The Annals of Statistics
1983, Vol. 11, No. 2, 640–653

A SHARP NECESSARY CONDITION FOR ADMISSIBILITY OF SEQUENTIAL TESTS—NECESSARY AND SUFFICIENT CONDITIONS FOR ADMISSIBILITY OF SPRT'S[1]

BY L. D. BROWN,[2] ARTHUR COHEN,[3] E. SAMUEL-CAHN

Cornell University, Rutgers University, and Hebrew University

Consider the problem of sequentially testing the hypothesis that the mean of a normal distribution of known variance is less than or equal to a given value versus the alternative that it is greater than the given value. Impose the linear combination loss function under which the risk becomes a constant c, times the expected sample size, plus the probability of error. It is known that all admissible tests must be monotone—that is, they stop and accept if S_n, the sample sum at stage n, satisfies $S_n \leq a_n$; stop and reject if $S_n \geq b_n$. In this paper we show that any admissible test must in addition satisfy $b_n - a_n \leq 2\bar{b}(c)$. The bound $2\bar{b}(c)$ is sharp in the sense that the test with stopping bounds $a_n \equiv -\bar{b}(c)$, $b_n \equiv \bar{b}(c)$ is admissible.

As a consequence of the above necessary condition for admissibility of a sequential test, it is possible to characterize all sequential probability ratio tests (SPRT's) regarding admissibility. In other words necessary and sufficient conditions for the admissibility of an SPRT are given. Futhermore, an explicit numerical upper bound for $\bar{b}(c)$ is provided.

1. Introduction and summary. Consider the problem of sequentially testing the composite null hypothesis that the mean of a normal distribution of known variance is less than a given value versus the alternative that it is greater than the given value. Impose the simple linear combination loss function under which the risk becomes c (expected sample size) + (probability of error).

It has previously been shown that all admissible tests for this problem are monotone—that is they stop and accept (reject; resp.) if S_n, the sample sum at stage n, satisfies $S_n \leq a_n$ ($S_n \geq b_n$). We show in Section 4 that an admissible test must in addition satisfy $b_n - a_n \leq 2\bar{b}(c)$. This bound is sharp in the sense that the test with stopping bounds $b_n \equiv \bar{b}(c)$, $a_n \equiv -\bar{b}(c)$ is admissible.

The proof of this result involves an examination of the properties of sequential probability ratio tests (SPRT's). Every SPRT is of the above form with

$$a_n = a + n\bar{\mu}, \qquad b_n = b + n\bar{\mu} \tag{1.1}$$

for some a, b, $\bar{\mu} \in R$. It is well known that for every SPRT there is a value of c for which this test is unique Bayes, and hence admissible among those which take at least one observation. On the other hand for given c not all SPRT's are admissible. In fact, we will show that if one considers only procedures which take at least one observation then an SPRT is admissible if and only if $0 < b - a \leq 2\bar{b}(c)$. (If one considers also procedures which are allowed to stop without taking any observations—as we do for most of what follows—then admissibility of an SPRT requires also that $|a + b|$ be not too large. See following Proposition 3.2.)

Symmetric SPRT's, i.e., those for which $\bar{\mu} = 0$ and $a = -b$, play a central role. The

Received August 1982; revised December 1982.

[1] This paper was presented at the International Meeting on Analysis of Sample Survey Data and on Sequential Analysis, Israel Statistical Association, June 1982.

[2] Research supported by National Science Foundation Grant MCS 7824175.

[3] Research supported by National Science Foundation Grant MCS 8118564.

AMS 1980 *subject classifications.* Primary 62L10; secondary 62C15.

Key words and phrases. Sequential tests, SPRT, admissibility, Wiener process.

constant $\bar{b}(c)$ is defined by examining the class of symmetric SPRT's. This is done in Section 3. An additional result proved there is that the symmetric SPRT with $a = -\bar{b}(c)$, $b = \bar{b}(c)$ dominates all symmetric SPRT's with $b - a > 2\bar{b}(c)$. (We have no parallel for this domination result for general monotone procedures—that is if $b_n - a_n > 2\bar{b}(c)$ for some n we know that the procedure is inadmissible, but do not know, in general, how to construct a procedure which dominates it.)

Consideration in Section 5 of a testing problem involving the Wiener process with constant drift yields $(.427\cdots)c^{-1/2}$ as an explicit numerical upper bound for $\bar{b}(c)$. If c is not large this upper bound is close to $\bar{b}(c)$. We also relate $\bar{b}(c)$ to the probabilities of error of a symmetric SPRT in one example. This relation should be of interest to the practitioner. Some generalizations are described in Corollary 4.5 and in Section 6.

The linear combination risk function used here is a natural and convenient choice. Such a risk has been used by Wald (1947), Le Cam (1955), Lehmann (1959), Ferguson (1967), and others. One other possibility is of course to evaluate procedures according to a two-component risk whose components are expected sample size and probability of error. Interestingly the monotone procedures form a complete class for the problem with linear combination risk function but not for problems with componentwise risk function. See Brown, Cohen, Strawderman (1979). Furthermore Bayes tests derived using the linear combination risk turn out to be very appealing on other intuitive grounds. See Schwartz (1962) and Berger (1980). This provides additional justification for using such a risk function.

2. Notation and preliminaries. Let $X_1, \cdots$ be a sequence of independent identically distributed normal random variables with mean θ and known variance, σ^2. Without loss of generality set $\sigma^2 = 1$.

In its customary form the problem under consideration involves sequentially testing the null hypothesis that $\theta \leq \theta_0$ versus the alternative that $\theta > \theta_0$. Without loss of generality take $\theta_0 = 0$. The terminal decision is either $\tau = 1$ ("accept") or $\tau = 2$ ("reject"). The value of the loss function corresponding to stopping at time n and making decision τ is

$$L(\theta, (n, \tau)) = \begin{cases} cn & \text{if } \tau = 1 \text{ and } \theta \leq 0 \\ cn & \text{if } \tau = 2 \text{ and } \theta > 0 \\ cn + 1 & \text{otherwise.} \end{cases}$$

Assume throughout that $c < \frac{1}{2}$.

For technical reasons a slightly more symmetric formulation is desirable. Let $\Theta = \Theta_1 \cup \Theta_2$ with Θ_1 and Θ_2 topologically disjoint components of Θ described as follows:

$$\Theta_1 = (-\infty, 0) \cup 0_1, \qquad \Theta_2 = 0_2 \cup (0, \infty). \tag{2.1}$$

Θ_1 (respectively Θ_2) has the obvious topology in which $\Theta_1(\Theta_2)$ is isomorphic to $(-\infty, 0]$ ($[0, \infty)$) (thus, neighborhoods of the point 0_1 are of the form $(-\varepsilon, 0) \cup 0_1$, etc). If $\theta \in \Theta$ is the true parameter point then $X_1, X_2, \cdots$ are independent identically distributed normal random variables with variance $\sigma^2 = 1$ and mean

$$\mu = \theta \quad \text{if} \quad \theta \neq 0_1 \quad \text{or} \quad 0_2; \qquad \mu = 0 \quad \text{if} \quad \theta = 0_1 \quad \text{or} \quad 0_2. \tag{2.2}$$

The loss function is

$$\begin{aligned} L(\theta, (n, \tau)) &= cn \qquad \text{if} \quad \theta \in \Theta_\tau; \\ &= cn + 1 \qquad \text{if} \quad \theta \notin \Theta_\tau. \end{aligned} \tag{2.3}$$

A procedure is admissible in the customary formulation if and only if it is admissible in the above symmetric formulation. Hence the two formulations are indeed equivalent from a practical point of view. In the following only the symmetric formulation is used.

To every prior in the customary formulation, there corresponds in an obvious way a

unique prior in the symmetric formulation which yields the same Bayes procedure. However there are priors in the symmetric formulation which have no counterpart in the customary one. These are the priors which give mass to the point 0_2 and (to avoid trivialities) to $\Theta_1 - \{0_1\}$. Further aspects of the symmetric formulation are discussed later.

The risk of a test procedure δ is of course $r(\theta, \delta) = E_\theta(L(\theta, (N_\delta, T_\delta))$ where (N_δ, T_δ) denote the (random Markovian) stopping time and terminal decision rule corresponding to δ. By convention set $r(\theta, \delta) = \infty$ if $P_\theta(N_\delta < \infty) < 1$. We note that the formulation allows the value $N = 0$.

Let Γ be a prior distribution on Θ. Γ can be written as $\Gamma = \pi_1\Gamma_1 + \pi_2\Gamma_2$ where Γ_i is the conditional distribution on Θ_i and $\pi_i = \Gamma(\Theta_i)$. To avoid trivialities assume throughout that each $\pi_i > 0$. The statistic $S_n = \sum_{i=1}^n X_i$ is sufficient given the observations $X_1, \cdots, X_n$. Hence the posterior distribution on Θ given $X_1, \cdots, X_n$ is only a function of n and the values, s, of S_n. It will be written $\Gamma_n(\cdot \mid s) = \pi_{n1}(s)\Gamma_{n1}(\cdot \mid s) + \pi_{n2}(s)\Gamma_{n2}(\cdot \mid s)$ where Γ_{ni} and π_{ni} represent the indicated posterior conditional distributions and probabilities of Θ_i. Let $r(\Gamma, \delta) = \int r(\theta, \delta)\Gamma(d\theta)$.

Here are some important facts:

To any prior distribution, Γ, corresponds a Bayes procedure. Call it δ_Γ. (Thus, $r(\Gamma, \delta_\Gamma) = \inf_\delta r(\Gamma, \delta) \leq \frac{1}{2}$.) The collection of Bayes procedures is a complete class (i.e. includes all admissible procedures). See Brown, Cohen and Strawderman (1980) and Berk, Brown, and Cohen (1981b). Even for $c \leq \frac{1}{2}$ there are certain priors for which the Bayes procedure must stop before taking the first observation. For the remainder of priors there is a Bayes procedure which takes a first observation with probability one. Such Bayes procedures are essentially uniquely determined, non-randomized and monotone. (See Brown, Cohen, and Strawderman, 1979). Thus such a δ_Γ is essentially uniquely determined by a set of stopping boundaries a_n^Γ, b_n^Γ, $n = 1, 2, \cdots$, according to the rule—if $S_n \leq a_n^\Gamma (S_n \geq b_n^\Gamma)$ stop at time n and accept (reject). Otherwise sample X_{n+1}. (The discussion following (4.2) slightly extends the definition of a_n^Γ, b_n^Γ.)

Consider values $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$, not both 0. The log likelihood ratio based on $X_1, \cdots, X_n$ is

$$\begin{aligned} \lambda_n(s) &= \log \textstyle\prod_{i=1}^n \{f_{\theta_2}(x_i)/f_{\theta_1}(x_i)\} = \log\{f_{\theta_2}^{(n)}(s)/f_{\theta_1}^{(n)}(s)\} \\ &= (\theta_2 - \theta_1)\{s - n(\theta_1 + \theta_2)/2\}, \end{aligned} \tag{2.4}$$

where $f_\theta^{(n)}(s)$ denotes the density under θ of S_n (normal with mean $n\theta$, variance n). Consequently, a sequential probability ratio test (SPRT) is a test with stopping boundaries of the form

$$a_n = a + n\bar{\mu}, \qquad b_n = b + n\bar{\mu} \tag{2.5}$$

for some constants a, b, $\bar{\mu}$. A symmetric SPRT is one for which $a = -b$, $\bar{\mu} = 0$.

It is well known that the Bayes procedure for a prior concentrated on the two points θ_1, θ_2 either takes no observations or is an SPRT. In particular if $\theta_1 = -\theta_2 \neq 0$ and $\pi_1 = \pi_2 = \frac{1}{2}$ then the Bayes procedure is a symmetric SPRT if it takes any observations. Let $b(\theta_2, c)$ denote the upper stopping boundary (b_n) of such a symmetric procedure, as a function of θ_2, c. By convention, if the Bayes procedure takes no observations, set $b(\theta_2, c) = 0$.

The main result of Section 3 is an upper bound on $b(\cdot, c)$, plus a related domination result for symmetric SPRT's whose b value exceeds this bound. As a preliminary, we state the exact form of a Bayes procedure for any two point prior.

For any prior Γ, and any sample size $n \geq 1$, define the neutral boundary, η_n, through the equation

$$\int_{\Theta_1} f_\theta^{(n)}(\eta_n)\Gamma(d\theta) = \int_{\Theta_2} f_\theta^{(n)}(\eta_n)\Gamma(d\theta). \tag{2.6}$$

Thus η_n is the unique value for which $\pi_{n1}(s) = \pi_{n2}(s) = \frac{1}{2}$. Where necessary we write $\eta_n = \eta_n^\Gamma$ and $\pi_{ni} = \pi_{ni}^\Gamma$ to indicate dependence on the prior Γ.

PROPOSITION 2.1. *Let Γ be a prior distribution concentrated on $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$ with not both $\theta_1 = 0_1$ and $\theta_2 = 0_2$. Suppose the Bayes procedure for Γ takes at least one observation. Then it is an* SPRT *with boundaries*

$$a_n = -b + \eta_1 + (n-1)\bar{\mu} = -b + \eta_n, \qquad b_n = b + \eta_1 + (n-1)\bar{\mu} = b + \eta_n \tag{2.7}$$

where $b = b((\theta_2 - \theta_1)/2, c)$ and $\bar{\mu} = (\theta_1 + \theta_2)/2$.

PROOF. Let Γ^0 be the prior giving mass $\frac{1}{2}$ to each of the points $-(\theta_2 - \theta_1)/2$ and $+(\theta_2 - \theta_1)/2$. The joint distribution under θ_1 (resp., θ_2) of $S_2 - S_1, S_3 - S_2, \cdots$ is the same as the joint distribution under $-(\theta_2 - \theta_1)/2$, $((\theta_2 - \theta_1)/2)$ of $S_2 - S_1 + \bar{\mu}, S_3 - S_2 + \bar{\mu}, \cdots$. Furthermore

$$\pi_{1i}^\Gamma(\eta_1) = \tfrac{1}{2} = \pi_{1i}^{\Gamma^0}(0). \tag{2.8}$$

Hence, given that a first observation (S_1) has been taken, the problem with prior Γ and cumulative sums $S_1, S_2, S_3, \cdots$ is equivalent to the problem with prior Γ^0 and cumulative sums $S_1 - \eta_1, S_2 - \eta_1 - \bar{\mu}, S_3 - \eta_1 - 2\bar{\mu}, \cdots$. The Bayes procedure for the latter problem is a symmetric SPRT with boundaries $\pm b$. Consequently the Bayes procedure for the former problem has boundaries $\pm b + \eta_1 + (n-1)\bar{\mu}$. This verifies the first expressions in (2.7). It also follows from (2.8), and the sentence preceding it that

$$\eta_n = \eta_1 + (n-1)\bar{\mu}, \tag{2.9}$$

which verifies the remaining expressions in (2.7). □

The value of η_1 which for given θ_1, θ_2 lead to Bayes procedures taking at least one observation are described in Proposition 3.2. It will also be noted in Section 3 that $b(\mu, c) = 0$ (and consequently the Bayes procedure takes no observations) if and only if $\Phi(-\mu) \leq \frac{1}{2} - c$. This fact was previously exploited in Cohen and Samuel-Cahn (1982).

3. Necessary and sufficient condition-symmetric case. We first review some well known facts about SPRT's. A reference for the following statements is Ferguson (1967). (See Chapter 7 and especially Exercise 7.5.3) Fix $\mu > 0$ and $-\theta_1 = \theta_2 = \mu$. For now, let $\Gamma^{(\pi)}$ denote the prior giving mass π, $(1 - \pi)$ to θ_1, θ_2 respectively. Let $V(\pi) = V(\pi, \mu) = r(\Gamma^{(\pi)}, \delta_{\Gamma^{(\pi)}})$. Let

$$\pi_U = \text{g.l.b.}\{\pi : V(\pi) = 1 - \pi\}. \tag{3.1}$$

(For $\pi < \pi_U$, $V(\pi) < 1 - \pi$. Note also that $\pi_U \geq \frac{1}{2}$, by symmetry.) The Bayes test for $\pi = \frac{1}{2}$ is a symmetric SPRT with upper boundary

$$b = b(\mu, c) = \log(\pi_U/(1 - \pi_U))/2\mu. \tag{3.2}$$

The Bayes test stops without taking an observation (and has Bayes risk $= \frac{1}{2}$) if and only if $b(\mu, c)$, as defined in (3.2), equals 0. This occurs if and only if $\mu \leq \theta^* = \theta^*(c)$, where $\theta^*(c)$ is the unique value for which

$$\Phi(\theta^*) - \Phi(-\theta^*) = 2c \tag{3.3}$$

(Φ denotes the standard normal c.d.f.).

It is easily seen that $\lim_{\mu\to\infty} V(\pi, \mu) = c$, which implies $\pi_U \to 1 - c$. Hence $\lim_{\mu\to\infty} b(\mu, c) = 0$.

As previously noted, the risk of any Bayes test is continuous in $\theta \in \Theta$. This fact can be used to show that $V(\pi, \mu)$ is continuous as a function of μ, and hence that $b(\cdot, c)$ is continuous.

The above facts directly yield the following proposition.

PROPOSITION 3.1 *The bound* $\bar{b}(c) = \sup_{\mu>0} b(\mu, c) < \infty$. *A symmetric* SPRT *is a Bayes test for some symmetric two point prior if and only if* $b \leq \bar{b}(c)$.

Proposition 2.1 can be combined with Proposition 3.1 to provide a description of all SPRT's which can be Bayes for two point priors. In particular, an SPRT of the form (2.7) can be Bayes for a two point prior only if $b \leq \bar{b}(c)$. Here is a precise statement of this result.

PROPOSITION 3.2. *Given* c, *an* SPRT *with boundaries of the form* (2.7) *is admissible Bayes for some two point prior if and only if* $b \leq \bar{b}(c)$ *and* $|\eta_1 - \bar{\mu}| \leq b$.

PROOF. Consider an SPRT of the form (2.7). If $b \leq \bar{b}(c)$ then there exists a μ such that $b(\mu, c) = b$. Let $\theta_1 = \bar{\mu} - \mu$, $\theta_2 = \bar{\mu} + \mu$. Consider the two point prior giving probabilities π, $(1 - \pi)$ to θ_1, θ_2 such that η_1 is as specified in (2.7)—that is

$$(1 - \pi) f_{\bar{\mu}+\mu}(\eta_1) = \pi f_{\bar{\mu}-\mu}(\eta_1),$$

or, equivalently,

$$(1 - \pi) f_{\mu}(\eta_1 - \bar{\mu}) = \pi f_{-\mu}(\eta_1 - \bar{\mu}). \tag{3.4}$$

Among procedures which take at least one observation, this SPRT is Bayes. Consult, for example, Ferguson (1967, Chapter 7) and Propositions 2.1 and 3.1. This SPRT will be Bayes among all procedures if and only if a Bayes test (Bayes among all procedures) may take an observation. This occurs if and only if $1 - \pi_U \leq \pi \leq \pi_U$.

An algebraic reduction of (3.4) yields as an equation for π:

$$\pi/(1 - \pi) = e^{2\mu(\eta_1 - \bar{\mu})}. \tag{3.5}$$

Hence π is monotone in η_1 and $1 - \pi_U \leq \pi \leq \pi_U$ if and only if

$$|\eta_1 - \bar{\mu}| \leq (\tfrac{1}{2}\bar{\mu}) \log\{\pi_U/(1 - \pi_U)\}.$$

According to (3.2) this is equivalent to the desired condition $|\eta_1 - \bar{\mu}| \leq b$. This proves that this condition is necessary and sufficient for a test of the form (2.7) with $b \leq \bar{b}(c)$ to be Bayes for a two point prior.

To complete the proof of the proposition, we need only note that if an SPRT is of the form (2.7) with $b > \bar{b}(c)$ then as a consequence of Propositions 3.1 and 2.1 it cannot be Bayes for a two point prior. □

We conjecture that an SPRT can be Bayes only for a two point prior. If this conjecture is valid then Proposition 3.2 would show that the only Bayes SPRT's are those of the form (2.7) with $0 < b \leq \bar{b}(c)$ and $|\eta_1 - \bar{\mu}| \leq b$. Consequently all admissible SPRT's would be of this form. (Theorem 3.4 treats the symmetric case and proves such an admissibility result for this case via a different line of argument.)

Let $\delta(b)$ denote the symmetric SPRT with stopping boundaries $\pm b$. It has risk

$$r(\theta, \delta(b)) = r(\theta, b) = cE_{\theta}(N_b) + \beta_{|\theta|}(b) \tag{3.6}$$

where N_b denotes the stopping time of the SPRT and $\beta_{\theta}(b)$ denotes the probability of error in the terminal decision. ($\beta_{\theta}(b)$ is symmetric in θ, as is $E_{\theta}(N_b)$.)

LEMMA 3.3. *Let* $\theta' \in \Theta$. *Then* $r(\theta', b)$ *is strictly increasing in* b *for* $b > b(\theta', c)$.

PROOF. Suppose the lemma is false. Since $\lim_{b\to\infty} r(\theta', b) = \infty$ and $r(\theta', \cdot)$ is continuous

there exist two values $b_2 > b_1 > b(\theta', c)$ such that $r(\theta', b_1) = r(\theta', b_2)$. From (3.6) one gets

$$c\{E_{\theta'}(N_{b_2}) - E_{\theta'}(N_{b_1})\} = \beta_{\theta'}(b_1) - \beta_{\theta'}(b_2). \tag{3.7}$$

Note that $E_{\theta'}(N_{b_2}) > E_{\theta'}(N_{b_1})$ since $b_1 < b_2$.

From Lehmann (1959, Lemma 6, page 107) it follows that $b(\theta', \cdot)$ is continuous, strictly decreasing, and $\lim_{c\to 0} b(\theta', c) = \infty$. Hence there exists a $c_1 < c$ such that $b_1 = b(\theta', c_1)$. This implies, because of symmetry, that

$$c_1\{E_{\theta'}(N_{b_2}) - E_{\theta'}(N_{b_1})\} \geq \beta_{\theta'}(b_1) - \beta_{\theta'}(b_2). \tag{3.8}$$

(3.8) contradicts (3.7) since $c > c_1$. □

We can now prove the main result of this section.

THEOREM 3.4. *The symmetric* SPRT, $\delta(b)$, *is admissible if and only if* $0 \leq b \leq \bar{b}(c)$.

PROOF. The quantity $b = 0$ corresponds to the test which takes no observation and decides to accept or reject at random. Since any test taking an observation has $\lim_{\mu\to 0}\{r(\mu, \delta) + r(-\mu, \delta)\}/2 = cE_0(N_b) + \frac{1}{2} > \frac{1}{2} \equiv r(\mu, \delta(0))$, the test $\delta(0)$ is always admissible. If $0 < b \leq \bar{b}(c)$ then, as previously noted, $b = b(\mu, c)$ for some $\mu > 0$. Thus $\delta(b)$ is unique Bayes for a two point prior and hence admissible. If $b > \bar{b}(c) = \bar{b}$ then

$$r(\theta, \delta(\bar{b})) < r(\theta, \delta(b))$$

for all $\theta \in \Theta$ by Lemma 3.3. Hence, $\delta(b)$ is inadmissible. □

4. A complete class theorem. It was noted in Sections 2 and 3 that all Bayes procedures for two point priors have stopping boundaries satisfying $b_n - a_n \leq 2\bar{b}(c)$, $n = 1, \cdots$. This fact, together with further basic facts concerning concavity and uniform convergence of Bayes risk functions, will be used to prove the main result of this section: all admissible procedures (= all Bayes procedures) have stopping boundaries satisfying

$$b_n - a_n \leq 2\bar{b}(c) \qquad n = 1, \cdots. \tag{4.1}$$

This result considerably improves on Lemma 3.3 of Brown and Cohen (1981) which states only that for each prior Γ the Bayes procedure has stopping boundaries which satisfy

$$\sup\{b_n^\Gamma - a_n^\Gamma : n = 1, \cdots\} < \infty. \tag{4.2}$$

We need first to discuss the nature of the Bayes stopping boundaries a_n^Γ, b_n^Γ. For the trivial case where $\Gamma(\{0_1\} + \{0_2\}) = 1$ the Bayes procedure stops with probability one at time 0. We need say no more about this case, and hence assume in the following that Γ is not of this form.

Let $\rho_n^\Gamma(s, \delta)$ denote the conditional integrated risk given that $S_n = s$ when using a procedure δ which continues sampling at least to stage n, and when the prior is Γ. Let $\delta_\Gamma^{(n)}$ denote the Bayes procedure for prior Γ under the restriction that sampling continues at least to stage n.

It follows from Brown, Cohen, Strawderman (1979) that $\delta_\Gamma^{(n)}$ accepts (rejects) for $S_n \leq a_n^\Gamma (S_n \geq b_n^\Gamma)$ and continues otherwise, for some $-\infty \leq a_n^\Gamma \leq b_n^\Gamma \leq \infty$. Hence

$$\rho_n^\Gamma(s, \delta_\Gamma^{(n)}) \begin{cases} = cn + \pi_{n2}^\Gamma(s) & \text{if } s \leq a_n^\Gamma \\ = cn + \pi_{n1}^\Gamma(s) & \text{if } s \geq b_n^\Gamma \\ < cn + \min\{\pi_{n1}^\Gamma(s), \pi_{n2}^\Gamma(s)\} & \text{if } a_n^\Gamma < s < b_n^\Gamma. \end{cases} \tag{4.3}$$

The above uniquely defines a_n^Γ, b_n^Γ even for values of n under which the unrestricted Bayes rule δ_Γ will have stopped with probability one before sampling reaches stage n. Of course, if δ_Γ can reach stage n with positive probability, then its stopping boundaries are

a_n^{Γ}, b_n^{Γ}, the same as those of $\delta_{\Gamma}^{(n)}$.

LEMMA 4.1. *Let* $\Gamma^{(i)}$, $i = 1, \cdots, I$ *be a collection of prior distributions. Let* $\lambda_i \geq 0$, $i = 1, \cdots I$, *satisfy* $\sum_{i=1}^{I} \lambda_i = 1$. *Define* $\Gamma' = \sum_{i=1}^{I} \lambda_i \Gamma^{(i)}$. *Suppose* $b_n^{\Gamma^{(i)}} \leq \beta_n (a_n^{\Gamma^{(i)}} \geq \alpha_n$; *respectively*). *Then* $b_n^{\Gamma'} \leq \beta_n (a_n^{\Gamma'} \geq \alpha_n)$.

PROOF. For any procedure $\delta^{(n)}$ which continues at least to stage n

$$\begin{aligned} \rho_n^{\Gamma'}(s, \delta^{(n)}) &= \sum \lambda_i \rho_n^{\Gamma^{(i)}}(s, \delta^{(n)}) \geq \sum \lambda_i \rho_n^{\Gamma^{(i)}}(s, \delta_{\Gamma'}^{(n)}) \\ &= cn + \pi_{n2}^{\Gamma'}(s) \quad \text{if } s \leq \alpha_n \\ &= cn + \pi_{n1}^{\Gamma'}(s) \quad \text{if } s \geq \beta_n \end{aligned} \tag{4.4}$$

by (4.3). If $\delta^{(n)}$ has stopping boundaries a_n, b_n and $\beta_n \leq s < b_n$ or $a_n < s \leq \alpha_n$, then strict inequality holds in (4.4) for these values of s. It follows from (4.3) and (4.4) that $\delta^{(n)}$ cannot be Bayes for Γ' at stage n unless $b_n \leq \beta_n$ and $a_n \geq \alpha_n$. □

Give $\{\Gamma\}$ the usual weak* topology under which $\Gamma^{(i)} \to \Gamma'$ if

$$\int c(\theta)\Gamma^{(i)}(d\theta) \to \int c(\theta)\Gamma'(d\theta)$$

for all continuous bounded functions $c: \Theta \to R$. Let Γ' be a given prior. A consequence of (4.2), as noted in Brown and Cohen (1981), is that $N_{\delta_{\Gamma'}}$ is uniformly exponentially bounded. It easily follows that also $N_{\delta_{\Gamma'}^{(n)}}$ is uniformly exponentially bounded. That is, for some $\varepsilon < 1$ (depending on Γ') and for each n

$$P_\theta(N_{\delta_{\Gamma'}^{(n)}} > k) = 0(\varepsilon^k) \quad \text{uniformly for every} \quad \theta \in \Theta. \tag{4.5}$$

It easily follows from (4.5) that

$$r(\cdot, \delta_{\Gamma'}^{(n)}) \quad \text{is bounded and continuous on } \Theta. \tag{4.6}$$

An argument showing that (4.5) implies (4.6) in a similar context appears, for example, in Berk, Brown, and Cohen (1981b).

The continuity of $r(\cdot, \delta_{\Gamma'}^{(n)})$ is an important feature of the symmetric formulation adopted in Section 2. (In the customary formulation described at the beginning of that section it is usually the case that $r(\theta, \delta_{\Gamma'})$ is discontinuous in θ at $\theta = 0$. Hence the desired weak* convergence result (4.7) does not hold in that formulation.)

As a consequence of (4.6) if $\Gamma^{(i)} \to \Gamma'$ then

$$\int r(\theta, \delta_{\Gamma'}^{(n)})\Gamma^{(i)}(d\theta) \to \int r(\theta, \delta_{\Gamma'}^{(n)})\Gamma'(d\theta). \tag{4.7}$$

Let $M < \infty$. Define the M-truncated risk $r^{(M)}$ by

$$r^{(M)}(\theta, \delta) = cE_\theta(N_\delta \wedge M) + P_\theta(\theta \notin \Theta_{T_\delta} \quad \text{and} \quad N_\delta \leq M).$$

Observe that $r^{(M)}(\theta, \delta) \nearrow r(\theta, \delta)$ as $M \to \infty$. Hence

$$\lim_{M\to\infty} \int r^{(M)}(\theta, \delta)\Gamma(d\theta) = \int r(\theta, \delta)\Gamma(d\theta). \tag{4.8}$$

Also $r^{(M)}(\cdot, \delta) \leq cM + 1$ and is continuous.

LEMMA 4.2. *Suppose* $\Gamma^{(i)} \to \Gamma'$. *Then*

$$a_n^{\Gamma^{(i)}} \to a_n^{\Gamma'}, \; b_n^{\Gamma^{(i)}} \to b_n^{\Gamma'}. \tag{4.9}$$

PROOF. Let $\{i_j\}$ be a subsequence such that $\{a_n^{\Gamma(i_j)}\}$ and $\{b_n^{\Gamma(i_j)}\}$ converge in $[-\infty, \infty]$, say

$$a_n^{\Gamma(i_j)} \to a_n,\ b_n^{\Gamma(i_j)} \to b_n, \quad n = 1, 2, \cdots, \tag{4.10}$$

with $-\infty \le a_n \le b_n \le \infty$. Let δ_0 denote the procedure with stopping bounds (a_n, b_n). Fix $1 \le n_0 < \infty$. It follows easily from (4.10) that

$$r^{(M)}(\theta, \delta_{\Gamma^{(i_j)}}^{(n_0)}) \to r^{(M)}(\theta, \delta_0^{(n_0)}), \quad \theta \in \Theta, \tag{4.11}$$

where, as before, $\delta_0^{(n_0)}$ denotes the procedure which continues to stage n_0, and then uses the boundaries a_n, b_n, $n = n_0, n_0 + 1 \cdots$. Furthermore, the convergence in (4.11) is uniform over bounded subsets of Θ. It is actually uniform over all Θ unless $-\infty = a_n < a_n^{\Gamma(i_j)}$ or $b_n^{\Gamma(i_j)} < b_n = \infty$. Hence, for any fixed M,

$$\int r^{(M)}(\theta, \delta_{\Gamma^{(i_j)}}^{(n_0)})\Gamma^{(i_j)}(d\theta) \to \int r^{(M)}(\theta, \delta_0^{(n_0)})\Gamma'(d\theta), \tag{4.12}$$

since, also, $r^M(\cdot, \delta_0^{(n_0)})$ is bounded and continuous, and $\Gamma^{(i_j)} \to \Gamma'$. Note that

$$\int r(\theta, \delta_{\Gamma'}^{(n_0)})\Gamma^{(i_j)}(d\theta) \ge \int r(\theta, \delta_{\Gamma^{(i_j)}}^{(n_0)})\Gamma^{(i_j)}(d\theta) \ge \int r^{(M)}(\theta, \delta_{\Gamma^{(i_j)}}^{(n_0)})\Gamma^{(i_j)}(d\theta). \tag{4.13}$$

Combining (4.7), (4.12), (4.13) yields

$$\int r^{(M)}(\theta, \delta_0^{(n_0)})\Gamma'(d\theta) \le \int r(\theta, \delta_{\Gamma'}^{(n_0)})\Gamma'(d\theta).$$

It then follows from (4.8) that

$$\int r(\theta, \delta_0^{(n_0)})\Gamma'(d\theta) \le \int r(\theta, \delta_{\Gamma'}^{(n_0)})\Gamma'(d\theta). \tag{4.14}$$

However, as noted, $\delta_{\Gamma'}^{(n_0)}$ is the essentially unique Bayes procedure, under the restriction of taking at least n observations. Hence $\delta_0^{(n_0)} = \delta_{\Gamma'}^{(n_0)}$. This completes the proof of the lemma since $n_0 \ge 1$ was arbitrary and also the subsequences $\{a_n^{\Gamma(i_j)}\}\{b_n^{\Gamma(i_j)}\}$ were arbitrary convergent subsequences. □

Here is the main theorem.

THEOREM 4.3. *Every admissible procedure satisfies*

$$b_n - a_n \le 2\bar{b}(c). \tag{4.15}$$

PROOF. As noted in Section 2 we need only prove that any Bayes procedure, $\delta_{\Gamma'}$, which takes a first observation satisfies

$$b_n^{\Gamma'} - a_n^{\Gamma'} \le 2\bar{b}(c). \tag{4.16}$$

Let Γ' be any prior for which $\delta_{\Gamma'}$ takes a first observation. Fix $1 \le n_0 < \infty$. Let $\eta = \eta_{n_0}^{\Gamma'}$. We now verify the existence of a sequence of simple priors $\Gamma^{(i)}$ such that

$$\eta_{n_0}^{\Gamma^{(i)}} = \eta, \quad i = 1, \cdots \text{ and } \Gamma^{(i)} \to \Gamma'. \tag{4.17}$$

Suppose supp $\Gamma' \subset I_B = [-B, 0_1] \cup [0_2, B]$, for some $B < \infty$, where supp means support of the prior. The set

$$S_B(\eta) = \{\Gamma : \text{supp } \Gamma \subset I_B, \eta_{n_0}^{\Gamma} = \eta\}$$

is convex and is compact in the weak* topology. The external points of S_B are the

probability measures concentrated on (at most) two points. By the Krein-Millman Theorem (see e.g. Royden, 1968, page 207) Γ' is a limit of finite convex combinations of such two-point probability measures. This proves (4.17) for such a Γ'. If supp Γ' is arbitrary, it is easy to see that Γ' may be approximated by a sequence of distributions, $\{\Gamma_{B_j}: B_j \to \infty\}$, say, each of which has support in I_{B_j} and satisfies $\eta_{n_0}^{\Gamma_{B_j}} = \eta$. Again (4.17) follows. Furthermore it follows that $\Gamma^{(i)}$ can be written as

$$\Gamma^{(i)} = \textstyle\sum_{j=1}^{J_i} \lambda_j \Gamma^{(i,j)} \tag{4.18}$$

with $\lambda_j > 0$, $\Sigma \lambda_j = 1$, and each $\Gamma^{(i,j)}$ concentrated on (at most) two points and satisfying

$$\eta_{n_0}^{\Gamma^{(i,j)}} = \eta. \tag{4.19}$$

By Propositions 2.1 and 3.2,

$$\eta - \bar{b}(c) \le a_{n_0}^{\Gamma^{(i,j)}} \le b_{n_0}^{\Gamma^{(i,j)}} \le \eta + \bar{b}(c).$$

By Lemma 4.1,

$$\eta - \bar{b}(c) \le a_{n_0}^{\Gamma^{(i)}} \le b_{n_0}^{\Gamma^{(i)}} \le \eta + \bar{b}(c).$$

Lemma 4.2 then implies

$$\eta - \bar{b}(c) \le a_{n_0}^{\Gamma'} \le b_{n_0}^{\Gamma'} \le \eta + \bar{b}(c).$$

This proves (4.16), and hence also (4.15). □

Here are some consequences of the preceding results.

COROLLARY 4.4. *The stopping times of all admissible procedures are uniformly exponentially bounded* – i.e. *for some* $\varepsilon < 1$,

$$P_\theta(N_\delta \ge k) = O(\varepsilon^k) \tag{4.20}$$

uniformly for all $\theta \in \Theta$ *and all admissible* δ.

Consequently, the risk functions $r(\cdot, \delta)$ of admissible procedures are bounded and uniformly continuous on Θ uniformly for all admissible δ.

These statements follow easily from (4.15) by arguments like those in Berk, Brown, and Cohen (1981b). □

One may also discuss the sequential problem in which the only allowable procedures are those which take at least one observation. (Equivalently one may modify the loss function so that the first observation is supplied free of charge—so that all admissible procedures again take at least one observation.) This formulation was also discussed in Brown, Cohen, Strawderman (1980), and it was shown that the complete class consists of Bayes procedures, certain generalized Bayes procedures, and certain partial truncations (at stage $n = 1$) of these. All these procedures become—a posteriori given S_1—ordinary Bayes procedures from stage 2 onwards. Theorem 4.3 thus immediately yields (4.15) for $n \ge 2$. Some care in applying the formulation of Brown, Cohen, Strawderman (1980) for this situation in the proof of Theorem 4.3 also yields (4.15) for $n = 1$. In summary:

COROLLARY 4.5. *Conclusion* (4.15) *is also valid for the sequential problem in which all allowable procedures take at least one observation.*

For the following let

$$\theta_1^\Gamma = \sup\{\theta: \theta \in \Theta_1, \theta \in \text{supp } \Gamma\}, \quad \theta_2^\Gamma = \inf\{\theta: \theta \in \Theta_2, \theta \in \text{supp } \Gamma\}.$$

PROPOSITION 4.6. *Suppose* Γ *has* $\theta_1 < 0$ *or* $\theta_2 > 0$. *Then as* $n \to \infty$

$$\eta_n^\Gamma - a_n^\Gamma \to b((\theta_2^\Gamma - \theta_1^\Gamma)/2), \quad b_n^\Gamma - \eta_n^\Gamma \to b((\theta_2^\Gamma - \theta_1^\Gamma)/2). \tag{4.21}$$

PROOF. Let Γ^*denote the prior giving mass ½ to each of θ_1^Γ and θ_2^Γ. Then $\Gamma_n(\cdot \mid \eta_n^\Gamma) \to \Gamma^*$ as $n \to \infty$. The proposition then follows from Lemma 4.2 and Proposition 2.1. □

In the case of a "gap" prior as defined in Brown and Cohen (1981) (one in which $\theta_i^\Gamma = 0_i$, $i = 1, 2$, and 0_i is an isolated point of the support of Γ_i for $i = 1$ or 2) a modified version of (4.21) is valid. We omit the details.

Note that if $\theta_i^\Gamma = 0_i$ for $i = 1, 2$ and the prior is not a gap prior, then $a_n^\Gamma = b_n^\Gamma = \eta_n^\Gamma$ for all n sufficiently large, by Berk, Brown, and Cohen (1981a). Hence the stopping time for the Bayes rule is bounded.

5. Testing for the Wiener process. Let $W(t)$ be a Wiener process with variance one and shift θ per unit time. We consider testing $H_0: \theta \leq 0$ versus $H_1: \theta > 0$, where the observation costs c per unit time. This problem is interesting in its own right, and a solution to this problem sheds light on our original (discrete time) problem. We (again) consider the risk function to be the error probability plus c times the expected duration of observation.

The main difference between our original problem and the present problem is that in the original problem when testing $H_0: \theta = -\theta_0$ one would not take any observations if $|\theta_0| < \Phi^{-1}(½ + c)$ (see (3.3)), while in the continuous case one would always start observing the process.

Consider testing the simple hypothesis that the parameter is $-\theta$, versus the simple alternative that it is θ, by means of the test which continues as long as $|W(t)| < b$, stops and rejects (accepts) when $W(t) = b$ ($W(t) = -b$). Recall $\beta(\theta, b)$, $E_\theta N_b$, and $r(\theta, b)$ denote error probability, expected stopping time and risk, respectively. (By symmetry they are the same for θ and $-\theta$). Then as in (3.6)

$$r_c(\theta, b) = cE_\theta(N_b) + \beta(\theta, b). \tag{5.1}$$

LEMMA 5.1. *For the above problem*

$$\beta(\theta, b) = \{1 + \exp(2\theta b)\}^{-1}, \quad \text{and for} \quad \theta \neq 0, \quad E_\theta N_b = b\{1 - 2\beta(\theta, b)\}/\theta.$$

The content of this lemma is well known, and the proof is straightforward.

For fixed c and θ it follows from (5.1) and Lemma 5.1, through differentiation, that $r_c(\theta, b)$ is minimized by $b = b^*(\theta, c)$ which satisfies the equation

$$2b^*(\theta, c) = \theta/c - [\sinh\{2\theta b^*(\theta, c)\}]/\theta. \tag{5.2}$$

We are interested in $\max_\theta b^*(\theta, c) = \bar{b}^*(c)$. The value $\theta' = \theta(c)$ for which this maximum is attained satisfies $db^*(\theta, c)/d\theta|_{\theta=\theta'} = 0$. Differentiating both sides of (5.2) therefore yields, (after multiplying by θ'^2)

$$\theta'^2/c - \{2\theta' b^*(\theta', c)\}\cosh\{2\theta' b^*(\theta', c)\} + \sinh\{2\theta' b^*(\theta', c)\} = 0. \tag{5.3}$$

Noting from (5.2) that $\theta'^2/c = 2\theta' b^*(\theta', c) + \sinh\{2\theta' b^*(\theta', c)\}$ and substituting $2\theta' b^*(\theta', c) = u$ in (5.3) yields the simple equation

$$u\{1 - \cosh(u)\} + 2\sinh(u) = 0 \tag{5.4}$$

which has the unique solution $u = u^* = 2.3994\cdots$. It therefore follows that $\theta' b^*(\theta', c)$ is constant and does not depend on c. Substituting back in (5.2) and solving for θ' yields $\theta' = \theta(c)$ given by

$$\theta(c) = c^{1/2}\{u^* + \sinh(u^*)\}^{1/2} = c^{1/2}\, 2.8040\cdots, \tag{5.5}$$

and

$$\bar{b}^*(c) = c^{-1/2}[u^*\{u^* + 2\sinh(u^*)\}^{-1/2}/2] = c^{-1/2}0.427\cdots. \tag{5.6}$$

It turns out that $\beta(\theta(c), \bar{b}^*(c))$ and $r_c(\theta(c), \bar{b}^*(c))$ do not depend on c. We summarize the above in

THEOREM 5.2. *Let θ be the shift of a Wiener process with variance one per unit time. For testing $H_0: \theta = -\theta_0$ versus $H_1: \theta = \theta_0$ let $b(\theta_0)$ be the value of b which minimizes $r_c(b, \theta_0)$. The maximal value of $b(\theta)$ is obtained for $\theta_0 = \theta(c)$ given by* (5.5) *and satisfies the definition of $\bar{b}^*(c)$ in* (5.6), *where u^* is the solution of* (5.4). *The values of $\beta(\theta(c), \bar{b}^*(c))$ and $r_c(\theta(c), \bar{b}^*(c))$ are independent of c and are*

$$\beta(\theta(c), \bar{b}^*(c)) = [1 + \exp(u^*)]^{-1} = 0.08322\cdots \tag{5.7}$$

$$r_c(\theta(c), \bar{b}^*(c)) = u^*\{u^* + \sinh(u^*)\}^{-1}[\tfrac{1}{2} - \{1 + \exp(u^*)\}^{-1}] + [1 + \exp(u^*)]^{-1} = 0.2104\cdots. \tag{5.8}$$

The bound $\bar{b}^*(c)$ of (5.6) is an upper bound for the key value $\bar{b}(c)$ of the preceding sections, as shown by the following corollary.

COROLLARY 5.3. *For each fixed θ $b^*(\theta, c) \geq b(\theta, c)$. Hence, in particular*

$$\bar{b}(c) \leq \bar{b}^*(c) = c^{-1/2}(0.427). \tag{5.9}$$

PROOF. Fix any prior distribution Γ. In parallel with definition (4.3)

$$\rho_t^{*\Gamma}(s, \delta_\Gamma^{*(n)}) \begin{cases} = ct + 1 - \pi_{t1}^{*\Gamma}(s) & \text{if } s \leq a_t^* \\ = ct + 1 - \pi_{t2}^{*\Gamma}(s) & \text{if } s \geq b_t^* \\ < ct + \min\{1 - \pi_{t1}^{*\Gamma}(s), 1 - \pi_{t2}^{*\Gamma}(s)\} & \text{if } a_t^{*\Gamma} < s < b_t^* \end{cases} \tag{5.10}$$

where the quantities with asterisks are the Wiener problem parallels of those in Section 4. Note that for $t = n$ $\Gamma_n^*(\cdot \mid s) = \Gamma_n(\cdot \mid s)$ and so, also $\pi_{ni}^{*\Gamma}(s) = \pi_{ni}^{\Gamma}(s)$, $i = 1, 2$. For any procedure δ in the discrete problem of Section 2 there is a corresponding procedure $\hat{\delta}$ in the Wiener problem which stops only at discrete times $t = 0, 1, 2, \cdots$, and at these times agrees exactly with δ_Γ. Clearly

$$\rho_n^{*\Gamma}(s, \delta_\Gamma^{*(n)}) \leq \rho_n^{*\Gamma}(s, \rho_\Gamma^{(n)}) = \rho_n^{\Gamma}(s, \delta_\Gamma^{(n)}). \tag{5.11}$$

It follows from (4.3), (5.11) and the above that $a_n^{*\Gamma} \leq a_n^{\Gamma}$ and $b_n^{\Gamma} \leq b_n^{*\Gamma}$. This yields the assertions of the corollary. (A closer examination shows that strict inequality holds in (5.11), and hence also in the assertion of the corollary.) □

REMARK 5.1. If the process has variance σ^2 per unit time $\theta(c)$ of (5.5) and $\bar{b}^*(c)$ of (5.6) should be replaced by $\theta(c)/\sigma$ and $\bar{b}^*(c)\sigma$ respectively.

REMARK 5.2. Note that $\bar{b}^*(c) = O(c^{-1/2})$ and $E_{\theta(c)}N_{\bar{b}^*(c)} = O(c^{-1})$ as $c \to 0$. This should be contrasted with the usual result, viz. that for a fixed testing problem the stopping boundary of a Bayes rule is $O(-\log c)$ as $c \to 0$.

REMARK 5.3. It is clear that the value of $\bar{b}(c)$ of our previous sections, can be approximated very well by $\bar{b}^*(c)$, when c is small. Thus (5.6) is useful also for our original problem, and $\bar{b}(c)$ is also $O(c^{-1/2})$ as $c \to 0$.

REMARK 5.4. It should, however be noted that for large values of c, ($c < \frac{1}{2}$), $\bar{b}(c)$ and

$\bar{b}^*(c)$ are very different. (Large values of c are usually not interesting in applications.) Note that for $c > .2995$ $\theta(c) < \Phi^{-1}(\frac{1}{2} + c)$, and we have seen that the best SPRT for the original problem of testing $-\theta$ vs. θ for any $|\theta| < \Phi^{-1}(\frac{1}{2} + c) = \theta_0(c)$, is not to take any observations. (Thus an improved approximation of $\bar{b}(c)$ would be to replace $\bar{b}^*(c)$ for $c > .2995$ by $b^*(\theta_0(c), c)$ where $b^*(c)$ satisfies (5.2)).

Simple considerations yield that $\bar{b}(c) < \log\{(1 - c)/c\}/\{2\Phi^{-1}(\frac{1}{2} + c)\}$. Note that for $c > .244$, $\bar{b}^*(c)$ exceeds this upper bound, and hence one is better off using this upper bound on $\bar{b}(c)$. The quantity $\bar{b}(c)$ can also be bounded below, if so desired, by comparison with the maximal width to be used in a one-step-look-ahead procedure. The upper and lower bounds are not close unless c is very large.

In Table 5.1 we list a few values of $\bar{b}^*(c)$ with corresponding $\theta(c)$.

REMARK 5.5. It is interesting to relate $\bar{b}^*(c)$ and the error probabilities of an SPRT. A symmetric SPRT for $H_0: \theta = -\theta_0$ vs $H_1: \theta = \theta_0$ determined by error probabilities β has approximate boundaries determined by $b = \{\log(1 - \beta)/\beta\}/2\theta_0$. So, for example, when $c = .04$ and $\theta_0 = 1$, using Table 5.1, we see that if $\beta < 1/(1 + e^{4.28}) \cong .0137$ the SPRT is inadmissible.

Corresponding to Theorem 4.3 we can state a theorem for the Wiener process. We have:

THEOREM 5.4. *Let the continuation region of a sequential test for testing $\theta \leq 0$ versus $\theta > 0$ be given by $a_t^* < W(t) < b_t^*$, $b > 0$. A necessary condition for the sequential test to be admissible is that for (almost) all $t > 0$, $b_t^* - a_t^* \leq 2\bar{b}^*(c) = c^{-1/2}\, 0.8557$. This bound is the best obtainable in as much as there exists an admissible test with $b_t^* - a_t^* = 2\bar{b}^*(c)$ for all t, namely the test with $-a_t^* = b_t^* = \bar{b}^*(c)$.*

6. Generalizations. The techniques applied in Section 4 are not restricted to the normal problem studied there or to the loss function (2.3). We now state a much more general version of Theorem 4.3. This version can be proven by arguments which exactly parallel those in Section 4, and so the proof of the following version is omitted.

The important features of the formulation needed for such a generalization are that the structure of the problem be such that (i) the monotone procedures form a complete class (see Brown, Cohen, Strawderman, 1979), (ii) the Bayes procedures (or the truncated generalized Bayes procedures) form a complete class (see Brown, Cohen, Strawderman, 1980), and (iii) the risk functions of Bayes procedures are continuous (see Berk, Brown, and Cohen, 1981b). Here is a formulation which yields these features.

Let $X_1, X_2, \cdots$ be independent identically distributed random variables with density p_θ from a one dimensional regular exponential family having natural parameter space (ℓ, u).

TABLE 5.1

c	$\bar{b}^*(c)$	(c)
0.001	13.53	.089
0.002	9.57	.125
0.004	6.76	.177
0.006	5.52	.217
0.008	4.78	.251
0.010	4.28	.280
0.020	3.03	.397
0.040	2.14	.561
0.060	1.75	.687
0.080	1.51	.793
0.100	1.14	.887

Let $\theta_0 \in (\ell, u)$ and let

$$\Theta_1 = (\ell, \theta_0) \cup \theta_{01}, \quad \Theta_2 = (\theta_0, u) \cup \theta_{02} \tag{6.1}$$

by analogy with (2.1).

Let the loss function be

$$L(\theta, (n, \tau)) = \begin{cases} cn & \text{if } \theta \in \Theta_\tau, \\ cn + w(\theta) & \text{if } \theta \notin \Theta_\tau, \end{cases} \tag{6.2}$$

where $c > 0$ and $w(\cdot)$ is bounded, continuous, and non-increasing (non-decreasing) in θ for $\theta \in \Theta_1 (\theta \in \Theta_2)$. Note: Brown and Cohen (1981) does not explicitly apply to a loss function as general as (6.2). However, it can be generalized to such loss functions without too much trouble. Undoubtedly even more general loss functions can be used. In particular, the boundedness condition imposed on $w(\cdot)$ can certainly be considerably relaxed.

Let $A(\eta)$, $B(\eta)$ be defined by

$$A(\eta) = \inf\{a_1^\Gamma : \Gamma \in S^{(2)}(\eta)\}, \quad B(\eta) = \sup\{b_1^\Gamma : \Gamma \in S^{(2)}(\eta)\} \tag{6.3}$$

where

$$S^{(2)}(\eta) = \{\Gamma : \Gamma \text{ is a two point prior, } \eta_1^\Gamma = \eta\}. \tag{6.4}$$

(The quantities a_n^Γ, b_n^Γ, η_1^Γ, etc., are defined by analogy with quantities already described in Sections 2–4. In particular

$$\int_{\Theta_1} p_\theta(\eta_1^\Gamma) w(\theta) \Gamma(d\theta) = \int_{\Theta_2} p_\theta(\eta_1^\Gamma) w(\theta) \Gamma(d\theta). \tag{6.5}$$

The set $S^{(2)}(\eta)$ appears implicitly in the proof of Theorem 4.3.)

THEOREM 6.1. *In the above setting the procedure with stopping boundaries* (a_n, b_n) *can be Bayes for prior* Γ *only if*

$$\eta_n^\Gamma + A(\eta_n^\Gamma) \le a_n \le b_n \le \eta_n^\Gamma + B(\eta_n^\Gamma), \quad n = 1, \cdots. \tag{6.6}$$

Consequently a procedure with stopping boundaries (a_n, b_n) *is admissible only if there exists values* η_n, $n = 1, \cdots$, *such that*

$$\eta_n + A(\eta_n) \le a_n \le b_n \le \eta_n + B(\eta_n). \tag{6.7}$$

The proof of the above theorem needs to stray in one respect from the patterns of proof in Section 4. In general (if the measure dominating p_θ is not continuous) Bayes procedures may be randomized. Such procedures are not uniquely determined. Hence one cannot conclude that $a_n^{\Gamma^{(i)}} \to a_n^{\Gamma'}$ and $b_n^{\Gamma^{(i)}} \to b_n^{\Gamma'}$ in the generalization of Lemma 4.2. Instead one shows that if $\Gamma^{(i)} \to \Gamma'$ and $a_n^{\Gamma^{(i)}} \to a_n$ and $b_n^{\Gamma^{(i)}} \to b_n$ then (a_n, b_n) are Bayes stopping boundaries for the prior Γ'. This is revealed by the analog of (4.10). (One needs also to pay attention to the fact that randomization is possible at $S_n = a_n$ or b_n, etc. That possibility affects, but does not disrupt, the proof.)

Of course, in order for a result like Theorem 6.1 to be ultimately useful, one needs to study more carefully the nature of $A(\eta)$, as was done in Sections 3 and 5 for the normal problem.

It appears reasonable to presume that Theorem 6.1 could be extended also to certain non-i.i.d. sequences $X_1, \cdots$ such as those treated in Brown, Cohen, Strawderman (1979), and Brown, Cohen, Strawderman (1980). The requisite version of result (iii), mentioned above, has not yet been proved, but it is reasonable to conjecture its validity. In such a result the boundaries $A(\cdot)$ and $B(\cdot)$ would of course now be functions of both n and η_n.

REFERENCES

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

BERK, R. H., BROWN, L. D., and COHEN, A. (1981a). Bounded stopping times for a class of sequential Bayes tests. *Ann. Statist.* **9** 834–845.

BERK, R. H., BROWN, L. D., and COHEN, A. (1981b). Properties of Bayes sequential tests. *Ann. Statist.* **9** 678–682.

BROWN, L. D., COHEN, A., and STRAWDERMAN, W. E. (1979). Monotonicity of Bayes sequential tests. *Ann. Statist.* **7** 1222–1230.

BROWN, L. D., COHEN, A., and STRAWDERMAN, W. E. (1980). Complete classes for sequential tests of hypotheses. *Ann. Statist.* **8** 377–398.

BROWN, L. D. and COHEN, A. (1981). Inadmissibility of large classes of sequential tests. *Ann. Statist.* **9** 1239–1247.

COHEN, A. and SAMUEL-CAHN, E. (1982). Necessary and sufficient conditions for bounded stopping time of sequential Bayes tests in one parameter exponential families. *Sequential Analysis*, Communications in Statistics **1** 89–99.

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

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

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

ROYDEN, H. L. (1968). *Real Analysis.* MacMillan, New York.

Schwartz, G. (1962). Asymptotic shapes of Bayes sequential testing regions. *Ann. Math. Statist.* **33** 224–236.

WALD, A. (1947). *Sequential Analysis.* Wiley, New York.

L. D. BROWN
DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

ARTHUR COHEN
DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903

ESTER SAMUEL-CAHN
DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM, ISRAEL