The Annals of Statistics
1980, Vol. 8, No. 5, 1110–1122

RISK OF ASYMPTOTICALLY OPTIMUM SEQUENTIAL TESTS[1]

BY GLORIA C. ZERDY
Cornell University

The problem considered is that of testing sequentially between two separated composite hypotheses concerning the mean of a normal distribution with known variance. The parameter space is the real line, on which is assumed an a priori distribution, W, with full support. A family $\{\delta(c)\}$ of sequential tests is defined and shown to be asymptotically Bayes, as the cost, c, per observation tends to zero, relative to a large class of fully supported a priori distributions. The ratio of the integrated risk of the Bayes procedure to that of $\delta(c)$ is shown to be $1 - 0(\log\log c^{-1}/\log c^{-1})$, as c tends to zero, for every W.

1. Introduction. For testing sequentially between two separated composite hypotheses, $\mu \leqslant \mu_0$ and $\mu \geqslant \mu_1$, where μ is a real parameter of a distribution of exponential type, Schwarz (1962) has given an asymptotic description, as the cost c per observation tends to zero, of the Bayes continuation region, $B_W(c)$, relative to a fixed a priori distribution, W. He showed that $B_W(c)/\log c^{-1}$ approaches an "asymptotic shape" B_0, which depends on the a priori distribution only through its support. Schwarz suggested using a family of procedures which have $B_0 \log c^{-1}$ as their continuation regions. The advantages of these procedures are that there is a specific formula for $B_0 \log c^{-1}$, and that $B_0 \log c^{-1}$ depends on the a priori distribution only through its support. Wong's Lemma 5.1 (1968) shows that Schwarz's procedures, $\delta'(c)$, are asymptotically Bayes. The order of the "efficiency" of $\delta'(c)$, that is, the ratio of the integrated risk of the Bayes procedure to that of $\delta'(c)$, has not been determined.

As an aid to finding an asymptotic description of $B(c)$, Schwarz proved that $C(\Delta c \log c^{-1}) \subset B(c) \subset C(c)$, where $C(c)$ is the set on which the a posteriori risk of stopping is at least c, and Δ is a constant. For certain a priori distributions W with compact support, Lorden (1967) improved this result to $C(M^*c) \subset B(c)$ for some constant M^*. This enabled Lorden to prove that if, for a fixed a priori distribution W, $\delta_W(Qc)$ is the procedure which has continuation region $C(Qc)$ for some $Q > 0$, and which chooses a terminal decision having minimum a posteriori risk, the efficiency of $\delta_W(Qc)$ is $1 - 0(1/\log c^{-1})$, as $c \to 0$.

When the number of possible states is finite, and the a priori distribution W has full support, Lorden's result extends to a class of procedures which do not depend on W. Lorden points out that in this case, a procedure $\delta(c)$ which stops no later than $\delta_W(Qc)$ and chooses a terminal decision whose a posteriori risk is at most Kc, for some $Q > 0$, $K > 0$, has efficiency $1 - 0(1/\log c^{-1})$, for every a priori

Received January 1976, revised March 1978.

[1]This paper is part of the author's doctoral dissertation submitted August, 1975 at Cornell University. The research was partially supported by the National Science Foundation under grant GP MPS 72-04998A02.

AMS 1970 *subject classification.* Primary 62L10, 62F05.

Key words and phrases. Integrated risk, asymptotically Bayes sequential tests, asymptotic efficiency.

distribution with full support. In case the underlying distribution is normal, Schwarz's procedures satisfy these conditions, as can be easily shown. However, for the continuous parameter space, no extension of Lorden's result to procedures independent of the a priori distribution has been proven.

This paper is concerned with independent and identically distributed normal random variables with unknown mean and known variance, the parameter space being the entire real line. We define a family $\{\delta(c)\}$ of procedures which are modified versions of Schwarz's procedures, and which are asymptotically Bayes relative to every a priori distribution with a bounded Lebesgue density having full support and bounded away from zero in a neighborhood of the endpoints of the indifference interval. The family $\{\delta(c)\}$ is shown to possess the property that the ratio of the integrated risk of the Bayes procedure to that of $\delta(c)$ is $1 - 0$ $(\log \log c^{-1}/ \log c^{-1})$, as $c \to 0$, for every a priori distribution whose support is the real line. Although the efficiency obtained is not as good as the $1 - 0(1/\log c^{-1})$ obtained by Lorden for procedures $\delta_W(Qc)$, the procedures $\delta(c)$ have the advantage of being independent of the a priori distribution, within the class of distributions whose support is the real line.

2. Unbounded parameter space. The parameter space, Ω, is assumed to be the real line, on which is defined a probability measure, W, with density, g, with respect to Lebesgue measure. It is assumed that g is bounded above on Ω.

The random variables $X_1, X_2, \cdots$ are independent and identically distributed with $X_1 \sim N(\mu, 1), \mu \in \Omega$, and $S_n = X_1 + \cdots + X_n$ denotes their cumulative sum. The density function of S_n is

$$f_n(s, \mu) = \exp(s\mu - n\mu^2/2)$$

relative to the measure

$$\mu_n(s) = \int_{-\infty}^{s} (2\pi n)^{-\frac{1}{2}} \exp(-y^2/2n) dy.$$

We are testing between $H_0 : \mu \leqslant -1$ and $H_1 : \mu \geqslant 1$. Assume g is bounded away from zero at $\mu = -1, 1$. Let $l(\mu)$ be the loss for making a wrong decision when μ is the true parameter value; $l(\mu) = 0$ on $(-1, 1)$ and $0 < L \leqslant l(\mu) \leqslant 1$ elsewhere. The a posteriori risk of stopping is defined by

$$R(n, S_n) = \left\{\min_{i=0,1} \int_{H_i} \exp(\mu S_n - n\mu^2/2) l(\mu) W(d\mu)\right\} / \left\{\int_\Omega \exp(\mu S_n - n\mu^2/2) W(d\mu)\right\}.$$

Let c denote the cost of each observation X_i, and let $C(c) = \{(x,y) : R(x,y) \geqslant c\}$. The Bayes continuation region with respect to W and c will be denoted by $B_W(c)$. The limiting region of Schwarz's paper is, in our case, defined by

$$B_0 = \left\{(x,y) : 1 + \min_{i=0,1}(\mu_i y - x/2) \geqslant \sup_\mu(\mu y - x\mu^2/2), \quad x \geqslant 0, \mu_0 = -1, \mu_1 = 1\right\}$$
$$= \left\{(x,y) : x - (2x)^{\frac{1}{2}} \leqslant -x + (2x)^{\frac{1}{2}}, x \geqslant 0\right\}.$$

Let $u_k(c)$ and n_k denote the x-coordinates of the points of intersection of the line $y = kx$ with the boundaries of $C(c)$ and B_0, respectively. Then $n_k = 2(1 + |k|)^{-2}$.

Let $T = \log c^{-1}$. We assume throughout that $c \leqslant c_0$, where c_0 is chosen so that $\log c_0^{-1} > 1$, and so that $c \leqslant c_0$ implies $2T^{-1}\log T \leqslant \varepsilon_0 < 1$, for some ε_0 satisfying $0 < \varepsilon_0 < 1$.

We begin by proving several lemmas. For positive α, we let $B_0\alpha = \{(\alpha x, \alpha y) : (x,y) \in B_0\}$.

LEMMA 1. *Let $N(c)$ be the first time (n, S_n) exits $B_0T(1 - \varepsilon)$, where $\varepsilon = 2T^{-1}\log T$, and let $\eta = [2(1 - \varepsilon)\varepsilon]^{\frac{1}{2}}$. For $\mu \in \Omega' = \{\mu : |\mu| + 1 < [(1 - \varepsilon)T/\log T]^{\frac{1}{2}}/4\}$.*

$$E_\mu N(c) \geqslant T\left(n_\mu - M_1 T^{-1}\log T\right) \qquad \text{if } |\mu| \geqslant \eta/2(1-\varepsilon)$$

$$\geqslant T\left(n_\mu - M_2(T^{-1}\log T)^{\frac{1}{2}}\right) \qquad \text{if } |\mu| < \eta/2(1-\varepsilon)$$

where M_1 and M_2 are constants independent of c, and M_1 (respectively M_2) is independent of $\mu \in \Omega', |\mu| \geqslant \eta/2(1 - \varepsilon)$ (respectively $|\mu| < \eta/2(1 - \varepsilon)$).

PROOF. Let $W(t)$ denote a standard Wiener process, i.e. $P[W(0) = 0] = 1$, $EW(t) = 0$ for all t satisfying $t \geqslant 0$, and $\mathrm{Cov}[W(s), W(t)] = \min(s, t)$. If $EX_1 = \mu$, then $S_n - \mu n$ and $W(n)$ have the same distribution. Using this fact, and a well-known result about Wiener processes which can be found in Doob (1953) page 392, we have for any $\mu \in \Omega$

$$\begin{aligned} P_\mu[\max\nolimits_{1 \leqslant n \leqslant 2(1-\varepsilon)T}|S_n - \mu n| \geqslant \eta T] &\leqslant P[\sup\nolimits_{0 \leqslant t \leqslant 2(1-\varepsilon)T}|W(t)| \geqslant \eta T] \\ &\leqslant 4P[W(2(1-\varepsilon)T) \geqslant \eta T] \\ &\leqslant \frac{4[2(1-\varepsilon)]^{\frac{1}{2}}}{\eta T^{\frac{1}{2}}} e^{-\eta^2 T/4(1-\varepsilon)} \\ &\leqslant \frac{4T^{-1}}{(2\log\log c_0^{-1})^{\frac{1}{2}}} \\ &= D_1 T^{-1} \text{ (say)}. \end{aligned} \tag{1}$$

By symmetry one need only consider $\mu \geqslant 0$. For $i = 1, 2$, let x_i be the larger of the x-coordinates of the intersections of $y = \mu x + (-1)^{i-1}\eta T$ with the (extended) upper boundary of $B_0(1 - \varepsilon)T$, that is, with $y = -x + (2(1 - \varepsilon)Tx)^{\frac{1}{2}}$. Let $\gamma = 1 - \varepsilon$. Then

$$x_1 = T\left\{n_\mu\gamma - \frac{\eta}{\mu+1} + \frac{\gamma}{(\mu+1)^2}\left[(1 - 2\eta(\mu+1)/\gamma)^{\frac{1}{2}} - 1\right]\right\}$$

and

$$x_2 = T\left\{n_\mu\gamma + \frac{\eta}{\mu+1} + \frac{\gamma}{(\mu+1)^2}\left[(1 + 2\eta(\mu+1)/\gamma)^{\frac{1}{2}} - 1\right]\right\}.$$

Let L be the line through the points (x_i, y_i) $i = 1, 2$, which lie on the (extended) upper boundary of $B_0(1-\varepsilon)T$, so that $y_i = \mu x_i + (-1)^{i-1}\eta T$. Then L has equation

$$y = \left(\mu - \frac{2\eta T}{x_2 - x_1}\right)x + T\left(\frac{2\eta x_1}{x_2 - x_1} + \eta\right).$$

Let N_L be the first time $S_n' \geqslant (2\eta x_1/(x_2 - x_1) + \eta)T$ where $S_n' = \sum_{i=1}^n Y_i$, $Y_i = X_i - (\mu - 2\eta T/(x_2 - x_1))$. Note that N_L is the first time S_n crosses L. The Y_i are i. i. d. $N(2\eta T/(x_2 - x_1), 1)$. Since $E_\mu N_L$ and $E|Y_1|$ are finite one can apply Wald's equation to get

$$\begin{aligned}[2\eta T/(x_2 - x_1)]E_\mu N_L &= E_\mu S'_{N_L} \\ &\geqslant (2\eta x_1/(x_2 - x_1) + \eta)T.\end{aligned}$$

Hence,

$$\begin{aligned}E_\mu N_L &\geqslant (x_1 + x_2)/2 \\ &= T\gamma\left\{n_\mu + \left[(1 + 2\eta(\mu+1)/\gamma)^{\frac{1}{2}}\right.\right. \\ &\quad \left.\left. + (1 - 2\eta(\mu+1)/\gamma)^{\frac{1}{2}} - 2\right]/2(\mu+1)^2\right\}.\end{aligned}$$

It is easily verified that

$$(1 + x)^{\frac{1}{2}} \geqslant 1 + x/2 - x^2/2 \qquad \text{for} \quad |x| < 1.$$

Thus for $\eta < \gamma/2(\mu + 1)$ or $(\mu + 1) < \frac{1}{2}[\gamma T/4 \log T]^{\frac{1}{2}}$, we have

$$\begin{aligned}E_\mu N_L &\geqslant T\gamma\left[n_\mu - 2\eta^2/\gamma^2\right] \\ &= T\left(n_\mu - \varepsilon n_\mu - 2\eta^2\gamma^{-1}\right) \\ &\geqslant T\left(n_\mu - D_2\eta^2\right),\end{aligned} \tag{2}$$

where D_2 is a constant independent of μ in Ω' and c, for $c \leqslant c_0$.

If $\mu \geqslant \eta/2\gamma$, the point of intersection, having the larger x-coordinate, of $y = \mu x + (-1)^i\eta T$ $(i = 0, 1)$ with the boundary of $B_0T(1-\varepsilon)$ is on the curve $y = -x + (2\gamma TX)^{\frac{1}{2}}$.

Let x_0 be the smaller of the x-coordinates of the intersections of $y = \mu x + \eta T$ with the upper boundary of $B_0(1-\varepsilon)T$, that is, with $y = -x + (2\gamma Tx)^{\frac{1}{2}}$. Then

$$x_0 = T\left\{n_\mu\gamma - \frac{\eta}{\mu+1} - \frac{\gamma}{(\mu+1)^2}\left[(1 - 2\eta(\mu+1)/\gamma)^{\frac{1}{2}} + 1\right]\right\}.$$

Let $A_1 = \{\max_{1 \leqslant n \leqslant x_0}|S_n - \mu n|/n < \eta T/x_0\}$. Then proceeding as in (1), and using the fact that $x_0 \leqslant 2T\eta^2/\gamma = 8 \log T$, whenever $(\mu + 1) < \gamma/2\eta$, we have

$$\begin{aligned}P_\mu(A_1^c) &\leqslant P_\mu\left[\max_{1 \leqslant n \leqslant x_0}|S_n - \mu n| \geqslant \eta T/x_0\right] \\ &\leqslant 4\frac{x_0^{\frac{3}{2}}}{\eta T}e^{-\eta^2T^2/2x_0^3} \\ &\leqslant 4\frac{x_0^{\frac{3}{2}}}{\eta T}e^{-(\log T + D_3)} \\ &\leqslant D_4T^{-1}\end{aligned}$$

where D_3 and D_4 are constants independent of μ in Ω' and c, for $c \leqslant c_0$.

Let $A_2 = \{\max_{x_0 < n \leqslant 2\gamma T} |S_n - \mu n| < \eta T\}$, and let $A = A_1 \cap A_2$. Then (1) implies $P_\mu(A_2^c) \leqslant D_1 T^{-1}$, so that

$$P_\mu(A^c) \leqslant (D_1 + D_4)T^{-1}. \tag{3}$$

Given the event A_2, if (n, S_n) has not exited from $B_0 T(1 - \varepsilon)$ before time x_0, then (n, S_n) must first cross L before exiting $B_0 T(1 - \varepsilon)$ because of the convexity of $B_0 T(1 - \varepsilon)$. Given A_1, (n, S_n) cannot exit $B_0 T(1 - \varepsilon)$ before time x_0 because of the convexity and the definition of x_0. Thus, for $\mu \geqslant \eta/2\gamma$, $N(c) \geqslant N_L$ with probability one, given the event $A = A_1 \cap A_2$.

Let $b = 2\eta x_1/(x_2 - x_1) + \eta$ and $m = 2\eta T/(x_2 - x_1)$. Note that $S_n' = S_n - n(\mu - m)$ and N_L is the first time $S_n' \geqslant bT$. It is easily seen that $b/m < 2$ and $m^{-1} < 4$.

Now

$$\begin{aligned} E_\mu N(c) &\geqslant E_\mu N_L - \textstyle\int_{(N(c) < N_L)} E\left[N_L - N(c) \mid S'_{N(c)}\right] dP_\mu \\ &\geqslant E_\mu N_L - \textstyle\int_{(N(c) < N_L)} \left(\frac{bT - S'_{N(c)}}{m} + 17\right) dP_\mu, \end{aligned}$$

using Wald's equation and the upper bound on excess over the boundary in Lorden (1970). Thus

$$\begin{aligned} E_\mu N(c) &\geqslant E_\mu N_L - (2T + 17)P_\mu(N(c) < N_L) + 4E_\mu \inf_n S_n' \\ &\geqslant E_\mu N_L - \text{const.}, \end{aligned} \tag{4}$$

using (3) to get an upper bound on the probability and also Kingman's (1962) inequality

$$E_\mu \inf_n S_n' \geqslant -\frac{\operatorname{Var}_\mu S_1'}{2E_\mu S_1'} = -\frac{1}{2m} > -2.$$

The inequality

$$E_\mu N(c) \geqslant T\left(n_\mu - M_1 T^{-1} \log T\right)$$

where M_1 is a constant which is independent of μ and c, $c \leqslant c_0$, follows immediately from (2) and (4).

If $0 \leqslant \mu < \eta/2\gamma$, the line $y = \mu x - \eta T$ intersects the boundary of $B_0 T(1 - \varepsilon)$ on the curve $y = x - (2\gamma T x)^{\frac{1}{2}}$ so that we do not have $N(c) \geqslant N_L$ on A with probability one. Let x_3 be the larger of the x-coordinates of the points of intersection of $y = \mu x - \eta T$ with $y = x - (2\gamma T x)^{\frac{1}{2}}$. Then

$$x_3 = T\left(n_\mu \gamma - \frac{\eta}{1 - \mu} + \frac{\gamma}{(1 - \mu)^2}\left[(1 - 2\eta(1 - \mu)\gamma^{-1})^{\frac{1}{2}} - 1\right]\right) \geqslant x_1$$

so that $N(c) \geqslant x_1$ on A with probability one. Hence, (1) and the fact that

$N(c) \leqslant 2\gamma T$ imply that

$$\begin{aligned} E_\mu N(c) &\geqslant E_\mu[N(c)|A] - 2\gamma T P_\mu(A^c) \\ &\geqslant x_1 - 2\gamma D_1 \\ &\geqslant T\left(n_\mu \gamma - \frac{2\eta}{1+\mu} - \frac{2\eta^2}{\gamma^2}\right) - 2\gamma D_1 \\ &\geqslant T\left(n_\mu - M_2(T^{-1}\log T)^{\frac{1}{2}}\right), \end{aligned}$$

where M_2 is a constant independent of μ and c for $c \leqslant c_0$. This establishes Lemma 1.

LEMMA 2. *Let $N(c)$ be the first time (n, S_n) exits $B_0 T(1 + \varepsilon)$ where $\varepsilon = T^{-1}\log T$. For any μ in Ω,*

$$E_\mu N(c) \leqslant T\left(n_\mu + KT^{-1}\log T\right),$$

where K is a constant independent of μ and c.

PROOF. Let L be the line through the point $(T(1 + \varepsilon)n_\mu, T(1 + \varepsilon)\mu n_\mu)$, tangent to the boundary of $B_0 T(1 + \varepsilon)$. Consider the case $\mu \geqslant 0$. The line L has equation

$$y = \tfrac{1}{2}(\mu - 1)x + T\frac{(1+\varepsilon)}{1+\mu}.$$

Let N_L be the first time n, that $S_n - \frac{1}{2}(\mu - 1)n > T(1 + \varepsilon)/(1 + \mu)$. Then

$$E_\mu\left(S_1 - \tfrac{1}{2}(\mu - 1)\right) = \tfrac{1}{2}(\mu + 1).$$

Since $E_\mu N_L < \infty$ and $E_\mu|X_1 - \frac{1}{2}(\mu - 1)| < \infty$, we may again use Wald's equation to get

$$E_\mu N_L = n_\mu T(1+\varepsilon) + \frac{2}{\mu+1} E_\mu\left[S'_{N_L} - \frac{T(1+\varepsilon)}{1+\mu}\right].$$

Applying Theorem 1 of Lorden (1970) we get

$$\begin{aligned} E_\mu\left[S'_{N_L} - \frac{T(1+\varepsilon)}{1+\mu}\right] &\leqslant \frac{2}{1+\mu} E_\mu(Y_1^+)^2 \\ &\leqslant \frac{2}{1+\mu}\left[1 + \left(\frac{\mu+1}{2}\right)^2\right], \end{aligned}$$

so that

$$\begin{aligned} E_\mu N_L &\leqslant n_\mu T(1+\varepsilon) + 5 \\ &\leqslant T\left(n_\mu + KT^{-1}\log T\right) \end{aligned}$$

where K is a constant independent of μ and c, for $\mu \geqslant 0$. Since $E_\mu N(c) \leqslant E_\mu N_L$, the desired result follows for $\mu \geqslant 0$.

The case $\mu < 0$ is analogous.

LEMMA 3. *Given $\varepsilon > 0$, there exists a positive number c^* such that for $c < c^*$ and all real k,*

$$\frac{u_k(c)}{T} < n_0 + \varepsilon = 2 + \varepsilon.$$

PROOF. This lemma follows from Lemmas 3.5 and 3.6 in Wong (1968).

LEMMA 4. *Let $\Omega'' = \{\mu : |\mu| + 1 < \frac{1}{4}[(1-\varepsilon)T/\log T]^{\frac{1}{2}} + \eta T\}$ where $\varepsilon = 2\log T/T$ and $\eta = [2(1-\varepsilon)\varepsilon]^{\frac{1}{2}}$. Let $\mathcal{K}(c) = \{(n, S_n) : S_n/n \in \Omega'', n$ is a positive integer$\}$. Then for sufficiently small c,*

$$B_0T(1-\varepsilon) \cap \mathcal{K}(c) \subset C(c) \cap \mathcal{K}(c).$$

PROOF. The boundary of $C(c)$ is defined by the two equations

$$\textstyle\int_{H_i} \exp(\theta S_n - n\theta^2/2)l(\theta)g(\theta)d\theta \int_\Omega \exp(\theta S_n - n\theta^2/2)g(\theta)d\theta = c,$$

$i = 0, 1$. Suppose $n^{-1}S_n = k$ with $k \geqslant 0$. Let $f(\theta) = \theta k - \theta^2/2$. Then $f(\theta) = f(-1) + (\theta+1)(k+1) - (\theta+1)^2/2$, and

$$\begin{aligned}\textstyle\int_{H_0} &\exp(\theta S_n - n\theta^2/2)l(\theta)g(\theta)d\theta \\ &= \textstyle\int_{-\infty}^{-1} \exp\{n[f(-1) + (\theta+1)(k+1) - (\theta+1)^2/2]\}l(\theta)g(\theta)d\theta.\end{aligned}$$

By the assumptions on g, there exists a $\rho > 0$ such that $g(\theta) \geqslant \rho$ whenever $\theta \in [a, -1]$, for some a satisfying $-2 < a < -1$. Let $\mu_1 = 4^{-1}[(1-\varepsilon)T/\log T]^{\frac{1}{2}} + \eta T - 1$. For $k \in \omega''$ and T sufficiently large so that $\mu_1 > \frac{3}{2}$,

$$\begin{aligned}\textstyle\int_{H_0} &\exp(\theta S_n - n\theta^2/2)l(\theta)g(\theta)d\theta \\ &\geqslant \exp[nf(-1)]L\rho \textstyle\int_a^{-1} \exp[n(\theta+1)(\mu_1 + (1-a)/2)]d\theta \\ &= \frac{\exp(nf(-1))L\rho}{n[\mu_1 + (1-a)/2]} \int_{n(a+1)(\mu_1+(1-a)/2)}^0 \exp(y)dy \\ &\geqslant \frac{\exp(nf(-1))L\rho}{2n\mu_1} \int_{(a+1)/2}^0 \exp(y)dy \\ &\geqslant \frac{\exp(nf(-1))}{n\mu_1} A_1,\end{aligned}$$

where A_1 depends only on g and l. We can also represent f as $f(\theta) = f(k) - (\theta - k)^2/2$, so that

$$\begin{aligned}\textstyle\int_\Omega \exp(\theta S_n - n\theta^2/2)g(\theta)\,d\theta &\leqslant \sup_{\theta\in\Omega} g(\theta)\exp(nf(k))\textstyle\int_\Omega \exp[-n(\theta-k)^2/2]\,d\theta \\ &= \sup_{\theta\in\Omega} g(\theta)\exp[nf(k)][n^{-\frac{1}{2}}\textstyle\int_{-\infty}^{+\infty}\exp(-y^2/2)\,dy] \\ &= n^{-\frac{1}{2}}A_2\exp[nf(k)],\end{aligned}$$

where A_2 depends only on g.

Thus, there exist constants C_i, $i = 1, 2$, which depend only on g and l, such that on the boundary of $C(c) \cap \mathcal{K}(c)$, we have

$$\log c \geqslant nf(-1) - \log n + C_1 - \log \mu_1 - nf(k) + \tfrac{1}{2}\log n + C_2$$

for $k \geqslant 0$. That is,

$$u_k(c)[f(k) - f(-1)] \geqslant T - \tfrac{1}{2}\log T - \tfrac{1}{2}\log(u_k(c)T^{-1}) - \log \mu_1 + (C_1 + C_2). \tag{5}$$

Now using Lemma 3 we get the right-hand side of (5)

$$\geqslant T - \tfrac{1}{2}\log T - \tfrac{1}{2}\log(n_0 + 1) - \log \mu_1 + (C_1 + C_2)$$

$$\geqslant T - \tfrac{1}{2}\log T - \tfrac{1}{2}\log C_3 - \log C_4 T$$

for c less than the c^* of Lemma 3, where C_3 and C_4 are constants depending only on g and l. Thus,

$$u_k(c)[f(k) - f(-1)] \geqslant T - 2\log T \tag{6}$$

for T sufficiently large, independent of k in Ω'', $k \geqslant 0$. Analogously, it can be shown that (6) holds for k in Ω'', $k < 0$.

Since $f(k) - f(-1) = 1/n_k$, we have

$$u_k(c) \geqslant n_k T(1 - 2T^{-1}\log T)$$

for sufficiently large T independent of k in Ω''. (T depends only on g, l and c_0). Therefore,

$$B_0 T(1 - \varepsilon) \cap \mathcal{K}(c) \subset C(c) \cap \mathcal{K}(c)$$

as desired.

LEMMA 5. *Let $\delta(c)$ be the procedure which stops the first time (n, S_n) exits $B_0T(1 + \varepsilon)$ and decides H_1 (respectively H_0) if $S_n \geqslant -n + (2T(1 + \varepsilon)n)^{\frac{1}{2}}$ (respectively $\leqslant n - (2T(1 + \varepsilon)n)^{\frac{1}{2}}$), where $\varepsilon = T^{-1}\log T$. Then there exists a constant B such that*

$$P_\mu[\delta(c) \text{ makes an error}] \leqslant Bc$$

for all μ in Ω.

PROOF. Suppose $\mu \leqslant -1$ and Y_i, $i = 1, 2, \cdots$, are independent and identically distributed random variables, with $Y_1 \sim N(0, 1)$. Then, with B_1 and B constants, and using the fact that $c^\varepsilon = T^{-1}$, we have

$$\begin{aligned} P_\mu[\delta(c) \text{ makes an error}] &\leqslant P_{-1}\left[\max_{1 \leqslant n \leqslant 2T(1+\varepsilon)} \frac{S_n + n}{n^{\frac{1}{2}}} \geqslant (2T(1 + \varepsilon))^{\frac{1}{2}}\right] \\ &\leqslant \Sigma_{n=1}^{2T(1+\varepsilon)} P\left[Y_n \geqslant (2T(1 + \varepsilon))^{\frac{1}{2}}\right] \\ &\leqslant B_1 T(1 + \varepsilon)\left[(T(1 + \varepsilon))^{-\frac{1}{2}} e^{-T(1+\varepsilon)}\right] \\ &\leqslant BT^{\frac{1}{2}} c^{1+\varepsilon} \\ &\leqslant Bc. \end{aligned}$$

The case $\mu \geqslant 1$ is analogous.

LEMMA 6. *Let $N(c)$ be as in Lemma 2. For sufficiently small c,*

$$\int_\Omega E_\mu N(c) W(d\mu) \geqslant M(g)T,$$

where $M(g)$ is a constant depending on g.

PROOF. Suppose $0 < \delta < \frac{1}{2}$. For $\mu \in \Omega$, Chebyshev's inequality yields

$$E_\mu N(c) \geqslant \delta T(1+\varepsilon) P_\mu[N(c) > \delta T(1+\varepsilon)]. \tag{7}$$

Let Y be a normal random variable such that $EY = 0$ and $EY^2 = 1$. For $-1 \leqslant \mu \leqslant 1$ and $1 \leqslant n \leqslant \delta(1+\varepsilon)T$, we have

$$\begin{aligned}
P_\mu\left[S_n \geqslant -n + (2nT(1+\varepsilon))^{\frac{1}{2}}\right] &= P_\mu\left[\frac{S_n + n}{n^{\frac{1}{2}}} \geqslant (2T(1+\varepsilon))^{\frac{1}{2}}\right] \\
&= P\left[Y \geqslant (2T(1+\varepsilon))^{\frac{1}{2}} - (1+\mu)n^{\frac{1}{2}}\right] \\
&\leqslant P\left[Y \geqslant (2T(1+\varepsilon))^{\frac{1}{2}}(1 - (2\delta)^{\frac{1}{2}})\right] \\
&\leqslant MT^{-\frac{1}{2}} e^{-T(1+\varepsilon)(1-(2\delta)^{\frac{1}{2}})^2} \\
&= MT^{-\frac{1}{2}}(cT^{-1})^{(1-(2\delta)^{\frac{1}{2}})^2}
\end{aligned}$$

where M is a constant.

Letting $r = \frac{1}{2}(1 - (2\delta)^{\frac{1}{2}})^{-2}$, we then get

$$\begin{aligned}
P_\mu\left[\max_{1 \leqslant n \leqslant \delta(1+\varepsilon)T} \frac{S_n + n}{n^{\frac{1}{2}}} \geqslant (2T(1+\varepsilon))^{\frac{1}{2}}\right] &\leqslant \delta(1+\varepsilon) M T^{\frac{1}{2}}(cT^{-1})^{(1-(2\delta)^{\frac{1}{2}})^2} \\
&= \delta(1+\varepsilon) M (cT^r T^{-1})^{(1-(2\delta)^{\frac{1}{2}})^2} \\
&\to 0 \text{ as } c \to 0,
\end{aligned}$$

where we have used the fact that $cT^\beta \to 0$ as $c \to 0$ for all β.

Similarly,

$$P_\mu\left[\inf_{1 \leqslant n \leqslant \delta(1+\varepsilon)T} \frac{S_n - n}{n^{\frac{1}{2}}} \leqslant -(2T(1+\varepsilon))^{\frac{1}{2}}\right] \to 0$$

as $c \to 0$.

Therefore, for $-1 \leqslant \mu \leqslant 1$,

$$\begin{aligned}
P_\mu[N(c) \leqslant \delta(1+\varepsilon)T] &\leqslant P_\mu\left[\max_{1 \leqslant n \leqslant \delta(1+\varepsilon)T} \frac{S_n + n}{n^{\frac{1}{2}}} \geqslant (2T(1+\varepsilon))^{\frac{1}{2}}\right] \\
&\quad + P_\mu\left[\inf_{1 \leqslant n \leqslant \delta(1+\varepsilon)T} \frac{S_n - n}{n^{\frac{1}{2}}} \leqslant -(2T(1+\varepsilon))^{\frac{1}{2}}\right] \qquad (8) \\
&\leqslant \tfrac{1}{2},
\end{aligned}$$

for small c, uniformly in μ.

From (7) and (8) we get

$$E_\mu N(c) \geqslant \tfrac{1}{2}\delta T(1+\varepsilon)$$

and

$$\begin{aligned}\textstyle\int_\Omega E_\mu N(c) W(d\mu) &\geqslant \textstyle\int_{[-1,1]} E_\mu N(c) W(d\mu)\\ &\geqslant M(g)T,\end{aligned}$$

where $M(g) = \frac{1}{2}\delta \int_{[-1,1]} g(\mu) d\mu$. This establishes Lemma 6.

We now prove the main result. We shall let $r(W, \delta)$ denote the integrated risk of a procedure δ with respect to an a priori distribution W.

THEOREM. *Let $\delta_W^*(c)$ denote a Bayes procedure with respect to W and c, and let $\delta(c)$ be the procedure defined in Lemma* 5. *Then*

$$\frac{r(W, \delta_W^*(c))}{r(W, \delta(c))} = 1 - 0\left(\frac{\log\log c^{-1}}{\log c^{-1}}\right)$$

as $c \to 0$.

PROOF. From Schwarz (1968) we know that if $B(c)$ denotes the Bayes continuation region, then

$$B(c) \supset C(\Delta c \log c^{-1}) \tag{9}$$

for some constant Δ. Let $\tilde{c} = \Delta c \log c^{-1}$, and let $\tilde{T}$ and $\tilde{\eta}$ be derived from T and η by replacing c by $\tilde{c}$, where η is as defined in Lemma 1. From (9) and Lemma 4 with ε replaced by $\varepsilon_1 = 2\tilde{T}^{-1}\log\tilde{T}$, we have

$$B(c) \cap \mathcal{K}(\tilde{c}) \supset C(\tilde{c}) \cap \mathcal{K}(\tilde{c}) \supset B_0\tilde{T}(1-\varepsilon_1) \cap \mathcal{K}(\tilde{c}). \tag{10}$$

Let A be the event that (n, S_n) is in $\mathcal{K}(\tilde{c})$ for all $n \leqslant 2(1-\varepsilon_1)\tilde{T}''$ and let $W(t)$ denote a standard Wiener process. For

$$|\mu| < \mu_0 = \tfrac{1}{4}\left(\frac{(1-\varepsilon_1)\tilde{T}}{\log\tilde{T}}\right)^{\frac{1}{2}} - 1,$$

we have, as was seen in the proof of Lemma 1,

$$\begin{aligned}P_\mu(A^c) &= P_\mu\left[\max_{1\leqslant n\leqslant 2(1-\varepsilon_1)\tilde{T}}\frac{|S_n|}{n} \geqslant \mu_0 + \tilde{\eta}\tilde{T}\right]\\ &\leqslant P_\mu\left[\max_{1\leqslant n\leqslant 2(1-\varepsilon_1)T}\frac{|S_n|}{n} \geqslant \mu + \tilde{\eta}\tilde{T}\right]\\ &\leqslant P_\mu\left[\max_{1\leqslant n\leqslant 2(1-\varepsilon_1)T}|S_n - \mu n| \geqslant \tilde{\eta}\tilde{T}\right] \qquad (11)\\ &\leqslant P\left[\sup_{0< t\leqslant 2(1-\varepsilon_1)T}|W(t)| \geqslant \tilde{\eta}\tilde{T}\right]\\ &\leqslant (\text{constant})\tilde{T}^{-1}.\end{aligned}$$

Let $N(\tilde{c})$ be the first time (n, S_n) exits $B_0\tilde{T}(1-\varepsilon_1)$. The definition of B_0 then implies that $N(\tilde{c}) \leqslant 2(1-\varepsilon_1)\tilde{T}$ with probability one. Using this fact together with

(11) and Lemma 1, we get, for $\mu \in \tilde{\Omega}'$ and $|\mu| \geqslant \tilde{\eta}/2(1-\varepsilon_1)$,

$$\begin{aligned} E_\mu[N(\tilde{c})|A]P_\mu[A] &= E_\mu N(\tilde{c}) - E_\mu[N(\tilde{c})|A^c]P_\mu[A^c] \\ &\geqslant \tilde{T}\big(n_\mu - M_1\tilde{T}^{-1}\log\tilde{T}\big) - 2(1-\varepsilon_1)(\text{constant}) \\ &\geqslant \tilde{T}\big(n_\mu - \tilde{M}_1\tilde{T}^{-1}\log\tilde{T}\big), \end{aligned} \tag{12}$$

where $\tilde{M}_1$ is independent of c and μ.

Similarly, for $|\mu| < \tilde{\eta}/2(1-\varepsilon_1)$, we get

$$E_\mu[N(\tilde{c})|A]P_\mu[A] \geqslant \tilde{T}\Big(n_\mu - \tilde{M}_2(\tilde{T}^{-1}\log\tilde{T})^{\frac{1}{2}}\Big), \tag{13}$$

where $\tilde{M}_2$ is independent of c and μ.

Let $N^*(c)$ and $N(c)$ denote the stopping times of the procedures $\delta_W^*(c)$ and $\delta(c)$, respectively. Lemma 2, (10) and (12) imply that for $\tilde{\eta}/2(1-\varepsilon_1) \leqslant |\mu| < \mu_0$,

$$\begin{aligned} E_\mu N(c) - E_\mu N^*(c) &\leqslant E_\mu N(c) - E_\mu[N^*(c)|A]P_\mu[A] \\ &\leqslant E_\mu N(c) - E_\mu[N(\tilde{c})|A]P_\mu[A] \\ &\leqslant (T-\tilde{T})n_\mu + K\log T + \tilde{M}_1\log\tilde{T}. \end{aligned}$$

Also, for $|\mu| < \tilde{\eta}/2(1-\varepsilon_1)$, Lemma 2, (10) and (13) imply that

$$E_\mu N(c) - E_\mu N^*(c) \leqslant (T-\tilde{T})n_\mu + K\log T + \tilde{M}_2(\tilde{T}\log\tilde{T})^{\frac{1}{2}}.$$

Note that

$$T - \tilde{T} = T - (\log\Delta^{-1} + T - \log T) = \log T - \log\Delta^{-1} = O(\log T),$$

and

$$\log\tilde{T} = \log(\log\Delta^{-1} + T - \log T) = O(\log T).$$

Hence, since $n_\mu \leqslant 2$, there exist constants K_1 and K_2 depending only on g and l, such that

$$E_\mu N(c) - E_\mu N^*(c) \leqslant K_1\log T, \tag{14}$$

for sufficiently small c independent of μ, when $\tilde{\eta}/2(1-\varepsilon_1) \leqslant |\mu| < \mu_0$. Also

$$E_\mu N(c) - E_\mu N^*(c) \leqslant K_2\log T + \tilde{M}_2(\tilde{T}\log\tilde{T})^{\frac{1}{2}}, \tag{15}$$

for sufficiently small c independent of μ, when $|\mu| < \tilde{\eta}/2(1-\varepsilon_1)$.

For $|\mu| \geqslant \mu_0$ we have from Lemma 2,

$$\begin{aligned} E_\mu N(c) - E_\mu N^*(c) &\leqslant E_\mu N(c) \\ &\leqslant T\left(\frac{2}{(|\mu|+1)^2} + KT^{-1}\log T\right) \\ &\leqslant T\big(K_3\tilde{T}^{-1}\log\tilde{T} + KT^{-1}\log T\big) \\ &< (K_3+K)\log T, \end{aligned} \tag{16}$$

where K_3 is a constant.

Let $e(W, \delta)$ denote the integrated risk due to error. Lemma 5 yields

$$\begin{aligned} r(W, \delta(c)) - r(W, \delta_W^*(c)) &= c\int_\Omega [E_\mu N(c) - E_\mu N^*(c)] W(d\mu) + e(W, \delta(c)) \\ &\quad - e(W, \delta_W^*(c)) \\ &\leqslant c\int_\Omega [E_\mu N(c) - E_\mu N^*(c)] W(d\mu) + Bc. \end{aligned} \tag{17}$$

We now have, by (14), (15), (16), (17) and Lemma 6

$$\begin{aligned} \frac{r(W, \delta(c)) - r(W, \delta_W^*(c))}{r(W, \delta(c))} &\leqslant \frac{1}{M(g)T}\left[\int_{|\mu| \leqslant \tilde{\eta}/2(1-\varepsilon_1)} (E_\mu N(c) - E_\mu N^*(c)) W(d\mu) \right. \\ &\qquad \left. + (K_1 + K_3 + K) \log T + B\right] \\ &\leqslant \frac{1}{M(g)T}\left[(\sup_{\theta \in \Omega} g(\theta))\left(K_2 \log T + \tilde{M}_2(\tilde{T} \log \tilde{T})^{\frac{1}{2}}\right)\right. \\ &\qquad \left. \times \frac{\tilde{\eta}}{1 - \varepsilon_1} + (K_1 + K_3 + K) \log T + B\right] \\ &\leqslant \frac{1}{M(g)T}\left[(\sup_{\theta \in \Omega} g(\theta))\left(K_2 \log T + \frac{2\tilde{M}_2}{(1 - \varepsilon_1)^{\frac{1}{2}}} \log \tilde{T}\right)\right. \\ &\qquad \left. + (K_1 + K_3 + K) \log T + B\right] \\ &\leqslant (\text{constant}) T^{-1} \log T, \end{aligned}$$

where the constant depends on g and l. Thus,

$$\frac{r(W, \delta(c)) - r(W, \delta_W^*(c))}{r(W, \delta(c))} = O\left(\frac{\log \log c^{-1}}{\log c^{-1}}\right)$$

as desired.

ACKNOWLEDGMENT. The author wishes to thank Professor Jack C. Kiefer for many helpful discussions.

REFERENCES

[1] DOOB, J. L. (1963). *Stochastic Processes*. Wiley, New York.
[2] KINGMAN, J. F. C. (1962). Some inequalities for the queue GI/G/1. *Biometrika* **49** 315–324.
[3] LORDEN, G. (1967). Integrated risk of asymptotically Bayes sequential tests. *Ann. Math. Statist.* **38** 1399–1422.
[4] LORDEN, G. (1970). On excess over the boundary. *Ann. Math. Statist.* **41** 520–527.
[5] SCHWARZ, G. (1962). Asymptotic shapes of Bayes sequential testing regions. *Ann. Math. Statist.* **33** 224–236.

[6] SCHWARZ, G. (1969). A second-order approximation to optimal sampling regions. *Ann. Math. Statist.* **40** 313–315.

[7] WALD, A. (1946). Differentiation under the expectation sign in the fundamental identity of sequential analysis. *Ann. Math. Statist.* **17** 493–497.

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

[9] WONG, S. P. (1968). Asymptotically optimum properties of certain sequential tests. *Ann. Math. Statist.* **39** 1244–1263.

1960 CHAPEL ST.
NEW HAVEN, CT 06515