The Annals of Statistics
1984, Vol. 12, No. 1, 380–386

EXTENDED OPTIMALITY OF SEQUENTIAL PROBABILITY RATIO TESTS

BY ALBRECHT IRLE

Westfälische Wilhelms-Universität, Münster

The problem of sequentially testing two simple hypotheses for a stochastic process is considered. It is shown that, for arbitrary distributions P_0 and P_1, the following optimality holds for an SPRT which stops on its boundaries: If α and β represent the error probabilities of the SPRT and a competing test has error probabilities $\alpha' \le \alpha$ and $\beta' \le \beta$ then $E_0 g(D_{\tau'}) \ge E_0 g(D_\tau)$ for any convex function g satisfying some minor requirement, provided $P_1(\tau' < \infty) = 1$ for the competing test. Here D_τ and $D_{\tau'}$ denote the terminal likelihood ratios under the SPRT and the competitor. An analogous statement holds for expectation under P_1, and several applications of this optimality result are given.

1. Introduction and preliminaries. We consider the problem of sequentially testing two simple hypotheses for a stochastic process. In the i.i.d. case, Wald and Wolfowitz (1948) established the optimality property of SPRT's, see e.g. also Lorden (1981), and this optimality carries over to continuous time stochastic processes having a log-likelihood ratio process of stationary and independent increments as noted by Dvoretzky, Kiefer and Wolfowitz (1953), see e.g. also Liptser and Shiryayev (1978), Irle and Schmitz (1981).

In this paper we treat a general model in the sense of Eisenberg, Ghosh and Simons (1976), i.e. there are no assumptions on the distributions under P_0 and P_1 of the observed stochastic process, and our basic result extends the optimality of SPRT's in the following way:

If α and β represent the error probabilities of an SPRT *which stops on its boundaries*, and a competing test has error probabilities $\alpha' \le \alpha$ and $\beta' \le \beta$, then

$$E_0 g(D_{\tau'}) \ge E_0 g(D_\tau)$$

for any convex function g satisfying some minor requirement, provided $P_1(\tau' < \infty) = 1$ for the competing test. Here D_τ and $D_{\tau'}$ denote the terminal likelihood ratios under the SPRT and the competitor; an analogous statement, of course, holds for expectation under P_1. This result and several applications are given in Section 2 of the paper.

The general model of sequential testing with which we are concerned takes the following form:

Let $(\Omega, \mathscr{A})$ be a measurable space, $T = \mathbb{N}$ or $T = [0, \infty)$ the time set and $(\mathscr{A}_t)_{t \in T}$ an increasing sequence of sub-σ-algebras of $\mathscr{A}$. Let P_0 and P_1 be probability measures on $(\Omega, \mathscr{A})$.

Received October 1982; revised June 1983.
AMS 1980 *subject classifications*. Primary 62L10; secondary 62L15.
Key words and phrases. Sequential probability ratio test, optimality, admissibility.

A sequential test for P_0 against P_1 is given by a stopping rule

$$\tau:\Omega \longrightarrow T \cup \{\infty\}$$

and a terminal decision function

$$\varphi:\{\tau < \infty\} \longrightarrow [0, 1],$$

which is $\mathscr{A}_\tau \cap \{\tau < \infty\}$-measurable.

Let $\mathscr{C}$ be a sub-σ-algebra of $\mathscr{A}$. Then a $\mathscr{C}$-measurable random variable U with values in $[0, \infty]$ is called a density of $P_1 \mid \mathscr{C}$ with respect to $P_0 \mid \mathscr{C}$ iff

$$P_1(C) = \int_C U\, dP_0 + P_1(C \cap \{U = \infty\}) \quad \text{for all} \quad C \in \mathscr{C},$$

U is denoted by $dP_1 \mid \mathscr{C}/dP_0 \mid \mathscr{C}$.

Let $D = (D_t)_{t\in T}$ denote the stochastic process with

$$D_t = dP_1 \mid \mathscr{A}_t/dP_0 \mid \mathscr{A}_t.$$

To obtain a stochastic process with well-behaved paths in the continuous time case, we make the usual assumption that the filtration $(\mathscr{A}_t)_{t\in T}$ is right-continuous and complete with respect to $(P_0 + P_1)/2$, and according to Jacod (1979), (7.2), we may then assume that D has right-continuous paths with left limits. Furthermore $D_\infty = \lim_{t\to\infty} D_t$ exists, $P_0(\sup_t D_t = \infty) = 0$ and

$$D_\tau = dP_1 \mid \mathscr{A}_\tau/dP_0 \mid \mathscr{A}_\tau \quad \text{for all stopping rules} \quad \tau.$$

Interchanging P_0 and P_1, we also consider the stochastic process $\bar{D} = (\bar{D}_t)_{t\in T}$, $\bar{D}_t = dP_0 \mid \mathscr{A}_t/dP_1 \mid \mathscr{A}_t$ with properties as above, and we may choose $\bar{D} = 1/D$.

We will now briefly discuss the Doob-Meyer decomposition of submartingales, as it will be used later; for this discussion we assume

(A1) $P_0 \mid \mathscr{A}_t$ and $P_1 \mid \mathscr{A}_t$ are mutually absolutely continuous for all $t < \infty$.

Let $g:[0, \infty) \longrightarrow \mathbb{R}$ be convex. If $g(D_t)$ is integrable for all $t \in T$, then $g(D)$ is a P_0-submartingale and we may now consider the Doob-Meyer decomposition

$$g(D) = M^g + A^g$$

with respect to P_0, where M^g is a local P_0-martingale with $E_0 M_0^g = 0$ and A^g is a process with increasing paths; see e.g. Jacod (1979), (1.53), (2.18). Likewise, we have the Doob-Meyer decomposition $g(\bar{D}) = \bar{M}^g + \bar{A}^g$ w.r.t. P_1. If, in the case of continuous time, D has uniformly bounded jumps P_0-a.s., then we obtain the Doob-Meyer decomposition for any convex g (without the above requirement of integrability), as $g(D)$ is then a locally bounded P_0-submartingale, i.e. there exists a sequence $(\tau_n)_n$ of stopping times such that $\tau_n \longrightarrow \infty$ P_0-a.s. and $g(D)^{\tau_n} = (g(D)_{\tau_n\wedge t})_{t\in T}$ is uniformly bounded P_0-a.s. (take $\tau_n = \inf\{t: D_t \geq n\}$).

Let us denote by $\mathscr{G}$ the set of all convex g with $g(x)/x \longrightarrow \infty$ for $x \longrightarrow \infty$, such that the Doob-Meyer decomposition of $g(D)$ exists. $\bar{\mathscr{G}}$ is defined likewise.

(1.1) LEMMA. *For any $g \in \mathscr{G}$ and any stopping rule τ with $E_0 A_\tau^g < \infty$ we have*

$$E_0 g(D)_\tau = E_0 A_\tau^g \quad \text{and} \quad E_0 D_\tau = 1.$$

PROOF. Let $(\tau_n)_n$ be a P_0-localizing sequence of stopping rules for D and M^g, i.e. $(\tau_n)_n$ is an increasing sequence tending to infinity P_0-a.s. such that D^{τ_n} and $(M^g)^{\tau_n}$ are uniformly integrable martingales. Then

$$E_0 g(D)_{\tau \wedge \tau_n} = E_0 A_{\tau \wedge \tau_n}^g, \quad E_0 D_{\tau \wedge \tau_n} = 1$$

holds for all n.

The result follows from this by an argument as in the proof of Wald's second moment identity by Chow, Robbins and Siegmund (1979), Theorem 2.3. We just remark that the condition $g(x)/x \to \infty (x \to \infty)$ ensures the uniform integrability of $(D_{\tau \wedge \tau_n})_n$. □

2. Optimality of sequential probability ratio tests.

We will now derive the extended optimality of SPRT's in the general situation as described in Section 1. For constants a, b, $0 < a < b$ let us denote by $\mathscr{G}(a, b)$ the set of all convex functions $g: [0, \infty) \to \mathbb{R}$ such that $g'(a) < g'(b)$ for some tangent slopes in a and b.

(2.1) THEOREM. *Let (τ, φ) be an* SPRT *with stopping bounds $a < b$, $0 < a \le 1 \le b$, and error probabilities α, β, such that $D_\tau \in \{a, b\}$ P_0-a.s. and P_1-a.s. Then for any sequential test (τ', φ') with error probabilities $\alpha' \le \alpha$, $\beta' \le \beta$ the following holds*:

$P_1(\tau' < \infty) = 1$ *implies*

$$E_0 g(D_{\tau'}) \ge E_0 g(D_\tau) \quad \textit{for all} \quad g \in \mathscr{G}(a, b),$$

$P_0(\tau' < \infty) = 1$ *implies*

$$E_1 g(D_{\tau'}) \ge E_1 g(D_\tau) \quad \textit{for all} \quad g \in \mathscr{G}(a, b),$$

and all the inequalities are strict if $(\alpha', \beta') \ne (\alpha, \beta)$.

PROOF. Let $g \in \mathscr{G}(a, b)$. We only consider the first inequality, since the second is treated in a similar manner. It is obviously sufficient to show the existence of u, $v > 0$ such that

$$E_0 g(D_\tau) + u\alpha + v\beta \le E_0 g(D_{\tau'}) + u\alpha' + v\beta'.$$

Using convexity and the assumption $g'(a) < g'(b)$ we easily obtain the existence of constants $u, v, c, d \in \mathbb{R}$, $u, v > 0$, such that

$$g(x) + \min\{u, vx\} \ge cx + d \quad \text{for all} \quad x \in [0, \infty),$$

$$\text{with equality for} \quad x = a, \quad x = b,$$

which implies $a < u/v < b$. By an argument used e.g. by Lorden (1981), we have

$$u\alpha' + v\beta' \ge E_0 \min\{u, vD_{\tau'}\} \quad \text{for any sequential test} \quad (\tau', \varphi'),$$

with equality for the SPRT (τ, φ). Now $P_1(\tau' < \infty) = 1$ is equivalent to $E_0 D_{\tau'} = 1$, and we thus obtain

$$\begin{aligned} E_0 g(D_{\tau'}) + u\alpha' + v\beta' &\geq E_0(g(D_{\tau'}) + \min\{u, vD_{\tau'}\}) \\ &\geq c + d = E_0(g(D_\tau) + \min\{u, vD_\tau\}) \\ &= E_0 g(D_\tau) + u\alpha + v\beta. \quad \square \end{aligned}$$

The following examples show that the assumptions on the convex function g, resp. on the competing test (τ', φ'), may not be omitted in general.

(i) Consider g which is linear on some interval $[a_1, b_1]$, such that $0 < a_1 < a < b < a_2$. Let (τ_1, φ_1) denote the SPRT with stopping bounds $a_1 < b_1$ and error probabilities α_1, β_1. Under the assumption $D_{\tau_1} \in \{a_1, b_1\}$ P_0-a.s. and P_1-a.s. we then obtain $E_0 g(D_{\tau_1}) = E_0 g(D_\tau)$, but $\alpha_1 < \alpha$ and $\beta_1 < \beta$.

(ii) Let $\varepsilon > 0$ and define (τ', φ') by $\tau' = \inf\{t \in T: D_t \leq \varepsilon\}$; $\varphi' = 0$, thus $\alpha' = 0$, $\beta' = P_1(\tau' < \infty)$. Consider g such that $g(x) > 0$ and $g(x) \to g(0) = 0$ for $x \to 0$. It follows $\beta' \to 0$ and $E_0 g(D_{\tau'}) \to 0$ for $\varepsilon \to 0$, and thus for suitably chosen $\varepsilon > 0$ we have $\alpha' < \alpha$, $\beta' < \beta$ and $E_0 g(D_{\tau'}) < E_0 g(D_\tau)$ for any SPRT (τ, φ) as in (2.1), provided that $P_0(D_\infty = 0) = 1$.

To obtain a result which holds for any competing test we now use the Doob-Meyer decomposition, and we assume validity of (A1) from now on.

(2.2) COROLLARY. *Let (τ, φ) be an* SPRT *as in* (2.1). *Then for any sequential test (τ', φ') with error probabilities $\alpha' \leq \alpha$, $\beta' \leq \beta$*

$$E_0 A^g_{\tau'} \geq E_0 A^g_\tau \quad \textit{for all} \quad g \in \mathscr{G}(a, b) \cap \mathscr{G},$$

$$E_1 \bar{A}^g_{\tau'} \geq E_1 \bar{A}^g_\tau \quad \textit{for all} \quad g \in \mathscr{G}(a, b) \cap \bar{\mathscr{G}},$$

and all the inequalities are strict if (α', β').

PROOF. Again, consider only the first inequality. If $E_0 A^g_{\tau'} = \infty$ then this is trivially true. For $E_0 A^g_\tau < \infty$ we obtain from (1.1) $E_0 A^g_{\tau'} = E_0 g(D_{\tau'})$ and $E_0 D_{\tau'} = 1$. Now obviously $E_0 A^g_\tau = E_0 g(D_\tau)$, and the result follows from (2.1). $\square$

Since the processes A^g, $\bar{A}^g$ are increasing the following admissibility result holds for an SPRT (τ, φ) as in (2.1):

There does not exist a sequential test (τ', φ') such that $\alpha' \leq \alpha$, $\beta' \leq \beta$ with $(\alpha', \beta') \neq (\alpha, \beta)$ and $\tau' \leq \tau$ P_0-a.s. *or* $\tau' \leq \tau$ P_1-a.s. Furthermore if A^g, $\bar{A}^{g'}$ are nonrandom for some g, g', then the condition "$\tau' \leq \tau$ P_0-a.s. *or* $\tau' \leq \tau$ P_1-a.s." may be replaced by

"$P_0(\{\tau' \leq t\}) \geq P_0(\{\tau \leq t\})$ f.a. t *or* $P_1(\{\tau' \leq t\}) \geq P_1(\{\tau \leq t\})$ f.a. t."

Let us remark that the admissibility result of Eisenberg, Ghosh and Simons (1976) in this situation only yields the nonexistence of a sequential test (τ', φ') with the properties $\alpha' \leq \alpha$, $\beta' \leq \beta$, $(\alpha', \beta') \neq (\alpha, \beta)$ and $\tau' \leq \tau$ P_0-a.s. *and* $\tau' \leq \tau$ P_1-a.s., but of course their result applies to more general tests (τ, φ).

For a further discussion we now consider the case of continuous time, and additionally to (A1) assume

(A2) $P_0 \mid \mathscr{A}_0 = P_1 \mid \mathscr{A}_0$, thus $D_0 = 1$ P_0-a.s. and P_1-a.s., and $P_0 \mid \mathscr{A}_\infty$ and $P_1 \mid \mathscr{A}_\infty$ are orthogonal, thus $P_0(D_\infty = 0) = 1$ and $P_1(\bar{D}_\infty = 0) = 1$.

(A3) D has continuous paths P_0-a.s. and P_1-a.s., so that the same is true for $\bar{D} = 1/D$.

In this situation we can apply (2.1), (2.2) to any SPRT and, as discussed in Section 1, for any convex g we obtain the Doob-Meyer decomposition for $g(D)$ and $g(\bar{D})$. Moreover, D may be represented in the following way: There exists a local P_0-martingale with continuous paths and $Y_0 = 0$ such that

$$D = \exp(Y - \tfrac{1}{2}\langle Y, Y\rangle),$$

see e.g. Jacod (1979), VI.1, where $\langle Y, Y\rangle$ denotes the previsible increasing process associated with Y under P_0. From the orthogonality of $P_0 \mid \mathscr{A}_\infty$ and $P_1 \mid \mathscr{A}_\infty$ one may infer $\langle Y, Y\rangle_\infty = \infty$ P_0-a.s. and P_1-a.s., see Jacod (1979), VIII.1, and Kabonov, Liptser and Shiryayev (1979), and furthermore an easy application of Ito's formula yields

$$E_0\langle Y, Y\rangle_\tau < \infty \quad \text{and} \quad E_1\langle Y, Y\rangle_\tau < \infty$$

for any SPRT.

We will now show that a special choice of g leads to the increasing process $\langle Y, Y\rangle$.

(2.3) Corollary. *Let (τ, φ) be an* SPRT *as in* (2.1). *Then for any sequential test (τ', φ') with $\alpha' \leq \alpha$, $\beta' \leq \beta$ we have*

$$E_0\langle Y, Y\rangle_{\tau'} \geq E_0\langle Y, Y\rangle_\tau, \quad E_1\langle Y, Y\rangle_{\tau'} \geq E_1\langle Y, Y\rangle_\tau,$$

and both inequalities are strict if $(\alpha, \beta) \neq (\alpha', \beta')$.

Proof. Consider $g(x) = 2x\log x$ for $x > 0$, $g(0) = 0$. By (2.2) it is sufficient to show that

$$E_0A_\sigma^g = E_1\langle Y, Y\rangle_\sigma \quad \text{and} \quad E_1A_\sigma^g = E_0\langle Y, Y\rangle_\sigma$$

for any stopping rule σ.

Define an increasing sequence of stopping rules $(\tau_n)_n$ by

$$\tau_n = \inf\{t : \langle Y, Y\rangle_t = n \text{ or } D_t \notin (1/n, n)\},$$

then $\tau_n < \infty$ and $\tau_n \to \infty$ $(n \to \infty)$ P_0-a.s. and P_1-a.s. Now for any stopping rule σ by (1.1)

$$\begin{aligned} E_0A^g_{\sigma\wedge\tau_n} &= E_0g(D_{\sigma\wedge\tau_n}) = E_1(2Y_{\sigma\wedge\tau_n} - \langle Y, Y\rangle_{\sigma\wedge\tau_n}) \\ &= 2E_1(Y_{\sigma\wedge\tau_n} - \langle Y, Y\rangle_{\sigma\wedge\tau_n}) + E_1\langle Y, Y\rangle_{\sigma\wedge\tau_n}. \end{aligned}$$

As a consequence of Girsanov's theorem we have $E_1(Y_{\sigma\wedge\tau_n} - \langle Y, Y\rangle_{\sigma\wedge\tau_n}) = 0$, see e.g. Elliott (1982), 13.14, 13.19, which yields $E_0A^g_{\sigma\wedge\tau_n} = E_1\langle Y, Y\rangle_{\sigma\wedge\tau_n}$, and by monotone convergence ($n \longrightarrow \infty$)

$$E_0 A^g_\sigma = E_1 \langle Y, Y \rangle_\sigma.$$

For the second relation we use $\bar{D} = 1/D = \exp(-Y + \frac{1}{2}\langle Y, Y \rangle)$ and obtain

$$E_1 A^g_{\sigma \wedge \tau_n} = E_1 g(D)_{\sigma \wedge \tau_n} = E_0(\langle Y, Y \rangle_{\sigma \wedge \tau_n} - 2Y_{\sigma \wedge \tau_n}) = E_0 \langle Y, Y \rangle_{\sigma \wedge \tau_n},$$

from which the assertion follows. □

Now (2.3) yields the well-known result on the optimality of the SPRT for sequentially testing about the drift parameter of a Wiener-process since in this case we have $\langle Y, Y \rangle_t = \gamma t$ for some constant $\gamma > 0$, and it also includes as a special case the result of Liptser and Shiryayev (1978), Theorem 17.8, on the sequential testing of two simple hypotheses for certain Ito processes.

Note that in the i.i.d. situation, this choice of $g(x) = 2x \log x$ yields optimality in the sense of Wald and Wolfowitz for an SPRT which stops on its boundaries.

For a different application, consider the situation where we have a family $(P_\theta)_{\theta \in \mathbb{R}}$ of probability measures with densities

$$dP_\theta \mid \mathscr{A}_t / dP_0 \mid \mathscr{A}_t = \exp(\theta Y_t - (\theta^2/2)\langle Y, Y \rangle_t),$$

for a local P_0-martingale Y_0 with continuous paths, $Y_0 = 0$. For $\theta_0 < \theta_1$ let (τ, φ) be an SPRT for testing $P_0 = P_{\theta_0}$ and $P_1 = P_{\theta_1}$.

Considering $g(x) = x^\gamma$, $\gamma > 1$, we easily obtain the following:

For each $\theta < \theta_0$ or $\theta > \theta_1$ there exists $r(\theta) > 0$ such that

$$E_\theta \exp(r(\theta)\langle Y, Y \rangle_{\tau'}) \geq E_\theta \exp(r(\theta)\langle Y, Y \rangle_\tau)$$

for any (τ', φ') with $\alpha' \leq \alpha$, $\beta' \leq \beta$ and strict inequality holds if $(\alpha, \beta) \neq (\alpha', \beta')$. Although exponential cost structure is not of particular interest by itself, this yields the following admissibility for an SPRT (τ, φ):

There does not exist a sequential test (τ', φ') such that $\alpha' \leq \alpha$, $\beta' \leq \beta$, $(\alpha', \beta) \neq (\alpha, \beta)$ and $\tau' \leq \tau$ P_θ-a.s. for at least one $\theta \leq \theta_0$ or $\theta \geq \theta_1$. "$\tau' \leq \tau P_\theta$-a.s." may be replaced by "$P_\theta(\tau' \geq t) \leq P_\theta(\tau \geq t)$ f.a. t" if $\langle Y, Y \rangle$ is nonrandom, as it is e.g. the case for Gaussian processes with common covariance kernel and mean value function $\theta m(t)$.

Acknowledgment. The author wishes to thank the referee for his helpful comments and suggestions.

REFERENCES

CHOW, Y. S., ROBBINS, H. and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton-Mifflin, Boston.

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.

EISENBERG, B., GHOSH, B. K. and SIMONS, G. (1976). Properties of generalized sequential probability ratio tests. *Ann. Statist.* **4** 237–251.

ELLIOTT, R. J. (1982). *Stochastic Calculus and Applications.* Springer, Berlin.

IRLE, A. and SCHMITZ, N. (1981). On the optimality of the SPRT for processes with continuous time parameter. *Math. Operationsforsch. Statist.* to appear.

JACOD, J. (1979). Calcul Stochastique et Problèmes de Martingale. *Lecture Notes in Mathematics* **714.** Springer, Berlin.

KABANOV, J. M., LIPTSER, R. S. and SHIRYAYEV, A. N. (1979). Absolute continuity and singularity of locally absolutely continuous probability distributions. *Math. USSR Sbornik* **35** 631–680.

LIPTSER, R. S. and SHIRYAYEV, A. N. (1978). *Statistics of Random Processes II.* Springer, Berlin.

LORDEN, G. (1980). Structure of sequential tests minimizing an expected sample size. *Z. Wahrsch. verw. Gebiete* **51** 291–302.

INSTITUT FÜR MATHEMATISCHE STATISTIK
UNIVERSITÄT MÜNSTER
EINSTEINSTRAßE 62
D-4400 MÜNSTER
WEST GERMANY