\omega, t)p(x, \omega)F(d\omega) \tag{2.27}$$

uniformly on compact sets of t. The right hand side of (2.27) is finite and continuous in t.

Proof. Let us first establish (2.26). Since $p(x, \omega)$ is continuous in ω we know that for any $0 < A < \infty$, $\lim_{n\to\infty} \int_{-A}^{A} p(x, \omega)F_n(d\omega) = \int_{-A}^{A} p(x, \omega)F(d\omega)$. Thus we have only to establish uniform integrability, i.e.,

$$\lim_{A\to\infty} \limsup_{n\to\infty} \int_{A}^{\infty} p(x, \omega)F_n(d\omega) = 0 \tag{2.28}$$

and

$$\lim_{A\to\infty} \limsup_{n\to\infty} \int_{-\infty}^{-A} p(x, \omega)F_n(d\omega) = 0. \tag{2.29}$$

Let X' be the set of $x \varepsilon (z_0, z_1)$ where (2.28) fails. We wish to prove that $\mu(X') = 0$. If $\mu(X') > 0$ then there is an x' such that $\mu(x', \infty) > 0$ and

$$\lim_{A\to\infty} \limsup_{n\to\infty} \int_{A}^{\infty} p(x', \omega)F_n(d\omega) > 0.$$

But, for $x > x'$,

$$\int_A^\infty p(x, \omega)F_n(d\omega) \geqq e^{(x-x')A} \int_A^\infty p(x', \omega)F_n(d\omega)$$

and, consequently,

$$\limsup_{n\to\infty} \int_{-\infty}^{+\infty} p(x, \omega)F_n(d\omega) \geqq \lim_{A\to\infty} \limsup_{n\to\infty} \int_A^\infty p(x, \omega)F_n(d\omega) = +\infty$$

for all $x > x'$. This, however, contradicts Lemma 2 and thereby establishes (2.28) for almost all $x < z_1$. Similarly (2.29) follows for almost all $x > z_0$ and, therefore, (2.26) is proved.

We next show that (2.27) holds for any t (the exceptional set of x's will not depend on t). Because W is continuous in ω it is again sufficient to prove uniform integrability. If x and t are such that

$$\lim_{A\to\infty} \limsup_{n\to\infty} \int_A^\infty W(\omega, t)p(x, \omega)F_n(d\omega) > 0$$

then observe that, for $y > x$,

$$\int_A^\infty W(\omega, t)p(x, \omega)F_n(d\omega) \leqq \sup_{\omega \geqq A} [W(\omega, t)e^{(x-y)\omega}] \int_A^\infty p(y, \omega)F_n(d\omega)$$

so that, by Assumption 3, $\limsup_{n\to\infty} \int_A^\infty p(y, \omega)F_n(d\omega) = +\infty$. Now use Lemma 2 or the first part of the argument of this Lemma to obtain the necessary contradiction. The finiteness of the right hand side of (2.27) is obvious.

To prove continuity in t of the limit, observe first that, for fixed $\epsilon > 0$,

$$\lim_{t\to t_0} \int_{t_0-\epsilon}^{t_0+\epsilon} W(\omega, t)p(x, \omega)F(d\omega) = \int_{t_0-\epsilon}^{t_0+\epsilon} W(\omega, t_0)p(x, \omega)F(d\omega)$$

by bounded convergence. Since $W(\omega, t) \leqq W(\omega, t_0 - \epsilon)$ for $t_0 + \epsilon > t > t_0 - \epsilon \geqq \omega$ (Assumption 2) we have by Lebesgue's dominated convergence theorem

$$\lim_{t\to t_0} \int_{-\infty}^{t_0-\epsilon} W(\omega, t)p(x, \omega)F(d\omega) = \int_{-\infty}^{t_0-\epsilon} W(\omega, t_0)p(x, \omega)F(d\omega)$$

with the same result for the integral over $[t_0 + \epsilon, \infty]$. Thus continuity of the limit is established.

Turning now to the question of uniformity in (2.27) let $C = [-\gamma, \gamma]$ with $\gamma > 0$. If the limit in (2.27) is not uniform for $t \varepsilon C$ then, for $\epsilon > 0$, there is a sequence $\{t_n\} \subset C$ with $t_n \to t_0$, $t_0 \varepsilon C$ such that

$$\left|\int_{-\infty}^{+\infty} W(\omega, t_n)p(x, \omega)F_n(d\omega) - \int_{-\infty}^{+\infty} W(\omega, t_n)p(x, \omega)F(d\omega)\right| > \epsilon. \tag{2.30}$$

Continuity in t shows that for n large enough

$$\left|\int_{-\infty}^{+\infty} W(\omega, t_n)p(x, \omega)F(d\omega) - \int_{-\infty}^{+\infty} W(\omega, t_0)p(x, \omega)F(d\omega)\right| < \epsilon/8. \tag{2.31}$$

For any A we have, from the continuity of W,

$$\left|\int_{t_0-A}^{t_0+A} [W(\omega, t_n) - W(\omega, t_0)]p(x, \omega)F_n(d\omega)\right| \leqq \int_{t_0-A}^{t_0+A} |W(\omega, t_n) - W(\omega, t_0)|p(x, \omega)F_n(d\omega) \leqq \epsilon/8 \tag{2.32}$$

for n large enough. Since $W(\omega, t_0 - \eta) \leqq W(\omega, t_n) \leqq W(\omega, t_0 + \eta)$ for $\omega \leqq t_0 - \eta \leqq t_n \leqq t_0 + \eta$ we have

$$\begin{aligned} \int_{-\infty}^{t_0-A} [W(\omega, t_0 - \eta) - W(\omega, t_0)]p(x, \omega)F_n(d\omega) &\leqq \int_{-\infty}^{t_0-A} [W(\omega, t_n) - W(\omega, t_0)]p(x, \omega)F_n(d\omega) \\ &\leqq \int_{-\infty}^{t_0-A} [W(\omega, t_0 + \eta) - W(\omega, t_0)]p(x, \omega)F_n(d\omega). \end{aligned}$$

Now, using (2.27) for fixed t (it is easy enough to check that (2.27) holds even when the range of integration is cut down to $(-\infty, t_0 - A])$ and the continuity in t of the right hand side of (2.27) (even when the range of integration is reduced) we obtain, for n large enough,

$$\left|\int_{-\infty}^{t_0-A} [W(\omega, t_n) - W(\omega, t_0)]p(x, \omega)F_n(d\omega)\right| < \epsilon/8. \tag{2.33}$$

Similarly, for n large enough,

$$\left|\int_{t_0+A}^{\infty} [W(\omega, t_n) - W(\omega, t_0)]p(x, \omega)F_n(d\omega)\right| < \epsilon/8. \tag{2.34}$$

Using (2.27) with $t = t_0$ and the range of integration $|\omega - t_0| \geqq A$ we obtain from (2.31), (2.32), (2.33), and (2.34) that (2.30) cannot hold for any n sufficiently large. This establishes the uniformity desired and completes the proof of Lemma 3.

THEOREM. *If Assumptions* 1, 2, 3 *are satisfied and* $\{\delta_n\}$ *is a sequence of Bayes solutions such that* $\delta_n \to \delta$ *regularly with* $\delta \varepsilon \mathcal{B}_0$ *then there exists a measure* F *on* $(-\infty, +\infty)$ *such that for almost all* $x(\mu)$ *in* (z_0, z_1) (*see above Lemma* 3 *for the definition of* z_0 *and* z_1) $\delta(B_x \mid x) = 1$ *where*

$$B_x = \left\{t \,\middle|\, \int_{-\infty}^{+\infty} W(\omega, t)p(x, \omega)F(d\omega) \text{ is minimized}\right\}.$$

PROOF. As noted following Lemma 1 we can assume that there are measures F_n and a measure F with $F_n \to F$. Lemma 3 is valid for $\{F_n\}$ and F and we take the exceptional set of x in the Theorem to be the exceptional set of Lemma 3. Henceforth all statements will be for $x \varepsilon (z_0, z_1)$ and not in the exceptional set. Put

$$H(t, x) = \int_{-\infty}^{+\infty} W(\omega, t)p(x, \omega)F(d\omega), \quad H(t, x, n) = \int_{-\infty}^{+\infty} W(\omega, t)p(x, \omega)F_n(d\omega),$$

$$m(x) = \inf_t H(t, x), \qquad m(x, n) = \inf_t H(t, x, n),$$

$$B_{x,r} = \{t \mid H(t, x) \leqq m(x) + r\}, \qquad B^n_{x,r} = \{t \mid H(t, x, n) \leqq m(x, n) + r\}$$

for $r \geqq 0$. Note that $B_{x,0} = B_x$ and that $B^n_{x,0}$ is the same as the set of t where $B(t, x, n)$ is minimized (see (2.6)). From Assumption 1 and Fatou's Lemma $H(t, x)(H(t, x, n)) \to +\infty$ as $t \to +\infty$ or $-\infty$. From Lemma 3 $H(t, x) < \infty$ for each t and the same is true for $H(t, x, n)$ for each n. It follows that $B_{x,r}$ is a bounded set and so is $B^n_{x,r}$. $B_{x,r}(B^n_{x,r})$ is a closed set because $H(t, x)(H(t, x, n))$ is continuous in t. Since δ_n is Bayes, we have, for almost all x, all $r \geqq 0$, and all n, $\delta_n(B^n_{x,r} \mid x) = 1$. If the theorem is false we can find positive numbers r_0, γ_0, ϵ_0, and a set E_0 of x, $E_0 \subset (z_0, z_1)$, and $\mu(E_0) > 0$ such that

$$\delta(B_{x,r} \mid x) \leqq 1 - \epsilon_0, \; B_{x,r} \subset [-\gamma_0, \gamma_0] \tag{2.35}$$

for all $x \varepsilon E_0$ and all $r \leqq r_0$.

The uniformity (for t in a compact set) of the convergence of $H(t, x, n)$ to $H(t, x)$ (Lemma 3) implies $m(x, n) \to m(x)$. In addition, the uniformity of convergence implies that there is a number γ_1 and a set E_1 with $E_1 \subset E_0$ and $\mu(E_1) > 0$ such that $B^n_{x,r} \subset [-\gamma_1, \gamma_1]$ all n, all $x \varepsilon E_1$, all $r \leqq r_0$. Again from the uniformity of the convergence in (2.27) $B^n_{x,r} \to B_{x,r}$ as $n \to \infty$ for each $x \varepsilon E_1$ and $0 < r \leqq r_0$. In fact (let $r = r_0/4$ for specificity) $B^n_{x,r/2} \subset B_{x,r}$ for $n \geqq n_0(x, r)$. If we let

$$E_{kr} = \{x \mid B^n_{x,r/2} \subset B_{x,r} \text{ for all } n \geqq k, \, x \varepsilon E_1\}$$

then, for some k sufficiently large, $\mu(E_{kr}) > 0$ and, for $x \varepsilon E_{kr}$,

$$1 = \delta_n(B^n_{x,r/2} \mid x) \leqq \delta_n(B_{x,r} \mid x). \tag{2.36}$$

Suppose there is an $x^* \varepsilon E_{kr}$ with $\mu(\{x^*\}) > 0$. By (2.36) $\lim_{n\to\infty} \delta_n(B_{x^*,r} \mid x^*) = 1$ and by (2.2) we must have $\delta(B_{x^*,r} \mid x^*) = 1$ but this contradicts (2.35).

If there are no atoms in E_{kr} then we can find an $x^* \varepsilon E_{kr}$ such that

$$[x^* - \eta, x^* + \eta] \cap E_{kr}$$

has positive μ-measure for all $\eta > 0$. Observe that $H(t, x)$ is continuous in x uniformly for t in a compact set—the argument is similar to that of the last part of Lemma 3 and, in fact, somewhat simpler—and consequently, $m(x)$ is continuous in x. Thus, for some positive η and all

$$y \,\varepsilon\, E^* = [x^* - \eta, x^* + \eta] \cap E_{kr}, \; B_{y,r} \subset B_{x^*,2r} \subset B_{y,4r}.$$

Note that $\mu(E^*) > 0$. Since $4r = r_0$ we have, by (2.35),

$$\delta(B_{x^*,2r} \mid y) \leqq \delta(B_{y,4r} \mid y) \leqq 1 - \epsilon_0 \tag{2.37}$$

for all $y \varepsilon E^*$. Since $E^* \subset E_{kr}$ we have for all $n \geqq k$

$$1 = \delta_n(B^n_{y,r/2} \mid y) \leqq \delta_n(B_{y,r} \mid y) \leqq \delta_n(B_{x^*,2r} \mid y) \tag{2.38}$$

for all $y \varepsilon E^*$. Use of (2.2) shows that (2.37) and (2.38) are incompatible with the regular convergence of δ_n to δ and the theorem is proved.

COROLLARY 1. *Let $\mathfrak{F}$ be the set of all measures F on the Borel sets of the real line which are finite on bounded sets and such that $H(t, x) < \infty$ for all t* a.e. (μ). *For $F \varepsilon \mathfrak{F}$ let $\mathcal{B}_F$ be the set of δ's which satisfy $\delta(B_x \mid x) = 1$* a.e. (μ). *If*

$$\mu(\{z_0\} \cup \{z_1\}) = 0$$

and Assumptions 1, 2, *and* 3 *are satisfied, then $\mathcal{B}_0 \subset \mathcal{B}^* = \bigcup \{\mathcal{B}_F \mid F \varepsilon \mathfrak{F}\}$ and hence $\mathcal{B}^*$ is essentially complete.*

COROLLARY 2. *If, in addition to the hypotheses of Corollary* 1, *W is strictly convex in t for each ω, then $\mathcal{B}_0 = \mathcal{B}^*$.*

PROOF. Corollary 1 is immediate from Theorem 1. To prove Corollary 2 observe that the convexity of W implies that, for any $F \varepsilon \mathfrak{F}$, $\mathcal{B}_F$ consists of a single (non-randomized) decision function—call it δ_F. Since $\delta_F(\pm\infty \mid x) = 0$ a.e. we need only show that $\delta_F \varepsilon \mathcal{B}'$ in order to conclude that $\delta_F \varepsilon \mathcal{B}_0$. For each positive integer n, define F_n to be the restriction of F to $[-n, n]$ (we might as well assume that F_n is non-trivial for all n). Let $\psi_n = F_n/F[-n, n]$. Then ψ_n is a probability measure concentrated on $[-n, n]$ and if δ_n is Bayes with respect to ψ_n then the convexity of W implies that $\delta_n(t_n(x) \mid x) = 1$ where $t_n(x)$ is that t which minimizes

$$H(t, x, n) = \int_{-n}^{n} W(\omega, t)p(x, \omega)F(d\omega)$$

It is easy to prove that $H(t, x, n) \to H(t, x)$ as $n \to \infty$ uniformly on compact sets of t and consequently $t_n(x) \to t(x)$ a.e. (μ), where $t(x)$ is that t which minimizes $H(t, x)$. That $\delta_n \to \delta_F$ regularly is an immediate consequence. Thus $\delta_F \varepsilon \mathcal{B}_0$ and hence $\mathcal{B}^* \subset \mathcal{B}_0$. Corollary 1 tells us that $\mathcal{B}_0 \subset \mathcal{B}^*$ and we are finished. Note that what was needed is the guarantee that $\mathcal{B}_F$ consists of a single decision function and convexity of W assures this.

REMARKS. 1. We do not know whether $\mathcal{B}^* = \mathcal{B}_0$ or not in the situations not covered by Corollary 2. If B_x consists of two points t_1, t_2 then it might be (as far as we know) that $B^n_x \cap [t_2 - \epsilon, t_2 + \epsilon]$ is empty for some $\epsilon > 0$ and all n and, consequently, δ_n could not converge regularly to δ if $\delta(\{t_2\} \mid x) > 0$ and $\mu(\{x\}) > 0$.

2. That we cannot rule out the condition $\mu(\{z_0\} \cup \{z_1\}) = 0$ in Corollaries 1 and 2 can be seen by the following example where we exhibit $\{\delta_n\}$ and δ with δ_n Bayes, $\delta_n \to \delta$ regularly, and $\delta \not\varepsilon \mathcal{B}^*$. Let μ assign uniform measure to $[0, 1)$ and let μ give measure 1 to $\{1\}$. Let F_n assign uniform measure to $[0, 1]$ and let $F_n(\{n\}) = 1/n$. Take $\xi_n = F_n/(1 + 1/n)$. Let δ_n be Bayes with respect to ξ_n so that $\delta_n(t_n(x) \mid x) = 1$ where

$$t_n(x) = \left[\int_0^1 \omega p(x, \omega)\, d\omega + p(x, n)\right] \Big/ \left[\int_0^1 p(x, \omega)\, d\omega + \frac{1}{n} p(x, n)\right].$$

It is easy to check that, as $n \to \infty$, $p(x, n) \to 0$ if $x < 1$ while $p(1, n) \to \frac{1}{2}$. Consequently $t_n(x) \to t(x)$ where

$$(2.39) \qquad t(x) = \left[\int_0^1 \omega p(x, \omega)\, d\omega + \phi(x)\, \frac{1}{2}\right] \Big/ \int_0^1 p(x, \omega)\, d\omega$$

where $\phi(x) = 0$ if $x < 1$, $\phi(x) = 1$ if $x = 1$ (we needn't worry about $x \not\in [0, 1]$). If δ is defined by $\delta(t(x) \mid x) = 1$ then $\delta_n \to \delta$ regularly and hence $\delta \in \mathcal{B}_0$. If $\delta \in \mathcal{B}^*$ then there is an F such that

$$(2.40) \qquad t(x) = \int_{-\infty}^{+\infty} \omega p(x, \omega) F(d\omega) \Big/ \int_{-\infty}^{+\infty} p(x, \omega) F(d\omega)$$

a.e. (μ). Let $L(x) = \int_0^1 p(x, \omega)\, d\omega = \int_0^1 e^{x\omega} \rho(\omega)\, d\omega$ and let

$$L^*(x) = \int_{-\infty}^{+\infty} p(x, \omega) F(d\omega).$$

$$t(x) = \frac{d \log L^*(x)}{dx} = \frac{d \log L(x)}{dx} + \frac{\frac{1}{2}\, \phi(x)}{L(x)}.$$

Hence, for $x < 1$ we have, for some constant $\lambda > 0$, $L^*(x) = \lambda L(x)$. Since L^* and L are Laplace transforms it must be that $F = \lambda \mathrm{U}$ where U is uniform measure on [0, 1]. But (2.39) and (2.40) are in conflict when $x = 1$ and consequently $\delta \not\in \mathcal{B}^*$.

3. The results above can be generalized to other than exponential densities. The key assumption is Assumption 3 which, for nonexponential densities would be modified by replacing $e^{\epsilon\omega}$ by $p(x + \epsilon, \omega)/p(x, \omega)$ (in case there is more than one observation we would consider

$$p(x_1 + \epsilon, x_2 + \epsilon, \cdots, x_n + \epsilon, \omega)/p(x_1, \cdots, x_n, \omega)).$$

For example, if $p(x, \omega) = e^{-(x-\omega)^4}$, $\Omega = [0, \infty)$, $W(\omega, t) = (\omega - t)^2$ then Assumption 3 holds and we could carry out the arguments to obtain the same conclusion. In fact, if Assumption 3 is modified as above, and $p(x, \omega)$ is continuous and positive for all ω and x, and if the support of μ is the entire real line (or all of n-dimensional space when there are n observations) then the arguments can be modified to give the same conclusions. The requirement that μ gives positive mass to every open interval enables us to consider in (2.22), for example, intervals of x's which lie far to the right of x^* so that the lack of monotonicity of α, β, and π is not essential to drawing the necessary contradictions.

Examples where these results fail are not too hard to provide but the only ones we know are when Assumption 3 is violated and p is not exponential. In particular take $\Omega = (-\infty, +\infty)$, $W(\omega, t) = (\omega - t)^2$, $p(x, \omega) = e^{-(x-\omega)}\phi(x, \omega)$ where $\phi(x, \omega) = 1$ if $x > \omega$, $\phi(x, \omega) = 0$ if $x < \omega$, and the sequence

$$\xi_n = F_n/(2n + e^n/n)$$

where F_n is Lebesgue measure on $(-n, n]$, puts mass e^n/n at $-n$, and gives the rest of the line 0 measure. It is simple to verify that δ_n (the Bayes solution with respect to ξ_n) $\to \delta$ regularly when $\delta(t(x) \mid x) = 1$ and

$$t(x) = \left(\int_{-\infty}^{x} \omega e^{\omega} \, d\omega - 1\right) \bigg/ \int_{-\infty}^{x} e^{\omega} \, d\omega = x - 1 - e^{-x}.$$

To show that there is no F such that

$$\int_{-\infty}^{x} \omega e^{\omega} \, F(d\omega) \bigg/ \int_{-\infty}^{x} e^{\omega} \, F(d\omega) = x - 1 - e^{-x} \tag{2.41}$$

integrate by parts to conclude that F is differentiable and then (putting $f = F'$)

$$\frac{d}{dx} \int_{-\infty}^{x} \omega e^{\omega} f(\omega) \, d\omega = \frac{d}{dx} (x - 1 - e^{-x}) \int_{-\infty}^{x} e^{\omega} f(\omega) d\omega$$

implies $xe^{x}f(x) = (x - 1 - e^{-x})e^{x}f(x) + \int_{-\infty}^{x} e^{\omega} f(\omega) \, d\omega \, (1 + e^{-x})$ so that $e^{x}f(x) = \int_{-\infty}^{x} e^{\omega} f(\omega) \, d\omega$ i.e., f is constant and this contradicts (2.41).

4. If p is a two-dimensional exponential density, i.e., if

$$p(x, \omega) = \rho(\omega_1, \omega_2) \exp(x_1 \omega_1 + x_2 \omega_2), \; \omega = (\omega_1, \omega_2), \text{ and } x = (x_1, x_2)$$

then the same results will hold with appropriate alteration of the assumptions. The argument proceeds as before except that $\alpha(x, n) + \beta(x, n)$ has to be broken up into four pieces corresponding to the four quadrants in the plane. Similarly the results can be generalized to n-dimensional exponential densities.

5. One of the questions that led us into these investigations is: what are the admissible procedures for estimating the mean of a normal distribution when the loss is squared error? In particular we became interested in the problem of finding an admissible minimax estimator when the parameter space is restricted to $[0, \infty)$ and, stimulated by Karlin's [1] work on admissibility, we announced (Abstract 75, *Ann. Math. Statist.*, (1960) p. 246) the result that

$$t^*(x) = \int_{0}^{\infty} \omega e^{-\frac{1}{2}(x-\omega)^2} \, d\omega \bigg/ \int_{0}^{\infty} e^{-\frac{1}{2}(x-\omega)^2} \, d\omega$$

is such an estimator. Independently of us Katz [2] obtained the same result and we refer the reader to Katz's work for a proof of this fact.

A consequence of Corollary 2 is that for $W(\omega, t) = (\omega - t)^2$ and p an exponential density, every admissible estimator must be of the form

$$t(x) = (d/dx) \log L_{\psi}(x) \tag{2.42}$$

where L_{ψ} is the Laplace transform of a measure ψ on Ω ($\psi(d\omega) = \rho(\omega)F(d\omega)$). Thus, for example, in the problem of the last paragraph $\hat{t}(x) = \max(0, x)$, which is the maximum likelihood estimator, is inadmissible. A sidelight of this occurs on comparing the risk functions of $\hat{t}$ (which is known to be minimax) and t^* which reveals that t^* is not an improvement of $\hat{t}$ and that, consequently, there are many admissible minimax estimators for this normal estimation problem.

That (2.42) does not answer the question raised at the beginning of this Remark can be seen by taking $F(d\omega) = e^{\alpha\omega}\, d\omega$ for any $\alpha > 0$, and observing that for $\Omega = (-\infty, +\infty)$ this results in an inadmissible t. R. Farrell has pointed out to us that the restriction of this F to $[0, \infty)$ gives an inadmissible estimator when $\Omega = [0, \infty)$.

REFERENCES

[1] KARLIN, SAMUEL (1958). Admissibility for estimation with quadratic loss. *Ann. Math. Statist.* **29** 406–436.

[2] KATZ, MORRIS W. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Ann. Math. Statist.* **32** 136–142.

[3] LECAM, L. (1955). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.