*The Annals of Statistics*
1984, Vol. 12, No. 3, 945–957

# OPTIMAL FIXED SIZE CONFIDENCE PROCEDURES FOR A RESTRICTED PARAMETER SPACE[1]

MEHMET ZEYTINOGLU AND MAX MINTZ

*University of Pennsylvania*

Optimal fixed size confidence procedures are derived for the mean of a normal random variable with known variance, when the mean is restricted to a compact interval. These confidence procedures are, in turn, based on the solution of a related minimax decision problem which is characterized by a zero-one loss function and a compact interval parameter space. The minimax rules obtained are nonrandomized, admissible, Bayes procedures. The decision-theoretic results are extended in two ways: (i) structurally similar (admissible) Bayes minimax rules are also obtained when the sampling distribution has a density function which is unimodal, symmetric about the location parameter and possesses a (strictly) monotone likelihood ratio; (ii) structurally similar minimax rules (minimax within the class of nonrandomized, odd, monotone procedures) are again obtained when the assumption of a monotone likelihood ratio is relaxed.

**Introduction.** We begin with the following minimax location parameter estimation problem: Let $Z$ denote a single observation of a scalar random variable, where $Z \in N(\theta, 1)$. Assume that $\theta$ is an unknown element of the given compact interval $\Omega = [-d, d]$. Let $\mathscr{A}$ denote the action space of the statistician. Here, $\mathscr{A} = [-d, d]$. Let $L(a, \theta)$ denote the zero-one loss function defined on $\mathscr{A} \times \Omega$:

$$L(a, \theta) = 0, \quad |a - \theta| \le e; \tag{1}$$
$$L(a, \theta) = 1, \quad |a - \theta| > e;$$

where $e > 0$, is given. Based on these previous definitions and assumptions, we wish to determine a minimax estimate $\delta^*(Z)$ for $\theta$. In this paper, we obtain minimax admissible Bayes estimates for $\theta$.

The interesting connection between the minimax rule $\delta^*$ and an optimal fixed size confidence procedure is obtained by noting that $C^*(Z) = [\delta^*(Z) - e, \delta^*(Z) + e]$ can be interpreted as a confidence procedure of size $2e$ which has the highest confidence coefficient, $\inf_\theta P_\theta[\theta \in C^*(Z)]$.

In viewing our present decision problem, in the context of the previous literature on minimax decision theory, we note that Wolfowitz (1950) solved the infinite interval version of the present problem. In particular, Wolfowitz showed that when $\Omega = E^1$, the equalizer rule $\delta^*(Z) = Z$, is minimax. A challenging aspect of this present problem is the compactness of $\Omega$. Recent papers by Bickel (1981),

Received December 1982; revised January 1984.

[1] The authors wish to express their gratitude for the following support. The research of the first author was supported in part by a Fulbright Grant. The research of the second author was supported in part by a National Research Council Associateship at the US Army Ballistic Research Laboratory.

*AMS* 1980 *subject classifications*. Primary 62F25; secondary 62C20.

*Key words and phrases*. Confidence procedures, minimax procedures, zero–one loss, restricted parameter space.

and Casella and Strawderman (1981) have explored the implications of a compact interval parameter space for minimax estimation problems under quadratic loss. Earlier research on minimax estimation theory dealing with compact parameter spaces has appeared in papers by Nelson (1966) and Ghosh (1964). The "blind" version of our present problem, i.e., the limiting case where the statistician makes no observation, has been considered by Egerland (1979). In this extreme case, the minimax strategy is a randomized procedure, which is characterized by a discrete distribution with a finite number of atoms in $[-d, d]$.

**Preliminary observations.**

OBSERVATION 1. We can restrict our attention to nonrandomized monotone decision rules. This follows by noting that the underlying decision problem is a monotone estimation problem in the sense of Karlin and Rubin (1956). (See also Brown, 1976, and Berger, 1980.) In particular, we note that:

(a) $L(a^*, \theta)$ attains its minimum, as a function of $a^*$, at a point $a^* = q(\theta)$, where $q$ is a nondecreasing function of $\theta$;

(b) $L(a^*, \theta)$, considered as a function of $a^*$, is nondecreasing as $a^*$ moves away from $q(\theta)$;

(c) $\mathscr{A}$ and $\Omega$ are closed intervals in $E^1$;

(d) the conditional distribution of $Z$ given $\theta$ is nonatomic;

(e) the distribution of $Z$ has a monotone likelihood ratio.

Conditions (a)–(e) above are sufficient to insure that the class of nonrandomized monotone decision rules is essentially complete.

OBSERVATION 2. We can further restrict our attention to the subclass of nonrandomized decision rules for which $\delta(Z) = -\delta(-Z)$ for all $Z$. This follows by noting that the underlying decision problem has a symmetric loss function, and that the sampling density is symmetric about its location parameter. (See Berger, 1980, and Brown, 1976.)

OBSERVATION 3. The range space of the decision rules can be restricted to the interval $I = [-(d-e), (d-e)]$. This follows by noting that $L(d\text{-}e, \theta) = 0$, for $\theta \in [d-e, d]$, with a corresponding result for negative arguments.

We denote the risk function of $\delta$ by $R(\delta, \theta)$.

OBSERVATION 4. In the present decision problem, $R(\delta, \theta)$ is given by:

$$R(\delta, \theta) = P_\theta[\delta(Z) > \theta + e] + P_\theta[\delta(Z) < \theta - e]. \tag{2}$$

Since $\delta$ is monotone, $R(\delta, \theta)$ can be expressed by:

$$R(\delta, \theta) = F(-\sup\{z: \delta(z) \le \theta + e\} + \theta) + F(\inf\{z: \delta(z) \ge \theta - e\} - \theta), \tag{3}$$

where $F$ is the CDF of the standard normal.

**A minimax decision rule $\delta^*$, when $d$ is an integer multiple of $e$.** In this section, we construct a minimax rule $\delta^*$, for the case when $d$ is an integer multiple of $e$. Let $d = (2n + 1)e + c$, where $n = 1, 2, \cdots$, and the parameter $c$ equals zero ($e$) if $d$ is an odd (even) multiple of $e$. Our construction of $\delta^*$ is based on a sequence of observations and lemmas which we develop in this section. Due to the existing symmetry in this problem setting, function definitions may be stated for nonnegative arguments.

Let $\Delta$ denote the following parameterized family of decision rules $\delta$:

$$(4) \qquad \delta(Z) = \begin{cases} d - e, & c + a_n + 2ne \le Z; \\ \vdots & \\ Z - a_2, & c + a_2 + 2e \le Z < c + a_2 + 4e; \\ 2e + c, & c + a_1 + 2e \le Z < c + a_2 + 2e; \\ Z - a_1, & c + a_1 \le Z < c + a_1 + 2e; \\ c, & c \le Z < c + a_1; \\ Z, & 0 \le Z < c; \end{cases}$$

where: $0 \le a_1 \le a_2 \le \cdots \le a_n < \infty$. The selection of the parameters $a_1, a_2, \cdots, a_n$ will be described in the sequel.

Let $\Lambda$ denote the following parameterized family of probability distributions $\lambda$, with density functions $f(\theta)$, defined on $\Omega$:

$$(5) \qquad f(\theta) = \textstyle\sum_{i=1}^{n+1} k_i[U(\theta + (2i - 1)e + c) - U(\theta - (2i - 1)e - c)];$$

where: $U(t) = 1$ if $t \ge 0$; $U(t) = 0$ if $t < 0$; and the parameters $\{k_i\}$ are nonnegative normalizing constants.

LEMMA 1. *Each $\delta \in \Delta$ (4) is Bayes with respect to a prior density function, $f \in \Lambda$, for some choice of parameters $\{k_i\}$.*

PROOF. See Appendix 1.

OBSERVATION 5. The risk function of each $\delta \in \Delta$ is given by:

$$(6) \quad R(\delta, \theta) = \begin{cases} F(a_n - e), & d - 2e < \theta \le d; \\ F(a_{n-1} - e), & \theta = d - 2e; \\ F(-a_n - e) + F(a_{n-1} - e), & d - 4e < \theta < d - 2e; \\ \vdots & \\ F(-a_2 - e) + F(a_1 - e), & c + e < \theta < c + 3e; \\ F(-a_2 - e) + (c/e)F(-e) + (1 - c/e)F(-a_1 - e), & \theta = c + e; \\ F(-a_1 - e) + (c/e)F(-e) + (1 - c/e)F(-a_1 - e), & 0 < \theta < c + e; \\ 2F(-a_1 - e), & \theta = 0. \end{cases}$$

We observe that $R(\delta, \theta)$ is piecewise constant over the sets of a finite partition of $\Omega$. We note specifically that $R(\delta, \theta)$ is piecewise constant over $n + 1$ nondegenerate subintervals of $[0, d]$, where a nondegenerate interval is an interval which consists of more than a single point.

Based on this latter observation, we define an $(n + 1)$-dimensional vector $J(\mathbf{a})$, with components equal to $R(\delta, \theta)$, over the $n + 1$ nondegenerate partition intervals of $[0, d]$:

$$J(\mathbf{a}) = [(F(-a_1 - e) + (c/e)F(-e) + (1 - c/e)F(-a_1 - e)), \quad (F(-a_2 - e) + F(a_1 - e)), \quad \vdots \quad (F(a_n - e))]^T, \tag{7}$$

where $\mathbf{a} = (a_1, a_2, \cdots, a_n)^T \in \mathscr{D} \subset E^n$, such that $0 \le a_1 \le a_2 \le \cdots \le a_n < \infty$.

OBSERVATION 6. For each $\delta \in \Delta$, the max $R(\delta, \theta)$ occurs at one or more of the nondegenerate subintervals in (6).

LEMMA 2. *The parameter vector* $\mathbf{a} \in \mathscr{D}$, *can be chosen to equalize the components of* $J(\mathbf{a})$ *to a common value* $M$, *where* $M < 2F(-e)$.

PROOF. See Appendix 2.

LEMMA 3. *If* $\mathbf{a} \in \mathscr{D}$, *satisfies the equalization condition of Lemma* 2, *then* $\delta^*$ *is minimax.*

PROOF. The proof follows by noting that $\delta^*$ is a Bayes rule with respect to a prior distribution $\lambda^*$, by Lemma 1; and that $\lambda^*$ assigns probability one to the subset of $\Omega$ on which $R(\delta^*, \theta)$ attains its maximum value, by Lemma 2 and Observation 6.

COROLLARY 1. *The prior distribution* $\lambda^*$ *referred to in Lemma* 3 *is a least favorable prior distribution.*

PROOF. See Ferguson (1967) or Berger (1980).

**An example.** We illustrate the previous results by means of the following example: let $d = 3e$, and $e = 0.1$. By referring to (4) and (6), the minimax rule $\delta^*$ and risk function $R(\delta^*, \theta)$ are, respectively:

$$\delta^*(Z) = \begin{cases} 2e, & a + 2e \le Z; \\ Z - a, & a \le Z < a + 2e; \\ 0, & 0 \le Z < a; \end{cases}$$

$$R(\delta^*, \theta) = \begin{cases} F(a - e), & e < \theta \le 3e; \\ F(-a - e), & \theta = e; \\ 2F(-a - e), & 0 \le \theta < e; \end{cases}$$

for some $a > 0$.

Let $J(a) = (2F(-a - e), F(a - e))^T$. Upon equating the two components of

$J(a)$, one obtains the relation $2F(-a-e) = F(a-e)$. Solving this last relation for $a$, we obtain $\mathring{a} = 0.4$. The corresponding minimax risk is 0.62.

By employing equation (A.5) (from Appendix 1), we obtain the corresponding least favorable prior distribution $\lambda^*$, with density $f^*$:

$$f^*(\theta) = 0.14[U(\theta + e) - U(\theta - e)] + 1.62[U(\theta + 3e) - U(\theta - 3e)].$$

It is illustrative to compare the minimax risk of $\delta^*$ with the maximum risk of the truncated maximum likelihood estimate $\delta_T$:

$$\delta_T(Z) = \begin{cases} (d-e), & (d-e) \le Z; \\ Z, & -(d-e) \le Z \le (d-e); \\ -(d-e), & Z \le -(d-e). \end{cases}$$

The risk function of $\delta_T$ is:

$$R(\delta_T, \theta) = \begin{cases} 2F(-e), & \theta \in (-e, e); \\ F(-e), & \theta \in [-3e, -e] \cup [e, 3e]. \end{cases}$$

Thus, the corresponding maximum risk is $2F(-e) = 0.92$.

**A minimax decision rule $\delta_M$, when $d$ is a noninteger multiple of $e$.** In this section, we obtain a minimax decision rule $\delta_M$, for the case when $d$ is a noninteger multiple of $e$. We first illustrate the main result by examining the special case when $d \in (e, 2e]$. The general case is then addressed in the sequel.

Let $d \in (e, 2e]$, and define a decision rule $\delta_M$ by:

$$\delta_M(Z) = \begin{cases} (d-e), & (d-e) \le Z; \\ Z, & -(d-e) \le Z \le (d-e); \\ -(d-e), & Z \le -(d-e). \end{cases} \tag{8}$$

OBSERVATION 7. The decision rule $\delta_M$ is Bayes with respect to the uniform distribution on $[-d, d]$. (See Appendix 1.)

OBSERVATION 8. The risk function $R(\delta_M, \theta)$ is:

$$R(\delta_M, \theta) = \begin{cases} 0, & \theta \in [d-2e, 2e-d]; \\ F(-e), & \theta \in [-d, d-2e) \cup (2e-d, d]. \end{cases}$$

OBSERVATION 9. The decision rule $\delta_M$ is Bayes and unique up to equivalence. Thus, $\delta_M$ is admissible. (See Ferguson, 1967.)

Define a prior distribution $\lambda^*$ with a density function $f^*$:

$$f^*(\theta) = \begin{cases} 0, & \theta \in [d-2e, 2e-d]; \\ 1/4(d-e), & \theta \in [-d, d-2e) \cup (2e-d, d]. \end{cases}$$

OBSERVATION 10. The admissible Bayes decision rule $\delta_M$ in (8), referred to in Observation 7, is also Bayes with respect to $\lambda^*$.

OBSERVATION 11. The prior distribution $\lambda^*$ assigns probability one to the subset of $\Omega$ on which $R(\delta_M, \theta)$ attains its maximum value. (See Observation 8.)

OBSERVATION 12. The decision rule $\delta_M$ and prior distribution $\lambda^*$ are, respectively, a minimax decision rule and a least favorable prior distribution. This observation is a consequence of Observations 10 and 11.

OBSERVATION 13. If $d \in (e, 2e)$, the uniform distribution on $[-d, d]$ is not a least favorable prior distribution, since the Bayes risk of $\lambda^*$ exceeds the Bayes risk of the uniform distribution on $[-d, d]$.

Thus, we have demonstrated the existence of a minimax admissible Bayes rule, which is Bayes with respect to a prior distribution which is *not* least favorable. This interesting result is a common characteristic of this class of minimax decision problems when $d$ is not an integer multiple of $e$.

Next, we address the general case.

Here, we let $d = (2n + 1)e + c - c_0$, $n = 1, 2, \cdots$, where the parameters $c$ and $c_0$ are chosen such that:

- $c \in [0, e]$ and $c_0 \in [0, e)$;
- $c = c_0 = 0$, when $d$ is an odd integer multiple of $e$;
- $c = e$, and $c_0 = 0$, when $d$ is an even integer multiple of $e$;
- $c = 0$, whenever $c_0 > 0$;
- $c_0 = 0$, whenever $c > 0$.

We generalize the definition of the decision rule $\delta_M$ by:

$$\delta_M(Z) = \begin{cases} d - e, & c - c_0 + a_n + 2ne \leq Z; \\ \vdots & \\ Z - a_2, & c - c_0 + a_2 + 2e \leq Z < c - c_0 + a_2 + 4e; \\ 2e + c - c_0, & c - c_0 + a_1 + 2e \leq Z < c - c_0 + a_2 + 2e; \\ Z - a_1, & c + a_1 \leq Z < c - c_0 + a_1 + 2e; \\ c, & c \leq Z < c + a_1; \\ Z, & 0 \leq Z < c; \end{cases} \tag{9}$$

where: $0 \leq a_1 \leq a_2 \leq \cdots \leq a_n < \infty$ as before.

For any $d$ value which is a noninteger multiple of $e$, the parameter vector $\mathbf{a}$ is selected to equal $\mathbf{a}'$, where $\mathbf{a}'$ denotes the parameter vector corresponding to the minimax decision rule $\delta^*$ associated with $\Omega' = [-d', d']$, where $d'$ is the smallest integer multiple of $e$ such that $d' > d$.

OBSERVATION 14. Let $M$ denote the maximum value of the risk function $R(\delta^*, \theta)$. Define $S_1 = [-e, -(e - c_0)] \cup [(e - c_0), e]$ and $S_2 = [-(e - c), (e - c)]$. The risk function corresponding to $\delta_M$ defined by (9) exhibits the following features: If $c = 0$ ($c_0 = 0$) and $c_0 > 0$ ($c > 0$), then

$$R(\delta_M, \theta) \leq M, \quad \theta \in \Omega - S_1 \quad (\theta \in \Omega - S_2);$$

and

$$R(\delta_M, \theta) < M, \quad \theta \in S_1 \qquad (\theta \in S_2).$$

OBSERVATION 15. The decision rule $\delta_M$ is Bayes with respect to a prior density of the form:

$$f(\theta) = k_1[U(\theta + e + c) - U(\theta - e - c)] \\ + \textstyle\sum_{i=2}^{n+1} k_i[U(\theta + (2i - 1)e + c - c_0) - U(\theta - (2i - 1)e - c + c_0)]. \tag{10}$$

Further, $\delta_M$ is unique up to equivalence. Thus $\delta_M$ is admissible.

LEMMA 4. *The decision rule $\delta_M$ (9) is a minimax rule.*

PROOF.

*Case* 1. $(0 < c_0 < e)$.

Let $d = (2n + 1)e - c_0$, and let $\Lambda'$ denote the following parameterized family of prior distributions $\lambda'$ with density functions $f'$:

$$f'(\theta; k_1, \cdots, k_{n+1}) = I_{(\Omega-\Gamma_0)}(\theta) f(\theta; k_1, \cdots, k_{n+1}), \tag{11}$$

where:

- $\Gamma_0 = \cup_{i=1}^{n} \{[-(2i - 1)e - c_0, -(2i - 1)e + c_0] \cup [(2i - 1)e - c_0, (2i - 1)e + c_0]\};$
- $I_A(\theta)$ denotes the indicator function of the set $A$;
- $f(\theta; k_1 \cdots, k_{n+1})$ is defined by (10);
- the nonnegative parameters $\{k_i\}$ are selected to normalize $f'$.

OBSERVATION 16. The admissible Bayes decision rule $\delta_M$ in (9) is also Bayes with respect to $f^* \in \Lambda'$ for some choice of $\{k_i\}$.

OBSERVATION 17. This prior distribution $\lambda^*$ (Observation 16) assigns probability one to the subset $\Omega - S_1$, where $S_1$ is defined in Observation 14. Further, $R(\delta_M, \theta) = M$ on $\Omega - S_1$, except on a finite point set. Thus, $\lambda^*$ assigns probability one to the subset of $\Omega$ on which $R(\delta_M, \theta)$ attains its maximum value.

OBSERVATION 18. The decision rule $\delta_M$ and prior distribution $\lambda^*$ are, respectively, a minimax decision rule and a least favorable prior distribution.

*Case* 2. $(0 < c < e)$.

Let $d = (2n + 1)e + c$, and let $\Lambda'$ denote the following parameterized family of prior distributions $\lambda'$ with density functions $f'$:

$$f'(\theta; k_1, \cdots, k_{n+1}) = I_\Gamma(\theta) f(\theta; k_1, \cdots, k_{n+1}), \tag{12}$$

where:

- $\Gamma = \bigcup_{i=1}^{n} \{[-(2i-1)e - c, -(2i-1)e + c] \cup [(2i-1)e - c, (2i-1)e + c]\}$;
- the remaining notation and conventions are identical to those of Case 1.

For Case 2, we obtain the analogous results—that $\delta_M$ and a specific $\lambda^*$ are, respectively, minimax and least favorable—by restating Observations 16–18 with $S_1$ replaced by $S_2$, where $S_2$ is defined in Observation 14.

REMARK. Although $\delta_M$ is Bayes with respect to $\lambda^*$, $\delta_M$ is generally *not* unique up to equivalence (with respect to $\lambda^*$). Thus, $\lambda^*$ does *not* provide a direct means to establish the admissibility of $\delta_M$. The admissibility of $\delta_M$ is, instead, a consequence of Observation 15. However, when $d$ is a noninteger multiple of $e$, no prior density of the form (10) is least favorable.

**Extensions.** Although this location parameter estimation problem was posed initially with a normal sampling distribution, the (admissible) Bayes minimax estimates and least favorable prior distributions obtained in this paper remain valid when the sampling distribution has a density function which is unimodal, symmetric about the location parameter, and possesses a (strictly) monotone likelihood ratio. If the sampling distribution is normal, then the problem of $N$ i.i.d. observations can be reduced to a problem in the format of the previous analysis by making use of the sufficiency and distributional properties of the sample mean. Further, a standard translation can be employed when $\Omega$ is asymmetric.

Finally, we address the question of relaxing the assumption of a monotone likelihood ratio: Let $\mathscr{L}$ denote the class of nonrandomized, odd, monotone (nondecreasing) decision rules with range space $I = [-(d-e), (d-e)]$.

LEMMA 5. *If the sampling distribution has a density function which is unimodal and symmetric about the location parameter, then* $\delta_M$ *in* (9) *is* $\mathscr{L}$*-minimax.*

PROOF. See Appendix 3.

## APPENDIX 1

LEMMA 1. *Each* $\delta \in \Delta$ (4) *is Bayes with respect to a prior density function* $f \in \Lambda$ (5), *for some choice of parameters* $\{k_i\}$.

PROOF. It is sufficient to prove that if the sampling CDF $F(Z - \theta)$ has a density which is unimodal, symmetric about $\theta$, and possesses a (strictly) monotone likelihood ratio, then each $\delta \in \Delta$ (4) is Bayes (and unique up to equivalence) with respect to a prior density function $f \in \Lambda$ (5), for some choice of parameters $\{k_i\}$.

We restrict our attention to the case when $d$ is an integer multiple of $e$, since it illustrates all the essential steps of a general proof.

Let $\lambda$ denote a prior distribution with corresponding density (5), or equivalently for $\theta \geq 0$,

$$f(\theta) = \begin{cases} K'_{n+1}, & c + (2n-1)e < \theta \leq c + (2n+1)e; \\ K'_n, & c + (2n-3)e < \theta \leq c + (2n-1)e; \\ \vdots & \\ K'_1, & 0 \leq \theta \leq c + e; \end{cases} \tag{A.1}$$

where:

$$K'_i = \sum_{j=i}^{n+1} k_j, \quad i = 1, 2, \cdots, n+1.$$

NOTATION. The posterior density $f(\theta \mid Z)$ which corresponds to the posterior distribution $\lambda(\theta \mid Z)$ is denoted by

$$f(\theta \mid Z) = G(Z) f_s(Z - \theta) f(\theta), \tag{A.2}$$

where $f_s(Z - \theta)$ denotes the sampling density, and $G(Z)$ denotes a normalizing factor.

OBSERVATION A1. A decision rule $\hat{\theta}(Z)$ is Bayes with respect to the prior distribution $\lambda$, if and only if it minimizes the posterior risk. Hence, for each $Z$, a Bayes rule must maximize:

$$A(\hat{\theta}, Z) = \int_{\hat{\theta}-e}^{\hat{\theta}+e} f(\theta \mid Z)\, d\theta. \tag{A.3}$$

OBSERVATION A2. The maximization of $A(\hat{\theta}, Z)$ over $\hat{\theta} \in [0, (d-e)]$ for fixed $Z \geq 0$ is obtained as follows:

(i) Consider a generic segment of $\delta(Z)$ (4) on the interval

$$[c + a_{m-1} + (2m-2)e,\ c + a_{m+1} + (2m)e], \quad \text{for fixed } m, \quad 1 \leq m \leq n,$$

where:

$$a_0 = 0 \quad \text{and} \quad a_{n+1} = \infty.$$

(ii) Define the contiguous subintervals:

$$I_1 = [c + a_{m-1} + (2m-2)e,\ c + a_m + (2m-2)e);$$

$$I_2 = [c + a_m + (2m-2)e,\ c + a_m + (2m)e];$$

$$I_3 = (c + a_m + (2m)e,\ c + a_{m+1} + (2m)e].$$

(iii) Evaluate $A(\hat{\theta}, Z)$ for $Z \geq 0$, and $\hat{\theta} \in [c + (2m-2)e, c + (2m)e]$:

$$A(\hat{\theta}, Z) = G(Z)\left[K'_m \int_{\hat{\theta}-e}^{c+(2m-1)e} f_s(Z-\theta)\,d\theta + K'_{m+1} \int_{c+(2m-1)e}^{\hat{\theta}+e} f_s(Z-\theta)\,d\theta\right]. \tag{A.4}$$

In order for $\delta$ (4) to be Bayes with respect to $f(\theta)$, it is necessary that $A(\hat{\theta}, Z)$ is maximized at $\hat{\theta} = Z - a_m$, when $Z \in I_2$. By solving $dA/d\hat{\theta} = 0$, we obtain a necessary condition for a local maximum at $\hat{\theta} = Z - a_m$:

$$f_s(-a_m + e)/f_s(-a_m - e) = K'_m/K'_{m+1}, \tag{A.5}$$

which must be satisfied by the parameters $K'_m$, and $K'_{m+1}$.

(iv) For a given parameter vector $\mathbf{a}$, the parameters $\{K'_i\}$ and hence $\{k_i\}$ can be chosen to simultaneously satisfy (A.5) and normalize $f(\theta)$.

(v) As a consequence of (A.5),

$$\frac{dA/d\hat{\theta}}{Gf_s(\hat{\theta} - e - Z)K'_{j+1}} = \frac{f_s(\hat{\theta} + e - Z)}{f_s(\hat{\theta} - e - Z)} - \frac{f_s(-a_j + e)}{f_s(-a_j - e)}, \tag{A.6}$$

where:

$$\hat{\theta} \in (c + (2j-2)e, c + (2j)e), \quad \text{and} \quad 1 \leq j \leq n.$$

Here, the derivative $dA/d\hat{\theta}$ is undefined when

$$\hat{\theta} \in Q = \{c + (2j-2)e: j = 1, \cdots, n+1\}.$$

(vi) If $Z \in I_1$, then $dA/d\hat{\theta} \geq 0$ when $\hat{\theta} \in (0, c + (2m-2)e) - Q$,
and $dA/d\hat{\theta} \leq 0$ when $\hat{\theta} \in (c + (2m-2)e, d - e) - Q$;
If $Z \in I_2$, then $dA/d\hat{\theta} \geq 0$ when $\hat{\theta} \in (0, Z - a_m) - Q$,
and $dA/d\hat{\theta} \leq 0$ when $\hat{\theta} \in (Z - a_m, d - e) - Q$;
If $Z \in I_3$, then $dA/d\hat{\theta} \geq 0$ when $\hat{\theta} \in (0, c + (2m)e) - Q$,
and $dA/d\hat{\theta} \leq 0$ when $\hat{\theta} \in (c + (2m)e, d - e) - Q$.

REMARK. These inequalities are strict when the likelihood ratio is strictly monotone.

(vii) It follows from (i–vi) that

$$\hat{\theta}(Z) = \begin{cases} c + (2m-2)e, & Z \in I_1; \\ Z - a_m, & Z \in I_2; \\ c + (2m)e, & Z \in I_3; \end{cases}$$

maximizes $A(\hat{\theta}, Z)$, for fixed $Z \in \{I_1 \cup I_2 \cup I_3\}$.

OBSERVATION A3. Since (vii) holds for $m = 1, 2, \cdots, n$, each $\delta$ (4) is Bayes with respect to a prior density $f$ (5) with normalizing parameters selected to satisfy (A.5). Further, if the likelihood ratio is strictly monotone, then $\delta$ is unique up to equivalence.

OBSERVATION A4. The assumption that $F$ possesses a monotone likelihood ratio has been employed in the proof of Lemma 1. However, it can be shown that the existence of a monotone likelihood ratio is not a necessary condition for a given $\delta \in \Delta$ to be Bayes with respect to some $f \in \Lambda$.

## APPENDIX 2

LEMMA 2. *The parameter vector* $\mathbf{a} \in \mathscr{D}$, *can be chosen to equalize the components of* $J(\mathbf{a})$ *to a common value* $M$, *where* $M \in (0, 2F(-e))$.

PROOF. We restrict our attention to the case where $c = 0$. The proof for the second case, $c = e$, follows in a similar fashion.

OBSERVATION A5. $F(X - e) + F(-X - e) > 2F(-e)$ for $X > 0$.

OBSERVATION A6. $F(X - e)$ $(F(-X - e))$ is an increasing (decreasing) function of $X$.

*Step* 1. Let $S_1(a_1) = 2F(-a_1 - e)$, and $S_2(a_1) = F(a_1 - e) + F(-a_2 - e)$, where $a_2 \geq 0$ is fixed, and $a_1 \in [0, a_2]$. Consider the following sequence of observations:

- $S_1(a_1)$ is a decreasing function of $a_1$.
- $S_2(a_1)$ is an increasing function of $a_1$.
- $S_1(0) = 2F(-e)$, and $S_2(0) = F(-e) + F(-a_2 - e)$.
- $S_1(0) > S_2(0)$.
- $S_1(a_2) = 2F(-a_2 - e)$, and $S_2(a_2) = F(a_2 - e) + F(-a_2 - e)$.
- $S_1(\mathring{a}_1) = S_2(\mathring{a}_1) < 2F(-e)$, for $0 < \mathring{a}_1 < a_2$.
- $\mathring{a}_1 = g_1(a_2)$, where $g(t)$ is a continuous increasing function of $t$ which passes through the origin.
- $S_1(\mathring{a}_1) = S_1(g_1(a_2))$ is a decreasing function of $a_2$.
- Since $S_1(g_1(a_2)) = S_2(g_1(a_2))$, $S_2(g_1(a_2))$ is a decreasing function of $a_2$.

*Step* 2. Let $T_1(a_2) = F(g_1(a_2) - e) + F(-a_2 - e)$, and $T_2(a_2) = F(a_2 - e) + F(-a_3 - e)$, where $a_3 > 0$ is fixed, and $a_2 \in [0, a_3]$.

Proceeding as in Step 1, we determine $\mathring{a}_2 = g_2(a_3)$, such that $T_1(\mathring{a}_2) = T_2(\mathring{a}_2) < 2F(-e)$.

We continue in this fashion, choosing at each stage, say the $i$th stage, the parameter $\mathring{a}_i$, as a function of the next parameter, i.e., $\mathring{a}_i = g_i(a_{i+1})$. At the $n$th stage, we have $Z_1(a_n) = F(-a_n - e) + F(g_{n-1}(a_n) - e)$, and $Z_2(a_n) = F(a_n - e)$. We observe that there exists an $\mathring{a}_n \in (0, \infty)$, such that $Z_1(\mathring{a}_n) = Z_2(\mathring{a}_n) < 2F(-e)$. Once $\mathring{a}_n$ is thus determined, the other parameters $\mathring{a}_1, \cdots, \mathring{a}_{n-1}$ can be uniquely evaluated by using the previously established relations.

Thus, we have obtained a parameter vector $\mathbf{\mathring{a}}$ which equalizes the components of $J(\mathbf{\mathring{a}})$ to a common value $M$, $M \in (0, 2F(-e))$.

## APPENDIX 3

OBSERVATION A7. If the sampling distribution has a density function which is unimodal and symmetric about the location parameter, then Lemma 2 applies and the rule $\delta_M$ (9), with maximum risk $M$, is well-defined.

Let $\mathscr{L}$ denote the class of nonrandomized, odd, monotone (nondecreasing) decision rules with range space $I = [-(d-e), (d-e)]$.

OBSERVATION A8. Any $\mathscr{L}$-minimax decision rule must attain $\pm(d-e)$ for finite values of $Z$.

LEMMA 5. *If the sampling distribution has a density function which is unimodal and symmetric about the location parameter, then $\delta_M$ (9) is $\mathscr{L}$-minimax.*

PROOF BY CONTRADICTION. Assume that $\delta_M$ is not $\mathscr{L}$-minimax, i.e., assume there exists a rule $\delta_0 \in \mathscr{L}$ such that:

$$\sup R(\delta_0, \theta) < \max R(\delta_M, \theta) = M.$$

By invoking monotonicity, it follows that $\delta_0$ is either a continuous mapping of $E^1$ into $I$, or $\delta_0$ has at most jump discontinuities.

OBSERVATION (a). In order that $R(\delta_0, d) < R(\delta_M, d) = M$, it is necessary that $Z_0 < Z_M$, where: $Z_0 = \inf\{Z\colon \delta_0(Z) = d - e\}$, and $Z_M = \min\{Z\colon \delta_M(Z) = d - e\}$. For otherwise, if either $Z_0$ is not defined, or $Z_0 \geq Z_M$, then $R(\delta_0, d) \geq M$.

OBSERVATION (b). From Observation (a), it follows that $\delta_0(Z) > \delta_M(Z)$, when $Z_0 < Z < Z_M$.

OBSERVATION (c). Assume $c_0 = 0$, and $c > 0$. In order that $R(\delta_0, e) < R(\delta_M, e) = M$, it is necessary that $\delta_0(Z_1^-) < \delta_M(Z_1^-)$, where $Z_1 = \sup\{Z\colon \delta_0(Z) \leq 2e\}$, and $h(t^-)$ $[h(t^+)]$ denotes the left-hand [right-hand] limit.

OBSERVATION (d). Assume $c_0 \geq 0$, and $c = 0$. In order that $R(\delta_0, 0) < R(\delta_M, 0) = M$, it is necessary that $\delta_0(Z_2^-) < \delta_M(Z_2^-)$, where $Z_2 = \sup\{Z\colon \delta_0(Z) \leq e\}$.

OBSERVATION (e). From Observations (b), (c), and (d), it follows that there exist values $Z'$ and $Z''$ such that $\delta_0(Z') > \delta_M(Z')$ and $\delta_0(Z'') < \delta_M(Z'')$. Thus, we can define $Z_3 = \sup\{Z\colon \delta_0(Z) < \delta_M(Z)\}$.

OBSERVATION (f). If $\theta_0 = \delta_M(Z_3) + e$, where $Z_3$ is defined in Observation (e), then it follows that:

$$R(\delta_0, \theta_0^-) \geq R(\delta_M, \theta_0^-) = M,$$

which contradicts the assumption that $\delta_M$ is not $\mathscr{E}$-minimax.

COROLLARY 2. *If the sampling density is unimodal and symmetric about $\theta$, and $d \in (e, 2e]$, then $\delta_M(8)$ is $\mathscr{E}$-minimax.*

PROOF. If $\delta \in \mathscr{E}$, then either: $R(\delta, e^+) \geq F(-e)$, or $R(\delta, e^-) \geq F(-e)$. Thus, since $R(\delta_M, \theta) \leq F(-e)$ for all $\theta \in [-d, d]$, it follows immediately that $\delta_M$ (8) is $\mathscr{E}$-minimax.

**Acknowledgment.** The authors wish to thank the referees and associate editor for their thoughtful suggestions which have enhanced this paper.

## REFERENCES

BERGER, J. O. (1980). *Statistical Decision Theory.* Springer, New York.

BICKEL, P. J. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. Statist.* **9** 1301–1309.

BROWN, L. D. (1976). Notes on statistical decision theory. (Unpublished lecture notes, Cornell University, Ithaca.)

CASELLA, G. and STRAWDERMAN, W. (1981). Estimating a bounded normal mean. *Ann. Statist.* **9** 868–876.

EGERLAND, W. (1979). A game-theoretic solution to an engagement problem. US Army Ballistic Research Laboratory Technical Report ARBRL-TR-02132. AD: A 066106.

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision-Theoretic Approach.* Academic, New York.

GHOSH, M. N. (1964). Uniform approximation of minimax point estimates. *Ann. Math. Statist.* **35** 1031–1047.

KARLIN, S. and RUBIN, H. (1956). The theory of decision procedures for distributions with monotone likelihood ratio. *Ann. Math. Statist.* **27** 272–299.

NELSON, W. (1966). Minimax solution of statistical decision problems by iteration. *Ann. Math. Statist.* **37** 1643–1657.

WOLFOWITZ, J. (1950). Minimax estimates of the mean of a normal distribution with known variance. *Ann. Math. Statist.* **21** 218–230.

DEPARTMENT OF SYSTEMS ENGINEERING
THE MOORE SCHOOL OF ELECTRICAL ENGINEERING, D2
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104