The Annals of Statistics
1975, Vol. 3, No. 2, 373–381

LOCALLY MOST POWERFUL SEQUENTIAL TESTS[1]

BY ROBERT H. BERK

Rutgers University

Sequential tests that are LMP for certain one-sided testing problems are discussed. In all cases considered, the stopping rule is the first time a certain random walk leaves a bounded interval. (Thus various inequalities and approximations due to Wald can be utilized in obtaining properties of these tests.) For models in a one-parameter exponential family, each LMP sequential test is shown to be a Wald SPRT for a family of paired (conjugate) simple hypotheses.

1. Introduction. Consider the following one-sided testing problem. We have a model, a family of pdf's $f(x \mid \theta)$ (with respect to a σ-finite measure λ on the sample space $(\mathscr{X}, \mathscr{A})$). The parameter θ ranges in Θ, a subinterval of the real line. $X, X_1, \cdots$ is a (i.i.d.) data sequence drawn from the model. Based on $X_1, X_2, \cdots$, we wish to test sequentially the hypothesis $H_1: \theta = \theta^*$ against the one-sided alternative $H_2: \theta > \theta^*$. It is well known that there is no one "best" sequential test for this problem, in general. Even among non-sequential tests, UMP tests do not necessarily exist. However, under certain regularity conditions, there are LMP non-sequential tests which (for a sample of size n) are of the form: reject H_1 if

$$S_n = \textstyle\sum_1^n r(X_j) \geqq a\,, \tag{1.1}$$

where $r(x) = \partial \log f(x \mid \theta)/\partial\theta|_{\theta=\theta^*}$.

We show that a similar phenomenon persists in the sequential case. Under regularity conditions, we delineate a certain class of sequential tests that are LMP for testing H_1 vs. H_2. That is, given a test in the class, let α denote its (exact) level and $\nu < \infty$ its expected stopping time under θ^*. Then among all level α sequential tests whose expected stopping times under θ^* do not exceed ν, the given test is LMP. These LMP tests are a sequential analog of (1.1); their stopping times are of the form

$$\sigma = \inf\{n : S_n \notin (-a_1, a_2)\}\,. \tag{1.2}$$

Typically, a_1 and a_2 are both positive. The terminal decision is, as one would expect: reject H_1 if $S_\sigma \geqq a_2$. For short, we shall refer to the sequential test as σ and denote by $\mathscr{S}$, the class of such tests. The boundary points a_1 and a_2 are determined (in principle) by specifying α and ν. That is,

$$\text{(1.3)}\qquad \text{(i)}\quad P(S_\sigma \geqq a_2) = \alpha$$
$$\text{(ii)}\quad E\sigma = \nu\,.$$

Received September 1972; revised March 1974.

[1] Prepared with the support of NSF grants GP-26066 and GP-36685X.

AMS 1970 *subject classifications.* Primary 62L10, 62L15, 62F05, 62F20; Secondary 60G40, 62E20.

Key words and phrases. Sequential analysis, sequential test, locally most powerful test, SPRT. exponential models, conjugate parameter pairs, optimal stopping.

(Probabilities and expectations, unless otherwise specified, are computed under θ^*.) Since such sequential tests are of the form studied by Wald (1947), the various inequalities and approximations for the SPRT are available. Indeed (as shown below in Section 5), for one-parameter exponential models, $f(x \mid \theta) = e^{\theta x - c(\theta)}$, an LMP sequential test is a Wald SPRT of $H_1': \theta = \theta_1$ vs. $H_2': \theta = \theta_2$, where $\theta_1 < \theta^* < \theta_2$. In the usual way, one argues, under suitable monotonicity conditions, that the tests in $\mathscr{S}$ are LMP for testing $H_1: \theta \leqq \theta^*$ vs. $H_2: \theta > \theta^*$.

This work solves a problem raised by Abraham (1969) in his thesis. He studied sequential tests of the form (1.2) and showed that in the symmetric binomial case, they are LMP. He left open the question of whether such tests are LMP more generally; a question to which we give here an affirmative answer. The class $\mathscr{S}$ was earlier considered by Dresselaers and Gilles (1951). They correctly surmised that such sequential tests are LMP, but their demonstration is heuristic, as it "neglects overshoot."

In the considerations below, we consider only sequential tests for which the stopping time, say t, satisfies $Et < \infty$. Together with the regularity conditions (on $f(\cdot \mid \cdot)$) assumed below, this entails that such sequential tests have power curves differentiable at θ^*. Hence a (level α) LMP test is one whose power curve has maximum slope at θ^*. (One can, of course, consider only right-hand derivatives, but we do not pursue that issue here.) If, on the other hand, the requirement $Et < \infty$ is dropped, some very curious phenomena can occur, even for the most regular of models. For example, in testing a normal mean, Darling and Robbins (1968) (see also Robbins (1970)) exhibit (level $\alpha < 1$) sequential tests of $H_1: \mu \leqq 0$ that have power 1 at all $\mu > 0$. The expected stopping times of these procedures are infinite for $\mu \leqq 0$ (in fact, for $\mu \leqq 0$, they terminate with probability not exceeding α). Clearly the Darling–Robbins tests qualify as being LMP (in fact, UMP) if the expected stopping time is unrestricted. Thus the requirement $Et < \infty$ eliminates from contention certain sequential procedures possessing features that, in some circumstances, may be deemed undesirable.

2. The main theorem. A sequential test of H_1 vs. H_2 is given by a stopping time, t, defined on $X_1, X_2, \cdots$, together with terminal decision rules $\{\varphi_n\}$. Here, for each n, φ_n is a possibly randomized (measurable) test function depending on $X_1, \cdots, X_n$. As mentioned above, we restrict attention to stopping times satisfying $Et < \infty$. For a given sequential test, its level is

$$\alpha = E\varphi_t \tag{2.1}$$

and more generally, its power at θ is

$$\beta(\theta) = E_\theta \varphi_t 1_{(t<\infty)} = \sum_1^\infty \int_{(t=n)} \varphi_n f_{n,\theta}\,, \tag{2.2}$$

where $f_{n,\theta}$ denotes the joint pdf of $X_1, \cdots, X_n$ under θ. The expression in (2.2) allows for the fact that we do not a priori require $P_\theta(t < \infty) = 1$ for $\theta \neq \theta^*$. In so doing, we interpret sampling forever ($t = \infty$) as neither accepting nor rejecting H_1. (We get the same result if $(t = \infty)$ is included in the acceptance region for H_1.) Assuming differentiability, the local slope (the slope of the power curve at θ^*) is

$$m = \partial E_\theta \varphi_t 1_{(t<\infty)}/\partial\theta|_{\theta=\theta^*} . \tag{2.3}$$

We refer to (α, m, Et) as the local characteristic (LC) of the sequential test.

We impose the following conditions on $f(\cdot \mid \cdot)$ which, as shown by Abraham (1969), are sufficient to insure that the derivative in (2.3) exists. Let

$$I(\theta \mid \theta^*) = E \log [f(X \mid \theta^*)/f(X \mid \theta)] . \tag{2.4}$$

ASSUMPTION 1. $\limsup_{\theta\to\theta^*} I(\theta \mid \theta^*)/(\theta - \theta^*)^2 < \infty$.

ASSUMPTION 2. $f(X \mid \cdot)$ is w.p. 1 differentiable at θ^*.

ASSUMPTION 3. The power function of every non-sequential test (of H_1 vs. H_2) is differentiable at θ^* under the integral sign.

As above, let

$$r(X) = \partial \log f(X \mid \theta)/\partial\theta|_{\theta=\theta^*} . \tag{2.5}$$

ASSUMPTION 4. $0 < Er^2(X) < \infty$.

(Various conditions that insure Assumption 3 are presented in the literature. Two that suffice are (i) For each θ in a bounded neighborhood U of θ^*, $f(X \mid \cdot)$ has, w.p. 1, a continuous derivative at θ. (ii) $\mathrm{Sup}\{\int |\partial f(x \mid \theta)/\partial\theta|\, d\lambda(x) : \theta \in U\} < \infty$. Note that condition (ii) holds if $\int \sup \{|\partial f(x \mid \theta)/\partial\theta| : \theta \in U\}\, d\lambda(x) < \infty$ or, more generally, if $\{\partial f(X \mid \theta)/\partial\theta : \theta \in U\}$ are uniformly integrable. For similar conditions, with further elaboration, see Bickel (1971, page 234).) We note that Assumptions 3 and 4 entail

$$Er(X) = 0 . \tag{2.6}$$

We have, then, the following result due to Abraham (1969).

1. PROPOSITION. *Suppose Assumptions* 1, 2 *and* 3 *hold. Then any sequential test* $(t, \{\varphi_n\})$ *with* $Et < \infty$ *has a power curve differentiable at* θ^*. *Moreover, the local slope is given by*

$$m = E\varphi_t S_t \tag{2.7}$$

where

$$S_n = \textstyle\sum_1^n r(X_j) . \tag{2.8}$$

REMARK. Equation (2.7) is obtained formally from (2.3) by differentiating across the expectation sign.

Since we are concerned only with tests for which $Et < \infty$, we assume in the sequel that (2.7) holds. Our main result is

2. THEOREM. *Suppose Assumptions* 1–4 *hold. Consider a sequential test* σ *in* $\mathscr{S}$ *and let* $(\alpha, m, E\sigma)$ *be its* LC. *If another sequential test of* H_1 *vs.* H_2 *has* LC $(\alpha, \hat{m}, Et)$,

then $Et \leqq E\sigma$ entails $\hat{m} \leqq m$. Moreover, if the first inequality is strict, so is the second. Thus σ is LMP *among all sequential tests of level $\hat{\alpha} \leqq \alpha$ and for which $Et \leqq E\sigma$.*

We note that under our assumptions, $P(r(X) = 0) < 1$. Thus a theorem of Stein (1946) insures that the mgf of σ is finite in a neighborhood of the origin. In particular, $E\sigma < \infty$.

To prove Theorem 2, we proceed analogously to the way one establishes the optimality of the Wald SPRT. That is, we reduce our considerations to solving a certain auxiliary optimal stopping problem. We begin by associating to the sequential test $(t, \{\varphi_n\})$, (for which $Et < \infty$), the value

$$(2.9) \qquad v(t, \{\varphi_n\}) = m - b\alpha - cEt = E\varphi_t S_t - bE\varphi_t - cEt\,.$$

(Compare (2.1) and (2.7). Note that since $Et < \infty$ and $E|r(X)| < \infty$, the three expectations in the last expression exist.) Here $b \in R$ and $c > 0$ and are otherwise arbitrary; c may be thought of as the cost per observation. Since $0 \leqq \varphi_t \leqq 1$, we have

$$(2.10) \qquad v(t, \{\varphi_n\}) = E(S_t - b)\varphi_t - cEt \leqq E(S_t - b)^+ - cEt = v(t)\,, \quad \text{say.}$$

Our auxiliary problem is whether there is a stopping time for which $v(t)$ is maximal.

3. An optimal stopping problem. The auxiliary problem formulated above is, in different notation, precisely the following optimal stopping problem: Let $Y, Y_1, \cdots$ be a sequence of i.i.d. random variables with $EY = 0$, $0 < EY^2 < \infty$. (Subsequently we identify Y with $r(X)$.) Let $S_n = Y_1 + \cdots + Y_n$. We define a reward sequence by

$$(3.1) \qquad w_n = (S_n - b)^+ - cn\,.$$

The optimal stopping problem is whether there exists a rule (= stopping time) defined on $Y_1, Y_2, \cdots$ that is optimal for the reward sequence $\{w_n\}$. That is, if C denotes the class of stopping time t (defined on $Y_1, Y_2, \cdots$) for which $Ew_t^- < \infty$, we then wish to know whether there is a stopping time, say τ, in C, for which

$$(3.2) \qquad Ew_\tau = \sup_{t \in C} Ew_t\,.$$

We restrict attention to rules that take at least one observation. The interested reader will have no difficulty taking into consideration the possibility $t = 0$. We note that although C could depend on b and c, in fact, it does not: For all b and c, $C = \{t : Et < \infty\}$. To see this, note that for any t,

$$(3.3) \qquad -ct \leqq w_t \leqq b^- + \sup_n (S_n - cn/2) - ct/2\,.$$

It follows from a result of Kiefer and Wolfowitz (1956) (see also [3], Theorem 4.13) that the supremum in (3.3) has finite expectation under our conditions.

Thus Et and Ew_t^- are finite or infinite together. Similar reasoning shows that $E\sup_n w_n^+ < \infty$ and hence, as an immediate consequence of Theorem 4.5 and the corollary to Theorem 5.2 of [3], there is an optimal rule for $\{w_n\}$ given by

$$\sigma_{b,c} = \inf\{n: u(S_n - b, c) \leqq (S_n - b)^+\}, \tag{3.4}$$

where

$$u(s, c) = \sup_{t\in C} E[(S_t + s)^+ - ct]. \tag{3.5}$$

Equivalently, we have

$$\sigma_{b,c} = \inf\{n: S_n \notin A_c + b\}, \tag{3.6}$$

where

$$A_c = \{s: u(s, c) > s^+\}. \tag{3.7}$$

We see from (3.6) that A_c is the optimal continuation set when $b = 0$ and that more generally, the continuation set for $\sigma_{b,c}$ is the translated set $A_c + b$. This particular optimal stopping problem is an instance of the so-called stationary Markov case. A good account of the known general results is given in Chapter 5 of [3].

We may characterize the optimal rule, via its continuation set, more specifically as follows.

3. THEOREM. *Under the conditions $EY = 0$ and $0 < EY^2 < \infty$, the continuation set A_c is a bounded open interval containing zero. Moreover, given any bounded open interval A, there are constants b and c so that A is the continuation set for $\sigma_{b,c}$; that is, $A = A_c + b$.*

REMARK. Making the identification $Y = r(X)$, we see that the stopping times of tests in $\mathscr{S}$ are optimal in the present sense.

PROOF. We first establish some properties of

$$u(s, c) - s^+ = \sup_{t\in C} E[(S_t + s)^+ - s^+ - ct] \tag{3.8}$$

as a function of s. For arbitrary values of t and S_t, the expression in brackets is convex increasing (= non-decreasing) for $s \leqq 0$ and convex decreasing for $s \geqq 0$. (Note that for $s \geqq 0$, $(S_t + s)^+ - s = \max(S_t, -s)$.) Since these properties are preserved in taking the indicated expectation and supremum, they hold for $u(s, c) - s^+$. It follows that A_c is empty or an interval containing zero, open because the above properties of $u(s, c) - s^+$ ensure it is continuous. If A_c were unbounded, then $E\sigma_{b,c}$ would be infinite (since $EY = 0$), contrary to the fact that $\sigma_{b,c} \in C$. (A more direct demonstration that A_c is bounded is given below in Lemma 5.)

We next establish some auxiliary results that give further properties of A_c. These enable us, *inter alia*, to prove the last part of Theorem 3. Let $\gamma = EY^+$.

4. LEMMA. *A_c is empty for $c \geqq \gamma$.*

PROOF. Clearly $u(s, c)$ is decreasing in c, so that A_c is also decreasing in c. (If $c < d$, $A_c \supset A_d$.) It thus suffices to show that $A_\gamma = 0$. Since $S_t^+ - \gamma t \leqq \sum_1^t (Y_j^+ - \gamma)$, it follows from Wald's lemma that for $t \in C$, $E(S_t^+ - \gamma t) \leqq 0$, hence $u(0, \gamma) \leqq 0$. In fact, $u(0, \gamma) = 0$ since $u(0, \gamma) \geqq E(S_t^+ - \gamma t)$, which vanishes for $t \equiv 1$. As is easily checked, $u(s, c) \leqq u(0, c) + s^+$, hence $u(s, \gamma) - s^+ \leqq u(0, \gamma) = 0$. It follows that $A_\gamma = \emptyset$. □

5. LEMMA. $A_c \subset (-EY^2/c, EY^2/c)$. *Also, there is a constant* $K > 0$ *so that for* c *sufficiently small*, $(-K/c, K/c) \subset A_c$.

PROOF. To establish the first assertion, we note that

$$E[(S_t + s)^+ - ct] - s = E[\max (S_t, -s) - ct] . \tag{3.9}$$

For $t \in C$ and $s > 0$,

$$E \max (S_t, -s) \leqq \int_{(S_t \geqq -s)} S_t = -\int_{(S_t < -s)} S_t \leqq ES_t^2/s = EY^2 Et/s .$$

(The first equality follows from Wald's first lemma: $ES_t = 0$; the second, from Wald's second lemma.) Thus $E[(S_t + s)^+ - ct] - s \leqq Et(EY^2/s - c) \leqq EY^2/s - c < 0$ if $s > EY^2/c$ and then $u(s, c) - s^+ < 0$. Similarly, for $s < -EY^2/c$,

$$\begin{aligned} E[(S_t + s)^+ - ct] &\leqq \int_{(S_t > -s)} S_t - cEt \\ &\leqq -ES_t^2/s - cEt \leqq -EY^2/s - c < 0 . \end{aligned}$$

Thus for $|s| > EY^2/c$, $u(s, c) < s^+$, which entails $A_c \subset (-EY^2/c, EY^2/c)$.

To establish the second part of the lemma, we see from (3.9) that for $s < 0$, $E[(S_t + s)^+ - ct] \geqq E[S_t^+ - ct] + s$, or that

$$u(s, c) \geqq u(0, c) + s , \qquad s < 0 , \tag{3.10}$$

while, since $u(s, c)$ is monotone increasing in s,

$$u(s, c) - s \geqq u(0, c) - s , \qquad s > 0 . \tag{3.11}$$

It follows that $A_c \supset (-u(0, c), u(0, c))$ for $c < \gamma$. Moreover, $u(0, c) \geqq \sup_n (ES_n^+ - cn)$. This lower bound may be approximated as follows. Since $S_n/n^{\frac{1}{2}} \to_{\mathscr{L}} Z \sim N(0, EY^2)$ and $S_n/n^{\frac{1}{2}}$ is uniformly integrable, it follows that $ES_n^+/n^{\frac{1}{2}} \to EZ^+ = (EY^2/2\pi)^{\frac{1}{2}}$. Hence for some $k > 0$, $ES_n^+ \geqq kn^{\frac{1}{2}}$, all $n \geqq 1$. Therefore $u(0, c) \geqq \max_n (kn^{\frac{1}{2}} - cn) > K/c$ for some $K > 0$, if c is sufficiently small. Thus for sufficiently small c, $A_c \supset (-K/c, K/c)$. □

We now establish the last part of Theorem 3. Since b merely translates the continuation set, it suffices to show that $|A_c|$, the length of A_c, increases continuously from 0 to ∞ as c decreases from γ to 0. We note first that $|A_c|$ increases as c decreases. Then from the proof of Lemma 5, we see that $u(s, c) - s^+ \leqq EY^2/s - c < 0$ if $|s| > EY^2/c$. By convexity, it follows that as a function of s, $u(s, c) - s^+$ is strictly increasing on $\{s < 0: u(s, c) - s^+ > -c\}$ and is strictly decreasing on $\{s > 0: u(s, c) - s^+ > -c\}$. Since also $u(s, c) - s^+$ is (jointly) continuous in s and c, it follows that $|A_c|$ increases continuously as c

decreases. It follows also from Lemma 5 that $|A_c| = O(1/c)$ and in particular, $\lim_{c\to 0} |A_c| = \infty$.

4. Proof of the main theorem. We now establish Theorem 2, using the results of Section 3. Let σ be a test in $\mathscr{S}$ with continuation interval $A = (-a_1, a_2)$ and let $(\alpha, m, E\sigma)$ be its LC. Making the identification $Y = r(X)$, by Theorem 3 there are numbers b and $c > 0$ so that σ is the optimal rule for $\{w_n\}$ and $A = A_c + b$. Since $0 \in A_c$, $b \in A$; i.e.

$$-a_1 < b < a_2 . \tag{4.1}$$

The terminal decision rule associated with σ is $1_{(S_\sigma \geqq a_2)}$ and since w.p. 1 either $S_\sigma \leqq -a_1$ or $S_\sigma \geqq a_2$, we see from (4.1) that w.p. 1

$$(S_\sigma \geqq a_2) = (S_\sigma > b) . \tag{4.2}$$

Then, in view of (4.2), it follows from (2.9) that

$$\begin{aligned} m - b\alpha - cE\sigma &= E(S_\sigma - b)1_{(S_\sigma \geqq a_2)} - cE\sigma \\ &= E(S_\sigma - b)^+ - cE\sigma . \end{aligned} \tag{4.3}$$

Let $(t, \{\varphi_n\})$ be another sequential test having LC $(\alpha, \hat{m}, Et)$ with $Et \leqq E\sigma$. From (2.9) and (2.10) we obtain

$$\begin{aligned} \hat{m} - b\alpha - cEt &\leqq E(S_t - b)^+ - cEt \\ &\leqq E(S_\sigma - b)^+ - cE\sigma = m - b\alpha - cE\sigma . \end{aligned} \tag{4.4}$$

The second inequality follows because σ is optimal for $\{w_n\}$ and the final equality is (4.3). Since $cEt \leqq cE\sigma$, we see that $\hat{m} \leqq m$ and moreover, if $Et < E\sigma$, then $\hat{m} < m$. Thus among level α tests satisfying $Et \leqq E\sigma$, σ is LMP. If a sequential test has LC $(\hat{\alpha}, \hat{m}, Et)$ with $Et < \infty$ and $\hat{\alpha} < \alpha$, the power function of this test, being differentiable at θ^* is necessarily continuous there. Thus, trivially, σ is more powerful locally. This completes the proof of Theorem 2.

5. Exponential models. We have noted above that an LMP sequential test (given by (1.2)) is like an SPRT (in that σ is the first time a certain random walk leaves an interval). For one-parameter exponential models, the LMP sequential tests *are* SPRT's. To show this, we parametrize the model so that $f(x \mid \theta) = \exp\{\theta x - c(\theta)\}$ and denote by Θ, the natural parameter set: $\Theta = \{\theta : \int e^{\theta x} d\lambda(x) < \infty\}$. Θ is necessarily an interval (which, to avoid trivialities, we assume is non-degenerate). For a given θ^* in the interior of Θ, $r(x) = x - c'(\theta^*)$, where $c'(\theta) = dc(\theta)/d\theta$. (It is well known that $c(\cdot)$ is infinitely differentiable (on the interior of Θ) and strictly convex (unless λ is a degenerate measure, a case we exclude from consideration). Thus $c'(\cdot)$ is strictly increasing. As an aside, we note that $c'(\theta) = E_\theta X$, so of course $Er(X) = 0$.) As these exponential models have a monotone likelihood ratio, it follows in the usual way [9, page 101] that a level α LMP sequential test of $H_1 : \theta = \theta^*$ vs. $H_2 : \theta > \theta^*$ is also level α LMP for testing $H_1 : \theta \leqq \theta^*$ vs. $H_2 : \theta > \theta^*$. Given an LMP sequential test of $H_1 : \theta \leqq \theta^*$ vs. $H_2 : \theta > \theta^*$, we show that there is a continuum of parameter pairs (θ_1, θ_2) in Θ

for which an SPRT of H_1': $\theta = \theta_1$ vs. H_2': $\theta = \theta_2$ coincides with the given LMP test.

We begin by noting that for given $\theta_1 < \theta_2$ in Θ, an SPRT of H_1' vs. H_2' stops the first time $\sum_1^n \log [f(X_j \mid \theta_2)/f(X_j \mid \theta_1)]$ leaves a certain interval. However, $\log [f(x \mid \theta_2)/f(x \mid \theta_1)] = (\theta_2 - \theta_1)x + c(\theta_1) - c(\theta_2)$, so equivalently, the SPRT terminates when $\sum_1^n X_j - n[c(\theta_2) - c(\theta_1)]/(\theta_2 - \theta_1)$ leaves a corresponding interval. The LMP sequential test of H_1 vs. H_2 stops the first time $\sum_1^n X_j - nc'(\theta^*)$ leaves a certain interval. We establish the desired result by showing that there is a continuum of pairs (θ_1, θ_2) of points in Θ for which

$$[c(\theta_2) - c(\theta_1)]/(\theta_2 - \theta_1) = c'(\theta^*) . \tag{5.1}$$

The following argument was suggested by Professor R. A. Wijsman. Let $f(\theta) = c(\theta) - c(\theta^*) - (\theta - \theta^*)c'(\theta^*)$; $f(\cdot)$ is strictly convex since $c(\cdot)$ is. The problem is solved if it can be shown that there is a continuum of pairs (θ_1, θ_2) with $\theta_1 < \theta^* < \theta_2$ so that $f(\theta_1) = f(\theta_2)$. For then it is easily verified that (5.1) holds for θ_1 and θ_2. Let f_1 and f_2 be f restricted to $(-\infty, \theta^*)$ and (θ^*, ∞), respectively. Since $f(\theta^*) = f'(\theta^*) = 0$, f_1 and f_2 are strictly monotone and tend to ∞ as $\theta \to \pm\infty$. Thus f_i^{-1} exists on $(0, \infty)$. Thus for any $y > 0$, $\theta_i = f_i^{-1}(y)$ provides a conjugate pair.

The above device of introducing $f(\cdot)$ is essentially the same argument used by Girshick (1946), who considered the phenomenon of conjugate pairs from a slightly different point of view. His motivation was to obtain the OC curve of an SPRT. In this context, see also Lechner (1964), who apparently rediscovered the idea of conjugate pairs and indicated that all such pairs lead to the same SPRT. In the context of SPRT's, θ^* is the exceptional point of the SPRT, the parameter value under which the random walk defining the SPRT has zero drift. The observation that an LMP sequential test is an SPRT thus has a dual formulation: For one-parameter exponential models, an SPRT of two simple (or one-sided) hypotheses is an LMP sequential test for its exceptional point.

Still in the context of one-parameter exponential models, Lechner further observed that if a given SPRT has stopping time σ and power function $\beta(\cdot)$ and if (θ_1, θ_2) is a conjugate pair for this test, then for any other sequential test, having stopping time $\hat{\sigma}$ and power function $\hat{\beta}(\cdot)$, say, the optimality property of the SPRT entails

$$\hat{\beta}(\theta_1) \leqq \beta(\theta_1) \quad \text{and} \quad \hat{\beta}(\theta_2) \geqq \beta(\theta_2) \Rightarrow E_{\theta_i}\hat{\sigma} \geqq E_{\theta_i}\sigma , \qquad i = 1, 2 . \tag{5.2}$$

Lechner also stated (without proof) that for the exceptional point, θ^*, of the given SPRT, if $\hat{\beta}(\theta^*) = \beta(\theta^*)$, then it cannot happen simultaneously that

$$\hat{m} > m \quad \text{and} \quad E\hat{\sigma} < E\sigma . \tag{5.3}$$

Here m and $\hat{m}$ denote the slopes of the respective power curves at θ^*. A proof of this last statement could presumably be based on (5.2) and the fact that there are conjugate pairs arbitrarily close to θ^* (y near zero). The argument appears to require the fact that if a sequential test satisfies $E\hat{\sigma} < \infty$, then its ASN function is continuous at θ^*. The author is unable to locate any general results to this effect in the literature, even for SPRT's. Note that the negation of (5.3) does not rule out the possibility that $\hat{m} > m$ and $E\hat{\sigma} = E\sigma$. Thus although tempting, it is not clear that the existence of conjugate pairs can be used to establish the LMP character of the tests (1.2) for one-parameter exponential models.

6. Acknowledgment. The author is indebted to Professor D. Siegmund for helpful discussions and to Professor R. A. Wijsman and a referee for improvements in the exposition.

REFERENCES

[1] ABRAHAM, J. K. (1969). The local power of sequential tests subject to an expected sample size restriction. Unpublished Stanford technical report.

[2] BICKEL, P. J. (1971). *Mathematical Statistics, Part I* (preliminary edition). Holden-Day, San Francisco.

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

[4] DARLING, D. A. and ROBBINS, H. (1968). Some further remarks on inequalities for sample sums. *Proc. Nat. Acad. Sci.* **60** 1175–1182.

[5] DRESSELAERS, C. and GILLES, P. (1951). Un test sequential unilateral. *Acad. Roy. Belg. Bull. Cl. Sci.* **37** 713–727.

[6] GIRSHICK, M. A. (1946). Contributions to the theory of sequential analysis III. *Ann. Math. Statist.* **17** 291–298.

[7] KIEFER, J. and WOLFOWITZ, J. (1956). On the characteristics of the general queuing process with applications to random walk. *Ann. Math. Statist.* **27** 147–161.

[8] LECHNER, J. A. (1964). Optimality and the OC curve for the Wald SPRT. *J. Amer. Statist. Assoc.* **59** 464–468.

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

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

[11] STEIN, C. (1946). A note on cumulative sums. *Ann. Math. Statist.* **17** 498–499.

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

STATISTICS CENTER
RUTGERS UNIVERSITY
NEW BRUNSWICK, N.J. 08903