The Annals of Mathematical Statistics
1972, Vol. 43, No. 6, 1861–1873

SEQUENTIAL PROBABILITY RATIO TEST FOR THE MEAN VALUE FUNCTION OF A GAUSSIAN PROCESS

BY P. K. BHATTACHARYA AND ROBERT P. SMITH

University of Arizona and Arkansas State University

Sequential probability ratio tests are defined for testing a simple hypothesis against a simple alternative for the mean value function of a real Gaussian process with known covariance kernel. Exact formulas are obtained for the error probabilities and the OC function using the fact that the log likelihood ratio process is a Gaussian process with independent increments and have continuous sample paths. An identity of a familiar nature holds for the expected value of the log likelihood ratio process at a random stopping time. In certain situations this identity yields an exact formula for the ASN function. Two examples are given. The analysis employs the theory of reproducing kernel Hilbert spaces.

1. Introduction. The theory of sequential tests of simple hypotheses against simple alternatives based on independent and identically distributed random variables was developed by Wald (1947) and the optimum character of such tests was established by Wald and Wolfowitz (1948). A few years later Dvoretzky, Kiefer and Wolfowitz (1953) pointed out that practically all of these results easily extend to the case of stochastic processes in continuous time provided that the latest observation is a sufficient statistic for the entire past and that the log likelihood ratios of these statistics at various points of time form a process with stationary and independent increments. In particular, this settled the problems of sequentially testing for the drift of a Brownian motion and the intensity of a Poisson process.

Wald's theory of sequential tests is based on sequential probability ratios, and if the problem of sequential testing in case of processes in continuous time is to be attacked in the same spirit, the first thing one would look for is the continuous-time analogue of sequential probability ratios. Such an attack would be feasible only when (i) for each $t > 0$ the measures P_0^t and P_1^t of the process up to time t corresponding to the null and the alternative hypotheses are equivalent, (ii) the Radon-Nikodym derivative process $\{dP_1^t/dP_0^t,\ t \geqq 0\}$ is analytically tractable, and (iii) the measures P_0 and P_1 of the entire process on $[0, \infty)$ corresponding to the null and the alternative hypotheses are orthogonal. The first two requirements are to ensure that the sequential probability ratios are defined and analytically tractable, while the last one is for the sequential probability ratio test to terminate with probability one.

Received August 2, 1971; revised March 1, 1972.

AMS 1970 *subject classifications.* Primary 62L10; Secondary 60G15.

Key words and phrases. Sequential probability ratio test, Gaussian processes, reproducing kernel Hilbert space.

The purpose of this paper is to carry out this project for testing the mean value function of a separable real Gaussian process with known covariance kernel. Feldman (1958) has proved that two Gaussian measures are either equivalent or orthogonal. Parzen (1959) has given conditions for equivalence or orthogonality when the two processes have a common covariance kernel differing only in their mean value functions, and has provided a formula for the Radon-Nikodym derivative of one measure with respect to the other in case of equivalence. These results of Parzen form the basis of our work by allowing us to understand dP_1^t/dP_0^t as a random variable for any given t (Theorem 1), and to set up the sequential probability ratio test (SPRT). We then look at the collection of all these random variables for different values of t as a stochastic process and observe a remarkable fact. The process $\{\log (dP_1^t/dP_0^t), t \geqq 0\}$ is a Gaussian process with independent (though possibly non-stationary) increments and under certain conditions has a sample continuous version (Theorem 2). Using these properties of the log likelihood ratio process we then show that a SPRT terminates with probability one (Theorem 3) and obtain formulas for the error probabilities (Theorem 4) and the operating characteristic (OC) function (Theorem 5) of a SPRT. We also obtain for certain situations a formula for the average sample number (ASN) function (Corollary to Theorem 6). Several examples are given. The theory of reproducing kernel Hilbert spaces is employed in our analysis. For a survey of this theory the reader may consult Aronszajn (1950) and Parzen (1959).

2. Preliminaries. Let $\{X(t), t \geqq 0\}$ be a separable real Gaussian process with known covariance kernel R and unknown mean value function m. We want to test the simple hypothesis $H_0: m = m_0$ against the alternative $H_1: m = m_1$ where m_0 and m_1 are given real-valued functions on $[0, \infty)$. For all θ real, let $m_\theta = (1 - \theta)m_0 + \theta m_1$.

We consider the stochastic process in its sample path formulation, i.e., we consider the sample space Ω to consist of all real-valued functions on $[0, \infty]$ and in this space we consider the smallest σ-field containing all sets $\{\omega \mid \omega(t) \leqq a\}$ for $t \geqq 0$ and a real. Then the random variables $X(t)$ are functions defined as $X(t, \omega) = \omega(t)$, and we denote by P_θ the probability measure on $(\Omega, \mathscr{A})$ corresponding to which $\{X(t), t \geqq 0\}$ is a Gaussian process with mean value function m_θ and covariance kernel R. By P_θ^t we denote the restriction of P_θ to the σ-field $\mathscr{A}^t$ of $\{X(s), 0 \leqq s \leqq t\}$.

By $H(R)$ we denote the reproducing kernel Hilbert space (RKHS) of real-valued functions on $[0, \infty)$ with reproducing kernel R and by $H(R^t)$ we denote the RKHS of real-valued functions on $[0, t]$ with reproducing kernel R^t which is the restriction of R to $[0, t] \times [0, t]$. When we mention a function f on $[0, \infty)$ in the context of $H(R^t)$, we actually mean its restriction on $[0, t]$.

We assume that the covariance kernel satisfies the following conditions.

CONDITION 1. R is continuous and for every positive integer n and $t_1 > \cdots > t_n \geqq 0$, the matrix $(\!(R(t_i, t_j))\!)$ is nonsingular.

CONDITION 2. For every $t > 0$, the restriction of $m_1 - m_0$ to $[0, t]$ is in $H(R^t)$.

CONDITION 3. Let $\psi(t) = \|m_1 - m_0\|^2_{H(R^t)}$. Then

(i) ψ is continuous.
(ii) ψ is strictly increasing and $\lim_{t\to\infty} \psi(t) = \infty$.

Let $\{T_n\}$ be an increasing sequence of finite sets converging to the set T of all nonnegative rationals, and for any positive integer n and any $\tau \geqq 0$, let $T_n{}^\tau = T_n \cap [0, \tau]$. By $R_{n,\tau}^{-1}(s, t)$, $s, t \in T_n{}^\tau$, we denote the elements of the inverse of the matrix $R_{n,\tau} = (\!(R(s, t))\!)_{s,t\in T_n{}^\tau}$ which exists by virtue of Condition 1.

We conclude this section with a theorem due to Parzen ((1959), Theorem 9A) which gives a formula for $dP_1{}^t/dP_0{}^t$.

THEOREM 1 (Parzen). *Under Conditions* 1 *and* 2, $P_1{}^\tau$ *is equivalent to* $P_0{}^\tau$ *for every* $\tau \geqq 0$, *and*

$$dP_1{}^\tau/dP_0{}^\tau = \exp\left[\langle X - m_0, m_1 - m_0\rangle_\tau - \tfrac{1}{2}\psi(\tau)\right] \tag{1}$$

where

$$\begin{aligned}\langle X - m_0, m_1 - m_0\rangle_\tau \\ = \lim_{n\to\infty} \textstyle\sum_{s,t\in T_n{}^\tau} \{X(s) - m_0(s)\}R_{n,\tau}^{-1}(s, t)\{m_1(t) - m_0(t)\}\,,\end{aligned} \tag{2}$$

the limit holding both in mean square and with probability one under P_0.

Since $m_\theta - m_0 = \theta(m_1 - m_0)$, Conditions 1, 2, 3 remain valid when m_1 is replaced by m_θ throughout. In view of this fact, the following is an immediate corollary to Theorem 1.

COROLLARY. *Under Conditions* 1 and 2, *the measures* $P_\theta{}^\tau$, θ *real, are equivalent for every* $\tau \geqq 0$.

3. Sequential probability ratio test. Following Wald (1947) we now define sequential probability ratio tests for $H_0: m = m_0$ against $H_1: m = m_1$.

DEFINITION 1. The sequential probability ratio test (SPRT) with boundaries $B < 1 < A$ for testing H_0 against H_1 is defined by the following stopping rule and terminal decision rule:

Continue observing as long as $B < dP_1{}^t/dP_0{}^t < A$. Stop and accept H_0 at the smallest t for which $dP_1{}^t/dP_0{}^t \leqq B$, or stop and reject H_0 at the smallest t for which $dP_1{}^t/dP_0{}^t \geqq A$, where $dP_1{}^t/dP_0{}^t$ is given by (1) and (2).

We call this test SPRT(A, B) for H_0 against H_1.

We shall call the process $\{Z(t) = \log(dP_1{}^t/dP_0{}^t), t \geqq 0\}$, the log likelihood ratio (LLR) process and we shall always consider a separable version of this process. In terms of the LLR process, SPRT(A, B) can be equivalently expressed as:

Continue sampling as long as $\log B < Z(t) < \log A$. Stop and accept H_0 at the smallest t for which $Z(t) \leqq \log B$, or stop and reject H_0 at the smallest t for which $Z(t) \geqq \log A$.

4. Properties of the LLR process. In this section we establish two important properties enjoyed by the LLR process. These two properties make it possible for us to investigate the behavior of SPRT(A, B).

THEOREM 2. (a) *Suppose Conditions* 1 *and* 2 *hold. Then under* P_θ, *the* LLR *process* $\{Z(t), t \geqq 0\}$ *is a Gaussian process with mean value function*

$$E_\theta[Z(t)] = (\theta - \tfrac{1}{2})\psi(t)$$

and covariance kernel

$$\operatorname{Cov}_\theta [Z(s), Z(t)] = \psi(\min(s, t))$$

where $\psi(t) = \|m_1 - m_0\|^2_{H(R^t)}$.

(b) *If Condition* 3(i) *also holds, then under every* P_θ, *almost all sample paths of the* LLR *process are continuous.*

REMARK. From an examination of the proof given below it can be seen that this theorem remains true (with obvious modifications) if in the definition of $Z(t)$ the function $m_1 - m_0$ is replaced by any function f whose restriction to $[0, t]$ is in $H(R^t)$ for all $t \geqq 0$. In fact, these ideas have been used by one of the authors (see Bhattacharya (1971), Theorem 1) to transform a class of Gaussian processes to the Brownian motion.

PROOF OF THEOREM 2. Argue under P_θ. Then $\{Y(t) = X(t) - m_\theta(t), t \geqq 0\}$ is a Gaussian process with mean 0 and covariance R. Let $d = m_1 - m_0$, $\langle Y, d\rangle_{n,\tau} = \sum_{s,t \in T_n^\tau} Y(s)R^{-1}_{n,\tau}(s, t)\, d(t)$, and $\langle Y, d\rangle_\tau = \lim_{n\to\infty} \langle Y, d\rangle_{n,\tau}$. By Theorem 1, $\langle Y, d\rangle_\tau$ exists a.s. and in mean square. Since $Z(\tau) = \langle Y, d\rangle_\tau + (\theta - \frac{1}{2})\psi(\tau)$, it suffices to show (for part (a) of the theorem) that $\{\langle Y, d\rangle_\tau, \tau \geqq 0\}$ is a Gaussian process with mean 0 and covariance ψ (min (s, t)). Now let $L_2(Y)$ denote the linear space of the Gaussian process $\{Y(t), t \geqq 0\}$ and let λ denote the congruence between $H(R)$ and $L_2(Y)$ for which $\lambda(R(\cdot, t)) = Y(t)$ holds for all t. Then by Theorem 9B of Parzen (1959),

$$\langle Y, d\rangle_{n,\tau} = \lambda(E^*[d \mid R_t, t \in T_n{}^\tau]) , \qquad n = 1, 2, \cdots$$

is a wide-sense martingale for every $\tau \geqq 0$ where E^* stands for projection and $R_t(\cdot) = R(\cdot, t)$, and converges to

$$\langle Y, d\rangle_\tau = \lambda(E^*[d \mid R_t, t \in \textstyle\bigcup_{n=1}^\infty T_n{}^\tau]) .$$

Since R is continuous and $\bigcup_{n=1}^\infty T_n{}^\tau$ is dense in $[0, \tau]$, the subspace of $H(R)$ spanned by $\{R_t, t \in \bigcup_{n=1}^\infty T_n{}^\tau\}$ is the same as the one spanned by $\{R_t, t \leqq \tau\}$. Thus

$$\langle Y, d\rangle_\tau = \lambda(E^*[d \mid R_t, t \leqq \tau]) . \tag{3}$$

Since the random variables $\langle Y, d\rangle_\tau$ are in $L_2(Y)$, all finite-dimensional distributions of $\{\langle Y, d\rangle_\tau, \tau \geqq 0\}$ are Gaussian with mean 0. Now consider $\sigma \leqq \tau$. Since

$$E^*\{E^*[d \mid R_t, t \leqq \tau] \mid R_t, t \leqq \sigma\} = E^*[d \mid R_t, t \leqq \sigma] ,$$

we have

$$(E^*[d \mid R_t, t \leqq \tau] - E^*[d \mid R_t, t \leqq \sigma], E^*[d \mid R_t, t \leqq \sigma])_{H(R)} = 0 .$$

This, in conjunction with (3) gives

$$\begin{aligned}\mathrm{Cov}\,[\langle Y, d\rangle_\sigma, \langle Y, d\rangle_\tau] &= (E^*[d \mid R_t, t \leqq \sigma], E^*[d \mid R_t, t \leqq \sigma]) \\ &= \|E^*[d \mid R_t, t \leqq \sigma]\|^2_{H(R)} \\ &= \|d\|^2_{H(R^\sigma)} = \psi(\sigma)\,,\end{aligned}$$

and that concludes the proof of part (a). To prove part (b), we note that $\{Z(\tau), \tau \geqq 0\}$ is a separable Gaussian process with independent increments and by Theorem 6, Chapter IV, Section 5 of Gikhman and Skorokhod (1965), has a.s. continuous sample paths by virtue of Condition 3(i).

5. Acceptance region, rejection region and stopping time of the SPRT.

DEFINITION 2. The acceptance region and the rejection region of SPRT(A, B) are defined respectively as $S_0 = \bigcup_{t \geqq 0} S_0(t)$ and $S_1 = \bigcup_{t \geqq 0} S_1(t)$ where $S_0(t)$ and $S_1(t)$ are given by

$$\begin{aligned} S_0(t) &= \{\omega \mid Z(0, \omega) \leqq \log B\}\,, & t = 0 \\ &= \{\omega \mid \log B < Z(s, \omega) < \log A,\ 0 \leqq s < t \text{ and } Z(t, \omega) \leqq \log B\}\,, & t > 0 \\ S_1(t) &= \{\omega \mid Z(0, \omega) \geqq \log A\}\,, & t = 0 \\ &= \{\omega \mid \log B < Z(s, \omega\} < \log A,\ 0 \leqq s < t \text{ and } Z(t, \omega) \geqq \log A\}\,, & t > 0\,. \end{aligned}$$

THEOREM 3. *Suppose Conditions* 1, 2, 3(ii) *hold. Then the* SPRT(A, B) *terminates with probability one under all* P_θ.

PROOF. Condition 3 implies the existence of an infinite sequence $\{t_n\}$ such that for some $\delta > 0$, $\psi(t_{n+1}) - \psi(t_n) = \delta$ for all n. Under $P_{\frac{1}{2}}$, $\zeta_n = Z(t_{n+1}) - Z(t_n)$, $n = 1, 2, \cdots$ are independent random variables, ζ_n being normally distributed with mean 0 and variance δ Following the same arguments as in the case of discrete sampling, we see that SPRT(A, B) terminates with probability one under $P_{\frac{1}{2}}$. Since all other P_θ are equivalent to $P_{\frac{1}{2}}$, the theorem is proved. In fact it is enough to have $\{t_n\}$ satisfy $\psi(t_{n+1}) - \psi(t_n) \geqq \delta$, and Condition 3(ii) guarantees the existence of such a sequence.

From now on we shall disregard those sample paths ω for which SPRT(A, B) does not terminate.

DEFINITION 3. The stopping time of SPRT(A, B) is a function T on Ω defined as

$$T(\omega) = t \qquad \text{if} \quad \omega \in S_0(t) \cup S_1(t)\,.$$

REMARK. From Theorems 2 and 3, the following conclusions are immediate.

(i) Both S_0 and S_1 are measurable and $S_1 = S_0^c$ (except for those ω that we are disregarding).

(ii) T is a random variable on Ω and for each $t \geqq 0$, $\{\omega \mid T(\omega) \leqq t\}$ belongs to the σ-field of $\{Z(s), 0 \leqq s \leqq t\}$, i.e., T is a random stopping time of the LLR process.

(iii) If $T(\omega) = 0$, then $Z(T(\omega), \omega) = Z(0, \omega)$, otherwise

$$\begin{aligned} Z(T(\omega), \omega) &= \log B \qquad \text{if} \quad \omega \in S_0 \\ &= \log A \qquad \text{if} \quad \omega \in S_1 . \end{aligned}$$

The subsequent theorems of this paper rely on Theorems 2 and 3 in which all the Conditions 1, 2, 3 have been used. From now on we shall not state these conditions explicitly though everything depends on them.

In the rest of this paper we shall study the error probabilities, the OC function and the ASN function of SPRT(A, B). These results will be derived by conditioning the process $\{Z(t), t \geqq 0\}$ by the random variable $Z(0)$ and then making a suitable transformation on the time axis. By means of this device we shall relate our problems to the corresponding ones for the Brownian motion and use the results of Dvoretzky, Kiefer and Wolfowitz (1953). The following lemmas are needed for this purpose.

LEMMA 1. *Under P_θ, the conditional process*

$$\{Z_1(t) = Z(t) - Z(0), t \geqq 0\}$$

given the random variable $Z(0)$ is a Gaussian process with mean value function

$$E_\theta[Z_1(t) \mid Z(0)] = (\theta - \tfrac{1}{2})\{\psi(t) - \psi(0)\}$$

and covariance kernel

$$\operatorname{Cov}_\theta [Z_1(s), Z_1(t) \mid Z(0)] = \psi(\min(s, t)) - \psi(0) .$$

PROOF. By Theorem 2, the LLR process has independent increments. Hence the distribution of the conditional process $\{Z_1(t), t \geqq 0\}$ given $Z(0)$ is the same as that of the unconditional process $\{Z_1(t), t \geqq 0\}$ which is a Gaussian process and $E_\theta[Z_1(t)] = (\theta - \frac{1}{2})\{\psi(t) - \psi(0)\}$ and $\operatorname{Cov}_\theta [Z_1(s), Z_1(t)] = \psi(\min(s, t)) - \psi(0)$. This completes the proof.

We now define a function γ by the relation

$$\psi(\gamma(t)) - \psi(0) = t , \qquad t \geqq 0 .$$

By Condition 3, γ is a well-defined, $1 - 1$, continuous mapping of $[0, \infty)$ onto $[0, \infty)$.

LEMMA 2. *Under P_θ, the conditional process*

$$\{Z_2(t) = Z_1(\gamma(t)), \ t \geqq 0\}$$

given the random variable $Z(0)$ is a Brownian motion with drift $(\theta - \frac{1}{2})$ per unit time.

PROOF. Due to the very nature of the transformation γ, it immediately follows from Lemma 1 that the conditional process $\{Z_2(t), t \geqq 0\}$ given $Z(0)$ is a Gaussian process with independent increments. The following computations now complete the proof.

$$E_\theta[Z_2(t) \mid Z(0)] = E_\theta[Z_1(\gamma(t)) \mid Z(0)] = (\theta - \tfrac{1}{2})\{\psi(\gamma(t)) - \psi(0)\} = (\theta - \tfrac{1}{2})t\,, \quad \text{and}$$

$$\operatorname{Var}_\theta [Z_2(t) \mid Z(0)] = \operatorname{Var}_\theta [Z_1(\gamma(t)) \mid Z(0)] = \psi(\gamma(t)) - \psi(0) = t\,.$$

6. Formulas for error probabilities and the OC function. Let us denote by α and β respectively, the probabilities of the type I and type II errors of SPRT(A, B), i.e., $\alpha = P_0(S_1)$ and $\beta = P_1(S_0)$. More generally, we denote the OC function of SPRT(A, B) by $L(\theta) = P_\theta(S_1)$. Formulas connecting α, β, A and B are given by Theorem 4, and Theorem 5 gives a formula for $L(\theta)$.

THEOREM 4. *The type I and type II error probabilities of* SPRT(A, B) *satisfy the relations*

$$\alpha = (1 - \beta)/A + f_0(A)$$

and

$$\beta = B(1 - \alpha) + f_1(B)\,,$$

where $f_0(A) = f_1(B) = 0$ *for* $m_1(0) = m_0(0)$ *and otherwise,*

$$f_0(A) = \{1 - I(A, 0)\} - \{1 - I(A, 1)\}/A\,,$$

$$f_1(B) = I(B, 1) - BI(B, 0)\,,$$

and for $C > 0$,

$$I(C, \theta) = P_\theta[Z(0) \leqq \log C] = \Phi(\{\log C - (\theta - \tfrac{1}{2})\psi(0)\}/(\psi(0))^{\frac{1}{2}})\,,$$

Φ *being the standard normal distribution function.*

THEOREM 5. *For* $\theta \neq \frac{1}{2}$, *the* OC *function of* SPRT(A, B) *is,*

$$L(\theta) = [1 - B^{1-2\theta} + h(A, B, \theta)]/(A^{1-2\theta} - B^{1-2\theta})\,,$$

where $h(A, B, \theta) = 0$ *for* $m_1(0) = m_0(0)$ *and otherwise,*

$$h(A, B, \theta) = A^{1-2\theta}\{1 - I(A, \theta)\} - \{1 - I(A, 1 - \theta)\} + B^{1-2\theta}I(B, \theta) - I(B, 1 - \theta)$$

with $I(C, \theta)$ *as in Theorem* 4,

Let us denote by $\alpha(z)$ and $\beta(z)$ respectively, the conditional probabilities of type I and type II errors and by $L(\theta, z)$ the conditional OC function given $Z(0) = z$, i.e., $\alpha(z) = P_0(S_1 \mid Z(0) = z)$, $\beta(z) = P_1(S_0 \mid Z(0) = z)$ and $L(\theta) = P_\theta(S_1 \mid Z(0) = z)$. These quantities are given by the following lemma.

LEMMA 3. (a) *The conditional etror probabilities* $\alpha(z)$ *and* $\beta(z)$ *satisfy the relations*

$$\begin{aligned} \alpha(z) &= 0\,, && \textit{if} \quad z \leqq \log B \\ &= \{1 - \beta(z)\}/Ae^{-z}\,, && \textit{if} \quad \log B < z < \log A \\ &= 1\,, && \textit{if} \quad z \geqq \log A\,, \end{aligned}$$

$$\begin{aligned} \beta(z) &= 1\,, && \textit{if} \quad z \leqq \log B \\ &= Be^{-z}\{1 - \alpha(z)\}\,, && \textit{if} \quad \log B < z < \log A \\ &= 0\,, && \textit{if} \quad z \geqq \log A\,. \end{aligned}$$

(b) *The conditional* OC *function* $L(\theta, z)$ *is given by*

$$\begin{aligned} L(\theta, z) &= 0\,, && \textit{if} \quad z \leqq \log B \\ &= (e^{(1-2\theta)z} - B^{1-2\theta})/(A^{1-2\theta} - B^{1-2\theta})\,, && \textit{if} \quad \log B < z < \log A \\ &= 1\,, && \textit{if} \quad z \geqq \log A\,. \end{aligned}$$

PROOF. We shall prove the lemma for the case when $m_1(0) \neq m_0(0)$. In this case $\psi(0) > 0$ and $Z(0)$ is a nondegenerate Gaussian random variable. When $m_1(0) = m_0(0)$, $\psi(0) = 0$ and $Z(0) = 0$ with probability 1 under all P_θ and the values of $\alpha(z)$, $\beta(z)$ and $L(\theta, z)$ for $z \neq 0$ are irrelevant, but for $z = 0$, the same proof works whether $m_1(0) \neq m_0(0)$ or $m_1(0) = m_0(0)$.

Since the values of $Z(0) \leqq \log B$ and $\geqq \log A$ lead to immediate acceptance or rejection of H_0, the cases for $z \leqq \log B$ and $\geqq \log A$ hold trivially. We therefore consider only the case when $Z(0)$ takes values $\log B < z < \log A$. In this case we shall prove the lemma by using Lemma 2 and the corresponding results for the Brownian motion in the following way. When $Z(0)$ takes one of these values z, SPRT(A, B) acts on the Z-process exactly as SPRT(Ae^{-z}, Be^{-z}) acts on the Z_2-process described in Lemma 2. The only difference is a change in the time-axis, but that does not affect the boundary-crossing probabilities. Since by Lemma 2, the conditional Z_2-process given $Z(0) = z$ is a Brownian motion under P_θ with drift $(\theta - \frac{1}{2})$ per unit time, $\alpha(z)$, $\beta(z)$, Ae^{-z} and Be^{-z} satisfy the usual relations for a Brownian motion and $L(\theta, z)$ is given in terms of Ae^{-z} and Be^{-z} in the same way as for a Brownian motion. The lemma now follows from the results of Dvoretzky, Kiefer and Wolfowitz (1953).

PROOF OF THEOREM 4. When $m_1(0) = m_0(0)$, the theorem follows trivially from Lemma 3(a) because in that case $\alpha = \alpha(0)$ and $\beta = \beta(0)$. We now consider the case $m_1(0) \neq m_0(0)$. Let ϕ denote the density function and Φ the distribution function of a standard normal random variable. Then the density function of $Z(0)$ under P_θ is given by

$$p_\theta(z) = \{\psi(0)\}^{-\frac{1}{2}}\phi(\{z - (\theta - \tfrac{1}{2})\psi(0)\}/(\psi(0))^{\frac{1}{2}})\,, \tag{4}$$

and

$$\alpha = \int_{-\infty}^{\infty} \alpha(z)p_\theta(z)\,dz\,, \qquad \beta = \int_{-\infty}^{\infty} \beta(z)p_\theta(z)\,dz\,.$$

Using Lemma 3, we thus get

$$\alpha = \int_{\log B}^{\log A} \left\{\frac{e^z - B}{A - B}\right\} p_\theta(z)\,dz + \int_{\log A}^{\infty} p_\theta(z)\,dz \tag{5}$$

and

$$\beta = \int_{-\infty}^{\log B} p_\theta(z)\,dz + \int_{\log B}^{\log A} \left\{\frac{B(Ae^{-z} - 1)}{A - B}\right\} p_\theta(z)\,dz\,. \tag{6}$$

From (5) and (6), the theorem follows.

PROOF OF THEOREM 5. Here also, when $m_1(0) = m_0(0)$, the theorem follows trivially from Lemma 3(b), because in that case $L(\theta) = L(\theta, 0)$. When $m_1(0) \neq$

$m_0(0)$, we use Lemma 3(b) to get

$$L(\theta) = \int_{-\infty}^{\infty} \left\{ \frac{e^{(1-2\theta)z} - B^{1-2\theta}}{A^{1-2\theta} - B^{1-2\theta}} \right\} p_\theta(z)\, dz + \int_{\log A}^{\infty} p_\theta(z)\, dz , \tag{7}$$

where $p_\theta(z)$ is given by (4). From (7), the theorem follows.

REMARKS. (i) Theorem 5 gives $L(\theta)$ for $\theta \neq \frac{1}{2}$. If $L(\theta)$ could be shown to be continuous at $\theta = \frac{1}{2}$, then by taking limits we would get $L(\frac{1}{2}) = \log B/(\log B - \log A)$ when $m_1(0) = m_0(0)$ and otherwise,

$$L(\tfrac{1}{2}) = 1 - \frac{I(A, \frac{1}{2}) \log A - I(B, \frac{1}{2}) \log B + (\psi(0))^{\frac{1}{2}} \{\phi(\log A/(\psi(0))^{\frac{1}{2}}) - \phi(\log B/(\psi(0))^{\frac{1}{2}})\}}{\log B - \log A} .$$

We now show that $L(\theta)$ is continuous at $\theta = \frac{1}{2}$.

Consider the set $S_1 \cap \{\omega \mid T(\omega) \leqq t\}$ where T is given by Definition 3. Since this set is in the σ-field $\mathscr{A}^t$ of the random variables $\{X(s), 0 \leqq s \leqq t\}$, we apply Theorem 1 to get

$$\begin{aligned} P_\theta[S_1 \cap \{T \leqq t\}] &= \int_{S_1 \cap \{T \leqq t\}} dP_\theta^{\,t} = \int_{S_1 \cap \{T \leqq t\}} (dP_\theta^{\,t}/dP_0^{\,t})\, dP_0^{\,t} \\ &= \int_{S_1 \cap \{T \leqq t\}} \exp\left[\theta \langle X - m_0, m_1 - m_0 \rangle_t - \tfrac{1}{2}\theta^2 \psi(t)\right] dP_0^{\,t} \end{aligned}$$

which is a continuous function of θ by Theorem 9, Chapter 2 of Lehmann (1959). We know from our Theorem 3 that $\lim_{t\to\infty} P_\theta[S_1 \cap \{T \leqq t\}] = L(\theta)$ for each θ.

To complete the proof we need this convergence to be uniform in a neighborhood of $\theta = \frac{1}{2}$. To show this, let $\{t_n\}$ be the sequence defined in the proof of Theorem 3, and let $t_0 = 0$. Choose $\varepsilon, \eta > 0$. Then for all $\theta \in (\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)$, $P_\theta[Z(t_{j+1}) - Z(t_j) > \eta] \geqq 1 - \Phi(\eta\delta^{-\frac{1}{2}} + \varepsilon\delta^{\frac{1}{2}}) = p > 0$. Let k be such that $k\eta \geqq \log A - \log B$. If we now look at the random walk $\{Z(t_j), j = 0, 1, 2, \cdots\}$, then it can be easily seen that

$$P_\theta[\log B < Z(t_{jk}) < \log A \mid \log B < Z(t_{(j-1)k}) < \log A] \leqq 1 - p^k < 1 , \qquad j = 1, \cdots, n .$$

Hence

$$\begin{aligned} L(\theta) - P_\theta[S_1 \cap \{T \leqq t_{nk}\}] &= P_\theta[S_1 \cap \{T > t_{nk}\}] \leqq P_\theta[T > t_{nk}] \\ &\leqq P_\theta[\log B < Z(t_{jk}) < \log A, j = 0, 1, \cdots, n] \leqq (1 - p^k)^n , \end{aligned}$$

and that concludes the proof.

(ii) It should be noted that our definition of the OC function is more restrictive than Wald's original definition according to which the domain of L should be all real-valued functions m on $[0, \infty)$. We have defined L only for $m \in \{(1 - \theta)m_0 + \theta m_1, \theta \text{ real}\}$.

7. ASN function. A random stopping time T of the LLR process can be treated exactly in the same way as it was treated in the case of stationary and independent increments by Dvoretzky, Kiefer and Wolfowitz (1953). It follows from Theorem 2 that the process $\{W(t) = Z(t) - (\theta - \frac{1}{2})\psi(t), t \geqq 0\}$ is of independent increments with mean 0, hence it is a martingale in continuous time. We assume $E_\theta(T) < \infty$ and note that by Condition 3 and Theorem 2(b) almost all sample paths of the process $\{W(t), t \geqq 0\}$ are continuous. Hence $E_\theta[W(T)] = 0$ by Theorem 11.8, Chapter VII of Doob (1953). We thus have the following theorem.

THEOREM 6. *Let T be a random stopping time of the* LLR *process with* $E_\theta(T) < \infty$. *Then*

$$E_\theta[Z(T)] = (\theta - \tfrac{1}{2})E_\theta[\psi(T)] .$$

In the special case when ψ is linear in t, Theorem 6 can be used in conjunction with Theorem 5 to obtain a formula for the ASN function of SPRT(A, B) whose value at θ is the expected value under P_θ of the stopping time T given by Definition 3. We have the following corollary to Theorem 6.

COROLLARY. *If* $\psi(t) = a + bt$ *and* T *is given by Definition* 3, *then for* $\theta \neq \frac{1}{2}$,

$$E_\theta(T) = \frac{E_\theta[Z(T)]}{b(\theta - \frac{1}{2})} - a ,$$

and

$$\begin{aligned} E_\theta[Z(T)] = L(\theta) \log A &+ \{1 - L(\theta)\} \log B + (\theta - \tfrac{1}{2})\psi(0) \\ &- (\theta - \tfrac{1}{2})\psi(0)\{I(A, \theta) - I(B, \theta)\} \\ &- \{1 - I(A, \theta)\} \log A - I(B, \theta) \log B \\ &+ (\psi(0))^{\frac{1}{2}}\{J(A, \theta) - J(B, \theta)\} , \end{aligned}$$

where $I(C, \theta)$ *is as in Theorem* 5, *and*

$$J(C, \theta) = \phi(\{\log C - (\theta - \tfrac{1}{2})\psi(0)\}/(\psi(0))^{\frac{1}{2}}) .$$

8. Examples. We consider two examples of stationary Gaussian processes. The mean value functions under the null and alternative hypotheses are constants in one example and polynomials in the other. To apply our results to these situations, we need to compute $||1||^2_{H(R^\tau)}$ in the first example and $(t^j, t^k)_{H(R^\tau)}$ for $j, k = 0, 1, 2, \cdots$ in the second example. For computing these inner products in $H(R^\tau)$ we shall use the well-known fact that

$$(f, g)_{H(R^\tau)} = (E^*[f \mid R_t, t \leqq \tau], E^*[g \mid R_t, t \leqq \tau])_{H(R)} .$$

Consequently, for $f \in L_2(R_t \mid t \leqq \tau)$ and $s \leqq \tau$,

$$(f, R_s)_{H(R^\tau)} = (f, R_s)_{H(R)} = f(s)$$

by reproducing kernel property of $H(R)$. However, it may be quite difficult to compute $(f, R_s)_{H(R^\tau)}$ in general.

(I) Consider the covariance kernel

$$\begin{aligned} R(s, t) &= \alpha - \beta|s - t| \quad \text{for} \quad |s - t| < \alpha/\beta \\ &= 0 \quad \text{otherwise} \end{aligned}$$

where α and β are positive constants and Condition 1 holds. We want to test the null hypothesis that the mean value function $m(t) \equiv \theta_0$ against the alternative $m(t) \equiv \theta_1$ where θ_0 and θ_1 are given constants. Let $\tau \geqq 0$ be given and find the positive integer n such that $(n-1)\alpha/\beta \leqq \tau < n\alpha/\beta$. We shall show that on $[0, \tau]$, the function 1 is a linear combination of functions $\{R_t, 0 \leqq t \leqq \tau\}$ and compute $||1||^2_{H(R^\tau)}$. Consider the function

$$f(t) = \sum_{j=0}^{n-1} (n-j)\left[R\left(\frac{j\alpha}{\beta}, t\right) + R\left(\tau - \frac{j\alpha}{\beta}, t\right)\right].$$

It is easy to verify that for $r\alpha/\beta \leqq t \leqq \tau - (n-r-1)\alpha/\beta$, $r = 0, 1, \cdots, n-1$,

$$f(t) = (n-r)R\left(\frac{r\alpha}{\beta}, t\right) + (n-r-1)R\left(\frac{(r+1)\alpha}{\beta}, t\right) + (r+1)R\left(\tau - \frac{(n-r-1)\alpha}{\beta}, t\right) + rR\left(\tau - \frac{(n-r)\alpha}{\beta}, t\right),$$

and for $\tau - r\alpha/\beta \leqq t \leqq (n-r)\alpha/\beta$, $r = 1, \cdots, n-1$,

$$f(t) = (r+1)R\left(\frac{(n-r-1)\alpha}{\beta}, t\right) + rR\left(\frac{(n-r)\alpha}{\beta}, t\right) + (n-r)R\left(\tau - \frac{r\alpha}{\beta}, t\right) + (n-r+1)R\left(\tau - \frac{(r-1)\alpha}{\beta}, t\right).$$

Examining these two cases separately, it can now be seen that

$$f(t) = 2n\alpha - \beta\tau, \qquad 0 \leqq t \leqq \tau.$$

Hence

$$\begin{aligned} 1^\tau &= f^\tau/(2n\alpha - \beta\tau) \\ &= \sum_{j=0}^{n-1} (n-j)[R^\tau_{j\alpha/\beta} + R^\tau_{\tau - j\alpha/\beta}]/(2n\alpha - \beta\tau), \end{aligned}$$

and Condition 2 holds. Now in view of the remark made earlier in this section,

$$\begin{aligned} (\theta_1 - \theta_0)^{-2}\psi(\tau) &= ||1^\tau||^2_{H(R^\tau)} \\ &= (2n\alpha - \beta\tau)^{-2}[2 \textstyle\sum_{j=0}^{n-1} (n-j)^2 R(0, 0) \\ &\qquad + 2 \textstyle\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (n-i)(n-j) R(j\alpha/\beta, \tau - j\alpha/\beta)] \\ &= n(n+1)/(2n\alpha - \beta\tau) \end{aligned}$$

by straightforward computation. ψ is continuous and it is easy to verify that it satisfies all other parts of Condition 3.

(II) Ornstein-Uhlenbeck process: Here the covariance kernel is

$$R(s, t) = \alpha \exp[-\beta|s-t|]$$

where α and β are positive constants and Condition 1 holds. We want to test the null hypothesis $m(t) = \sum_{k=0}^{p} \theta_k t^k$ against the alternative $m(t) = \sum_{k=0}^{p} \theta_k' t^k$ where $\theta_0, \theta_1, \cdots, \theta_p, \theta_0', \theta_1', \cdots, \theta_p'$ are given constants. Now for all $t \in [0, \tau]$,

$$\begin{aligned} (\beta^{k+1}/k!) \textstyle\int_0^\tau s^k R(s, t)\, ds &= (-1)^{k+1} R(0, t) - \{\textstyle\sum_{r=0}^{k} (\beta\tau)^r/r!\} R(\tau, t) \\ &\quad + \alpha\{\textstyle\sum_{r=0}^{k} (\beta t)^r/r! + \sum_{r=0}^{k} (-1)^{k-r} (\beta t)^r r!\}. \end{aligned}$$

Hence

$$[(\beta^{k+1}/k!) \int_0^\tau s^k R_s\, ds + (-1)^k R_0 + \{\sum_{r=0}^k (\beta\tau)^r/r!\} R_\tau]/2\alpha$$
$$= \sum_{r=0}^{[k/2]} (\beta t)^{k-2r}/(k-2r)!\,,$$

where $[k/2]$ is the largest integer $\leqq k/2$. From this, we have

$$1 = [\beta \int_0^\tau R_s\, ds + R_0 + R_\tau]/2\alpha\,,$$
$$t = [\beta^2 \int_0^\tau sR_s\, ds - R_0 + (1+\beta\tau)R_\tau]/2\alpha\beta\,,$$

and

$$t^k = k[\{k(k+1)\}^{-1} \int_0^\tau \{\beta^2 s^k - k(k+1)s^{k-2}\} R_s\, ds + k^{-1}(\beta\tau^k - \tau^{k-1})R_\tau]/2\alpha\beta$$

for $k = 3, 4, \cdots$. Thus the restrictions of 1 and all positive integral powers of t are in $H(R^\tau)$ and Condition 2 holds. Furthermore, in view of our earlier remark, $(t^j, t^k)_{H(R^\tau)}$ is easily seen to be a polynomial of degree $\max(j, k) + 1$ in τ and therefore, ψ satisfies Condition 3. In particular, if $m_0(t) \equiv \theta_0$ and $m_1(t) \equiv \theta_0'$, then

$$(\theta_0' - \theta_0)^{-2}\psi(\tau) = \|1\|^2_{H(R^\tau)} = (1, R_0 + R_\tau + \beta \int_0^\tau R_s\, ds)_{H(R^\tau)}/2\alpha$$
$$= [(1, R_0)_{H(R^\tau)} + (1, R_\tau)_{H(R^\tau)} + \beta \int_0^\tau (1, R_s)_{H(R^\tau)}\, ds]/2\alpha$$
$$= [1 + 1 + \beta \int_0^\tau 1 \cdot ds]/2\alpha = (2 + \beta\tau)/2\alpha\,.$$

In this special case, ψ is linear and the corollary to Theorem 6 becomes applicable.

Smith (1971) has considered various other examples where our results are applicable. He has also shown that for a certain class of stationary covariance kernels, $1^\tau \in H(R^\tau)$, $\tau \geqq 0$ giving rise to a linear ψ-function.

Acknowledgments. We gratefully acknowledge the suggestions from a referee which were very helpful in revising the paper. Among other things, his comments helped us simplify the proof of Theorem 2. We also thank William D. Sudderth for some comments.

REFERENCES

[1] ARONSZAJN, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **69** 337–404.

[2] BHATTACHARYA, P. K. (1971). Transformations of Gaussian processes to the Brownian motion. *Ann. Math. Statist.* **42** 2008–2017.

[3] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[4] DVORETZKY, A., KIEFER, J., and WOLFOWITZ, J. (1953). Sequential decision problems for processes with continuous time parameter. Testing hypotheses. *Ann. Math. Statist.* **24** 254–264.

[5] FELDMAN, J. (1958). Equivalence and perpendicularity of Gaussian processes. *Pacific J. Math.* **8** 699–708.

[6] GIKHMAN, I. I. and SKOROKHOD, A. V. (1965). *Introduction to the Theory of Random Processes*. Nauka Press, Moskow. (Eng. trans. (1969). Saunders, Philadelphia.)

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

[8] PARZEN, EMANUEL (1959). Statistical analysis on time series by Hilbert space methods, I. Technical Report No. 23, Applied Mathematics and Statistical Laboratory, Stanford University. (Reprinted, Parzen (1967) *Time Series Analysis Papers*, Holden Day, San Francisco.)

[9] SMITH, ROBERT P. (1971). Contributions in sequential analysis. Ph. D. dissertation, University of Arizona.

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

[11] WALD, A., and WOLFOWITZ, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Statist.* **19** 326-329.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721