Technical Report No. 74

GENERALIZED BAYES DECISION FUNCTIONS, ADMISSIBILITY AND THE EXPONENTIAL FAMILY

By

M. Stone

University of Wisconsin

This research was partially supported by the United States Navy through the Office of Naval Research under Contract Nonr-1202, Project NR 042 222.
GENERALIZED BAYES DECISION FUNCTIONS, ADMISSIBILITY AND THE EXPONENTIAL FAMILY

By

M. Stone

University of Wisconsin

1. Introduction and summary. For the experiment $\mathcal{E}$, given by $\mathcal{E} = \{ x \text{ distributed with probability element } f(x|\theta) \, d\mu(x), \, \{\theta\} = \emptyset \text{ the parameter space, } W(d, \theta) \text{ a non-negative loss function for decision } d \in D \}$, the definition of a (non-randomized) strict Bayes decision function (BDF) can, following Wald [7], be stated: $\delta^*_{\pi}$, given by the decision function $d(\pi)$, is a strict BDF with respect to the proper prior probability measure $\pi$ on $\emptyset$ if $\delta^*_{\pi}$ minimizes

$$R(\pi, \delta^*) = \int d\pi(\theta) \, r(\theta, \delta^*) = \int d\pi(\theta) \int d\mu(x) \, W[d(x), \theta] \, f(x|\theta)$$

[requiring, of course, that $R(\pi, \delta^*) < \infty$.] (As is well known, randomized decision functions can be excluded from standard Bayesian methods. To maintain conciseness, we will not even attempt to justify their exclusion in the below deviations from the standard Bayesian formulations.)

Relaxing the restriction that the prior measure be proper, we have:

DEFINITION 1.1 $\delta_m$ is a normal generalized BDF (NGBDF) with respect to the generalized measure $\pi_m$ on $\emptyset$ if $\delta_m$ minimizes $R(\pi, \delta)$ [requiring, of course, that $R(\pi, \delta_m) < \infty$].

DEFINITION 1.2. $\delta_m$ is an extensive generalized BDF (EGBDF) with respect to $\pi_m$ if $\delta_m = \{d_m(x)\}$ where $d_m(x)$ minimizes $\int d\mu(\theta) \, W(d(x), \theta) \, f(x|\theta) \, d\pi_m$. (The use of the epithets "extensive" and "normal" is consistent with their use in Raiffa and Schlaifer [5].) Definition 1, 2 differs from the
definition of Sacks [6] in not requiring the finiteness of \( \int dm(\theta) f(x|\theta) \).

Theorem 2.1 shows that if \( \delta_m \) is an NGBDF then it is also an EGBDF. For some cases in which \( \min_\delta R(m,\delta) = \infty \), the following generalization of EGBDF is useful:

**DEFINITION 1.3.** \( \delta_m^* \) is a **comparative generalized BDF (CGBDF)** with respect to \( m \) if the quantity \( \Delta_m(\delta^*,\delta_m^*) \) defined by

\[
\Delta_m(\delta^*,\delta_m^*) = \int dm(\theta) \left[ r(\theta,\delta) - r(\theta,\delta_m^*) \right]
\]

is non-negative for all \( \delta \).

Section 2 states one theorem and four examples concerning these definitions.

Admissibility considerations are unaffected by multiplying \( W(d,\theta) \) by an arbitrary positive function of \( \theta \). So, since the above definitions involve \( m \) and \( W \) in the composite element \( W(d,\theta) dm(\theta) \), it is clear that any general sufficient condition for admissibility of \( \delta_m \) or \( \delta_m^* \) for \( m \) proper may also be stated for \( m \) general (that is, possibly improper). (The same point is made by Stein [4].) Theorem 3.1 gives such a sufficient condition, suggested by the Lehmann-Blyth technique for proving admissibility [1], while Corollary 3.1 is an extension of the well known admissibility of strict BDFs under certain conditions.

The above ideas are applied in Section 4 to the estimation with quadratic loss of the mean of the one dimensional exponential family. The very close links with Karlin's technique [3] are immediately apparent. The
application clarifies Karlin's remark (p. 411 of [3]) that his results may be regarded as a refinement of the Lehmann-Blyth technique and also, finally, lends some support to his conjecture (p. 415 of [3]) that a certain condition for admissibility of the contracted estimator $\gamma x$, $0 < \gamma \leq 1$, is necessary as well as sufficient. An associated reference is Cheng Ping [2].

2. **EGBDF, NGBDF and CGBDF connections.** If $\delta_m$ is an NGBDF with respect to $m$, it is clearly a CGBDF with respect to $m$. In addition:

**THEOREM 2.1.** If $\delta_m$ is an NGBDF with respect to $m$ then it is an EGBDF with respect to $m$.

**PROOF.** Suppose $\delta_m$ is not an NGBDF. Then there exists a $\delta = \{d(x)\}$ such that

$$\int d\mu(x) \int dm(\theta) W[d(x), \theta] f(x | \theta) < \int d\mu(x) \int dm(\theta) W[d_m(x), \theta] f(x | \theta)$$

or, since, by Tonelli, non-negativity of $W$ allows inversion of integrations, $R(m, \delta) < R(m, \delta_m)$. Whence, since $R(m, \delta_m) < \infty$, there is a contradiction. So $\delta_m$ is an NGBDF.

**Example 2.1.** For $E = \{x \in \text{Binomial}(n, \theta), \ \theta = [0, 1] = D, \ W(d, \theta) = (d-\theta)^2\}$, $\delta_m = \{d_m(x) = x/n\}$ is both an NGBDF and an EGBDF with respect to $m$ given by $dm(\theta) = d\theta/\theta(1-\theta)$.

**Example 2.2.** For $E = \{x \in \text{N}(\theta, 1), \ \theta = [D, \ W(d, \theta) = (d-\theta)^2\}$, $\delta_m = \{d_m(x) = x\}$ is an EGBDF with respect to $m$ given by $dm(\theta) = d\theta$ but no NGBDF with respect to the same $m$ is definable.

The following examples concern CGBDFs.
Example 2.3. In Example 2.2, \( \delta_m \) is a CGBDF. This result is a special case of Theorem 4.1 below.

Example 2.4. For \( \mathcal{EUR} \) having \( X = \{x_1, x_2, \ldots\} \), \( \Theta = \{\theta_1, \theta_2, \ldots\} \), \( D = \{d_1, d_2\} \), suppose

\[
[W(l, j) - W(2, j)] f(x_i | \theta_j) \, d \mu \, (x_1) \, dm(\theta_j)
\]

\[
= \frac{1}{(l+1)} \left( \frac{1}{i+1} \right)^j - \frac{1}{(l+2)} \left( \frac{i+1}{i+2} \right)^j = \Delta_{ij}, \quad \text{say.}
\]

Now \( \sum \Delta_{ij} < 0, \; i = 1, 2, \ldots \), so that the EGBDF is \( \delta_m = \{d_m(x_i) = d_1\} \).

However \( \sum \Delta_{ij} > 0, \; j = 1, 2, \ldots \), while \( \sum \Delta_{ij} < \infty \), whence it can be seen that the CGBDF is \( \delta^*_m = \{d^*_m(x_i) = d_2\} \). So this is an example where both EGBDF and CGBDF are definable and are drastically different.

3. Admissibility. THEOREM 3.1. Suppose \( \Theta \) has a topology \( \mathcal{J} \) of open sets. If

(i) given a decision function \( \delta_0 \), the risk function \( r(\theta, \delta) \) of any \( \delta \) better than or equivalent to \( \delta_0 \) is continuous with respect to \( \mathcal{J} \)

(ii) there exists a sequence of generalized priors \( \{\mu_i\} \) such that

(a) \( \liminf_{i \to \infty} m_i(T_0) > 0 \) for all non-null open sets \( T_0 \) in \( \mathcal{J} \) (b) \( \lim_{i \to \infty} \Delta_{m_i} (\delta_0, \delta_i) = 0 \) where \( \delta_i \) is a CGBDF with respect to \( m_i \)

then \( \delta_0 \) is admissible.

PROOF. Suppose \( \delta_0 \) is inadmissible. Then there exists a \( \delta^+ \) such that

\( r(\theta, \delta^+) \leq r(\theta, \delta_0), \; \theta \in \Theta \), and \( r(\theta, \delta^+) < r(\theta, \delta_0) \) for some \( \theta_0 \). But, by assumption, \( r(\theta, \delta^+) \) and \( r(\theta, \delta_0) \) are continuous. So there is a neighbourhood
of $\theta_0$, say $T_{\theta}$, with $r(\theta, \delta_0) > r(\theta, \delta^+) + \epsilon$, $\epsilon > 0$, $\theta \in T_{\theta}$. Whence
\[ \Delta_{m_1}(\delta_0, \delta^+) \geq \epsilon \Delta_{m_1}(T_{\theta}), \quad i = 1, 2, \ldots \]

But
\[ 0 \leq \Delta_{m_1}(\delta^+, \delta) = \Delta_{m_1}(\delta_0, \delta) - \Delta_{m_1}(\delta_0, \delta^+) \]

as $i \to \infty$, the lim inf of the right hand side of (3.1) is negative. The contradiction proves $\delta_0$ admissible.

**COROLLARY 3.1.** If (i) $\delta$ is a CGBDF with respect to $m$ (ii) the $r(\theta, \delta)$ of any $\delta$ better than or equivalent to $\delta_0$ is continuous with respect to $J$
(iii) $m(T_{\theta}) > 0$ for all open non-null sets $T_{\theta}$ in $J$ then $\delta_0$ is admissible.

**PROOF.** In Theorem 3.1, set $m_i = m$ ($i=1,2,\ldots$) and $\delta_0 = \delta_0$.

**4. The exponential family.** This specialization is given by:
\[ f(x | \theta) = \beta(\omega) \exp(\omega x) \]
\[ \theta = \theta(\omega) = E(x | \omega) = \beta'(\omega)/\beta(\omega) \]
\[ \Omega = \{ \omega | \beta(\omega)^{-1} = \int \exp(\omega x) \, d\mu(x) < \infty \} = (\omega, \omega) \]
\[ \Theta = \{ \theta = \theta(\omega) | \omega \in \Omega \} = (\theta, \theta) \]
\[ W(d\theta, \theta) = (d\theta)^2 \]

**LEMMA 4.1.** If $r(\theta, \delta) < \infty$ for all $\theta \in \Theta$, $r(\theta, \delta)$ is continuous in $\theta$.

**PROOF.** Now $r(\theta, \delta) = E[d(x)^2 | \theta] - 2 \theta E[d(x) | \theta] + \theta^2$. Observe that
for $\omega_0 = \omega(\theta_0)$ and $\omega = \omega(\theta)$ in $\Omega$
\[ |E[d(x)^2 | \theta] - E[d(x)^2 | \theta_0]| \]
\[ \leq \int_{-\infty}^\infty \beta(\omega_0) d(x)^2 e^{\omega_0 x} \beta(\omega_0)^{-1} \beta(\omega) e^{(\omega-\omega_0)x-1} \, d\mu(x) \]
\[ + \int_{-\infty}^\infty \beta(\omega) d(x)^2 e^{\omega x} \beta(\omega)^{-1} \beta(\omega) e^{-(\omega-\omega_0)x} \, d\mu(x) \].
Since $\beta(\omega)$ is continuous in $\Omega$, $|\beta(\omega)^{-1} \beta(\omega) e^{(\omega - \omega_0)x} - 1|$ and $|1 - \beta(\omega)^{-1} \beta(\omega) e^{-(\omega - \omega_0)x}|$ are, in $(-\infty, 0)$ and $(0, \infty)$ (resp.), bounded functions of $x$ tending to zero as $\omega \downarrow \omega_0$. So, since $\theta(\omega)$ is continuous and monotone increasing and $\int \beta(\omega) d^2(x) e^{\omega_0 x} d\mu(x) < \infty$ by assumption, we have $E[d(x)^2 | \theta] \to E[d(x)^2 | \theta_0]$ as $\theta \searrow \theta_0$. Similarly as $\theta \nearrow \theta_0$.

So $E[d(x)^2 | \theta]$ is continuous in $\theta$. An almost similar argument establishes the continuity of $E[d(x) | \theta]$ and the lemma follows.

Karlin's sufficient condition [3], $K_{\lambda}$, say, for the admissibility of the estimator $x/(\lambda + 1)$ of $\theta$ is that $\int_a^b \beta^{-\lambda}_x(\omega) d\omega \to \infty$ as $a \to \omega$ and as $b \to \omega$.

THEOREM 4.1. If $K_{\lambda}$ holds, $\delta_{\lambda} = \{d(x) = x/(\lambda + 1)\}$ is a CGBDF with respect to $m_{\lambda}$ given by $dm_{\lambda} = \beta(\omega)^{\lambda} d\omega$.

PROOF. Writing $\gamma = 1/(\lambda + 1)$,

$$r(\theta, \delta) - r(\theta, \delta_{\lambda}) = E\{[d(x) - \theta]^2 | \theta\} - E\{(\gamma x - \theta)^2 | \theta\}$$

$$= T(\omega) + 2 E\{[d(x) - \gamma x][\gamma x - \theta] | \theta\}$$

where $T(\omega) = E\{[d(x) - \gamma x]^2 | \theta\}$. If $\Delta_{m_{\lambda}}(\delta_x, \delta_{\lambda}) = \infty$, $\Delta_{m_{\lambda}}(\delta_x, \delta_{\lambda}) > 0$ automatically. While $\Delta_{m_{\lambda}}(\delta_x, \delta_{\lambda}) < \infty$ implies $\int_a^b d\omega \beta(\omega)^{\lambda} [r(\theta, \delta) - r(\theta, \delta_{\lambda})] < \infty$ for arbitrary $a > \omega$, $b < \bar{\omega}$, whence by the continuity of $\beta(\omega)$ and $\theta(\delta_x, \delta_{\lambda})$ in $\Omega$, $\int_a^b d\omega \beta(\omega)^{\lambda} r(\theta, \delta_{\lambda}) < \infty$ and $\int_a^b d\omega \beta(\omega)^{\lambda} r(\theta, \delta) < \infty$; which justify the inversions of order of integration that give

$$\int_a^b d\omega \beta(\omega)^{\lambda} [r(\theta, \delta) - r(\theta, \delta_{\lambda})]$$

$$= \int_a^b d\omega \beta(\omega)^{\lambda} T(\omega) + 2 \int d\mu(x) [d(x) - \gamma x] \int_a^b d\omega \beta(\omega)^{\lambda + 1} e^{\omega_0 (\gamma x - \theta)}.$$
Reference to pp. 413-414 of [3] shows that, if $K_{\lambda}$ holds, the right hand side of (4.1) is non-negative, that is, letting $a \to \omega$, $b \to \bar{\omega}$, $\Delta_{m_{\lambda}\alpha}(\delta, \delta, \lambda) \geq 0$ if $K_{\lambda}$ holds.

Define $K_{\lambda,\alpha} = \int_a^b \beta(\omega)^{-\lambda} e^{-\alpha\omega} d\omega \to \infty$ as $a \to \omega$ and as $b \to \bar{\omega}$.

THEOREM 4.2. If $K_{\lambda,\alpha}$ holds, $\delta_{\lambda,\alpha} = \{d(x) = (x+\alpha)/(\lambda+1)\}$ is a CGBDF with respect to $m_{\lambda,\alpha}$ given by $dm_{\lambda,\alpha} = \beta(\omega)^{\lambda} e^{\alpha\omega} d\omega$.

PROOF. It is readily verified that

\begin{equation}
\Delta_{m_{\lambda,\alpha}}(\delta, \delta, \lambda, \alpha) = \int d\omega \beta_y(\omega)^{\lambda} \int d\mu_y(y) \beta_y(\omega) e^{\omega y} \{[d(y) - \phi]^2 - [y/(\lambda+1) - \phi]^2\}
\end{equation}

where $y = x - \alpha/\lambda$, $\phi = \theta - \alpha/\lambda = 2(1/\omega)$, $d\mu_y(y) = d\mu(y+\alpha/\lambda)$, $d(y) = d(y+\alpha/\lambda) - \alpha/\lambda$ and $\beta_y(\omega) = \beta(\omega) \exp(\alpha \omega/\lambda)$. Theorem 4.1 then applies to show that the right hand side of (4.2) is non-negative if $K_{\lambda}$ holds with $\beta_y(\omega)$ for $\beta(\omega)$, that is, if $K_{\lambda,\alpha}$ holds. So $\Delta_{m_{\lambda,\alpha}}(\delta, \delta, \lambda, \alpha) \geq 0$ if $K_{\lambda,\alpha}$ proving the theorem.

THEOREM 4.3. If $K_{\lambda,\alpha}$ holds then $\delta_{\lambda,\alpha}$ is admissible.

PROOF. $\tau(\theta, \delta, \alpha)$ is readily verified continuous in $\Omega$. Theorem 4.2, Lemma 4.1 and Corollary 3.1 then give the result.

[Theorem 4.3 is also proved in Cheng Ping and is a simple extension of Karlin's Theorem 1. We reprove it to complete and illustrate our approach.]

The possibility of finding a counter example to Karlin's conjecture
that $K_{\lambda}$ is necessary for the admissibility of $\delta_{\lambda}$, involving for some $\mu$
a sequence of $\alpha$ values tending to zero for which

(i) $K_{\lambda, \alpha}$ holds but $K_{\lambda}$ does not

(ii) $\lim_{\alpha \to 0} \Delta m_{\lambda, \alpha} (\delta_{\lambda}, \delta_{\lambda}, \alpha) = 0$

is dismissed by Theorem 4.4 below. [The particular choice $\mu = \{ d \mu(x) = x^2 \exp(-1/x) \, dx, \, 0 < x < 1 ; \, d \mu(x) = 0 \text{ otherwise} \}$ does in fact obey $K_{\lambda, \alpha}$
for $\lambda > 0$ and $0 < \alpha < \frac{1}{2} \lambda$ but not $K_{\lambda}$.

THEOREM 4.4. If there is a sequence $\{ \alpha_i \}$ of $\alpha$-values tending to zero,
such that $K_{\lambda, \alpha_i}$ holds for $i = 1, 2, \ldots$ and $\lim_{\alpha_i \to 0} \Delta m_{\lambda, \alpha_i} (\delta_{\lambda}, \delta_{\lambda}, \alpha_i) = 0$
then $K_{\lambda}$ holds.

PROOF. It is verifiable, using integration by parts that $\Delta m_{\lambda, \alpha_i} (\delta_{\lambda}, \delta_{\lambda}, \alpha_i) 
\to 0$ $\Rightarrow$ $\int_0^{\alpha_i} d\omega \beta(\omega)^{\lambda} e^{\sigma \omega} \to 0$.

By Schwarz

$\int_a^b \beta(\omega)^{\lambda} d\omega \left[ \int_a^b \beta(\omega)^{\lambda} e^{\sigma \omega} d\omega \right] \geq \left[ \int_a^b e^{\sigma \omega} d\omega \right]^2
= \left( e^{\frac{1}{2} \sigma a} - e^{\frac{1}{2} \sigma b} \right)^2$

If $\bar{\omega} < \infty$ then, for any $i$, $K_{\lambda, \alpha_i} \Rightarrow \int_a^b \beta(\omega)^{-\lambda} d\omega \to \infty$ as $b \to \bar{\omega}$. If $\bar{\omega} = \infty$
then the choice of $b = |\alpha_i|^{-1}$ and fixed $\bar{a}$ in (4.7) establishes that $\int_a^b \beta(\omega)^{-\lambda} d\omega 
\to \infty$ as $b \to \bar{\omega}$. Similarly $\int_a^b \beta(\omega)^{-\lambda} d\omega \to \infty$ as $a \to \bar{\omega}$. Hence $K_{\lambda}$ holds.
REFERENCES


13. Three types of generalized Bayes decision functions NGBDF, EGBDF, CGBDF are defined and their relationships examined. A sufficient condition is given for the admissibility of a CGBDF. In the context of the exponential family and quadratic loss estimation, a sufficient condition that $x/(\lambda+1)$ be a CGBDF for the mean is shown to be the same as Karlin's condition for the admissibility of $x/(\lambda+1)$. 