The Annals of Statistics
1976, Vol. 4, No. 5, 891–911

ASYMPTOTIC EFFICIENCIES OF SEQUENTIAL TESTS[1]

BY ROBERT H. BERK

Rutgers University

Some concepts of relative and absolute efficiency for sequential tests are considered. These are sequential analogs of Hodges–Lehmann and Chernoff efficiencies. Using these criteria, several sequential tests for the mean of a normal distribution are evaluated. Among them are Wald's SPRT, Anderson's triangular boundary, Bayes and APO tests and a repeated significance test. Truncated versions of these tests are also considered. Asymptotic expressions for the (expected) stopping times and error rates are given.

1. Introduction and summary. In this paper we study sequential tests of H_1: $\theta \leqq 0$ vs. H_2: $\theta > 0$ for an unknown normal mean θ, the variance being assumed known. Starting with Wald's SPRT, many sequential tests for this problem have been put forth. We compare some of these tests asymptotically, letting the stopping times approach infinity in a suitable way. Inter alia, we suggest some measures of relative and absolute efficiency for sequential tests. These are in the spirit of Hodges–Lehmann and Chernoff efficiencies: one compares the rates at which error probabilities tend to zero. To an extent, the proposed measures allow for the added complication met in the sequential case, that the behavior of the sample size depends on the unknown parameter.

In our considerations, we actually deal with the related problem of sequentially testing hypotheses about the drift of a Wiener process. In so doing, two benefits accrue. First, we can dispense with considerations involving "overshoot." Such considerations are secondary to the main purpose of this study. By combining the techniques used here for continuous time with suitable methods of handling overshoot (see e.g. Berk (1973), Lorden (1970)), corresponding results for the discrete-time problem can be obtained. Of course, as has been noted by Anderson (1960) and others, results for continuous time sometimes give bounds for the corresponding discrete-time quantities. This applies especially to the (expected) stopping times. The second benefit arises from the fact that most of the procedures considered here have straight-line boundaries, for which exact formulas are available; see Anderson (1960). Thus we have a means of investigating numerically the qualitative conclusions

Received August 1974; revised February 1976.

[1] Prepared with the support of NSF Grant GP-36685X1. Work done while on sabbatical leave at the University of Tel Aviv.

AMS 1970 *subject classifications*. Primary 62L10, 62F05, 62F20; Secondary 60G40, 62E20.

Key words and phrases. Sequential test, asymptotic efficiency, SPRT, truncated sequential test, Anderson's boundaries, APO test, Bayes sequential test, repeated significance test, Wiener process.

suggested by the asymptotic analyses. We hope to report on such a numerical investigation separately.

We have chosen to consider the one-sided normal sequential testing problem, since next to the SPRT for simple hypotheses, it has probably received more attention than any other sequential testing problem. Our considerations do not provide us with a single-number summary of each test. Rather, we obtain an asymptotic profile that hopefully points to the good and bad features in each case.

In the next section, we consider some measures of asymptotic efficiency for sequential tests. Then, in subsequent sections, we treat various specific sequential tests, the analysis being directed toward computing the measures of efficiency discussed below. In the sequel, we adopt the following notation and conventions: $X(t)$, $t \geqq 0$ denotes an observable continuous-path Wiener process with unknown drift θ and variance 1 per unit time. For a sequential test of H_1 vs. H_2, let τ be the stopping time, $A =$ (accept H_1) and $R =$ (reject H_1). Let P_θ denote the probability distribution of $X(t)$, $t \geqq 0$ under θ and let $\varepsilon(\theta)$ be the corresponding probability of error. We write $\bar{\varepsilon}(\theta) = 1 - \varepsilon(\theta)$. Usually, $\varepsilon(0)$ is the level of the test. We consider only tests for which $E_\theta \tau < \infty$ for all θ. This entails $P_\theta A + P_\theta R = 1$. As is usual in asymptotic considerations, each test we consider is embedded in a corresponding family of tests, indexed by a parameter a, so that as $a \to \infty$, $\tau \to \infty$ w.p. 1. The embedding is done so that the family corresponding to a given test is consistent: For $\theta \neq 0$, $\varepsilon(\theta) \to 0$ as $a \to \infty$. We shall elliptically say that the test is consistent. In the sequel, $\lim_a$ etc. means as $a \to \infty$. Especially in the sequential case, the manner in which this embedding is done is, to a large extent, arbitrary. This is evident from the considerations in Sections 5 and 6 below.

2. Asymptotic efficiency. The problem in comparing sequential tests, asymptotically or otherwise, is that their operating characteristics cannot be made to match very well. This is in contrast to the nonsequential asymptotic theory giving rise to Pitman efficiency. There, two different tests can be made, asymptotically, to have the same power curve and the corresponding sample sizes required to do this are compared. Recently, Pitman efficiency has been extended to the sequential case; see e.g. Hall and Loynes (1974). However, as in the nonsequential theory, the comparison applies to two different sequences of statistics (normal scores and Wilcoxon, e.g.), using similar stopping boundaries for both sequences. For the tests we consider, the basic stochastic sequence is the same in every case; it is the stopping boundaries which differ.

A possible approach to comparing such sequential tests is suggested by the notions of Hodges–Lehmann and Chernoff efficiencies. The former [Hodges and Lehmann (1956)] is determined by fixing $\varepsilon(0)$ and considering, for $\theta \neq 0$, the rate at which $\varepsilon(\theta)$ tends to zero as the sample size increases. There is a theoretical upper bound to this rate, which often can be achieved. This provides then a notion of absolute as well as relative asymptotic efficiency. Chernoff's (1952) notion of efficiency was developed for testing one simple hypothesis against another, $-\theta$ vs. θ say, in our context. Chernoff's efficiency is obtained from the rate at which the total error $\gamma(\theta) = \varepsilon(\theta) + \varepsilon(-\theta)$ tends to zero as the sample size becomes infinite. There is also a theoretical upper bound available here, giving again a notion of absolute as well as relative efficiency. For both efficiencies, the basic quantity treated is a log error, divided by the sample size. We consider analogs for the sequential case, normalizing a log error by an appropriate expected sample size. We consider more than one expected sample size as the normalizing factor, which affords some insight into the behavior of the ASN function. There are theoretical upper bounds available for some of the quantities we consider. With one exception, the examples below show that the bounds can be attained by consistent tests.

In analogy with Hodges–Lehmann efficiency, we consider (limiting values of) the ratio $[-\log \varepsilon(\theta)]/E_\theta \tau$. This ratio in a sense standardizes the error rate at θ for the amount of sampling required to achieve that error rate. We also consider the ratio $[-\log \varepsilon(\theta)]/E_{-\theta} \tau$, which appears to have little intuitive appeal as a measure of efficiency at θ. However (unlike the first ratio), there is a theoretical upper bound for the limiting value of this second ratio. For symmetric procedures, this provides a corresponding upper bound for the first ratio. We also consider the corresponding ratio with $E_0 \tau$ in the denominator. This suggests itself as a sequential analog of standardizing competing tests by matching their type I error rates. (Presumably, one wants also to match values of $\varepsilon(0)$ here. However, for most of the tests we consider, the limiting behavior of these ratios does not depend on $\varepsilon(0)$, at least if it does not degenerate (to zero or one). For that reason, the selected value of $\varepsilon(0)$ is not indicated explicitly in our expressions.) For most of the procedures discussed in this paper, $E_0 \tau$ is, at least asymptotically, the maximum expected stopping time. (For symmetric procedures, this is true nonasymptotically as well. See Anderson (1960) and Hoeffding (1960).) This last ratio then gives a more conservative measure of efficiency than the first. (But, of course, it is the relative behavior of the same ratio for different procedures that is of interest.) The first ratio suffers from the theoretical defect that its limit can assume any nonnegative value, including $+\infty$. Thus it does not provide an interesting notion of absolute efficiency (except for symmetric procedures). By contrast, the ratio utilizing $E_0 \tau$ has a nontrivial limiting upper bound and thus provides a notion of absolute efficiency for all procedures. We also consider ratios with the total error $\gamma(\theta)$ replacing $\varepsilon(\theta)$. In this case, there are limiting upper bounds with both $E_\theta \tau$ and $E_0 \tau$ in the denominator. Since $\gamma(\cdot)$ is an even function, the ratio $[-\log \gamma(\theta)]/E_\theta \tau$ gives, inter alia, a comparison between $E_\theta \tau$ and $E_{-\theta} \tau$. In most cases, the analyses for $\theta > 0$ and $\theta < 0$ are similar, so that one is omitted.

It appears that no one ratio gives a completely adequate profile of a sequential test. For example, in the symmetric case, the SPRT is asymptotically optimal when judged by the $E_\theta\tau$-ratio. However, with $E_0\tau$ in the denominator, the corresponding limit is worst possible (zero). This reflects the fact that asymptotically, $E_0\tau$ for the SPRT is unduly large. There is, to be sure, a degree of arbitrariness in the choice of denominators for the above ratios, reflecting the fact that one usually cannot summarize the characteristics of a sequential test, even asymptotically, with a single expression. One might even contemplate a more complex profile, a two-parameter efficiency obtained by considering $[-\log \varepsilon(\theta)]/E_\nu\tau$. The analyses given below allow one to obtain the limit of this ratio for all values of θ and ν. To avoid over-burdening the discussion, we confine attention to the cases indicated above: $\nu = \theta$, 0 or $-\theta$. It is hoped that some numerical investigation will shed light on the extent to which these measures of efficiency are useful, quantitatively and/or qualitatively, for choosing among competing sequential tests.

In the remainder of this section we obtain some upper bounds for the limiting values of the above ratios. The following idea is used often: Let σ be a stopping time for $\{X(t): t \geqq 0\}$ which, for ν and θ in R, $P_\theta(\sigma < \infty) = 1 = P_\nu(\sigma < \infty)$. Let $P_\theta{}^\sigma$ denote the distribution under θ of the stopped process $\{X(t): 0 \leqq t \leqq \sigma\}$. Then $P_\theta{}^\sigma \equiv P_\nu{}^\sigma$ and

$$dP_\theta{}^\sigma/dP_\nu{}^\sigma = e^{(\theta-\nu)X(\sigma)-\frac{1}{2}(\theta^2-\nu^2)\sigma}[P_\nu] \tag{2.1}$$

(cf. Freedman (1971), Theorem 159). The following two lemmas give bounds for $[-\log \varepsilon(\theta)]/E_0\tau$.

2.1. LEMMA. *For a consistent test, provided* $\varepsilon(0)$ *does not degenerate,*

$$\limsup_a [-\log \varepsilon(\theta)]/E_0\tau \leqq \tfrac{1}{2}\theta^2\{\limsup_a E_0(\tau \mid A)/E_0\tau\}\,, \qquad \theta > 0\,. \tag{2.2}$$

PROOF. Using (2.1), for $\theta > 0$,

$$\begin{aligned} \varepsilon(\theta) &= P_\theta A = \textstyle\int_A e^{\theta X(\tau)-\frac{1}{2}\theta^2\tau}\, dP_0 \\ &= \bar\varepsilon(0)E_0(e^{\theta X(\tau)-\frac{1}{2}\theta^2\tau} \mid A) \\ &\geqq \bar\varepsilon(0)\exp\{E_0(\theta X(\tau) - \tfrac{1}{2}\theta^2\tau \mid A)\}\,. \end{aligned} \tag{2.3}$$

Using Wald's second lemma, we see that

$$|E_0(X(\tau) \mid A)| \leqq E_0|X(\tau)|/\bar\varepsilon(0) \leqq (E_0\tau)^{\frac{1}{2}}/\bar\varepsilon(0)\,. \tag{2.4}$$

Since for $\theta > 0$, $\lim_a \varepsilon(\theta) = 0$, it follows from (2.3) that $\lim_a E_0(\tau \mid A) = \infty$; hence also $\lim_a E_0\tau = \infty$. The lemma now follows from (2.3) and (2.4). □

Lemma 1 can presumably be extended generally to tests of simple null hypotheses. By contrast, the following bound utilizes the composite nature of H_1. (On the other hand, $\varepsilon(0)$ is completely unrestricted.)

2.2. LEMMA. *For consistent tests,*

$$\limsup_a [-\log \varepsilon(\theta)]/E_0\tau \leqq \tfrac{1}{2}\theta^2\{\liminf_\nu \limsup_a E_{-\nu}\tau/E_0\tau\}\,, \qquad \theta \neq 0\,; \tag{2.5}$$

the limit on ν *being taken as* $\nu \to 0$ *through nonzero values of the same sign as* θ.

PROOF. For $\nu\theta > 0$, Wald ((1947), equation A: 205) gave the following lower bound for $E_{-\nu}\tau$:

$$(2.6)\qquad E_{-\nu}\tau \geqq 2(\theta+\nu)^{-2}\left\{\bar{\varepsilon}(-\nu)\log\left[\frac{\bar{\varepsilon}(-\nu)}{\varepsilon(\theta)}\right] + \varepsilon(-\nu)\log\left[\frac{\varepsilon(-\nu)}{\bar{\varepsilon}(\theta)}\right]\right\}.$$

For $\nu\theta \neq 0$, both $\varepsilon(\theta)$ and $\varepsilon(-\nu)$ approach zero, so that

$$(2.7)\qquad E_{-\nu}\tau \geqq 2(\theta+\nu)^{-2}[-\log\varepsilon(\theta)]\{1+o(1)\}$$

and hence

$$(2.8)\qquad \limsup_a\,[-\log\varepsilon(\theta)]/E_0\tau \leqq \tfrac{1}{2}(\theta+\nu)^2\{\limsup_a E_{-\nu}\tau/E_0\tau\}.$$

The lemma follows on letting $\nu \to 0$ in (2.8). □

REMARK. It follows from (2.7) that

$$(2.9)\qquad \limsup_a\,[-\log\varepsilon(\theta)]/E_{-\nu}\tau \leqq \tfrac{1}{2}(\theta+\nu)^2, \qquad \nu\theta > 0.$$

The bound given in Lemma 2.1 is unsatisfactory in that it involves the intrinsic behavior of the stopping time. (The same may be said for Lemma 2.2.) A cruder bound, not suffering from this drawback, may be obtained as a consequence:

$$(2.10)\qquad \limsup_a\,[-\log\varepsilon(\theta)]/E_0\tau \leqq \tfrac{1}{2}\theta^2\liminf_a\bar{\varepsilon}(0), \qquad \theta > 0.$$

This follows from Lemma 2.1 on noting that $E_0\tau \geqq \bar{\varepsilon}(0)E_0(\tau\mid A)$. Unlike the previous bounds, this last bound apparently cannot be attained by a consistent test. (However, we have no proof of this assertion to offer.) It often happens that the LHS of (2.10) does not depend on $\varepsilon(0)$. Then, by taking the infimum of the RHS over all values of $\varepsilon(0)$, we see that

$$(2.11)\qquad \limsup_a\,[-\log\varepsilon(\theta)]/E_0\tau \leqq \tfrac{1}{2}\theta^2, \qquad \theta \neq 0.$$

For the tests considered in this paper, this last bound is violated only for the repeated significance tests considered in Section 9. Clearly (2.11) holds if the respective terms in curly brackets of Lemmas 2.1 or 2.2 do not exceed one. In particular, (2.11) holds for nonsequential tests. This also follows from a general theorem of Stein (cf. Bahadur (1971) or Chernoff (1956)) and can also be shown directly in the present case by considering the error-rates of the UMP nonsequential tests of H_1 vs. H_2. More generally, (2.11) holds for tests which are, in a sense, symmetric. This is brought out in the following.

2.3. PROPOSITION. *For consistent tests, if*

$$(2.12)\qquad \limsup_a E_0(\tau\mid A)/E_0\tau \leqq 1 \geqq \limsup_a E_0(\tau\mid R)/E_0\tau,$$

then (2.11) *holds. Moreover, the same conclusion follows if*

$$(2.13)\qquad \lim_a\,[-\log\varepsilon(\theta)]/E_0\tau = \lim_a\,[-\log\varepsilon(-\theta)]/E_0\tau.$$

Conversely, if (2.13) *holds and the common value is* $\frac{1}{2}\theta^2$, *then provided* $\varepsilon(0)$ *does not degenerate, necessarily* (2.12) *holds.*

REMARK. Since $E_0\tau = \bar{\varepsilon}(0)E_0(\tau \mid A) + \varepsilon(0)E_0(\tau \mid R)$, provided $\varepsilon(0)$ does not degenerate, (2.12) is equivalent to

$$\lim_a E_0(\tau \mid A)/E_0\tau = 1 = \lim_a E_0(\tau \mid R)/E_0\tau \quad \text{or} \tag{2.14}$$
$$\liminf_a E_0(\tau \mid A)/E_0\tau \geqq 1 \leqq \liminf_a E_0(\tau \mid R)/E_0\tau .$$

The proof of Proposition 2.3 is given after Lemma 2.4 below.

The following special case of (2.9) is of interest: On taking $\nu = \theta$, we obtain

$$\limsup_a [-\log \varepsilon(\theta)]/E_{-\theta}\tau \leqq 2\theta^2 , \qquad \theta \neq 0 . \tag{2.15}$$

This suggests that $E_{-\theta}\tau$ is a natural yardstick for measuring $\log \varepsilon(\theta)$. By contrast, there is no finite upper bound for $\limsup_a [-\log \varepsilon(\theta)]/E_\theta\tau$; see (4.4) and (9.8). This difficulty is cvercome by considering total error:

2.4. LEMMA. *For consistent tests,*

$$\limsup_a [-\log \gamma(\theta)]/E_0\tau \leqq \tfrac{1}{2}\theta^2 , \qquad \theta \neq 0 \tag{2.16}$$

and

$$\limsup_a [-\log \gamma(\theta)]/E_\theta\tau \leqq 2\theta^2 , \qquad \theta \neq 0 . \tag{2.17}$$

PROOF. Hoeffding ((1960), equation (1.4)) gives a lower bound for $E_\nu\tau$, which, on taking (in his notation) $\theta_0 = \nu$, $\theta_1 = -\theta_2 = \theta$, becomes here

$$E_\nu\tau \geqq \{[c^2 - \zeta \log \gamma(\theta)]^{\frac{1}{2}} - c\}^2/\zeta^2 , \tag{2.18}$$

where $\zeta = \frac{1}{2}(|\theta| + |\nu|)^2$ and c is a constant not depending on a. We obtain (2.16) on taking $\nu = 0$ and dividing across in (2.18) by $E_0\tau$. (Note that (2.18) entails $E_\nu\tau \to \infty$ for consistent tests.) Similarly, (2.17) follows on setting $\nu = \theta$. □

The bound in (2.17) can be attained (by an SPRT, e.g.) but not by a nonsequential test. This last fact follows from the nonsequential upper bound given by Chernoff (1952), which in this case is $\frac{1}{2}\theta^2$. Chernoff (1956) has noted the factor 4 between the sequential and nonsequential bounds in (2.17). An unsatisfactory feature of the ratio in (2.17) is that it loses track of the smaller error rate, since for $\theta \neq 0$, $[\log \gamma(\theta)]/(\log \max [\varepsilon(\theta), \varepsilon(-\theta)]) \to 1$.

In the following sections, we evaluate various tests of H_1 vs. H_2 according to the above criteria, beginning in Section 3 with the UMP nonsequential test. Appended here is a proof of Proposition 2.3.

PROOF OF PROPOSITION 2.3. The first assertion follows from Lemma 1; the second, from (2.16). Finally, if the common limit in (2.13) is $\frac{1}{2}\theta^2$, it follows from (2.3) and (2.4) that [the second relation in] (2.14) and hence also (2.12) hold. □

3. UMP nonsequential test. This test being exceedingly familiar, we simply list the formulas we require. We take $\tau \equiv a$ and $A = (X(a) \leqq \eta a^{\frac{1}{2}})$, where η is an appropriate fractile of the $N(0, 1)$ distribution. In the sequel, $\varphi(\cdot)$ and $\Phi(\cdot)$

denote the $N(0, 1)$ pdf and df, respectively. We have

$$\varepsilon(\theta) = \Phi(\eta \operatorname{sgn} \theta - |\theta| a^{\frac{1}{2}}) = e^{-\frac{1}{2}\theta^2 a + o(a)}, \qquad \theta \neq 0, \tag{3.1}$$

the last equality holding if $|\eta| = o(a^{\frac{1}{2}})$. Thus

$$\lim_a [-\log \varepsilon(\theta)]/a = \tfrac{1}{2}\theta^2, \qquad \theta \neq 0. \tag{3.2}$$

We then see that (2.11) holds as an equality (alternatively, that the bounds in (2.2) and (2.5) are attained); i.e., judged against its operating characteristics at zero, $[\varepsilon(0), E_0\tau]$, the UMP nonsequential test is asymptotically as powerful as any sequential test exhibiting the symmetry conditions of Proposition 2.3. We note too that the bound in (2.16) is attained but that those in (2.15) and (2.17) are not.

4. SPRT. For fixed positive λ and μ, we take

$$\tau = \inf\{t : X(t) \notin (-\lambda a, \mu a)\}$$

and $A = (X(\tau) = -\lambda a)$. This is an SPRT for testing $-\theta$ vs. θ with corresponding error levels $\varepsilon(-\theta)$ and $\varepsilon(\theta)$ and is LMP (locally most powerful) for testing H_1 vs. H_2 for the given $[\varepsilon(0), E_0\tau]$; see Berk (1975). The asymptotic behavior of τ is as follows.

4.1. THEOREM. *Under P_θ,* w.p. 1

$$\begin{aligned} \tau/a &\to \mu/\theta, \quad \theta > 0 \\ &\to \lambda/|\theta|, \quad \theta < 0 \end{aligned}$$

and $E_\theta\tau/a$ has the same limit. For $\theta = 0$, $\tau \sim a^2\tau(1)$, where $\tau(1)$ is the stopping time when $a = 1$. In particular, $E_0\tau = \lambda\mu a^2$.

PROOF. We consider $\theta > 0$. Let $X_* = \inf_{t \geq 0} X(t)$. Since $P_\theta(X(t) \to \infty) = 1$, $P_\theta(X_* > -\infty) = 1$. In turn, this entails $P_\theta(1_R \to 1) = 1$. We have

$$X(\tau) = \mu a 1_R - \lambda a 1_A. \tag{4.1}$$

We divide across by τ and let $a \to \infty$. Clearly $P_\theta(\lim_a \tau = \infty) = 1$, so that $P_\theta(\lim_a X(\tau)/\tau = \theta) = 1$. It follows that $P_\theta(\lim_a \mu a/\tau = \theta) = 1$. The corresponding convergence of $E_\theta\tau/a$ follows from Wald's first lemma: $E_\theta\tau = E_\theta X(\tau) = \bar{\varepsilon}(\theta)\mu a - \varepsilon(\theta)\lambda a$; note that $\varepsilon(0) \to 0$ as $a \to \infty$ (see (4.2) below). The result for $\theta = 0$ follows on replacing $X(t)$ by $aX(t/a^2)$, which has the same distribution when $\theta = 0$. □

Wald's formula for the error rate of an SPRT, which is exact in this case, is

$$\begin{aligned} \varepsilon(\theta) &= (e^{-2\theta\lambda a} - e^{-2\theta(\lambda+\mu)a})/(1 - e^{-2\theta(\lambda+\mu)a}) \\ &= e^{-2\theta\lambda a\{1+o(1)\}}, \qquad \theta > 0, \end{aligned} \tag{4.2}$$

so that

$$\lim_a [-\log \varepsilon(\theta)]/a = 2\lambda\theta, \qquad \theta \geqq 0. \tag{4.3}$$

It follows that the bound in (2.15) is attained, as is (2.17) if $\lambda \leqq \mu$. (Note that

$\varepsilon(0) = \lambda/(\lambda + \mu)$ so that $\lambda \leqq \mu \equiv \varepsilon(0) \leqq \frac{1}{2}$.) Also,

$$\lim_a [-\log \varepsilon(\theta)]/E_\theta \tau = 2\lambda\theta^2/\mu\,, \qquad \theta \geqq 0\,. \tag{4.4}$$

All ratios involving $E_0\tau$ tend to zero, as $E_\theta\tau = o(E_0\tau)$. Thus the SPRT is very inefficient when judged by $E_0\tau$, a phenomenon that is well known (cf. Bechhofer (1960) and Berk (1973)). This suggests that if sequential tests are matched by their operating characteristics at zero, the SPRT will not be as powerful as other sequential tests. In particular, the SPRT should not compare favorably in this way with the UMP nonsequential test. Limited numerical studies (not reported here) bear out this contention: The nonsequential test is more powerful, except for θ near zero (recall that the SPRT is LMP).

5. TPRT: Truncated SPRT. We consider here a truncated SPRT; the effect of the truncation being, inter alia, to make $E_0\tau$ grow at the same rate as $E_\theta\tau$, $\theta \neq 0$. For fixed positive λ and μ and fixed η, let

$$\sigma = \inf\{t : X(t) \notin (-\lambda a, \mu a)\}\,,$$

$\tau = \sigma \wedge a$ and $A = (X(\tau) < \eta a^{\frac{1}{2}})$. The subsequent analysis shows that $\varepsilon(0)$ converges to $P_0(X(a) \geqq \eta a^{\frac{1}{2}}) = 1 - \Phi(\eta)$. Hence η may be chosen to yield the desired level. (In particular, for the "natural" choice $\eta = 0$, $\varepsilon(0) \to \frac{1}{2}$, even if $\lambda \neq \mu$.) Following is the asymptotic behavior of τ.

5.1. THEOREM. *Under P_θ*, w.p. 1

$$\begin{aligned} \tau/a &\to \mu/\theta\,, \quad \theta > \mu\,,\\ &\to 1\,, \quad -\lambda \leqq \theta \leqq \mu\,,\\ &\to \lambda/|\theta|\,, \quad \theta < -\lambda\,. \end{aligned}$$

The corresponding limit for $E_\theta\tau/a$ is the same.

PROOF. For $\theta \neq 0$, the pointwise results follow from corresponding results for σ, which are given in Theorem 4.1. For $\theta = 0$, the pointwise result follows from (4.1) which, on dividing across by σ, entails $P_0(a/\sigma \to 0) = 1$. The corresponding results for $E_\theta\tau/a$ follow by dominated convergence, since $\tau/a \leqq 1$. □

The corresponding behavior of $\varepsilon(\theta)$ is given by

$$\begin{aligned} \lim_a [-\log \varepsilon(\theta)]/a &= \tfrac{1}{2}\theta^2\,, \quad 0 < \theta \leqq 4\lambda\,,\\ &= 2\lambda\theta\,, \quad \theta > 4\lambda\,. \end{aligned} \tag{5.1}$$

We establish (5.1) with the aid of Lemmas 2 and 3 below. We begin by noting that

$$\varepsilon(\theta) = P_\theta A(\sigma < a) + P_\theta A(\sigma \geqq a)\,, \qquad \theta > 0 \tag{5.2}$$

and that $A(\sigma \geqq a)$ is an event determined by $\{X(t) : 0 \leqq t \leqq a\}$. Then using

(2.1) with $\nu = 0$, we have

$$(5.3)\qquad \begin{aligned} \varepsilon(\theta) &= \int_{A(\sigma<a)} e^{\theta X(\sigma)-\frac{1}{2}\theta^2\sigma}\, dP_0 + \int_{A(\sigma\geqq a)} e^{\theta X(a)-\frac{1}{2}\theta^2 a}\, dP_0 \\ &= e^{-\lambda\theta a}\int_{A(\sigma<a)} e^{-\frac{1}{2}\theta^2\sigma}\, dP_0 + e^{-\frac{1}{2}\theta^2 a}\int_{A(\sigma\geqq a)} e^{\theta X(a)}\, dP_0\,, \qquad \theta > 0\,. \end{aligned}$$

We study the behavior of these last two integrals in the following lemmas.

5.2. LEMMA. *We have*

$$(5.4)\qquad \int_{A(\sigma<a)} e^{-\frac{1}{2}\theta^2\sigma}\, dP_0 \leqq e^{-\lambda\theta a}\,, \qquad \theta > 0\,.$$

Moreover,

$$(5.5)\qquad \lim_a e^{\lambda\theta a}\int_{A(\sigma<a)} e^{-\frac{1}{2}\theta^2\sigma}\, dP_0 = 1\,, \qquad \theta > \lambda$$

and

$$(5.6)\qquad \int_{A(\sigma<a)} e^{-\frac{1}{2}\lambda^2\sigma}\, dP_0 = e^{-\lambda^2 a + o(a)}\,.$$

PROOF. Take $\theta > 0$. Using (2.1) we have that $1 \geqq P_{-\theta}A(\sigma < a) = e^{\lambda\theta a}\int_{A(\sigma<a)} e^{-\sigma\theta^2/2}\, dP_0$, which gives (5.4). We note next that

$$(5.7)\qquad \begin{aligned} P_{-\theta}A(\sigma \geqq a) &\leqq P_{-\theta}(X(a) > -\lambda a) \\ &\leqq \tfrac{1}{2}\,, \qquad \theta = \lambda \\ &\leqq e^{-\frac{1}{2}a(\theta-\lambda)^2}\,, \qquad \theta > \lambda \text{ and } a \text{ large.} \end{aligned}$$

Moreover, since for $\theta > 0$, $\bar{\varepsilon}(-\theta) \to 1$, we see from the above that

$$(5.8)\qquad \begin{aligned} \liminf_a e^{\lambda\theta a}\int_{A(\sigma<a)} e^{-\frac{1}{2}\theta^2\sigma}\, dP_0 &\geqq \tfrac{1}{2}\,, \quad \theta = \lambda\,, \\ &\geqq 1\,, \quad \theta > \lambda\,, \end{aligned}$$

Together, (5.8) and (5.4) entail (5.5) and (5.6). □

5.3. LEMMA. $\int_{A(\sigma\geqq a)} e^{\theta X(a)}\, dP_0 = e^{o(a)}$, $\theta > 0$.

PROOF. We note that

$$(5.9)\qquad 0 \leqq \int_{(X(a)<\eta a^{\frac{1}{2}})} e^{\theta X(a)}\, dP_0 - \int_{A(\sigma\geqq a)} e^{\theta X(a)}\, dP_0 \leqq e^{\theta\eta a^{\frac{1}{2}}} P_0(\sigma < a)\,.$$

We estimate the RHS of (5.9). Suppose $\lambda \leqq \mu$. Then

$$(5.10)\qquad \begin{aligned} P_0(\sigma < a) &= P_0(\sup_{0<t<a} |X(t)| \geqq \lambda a) \\ &= P_0(\sup_{0<s<1} |X(s)| \geqq \lambda a^{\frac{1}{2}}) \leqq 4P_0(X(1) \geqq \lambda a^{\frac{1}{2}}) \\ &\leqq e^{-\frac{1}{2}\lambda^2 a} \qquad \text{for } a \text{ large.} \end{aligned}$$

It follows by an elementary calculation that

$$(5.11)\qquad \int_{(X(a)<\eta a^{\frac{1}{2}})} e^{\theta X(a)}\, dP_0 = e^{\frac{1}{2}\theta^2 a}[1 - \Phi(\theta a^{\frac{1}{2}} - \eta)] = e^{o(a)}\,.$$

The lemma follows from (5.9)—(5.11). □

Equation (5.1) now follows from (5.3) and Lemmas 5.2 and 5.3.

Combining Theorem 5.1 and (5.1), we find

$$(5.12)\qquad \begin{aligned} \lim_a [-\log \varepsilon(\theta)]/E_0\tau &= \tfrac{1}{2}\theta^2\,, \quad 0 \leqq \theta \leqq 4\lambda\,, \\ &= 2\lambda\theta\,, \quad 4\lambda < \theta\,; \end{aligned}$$

and (2.11) holds for $-4\mu \leqq \theta \leqq 4\lambda$. This result appears somewhat anomalous in that one would expect the TPRT to behave asymptotically like a nonsequential test for $-\lambda < \theta < \mu$. Presumably, the behavior of vanishingly small error probabilities is not easily assessed by the intuition. We have further

$$\begin{aligned} \lim_a [-\log \varepsilon(\theta)]/E_\theta \tau &= \tfrac{1}{2}\theta^2 , \quad 0 \leqq \theta \leqq \mu \wedge 4\lambda , \\ &= 2\lambda\theta^2/\mu , \quad \mu \vee 4\lambda \leqq \theta . \end{aligned} \tag{5.13}$$

If $4\lambda \neq \mu$, the limit for $\mu \wedge 4\lambda < \theta < \mu \vee 4\lambda$ depends on whether $4\lambda < \mu$ or vice versa. At any rate, for the symmetric case $\mu = \lambda$, the best possible rate is attained in (5.13) for $|\theta| \geqq 4\lambda$; cf. (2.15). Also,

$$\begin{aligned} \lim_a [-\log \varepsilon(\theta)]/E_{-\theta} \tau &= \tfrac{1}{2}\theta^2 , \quad 0 \leqq \theta \leqq \lambda , \\ &= \theta^3/2\lambda , \quad \lambda < \theta < 4\lambda , \\ &= 2\theta^2 , \quad 4\lambda \leqq \theta , \end{aligned} \tag{5.14}$$

so the bound in (2.15) is attained for $\theta \notin (-4\mu, 4\lambda)$. If $\lambda = \mu$, the RHS of (5.14) gives also $\lim_a [-\log \gamma(\theta)]/E_\theta \tau$ (for $\theta \geqq 0$), showing that the bound in (2.17) is also attained in that case, for $|\theta| \geqq 4\lambda$. By contrast, if $4\lambda \leqq \mu$,

$$\begin{aligned} \lim_a [-\log \gamma(\theta)]/E_\theta \tau &= 2\theta^2 , \quad \theta \leqq -4\lambda\mu , \\ &= |\theta|^3/2\lambda , \quad -4\mu < \theta < -\lambda , \\ &= \tfrac{1}{2}\theta^2 , \quad -\lambda \leqq \theta \leqq 4\lambda , \\ &= 2\lambda\theta , \quad 4\lambda < \theta < \mu , \\ &= 2\lambda\theta^2/\mu , \quad \mu \leqq \theta , \end{aligned} \tag{5.15}$$

so that the bound in (2.17) is attained only for $\theta \leqq -4\mu$.

6. TPRT II. It may be objected that in Section 5, the truncation point is arbitrarily allowed to increase at the same rate as the horizontal boundaries. (This reflects the presumably practical objective of not wanting to have unduly large stopping times for certain parameter values.) Embeddings for which the truncation point grows at a different rate are also possible. Presumably, if the truncation point increases more slowly, one obtains, asymptotically, a nonsequential test. If it grows more quickly, one obtains an SPRT. We show here that this last statement is essentially correct. The particular rates are chosen to facilitate comparison with the repeated significance test treated in Section 9. We let

$$\sigma = \inf\{t : X(t) \notin (-\lambda a, \mu a)\} ,$$

$\tau = \sigma \wedge a^2$ and $A = (X(\tau) < 0)$. The asymptotic behavior of τ is given by

6.1. THEOREM. *Under P_θ, w.p. 1*

$$\begin{aligned} \tau/a &\to \mu/\theta , \quad \theta > 0 , \\ &\to \lambda/|\theta| , \quad \theta < 0 . \end{aligned}$$

$E_\theta \tau/a$ has the same limit. When $\theta = 0$, $\tau \sim a^2\tau(1)$. In particular, $E_0\tau = a^2E_0\tau(1)$.

We remark that $E_0\tau(1) < 1$, since $\tau(1) \leqq 1$ and $P_0(\tau(1) < 1) > 0$.

PROOF. For $\theta \neq 0$, the pointwise results follow from corresponding results for σ, which are given in Theorem 4.1. To get the expectation, we have $E_\theta\tau/a \leqq E_\theta\sigma/a$; as the latter converges to the desired limit, we have half of the result. To get the other half, apply Fatou's lemma to $E_\theta\tau/a$. The result for $\theta = 0$ follows as in Theorem 4.1. □

6.2. THEOREM. *For $\theta > 0$, $\lim_a[-\log\varepsilon(\theta)]/a = 2\lambda\theta$.*

PROOF. The decomposition (5.3) becomes here

$$\varepsilon(\theta) = e^{-\lambda\theta a}\int_{A(\sigma<a^2)} e^{-\frac{1}{2}\theta^2\sigma}\,dP_0 + e^{-\frac{1}{2}\theta^2a^2}\int_{A(\sigma\geqq a^2)} e^{\theta X(a^2)}\,dP_0\,. \tag{6.1}$$

In this case, the second term on the RHS is negligible compared with the first: We again note that $\bar\varepsilon(-\theta)\to 1$ and since (for a large), $P_{-\theta}A(\sigma \geqq a^2) \leqq P_{-\theta}(X(a^2) > -\lambda a) \leqq \exp\{-\frac{1}{2}(\theta a - \lambda)^2\}\to 0$, $P_{-\theta}A(\sigma < a^2)\to 1$. That is,

$$\lim_a e^{\lambda\theta a}\int_{A(\sigma<a^2)} e^{-\frac{1}{2}\theta^2\sigma}\,dP_0 = 1\,, \qquad \theta > 0 \tag{6.2}$$

(cf. (5.5)). Since $X(a^2) < 0$ on $A(\sigma \geqq a^2)$, the second term on the RHS of (6.1) is manifestly of smaller order than the first. The theorem now follows from (6.1) and (6.2). □

It follows that all the ratios considered in Section 4 have the same limits here. The only effect of truncation, asymptotically, is to reduce $E_0\tau$, but it is still of larger order than $E_\theta\tau$, $\theta \neq 0$.

7. Anderson's triangular boundary. For fixed positive λ and μ, we let

$$\tau = \inf\{t : X(t) \notin (\lambda(t-a), \mu(a-t))\}$$

and $A = (X(\tau) < 0)$. Note that $\tau \leqq a$. Such boundaries were considered by Anderson (1960), who gives exact formulas for $\varepsilon(\theta)$ and $E_\theta\tau$. Anderson's expressions appear somewhat unwieldy and we obtain asymptotic expressions using other methods.

7.1. THEOREM. *Under P_θ, w.p. 1*

$$\begin{aligned}\tau/a &\to \mu/(\mu+\theta)\,, \quad \theta \geqq 0\,,\\ &\to \lambda/(\lambda+|\theta|)\,, \quad \theta < 0\,.\end{aligned}$$

The corresponding limit for $E_\theta\tau/a$ is the same.

PROOF. We consider $\theta > 0$, the argument for $\theta < 0$ being similar. We have

$$X(\tau) = \mu(a-\tau)1_R + \lambda(\tau - a)1_A\,. \tag{7.1}$$

As in Theorem 4.1, $P_\theta(1_R \to 1) = 1$. The pointwise result follows on dividing across in (7.1) by τ and letting $a\to\infty$, noting that $X(\tau)/\tau\to\theta[P_\theta]$. For the case $\theta = 0$, we note first that $\tau/a \leqq 1$, hence $P_0(\limsup_a \tau/a \leqq 1) = 1$. Suppose $\lambda \leqq \mu$, so that $|X(\tau)| \geqq \lambda(a-\tau)$. On dividing by τ and noting that $X(\tau)/\tau \to 0[P_0]$, we find also that $\limsup_a a/\tau \leqq 1[P_0]$ or $\liminf_a \tau/a \geqq 1\,[P_0]$. The results for the expectations follow by dominated convergence. □

The corresponding behavior of the error rate is given by

$$[-\log \varepsilon(\theta)]/a \to \tfrac{1}{2}\theta^2, \quad 0 \leqq \theta \leqq 2\lambda, \qquad \to 2\lambda(\theta - \lambda), \quad 2\lambda < \theta. \tag{7.2}$$

This follows from Lemmas 7.2 and 7.4 below and the following representation: Using (2.1) and (7.1), we see that

$$P_\theta A = \textstyle\int_A e^{\theta X(\tau) - \frac{1}{2}\theta^2\tau}\, dP_0 = e^{-\lambda\theta a} \int_A e^{(\theta\lambda - \frac{1}{2}\theta^2)\tau}\, dP_0, \qquad \text{all } \theta. \tag{7.3}$$

Of course for $\theta > 0$, $P_\theta A = \varepsilon(\theta)$. The following lemmas give the behavior of the last integral in (7.3).

7.2. Lemma. *For $\rho \geqq 0$, $\lim_a \{\int_A e^{\rho\tau}\, dP_0\}^{1/a} = e^\rho$.*

To prove this, we need the following:

7.3. Lemma. *Let Y_a be a sequence of random variables so that for some constant y, $Y_a/a \to_P y$ as $a \to \infty$. Let B_a be any sequence of events for which* $\liminf_a PB_a >$ 0. *Then,*

$$\textstyle\int_{B_a} \exp\{Y_a\}\, dP \geqq e^{ya + o(a)}.$$

Proof. We drop the subscript a for convenience. Choose $\delta > 0$. $\int_B e^Y\, dP \geqq \int_{B(Y > ay - a\delta)} dP \geqq e^{ay - a\delta} PB(Y > ay - a\delta)$. Since $PB = e^{O(1)}$ and $P(Y > ay - a\delta) \to 1$, also $PB(Y > ay - a\delta) = e^{O(1)}$, hence $\int_B e^Y\, dP \geqq e^{ay - a\delta + O(1)}$. Since δ is arbitrary, the lemma follows. □

Proof of Lemma 7.2. One half of the result is immediate, since $\int_A e^{\rho\tau}\, dP_0 \leqq e^{\rho a}$. For the other half, we use (2.1) to write

$$\begin{aligned}\textstyle\int_A e^{\rho\tau}\, dP_0 &= \textstyle\int_A e^{\rho\tau} e^{\nu X(\tau) + \frac{1}{2}\nu^2\tau}\, dP_{-\nu} \\ &= e^{-a\lambda\nu} \textstyle\int_A e^{(\rho + \lambda\nu + \frac{1}{2}\nu^2)\tau}\, dP_{-\nu}.\end{aligned}$$

Since $P_{-\nu}(\tau/a \to \lambda/(\lambda + \nu)) = 1$ and $P_{-\nu}A \to 1$ for $\nu > 0$, it follows from Lemma 7.3 that

$$\liminf_a (\textstyle\int_A e^{\rho\tau}\, dP_0)^{1/a} \geqq \exp\{-\lambda\nu + \lambda(\rho + \lambda\nu + \frac{1}{2}\nu^2)/(\lambda + \nu)\}, \qquad \text{all } \nu > 0,$$

hence that

$$\liminf_a (\textstyle\int_A e^{\rho\tau}\, dP_0)^{1/a} \geqq e^\rho. \qquad \square$$

Remark. The preceding proof could presumably be shortened by applying Lemma 7.3 directly to $\int_A e^{\rho\tau}\, dP_0$. To do this, one must demonstrate that $P_0 A$ remains bounded away from zero. In fact, it may be shown (using Anderson's (1960) formulas, e.g.) that for all positive λ and μ, $P_0 A \to \frac{1}{2}$. To achieve a test whose level is asymptotically less than $\frac{1}{2}$, one can modify A (as done for the TPRT in Section 5) letting, for example, $A = (X(\tau) < \eta a^{\frac{1}{2}})$ for some $\eta > 0$.

7.4. Lemma. *For $\theta > 2\lambda$,*

$$\lim_a e^{\lambda(\theta - 2\lambda)a} \textstyle\int_A e^{(\theta\lambda - \frac{1}{2}\theta^2)\tau}\, dP_0 = 1.$$

Proof. Choose $\theta > 2\lambda$. We note that the quadratic $\theta\lambda - \frac{1}{2}\theta^2 = \frac{1}{2}\theta(2\lambda - \theta)$

is unchanged if θ is replaced by $2\lambda - \theta$. From (7.3) we have that

$$P_{2\lambda-\theta}A = e^{\lambda(\theta-2\lambda)a} \int_A e^{(\theta\lambda-\frac{1}{2}\theta^2)\tau}\, dP_0 . \tag{7.4}$$

The proof is concluded by noting that $P_{2\lambda-\theta}\, A \to 1$ for $\theta > 2\lambda$. □

REMARK. Combining (7.3) and (7.4), we obtain

$$P_{2\lambda-\theta}A = e^{2\lambda(\theta-\lambda)a}P_\theta A , \tag{7.5}$$

a relation that has been previously noted by Lawing and David (1966). In particular $\varepsilon(2\lambda) = \frac{1}{2}e^{-2\lambda^2 a}$ in the symmetric case.

From Theorem 7.1 and (7.2) we find that

$$\begin{aligned} \lim_a [-\log \varepsilon(\theta)]/E_0\tau &= 2\mu(|\theta| - \mu)\,, \quad \theta < 2\mu\,, \\ &= \tfrac{1}{2}\theta^2\,, \quad -2\mu \leqq \theta \leqq 2\lambda\,, \\ &= 2\lambda(\theta - \lambda)\,, \quad 2\lambda < \theta\,, \end{aligned} \tag{7.6}$$

hence that the bound in (2.11) is attained for $-2\mu \leqq \theta \leqq 2\lambda$. This result seems surprising at first, since the test in no way appears to resemble a nonsequential test. However, it accords with the findings of Lai (1973) that the boundaries minimizing $E_0\tau$ for given $\varepsilon(\theta) = \varepsilon(-\theta)$ and fixed θ are asymptotically (as $\varepsilon(\theta) \to 0$) triangular.

Other limiting ratios can be obtained similarly. For example, we have

$$\begin{aligned} \lim_a [-\log \varepsilon(\theta)]/E_\theta\tau &= 2\mu(|\theta| + \mu)(|\theta| - \lambda)/\lambda\,, \quad \theta < 2\mu\,, \\ &= \theta^2(|\theta| + \lambda)/2\lambda\,, \quad -2\mu \leqq \theta < 0\,, \\ &= \theta^2(\theta + \mu)/2\mu\,, \quad 0 \leqq \theta \leqq 2\lambda\,, \\ &= 2\lambda(\theta + \mu)(\theta - \lambda)/\mu\,, \quad 2\lambda < \theta\,. \end{aligned} \tag{7.7}$$

In particular, the bound in (2.15) is not attained in the symmetric case $\lambda = \mu$, although it is approached as θ increases. A similar examination of $\lim_a [-\log \gamma(\theta)]/E_\theta\tau$ shows that the bound in (2.17) is also not attained.

8. A square-root boundary. We consider here a number of tests, arising out of proposals due to various authors, that have similar asymptotic properties. A prototype of these tests is given by the stopping time

$$\sigma = \inf\{t: X(t) \notin (-\lambda[a(t+1)]^{\frac{1}{2}}, \mu[a(t+1)]^{\frac{1}{2}})\}\,,$$

where λ and μ are fixed positive constants. For all tests in this section, the terminal decision is to accept H_1 iff the stopped process is negative. We will study the truncated test $\tau = \sigma \wedge a$, but as seen below, there is also theoretical interest in the behavior of the test using σ.

A version of the untruncated test σ arises from the theory of APO tests developed by Bickel and Yahav (1967a). For this, one utilizes a Bayesian formulation, introducing a prior distribution for θ, a regret function $r(\theta)$, representing the loss incurred on taking the wrong decision when θ obtains and a cost of sampling c. An APO test (i.e., stopping time) T asymptotically (as

$c \to 0$) minimizes the posterior risk $Y_T + cT$, where Y_t is the posterior expected loss for making a terminal decision, having observed $[X(s): 0 \leqq s \leqq t]$. Under certain conditions, satisfied in the present case, Bickel and Yahav showed that the stopping time

$$T = \inf\{t: Y_t \leqq c\} \tag{8.1}$$

is APO. For prior pdf $\varphi(\theta)$ and the usual 0–1 testing loss structure ($r(\theta) \equiv 1$), it is straightforward to calculate that $Y_t = \Phi(-(t+1)^{-\frac{1}{2}}|X(t)|)$. If we take $a = [\Phi^{-1}(c)]^2$ [so that $a \sim -2\log c$ for small c] and $\lambda = \mu = 1$, then $T = \sigma$. In this case τ is not APO; the truncation destroys the APO property. (However, if σ is truncated at s_a, where $a = o(s_a)$, the APO property is preserved.) Nevertheless, we will see that σ, τ and the other stopping times considered in this section share interesting asymptotic properties and that, if anything, τ seems better (although not necessarily good), judged by our criteria.

Naturally closed APO rules for the above specifications are contained in the following class of rules. For $m > 0$, let

$$T_m = \inf\{t: |X(t)| \geqq (at - mt\log t)^{\frac{1}{2}}\}\,, \tag{8.2}$$

where $a = -2\log c$. Note that $T_m \leqq e^{a/m}$. If one replaces Y_t above by the (large t) approximation $Z_t = t^{-\frac{1}{2}}e^{-X^2(t)/2t}$, the corresponding rule, $\inf\{t: Z_t \leqq c\}$, is just T_1. It may be verified that for $0 < m < 2$, T_m is APO for the problem with 0–1 loss. If the regret function is changed to $r(\theta) = |\theta|$, one obtains from (8.1) the rule T_3; see Bickel and Yahav (1967b).

Taking a fully Bayesian approach, Chernoff considered Bayes rules for $r(\theta) = |\theta|$ and normal priors; see Chernoff (1965) for references. We consider here the Bayes test corresponding to a Lebesgue prior. This is only a technical convenience and our results hold for proper normal priors as well. The Bayes rule is of the form

$$\sigma_B = \inf\{t: |X(t)| \geqq h(t, a)\}\,,$$

where again $a = -2\log c$. Unlike the corresponding APO rule T_3, σ_B is not closed (although it is also APO). It does not seem possible to give an explicit expression for $h(t, a)$. However, Chernoff (1965) showed that

$$h(t, a) = (at - 3t\log t + O(1))^{\frac{1}{2}}\,, \qquad t = o(e^{a/3})\,. \tag{8.3}$$

(In utilizing Chernoff's results, one must keep in mind that he considers a rescaled problem in which time is measured in units of $c^{\frac{2}{3}} = e^{-a/3}$.) The similarity between σ_B and T_3 is apparent and in fact, (8.3) suffices to show that σ_B and the APO rules all behave the same, asymptotically, judged by our criteria. We note that σ (and τ) differ from the other rules discussed in that their boundaries lack a logarithmic term. The presence (or absence) of this term has no effect, asymptotically, in our considerations.

The test based on σ is also a modification of a test suggested by Darling and Robbins; see Robbins (1970), pages 1404–1405. They obtained tests having uniformly small rates for testing H_1': $\theta < 0$ ($\theta = 0$ being excluded). One of their stopping times (adapted to continuous time) is of the form

$$\sigma^* = \inf\{t: |X(t)| \geqq [a(t+1) + (t+1)\log(t+1)]^{\frac{1}{2}}\}.$$

The presence of the logarithmic term assures that the error rates are uniformly small, but it also entails $P_0(\sigma^* = \infty) > 0$. This latter property seems undesirable for our problem. Of course a truncation of σ^* avoids this difficulty. The interested reader can check that the asymptotics developed below for the test based on τ apply as well to the procedure based on $\tau^* = \sigma^* \wedge a$. It can also be shown that σ^* is APO for the 0–1 loss structure discussed above.

We note one other context in which σ arises. Wald (1947) suggested a method of weight-functions for obtaining sequential tests of composite hypotheses. By taking the weight-function to be $\varphi(\cdot)$ on $(-\infty, 0)$ and $(0, \infty)$, it is straightforward to verify that the stopping time of the resulting test is σ.

We turn now to asymptotic considerations, beginning with the behavior of σ and τ. Results for σ are included, not necessarily because there is inherent interest in the untruncated test, but because the results apply as well to the other tests mentioned above. Results are given for $\theta \geqq 0$, the case $\theta < 0$ being analogous.

8.1. THEOREM. W.p. 1,

$$\begin{aligned} \sigma/a &\to \mu^2/\theta^2, \quad \theta > 0, \\ &\to \infty, \quad \theta = 0. \end{aligned}$$

For $\theta > 0$, $E_\theta \sigma/a$ has the same limit, while for large a $E_0\sigma = \infty$. Consequently, w.p. 1,

$$\begin{aligned} \tau/a &\to 1, \quad 0 \leqq \theta \leqq \mu, \\ &\to \mu^2/\theta^2, \quad \mu < \theta \end{aligned}$$

and $E_\theta \tau/a$ has the same limit.

PROOF. Choose $\theta > 0$. We have

$$X(\sigma) = \mu[a(\sigma+1)]^{\frac{1}{2}} 1_R - \lambda[a(\sigma+1)]^{\frac{1}{2}} 1_A \tag{8.4}$$

and, as in Theorem 4.1, $P_\theta(1_R \to 1) = 1$. On dividing across by σ and letting $a \to \infty$, we see that $\mu(a/\sigma)^{\frac{1}{2}} \to \theta$ w.p. 1. To do $E_\theta \sigma/a$, we note first that by Fatou, $\liminf_a E_\theta \sigma/a \geqq \mu^2/\theta^2$. Also, $X(\sigma) \leqq \mu[a(\sigma+1)]^{\frac{1}{2}}$, so that by the version of Wald's lemma due to Robbins and Samuel (1966), $E_\theta X(\sigma) = \theta E_\theta \sigma \leqq \mu[a(E\sigma + 1)]^{\frac{1}{2}}$, which entails $\limsup_a (E_\theta \sigma/a)^{\frac{1}{2}} \leqq \mu/\theta$.

Now consider $\theta = 0$. Suppose $\lambda \leqq \mu$. Then $|X(\sigma)| \geqq \lambda[a(\sigma+1)]^{\frac{1}{2}}$, so on dividing by σ and letting $a \to \infty$, we find $a/\sigma \to 0$ w.p. 1. That $E_0\sigma = \infty$ for large a follows from Shepp (1967). The pointwise results for τ follow from those for σ. Convergence of $E_\theta \tau/a$ then follows by dominated convergence. □

REMARK. The first part of the theorem remains true with σ replaced by T_m,

σ_B or σ^*. When $\theta = 0$, on replacing $X(t)$ by $e^{a/2m}X(te^{-a/m})$, one sees that $T_m \sim e^{a/m}T_m(0)$ and similarly, from Chernoff's results, that $\sigma_B \sim e^{a/3}\sigma_B(0)$.

The following lemma enables us to estimate the error rates for σ and τ. An examination of the proof shows that it holds as well for T_m, σ_B and σ^*.

8.2. LEMMA. *For $\theta \geqq 0$ and fixed $r \geqq 1$, $P_\theta A_\sigma(\sigma < a^r) = e^{-\frac{1}{2}\lambda^2 a + o(a)}$. Here $A_\sigma = (X(\sigma) < 0)$.*

PROOF. Clearly $P_\theta A_\sigma(\sigma < a^r) \leqq P_0(X(t) < \lambda[a(t+1)]^{\frac{1}{2}}$, some $0 \leqq t \leqq a^r) \leqq \sum_0^{a^r} p_k$, where $p_k = P_0(X(t) < -\lambda[a(t+1)]^{\frac{1}{2}}$, some $k \leqq t \leqq k+1)$. Since $X(t)$ has independent increments, $\inf\{X(t) : k \leqq t \leqq k+1\} \sim X(k) + \Delta$, where $\Delta \sim \inf\{X(t) : 0 \leqq t \leqq 1\}$ and is independent of $X(k)$. It is well known that for $x > 0$, $P_0(\Delta < -x) = 2P_0(X(1) < -x)$, so by a simple convolution argument, $P_0(X(k) + \Delta < -x) \leqq 2P_0(X(k+1) < -x)$. Thus $p_k \leqq 2P_0(X(k+1) < -\lambda[a(k+1)]^{\frac{1}{2}}) \leqq e^{-\frac{1}{2}\lambda^2 a}$ for large a and

$$P_\theta A_\sigma(\sigma < a^r) \leqq a^r e^{-\frac{1}{2}\lambda^2 a}, \qquad \text{large} \quad a, \tag{8.5}$$

which provides half of the result.

To obtain the other half, we use (2.1) with $-\nu < 0$ to obtain

$$\begin{aligned} P_\theta A(\sigma < a^r) &= \textstyle\int_{A_\sigma(\sigma < a^r)} \exp[-\lambda(\nu + \theta)[a(\sigma+1)]^{\frac{1}{2}} + \frac{1}{2}(\nu^2 - \theta^2)\sigma]\, dP_{-\nu} \\ &\geqq \textstyle\int_{A_\sigma(\varepsilon^{-1} < \sigma < a^r)} \exp[-\lambda(\nu+\theta)[a\sigma(1+\varepsilon)]^{\frac{1}{2}} \\ &\qquad + \tfrac{1}{2}(\nu^2 - \theta^2)\sigma]\, dP_{-\nu}\,, \qquad\qquad \varepsilon > 0 \\ &\geqq \exp[-\tfrac{1}{2}\lambda^2 a(1+\varepsilon)(\nu+\theta)/(\nu-\theta)]P_{-\nu}A_\sigma(\varepsilon^{-1} < \sigma < a^r)\,, \end{aligned} \tag{8.6}$$

the exponential term in the last expression being the minimum value of the second integrand (for σ ranging in $(0, \infty)$), provided $\nu > \theta$. Also, for $\nu > \lambda$, $\lim_a P_{-\nu}A_\sigma(\varepsilon^{-1} < \sigma < a^r) = 1$. Since ν in (8.6) can be taken arbitrarily large and ε, arbitrarily small, this, together with (8.5), establishes the lemma. □

The behavior of the error rates is given in the following theorems. The error probability for σ is denoted by $\varepsilon_\sigma(\theta)$.

8.3. THEOREM. *For $\theta > 0$, $\lim_a [-\log \varepsilon_\sigma(\theta)]/a = \frac{1}{2}\lambda^2$.*

PROOF. For $\theta > 0$, $\varepsilon_\sigma(\theta) = P_\theta A_\sigma(\sigma < a^2) + P_\theta A_\sigma(\sigma \geqq a^2)$. The behavior of the first term is given in Lemma 8.2, so it remains to dispose of the second:

$$\begin{aligned} P_\theta A_\sigma(\sigma \geqq a^2) &\leqq P_0(\inf_{t > a^2}(X(t) + \theta t) < 0) \\ &= P_0(X(a^2) + \inf_{t>0}(X(t) + \theta t) < -\theta a^2) \\ &\leqq P_0(X(a^2) < -\tfrac{1}{2}\theta a^2) + P_0(\inf_{t>0}(X(t) + \theta t) < -\tfrac{1}{2}\theta a^2) \\ &= \Phi(-\tfrac{1}{2}\theta a) + e^{-\theta^2 a^2} \leqq 2e^{-\theta^2 a^2/8}\,, \qquad a \quad \text{large}\,. \end{aligned}$$

The theorem then follows from Lemma 8.2. □

We again remark that the above proof and result applies as well to T_m, σ_B and σ^*. We note that the limit given in Theorem 8.3 does not depend on θ. (For the other tests considered in this paper, the corresponding limit increases with θ.) This seems to suggest that the power-function of the test based on σ (or T_m, σ_B or σ^*) increases relatively slowly to 1 as $\theta \to \infty$. Some numerical evidence might help in providing a more definitive interpretation of this phenomenon. The corresponding behavior of $\varepsilon(\theta)$, the error probability for τ, is given in the next theorem.

8.4. THEOREM.

$$\lim_a [-\log \varepsilon(\theta)]/a = \tfrac{1}{2}\theta^2, \quad 0 \leqq \theta \leqq \lambda,$$
$$= \tfrac{1}{2}\lambda^2, \quad \lambda < \theta.$$

PROOF. For $\theta > 0$, we have $\varepsilon(\theta) = P_\theta A(\sigma < a) + P_\theta A(\sigma \geqq a)$. Again because of Lemma 8.2, we need only analyze the behavior of the second term. Using (2.1),

$$P_\theta A(\sigma \geqq a) = e^{-\frac{1}{2}\theta^2 a} \int_{A(\sigma \geqq a)} e^{\theta X(a)}\, dP_0 . \tag{8.7}$$

Since $A(\sigma \geqq a) = (X(a) < 0, \sigma \geqq a)$,

$$\int_{(X(a)<0)} e^{\theta X(a)}\, dP_0 - \int_{A(\sigma \geqq a)} e^{\theta X(a)}\, dP_0 \leqq P_0(\sigma < a) < 2ae^{-\frac{1}{2}(\lambda \wedge \mu)^2 a}$$

by (8.5). As in (5.11) with $\eta = 0$, $\int_{(X(a)<0)} e^{\theta X(a)}\, dP_0 = e^{o(a)}$, hence by Lemma 8.2 and the preceding,

$$\varepsilon(\theta) = e^{-\frac{1}{2}\lambda^2 a + o(a)} + e^{-\frac{1}{2}\theta^2 a + o(a)},$$

from which the theorem follows. □

As contrasted with Theorem 8.3, the limit has some dependence on θ. This arises from the truncation, since for $0 \leqq \theta \leqq \lambda$, the test behaves, asymptotically, like a nonsequential test. If the truncation were to increase at a rate faster than a, we would obtain the limit given in Theorem 8.3.

Combining Theorems 8.1 and 8.4, we find

$$\begin{aligned} \lim_a [-\log \varepsilon(\theta)]/E_0\tau &= \tfrac{1}{2}\theta^2, \quad 0 \leqq \theta \leqq \lambda, \\ &= \tfrac{1}{2}\lambda^2, \quad \lambda < \theta, \end{aligned} \tag{8.8}$$

confirming that for $-\mu \leqq \theta \leqq \lambda$, the test based on τ behaves asymptotically like a nonsequential test. Also, for $\lambda \leqq \mu$,

$$\begin{aligned} \lim_a [-\log \gamma(\theta)]/E_\theta\tau &= \tfrac{1}{2}\theta^2, \quad 0 \leqq \theta \leqq \lambda, \\ &= \tfrac{1}{2}\lambda^2, \quad \lambda \leqq \theta \leqq \mu, \\ &= \lambda^2\theta^2/2\mu^2, \quad \mu \leqq \theta, \end{aligned} \tag{8.9}$$

and we see that the bound in (2.17) is missed by at least a factor of 4. For σ we have, correspondingly,

$$\begin{aligned} &\lim_a [-\log \varepsilon_\sigma(\theta)]/E_0\sigma = 0 \quad \text{and} \\ &\lim_a [-\log \gamma_\sigma(\theta)]/E_\theta\sigma = \lambda^2\theta^2/2\mu^2, \quad \text{all } \theta . \end{aligned} \tag{8.10}$$

Judging by these last relations, the tests based on τ, σ, T_m, σ_B and σ^* have poor asymptotic performance; this despite the fact that σ_B is an "optimal" test.

It can only be concluded that the Bayes criterion of optimality is unrelated to the measures of performance we consider here. In this connection, we remark that often the components of the total risk, Y_T and cT (or their expectations), tend to zero at different rates (as $c \to 0$). As the latter term is usually dominant, the Bayes criterion loses sight of the error probabilities, so to speak. This suggests exercising caution in embracing the Bayes critierion as an overall measure of performance of sequential tests.

9. Repeated significance test. Samuel-Cahn (1974) discusses repeated significance tests for a normal mean. The continuous-time version of one of these procedures is the following: For a fixed positive μ, let $\sigma = \inf\{t: X(t) \geqq \mu a^{\frac{1}{2}}\}$, $\tau = \sigma \wedge a$ and $A = (X(\tau) < \mu a^{\frac{1}{2}})$. Here μ is an appropriate fractile of the standard normal distribution, chosen to give a desired $\varepsilon(0)$. The relation between μ and $\varepsilon(0)$ is given in (9.1) below; see also Samuel-Cahn (1974). This test may be described informally as follows. A total observation time a is chosen in advance. If, at any time prior to a, $X(t)$ is significant, as judged by the critical value appropriate for time a, sampling is terminated and H_1 rejected. Only if $X(t)$ fails to be significant for any $0 \leqq t \leqq a$ is H_1 accepted.

Following are some asymptotic properties of the test. It is natural here that the (upper) boundary and the truncation point increase at different rates; cf. Section 6.

9.1. THEOREM. *For* $\theta > 0$,

$$\tau/a^{\frac{1}{2}} \to \mu/\theta[P_\theta] \quad \textit{and}$$
$$\tau/a \to 1[P_{-\theta}] .$$

The corresponding expectations have the same limits. When $\theta = 0$, $\tau \sim a\tau(1)$. *In particular,* $E_0\tau = aE_0\tau(1) < a$.

PROOF. For positive drift, the pointwise result follows from the fact that $X(\sigma) = \mu a^{\frac{1}{2}}$ and that $X(\sigma)/\sigma \to \theta[P_\theta]$. Thus $P_\theta(\sigma/a^{\frac{1}{2}} \to \mu/\theta) = 1$ and similarly for τ. For the expectation, using the version of Wald's lemma due to Robbins and Samuel (1966), $E_\theta X(\sigma) = \theta E_\theta \sigma = \mu a^{\frac{1}{2}}$, hence $E_\theta\sigma/a^{\frac{1}{2}} = \mu/\theta$ and $E_\theta\tau/a^{\frac{1}{2}} \leqq \mu/\theta$. Also, by Fatou, $\liminf_a E_\theta\tau/a^{\frac{1}{2}} \geqq \mu/\theta$.

For negative drift, $P_{-\theta}(\sigma < a, \text{ some } a > x) \leqq P_0(\sup_{0<t<\infty}[X(t) - \theta t] > \mu x^{\frac{1}{2}}) \to 0$ as $x \to \infty$. Hence $P_{-\theta}(\tau/a \to 1) = 1$. Convegence of the expectation follows by dominated convergence. For $\theta = 0$, replacing $X(t)$ by $a^{\frac{1}{2}}X(t/a)$ gives the desired result. □

REMARK. Replacing $X(t)$ by $a^{\frac{1}{2}}X(t/a)$ similarly yields

$$\varepsilon(0) = P_0(\sup_{0<t<1} X(t) \geqq \mu) = 2[1 - \Phi(\mu)] . \tag{9.1}$$

We note another consequence of Theorem 9.1. Clearly $E_0(\tau \mid A) = a$, so that $\lim_a E_0(\tau \mid A)/E_0\tau = 1/E_0\tau(1) > 1$. Since $E_0\tau = \bar{\varepsilon}(0)E_0(\tau \mid A) + \varepsilon(0)E_0(\tau \mid R)$, it follows that $\lim_a E_0(\tau \mid R)/E_0\tau < 1$. Hence by Lemma 2.1, for $\theta > 0$,

$[-\log \varepsilon(-\theta)]/E_0\tau$ cannot, in the limit, attain the value $\frac{1}{2}\theta^2$. That it does not is also evident from the evaluations given below.

From Samuel-Cahn (1974), we have that

$$E_0\tau(1) = [2\Phi(\mu) - 1] + \mu^2[1 - \Phi(\mu)] + 2\mu\varphi(\mu) \tag{9.2}$$

and

$$\begin{aligned} P_\theta(\sup\nolimits_{0<t<a} X(t) < \mu a^{\frac{1}{2}}) &= \Phi(\mu - \theta a^{\frac{1}{2}}) - e^{2\mu\theta a^{\frac{1}{2}}}\Phi(-\mu - \theta a^{\frac{1}{2}}) \\ &= e^{-\frac{1}{2}\theta^2 a + o(a)} \qquad \text{if} \quad \theta > 0\,. \end{aligned} \tag{9.3}$$

For $\theta > 0$, (9.3) is $\varepsilon(\theta)$, thus

$$\lim\nolimits_a [-\log \varepsilon(\theta)]/E_0\tau = \theta^2/2E_0\tau(1) > \tfrac{1}{2}\theta^2\,, \qquad \theta > 0\,, \tag{9.4}$$

so that (2.11) is not valid for this test. In fact, the bound given by Lemma 2.1 is attained in this case, although, as can be seen from (9.1) and (9.2), that of (2.10) is not; the bound there being $\theta^2/2[2\Phi(\mu) - 1]$ in this case.

For $-\theta < 0$, (9.3) gives

$$\varepsilon(-\theta) = \exp[-2\mu\theta a^{\frac{1}{2}}[1 + o(1)]]\,, \tag{9.5}$$

hence

$$\lim\nolimits_a [-\log \varepsilon(-\theta)]/E_0\tau = 0\,, \qquad \theta > 0\,. \tag{9.6}$$

The contrast between (9.4) and (9.6) reflects the markedly asymmetric nature of this test. On the other hand, from (9.5) and Theorem 9.1, we see that

$$\lim\nolimits_a [-\log \varepsilon(-\theta)]/E_0\tau = 2\theta^2\,, \qquad \theta > 0\,, \tag{9.7}$$

so that the bound in (2.15) is attained. This last result just indicates that, as far as hitting the upper boundary is concerned, this test behaves asymptotically like the SPRT of Section 4. Further indication of the asymmetric nature of the test is given by

$$\begin{aligned} \lim\nolimits_a [-\log \varepsilon(\theta)]/E_\theta\tau &= \infty\,, \quad \theta > 0\,, \\ &= 0\,, \quad \theta < 0\,. \end{aligned} \tag{9.8}$$

Finally, we have

$$\lim\nolimits_a [-\log \gamma(\theta)]/E_0\tau = 0 = \lim\nolimits_a [-\log \gamma(\theta)]/E_{-\theta}\tau\,, \qquad \theta > 0 \tag{9.9}$$

and

$$\lim\nolimits_a [-\log \gamma(\theta)]/E_\theta\tau = 2\theta^2\,, \qquad \theta > 0\,. \tag{9.10}$$

10. Some comparisons. As a partial synthesis of the foregoing, we collect here, for purposes of comparison, results for the symmetric versions of the procedures considered above. Naturally, the repeated significance test of Section 9 must be excluded from such a comparison. Thus, for the procedures considered in Sections 3–8, we take $\lambda = \mu$ and (where relevant) $\eta = 0$. For symmetric procedures, many of the ratios discussed in Section 2 coalesce. In particular,

$\gamma(\theta) = 2\varepsilon(\theta)$ and the corresponding ratios based on these quantities have the same limits. Also, (2.11) necessarily holds for symmetric procedure (cf. Proposition 2.3). In the table below, the number(s) in parenthesis after (the abbreviated name for) a procedure gives the relevant section(s).

TABLE 1

Procedure	limit of		
	$[-\log \varepsilon(\theta)]/E_0\tau$		$[-\log \varepsilon(\theta)]/E_\theta\tau$
NS (3)	$\frac{1}{2}\theta^2$		$\frac{1}{2}\theta^2$
TPRT (5)	$\frac{1}{2}\theta^2$,	$\|\theta\| \leqq \lambda$,	$\frac{1}{2}\theta^2$
	$\frac{1}{2}\theta^2$,	$\lambda \leqq \|\theta\| \leqq 4\lambda$,	$\theta^3/2\lambda$
	$2\lambda\theta$,	$\|\theta\| > 4\lambda$,	$2\theta^2$
SPRT and TPRT II (4, 6)	0		$2\theta^2$
AND (7)	$\frac{1}{2}\theta^2$,	$\|\theta\| \leqq 2\lambda$,	$[(\theta+\lambda)/2\lambda]\theta^2$
	$2\lambda(\theta - \lambda)$,	$\|\theta\| > 2\lambda$,	$2(\theta^2 - \lambda^2)$
TAPO (8)	$\frac{1}{2}\theta^2$,	$\|\theta\| \leqq \lambda$,	$\frac{1}{2}\theta^2$
	$\frac{1}{2}\lambda^2$,	$\|\theta\| > \lambda$,	$\frac{1}{2}\theta^2$
APO, Bayes and Darling-Robbins (8)	0		$\frac{1}{2}\theta^2$
Upper bound (2.16, 2.17)	$\frac{1}{2}\theta^2$,		$2\theta^2$

From the table, it appears that the untruncated procedures having a square root boundary (the APO, Bayes and Darling-Robbins procedures) fare poorly, asymptotically. The limiting values are dominated by those for the SPRT and TPRT. Truncating does not help either; the TAPO is also dominated by the TPRT. (To effect this comparison, if λ is chosen for the TAPO, the same value should be chosen for the TPRT.) Although the SPRT itself is not dominated in this way, and is in fact optimal when judged by the criterion $\lim_\alpha [-\log \varepsilon(\theta)]/E_\theta\tau$, one may wish to exclude it from consideration for other reasons: The fact that τ is unbounded and/or that $E_\theta\tau$ is unduly large for θ near zero. In choosing among these procedures, the author's inclination is to restrict attention to the nonsequential test, the TPRT and Anderson's boundaries. (None of these three procedures dominates the others.) Of course these conclusions should, at best, be viewed as tentative, pending corroborative finite-sample information.

11. Acknowledgment. The author is grateful to Professor J. Yahav for patiently listening as the present work developed and for providing helpful comments and discussion.

REFERENCES

ANDERSON, T. W. (1960). A modification of the sequential probability ratio test to reduce the sample size. *Ann. Math. Statist.* **31** 165–197.

BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics.* SIAM, Philadelphia.

BECHHOFER, R. (1960). A note on the limiting relative efficiency of the Wald sequential probability ratio test. *J. Amer. Statist. Assoc.* **55** 660–663.

BERK, R. H. (1973). Some asymptotic aspects of sequential analysis. *Ann. Statist.* **1** 1126–1138.

BERK, R. H. (1975). Locally most powerful sequential tests. *Ann. Statist.* **3** 373-381.

BICKEL, P. J. and YAHAV, J. (1967a). Asymptotically pointwise optimal procedures in sequential analysis. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 401-414. Univ. of California Press.

BICKEL, P. J. and YAHAV, J. (1967b). On testing sequentially the mean of a normal distribution. Technical report no. 26, Stanford Univ., No. 10.

CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493-507.

CHERNOFF, H. (1956). Large sample theory: parametric case. *Ann. Math. Statist.* **27** 1-22.

CHERNOFF, H. (1965). Sequential test for the mean of a normal distribution III (Small t). *Ann. Math. Statist.* **36** 28-54.

FREEDMAN, D. (1971). *Brownian Motion and Diffusion.* Holden-Day, San Francisco.

HALL, W. J. and LOYNES, R. (1975). Weak convergence of processes related to likelihood ratios. Unpublished.

HODGES, J. L. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324-335.

HOEFFDING, W. (1960). Lower bounds for the expected sample size and the average risk of a sequential procedure. *Ann. Math. Statist.* **31** 352-368.

LAI, T. L. (1973). Optimal stopping and sequential tests which minimize the maximum expected sample size. *Ann Statist.* **1** 659-673.

LAWING, W. D. and DAVID, H. T. (1966). Likelihood ratio computations of operating characteristics. *Ann. Math. Statist.* **37** 1704-1716.

LORDEN, G. (1970). On excess over the boundary. *Ann. Math. Statist.* **41** 520-527.

ROBBINS, H. (1970). Statistical methods related to the law of the iterated logarithm. *Ann. Math. Statist.* **41** 1397-1409.

ROBBINS, H. and SAMUEL, E. (1966). An extension of a lemma of Wald. *J. Appl. Probability* **3** 272-273.

SAMUEL-CAHN, E. (1974). Repeated significance test II, for hypotheses about the normal distribution. *Comm. Statist.* **3** (8) 711-733.

SHEPP, L. A. (1967). A first passage problem for the Wiener process. *Ann. Math. Statist.* **38** 1912-1914.

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

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903