

Sequential change-point detection in continuous time when the post-change drift is unknown

M. BEIBEL

Institut für Mathematische Stochastik, Albert-Ludwigs-Universität Freiburg, Eckerstrasse 1, 79104 Freiburg, Germany
e-mail: beibel@pollux.mathematik.uni-freiburg.de

Let $W_t(0 \leq t < \infty)$ denote a Brownian motion process which has zero drift during the time interval $[0, \nu)$ and drift θ during the time interval $[\nu, \infty)$, where θ and ν are unknown. The process W is observed sequentially. The general goal is to find a stopping time T of W that ‘detects’ the unknown time point ν as soon and as reliably as possible on the basis of this information. Here stopping always means deciding that a change in the drift has already occurred. We discuss two particular loss structures in a Bayesian framework. Our first Bayes risk is closely connected to that of the Bayes tests of power one of Lerche. The second Bayes risk generalizes the disruption problem of Shiriyayev to the case of unknown θ .

Keywords: Bayes problems; Brownian motion; change point; sequential detection; tests of power one

1. Introduction and main results

Let $W_t(0 \leq t < \infty)$ denote a Brownian motion process which has zero drift during the time interval $[0, \nu)$ and drift θ during the time interval $[\nu, \infty)$, where θ and ν are unknown parameters. Let $P_{(\theta, \nu)}$ denote the corresponding probability. Let P_∞ denote the probability measure when no change in the drift occurs. Note that $P_\infty = P_{(\theta, \infty)}$ for all $\theta \in (-\infty, +\infty)$. Let $E_{(\theta, \nu)}$ denote the expectation with respect to $P_{(\theta, \nu)}$ and E_∞ the expectation with respect to P_∞ . We observe W sequentially and are able to stop this process at any given instant. Let $\mathcal{F}_t = \mathcal{F}_t^W = \sigma(W_s; 0 \leq s \leq t)$. The process W is ‘in control’ or in a ‘favourable’ state as long as the drift is zero and we do not want to interrupt it during this period. As soon as the drift becomes $\theta \neq 0$ the process W is ‘out of control’ or in an ‘unfavourable’ state and we now want to stop and take some action. We have no precise information about ν and are not able to anticipate the future development of W . We thus seek a stopping rule T of W which will stop soon after ν without too many ‘false alarms’. These two conflicting goals are still rather vague and have to be further specified. We shall investigate two particular loss structures in a Bayesian framework (see (8) and (19) below).

Shiriyayev (1963) developed the following Bayes approach. Let $\theta \in R$ be a fixed and known constant. Let τ be a random variable with $P(\{\tau = 0\}) = p$ and $P(\{\tau > t\}) =$

$(1 - p)e^{-\lambda t}$ for $t \geq 0$ where $p \in [0, 1)$ and $\lambda > 0$. The observed process W is given by $W_t = B_t + \theta(t - \tau)^+$, where B is a standard Brownian motion independent of τ . The distribution P of W therefore is given by

$$P = pP_{(\theta,0)} + (1 - p) \int_0^\infty P_{(\theta,v)} \lambda e^{-\lambda v} dv.$$

The quality of a stopping time T is measured by the risk function $\mathcal{R}(c, T)$ given as

$$\mathcal{R}(c, T) = P(T < \tau) + cE(T - \tau)^+, \tag{1}$$

with costs $c > 0$. Let \mathcal{R}_c^* denote the minimal Bayes risk $\mathcal{R}_c^* = \inf_T \mathcal{R}(c, T)$, where the infimum is taken over all stopping times T of W . Let

$$\begin{aligned} \pi_t &= P(\tau \leq t | \mathcal{F}_t) \\ &= \frac{p e^{\theta W_t - (\theta^2/2)t} + (1 - p) \int_0^t e^{\theta(W_t - W_s) - (\theta^2/2)(t-s)} \lambda e^{-\lambda s} ds}{p e^{\theta W_t - (\theta^2/2)t} + (1 - p) \int_0^t e^{\theta(W_t - W_s) - (\theta^2/2)(t-s)} \lambda e^{-\lambda s} ds + (1 - p) e^{-\lambda t}}. \end{aligned}$$

Shiryayev (1963) obtained the following optimality result. Let p_c^* denote the unique solution in $(0, 1)$ of the implicit equation

$$c \frac{2}{\theta^2} \exp \{-\Lambda H(x)\} \int_0^x \exp \{\Lambda H(w)\} w^{-1} (1 - w)^{-2} dw = 1,$$

where

$$\Lambda := \frac{2\lambda}{\theta^2}, \quad H(w) := \log \left(\frac{w}{1 - w} \right) - \frac{1}{w}.$$

Then

$$T_c^* := \inf \{t > 0 | \pi_t \geq p_c^*\} \tag{2}$$

minimizes $\mathcal{R}(c, T)$ among all stopping times T of W , i.e., $\mathcal{R}(c, T_c^*) = \mathcal{R}_c^*$.

By modifying the proof given by Shiryayev (1973), one can transform the problem of finding a Bayes solution into a generalized parking problem (Beibel 1994b). This means that for a sufficiently large class of stopping times T the Bayes risk $\mathcal{R}(c, T)$ can be written as $E(g_c(\pi_T))$, where g_c is a convex function with a unique minimum at $x = p_c^*$. Since π_t is continuous in t , one can stop exactly in the minimum. A similar approach can be applied to the continuous-time version of the Bayes problems of Ritov (1990) (Beibel 1996). This method also works for many other cases (see, for example, Beibel and Lerche 1997, Lerche 1986a and Woodroffe *et al.* 1994).

Let $p = 0$. The optimal solution T_c^* in (2) can then be rewritten as

$$T_c^* = \inf \left(t > 0 \mid \int_0^t e^{\theta(W_t - W_s) - (\theta^2/2)(t-s)} e^{\lambda(t-s)} ds \geq A \right),$$

with $A = p_c^*/[\lambda(1 - p_c^*)]$. For λ close to 0 this stopping rule does not differ much from the stopping rule

$$T_A = \inf \left(t > 0 \mid \int_0^t e^{\theta(W_t - W_s) - (\theta^2/2)(t-s)} ds \geq A \right). \tag{3}$$

In fact p_c^*/λ converges to a non-degenerate limit as λ tends to zero (Pollak and Siegmund 1985).

The stopping rule T_A has been studied in detail by Pollak and Siegmund (1985). Pollak and Siegmund also proposed the following generalization of (3) to the case of unknown $\theta > 0$. Let G denote some probability measure on $(0, \infty)$. A new stopping rule \hat{T}_A is then defined by

$$\hat{T}_A = \inf \left\{ t > 0 \mid \int_0^\infty \left(\int_0^t e^{y(W_t - W_s) - (y^2/2)(t-s)} ds \right) G(dy) \geq A \right\}. \tag{4}$$

Pollak and Siegmund noted that

$$\int_0^\infty \left(\int_0^t e^{y(W_t - W_s) - (y^2/2)(t-s)} ds \right) G(dy) - t$$

is a P_∞ martingale with mean 0 and hence $E_\infty(\hat{T}_A) = A$. They also discuss the average run length of \hat{T}_A under $P_{(\theta,0)}$ when the threshold A is large. Let κ denote Euler's constant. Then they stated that

$$\frac{\theta^2}{2} E_{(\theta,0)}(\hat{T}_A) = \log A + \frac{1}{2} \log \log A - \frac{1}{2} - \kappa - \frac{1}{2} \log \{2\pi g^2(\theta)\} - \frac{1}{2} \log \left(\frac{2}{\theta^2} \right) + o(1) \tag{5}$$

as $A \rightarrow \infty$ if $\theta > 0$ and G has a positive continuous density g in some neighbourhood of θ .

The stopping times T_c^* , T_A and \hat{T}_A all have finite average run lengths under P_∞ . This means in particular that they will eventually stop even if no change in the drift occurs. In some applications this might be considered as a disadvantage. One way to remedy this is to impose a probability constraint such as $P_\infty(T < \infty) \leq \alpha < 1$ for some $\alpha \in (0, 1)$ on T (Pollak and Siegmund 1975). Any stopping time T which satisfies this condition as well as $P_{(\theta,\nu)}(T < \infty) = 1$ for $\theta \neq 0$ and $0 \leq \nu < \infty$ can be considered as a test of power one for the testing problem

$$H_0: P_\infty \text{ against } H_1: P_{(\theta,\nu)} \text{ for some } \theta \neq 0 \text{ and } \nu \in [0, +\infty).$$

Tests of power one for

$$H'_0: P_\infty \text{ against } H'_1: P_{(\theta,0)} \text{ for some } \theta \neq 0$$

were introduced by Darling and Robbins (1967). A test of power one for H'_0 versus H'_1 is by definition a stopping time T which satisfies the conditions

$$(I) P_\infty(T < \infty) < 1, \quad (II) P_{(\theta,0)}(T < \infty) = 1 \text{ for all } \theta \neq 0.$$

Here stopping always means deciding in favour of H'_1 . Lerche (1986b) determined Bayes tests of power one for H'_0 versus H'_1 when the sampling costs are proportional to θ^2 . Let

$\sigma^2 > 0$ and $\mu \in R$ be known constants. Let G denote the normal distribution with mean μ and variance σ^2 . Lerche (1986b) considered the Bayes risk

$$\mathcal{L}(c, T) = \gamma P_\infty(T < \infty) + (1 - \gamma)c \int_{-\infty}^{+\infty} \theta^2 E_{(\theta,0)}(T) G(d\theta), \tag{6}$$

with $\gamma \in (0, 1)$ and small observation costs c . The expression on the right-hand side of (6) is proportional to

$$P_\infty(T < \infty) + c \frac{2(1 - \gamma)}{\gamma} \int_{-\infty}^{+\infty} \frac{\theta^2}{2} E_{(\theta,0)}(T) G(d\theta). \tag{7}$$

Our first loss structure is a natural generalization of (7). Let ρ denote some probability measure on $[0, \infty)$ with

$$\int_0^\infty s \rho(ds) < \infty.$$

Let

$$L(c, T) = P_\infty(T < \infty) + c \int_{-\infty}^{+\infty} \left(\frac{\theta^2}{2} \int_{[0,\infty)} E_{(\theta,\nu)}(T - \nu)^+ \rho(d\nu) \right) G(d\theta). \tag{8}$$

The costs for observations taken after the change point ν are proportional to θ^2 . The corresponding Bayes risk for the case of θ known has been studied by Beibel (1994b) and Keener *et al.* (1995). For $\rho(\{0\}) = 1$ the Bayes risk (8) reduces to (7) if we replace c by $2c(1 - \gamma)/\gamma$. Let L_c^* denote the minimal Bayes risk $L_c^* = \inf_T L(c, T)$. Let P now denote the probability measure

$$P = \int_{-\infty}^{+\infty} \int_{[0,\infty)} P_{(\theta,\nu)} \rho(d\nu) G(d\theta). \tag{9}$$

Let

$$\begin{aligned} \psi_t &= \frac{dP}{dP_\infty} \Big|_{\mathcal{F}_t} = \int_{-\infty}^{+\infty} \int_{[0,\infty)} \frac{dP_{(\theta,\nu)}}{dP_\infty} \Big|_{\mathcal{F}_t} \rho(d\nu) G(d\theta) \\ &= \frac{1}{\sigma} e^{-\mu^2/2\sigma^2} \int_{[0,\infty)} \frac{1}{\{(t-s)^+ + \sigma^{-2}\}^{1/2}} \exp\left(\frac{(W_t - W_{t \wedge s} + \mu/\sigma^2)^2}{2\{(t-s)^+ + \sigma^{-2}\}}\right) \rho(ds). \end{aligned} \tag{10}$$

Note that

$$\frac{dP_{(\theta,\nu)}}{dP_\infty} \Big|_{\mathcal{F}_t} = \begin{cases} \exp\left(\theta(W_t - W_\nu) - \frac{\theta^2}{2}(t - \nu)\right), & \text{if } \nu < t, \\ 1, & \text{if } \nu \geq t. \end{cases}$$

Let $S_b = \inf \{t > 0 | \psi_t \geq b\}$. S_b stops the process as soon as the likelihood ratio $dP/dP_\infty|_{\mathcal{F}_t}$ attains or exceeds b . Let $\beta(c) = 1/c$. Theorem 1 below states that the stopping rules $S_{\beta(c)}$ are approximately optimal for $L(c, \cdot)$. Their expected loss $L(c, S_{\beta(c)})$ approaches the minimal

Bayes risk L_c^* within an error of $o(c)$ as c tends to zero. Theorem 2 below gives an explicit expansion of $L(c, S_{\beta(c)})$ up to an error term of magnitude $o(c)$ for small costs.

Theorem 1.

$$L_c^* = L(c, S_{\beta(c)}) + o(c) \text{ when } c \rightarrow 0.$$

Theorem 2.

$$L(c, S_{\beta(c)}) = c \left\{ \log \left(\frac{1}{c} \right) + \frac{1}{2} \log \log \left(\frac{1}{c} \right) + 1 - \frac{1}{2} \int_{-\infty}^{+\infty} \log \left(\frac{\theta^2}{2} \right) G(d\theta) + \log \sigma \right. \\ \left. + \left(\int_{-\infty}^{+\infty} \frac{\theta^2}{2} G(d\theta) \right) \left(\int_0^{+\infty} s \rho(ds) \right) - K_\infty \right\} + o(c)$$

when $c \rightarrow 0$. Here K_∞ is given by

$$K_\infty = \int_{-\infty}^{\infty} \int_{[0, \infty)} \left\{ E_{(\theta, \nu)} \log \left(\int_{[0, \infty)} e^{-\theta W_s + (\theta^2/2)s} \rho(ds) \right) \right\} \rho(d\nu) G(d\theta).$$

Remark 1. For $\rho(ds) = \lambda e^{-\lambda s} ds$, it is possible to evaluate K_∞ explicitly. Using the results of Yor (1992), one then obtains

$$K_\infty = \int_{-\infty}^{\infty} \int_0^{\infty} \left\{ \frac{2\lambda}{y^2} \frac{1}{u} \log \left(\frac{2\lambda}{y^2} \frac{1}{u} \right) \frac{u^{2\lambda/y^2} e^{-u}}{\Gamma(1 + 2\lambda/y^2)} du \right\} G(dy). \tag{11}$$

Here Γ denotes the usual gamma function, i.e., $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$.

Remark 2. For $\rho(\{0\}) = 1$, $\mu = 0$, and $\sigma^2 = 1$ the stopping time S_b becomes

$$\inf(t > 0 \mid |W_t| \geq [(t + 1)\{\log(t + 1) + 2 \log b\}]^{1/2}).$$

Moreover K_∞ and $\int_0^{\infty} s \rho(ds)$ vanish. Let $\tilde{c} = 2(1 - \gamma)c/\gamma$. $L(\tilde{c}, T)$ equals

$$P_\infty(T < \infty) + \frac{2(1 - \gamma)}{\gamma} c \int_{-\infty}^{+\infty} \frac{\theta^2}{2} E_{(\theta, 0)}(T) G(d\theta) = \frac{1}{\gamma} \mathcal{L}(c, T).$$

Let \mathcal{L}_c^* denote the minimal Bayes risk for (6)

$$\mathcal{L}_c^* = \inf_T \left(\gamma P_\infty(T < \infty) + (1 - \gamma)c \int_{-\infty}^{+\infty} \theta^2 E_{(\theta, 0)}(T) G(d\theta) \right).$$

Theorem 1 and Theorem 2 now yield together

$$\mathcal{L}_c^* = 2(1 - \gamma)c \left\{ \log \left(\frac{1}{\tilde{c}} \right) + \frac{1}{2} \log \log \left(\frac{1}{\tilde{c}} \right) + 1 + \frac{1}{2} \log 2 \right. \\ \left. - \frac{2}{(2\pi)^{1/2}} \int_0^{+\infty} \log(\theta) e^{-\theta^2/2} d\theta + o(1) \right\} \tag{12}$$

when $c \rightarrow 0$. Hence Theorem 5 of Lerche (1986b) can be obtained as a special case of Theorem 1 and Theorem 2 above. Note, however, that Lerche (1986b) gives for fixed costs c upper and lower bounds for the optimal stopping rules which are asymptotically tight. Analogous results on the structure of the optimal solutions are no longer feasible in our set-up since in general we do not have a Markovian structure.

The main point in our approach to prove Theorem 1 and Theorem 2 is similar to that of Lerche (1986b). In a first step we derive an alternative representation of the Bayes risk $L(c, \cdot)$. We show (see Proposition 2 and Proposition 3 below) for all stopping times T of W with $L(c, T) < \infty$ that

$$L(c, T) = E\{g_c(\psi_T) + cV_T\}, \tag{13}$$

where $g_c(x) = 1/x + c \log x$ and $(V_t; 0 \leq t < \infty)$ is a non-negative increasing process which grows for large t as $\frac{1}{2} \log t$ (see (24), Proposition 3 and Lemma 2 below). Then (13) transforms the initial optimal stopping problem into a ‘perturbed’ generalized parking problem. This approach has also been employed by Beibel (1994b) and Keener *et al.* (1995). The main point in proving (13) is to show that the mean delay can be rewritten for any stopping time T of W with $L(c, T) < \infty$ as

$$\int_{-\infty}^{+\infty} \left(\frac{\theta^2}{2} \int_0^{+\infty} E_{(\theta, \nu)}(T - \nu)^+ \rho(d\nu) \right) G(d\theta) = E(\log \psi_T + V_T), \tag{14}$$

where ψ_t is defined in (10). We recall that $P = \int_{-\infty}^{+\infty} \int_{[0, \infty)} P_{(\theta, \nu)} \rho(d\nu) G(d\theta)$. Using (13) we shall first prove Theorem 2 and then Theorem 1.

Assuming (13) we obtain, after verifying $L(c, S_{\beta(c)}) < \infty$ (see Lemma 3 below), that

$$L(c, S_{\beta(c)}) = E\left(\frac{1}{\psi_{S_{\beta(c)}}} + c \log \psi_{S_{\beta(c)}} + cV_{S_{\beta(c)}} \right).$$

Obviously $E(1/\psi_{S_{\beta(c)}} + c \log \psi_{S_{\beta(c)}}) = c + c \log(1/c)$ for $0 < c \leq 1$. In order to prove Theorem 2, it therefore suffices to show that

$$\begin{aligned} E(V_{S_b}) &= \frac{1}{2} \log \log b - \frac{1}{2} \int_{-\infty}^{+\infty} \log \left(\frac{\theta^2}{2} \right) G(d\theta) + \log \sigma \\ &\quad + \left(\int_{-\infty}^{+\infty} \frac{\theta^2}{2} G(d\theta) \right) \left(\int_0^{+\infty} s \rho(ds) \right) - K_\infty + o(1) \end{aligned} \tag{15}$$

as b tends to infinity. This is done in Proposition 5 below.

The function g_c assumes a unique minimum over the interval $(0, \infty)$ at $x = \beta(c)$. The stopping times $S_{\beta(c)}$ stop the process ψ in the global minimum of g_c and thus minimize $Eg_c(\psi_T)$ over all stopping times T of W . Let \tilde{S}_c denote a c^2 -optimal solution with respect to $L(c, \cdot)$. That is $L(c, \tilde{S}_c) \leq L_c^* + c^2$. Theorem 2 implies in particular that

$$L_c^* \leq c \log\left(\frac{1}{c}\right) + \frac{c}{2} \log \log\left(\frac{1}{c}\right) + O(c)$$

as $c \rightarrow 0$. For sufficiently small costs c , therefore $L_c^* < 1$. This means that the problem of minimizing $L(c, \cdot)$ is non-trivial. Note that $L(c, T) = 1$ for $T \equiv 0$. Since V_t is increasing in t , we may assume without loss of generality that $\tilde{S}_c \leq S_{\beta(c)}$. To obtain the approximate optimality of $S_{\beta(c)}$ it only remains to show that $E(V_{S_{\beta(c)}} - V_{\tilde{S}_c}) = o(1)$ as c tends to zero. If, as in the work of Beibel (1994b) and Keener *et al.* (1995), $E(\lim_{t \rightarrow \infty} V_t) < \infty$, it is sufficient to show that \tilde{S}_c converges to infinity in probability as $c \rightarrow 0$. In our case V_t has logarithmic growth and so we need a stronger condition. It is sufficient to show that

$$E \log(S_{\beta(c)} + \sigma^{-2}) - E \log(\tilde{S}_c + \sigma^{-2}) = o(1)$$

as $c \rightarrow 0$.

Remark 3. Note that (14) and (15) yield together (see Proposition 6 below)

$$\begin{aligned} \int_{-\infty}^{+\infty} \left(\frac{\theta^2}{2} \int_0^{+\infty} E_{(\theta, \nu)}(S_b - \nu)^+ \rho(d\nu) \right) G(d\theta) &= \log b + \frac{1}{2} \log \log b \\ &+ \frac{1}{2} \int_{-\infty}^{+\infty} \log\left(\frac{2}{\theta^2}\right) G(d\theta) + \log \sigma \\ &+ \left(\int_{-\infty}^{+\infty} \frac{\theta^2}{2} G(d\theta) \right) \left(\int_0^{+\infty} s \rho(ds) \right) - K_\infty + o(1) \end{aligned} \tag{16}$$

as b tends to infinity. The asymptotic expansions (5) and (16) are related. Following the lines of (A7) in the paper by Pollak and Siegmund (1985) one obtains for sufficiently large b and fixed θ

$$\begin{aligned} \log b &= \log(\psi_{S_b}) \\ &\approx \log \left(\int_{-\infty}^{+\infty} \int_0^{S_b} e^{y(W_{S_b} - W_s) - (y^2/2)(S_b - s)} \rho(ds) G(dy) \right) \\ &= \theta W_{S_b} - \frac{\theta^2}{2} S_b + \frac{1}{2} \frac{(W_{S_b} - \theta S_b)^2}{S_b} - \log \left\{ \left(\frac{S_b}{2\pi} \right)^{1/2} \right\} \\ &+ \log \left(\int_{-\infty}^{+\infty} \left[S_b^{1/2} \phi \left\{ S_b^{1/2} \left(\frac{W_{S_b}}{S_b} - y \right) \right\} \int_0^{S_b} e^{-yW_s + (y^2/2)s} \rho(ds) \right] G(dy) \right), \end{aligned}$$

where $\phi(x) = \{1/(2\pi)^{1/2}\} \exp(-\frac{1}{2}x^2)$. Arguing in a similar way as Pollak and Siegmund (1985) one then obtains for fixed θ and ν (see also Kohler 1995)

$$\begin{aligned} \frac{\theta^2}{2} E_{(\theta,\nu)}(S_b - \nu)^+ &= \log b + \frac{1}{2} \log \log b - \frac{1}{2} - \frac{1}{2} \log \left\{ 2\pi\phi^2 \left(\frac{\theta - \mu_0}{\sigma} \right) \right\} + \frac{1}{2} \log \left(\frac{2}{\theta^2} \right) \\ &+ \frac{1}{2}\theta^2\nu - E_{(\theta,\nu)} \log \left(\int_{[0,\infty)} e^{-\theta W_s + (\theta^2/2)s} \rho(ds) \right) + o(1) \end{aligned} \tag{17}$$

as $b \rightarrow \infty$. Formally integrating (17) with respect to $\rho(d\nu)$ and $G(d\theta)$ yields (16). To justify this operation one would of course have to show that (17) holds in some sense uniformly in θ and ν .

In order to motivate our second loss structure, we shall now have a brief look at Shiryaev's problem for small costs c . One can show (see Proposition 1 of Beibel 1994b) that, as $c \rightarrow 0$,

$$\mathcal{R}_c^* = \frac{c}{\lambda + \theta^2/2} \log \left(\frac{\lambda + \theta^2/2}{c} \right) + O(c) \tag{18}$$

The asymptotic expansion (18) has an important consequence. If we rewrite $\mathcal{R}(c, T)$ as

$$\mathcal{R}(c, T) = P(T < \tau) + \tilde{c} E \left\{ \left(\lambda + \frac{\theta^2}{2} \right) (T - \tau)^+ \right\},$$

where $\tilde{c} = c/(\lambda + \theta^2/2)$, (18) becomes $\mathcal{R}_c^* = \tilde{c} \log(1/\tilde{c}) + O(\tilde{c})$. The leading term $\tilde{c} \log(1/\tilde{c})$ does not depend on θ . This means that taking the costs for the delay proportional to $\lambda + \theta^2/2$ standardizes the problems with different values of θ such that they are asymptotically of equal difficulty for small costs $\tilde{c} = c/(\lambda + \theta^2/2)$. We therefore take the costs for observations after the change point proportional to $\lambda + \theta^2/2$ and consider now

$$R(c, T) = \int_{[0,\infty)} P_\infty(T < \nu) \lambda e^{-\lambda\nu} d\nu + c \int_{-\infty}^{+\infty} \left\{ \left(\lambda + \frac{\theta^2}{2} \right) \int_{[0,\infty)} E_{(\theta,\nu)}(T - \nu)^+ \lambda e^{-\lambda\nu} d\nu \right\} G(d\theta). \tag{19}$$

Note that we have now put $\rho(\{0\}) = p = 0$ to keep our formulas simple. Let R_c^* denote the minimal Bayes risk $R_c^* = \inf_T R(c, T)$, where the infimum is taken over all stopping times T with respect to the observed process W . Let $dP = dP_{(\theta,\nu)} \lambda e^{-\lambda\nu} G(d\theta)$. Let π now denote the process

$$\begin{aligned} \pi_t &= P(\tau \leq t | \mathcal{F}_t) = \left\{ \int_{[0,t)} \frac{1}{\{(t-s)^+ + \sigma^{-2}\}^{1/2}} \exp \left(\frac{(W_t - W_{t\wedge s} + \mu/\sigma^2)^2}{2\{(t-s)^+ + \sigma^{-2}\}} \right) \lambda e^{-\lambda s} ds \right\} \\ &\times \left\{ \int_{[0,t)} \frac{1}{\{(t-s)^+ + \sigma^{-2}\}^{1/2}} \exp \left(\frac{(W_t - W_{t\wedge s} + \mu/\sigma^2)^2}{2\{(t-s)^+ + \sigma^{-2}\}} \right) \lambda e^{-\lambda s} ds \right. \\ &\left. + \sigma e^{\mu^2/2\sigma^2} e^{-\lambda t} \right\}^{-1} \end{aligned} \tag{20}$$

for $0 \leq t < \infty$. Let T_a for $a \in (0, 1)$ denote the stopping time $T_a = \inf \{t > 0 | \pi_t \geq a\}$. These stopping times are simple Bayes rules in the sense that they stop as soon as the posterior probability π_t rises above the threshold a . Let $\alpha(c) = 1/(1 + c)$ for $0 < c < 1$. Theorem 3 below states that the stopping rules $T_{\alpha(c)}$ are approximately optimal for $R(c, \cdot)$. Their risk $R(c, T_{\alpha(c)})$ approximates the minimal Bayes risk R_c^* within an error of $o(c)$ as c tends to zero. Theorem 4 below gives an explicit expansion of the Bayes risk $R(c, T_{\alpha(c)})$ up to a remainder term of $o(c)$ for small c .

Theorem 3.

$$R_c^* = R(c, T_{\alpha(c)}) + o(c) \text{ when } c \rightarrow 0.$$

Theorem 4.

$$R(c, T_{\alpha(c)}) = c \left\{ \log \left(\frac{1}{c} \right) + \frac{1}{2} \log \log \left(\frac{1}{c} \right) - \frac{1}{2} \int_{-\infty}^{\infty} \log \left(\lambda + \frac{\theta^2}{2} \right) G(d\theta) \right. \\ \left. + \log \sigma + \frac{\mu^2 + \sigma^2}{\lambda} - K_{\infty} \right\} + o(c)$$

when $c \rightarrow 0$. Here K_{∞} is the same constant as in (11).

The proofs of Theorem 3 and Theorem 4 are similar to the proofs of Theorem 1 and Theorem 2 above (Beibel 1994a). We therefore omit the details. One can show that

$$R(c, T) = E \left[1 - \pi_T + c \left\{ \log \left(\frac{1}{1 - \pi_T} \right) - \pi_T \right\} + cV_T \right]$$

for all stopping times T of \mathcal{W} with $R(c, T) < \infty$. The function

$$h_c(x) = 1 - x + c \left\{ \log \left(\frac{1}{1 - x} \right) - x \right\}$$

is convex and assumes its unique minimum over the interval $(0, 1)$ at $x = \alpha(c)$. $T_{\alpha(c)}$ stops the process π in the global minimum of h_c and thus minimizes

$$E \left[1 - \pi_T + c \left\{ \log \left(\frac{1}{1 - \pi_T} \right) - \pi_T \right\} \right]$$

over all stopping times T of \mathcal{W} . The remainder term V_T is that of (13).

The rest of this paper is organized as follows. Section 2 is devoted to the proof of representation (13). Section 3 deals with the remainder term $E(V_{S_b})$ which appeared above. In particular we derive the asymptotic expansion (15). This result then yields a proof of Theorem 2. Section 4 covers the asymptotic Bayes optimality of the stopping rules $S_{\beta(c)}$ and gives a proof of Theorem 1. In Section 5 we collect some supporting lemmas which are of a more technical nature.

2. The Bayes risk

In this section we shall prove the representation (13) (see Propositions 1 and 2 below). From now on we shall assume without loss of generality that $\mu = 0$ and $\sigma^2 = 1$. The definition of ψ_t in (10) then simplifies to

$$\psi_t = \int_{[0, \infty)} \frac{1}{\{(t - s)^+ + 1\}^{1/2}} \exp\left(\frac{(W_t - W_{t \wedge s})^2}{2\{(t - s)^+ + 1\}}\right) \rho(ds).$$

In a first step we introduce random variables Y and τ that correspond to the prior distributions G and ρ . Let B denote a standard Brownian motion. Let Y be a standard normal random variable and τ be a random variable with $P(\tau > t) = \int_{(t, \infty)} \rho(ds)$ for all $t \geq 0$. Let B , Y and τ be independent. Put

$$W_t = B_t + Y(t - \tau)^+ = B_t + \int_0^t R_s ds,$$

where $R_s = Y1_{\{\tau \leq s\}}$. The distribution of W now is given by $\int_{-\infty}^{+\infty} \int_{[0, \infty)} P_{(\theta, \nu)} \rho(d\nu) G(d\theta)$ and so agrees with (9). Therefore $\psi_t = dP/dP_\infty|_{\mathcal{F}_t}$.

Let us recall that $\mathcal{F}_t = \mathcal{F}_t^W = \sigma(W_s; 0 \leq s \leq t)$. We also recall that P_∞ denotes the probability measure under which W is a standard Brownian motion. It can be considered as the measure with change at $\nu = \infty$. For any stopping time T of W the Bayes risk L is now equal to

$$L(c, T) = P_\infty(T < \infty) + cE\left(\frac{Y^2}{2}(T - \tau)^+\right). \tag{21}$$

We shall rewrite $L(c, \cdot)$ further. To do so, we first need some new notation. Let

$$\hat{R}_t = \psi_t^{-1} \int_{[0, t]} \frac{W_t - W_s}{(t - s + 1)^{3/2}} \exp\left(\frac{(W_t - W_s)^2}{2(t - s + 1)}\right) \rho(ds)$$

and

$$\widehat{R}_t^2 = \psi_t^{-1} \int_{[0, t]} \frac{\{(W_t - W_s)/(t - s + 1)\}^2 + 1/(t - s + 1)}{(t - s + 1)^{1/2}} \exp\left(\frac{(W_t - W_s)^2}{2(t - s + 1)}\right) \rho(ds).$$

Straightforward calculations show that for all $t \geq 0$

$$\hat{R}_t = E(R_t | \mathcal{F}_t), \quad \widehat{R}_t^2 = E(R_t^2 | \mathcal{F}_t).$$

Fubini's theorem yields for all stopping times T of W

$$E\left(\int_0^T \widehat{R}_s^2 ds\right) = E\{Y^2(T - \tau)^+\}.$$

Moreover $P_\infty(T < \infty) = E(1/\psi_T 1_{\{T < \infty\}})$. Therefore we have for all stopping times T of W with $L(c, T) < \infty$

$$L(c, T) = E \left(\frac{1}{\psi_T} + \frac{c}{2} \int_0^T \widehat{R}_s^2 ds \right). \tag{22}$$

Now let \overline{W} denote the innovation process

$$\overline{W}_t := W_t - \int_0^t \hat{R}_s ds.$$

This process is a standard Brownian motion under the probability measure P relative to the filtration $\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$ (Liptser and Shirayev 1977, pp. 297–299). For the log-likelihood $\log \psi$ we then obtain the following representation.

Proposition 1.

$$d(\log \psi_t) = \frac{1}{2}(\hat{R}_t)^2 dt + \hat{R}_t d\overline{W}_t. \tag{23}$$

Proof. Let

$$\tilde{Z}_t^\infty = \exp \left(- \int_0^t \hat{R}_s d\overline{W}_s - \frac{1}{2} \int_0^t (\hat{R}_s)^2 ds \right).$$

The Jensen inequality yields

$$E \exp \left\{ \frac{1}{4}(\hat{R}_s)^2 \right\} \leq E \exp \left(\frac{1}{4}Y^2 \right) < \infty.$$

Let \tilde{P}_∞ denote the probability measure on $\sigma(W_s; 0 \leq s < \infty)$ with

$$d\tilde{P}_\infty|_{\mathcal{F}_t} = \tilde{Z}_t^\infty dP|_{\mathcal{F}_t} \text{ for } t \in [0, \infty).$$

Girsanov’s theorem implies that $(W_t, \mathcal{F}_t; 0 \leq t < \infty)$ is a standard Brownian motion under the probability measure \tilde{P}_∞ . From Kolmogorov’s consistency theorem it follows that the two measures P_∞ and \tilde{P}_∞ agree on the σ -algebra $\sigma(W_s; 0 \leq s < \infty)$. Therefore $(1/\psi_t; 0 \leq t < \infty)$ and $(\tilde{Z}_t^\infty; 0 \leq t < \infty)$ are modifications of each other. Since both processes are continuous, they are also indistinguishable. \square

Proposition 1 and (22) yield together the following.

Proposition 2. For all stopping times T of W with $E\{Y^2(T - \tau)^+\} < \infty$

$$E \left(\frac{Y^2}{2}(T - \tau)^+ \right) = E \left(\log \psi_T + \frac{1}{2} \int_0^T \{ \widehat{R}_s^2 - (\hat{R}_s)^2 \} ds \right)$$

and

$$L(c, T) = E \left(\frac{1}{\psi_T} + c \log \psi_T + \frac{c}{2} \int_0^T \{ \widehat{R}_s^2 - (\hat{R}_s)^2 \} ds \right).$$

Let

$$V_t = \frac{1}{2} \int_0^t \{ \widehat{R}_s^2 - (\widehat{R}_s)^2 \} ds \tag{24}$$

denote the remainder term in the preceding formula. We now investigate $E(V_T)$.

To state our next result, we need some further notation. Let $\mathcal{F}_t^{W,\tau}$ denote the σ -algebra $\sigma(W_s; 0 \leq s \leq t, \tau)$. The filtration $\mathcal{F}^{W,\tau}$ corresponds to a hypothetical observer who knows τ beforehand and is ignorant about the drift Y . Let

$$\widehat{R}_t^{(\tau)} = \frac{W_t - W_\tau}{t - \tau + 1} 1_{\{\tau \leq t\}}.$$

We have $E(R_t | \mathcal{F}_t^{W,\tau}) = \widehat{R}_t^{(\tau)}$.

It is convenient to introduce one more probability measure on $\sigma(W_s; 0 \leq s < \infty, Y, \tau)$. Let P_0 denote the probability measure under which $(W_t - Yt; 0 \leq t < \infty)$ is a standard Brownian motion. The notation may seem cumbersome on the first view, but P_0 can be considered as the measure with ‘immediate change’ at $\nu = 0$. The distribution of W under P_0 is given by $\int_{-\infty}^{+\infty} P_{(\theta,0)} G(d\theta)$. We shall use the likelihood ratio of P with respect to P_0 relative to the filtrations \mathcal{F} and $\mathcal{F}^{W,\tau}$ later. The probability measures P and P_0 are equivalent on the σ -algebra $\sigma(W_s; 0 \leq s < \infty, \tau)$. Let

$$N_\infty = \frac{dP}{dP_0} \Big|_{\sigma(W_s; 0 \leq s < \infty)}, \quad N_\infty^{(\tau)} = \frac{dP}{dP_0} \Big|_{\sigma(W_s; 0 \leq s < \infty, \tau)}.$$

The quantities $E(\log N_\infty)$ and $E(\log N_\infty^{(\tau)})$ appear below (see (28)). We have

$$N_\infty = \int_0^\infty e^{-YW_s + (Y^2/2)s} \rho(ds), \quad N_\infty^{(\tau)} = e^{-YW_\tau + (Y^2/2)\tau}.$$

It is easy to see that

$$E(\log N_\infty^{(\tau)}) = \left(\int_{-\infty}^{+\infty} \frac{y^2}{2} G(dy) \right) \left(\int_0^{+\infty} s \rho(ds) \right) < \infty. \tag{25}$$

Since $\mathcal{F}_t = \sigma(W_s; 0 \leq s < \infty) \subset \sigma(W_s; 0 \leq s < \infty, \tau)$, we get by the monotonicity of the Kullback–Leibler information that $E(\log N_\infty) \leq E(\log N_\infty^{(\tau)})$. Moreover

$$E(\log N_\infty) = \int_{-\infty}^\infty \int_{[0, \infty)} \left\{ E_{(\theta, \nu)} \log \left(\int_{[0, \infty)} e^{-\theta W_s + (\theta^2/2)s} \rho(ds) \right) \right\} \rho(d\nu) G(d\theta). \tag{26}$$

Proposition 3. *For all stopping times T of W with $E\{Y^2(T - \tau)^+\} < \infty$*

$$E(V_T) = \frac{1}{2} E[\log \{(T - \tau)^+ + 1\}] + \frac{1}{2} E \left(\int_0^T (\widehat{R}_s^{(\tau)} - \widehat{R}_s)^2 ds \right). \tag{27}$$

The remainder term $\frac{1}{2} E \int_0^T (\widehat{R}_s^{(\tau)} - \widehat{R}_s)^2 ds$ in (27) stays bounded. We have

$$\frac{1}{2} E \left(\int_0^\infty (\widehat{R}_s^{(\tau)} - \widehat{R}_s)^2 ds \right) = E(\log N_\infty^{(\tau)}) - E(\log N_\infty) < \infty. \tag{28}$$

Proof. The following equation holds:

$$E\{(R_t - \hat{R}_t^{(\tau)})^2 | \mathcal{F}_t^{W, \tau}\} = \frac{1}{t - \tau + 1} 1_{\{\tau \leq t\}}. \tag{29}$$

Since

$$E\{(R_t - \hat{R}_t^{(\tau)})(\hat{R}_t^{(\tau)} - \hat{R}_t) | \mathcal{F}_t^{W, \tau}\} = 0,$$

a Fubini type of argument provides

$$E \int_0^T \{\widehat{R}_s^2 - (\hat{R}_s)^2\} ds = E \int_0^T (R_s - \hat{R}_s^{(\tau)})^2 ds + E \int_0^T (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds.$$

Note that $\hat{R}^{(\tau)} = E(R_t | \mathcal{F}_t^{W, \tau})$. Then (29) implies that

$$E \int_0^T (R_s - \hat{R}_s^{(\tau)})^2 ds = E \int_0^{(T-\tau)^+} \frac{1}{s+1} ds = E \log \{(T - \tau)^+ + 1\}.$$

This yields (27). Lemma 2 below provides (28). □

3. The stopping times S_b

In this section we prove the asymptotic expansion of the remainder term V_{S_b} as $b \rightarrow \infty$ which we have already stated in (15) (see Proposition 3 below). This result immediately yields a proof of Theorem 2. We also obtain an asymptotic expansion of the weighted mean delay

$$E\left(\frac{Y^2}{2}(S_b - \tau)^+\right)$$

for $b \rightarrow \infty$ (see Proposition 6 below).

Obviously $P(\lim_{b \rightarrow \infty} S_b = +\infty) = 1$. According to Proposition 3 above we have to evaluate $E \log \{(S_b - \tau)^+ + 1\}$ for large b . In a first step we study the asymptotic behaviour of ψ_t as t tends to infinity. One easily obtains

$$\log \psi_t = \frac{W_t^2}{2(t+1)} - \frac{1}{2} \log(t+1) + \log\left(\frac{dP}{dP_0} \Big|_{\mathcal{F}_t}\right).$$

This shows that the process $\log \psi_t$ behaves for large times t in first order as $(Y^2/2)t$. Note that the P martingale $dP_0/dP | \mathcal{F}_t$ converges to N_∞ as $t \rightarrow \infty$ (see Lemma 2 below). On the set $\cap_{k=1}^\infty \{S_k < \infty\}$ we have $\psi_{S_b} = b$ for all $b \geq 1$. A straightforward argument therefore yields

$$\frac{S_b}{\log b} \rightarrow \frac{2}{Y^2} P\text{-almost sure} \tag{30}$$

as $b \rightarrow \infty$. The following result is now intuitively plausible.

Proposition 4. As $b \rightarrow \infty$

$$E[\log \{(S_b - \tau)^+ + 1\}] = \log \log b - \int_{-\infty}^{+\infty} \log \left(\frac{y^2}{2} \right) G(dy) + o(1).$$

Proof. Let

$$\xi_b = \frac{(Y^2/2)\{(S_b - \tau)^+ + 1\}}{\log b}, \quad E_b = \left\{ \frac{Y^2}{2}(S_b - \tau)^+ \geq \frac{1}{12} \log b \right\}.$$

The assertion is equivalent to $\lim_{b \rightarrow \infty} E(\log \xi_b) = 0$. On the event E_b holds $\xi_b \geq \frac{1}{12}$. Lemma 3 below implies that

$$\sup_{2 \leq b < \infty} E(e^{1_{E_b} \log \xi_b}) \leq 1 + \sup_{2 \leq b < \infty} E(\xi_b) < \infty.$$

These bounds yield together with (30) $\lim_{b \rightarrow \infty} E(1_{E_b} \log \xi_b) = 0$. Lemma 4 below provides $\lim_{b \rightarrow \infty} E(1_{E_b^c} \log \log b) = 0$ and

$$\lim_{b \rightarrow \infty} E \left\{ 1_{E_b^c} \log \left(\frac{Y^2}{2} \right) \right\} = 0 = \lim_{b \rightarrow \infty} P(E_b^c).$$

Therefore $\lim_{b \rightarrow \infty} E(1_{E_b^c} \log \xi_b) = 0$. □

We now come to the asymptotic expansion of $E(V_{S_b})$.

Proposition 5. As $b \rightarrow \infty$

$$E(V_{S_b}) = \frac{1}{2} \log \log b - \frac{1}{2} \int_{-\infty}^{\infty} \log \left(\frac{y^2}{2} \right) G(dy) + E(\log N_{\infty}^{(\tau)}) - E(\log N_{\infty}) + o(1).$$

Proof. Lemma 3 below provides that $E\{Y^2(S_b - \tau)^+\} < \infty$. From Proposition 3 we obtain for $b > 1$

$$E(V_{S_b}) = \frac{1}{2} E[\log \{(S_b - \tau)^+ + 1\}] + \frac{1}{2} E \left(\int_0^{S_b} (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds \right).$$

Since $P(\lim_{b \rightarrow \infty} S_b = +\infty) = 1$, Lemma 2 below implies that

$$\lim_{b \rightarrow \infty} \frac{1}{2} E \left(\int_0^{S_b} (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds \right) = E(\log N_{\infty}^{(\tau)}) - E(\log N_{\infty}).$$

Proposition 4 now yields the assertion. □

Proof of Theorem 2. Proposition 2 yields

$$L(c, S_{\beta(c)}) = E \left(\frac{1}{\psi_{S_{\beta(c)}}} + c \log \psi_{S_{\beta(c)}} + c V_{S_{\beta(c)}} \right).$$

For sufficiently small c it holds that

$$E\left(\frac{1}{\psi_{S_{\beta(c)}}} + c \log \psi_{S_{\beta(c)}}\right) = c + c \log c.$$

Since $\lim_{c \rightarrow 0} \beta(c) = +\infty$, we obtain from Proposition 5 above that

$$E(V_{S_{\beta(c)}}) = \frac{1}{2} \log \log \left(\frac{1}{c}\right) - \frac{1}{2} \int_{-\infty}^{\infty} \log \left(\frac{y^2}{2}\right) G(dy) + E(\log N_{\infty}^{(\tau)}) - E(\log N_{\infty}) + o(1)$$

as $c \rightarrow 0$. The quantities $E(\log N_{\infty}^{(\tau)})$ and $E(\log N_{\infty})$ are evaluated in (25) and (26) above. This yields the assertion. □

Proposition 2 and Proposition 5 also immediately yield the following.

Proposition 6. *As $b \rightarrow \infty$*

$$E\left(\frac{Y^2}{2}(S_b - \tau)^+\right) = \log b + \frac{1}{2} \log \log b - \frac{1}{2} \int_{-\infty}^{\infty} \log \left(\frac{y^2}{2}\right) G(dy) + E(\log N_{\infty}^{(\tau)}) - E(\log N_{\infty}) + o(1).$$

4. The minimal Bayes risk

In order to prove the asymptotic Bayes optimality of the stopping times $S_{\beta(c)}$ we compare their performance with the performance of c^2 -optimal solutions. Let \tilde{S}_c for $0 < c \leq 1$ be a c^2 -optimal stopping rule, i.e., a stopping time with $L(c, \tilde{S}_c) \leq L_c^* + c^2$. Without loss of generality we may assume that $\tilde{S}_c \leq S_{\beta(c)}$.

Proof of Theorem 1. The function $g_c(x) = 1/x + c \log x$ assumes a unique minimum over the interval $(0, \infty)$ at $x = \beta(c)$. Proposition 2 and Proposition 3 therefore yield together

$$\begin{aligned} 0 &\leq L(c, S_{\beta(c)}) - L_c^* \leq L(c, S_{\beta(c)}) - L(c, \tilde{S}_c) + c^2 \\ &\leq E\left(\frac{1}{\psi_{S_{\beta(c)}}} + c \log \psi_{S_{\beta(c)}} - \frac{1}{\psi_{\tilde{S}_c}} - c \log \psi_{\tilde{S}_c}\right) + cE(V_{S_{\beta(c)}} - V_{\tilde{S}_c}) + c^2 \\ &\leq cE(V_{S_{\beta(c)}} - V_{\tilde{S}_c}) + c^2 \\ &\leq cE\left(\int_{\tilde{S}_c}^{\infty} (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds\right) + cE\left\{\log \left(\frac{(S_{\beta(c)} - \tau)^+ + 1}{(\tilde{S}_c - \tau)^+ + 1}\right)\right\} + c^2. \end{aligned}$$

Note that Lemma 3 below provides that $E(Y^2(S_{\beta(c)} - \tau)^+) < \infty$. According to Lemma 5 below we have $\lim_{c \rightarrow 0} P(\tilde{S}_c < M) = 0$ for all $M > 0$ and so Lemma 2 gives

$$\lim_{c \rightarrow 0} \mathbb{E} \left(\int_{\hat{S}_c}^{\infty} (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds \right) = 0.$$

Lemma 6 below now yields the assertion. □

5. Supporting lemmata

In this section we collect some useful lemmas which are of a more technical nature.

Let

$$\bar{W}_t^{(\tau)} = W_t - \int_0^t \hat{R}_s^{(\tau)} ds.$$

The process $(\bar{W}_t^{(\tau)}, \mathcal{F}_t^{W, \tau})$ is a standard Brownian motion under P . We recall that $W_t - Yt$ is a standard Brownian motion under P_0 . The distribution of W under P_0 equals $\int_{-\infty}^{+\infty} P_{(\theta, 0)} G(d\theta)$. We now have a closer look at the likelihood ratio of P with respect to P_0 relative to the filtrations \mathcal{F} and $\mathcal{F}^{W, \tau}$. Let

$$N_t = \frac{dP}{dP_0} \Big|_{\mathcal{F}_t}, \quad N_t^{(\tau)} = \frac{dP}{dP_0} \Big|_{\mathcal{F}_t^{W, \tau}}.$$

Lemma 1. For all $0 \leq t < \infty$

$$N_t = \exp \left\{ \int_0^t \left(\hat{R}_s - \frac{W_s}{s+1} \right) d\bar{W}_s - \frac{1}{2} \int_0^t \left(\hat{R}_s - \frac{W_s}{s+1} \right)^2 ds \right\}$$

and

$$N_t^{(\tau)} = \exp \left\{ \int_0^t \left(\hat{R}_s^{(\tau)} - \frac{W_s}{s+1} \right) d\bar{W}_s^{(\tau)} - \frac{1}{2} \int_0^t \left(\hat{R}_s^{(\tau)} - \frac{W_s}{s+1} \right)^2 ds \right\}.$$

Proof. Since

$$\frac{dP}{dP_0} \Big|_{\mathcal{F}_t} = \frac{dP}{dP_\infty} \Big|_{\mathcal{F}_t} \frac{dP_\infty}{dP_0} \Big|_{\mathcal{F}_t}, \quad \frac{dP}{dP_0} \Big|_{\mathcal{F}_t^{W, \tau}} = \frac{dP}{dP_\infty} \Big|_{\mathcal{F}_t^{W, \tau}} \frac{dP_\infty}{dP_0} \Big|_{\mathcal{F}_t^{W, \tau}},$$

we obtain

$$\frac{dP}{dP_0} \Big|_{\mathcal{F}_t} = (t+1)^{1/2} \exp \left(-\frac{W_t^2}{2(t+1)} \right) \psi_t \tag{31}$$

and

$$\frac{dP}{dP_0} \Big|_{\mathcal{F}_t^{W, \tau}} = \left(\frac{t+1}{(t-\tau)^+ + 1} \right)^{1/2} \exp \left(\frac{(W_t - W_{t \wedge \tau})^2}{2[(t-\tau)^+ + 1]} \right) \exp \left(-\frac{W_t^2}{2(t+1)} \right). \tag{32}$$

Then (31) yields after some algebra

$$d(\log N_t) = \frac{1}{2} \left(\hat{R}_t - \frac{W_t}{t+1} \right)^2 dt + \left(\hat{R}_t - \frac{W_t}{t+1} \right) d\bar{W}_t, \tag{33}$$

and (32) yields after some algebra

$$d(\log N_t^{(\tau)}) = \frac{1}{2} \left(\hat{R}_t^{(\tau)} - \frac{W_t}{t+1} \right)^2 dt + \left(\hat{R}_t^{(\tau)} - \frac{W_t}{t+1} \right) d\bar{W}_t^{(\tau)}. \tag{34}$$

□

Lemma 2. For all stopping times T of W with $E\{Y^2(T - \tau)^+\} < \infty$ it holds that

$$E \left(\int_0^T \hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds \right) = E \left\{ \int_0^T \left(\hat{R}_s^{(\tau)} - \frac{W_s}{s+1} \right)^2 ds \right\} - E \left\{ \int_0^T \left(\hat{R}_s - \frac{W_s}{s+1} \right)^2 ds \right\}. \tag{35}$$

Moreover

$$\frac{1}{2} E \left\{ \int_0^\infty \left(\hat{R}_s^{(\tau)} - \frac{W_s}{s+1} \right)^2 ds \right\} = E(\log N_\infty^{(\tau)}) < \infty \tag{36}$$

and

$$\frac{1}{2} E \left\{ \int_0^\infty \left(\hat{R}_s - \frac{W_s}{s+1} \right)^2 ds \right\} = E(\log N_\infty) < \infty. \tag{37}$$

Proof. Let $U_s = W_s/(s + 1)$ for $s \in [0, \infty)$. For all $s \in [0, \infty)$ and $A \in \mathcal{F}_s$ it holds that

$$\begin{aligned} E\{1_A(\hat{R}_s^{(\tau)} - \hat{R}_s)^2\} &= E\{1_A(\hat{R}_s^{(\tau)} - U_s + U_s - \hat{R}_s)^2\} \\ &= E\{1_A(\hat{R}_s^{(\tau)} - U_s)^2\} - 2E\{1_A(U_s - \hat{R}_s^{(\tau)})(U_s - \hat{R}_s)\} \\ &\quad + E\{1_A(U_s - \hat{R}_s)^2\} \\ &= E\{1_A(\hat{R}_s^{(\tau)} - U_s)^2\} - E\{1_A(U_s - \hat{R}_s)^2\}. \end{aligned}$$

Note that $E(\hat{R}_s^{(\tau)} | \mathcal{F}_s) = \hat{R}_s$. If we show that

$$E \left(\int_0^\infty (\hat{R}_s - U_s)^2 ds \right) < \infty, \quad E \left(\int_0^\infty (\hat{R}_s^{(\tau)} - U_s)^2 ds \right) < \infty,$$

a Fubini type of argument then yields (35). Therefore it is sufficient to prove (36) and (37). The submartingales $(\log N_t, \mathcal{F}_t)$ and $(\log N_t^{(\tau)}, \mathcal{F}_t^{W, \tau})$ are uniformly integrable. Hence $\lim_{t \rightarrow \infty} E(\log N_t) = E(\log N_\infty)$ and $\lim_{t \rightarrow \infty} E(\log N_t^{(\tau)}) = E \log(N_\infty^{(\tau)})$. A straightforward calculation shows that, for all $t \geq 0$,

$$E \left(\int_0^t (\hat{R}_s - U_s)^2 ds \right) < \infty, \quad E \left(\int_0^t (\hat{R}_s^{(\tau)} - U_s)^2 ds \right) < \infty.$$

Lemma 1 therefore implies that

$$\frac{1}{2}E\left(\int_0^t (\hat{R}_s - U_s)^2 ds\right) = E(\log N_t), \quad \frac{1}{2}E\left(\int_0^t (\hat{R}_s^{(\tau)} - U_s)^2 ds\right) = E(\log N_t^{(\tau)}).$$

□

Lemma 3. *There exists a constant $A \geq 0$ such that for all $b \geq 1$*

$$E\left(\frac{Y^2}{2}(S_b - \tau)^+\right) \leq 2 \log b + A.$$

Proof. Proposition 2 yields for all positive integers n

$$E\left(\frac{Y^2}{2}(S_b \wedge n - \tau)^+\right) \leq \log b + \frac{1}{2}E\left(\int_0^{S_b \wedge n} \{\widehat{R}_s^2 - (\hat{R}_s)^2\} ds\right).$$

Proposition 3 and Lemma 2 now give

$$E\left(\frac{Y^2}{2}(S_b \wedge n - \tau)^+\right) \leq \log b + \frac{1}{2}E \log \{(S_b \wedge n - \tau)^+ + 1\} + \frac{1}{2} \int_0^\infty s \rho(ds).$$

Since $\log x \leq x$ and $|E \log(Y^2)| < \infty$, we obtain the assertion.

□

Lemma 4. *As $b \rightarrow \infty$*

$$P\left(\frac{Y^2}{2}(S_b - \tau)^+ \leq \frac{1}{12} \log b\right) = o\left(\frac{1}{\log b}\right).$$

Proof. Our arguments are similar to those of Pollak (1987, p. 772). Since $E(Y^2\tau) < \infty$, it is more than sufficient to prove that

$$P\left(\frac{Y^2}{2}S_b \leq \frac{1}{6} \log b\right) \leq 2b^{-1/3}$$

for $b > 1$. Let $\eta_b = (1/3Y^2) \log b$ and $\eta_{b,y} = (1/3y^2) \log b$. Let $\mathcal{F}_t^{W,\tau,Y}$ denote the σ -algebra $\sigma(W_s; 0 \leq s \leq t, \tau, Y)$. For arbitrary $M > 0$ it holds that

$$\begin{aligned} P(S_b \leq \eta_b) &\leq P\left(\frac{dP}{dP_\infty} \Big|_{\mathcal{F}_{\eta_b}^{W,\tau,Y}} \geq M\right) + MP_\infty(S_b \leq \eta_b) \\ &\leq \frac{1}{M} E_\infty \left(\frac{dP}{dP_\infty} \Big|_{\mathcal{F}_{\eta_b}^{W,\tau,Y}}\right)^2 + M \int_{-\infty}^{+\infty} P_\infty(S_b \leq \eta_{b,y}) G(dy). \end{aligned}$$

The process ψ_t is a positive martingale under P_∞ . Therefore the Doob inequality yields

$$P_\infty(S_b \leq \eta_{b,y}) = P_\infty\left(\max_{0 \leq s \leq \eta_{b,y}} \psi_s \geq b\right) \leq \frac{1}{b}. \tag{38}$$

Since

$$\frac{dP}{dP_\infty} \Big|_{\mathcal{F}_{\eta_b}^{W,\tau,Y}} = e^{Y(W_{\eta_b} - W_{\eta_b \wedge \tau}) - (Y^2/2)(\eta_b - \tau)^+}, \quad Y^2 \eta_b = \frac{1}{3} \log b,$$

we have

$$E_\infty \left(\frac{dP}{dP_\infty} \Big|_{\mathcal{F}_{\eta_b}^{W,\tau,Y}} \right)^2 \leq b^{1/3} E_\infty (e^{2Y(W_{\eta_b} - W_{\eta_b \wedge \tau}) - (4Y^2/2)(\eta_b - \tau)^+}) = b^{1/3}. \tag{39}$$

The inequalities (38) and (39) yield for all $M > 0$ and $b > 1$

$$P(S_b \leq \eta_b) \leq M \frac{1}{b} + \frac{1}{M} b^{1/3}.$$

With $M = b^{2/3}$ we obtain the assertion. □

Let $\tilde{\beta}(c) = 1/c(1 - \log c) = \beta(c)/(1 - \log c)$. In the next lemma we compare c^2 -optimal stopping rules \tilde{S}_c with $S_{\tilde{\beta}(c)}$ in order to bound such stopping rules from below.

Lemma 5. *Let $(\tilde{S}_c; 0 < c \leq 1)$ be stopping times of W with $L(c, \tilde{S}_c) \leq L_c^* + c^2$. Then, as $c \rightarrow 0$,*

$$P(\tilde{S}_c \leq S_{\tilde{\beta}(c)}) = O\left(\frac{\log \log (1/c)}{\log (1/c)}\right).$$

Proof. On the event $\{\tilde{S}_c \leq S_{\tilde{\beta}(c)} < \infty\}$ it holds that $\psi_{\tilde{S}_c} \leq \psi_{S_{\tilde{\beta}(c)}} \leq 1/c$ and so

$$\frac{1}{\psi_{\tilde{S}_c}} + c \log \psi_{\tilde{S}_c} \geq \frac{1}{\psi_{S_{\tilde{\beta}(c)}}} + c \log \psi_{S_{\tilde{\beta}(c)}}.$$

This gives

$$\begin{aligned} L_c^* + c^2 &\geq L(c, \tilde{S}_c) \geq E \left\{ \frac{1}{\psi_{\tilde{S}_c}} + c \log \psi_{\tilde{S}_c} \right\} \\ &\geq c + \log \left(\frac{1}{c} \right) + \left\{ \frac{1}{\tilde{\beta}(c)} + c \log \tilde{\beta}(c) - c - c \log \left(\frac{1}{c} \right) \right\} P(\tilde{S}_c \leq S_{\tilde{\beta}(c)}) \\ &= c + c \log \left(\frac{1}{c} \right) + \{1 + o(1)\} c \log \left(\frac{1}{c} \right) P(\tilde{S}_c \leq S_{\tilde{\beta}(c)}) \end{aligned}$$

when $c \rightarrow 0$ and so

$$P(\tilde{S}_c \leq S_{\tilde{\beta}(c)}) \leq \{1 + o(1)\} \frac{(1/c)\{L_c^* + c^2 - c - c \log (1/c)\}}{\log (1/c)}.$$

Theorem 1 yields for $c \rightarrow 0$

$$L(c, S_{\beta(c)}) = c + c \log \left(\frac{1}{c} \right) + \frac{1}{2} c \log \log \left(\frac{1}{c} \right) + O(c),$$

and so

$$L_c^* \leq c + c \log \left(\frac{1}{c} \right) + \frac{c}{2} \log \log \left(\frac{1}{c} \right) + O(c) \text{ when } c \rightarrow 0. \tag{40}$$

Therefore

$$\frac{1}{c} \left\{ L_c^* + c^2 - c - c \log \left(\frac{1}{c} \right) \right\} \leq \frac{1}{2} \log \log \left(\frac{1}{c} \right) + O(1)$$

as $c \rightarrow 0$. □

Lemma 6. Let $(\tilde{S}_c; 0 < c \leq 1)$ be stopping times of W with $L(c, \tilde{S}_c) \leq L_c^* + c^2$ and $\tilde{S}_c \leq S_{\beta(c)}$. Then

$$\lim_{c \rightarrow 0} E \left\{ \log \left(\frac{(S_{\beta(c)} - \tau)^+ + 1}{(\tilde{S}_c - \tau)^+ + 1} \right) \right\} = 0.$$

Proof. Proposition 4 implies that

$$E[\log \{(S_{\beta(c)} - \tau)^+ + 1\}] = \log \log \left(\frac{1}{c} \right) - \int_{-\infty}^{+\infty} \log \left(\frac{y^2}{2} \right) G(dy) + o(1)$$

as $c \rightarrow 0$. Since $\tilde{S}_c \leq S_{\beta(c)}$, it is therefore sufficient to show that

$$\liminf_{c \rightarrow 0} \left\{ E[\log \{(\tilde{S}_c - \tau)^+ + 1\}] - \log \log \left(\frac{1}{c} \right) + \int_{-\infty}^{+\infty} \log \left(\frac{y^2}{2} \right) G(dy) \right\} \geq 0.$$

It holds that

$$\begin{aligned} E[\log \{(\tilde{S}_c - \tau)^+ + 1\}] &\geq E[1_{\{S_{\beta(c)} \leq \tilde{S}_c\}} \log \{(S_{\beta(c)} - \tau)^+ + 1\}] \\ &= E[\log \{(S_{\beta(c)} - \tau)^+ + 1\}] \\ &\quad - E[1_{\{S_{\beta(c)} > \tilde{S}_c\}} \log \{(S_{\beta(c)} - \tau)^+ + 1\}]. \end{aligned}$$

Lemma 5 yields $\lim_{c \rightarrow 0} P(S_{\beta(c)} > \tilde{S}_c) = 0$ and Proposition 4 implies that

$$E[\log \{(S_{\beta(c)} - \tau)^+ + 1\}] = \log \log \left(\frac{1}{c} \right) - \int_{-\infty}^{+\infty} \log \left(\frac{y^2}{2} \right) G(dy) + o(1).$$

It therefore suffices to show that

$$\limsup_{c \rightarrow 0} E[1_{\{S_{\beta(c)} > \tilde{S}_c\}} \log \{(S_{\beta(c)} - \tau)^+ + 1\}] \leq 0.$$

There exists a constant $A > 0$ such that $\log(x + 1) \leq Ax^{1/4}$ for all $x \geq 0$. Hence

$$\begin{aligned} E \left\{ 1_{\{S_{\tilde{\beta}(c)} > \tilde{s}_c\}} \log \left(\frac{Y^2}{2} \{(S_{\tilde{\beta}(c)} - \tau)^+ + 1\} \right) \right\} &\leq AE \left\{ 1_{\{S_{\tilde{\beta}(c)} > \tilde{s}_c\}} \left(\frac{Y^2}{2} (S_{\tilde{\beta}(c)} - \tau)^+ + 1 \right)^{1/4} \right\} \\ &\leq A \left\{ E \left(\frac{Y^2}{2} (S_{\tilde{\beta}(c)} - \tau)^+ + 1 \right) \right\}^{1/4} P(S_{\tilde{\beta}(c)} > \tilde{s}_c)^{3/4}. \end{aligned}$$

Lemma 3 and Lemma 5 now yield the assertion. Note that $|\int_{-\infty}^{+\infty} \log(y^2/2)G(dy)| < \infty$ and so $\lim_{c \rightarrow 0} E(1_{\{S_{\tilde{\beta}(c)} > \tilde{s}_c\}} \log(Y^2)) = 0$.

Acknowledgement

This work is based on part of the author's Ph.D. thesis written under the guidance of Professor H.R. Lerche. The author is deeply indebted to him for posing the problem and for his constant advice and encouragement.

References

- Beibel, M. (1994a) Bayes-Optimalität in Trendänderungsmodellen mit kontinuierlicher Zeit. Ph.D. thesis, Albert-Ludwigs Universität, Freiburg.
- Beibel, M. (1994b) Bayes problems in change-point models for the Wiener process. In E. Carlstein, H.G. Müller and D. Siegmund (eds), *Change-point Problems*, pp. 1–6. Lecture Notes Monograph Series 23. Hayward, CA: Institute of Mathematical Statistics.
- Beibel, M. (1996) A note on Ritov's Bayes approach to the minimax property of the cusum procedure. *Ann. Statist.*, **24**, 1804–1812.
- Beibel, M. and Lerche, H.R. (1997) A new look at optimal stopping problems related to mathematical finance. *Statist. Sinica.*, **7**, 93–108.
- Darling, D. and Robbins, H.R. (1967) Iterated logarithm inequalities. *Proc. Nat. Acad. Sci. USA*, **57**, 1188–1192.
- Keener, R., Lerche, H.R. and Woodroffe, M. (1995) A nonlinear parking problem. *Sequential Anal.*, **14**, 247–272.
- Kohler, O. (1995) Erwartete Verzögerungen bei Trenderkennung in Bayes-Modellen mit Brownscher Bewegung. Master's thesis, Albert-Ludwigs Universität, Freiburg.
- Lerche, H.R. (1986a) An optimality property of the repeated significance test. *Proc. Nat. Acad. Sci. USA*, **83**, 1546–1548.
- Lerche, H.R. (1986) The shape of Bayes tests of power one. *Ann. Statist.*, **14**, 1030–1048.
- Liptser, R.S. and Shiriyayev, A.N. (1977) *Statistics of Random Processes*, Vol. 1. Berlin: Springer-Verlag.
- Pollak, M. (1987) Average run lengths of an optimal method of detecting a change in distribution. *Ann. Statist.*, **15**, 749–779.
- Pollak, M. and Siegmund, D. (1975) Approximations to the expected sample size of certain sequential tests. *Ann. Statist.*, **3**, 1267–1282.
- Pollak, M. and Siegmund, D. (1985) A diffusion process and its application to detecting a change in

- the drift of Brownian motion. *Biometrika*, **72**, 267–280.
- Ritov, Y. (1990) Decision theoretic optimality of the cusum procedure. *Ann. Statist.*, **18**, 464–469.
- Shiryayev, A.N. (1963) On optimum methods in quickest detection problems. *Theory Probab. Appl.*, **8**, 22–46.
- Shiryayev, A.N. (1973) *Statistical Sequential Analysis*. Translations of Mathematical Monographs 8. Providence, RI: American Mathematical Society.
- Woodroffe, M., Lerche, R. and Keener, R. (1994) A generalized parking problem. In S.S. Gupta and J.O. Berger (eds), *Statistical Decision Theory and Related Topics V*, pp. 523–532. Berlin: Springer-Verlag.
- Yor, M. (1992) Sur certaines fonctionelles exponentielles du mouvement brownien réel. *J. Appl. Probab.*, **29**, 202–208.

Received March 1996 and revised November 1996