The Annals of Statistics
1986, Vol. 14, No. 3, 1030–1048

THE SHAPE OF BAYES TESTS OF POWER ONE[1]

BY HANS RUDOLF LERCHE

University of Heidelberg

The problem of determining Bayes tests of power one (without an indifference zone) is considered for Brownian motion with unknown drift. When we let the unit sampling cost depend on the underlying parameter in a natural way, it turns out that a simple Bayes rule is approximately optimal. Such a rule stops sampling when the posterior probability of the hypothesis is too small.

1. Introduction. Let $W(t)$ denote Brownian motion with unknown drift $\theta \in \mathbb{R}$ and P_θ the associated measure. We consider the following sequential decision problem. Let F be prior on $\mathbb{R}$ given by $F = \gamma\delta_0 + (1-\gamma)\int\cdot\phi(\sqrt{r}\,\theta)\sqrt{r}\,d\theta$ with $0 < \gamma < 1$ and $\phi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$, consisting of a point mass at $\{\theta = 0\}$ and a smooth normal part on $\{\theta \neq 0\}$. Let the sampling cost be $c\theta^2$, with $c > 0$ for the observation of W per unit time when the underlying measure is P_θ. We assume also a loss function which is equal to 1 if $\theta = 0$ and we decide in favor of "$\theta \neq 0$" and which is identically 0 if $\theta \neq 0$. A statistical test consists of a stopping time T of Brownian motion where stopping means a decision in favor of "$\theta \neq 0$."

The Bayes risk for this problem is then given by

$$\rho(T) = \gamma P_0(T < \infty) + (1-\gamma)c\int_{-\infty}^{\infty}\theta^2 E_\theta T\phi(\sqrt{r}\,\theta)\sqrt{r}\,d\theta. \tag{1.1}$$

In this paper we investigate the "optimal" stopping rule T_c^* which minimizes $\rho(T)$.

For the cost c sufficiently small, T_c^* is a test of power one for the decision problem H_0: $\theta = 0$ versus H_1: $\theta \neq 0$. This is by definition a stopping time T which satisfies the conditions

$$P_0(T < \infty) < 1, \tag{1.2}$$

$$P_\theta(T < \infty) = 1 \quad \text{if } \theta \neq 0. \tag{1.3}$$

Here stopping also means a decision in favor of "$\theta \neq 0$." For a discussion of tests of power one see Robbins (1970). A similar problem has been studied by Pollak (1978) who assumed an indifference zone in the parameter space. The type of prior assumed here was once proposed by Jeffreys (1948).

A basic idea of this paper is to let the sampling cost depend on the underlying parameter in a natural way. At the first view the cost term "$c\theta^2$" has an unusual

Received May 1984; revised October 1985.

[1]This work was supported by the Deutsche Forschungsgemeinschaft and the National Science Foundation at MSRI, Berkeley, under NSF Grant MCS81-20790.

AMS 1980 *subject classifications*. Primary 62L15; secondary 62C10.

Key words and phrases. Sequential Bayes tests for composite hypotheses, tests of power one, simple Bayes rules.

structure. The factor $\theta^2/2$ is the Kullback–Leibler information number $E_\theta \log(dP_{\theta,1}/dP_{0,1})$ which quantifies the separability of the measures P_0 and P_θ. Its meaning becomes apparent by the following consideration. Let us consider two testing problems with simple hypotheses:

(1) H_0: $\theta = 0$ versus H_1: $\theta = \theta_1$,
(2) H_0: $\theta = 0$ versus H_1: $\theta = \theta_2$

with $\theta_i > 0$, $i = 1,2$. Let t_i, $i = 1,2$ denote the sampling lengths. Then the level-α Neyman–Pearson tests for both problems have the same error probabilities if and only if $\theta_1^2 t_1 = \theta_2^2 t_2$. [This follows from the power function of a Neyman–Pearson test of level α: $\Phi(-c_\alpha + \theta\sqrt{t})$.] Thus the factor θ^2 standardizes the sampling lengths in such a way that the embedded simple testing problems are of equal difficulty. Beside this statistical aspect there is a basic mathematical reason for this choice of the sampling costs. Since in our decision problem (1.1) an indifference zone does not occur and since $E_0 T = \infty$ [as $P_0(T < \infty) < 1$] we have $\lim_{\theta \to 0} E_\theta T = \infty$. More information about the singularity is provided by a lemma of Darling and Robbins (1967) [see also Robbins and Siegmund (1973) and Wald (1947), page 197]. It states that for every stopping time T with $P_0(T < \infty) < 1$

$$E_\theta T \geq 2b/\theta^2, \quad \text{where } b = -\log P_0(T < \infty). \tag{1.4}$$

Equality in (1.4) holds for the special stopping rule

$$T_\theta = \inf\left\{ t > 0 \,\middle|\, \frac{dP_{\theta,t}}{dP_{0,t}} \geq e^b \right\}. \tag{1.5}$$

Here $dP_{\theta,t}/dP_{0,t}$ denotes the likelihood ratio (Radon-Nikodym derivative) of P_θ with respect to P_0 given the path $W(u)$, $0 \leq u \leq t$. It is given by

$$\frac{dP_{\theta,t}}{dP_{0,t}} = \exp\left(\theta W(t) - \tfrac{1}{2}\theta^2 t\right).$$

According to (1.4) the expected sample size $E_\theta T$ of a test of power one, considered as a function of θ has a pole at $\theta = 0$. The choice of "c" or "$c|\theta|$" instead of "$c\theta^2$" would imply that tests of power one have an infinite Bayes risk since

$$\int |\theta|^i E_\theta T \phi(\sqrt{r}\,\theta)\sqrt{r}\, d\theta = \infty \quad \text{for } i = 0, 1.$$

A precise description of the pole of $E_\theta T$ is given by Robbins and Siegmund (1973) and Jennen and Lerche (1982). The sampling costs "$c\theta^2$" remove the nonintegrability of the singularity of $E_\theta T$ for a large class of tests of power one, although $\lim_{\theta \to 0} \theta^2 E_\theta T = \infty$ still holds [by the corollary on page 102 of Robbins and Siegmund (1973)]. For instance for all tests of power one defined by

$$T = \inf\left\{ t > 0 \,\middle|\, |W(t)| \geq \psi(t) \right\},$$

where the function $\psi(t)$ is concave and $\psi(t) = o(t^{2/3-\varepsilon})$ when $t \to \infty$ (with $\varepsilon > 0$

arbitrary small), the Bayes risk (1.1) is finite. This follows from the inequality $|\theta|E_\theta T \leq \psi(E_\theta T)$, which is a consequence of Wald's lemma and Jensen's inequality. Therefore by the choice of the sampling costs as "$c\theta^2$" the concept of Bayes tests of power one becomes an interesting topic to study.

The related problem for simple hypotheses can be solved easily. The Bayes risk given by

$$\rho(T) = \gamma P_0(T < \infty) + (1 - \gamma)c\theta^2 E_\theta T, \tag{1.6}$$

using statement (1.4), is minimized by the stopping rule

$$T_c^* = \inf\left\{t > 0 \middle| W(t) \geq \log a/\theta + \tfrac{1}{2}\theta t\right\} \tag{1.7}$$

with $a = \gamma(2(1 - \gamma)c)^{-1}$ provided $a > 1$. In this case the minimal Bayes risk is given by

$$\rho(T_c^*) = 2(1 - \gamma)c[\log a + 1]. \tag{1.8}$$

When $a \leq 1$, $T_c^* = 0$, and $\rho(T_c^*) = \gamma$. (For more details see the end of the proof of Theorem 2.) Here the choice of the sampling costs leads to a solution not depending on θ. This becomes obvious when one expresses T_c^* in another way. It can be rewritten as

$$T_c^* = \inf\left\{t > 0 \middle| \gamma(W(t), t) \leq \frac{2c}{1 + 2c}\right\},$$

where

$$\gamma(x, t) = \frac{\gamma}{\gamma + (1 - \gamma)\dfrac{dP_{\theta, t}}{dP_{0, t}}(x)}$$

denotes the posterior mass of the parameter "0" at (x, t) with respect to the prior $F = \gamma\delta_0 + (1 - \gamma)\delta_\theta$. Thus T_c^* has the intuitive meaning "stop when the posterior mass of the hypothesis "0" is too small". This is a simple Bayes rule or equivalently the one-sided sequential probability ratio test (1.5).

The following study shows that a simple Bayes rule which stops when the posterior probability of the hypothesis "$\theta = 0$" is too small, is approximately optimal for the risk (1.1). For precise statements see the Theorems 2 and 3 and the corollaries. This simple Bayes rule is of the type (1.5) with a boundary equal to

$$\psi(t) = \left((t + r)\left(\log\left(\frac{t + r}{r}\right) + 2\log b\right)\right)^{1/2}, \quad \text{where } b = \frac{\gamma}{2(1 - \gamma)c}.$$

For large t this boundary asymptotically grows like $(t \log t)^{1/2}$, which is faster than the limiting growth rate $(2t \log\log t)^{1/2}$ of the law of the iterated logarithm. As a consequence of our results the minimal Bayes risk can be approximated by that of simple Bayes rules within $o(c)$ when $c \to 0$ (Theorem 4). Simple Bayes rules for the same type of prior which we use (Jeffreys' priors) were already discussed by Cornfield (1966).

Similar results hold for exponential families with general priors, although one has to make a careful analysis of the overshoot effect following the ideas of

Lorden (1977) to derive an $o(c)$-approximation for the minimal Bayes risk. The results will be published elsewhere. The proofs for the case of exponential families are more technical, since special approximation arguments are needed. The nice feature of the Brownian motion case is, that most expressions can be calculated exactly. Thus no approximations are needed and the proofs become simple.

This paper is organized as follows: Theorem 1 states the existence of an optimal (Bayes) stopping rule T_c^*. Theorem 2 gives upper and lower bounds for T_c^*, which makes it possible to derive its asymptotic shape when $c \to 0$ or $t \to \infty$. Theorem 3 refines these bounds, which yields the above mentioned $o(c)$-approximation of the minimal Bayes risk. Theorem 5 treats the one-sided case.

The results have some meaning for sequential clinical trials. These aspects are discussed in more detail in a subsequent paper. Historical facts are mentioned in Lerche (1985). A further result connected with the costs $c\theta^2$ is the exact Bayes property of the repeated significance test. For that see Lerche (1985, 1986).

2. Preliminaries. We need the following notations. The Brownian motion W with drift θ starting at time t in point x is understood as a measure $P_\theta^{(x,t)}$ on the space $C[t, \infty)$ of continuous functions on $[t, \infty)$. $\mathscr{F}_s^t$ denotes the σ-algebra on $C[t, \infty)$ which is generated by $W(u)$, $t \le u \le s$. The restriction of the measure $P_\theta^{(x,t)}$ on $\mathscr{F}_s^t$ is denoted by $P_{\theta,s}^{(x,t)}$. This notation is also used for stopping times S instead of fixed times s. When the process starts at 0 at time 0, then we very often skip the superindex and write just $\mathscr{F}_s$, $P_{\theta,s}$ etc. The Borel σ-algebra on the parameter space $\mathbb{R}$ is denoted by $\mathscr{B}$. For $F = \gamma\delta_0 + (1-\gamma)\int\cdot\sqrt{r}\,\phi(\sqrt{r}\,\theta)\,d\theta$, let $dP = dP_\theta F(\mathrm{d}\theta)$ and $d\bar{P} = d(\int P_\theta F(d\theta))$ be its projection. Let $F_{x,t}$ denote the posterior distribution given that the process $W(t) = x$. This means that for $A \times B \in \mathscr{F}_t \oplus \mathscr{B}$ $\int_A F_{W(t),t}(B)\bar{P}(dW) = P(A \times B)$ holds. Thus the Bayes risk (1.1) can be rewritten as

$$\rho(T) = \int_{\{T<\infty\}} \left(F_{W(T),T}(\{0\}) + cT\int_{-\infty}^{\infty} \theta^2 F_{W(T),T}(d\theta) \right) d\bar{P}. \tag{2.1}$$

Let $\bar{P}^{(x,t)}$ denote the conditional distribution of the process under $\bar{P}$ given $W(t) = x$. It can be represented as $\bar{P}^{(x,t)} = \int P_\theta^{(x,t)} F_{x,t}(d\theta)$.

We define the posterior risk at the space-time point (x, t) for a stopping rule $T \ge t$ as

$$\rho(x,t,T) = \int_{\{T<\infty\}} \Big(F_{W(T),T}(\{0\}) + c(T-t) \times \int_{-\infty}^{\infty} \theta^2 F_{W(T),T}(d\theta) \Big) d\bar{P}^{(x,t)}. \tag{2.2}$$

The minimal posterior risk at (x, t) is defined as

$$\rho(x,t) = \inf_T \rho(x,t,T), \tag{2.3}$$

where the infimum is taken over all stopping times of the process $(W(s), s)$

starting at (x, t), including $T_t \equiv t$. For T_t the risk is given by

$$\gamma(x, t) = \rho(x, t, T_t) = F_{x,t}(\{0\}) \tag{2.4}$$

and therefore the inequality $\rho(x, t) \leq \gamma(x, t)$ holds. The quantity $\rho(x, t, T) + ct\int\theta^2 F_{x,t}(d\theta)$ represents the loss when the process runs without stopping up to (x, t) and is stopped at $T \geq t$.

The following theorem states that an optimal (Bayes) stopping rule exists which minimizes (2.1) and characterizes it. Let $\mathscr{C}^*(c) = \{(y, s)|\rho(y, s) < \gamma(y, s)\}$ and

$$T_c^* = \inf\{s|(W(s), s) \notin \mathscr{C}^*(c)\}. \tag{2.5}$$

THEOREM 1. *The stopping rule $T_c^*(\geq t)$ of the space-time process $(W(t), t)$ minimizes the risk* (2.2) *for all starting points (x, t).*

This type of result is well known. Its statement is usually called the principle of dynamic programming. The result follows from the theory of optimal stopping for Markov processes [cf. Shiryayev (1978), page 127] applied to the space-time process $(W(t), t)$. We note that $W(t)$ under the measure $\bar{P}$ is a diffusion process which satisfies the stochastic differential equation $dW(t) = (1 - \gamma(W(t), t))W(t)/(t + r)\,dt + dX(t)$ where $X(t)$ is a standard Brownian motion [cf. Liptser and Shiryayev (1977), page 258].

The stopping risk can be calculated by (2.4) as

$$\gamma(x, t) = \frac{\gamma}{\gamma + (1 - \gamma)g(x, t)} \tag{2.6}$$

with

$$\begin{aligned} g(x, t) &= \int_{-\infty}^{\infty} \frac{dP_{\theta,t}}{dP_{0,t}}(x)\phi(\sqrt{r}\theta)\sqrt{r}\,d\theta \\ &= \int_{-\infty}^{\infty} \exp\left(\theta x - \tfrac{1}{2}\theta^2 t\right)\phi(\sqrt{r}\theta)\sqrt{r}\,d\theta \\ &= \sqrt{\frac{r}{t+r}}\exp\left(\frac{x^2}{2(t+r)}\right). \end{aligned} \tag{2.7}$$

We note that on $\{\theta \neq 0\}$

$$F_{x,t}(d\theta) = (1 - \gamma(x, t))G_{x,t}(d\theta), \tag{2.8}$$

where

$$G_{x,t} = N\left(\frac{x}{t+r}, \frac{1}{t+r}\right)$$

holds.

Here $N(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. The exact calculation of the minimal posterior risk $\rho(x, t)$ seems to be impossible for this problem. We can only derive upper and lower bounds for it. To get those we will rewrite the posterior risk in an appropriate form.

LEMMA 1.

$$\rho(x,t,T) = \gamma(x,t)P_0^{(x,t)}(t \le T < \infty) + (1-\gamma(x,t))c\int \theta^2 E_\theta^{(x,t)}(T-t)G_{x,t}(d\theta). \tag{2.9}$$

The posterior risk has the same form as the Bayes risk (1.1), with the slight difference that the process starts in the space-time point (x, t), stops at $T \ge t$, and has as prior $F_{x,t} = \gamma(x,t)\delta_0 + (1-\gamma(x,t))G_{x,t}$, the posterior at the point (x,t).

The proof of the lemma is a direct consequence of the preceding definitions and of the following basic fact about posterior distributions: the posterior of Brownian motion starting at (x, t) with prior $F_{x,t}$ at point $(W(S), S)$ is given by $F_{W(S),S}$.

3. Results for two-sided tests. The continuation region $\mathscr{C}^*(c)$ of the optimal stopping rule for the Bayes risk (1.1) is now approximated by upper and lower bounding regions of the space-time plane. These bounds are refined in Theorem 3. The bounding regions are given by sets of the type $\mathscr{C}(\lambda) := \{(x,t) | \gamma(x,t) > \lambda\}$.

THEOREM 2. *There exists a constant* $M > 2$ *such that for every* $c > 0$

$$\mathscr{C}\left(\frac{Mc}{1+Mc}\right) \subset \mathscr{C}^*(c) \subset \mathscr{C}\left(\frac{2c}{1+2c}\right) \tag{3.1}$$

holds.

REMARK 1. Let $T_\lambda = \inf\{t > 0 | (W(t), t) \notin \mathscr{C}(\lambda)\}$. Then (3.1) translates to

$$T_{Mc/(1+Mc)} \le T_c^* \le T_{2c/(1+2c)}.$$

REMARK 2. The theorem holds also for the more general prior

$$\sqrt{r}\,\phi\big(\sqrt{r}(\theta-\mu)\big)\,d\theta$$

by exactly the same arguments.

PROOF. At first we prove the lower inclusion of (3.1), which is the more difficult part. We show that for all points $(x,t) \in \mathscr{C}(Mc/(1+Mc))$ (M will be specified during the proof) there exist stopping times $S_{(x,t)}$ of the process $(W(s), s)$ starting at (x, t), such that

$$\rho\big(x,t,S_{(x,t)}\big) < \gamma(x,t) \tag{3.2}$$

holds.

Since by definition $\rho(x,t) \le \rho(x,t,S_{(x,t)})$, it follows from (3.2) and Theorem 1 that $(x,t) \in \mathscr{C}^*(c)$. We choose the stopping times as

$$S_{(x,t)} = \inf\{s > t | \gamma(W(s), s) \le Qc\},$$

where the constant $Q > 1$ will be defined below in such a way that $Qc < 1$. [In fact the stopping times $S_{(x,t)}$ all arise from the same stopping time T_{Qc} by changing the starting point of the process.] We need several representations of $S_{(x,t)}$ during the proof.

$$\begin{aligned} S_{(x,t)} &= \inf\{s > t \mid F_{W(s),s}(\{0\}) \le Qc\} \\ &= \inf\left\{s > t \,\middle|\, \int \frac{dP_{\theta,s}^{(x,t)}}{dP_{0,s}^{(x,t)}} G_{x,t}(d\theta) \ge b(x,t)\right\} \\ &= \inf\left\{s > t \,\middle|\, \sqrt{\frac{t+r}{s+r}} \exp\left(\frac{1}{2}\left(\frac{W(s)^2}{s+r} - \frac{x^2}{t+r}\right)\right) \ge b(x,t)\right\}, \end{aligned} \tag{3.3}$$

where

$$b(x,t) = \frac{\gamma(x,t)(1-Qc)}{(1-\gamma(x,t))Qc}.$$

The first equality holds by definition. The second equality follows by the calculation

$$F_{W(s),s}(\{0\}) = \frac{\gamma(x,t)}{\gamma(x,t) + (1-\gamma(x,t))h(x,t,w,s)}$$

with

$$h(x,t,w,s) = \int \frac{dP_{\theta,s}^{(x,t)}}{dP_{0,s}^{(x,t)}} G_{x,t}(d\theta).$$

The third equality follows by the following calculation [note here that $G_{x,t} = N(x/t+r), 1/(t+r)$],

$$\begin{aligned} &\int \frac{dP_{\theta,s}^{(x,t)}}{dP_{0,s}^{(x,t)}} G_{x,t}(d\theta) \\ &\quad = \int \exp\left(\theta(W(s)-x) - \tfrac{1}{2}\theta^2(s-t)\right)\sqrt{\frac{t+r}{2\pi}} \\ &\qquad \times \exp\left(-\frac{(t+r)}{2}\left(\theta - \frac{x}{t+r}\right)^2\right) d\theta \\ &\quad = \sqrt{(t+r)} \exp\left(-\frac{x^2}{2(t+r)}\right) \int \exp\left(\theta W(s) - \tfrac{1}{2}\theta^2(s+r)\right)\frac{d\theta}{\sqrt{2\pi}} \\ &\quad = \sqrt{\frac{t+r}{s+r}} \exp\left(\frac{1}{2}\left(\frac{W(s)^2}{s+r} - \frac{x^2}{t+r}\right)\right). \end{aligned}$$

We start now to estimate the posterior risk for $S_{(x,t)}$, which is given according to

Lemma 1 by

$$\rho(x, t, S_{(x,t)}) = \gamma(x, t) P_0^{(x,t)}(t < S_{(x,t)} < \infty) \tag{3.4}$$
$$+ (1 - \gamma(x, t)) c \int \theta^2 (S_{(x,t)} - t) \, dQ^{(x,t)}$$

with the new notation $Q^{(x,t)}(dW, d\theta) = P_\theta^{(x,t)}(dW) G_{x,t}(d\theta)$. We will also use $\bar{Q}^{(x,t)}(dW) = \int_\Theta P_\theta^{(x,t)}(dW) G_{x,t}(d\theta)$ and will from now on simply write S instead of $S_{(x,t)}$. Then we get for the first term

$$P_0^{(x,t)}(t < S < \infty) = \frac{Qc}{1 - Qc} \frac{1 - \gamma(x, t)}{\gamma(x, t)} = b(x, t)^{-1}. \tag{3.5}$$

This follows from a well known martingale argument [see Robbins and Siegmund (1970), Lemma 1] by using (3.3):

$$P_0^{(x,t)}(t < S < s_0) = \int_{\{t < S < s_0\}} \frac{dP_{0,S}^{(x,t)}}{d\bar{Q}_S^{(x,t)}} \, d\bar{Q}^{(x,t)}$$
$$= b(x, t)^{-1} \bar{Q}^{(x,t)}\{t < S < s_0\}.$$

Since $\bar{Q}^{(x,t)}\{t < S < s_0\} \to 1$ as $s_0 \to \infty$ (3.5) follows. We note that for $(x, t) \in \mathscr{C}(Qc)$, $b(x, t) > 1$ holds and that thus the probability in (3.5) is less than one.

To estimate the second term of (3.4) we rewrite the integral. Since on $\mathscr{F}_S^t$ we have

$$Q^{(x,t)}(dW, d\theta) = N\left(\frac{W(S)}{S + r}, \frac{1}{S + r}\right)(d\theta) \bar{Q}^{(x,t)}(dW),$$

we get for the integral

$$\int \theta^2 (S - t) \, dQ^{(x,t)} = \int \theta^2 (S + r) \, dQ^{(x,t)} - \int \theta^2 (t + r) \, dQ^{(x,t)}$$
$$= \int \int \theta^2 (S + r) N\left(\frac{W(S)}{S + r}, \frac{1}{S + r}\right)(d\theta) \, d\bar{Q}^{(x,t)} \tag{3.6}$$
$$- \int \theta^2 (t + r) G_{x,t}(d\theta)$$
$$= \int \left(\frac{W(S)^2}{S + r} - \frac{x^2}{t + r}\right) d\bar{Q}^{(x,t)}.$$

Using now the third form in (3.3) of the stopping rule S yields

$$\int \left(\frac{W(S)^2}{S + r} - \frac{x^2}{t + r}\right) d\bar{Q}^{(x,t)} = 2 \log b(x, t) + \int \log\left(\frac{S + r}{t + r}\right) d\bar{Q}^{(x,t)}. \tag{3.7}$$

Let $\alpha > 3$. We show now that there exists a constant $0 < C_\alpha < \infty$ with

$$\int \log\left(\frac{S+r}{t+r}\right) d\bar{Q}^{(x,t)} \le C_\alpha\left[\int \theta^2(S-t)\, dQ^{(x,t)}\right]^{1/\alpha}. \tag{3.8}$$

Then (3.6), (3.7), and (3.8) yield

$$\int \theta^2(S-t)\, dQ^{(x,t)} \le 2\log b(x,t) + C_\alpha\left[\int \theta^2(S-t)\, dQ^{(x,t)}\right]^{1/\alpha}, \tag{3.9}$$

from which one can derive (3.2), as will be explained below.

The proof of (3.8) runs as follows. By using the inequality $\log(1+x) \le K_\alpha x^{1/\alpha}$ for $x \ge 0$ we get by Hölder's inequality

$$\begin{aligned}
\int \log\left(\frac{S+r}{t+r}\right) d\bar{Q}^{(x,t)} &= \int \log\left(1 + \frac{S-t}{t+r}\right) d\bar{Q}^{(x,t)} \\
&\le K_\alpha \int \left(\frac{S-t}{t+r}\right)^{1/\alpha} d\bar{Q}^{(x,t)} \\
&= K_\alpha \int (\theta^2(S-t))^{1/\alpha}(\theta^2(t+r))^{-1/\alpha}\, dQ^{(x,t)} \\
&\le K_\alpha\left[\int \theta^2(S-t)\, dQ^{(x,t)}\right]^{1/\alpha} \\
&\quad \times\left[\int (\theta^2(t+r))^{-1/(\alpha-1)} G_{x,t}(d\theta)\right]^{(\alpha-1)/\alpha}
\end{aligned} \tag{3.10}$$

But since $G_{x,t} = N(x/(t+r), 1/(t+r))$ we get for $\alpha > 3$

$$\begin{aligned}
\int (\theta^2(t+r))^{-1/(\alpha-1)} G_{x,t}(d\theta) &\le \int (\theta^2(t+r))^{-1/(\alpha-1)} N\left(0, \frac{1}{t+r}\right)(d\theta) \\
&= \int y^{-2/(\alpha-1)} N(0,1)(dy) < \infty.
\end{aligned}$$

We now put

$$C_\alpha = K_\alpha\left[\int y^{-2/(\alpha-1)} N(0,1)(dy)\right]^{(\alpha-1)/\alpha}$$

and get finally (3.8) and (3.9).

Let $b > 1$ be given. Then by (3.9) there exists a constant $B > 2$ such that for all $b(x,t) \ge b$

$$\int \theta^2(S-t)\, dQ^{(x,t)} \le B\log b(x,t) \tag{3.11}$$

holds. Now we choose $Q = B/(1+Bc)$ and $M = bB$. Then for $(x,t) \in \mathscr{C}(Mc/(1+Mc))$ we have

$$b(x,t) = \frac{1-Qc}{Qc}\,\frac{\gamma(x,t)}{1-\gamma(x,t)} = \frac{\gamma(x,t)}{Bc(1-\gamma(x,t))} > \frac{Mc}{Bc} = b,$$

and by (3.4), (3.5), and (3.11) we get further

$$\begin{aligned}\rho(x,t,S) &\le \gamma(x,t)b(x,t)^{-1} + (1-\gamma(x,t))Bc\log b(x,t)\\ &= \gamma(x,t)b(x,t)^{-1}(1+\log b(x,t))\\ &< \gamma(x,t).\end{aligned}$$

The last inequality follows from the inequality $x(1+\log x^{-1}) < 1$ for $x < 1$ since $b(x,t) > b > 1$. This proves (3.2).

Now we prove the upper inclusion. We show

$$\rho(x,t) = \gamma(x,t) \quad \text{if } \gamma(x,t) \le \frac{2c}{1+2c}. \tag{3.12}$$

This implies the upper inclusion of statement (3.1). The method of proof consists in comparing the Bayes rule T_c^* with the best rule if θ were known.

For the Bayes rule T_c^* we always have

$$\begin{aligned}\gamma(x,t) \ge \rho(x,t) &= \rho(x,t,T_c^*)\\ &= \gamma(x,t)P_0^{(x,t)}(t \le T_c^* < \infty)\\ &\quad + (1-\gamma(x,t))c\int \theta^2(T_c^* - t)\,dQ^{(x,t)}\\ &= \int_{-\infty}^{\infty}\big[\gamma(x,t)P_0^{(x,t)}(t \le T_c^* < \infty)\\ &\qquad + (1-\gamma(x,t))c\theta^2 E_\theta^{(x,t)}(T_c^* - t)\big]G_{x,t}(d\theta).\end{aligned}$$

Let the process W start in x and let $W(u) = z$. Under the transformation $y = \theta(z-x)$, $s = \theta^2(u-t)$ Brownian motion with drift θ (resp. 0) goes over into Brownian motion with drift 1 (resp. 0). With $S_\theta = \theta^2(T_c^* - t)$ we get

$$\begin{aligned}&= \int_{-\infty}^{\infty}\big[\gamma(x,t)P_0^{(0,0)}(0 \le S_\theta < \infty) + (1-\gamma(x,t))cE_1^{(0,0)}S_\theta\big]G_{x,t}(d\theta)\\ &\ge \inf_S\big[\gamma(x,t)P_0^{(0,0)}(0 \le S < \infty) + (1-\gamma(x,t))cE_1^{(0,0)}S\big] =: \tilde{\rho}(x,t).\end{aligned}$$

But $\tilde{\rho}(x,t)$ is the minimal Bayes risk of (1.6) with $\gamma = \gamma(x,t)$. We determine now its Bayes stopping set. By (1.4)

$$\begin{aligned}\tilde{\rho}(x,t) &= \min_{0\le p\le 1}\left(\gamma p + 2(1-\gamma)c\log p^{-1}\right)\\ &= \gamma p_0 + 2(1-\gamma)c\log p_0^{-1}\end{aligned}$$

with $p_0 = (2(1-\gamma)c)/\gamma \wedge 1$ (p_0 denotes the stopping probability). Thus

$$\tilde{\rho}(x,t) = 2(1-\gamma)c\left[1 + \log\left(\frac{2(1-\gamma)c}{\gamma}\right)^{-1}\right] \quad \text{if } \frac{2(1-\gamma)c}{\gamma} \le 1$$

and

$$\tilde{\rho}(x,t) = \gamma \quad \text{if } \frac{2(1-\gamma)c}{\gamma} > 1.$$

The Bayes stopping region is therefore equal to

$$\{(x,t)|\gamma(x,t)=\tilde{\rho}(x,t)\}=\{(x,t)|\gamma(x,t)\le\gamma_0\},$$

where $\gamma_0 = 2c(1+2c)^{-1}\gamma_0$ is determined by the equation

$$p_0=\frac{2(1-\gamma_0)c}{\gamma_0}=1.\ \square$$

We now derive a refinement of the statement of Theorem 2. For this we need a somewhat more general notation. If $h(t)$ is a positive function of time we shall denote by $\mathscr{C}(h(\cdot))=\{(x,t)|\gamma(x,t)>h(t)\}$.

THEOREM 3. *For every $c>0$ there exists a bounded function $\tilde{c}(\cdot)\ge c$ with*

(a) *$\tilde{c}(t)/c\to 1$ when $t\to\infty$ for every fixed c, and*
(b) *$\sup_{0<t<\infty}\tilde{c}(t)/c\to 1$ when $c\to 0$, such that*

$$\mathscr{C}\left(\frac{2\tilde{c}(\cdot)}{1+2\tilde{c}(\cdot)}\right)\subset\mathscr{C}^*(c)\subset\mathscr{C}\left(\frac{2c}{1+2c}\right)\tag{3.13}$$

holds.

The theorem states that for c small or t large the optimal stopping region is very near to its upper bound $\mathscr{C}(2c/(1+2c))$. The proof of Theorem 3, which is deferred to the end of this section, will show that the upper bound of $\tilde{c}(\cdot)/c$ is a bit larger than $M/2$ where M is the constant appearing in Theorem 2.

Several conclusions can be drawn from the theorem. Let $\psi_c^*(t)=\inf\{x>0|\rho(x,t)=\gamma(x,t)\}$. By Theorem 2 this definition makes sense. Thus by the symmetry of the problem

$$T_c^*=\inf\{t||W(t)|\ge\psi_c^*(t)\}.$$

COROLLARY 1.

$$\psi_c^*(t)=\left[(t+r)\left(\log\left(\frac{t+r}{r}\right)+2\log\frac{\gamma}{2(1-\gamma)c}+o(1)\right)\right]^{1/2}$$

when $t\to\infty$.

COROLLARY 2.

$$\psi_c^*(t)=\left[(t+r)\left(2\log\frac{\gamma}{2(1-\gamma)c}+\log\left(\frac{t+r}{r}\right)+o(1)\right)\right]^{1/2}$$

uniformly in t when $c\to 0$.

COROLLARY 3. *For every $\varepsilon>0$ there exists a $c_0>0$ such that*

$$T_{2c(1+\varepsilon)/(1+2c(1+\varepsilon))}\le T_c^*\le T_{2c/(1+2c)}\quad \textit{for all } c_0\ge c>0.$$

We can combine Corollary 3 with some recent results about boundary crossing distributions to get the minimal Bayes risk for (1.1) up to an $o(c)$-term. A related $O(c)$-result for the Bayes risk has been obtained by Pollak (1978), when there is an indifference zone in the parameter space.

THEOREM 4.

$$0 \le \rho(T_{2c/(1+2c)}) - \rho(T_c^*) = o(c) \quad \textit{when } c \to 0. \tag{3.14}$$

The minimal Bayes risk for (1.1) *is given by*

$$\rho(T_c^*) = 2(1-\gamma)c\left[\log b + \tfrac{1}{2}\log\log b + 1 + \tfrac{1}{2}\log 2 - A + o(1)\right] \tag{3.15}$$

when $c \to 0$. *Here*

$$b = \frac{\gamma}{2(1-\gamma)c} \quad \textit{and} \quad A = 2\int_0^\infty \log x\, \phi(x)\, dx.$$

REMARK. Comparing statement (3.15) with the related formula (1.8) for the simple testing problem shows that the additional term $2(1-\gamma)c[\frac{1}{2}\log(2\log b) - A + o(1)]$ appears in the minimal Bayes risk. This is caused by the ignorance about the parameter $\theta \neq 0$.

PROOF. From Corollary 3 it follows

$$\begin{aligned} &\rho(T_{2c(1+\varepsilon)/(1+2c(1+\varepsilon))}) \\ &\quad -\gamma\left[P_0(T_{2c(1+\varepsilon)/(1+2c(1+\varepsilon))} < \infty) - P_0(T_{2c/(1+2c)} < \infty)\right] \\ &\quad \le \rho(T_c^*) \le \rho(T_{2c/(1+2c)}). \end{aligned} \tag{3.16}$$

We now show that the right- and left-hand sides of (3.16) differ from each other only by a $o(c)$-term. Formula (3.5) yields

$$P_0(T_{2c(1+\varepsilon)/(1+2c(1+\varepsilon))} < \infty) - P_0(T_{2c/(1+2c)} < \infty) \le \varepsilon b^{-1} = O(\varepsilon c). \tag{3.17}$$

Now we compute $\rho(T_{2c/(1+2c)})$. We write from now on for simplicity T instead of $T_{2c/(1+2c)}$. By (3.5) and (3.7) for $x = 0$, $t = 0$

$$\rho(T) = 2(1-\gamma)c\left[1 + \log b + \frac{1}{2}\int \log\left(\frac{T+r}{r}\right) dQ\right]. \tag{3.18}$$

The integral on the right-hand side can be calculated by using Theorem 5 of Jennen and Lerche (1981). The following result is intuitively plausible by virtue of the relation

$$P_\theta\{T/\log b \to 2\theta^{-2}\} = 1.$$

We note

$$\int \log\left(\frac{T+r}{r}\right) dQ = \log(2\log b) - 2A + o(1). \tag{3.19}$$

Combining (3.19) with (3.18) yields

$$(3.20)\quad \rho\big(T_{2c/(1+2c)}\big) = 2(1-\gamma)c\left[\log b + \frac{\log(2\log b)}{2} + 1 - A + o(1)\right].$$

From (3.17) and (3.20) it follows also that

$$(3.21)\qquad \rho\big(T_{2c(1+\varepsilon)/(1+2c(1+\varepsilon))}\big) = \rho\big(T_{2c/(1+2c)}\big) + O(\varepsilon c).$$

Statement (3.21) together with (3.16) and (3.17) yields (3.14) and (3.14) together with (3.20) yields (3.15). □

PROOF OF THEOREM 3. The upper inclusion of (3.13) is already proved by (3.12). Now we prove the lower one. For the stopping times

$$(3.22)\qquad S_{(x,t)} = \inf\left\{s \ge t \,|\, \gamma(W(s), s) \le \frac{2c}{1+2c}\right\},$$

we show that

$$(3.23)\quad \rho\big(x, t, S_{(x,t)}\big) < \gamma(x,t) \quad \text{for } (x,t) \in \mathscr{C}\left(\frac{2\tilde{c}(\cdot)}{1+2\tilde{c}(\cdot)}\right)\Big\backslash \mathscr{C}\left(\frac{Mc}{1+Mc}\right),$$

where $\tilde{c}(\cdot)$ will be specified below. M is the constant of Theorem 2. Then (3.1) together with (3.23) implies the lower inclusion of (3.13).

Now we define $\tilde{c}(t)$. We note that for the stopping times (3.22) by (3.9) with $\alpha = 4$ and

$$b(x,t) = \frac{\gamma(x,t)}{2(1-\gamma(x,t))c}$$

the inequality

$$(3.24)\quad \int \theta^2(S-t)\,dQ^{(x,t)} \le 2\log b(x,t) + C(x,t)\left[\int \theta^2(S-t)\,dQ^{(x,t)}\right]^{1/4}$$

holds. The constants $C(x,t)$ are given by

$$(3.25)\qquad \begin{aligned} C(x,t) &= K\left(\int \big(\theta^2(t+r)\big)^{-1/3} G_{x,t}(d\theta)\right)^{3/4} \\ &= K\left(\int y^{-2/3} N(\hat{\theta}\sqrt{t+r}, 1)(dy)\right)^{3/4} \end{aligned}$$

with $\hat{\theta} = x/(t+r)$.

Let ψ_c^+ and ψ_c^- denote the positive and negative branches of the solution of the implicit equation $\gamma(\psi_c^{\pm}(t), t) = Mc/(1+Mc)$. By symmetry $\psi_c^{\pm} = \pm\psi_c$ where ψ_c is given by

$$\psi_c(t) = \left[(t+r)\left(\log\left(\frac{t+r}{r}\right) + 2\log\frac{\gamma}{(1-\gamma)Mc}\right)\right]^{1/2}$$

We choose

$$(3.26)\qquad e(t,c) = -\log\Big(1 - C(\psi_c(t), t)^{1/4}\Big) \wedge \log M/2$$

and put $\bar{c}(t) = c\exp(e(t, c))$. Let $a > 1$. Let $d(a) = \inf\{y > 1 | a\log(ay) < ay - 1\}$. $d(a)$ is uniquely determined. We define $\tilde{c}(t) = d(\bar{c}(t)/c)\bar{c}(t)$.

Now we claim that $\tilde{c}(\cdot)/c$ has the demanded properties (a) and (b). By (3.25) $C(x, t)$ depends only on $|\hat{\theta}\sqrt{t+r}| = |x|/(\sqrt{t+r})$. Evaluating $|\hat{\theta}\sqrt{t+r}|$ at the graphs $(\pm\psi_c(t), t)$ yields

$$|\hat{\theta}\sqrt{t+r}| = \left[\log(t+r) + 2\log\frac{\gamma}{(1-\gamma)Mc}\right]^{1/2},$$

which tends to infinity, uniformly in t when $c \to 0$, or when $t \to \infty$.

Consequently, $C(\pm\psi_c(t), t) \to 0$ and therefore by (3.26) $e(t, c) \to 0$, uniformly in t when $c \to 0$, or when $t \to \infty$. Since $d(a) \to 1$ as $a \to 1$ the properties (a) and (b) follow.

Now we show (3.23). As a first step we prove

$$c\int \theta^2(S-t)\,dQ^{(x,t)} \le 2\bar{c}(t)\log b(x,t) \tag{3.27}$$

for

$$(x, t) \in \mathscr{C}\left(\frac{2\bar{c}(\cdot)}{1+2\bar{c}(\cdot)}\right)\Big\backslash\mathscr{C}\left(\frac{Mc}{1+Mc}\right).$$

By (3.26) we can assume that $\bar{c}(t) < Mc/2$. Let $H(x, t) = \int\theta^2(S-t)\,d\overline{Q}^{(x,t)}$. Then we have from (3.6), (3.7), and (3.24) (with the x and t variables suppressed)

$$2\log b \le H \le 2\log b + CH^{1/4}. \tag{3.28}$$

Let $C_1(t) = C(\psi_c(t), t)$. Then $0 \le C(x, t) \le C_1(t)$ holds on $\mathscr{C}(Mc/(1+Mc))^c$ and therefore

$$\left(1 - C_1/H^{3/4}\right)H \le 2\log b. \tag{3.29}$$

If

$$b(x, t) \ge \exp\left(\tfrac{1}{2}C_1(t)\right) \tag{3.30}$$

holds for

$$(x, t) \in \mathscr{C}\left(\frac{2\bar{c}(\cdot)}{1+2\bar{c}(\cdot)}\right)\Big\backslash\mathscr{C}\left(\frac{Mc}{1+Mc}\right)$$

then we get from the left-hand side of (3.28) $C_1(t) \le H(x, t)$ and therefore from (3.29)

$$H(x, t) \le 2\log(b(x, t))\left(1 - C_1(t)^{1/4}\right)^{-1}.$$

But this yields (3.27).

It is left to show that (3.30) holds. Let

$$c(t) = \frac{c}{1 - C_1(t)^{1/4}}\Big/\exp\left(\tfrac{1}{2}C_1(t)\right).$$

An elementary calculation shows that $c(t) \geq c$ for $0 \leq C_1(t) \leq 1$. Then

$$\begin{aligned} b(x,t) &= \frac{\gamma(x,t)}{2(1-\gamma(x,t))c} \geq \frac{\gamma(x,t)}{2(1-\gamma(x,t))c(t)} \\ &= \frac{\gamma(x,t)\exp\left(\frac{1}{2}C_1(t)\right)}{2(1-\gamma(x,t))\bar{c}(t)} \\ &\geq \exp\left(\tfrac{1}{2}C_1(t)\right). \end{aligned}$$

The second equation holds since

$$\bar{c}(t) = c\left(1 - C_1(t)^{1/4}\right)^{-1} < M/2,$$

and the last inequality follows from the definition of $\mathscr{C}(2\bar{c}(\cdot)/(1+2\bar{c}(\cdot))$. This proves (3.30) and completes the proof of (3.27).

Combining now (3.4) and (3.5) with (3.27) yields for the stopping times (3.22) the estimate for the Bayes risks

$$\rho(x,t,S) \leq \gamma(x,t)b(x,t)^{-1} + 2(1-\gamma(x,t))\bar{c}(t)\log b(x,t) \tag{3.31}$$

with

$$b(x,t) = \frac{\gamma(x,t)}{2(1-\gamma(x,t))c} \quad \text{on } \mathscr{C}\left(\frac{2\bar{c}(\cdot)}{1+2\bar{c}(\cdot)}\right)\Big\backslash \mathscr{C}\left(\frac{Mc}{1+Mc}\right).$$

We assume now that

$$(x,t) \in \mathscr{C}\left(\frac{2\tilde{c}(\cdot)}{1+2\tilde{c}(\cdot)}\right)\Big\backslash \mathscr{C}\left(\frac{Mc}{1+Mc}\right)$$

and estimate the right-hand side of (3.31) further. It is equal to

$$\begin{aligned} &\gamma(x,t)b(x,t)^{-1}[1 + (\bar{c}(t)/c)\log b(x,t)] \\ &\quad = \gamma(x,t)[h(x,t)\bar{c}(t)/c]^{-1}[1 + (c(t)/c)\log(h(x,t)\bar{c}(t)/c)] \end{aligned} \tag{3.32}$$

with $h(x,t) = \gamma(x,t)/(2(1-\gamma(x,t)\bar{c}(t))$. Since $(ay)^{-1}(1 + a\log(ay)) < 1$ for $y > d(a)$ and since on $\mathscr{C}(2\tilde{c}(\cdot)/(1+2\tilde{c}(\cdot))$, $h(x,t) > d(\bar{c}(t)/c)$ by the definition of $\tilde{c}$, it follows that expression (3.32) is strictly less than $\gamma(x,t)$. This yields (3.23) and completes the proof of Theorem 3. □

4. Results for one-sided tests. In this section we consider the Bayes risk given by

$$\rho(T) = \gamma P_0(T < \infty) + (1-\gamma)c\int_0^\infty \theta^2 E_\theta T \phi(\sqrt{r}\,\theta)2\sqrt{r}\,d\theta. \tag{4.1}$$

For it we can characterize the minimizing stopping rule T_c^* (it also exists) by results similar to those for the two-sided case.

If not mentioned otherwise we will use the same notation as in the preceding sections for the corresponding objects here. For instance,

$$\gamma(x,t) = F_{x,t}(\{0\}) = \left[1 + \frac{1-\gamma}{\gamma}\int_0^\infty \exp\left(\theta x - \tfrac{1}{2}\theta^2 t\right)\phi(\theta\sqrt{r})2\sqrt{r}\,d\theta\right]^{-1},$$

and also $\rho(x, t, T)$, $\rho(x, T)$, $\mathscr{C}^*(c)$, $\mathscr{C}(Kc)$, etc. The prior on $[0, \infty)$ is given by

$$F = \gamma\delta_0 + (1-\gamma)\int_\cdot \phi(\sqrt{r}\,\theta)2\sqrt{r}\,d\theta.$$

The posterior at (x, t) can be represented as

$$F_{x,t} = \gamma(x,t)\delta_0 + (1-\gamma(x,t))H_{x,t},$$

where

$$H_{x,t} = N\left(\frac{x}{t+r}, \frac{1}{t+r}\right)\bigg/\Phi\left(\frac{x}{\sqrt{t+r}}\right)$$

on $(0, \infty)$. We only state the analogous result to Theorem 2. The counterpart to Theorem 3 holds also and can be proved in exactly the same way as Theorem 3.

THEOREM 5. *There exists a constant $K > 2$ such that for every $c > 0$*

$$\mathscr{C}\left(\frac{Kc}{1+Kc}\right) \subset \mathscr{C}^*(c) \subset \mathscr{C}\left(\frac{2c}{1+2c}\right). \tag{4.2}$$

PROOF. The proof of the upper inclusion of (4.2) runs exactly along the same lines as that of (3.1). For the lower inclusion we show that for all $(x, t) \in \mathscr{C}(Kc/(1 + Kc))$ there exists a stopping time $S_{(x,t)}$ of the processes $(W(s), s)$ starting at (x, t) such that

$$\rho(x, t, S_{(x,t)}) < \gamma(x,t) \tag{4.3}$$

holds. Let Q denote a constant which satisfies $Qc < 1$. We choose

$$S_{(x,t)} = \inf\{s > t \,|\, \gamma(W(s), s) \le Qc\},$$

which can be rewritten as

$$\begin{aligned} S_{(x,t)} &= \inf\left\{s > t \,\middle|\, \int_0^\infty \frac{dP_{\theta,s}^{(x,t)}}{dP_{0,s}^{(x,t)}} H_{x,t}(d\theta) \ge b(x,t)\right\} \\ &= \inf\left\{s > t \,\middle|\, \sqrt{\frac{t+r}{s+r}}\exp\left(\frac{1}{2}\left(\frac{W(s)^2}{s+r} - \frac{x^2}{t+r}\right)\right)\frac{\Phi\left(\frac{W(s)}{\sqrt{s+r}}\right)}{\Phi\left(\frac{x}{\sqrt{t+r}}\right)} \ge b(x,t)\right\} \end{aligned} \tag{4.4}$$

with

$$b(x,t) = \frac{\gamma(x,t)(1-Qc)}{(1-\gamma(x,t))Qc} \quad \text{and} \quad \Phi(y) = \int_{-\infty}^{y} \phi(x)\,dx.$$

The posterior risk at (x, t) can be represented as

$$\rho(x,t,S_{(x,t)}) = \gamma(x,t)P_0^{(x,t)}(t < S_{(x,t)} < \infty) + (1-\gamma(x,t))c\int_0^{\infty} \theta^2 E_\theta^{(x,t)}(S_{(x,t)} - t)H_{x,t}(d\theta). \tag{4.5}$$

From here on we write S instead of $S_{(x,t)}$. The same martingale argument as for (3.5) yields

$$P_0^{(x,t)}(t < S < \infty) = b(x,t)^{-1}.$$

The estimate of the other part of the Bayes risk (4.5) is a bit more complicated than that of the corresponding part of (3.4). It can be expressed after some calculations similar to those of (3.6) as follows:

$$\int \theta^2 E_\theta^{(x,t)}(S-t)H_{x,t}(d\theta) = \int\left(\frac{W(S)^2}{S+r} - \frac{x^2}{t+r}\right)d\overline{Q}^{(x,t)} + \int\left(h\left(\frac{W(S)}{\sqrt{S+r}}\right) - h\left(\frac{x}{\sqrt{t+r}}\right)\right)d\overline{Q}^{(x,t)} \tag{4.6}$$

with

$$h(y) = y\phi(y)/\Phi(y) \quad \text{and} \quad \overline{Q}^{(x,t)} = \int_0^{\infty} P_\theta^{(x,t)}H_{x,t}(d\theta).$$

Using the defining equation (4.4) of the stopping time S yields

$$\begin{aligned}\int_0^{\infty} &\theta^2 E_\theta^{(x,t)}(S-t)H_{x,t}(d\theta) \\ &= 2\log b + \int \log\left(\frac{S+r}{t+r}\right)d\overline{Q}^{(x,t)} \\ &\quad + \int\left[h\left(\frac{W(S)}{\sqrt{S+r}}\right) - h\left(\frac{x}{\sqrt{t+r}}\right)\right]d\overline{Q}^{(x,t)} \\ &\quad -2\int\left[g\left(\frac{W(S)}{\sqrt{S+r}}\right) - g\left(\frac{x}{\sqrt{t+r}}\right)\right]d\overline{Q}^{(x,t)}\end{aligned} \tag{4.7}$$

with $g(y) = \log\Phi(y)$.

Now after some calculations we get

$$\begin{aligned}h&\left(\frac{W(S)}{\sqrt{S+r}}\right) - h\left(\frac{x}{\sqrt{t+r}}\right) - 2\left[g\left(\frac{W(S)}{\sqrt{S+r}}\right) - g\left(\frac{x}{\sqrt{t+r}}\right)\right] \\ &= \int_{x/\sqrt{t+r}}^{W(S)/\sqrt{S+r}} (h'(y) - 2g'(y))\,dy \\ &= \int_{x/\sqrt{t+r}}^{W(S)/\sqrt{S+r}} \frac{\phi(y)}{\Phi(y)}\left[-(1+y^2) - h(y)\right]dy.\end{aligned} \tag{4.8}$$

But this integral is always negative. It is obvious that the integrand is negative for positive values of y. That it is also negative for negative y-values can be seen as follows. We have to show that

$$-(1 + y^2) - y\phi(y)/\Phi(y) \le 0 \quad \text{for } y < 0, \tag{4.9}$$

which is equivalent to

$$-(1 + y^2)(1 - \Phi(y)) + y\phi(y) \le 0 \quad \text{for } y > 0$$

and to

$$1 - \Phi(y) \ge (y/(1 + y^2))\phi(y) \quad \text{for } y > 0. \tag{4.10}$$

Both sides of (4.10) vanish at $y = \infty$ and the derivative of the left-hand side is always smaller than that of the right-hand side and both are negative, i.e.,

$$-\phi(y) \le -\phi(y)\left(1 - 2/(1 + y^2)^2\right) \quad \text{for all } y > 0.$$

This yields (4.9) and therefore the integrand in (4.8) is always negative.

It is left to show that $W(S)/\sqrt{S + r} > x/\sqrt{t + r}$. Now let $K/(1 + Kc) > Q$. Then $(x, t) \in \mathscr{C}(Kc/(1 + Kc))$ implies $\gamma(x, t) > Qc$, which yields $b(x, t) > 1$. This together with (4.4) implies for $S > t$ the inequality

$$\exp\left(\frac{x^2}{2(t + r)}\right)\Phi\left(\frac{x}{\sqrt{t + r}}\right) < \sqrt{\frac{t + r}{S + r}}\exp\left(\frac{W(S)^2}{2(S + r)}\right)\Phi\left(\frac{W(S)}{\sqrt{S + r}}\right).$$

Since the function $\lambda \rightsquigarrow e^{\lambda^2/2}\Phi(\lambda)$ is increasing this yields $W(S)/\sqrt{S + r} > x/\sqrt{t + r}$. Thus the expression in (4.8) is always negative, which by (4.7) yields

$$\int_0^\infty \theta^2 E_\theta^{(x,t)}(S - t)H_{x,t}(d\theta) \le 2\log b + \int \log\left(\frac{S + r}{t + r}\right) d\overline{Q}^{(x,t)}.$$

The rest of the proof is similar to that of Theorem 2 from (3.7) on. □

Acknowledgments. The author wishes to thank W. Ehm, C. Jennen, I. Johnstone, C. Klaasen, S. Lalley, T. L. Lai, D. W. Müller, T. Sellke, and D. Siegmund for stimulating discussions and the Associate Editor and the referees for their careful work.

REFERENCES

CORNFIELD, J. (1966). A Bayesian test of some classical hypotheses—with application to sequential clinical trials. *J. Amer. Statist. Assoc.* **61** 577–594.

DARLING, D. and ROBBINS, H. (1967). Iterated logarithm inequalities. *Proc. Nat. Acad. Sci U.S.A.* **57** 1188–1192.

JEFFREYS, H. (1948). *Theory of Probability.* 2nd ed. Clarendon, Oxford.

JENNEN, C. and LERCHE, H. R. (1981). First exit densities of Brownian motion through one-sided moving boundaries. *Z. Wahrsch. verw. Gebiete* **55** 133–148.

JENNEN, C. and LERCHE, H. R. (1982). Asymptotic densities of stopping times associated with tests of power one. *Z. Wahrsch. verw. Gebiete* **61** 501–511.

LERCHE, H. R. (1985). On the optimality of sequential tests with parabolic boundaries. In *Proc. of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer.* (L. Le Cam and R. Olshen, eds.) **2** 579–597. Wadsworth, Monterey, Calif.

LERCHE, H. R. (1986). An optimal property of the repeated significance test. *Proc. Nat. Acad. Sci. U.S.A.* **83** 1546–1548.

LIPTSER, R. S. and SHIRYAYEV, A. N. (1977). *Statistics of Random Processes* **1**. Springer, Berlin.

LORDEN G. (1977). Nearly-optimal sequential tests for finitely many parameter values. *Ann. Statist.* **5** 1–21.

POLLAK, M. (1978). Optimality and almost optimality of mixture stopping rules. *Ann. Statist.* **6** 910–916.

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

ROBBINS, H. and SIEGMUND, D. (1970). Boundary crossing probabilities for the Wiener process and partial sums. *Ann. Math. Statist.* **41** 1410–1429.

ROBBINS, H. and SIEGMUND, D. (1973). Statistical tests of power one and their integral representation of solutions of certain partial differential equations. *Bull. Acad. Sinica* **1** 93–120.

SHIRYAYEV, A. N. (1978). *Optimal Stopping Rules*. Springer, Berlin.

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

UNIVERSITÄT HEIDELBERG
INSTITUT FÜR ANGEWANDTE MATHEMATIK
IM NEUENHEIMER FELD 294
6900 HEIDELBERG 1
WEST GERMANY