The Annals of Statistics
1979, Vol. 7, No. 6, 1222–1230

MONOTONICITY OF BAYES SEQUENTIAL TESTS[1]

BY L. D. BROWN, ARTHUR COHEN, AND W. E. STRAWDERMAN
Rutgers University

Consider the problem of sequential testing of a one sided hypothesis when the risk function is a linear combination of a probability of an error component and an expected cost component. Sobel's results on monotonicity of Bayes procedures and essentially complete classes are extended. Sufficient conditions are given for every Bayes test to be monotone. The conditions are satisfied when the observations are from an exponential family. They are also satisfied for orthogonally invariant tests of a mean vector of a multivariate normal distribution and for scale invariant tests of two normal variances. Essentially complete classes of tests are the monotone tests for all situations where these sufficient conditions are satisfied.

1. Introduction and summary. Consider the problem of sequential testing of a one sided hypothesis. If the observations are drawn from an exponential family and if the risk function is the sum of the probability of an error component, and an expected cost of observation component, then Sobel [8] found an essentially complete class of tests. These tests are what we call monotone tests. In finding the essentially complete class, Sobel proved that for every prior distribution there is a monotone procedure which is Bayes with respect to this prior.

In this paper we extend Sobel's results in several directions. Sufficient conditions are given for every Bayes test to be monotone. The conditions are satisfied when the observations are from an exponential family, but work in other cases as well. For example, they are satisfied for orthogonally invariant tests of a mean vector of a multivariate normal distribution, where the maximal invariant sufficient statistic has a noncentral chi-square distribution. Another example is a test for two normal variances where the maximal invariant sufficient statistic is scaled central F.

The essential completeness of the class of monotone procedures will in all our cases follow from the arguments of Sobel and will not be repeated here. The basis of this argument is that the limits of Bayes procedures form an essentially complete class and that limits of monotone procedures are monotone. A complete class, as opposed to an essentially complete class, for this problem is discussed in Brown, Cohen, and Strawderman [3].

We remark here that the main results apply to symmetric two sided hypothesis testing problems. Also no assumption is made requiring the sum of the costs of observations to tend to infinity as the number of observations tends to infinity.

Received November 1977; revised October 1978.
[1]Research Supported by N.S.F. Grant No. MCS-75-23343.
AMS 1970 *subject classifications.* Primary 62L10; secondary 62C10, 62C07.
Key words and phrases. sequential tests, hypothesis testing, Bayes test, complete class, monotone likelihood ratio, exponential family, invariant tests.

In the next section we give preliminaries and define monotone procedures. At the end of Section 3 we prove the main theorem which states the conditions under which Bayes tests are monotone. Section 4 contains examples.

2. Preliminaries. The elements of the problem are as follows: The parameter space is Θ with typical element θ. The null space is $\Theta_1 \subset \Theta$ and the alternative space is $\Theta_2 \subset \Theta$. The action space $\mathcal{A}$ consists of pairs (n, τ), where $n = 0, 1, 2, \cdots, \infty$ is the stopping time and τ is 1 or 2, depending on whether the null hypothesis is accepted or rejected. The loss function is denoted by $L(\theta, (n, \tau))$. For $\theta \in \Theta_1$, $L(\theta, (n, 1)) = C(n)$. For $\theta \in \Theta_2$, $L(\theta, (n, 1)) = C(n) + d_2$. For $\theta \in \Theta_1$, $L(\theta, (n, 2)) = C(n) + d_1$. For $\theta \in \Theta_2$, $L(\theta, (n, 2)) = C(n)$. Here $C(n)$ represents the cost of taking n observations. Let c_j represent the cost of taking the jth observation so that $C(n) = \sum_{j=1}^{n} c_j$. We assume $c_1 = 0$ so that all procedures take at least one observation. (We could avoid this assumption by stating what Bayes procedures do at stage 0.) Also assume $c_j > 0$, for $j = 2, 3, \cdots$. It will not be necessary to assume that $C(n) \to \infty$ as $n \to \infty$, so we will be considering procedures that do not stop with a finite number of observations with probability 1. For such procedures we continue to require that a terminal decision τ be made. When $\lim_{n\to\infty} C(n)$ is finite, we will assume that the observations, to be discussed below, can provide a strongly consistent estimate of θ. This will insure that the closure of the class of Bayes solutions is an essentially complete class by virtue of an application of Brown [1], Theorem 3.18. The constants d_1 and d_2 are positive.

Random variables available to the statistician for observation are a set $\mathbf{X} = (X_1, X_2, \cdots, X_\infty)$. $\mathbf{X}$ lies in an infinite product space. Assume that there exists a sequence $\{S_j\}, j = 1, 2, \cdots, \infty$, where S_j is a real valued function of $\underline{X}_j = (X_1, X_2, \cdots, X_j)$, such that $\{S_j\}$ is a sufficient, transitive sequence for θ. The point $\mathbf{S} = (S_1, S_2, \cdots, S_\infty)$ lies in an infinite product space. We assume that there is a σ-finite measure μ defined on this space which dominates the family $\{P_\theta(\cdot), \theta \in \Theta\}$ of probability measures for $\mathbf{S}$ in the following sense: For each $j = 1, 2, \cdots$, over the σ-field generated by $(S_1, S_2, \cdots, S_j)$, the measure P_θ is dominated by μ. Write $f_\theta^{(j)}(s_1, s_2, \cdots, s_j) = dP_\theta / d\mu$, where the Radon-Nikodym derivative is taken relative to this σ-field.

A prior probability measure on Θ denoted by $\Gamma(\cdot)$ will be represented by a mixture expressed as $\pi_1\Gamma_1(\cdot) + \pi_2\Gamma_2(\cdot)$. Here, if T is a random variable with distribution Γ, then π_1 is the probability that $T \in \Theta_1$, and Γ_1 represents the conditional distribution of T, given $T \in \Theta_1$. Similarly for Γ_2. We assume $0 < \pi_1 < 1$.

We consider only procedures based on $\{S_j\}$. (See Ferguson [7], Theorem 4, page 337.) A decision function δ will consist of a set of nonnegative functions $\delta_{ij}(s_j)$, $(i = 0, 1, 2,; j = 1, 2, \cdots, \infty)$ defined for all s_j, such that $\sum_{i=0}^{2} \delta_{ij}(s_j) = 1$. The quantities $\delta_{ij}(s_j)$ represent, respectively, the probability of taking another observation, accepting H_1, and accepting H_2, when j observations have been taken.

The risk function is denoted by $R(\theta, \delta) = E_\theta\{L(\theta, \delta)\}$ and the expected risk is $R(\Gamma, \delta) = ER(\theta, \delta)$.

Now we define monotone tests.

DEFINITION 2.1. A sequential test of a one sided hypothesis is said to be monotone if for every j, there exist numbers (a_{1j}, a_{2j}), $-\infty \leqslant a_{1j} \leqslant a_{2j} \leqslant \infty$, such that $\delta_{1j}(s_j) = 0$ for $s_j > a_{1j}$; $\delta_{0j} = 0$ if $s_j < a_{1j}$ or $s_j > a_{2j}$; $\delta_{2j}(s_j) = 0$ if $s_j < a_{2j}$. If $s_j = a_{1j}$, then $\delta_{2j}(s_j) = 0$ but δ_{0j} and δ_{1j} are arbitrary. Similarly if $s_j = a_{2j}$, $\delta_{1j} = 0$ and δ_{0j} and δ_{2j} are arbitrary.

DEFINITION 2.2. A sequential test of a two sided hypothesis is said to be monotone if for every j there exist numbers $(a_{1j}, a_{2j}, a_{3j}, a_{4j})$, $-\infty \leqslant a_{1j} \leqslant a_{2j} \leqslant a_{3j} \leqslant a_{4j} \leqslant \infty$, such that $\delta_{1j}(s_j) = 0$ for $s_j < a_{2j}$ or $s_j > a_{3j}$; $\delta_{2j}(s_j) = 0$ if $a_{1j} < s_j < a_{4j}$; $\delta_{0j}(s_j) = 0$ if $s_j < a_{1j}$, $s_j > a_{4j}$, or $a_{2j} < s_j < a_{3j}$. Certain obvious randomizations are permitted when s_j equals some a_{ij}, $i = 1, 2, 3, 4$.

3. Bayes tests are monotone. Before stating the main theorem, we offer a series of lemmas. The first lemma is concerned with the existence of a dominating probability measure having a Markov chain property, and with the resulting form of the densities of the sufficient statistics. The second lemma shows how the composite vs composite hypothesis testing problem is reduced, for purposes of determining Bayes tests, to testing a simple hypothesis against a simple alternative. Two of the last four lemmas require the conditions for Bayes tests to be monotone and the four lemmas are offered to facilitate the proof of the main theorem. Now let

$$f^{(j)}_{(i)}(s_1, s_2, \cdots, s_j) = \int_{\Theta_i} f^{(j)}_\theta(s_1, s_2, \cdots, s_j)\Gamma_i(d\theta), \tag{3.1}$$

$$i = 1, 2; j = 1, 2, \cdots.$$

Note $f^{(j)}_{(i)}(s_1, s_2, \cdots, s_j)$ are densities equal to $dP_{(i)}/d\mu$ over the σ-field generated by $S_1, S_2, \cdots, S_j$, where $P_{(i)}(\cdot) = \int_{\Theta_i} P_\theta(\cdot)\Gamma_i(d\theta)$.

LEMMA 3.1. *$S_1, S_2, \cdots, S_\infty$ is a sufficient transitive sequence for the probability measures $P_{(1)}$, $P_{(2)}$. These measures are dominated by the probability measure $\nu(\cdot) = \pi_1 P_{(1)}(\cdot) + \pi_2 P_{(2)}(\cdot)$. The probability measure ν represents a marginal probability law for $(S_1, S_2, \cdots, S_\infty)$ and the process it defines is a Markov process.* (i.e., *$(S_1, S_2, \cdots, S_\infty)$ is transitive for this process as well.*)

PROOF. The proof is omitted.

At this point it is essential to introduce some further notation and make some observations that will be used later. Let $g^{(j)}_{(i)}(s_1, s_2, \cdots, s_j) = dP_{(i)}/d\nu, j = 1, 2, \cdots, \infty$, over the σ-field generated by $S_1, S_2, \cdots, S_j$. Relative to the σ-field

generated by $S_1, S_2, \cdots, S_n$, $[d\nu/d\mu] = \pi_1 f_{(1)}^{(n)}(s_1, s_2, \cdots, s_n) + \pi_2 f_{(2)}^{(n)}(s_1, s_2, \cdots, s_n)$ for $n = 1, 2, \cdots$, and

$$
\begin{aligned}
g_{(i)}^{(n)}(s_1, s_2, \cdots, s_n) &= dP_{(i)}/d\nu = [dP_i/d\mu][d\mu/d\nu] \\
&= f_{(i)}^{(n)}(s_1, s_2, \cdots, s_n)/[\pi_1 f_{(1)}^{(n)}(s_1, s_2, \cdots, s_n) + \pi_2 f_{(2)}^{(n)}(s_1, s_2, \cdots, s_n)] \\
&= v_{(i)}^{(n)}(s_n)/[\pi_1 v_{(1)}^{(n)}(s_n) + \pi_2 v_{(2)}^{(n)}(s_n)] \\
&= g_{(i)}^{(n)}(s_n), \quad \text{say, where}
\end{aligned} \tag{3.2}
$$

$v_{(i)}^{(n)}(s_n) = \int_{\Theta_i} v_\theta^{(n)}(s_n)\Gamma_{(i)}(d\theta)$, $i = 1, 2$, and $f_\theta^{(n)}(s_1, s_2, \cdots, s_n) = v_\theta^{(n)}(s_n)h(s_1, \cdots, s_n)$, for some function h. Note that $\pi_1 g_{(1)}^{(j)} + \pi_2 g_{(2)}^{(j)} \equiv 1$. For the process defined by $P_{(i)}$ the conditional density of S_{n+1} given $S_1 = s_1, \cdots, S_n = s_n$, relative to $\nu^{((n+1)|n)}(\cdot|s_n)$ is given by $g_{(i)}^{((n+1)|n)}(s_{n+1}|s_n) = g_{(i)}^{(n+1)}(s_{n+1})/g_{(i)}^{(n)}(s_n)$. Note that Lemma 3.1 permitted us, to write the conditional measure $\nu^{((n+1)|n)}(\cdot|s_1, s_2, \cdots, s_n)$ as $\nu^{((n+1)|n)}(\cdot|s_n)$.

Consider the Markov process $S_1, S_2, \cdots, S_\infty$ governed by the measure ν. The functions $g_{(i)}^{(1)}(s_1), g_{(i)}^{(2)}(s_2), \cdots, i = 1, 2$, define two bounded martingales over this process since

$$\int g_{(i)}^{(n+1)}(s_{n+1})\nu^{((n+1)|n)}(ds_{n+1}|s_n) = g_{(i)}^{(n)}(s_n). \tag{3.3}$$

It will prove useful to define $V^{(n)}(\cdot)$ on the infinite product sample space by

$$V^{(n)}(\mathbf{s}) = \inf_{k \geqslant n} \min_{i=1,2} d_i \pi_i g_{(i)}^{(k)}(s_k). \tag{3.4}$$

Note that $\lim_{n\to\infty} V^{(n)}(\mathbf{S}) = \lim_{n\to\infty} \min_{i=1,2} d_i \pi_i g_{(i)}^{(n)}(S_n)$ with probability one (ν) by the martingale convergence theorem. As a consequence

$$\lim_{k\to\infty} E\{\min_{i=1,2} d_i \pi_i g_{(i)}^{(k)}(s_k) - V^{(k)}(\mathbf{S})|S_k = s_k\} = 0 \tag{3.5}$$

with probability one. Note also that the $g_{(i)}^{(\infty)} = dP_{(i)}/d\nu$ satisfies $g_{(i)}^{(\infty)}(\mathbf{s}) = g_{(i)}^{(\infty)}(s_\infty) = \lim g_{(i)}^{(n)}(s_n)$ with probability 1.

LEMMA 3.2. *The Bayes test for* $H_1 : \theta \in \Theta_1$ *vs* $H_2 : \theta \in \Theta_2$ *with respect to* $\Gamma(\cdot) = \pi_1\Gamma_1(\cdot) + \pi_2\Gamma_2(\cdot)$, *is the same as the Bayes test of* $H_1' : P(\cdot) = P_1(\cdot)$ *vs* $H_2' : P(\cdot) = P_2(\cdot)$ *with respect to the prior* (π_1, π_2).

PROOF. The proof is omitted.

Note that the combination of Lemmas 3.1 and 3.2 implies that for a given prior, $\Gamma = \pi_1\Gamma_1 + \pi_2\Gamma_2$, the Bayes test of Θ_1 vs Θ_2, is the same as the Bayes test of $\{g_{(1)}^{(j)}\}$ vs $\{g_{(2)}^{(j)}\}$, with prior (π_1, π_2). Now assume

(3.6) For each $n = 1, 2, \cdots, g_{(i)}^{(n)}(s_n)$ has strict monotone likelihood ratio (s.m.l.r.) in s_n and $i = 1, 2$.

(3.7) $P_{(i)}^{((n+1)|n)}(\cdot|s_n)$ is stochastically strictly increasing in s_n.

That is, for $s_n < s_n'$, $P_{(i)}^{((n+1)|n)}((c, \infty)|s_n) \leqslant P_{(i)}^{((n+1)|n)}((c, \infty)|s_n')$, with strict inequality unless $P_{(i)}^{((n+1)|n)}((c, \infty)|s_n) = 1$ or $P_{(i)}^{((n+1)|n)}((c, \infty)|s_n') = 0$.

In connection with (3.6) we note

LEMMA 3.3. *Suppose* $\Theta_1 \subset (-\infty, \theta_0]$, $\Theta_2 \subset (\theta_0, \infty)$ *and the family of distributions of* S_j *has* s.m.l.r. *in* θ. *Then the densities* $g_{(i)}^{(j)}(s_j)$ *have* s.m.l.r. *in* $i = 1, 2$.

PROOF. The lemma is easily verified using (3.2) in the definition of $g_{(i)}^{(j)}(s_j)$, and using the definition of s.m.l.r. □

LEMMA 3.4. *Under assumptions* (3.6) *and* (3.7) $\nu^{((n+1)|n)}(\cdot|s_n)$ *is stochastically strictly increasing in* s_n.

PROOF. Fix c and n and note

$$\nu^{(n+1|)}((c, \infty)|s_n) = P_1^{((n+1)|n)}((c, \infty)|s_n)\Pr\{i = 1|s_n\} + P_2^{((n+1)|n)}((c, \infty)|s_n)\Pr\{i = 2|s_n\}. \tag{3.8}$$

Use (3.6) and (3.7) in (3.8) to find

$$D = \nu((c, \infty)|s_n') - \nu((c, \infty)|s_n) \geqslant 0. \tag{3.9}$$

By delineating all possible cases it is verified that in fact $D > 0$. □

Now for given prior $\Gamma(\cdot)$ with a π_1 such that $0 < \pi_1 < 1$, let $\beta_n(s_n)$ be the minimum conditional expected risk given $S_n = s_n$ is observed at stage n and sampling continues at least to stage $(n + 1)$. In light of Lemmas 3.1 and 3.2 we have

$$\beta_n(s_n) = \int \min\big[\beta_{n+1}(s_{n+1}), C(n + 1) + d_2\pi_2 g_{(2)}^{(n+1)}(s_{n+1}), C(n + 1) + d_1\pi_1 g_{(1)}^{(n+1)}(s_{n+1})\big]\nu^{((n+1)|n)}(ds_{n+1}|s_n). \tag{3.10}$$

Similarly define $\beta_n^M(s_n)$, $n = 1, 2, \cdots, M - 1$, to be the minimum conditional expected risk given $S_n = s_n$ is observed at stage n and sampling continues at least to stage $(n + 1)$ but only procedures that sample at most M observations are considered. For $n = 1, 2, \cdots, M - 2$, determine $\beta_n^M(s_n)$ as in (3.10). For $n = M - 1$,

$$\beta_{M-1}^M(s_{M-1}) = C(M) + \int \min\big[d_2\pi_2 g_{(2)}^{(M)}(s_M), d_1\pi_1 g_{(1)}^{(M)}(s_M)\big]\nu(ds_M|s_{M-1}). \tag{3.11}$$

Let us study, for $n = 1, 2, \cdots, M - 1$,

$$D_n^M(s_n) = \beta_n^M(s_n) - C(n) - d_2\pi_2 g_{(2)}^{(n)}(s_n) \tag{3.12}$$

and

$$E_n^M(s_n) = \beta_n^M(s_n) - C(n) - d_1\pi_1 g_{(1)}^{(n)}(s_n). \tag{3.13}$$

LEMMA 3.5. *Let assumptions* (3.6) *and* (3.7) *be satisfied. Then* $D_n^M(s_n)$ *in* (3.12) ($E_n^M(s_n)$ *in* (3.13)) *is nonincreasing* (*nondecreasing*) *in* s_n *and is either strictly decreasing* (*strictly increasing*) *or nonincreasing and strictly positive* (*or nondecreasing and strictly negative*).

PROOF. We use backward induction. Assume that for $n = N + 1 \leqslant M - 1$, D_{N+1}^M is strictly decreasing in s_{N+1} or is nonincreasing and strictly positive. From (3.10) and (3.12) and (3.3) we have

$$(3.14)\quad \beta_N^M(s_N) - C(N) - d_2\pi_2 g_{(2)}^{(N)}(s_N)$$

$$= C_{N+1} + \int \min\Big[\beta_{N+1}^M(s_{N+1}) - C(N+1) - d_2\pi_2 g_{(2)}^{(N+1)}(s_{N+1}), 0,$$

$$d_1\pi_1 g_{(1)}^{(N+1)}(s_{N+1}) - d_2\pi_2 g_{(2)}^{(N+1)}(s_{N+1})\Big]$$

$$\times \nu^{((N+1)|N)}(ds_{N+1}|s_N).$$

Observe that since $g_{(i)}^{N+1}(s_{N+1})$ is s.m.l.r., it follows that $[d_1\pi_1 g_{(1)}^{(N+1)}(s_{N+1}) - d_2\pi_2 g_{(2)}^{(N+1)}(s_{N+1})]$ is strictly decreasing. The induction hypothesis implies that $\beta_{N+1}^M(s_{N+1}) - C(N+1) - d_2\pi_2 g_{(2)}^{(N+1)}(s_{N+1})$ is strictly decreasing or nonincreasing and strictly positive. Hence the integrand in the final expression for (3.14) is either strictly decreasing or is 0 for $s_{N+1} \leqslant s_{N+1}^*$ say, $s_{N+1}^* < \infty$, and then strictly decreasing. If $s_{N+1}^* = \infty$, then $D_N^M(s_N) = c_{N+1}$. Now apply Lemma 3.4 to conclude that (3.14) is strictly decreasing in s_N or is nonincreasing and strictly positive. To verify the conclusion for $n = M - 1$, use (3.11) instead of (3.10) in the above argument.

Similarly it can be shown that $E_n^M(s_n)$ is nondecreasing and is either strictly positive or strictly increasing. □

THEOREM 3.1. *Consider the one sided hypothesis testing problem under assumptions* (3.6) *and* (3.7). *Assume only procedures that sample at most M observations are considered. Then every Bayes test is monotone.*

PROOF. Let the prior distribution be denoted by $\Gamma(\cdot)$ with $0 < \pi_1 < 1$, and use Lemmas 3.1, 3.2, 3.3 to reduce the problem to testing $H_1' : \{g_{(1)}^{(J)}\}$ vs $H_2' : \{g_{(2)}^{(J)}\}$. It can be verified that every Bayes procedure at stage n is as follows: Continue sampling if for the observed s_N, (3.12) < 0, and (3.13) < 0. Stop and reject H_1' if (3.13) > 0 and $d_1\pi_1 g_{(1)}^{(n)}(s_n) < d_2\pi_2 g_{(2)}^{(n)}(s_n)$, while acceptance of H_1' occurs if $d_1\pi_1 g_{(1)}^{(n)}(s_n) > d_2\pi_2 g_{(2)}^{(n)}(s_n)$ and (3.12) > 0. Randomizations may occur if equalities occur. Lemma 3.5 and the fact that $g_{(i)}^{(n)}(s_n)$ has s.m.l.r. imply that every Bayes procedure is monotone according to Definition 2.1. □

The main theorem will follow

LEMMA 3.6. *For every* $n = 1, 2, \cdots, \beta_n^M(s_n)$ *converges to* $\beta_n(s_n)$ *with probability one as* $M \to \infty$.

PROOF. The proof follows from Theorems 4.4 and 4.7 of Chow, Robbins, and Siegmund [5].

THEOREM 3.2. *Consider the testing problem of* H_1' *vs* H_2' *under assumptions* (3.6) *and* (3.7). *Then every Bayes test is monotone.*

PROOF. For any prior a Bayes test exists. (See Brown and Doshi [4]). Any Bayes test must satisfy the relations concerning (3.12) and (3.13), with β_n^M replaced by β_n, given in the proof of Theorem 3.1. By Lemmas 3.5 and 3.6 we have $D_n(s_n)$ is nonincreasing and $E_n(s_n)$ is nondecreasing. By the s.m.l.r. property, $D_n(s_n)$, for $n = n^*$, is strictly decreasing or strictly positive unless $D_{n^*+1}(s_{n^*+1}) \equiv k_{n^*+1}$, where $k_{n^*+1} \leqslant -c_{n^*+1}$. Continuing by induction this can be true only if $D_j(s_j) \equiv k_j \leqslant -\sum_{i=n^*+1}^{j} c_i \to -(C(\infty) - C(n^*)) < 0$ as $j \to \infty$. However $D_j(s_j) \geqslant E\{V^{(J)}(\mathbf{S}) - \min_{i=1,2} d_i \pi_i g_{(i)}^{(J)}(s_j) | S_j = s_j\} \to 0$ with probability one as $j \to \infty$ by (3.5). This contradiction shows that the conclusion of Lemma 3.5 holds for D_n and similarly for E_n. The proof of Theorem 3.2 can be completed just like the proof of Theorem 3.1. □

We conclude this section with four remarks.

REMARK 3.1. The fact that the monotone tests form an essentially complete class depends strongly on the loss function being a linear combination of cost plus loss due to terminal decision. If the risk function is a pair of components, namely expected sample size and probability of error, the monotone tests are not necessarily an essentially complete class. In Brown, Cohen, and Strawderman [2] an example is given when P_θ is normal with unknown θ and the monotone tests are not essentially complete.

REMARK 3.2. Consider a two sided hypothesis problem in which the hypothesis, distribution, loss function, and prior distribution are all symmetric. Then the development of Section 3 implies that every such Bayes test is monotone according to Definition 2.2.

REMARK 3.3. The constant losses d_1 and d_2 can be replaced by special monotone functions of θ as in Sobel [8] and DeGroot [6]. The monotone procedures still form an essentially complete class.

REMARK 3.4. Generalized Bayes tests with respect to generalized priors such that the integrated risks are finite, are also monotone.

4. Examples. In this section we offer three examples. In all examples we verify assumptions (3.6) and (3.7). The first example considers the exponential family. The second is concerned with testing a multivariate normal mean by an orthogonally invariant test based on a statistic which has a noncentral chi-square distribution. The third example is appropriate for testing two normal variances.

For all succeeding examples assumption (3.6) is easily verified using the fact that in all cases the underlying densities $f_\theta^{(J)}(s_j)$ have s.m.l.r. in s_j and θ.

EXAMPLE 4.1. *Exponential family.* Let X_i, $i = 1, 2, \cdots$ be independent, identically distributed according to a distribution in the exponential family; i.e., the density is of the form $f(x|\theta) = c(\theta)e^{x\theta}d\mu(x)$. The problem is to test $H_1 : \theta \leqslant \theta_0$ against $H_2 : \theta > \theta_0$. Let $S_k = \sum_{i=1}^{k} X_i$. For this model Sobel [8] proved that for

every prior distribution, there exists a monotone procedure which is Bayes. We prove

THEOREM 4.1. *Every Bayes test for the problem described above is monotone.*

PROOF. The proof is omitted.

EXAMPLE 4.2. *Noncentral chi-square.* Let $\mathbf{X}(px1)$ be multivariate normal with mean vector $\boldsymbol{\theta}$ and covariance matrix I. We wish to use a sequential invariant test to test $H_1 : \boldsymbol{\theta} = 0$ vs $H_2 : \boldsymbol{\theta} \neq 0$. The problem is invariant under the orthogonal group. We let $\mathbf{X}_i$ be independent observations on the process and define $\mathbf{S}_k = \sum_{i=1}^{k} \mathbf{X}_i$. Also let $\Gamma(\theta)$ denote an invariant prior distribution. It is clear that the sequential Bayes invariant test depends on the invariantly sufficient transitive sequence $\chi_k^2 = \mathbf{S}_k'\mathbf{S}_k = \|\mathbf{S}_k\|^2$, whose distribution depends on $\|\boldsymbol{\theta}\|$. Our claim is that every such Bayes invariant test is monotone. To do this we apply Theorem 3.2 to the process $\|t_1\|, \|t_2\|, \cdots$, where $t_n = \mathbf{S}_n$. To verify the claim we need to verify (3.7) for this process since it is clear that (3.6) is true for this process. Let $s = \mathbf{S}_{n-1}$. Assumption (3.7) will follow from the fact that for every $0 < c < \infty$, $P\{\|t\| > c\|\|s\|\}$ is increasing in $\|s\|$. (Note that there is no dependence on θ in the above statement.) To show this we first prove

LEMMA 4.1. *The conditional density of $t = \mathbf{S}_n$ given $\mathbf{S}_{n-1} = s$ is exponential family of the form*

$$h(t|s) = c(\|s\|)q(\|t\|)e^{t's}. \tag{4.1}$$

PROOF. The proof is omitted.

By analogy to the relation of the noncentral chi-square to the multivariate normal, Lemma 4.1 implies that the conditional density $h(\|t\|\|\|s\|)$ is s.m.l.r. Therefore let $r(\|t\|) = 1$, if $\|t\| > c$, and $r(\|t\|) = 0$ if $0 \leqslant \|t\| \leqslant c$. It follows that

$$P\{\|t\| > c\|\|s\|\} = \int r(\|t\|)h(\|t\|\|\|s\|), \tag{4.2}$$

is increasing in $\|s\|$ since $r(\|t\|)$ is increasing and h is s.m.l.r. This completes the verification.

EXAMPLE 4.3. *Central F.* In this example we assume (X_i, Y_i) are independent observations on a bivariate normal distribution with mean vector $\mathbf{O}$ and covariance matrix with off diagonal elements o and diagonal elements σ_x^2 and σ_y^2. We test $H_1 : \sigma_x^2 \geqslant \sigma_y^2$ vs $H_2 : \sigma_y^2 > \sigma_x^2$. We seek the test which is Bayes among scale invariant tests. That is, limit consideration to tests which depend on the maximal invariant, transitive sequence (reduced by sufficiency), $S_n = \sum_{i=1}^{n} Y_i^2 / \sum_{i=1}^{n} X_i^2$. Let Γ be any invariant prior distribution on (σ_x^2, σ_y^2) and find the test which is Bayes among all those depending only on S_n. (This test really depends only on the prior distribution of the maximal invariant σ_x^2/σ_y^2.) The distribution of S_n is scaled F, where the scale factor is σ_x^2/σ_y^2. Hence assumption (3.6) is satisfied.

To verify assumption (3.7) note

$$S_{n-1} = S_n(V_1/V_2), \quad \text{where} \quad V_1 = \left[\Sigma_{i=1}^{n-1} X_i^2 / \Sigma_{i=1}^{n} X_i^2\right] \tag{4.3}$$

$$\text{and} \quad V_2 = \Sigma_{i=1}^{n-1} Y_i^2 / \Sigma_{i=1}^{n} Y_i^2 / \Sigma_{i=1}^{n} Y_i^2.$$

Since S_n, V_1, and V_2 are independent and V_1, V_2 are a pair of beta variables it is easy to derive the conditional distribution of S_{n-1} given S_n. This in turn makes it easy to derive the distribution of S_n given S_{n-1} and to verify that (3.7) holds.

Acknowledgment. The authors are grateful to Professor Robert H. Berk for many helpful discussions.

REFERENCES

[1] BROWN, L. D. (1977). Closure theorems for sequential-design processes. In *Statistical Decision Theory and Related Topics II* 57–91, Academic Press.

[2] BROWN, L. D., COHEN, ARTHUR, and STRAWDERMAN, W. E. (1979). On the admissibility or inadmissibility of fixed sample size tests in a sequential setting. *Ann. Statist.* 7 569–578.

[3] BROWN, L. D., COHEN, ARTHUR, and Strawderman, W. E. (1980). Complete classes for sequential tests of hypotheses. To appear in *Ann. Statist.*

[4] BROWN, L. D. and DOSHI, B. T. (1978). Existence of optimal policies in stochastic dynamic programming. To appear in *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.*

[5] CHOW, Y. S., ROBBINS, H., SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton, Mifflin Company.

[6] DEGROOT, M. H. (1961). The essential completeness of the class of generalized sequential probability ratio tests. *Ann. Math. Statist.* **32** 602–605.

[7] FERGUSON, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach*. Academic Press.

[8] SOBEL, M. (1952). An essentially complete class of decision functions for certain standard sequential problems. *Ann. Math. Statist.* **23** 319–337.

L. D. BROWN
DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850

ARTHUR COHEN
W. E. STRAWDERMAN
DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903