SEQUENTIAL ANALYSIS AND THE LAW OF THE ITERATED LOGARITHM*

Hans Rudolf Lerche

University of Heidelberg

Some recent developments in sequential statistics are reviewed which are connected with the law of the iterated logarithm. On the one hand optimal properties of sequential tests with parabolic boundaries are discussed, on the other hand approximations to curved boundary crossing distributions of Brownian motion. The connection of both topics is also indicated.

1. Introduction.

H. Robbins was one of the first who noted that the law of the iterated logarithm (LIL) has some statistical meaning. In his 1952 paper he observed that the repeated significance test (RST) gives false alarm eventually with probability one. He further asked for the operating characteristics of the RST if it is truncated.

Later Robbins found a way to correct the misbehavior of the RST. He discovered the positive side of the LIL. His idea was that, to control the error probabilities over all sample sizes, one has just to make the boundaries a little bit wider than $\sqrt{n}$. However this meant the construction of tests of power one (cf. Robbins, (1970)).

Together with Darling and Siegmund, Robbins studied the error probabilities and expected sample sizes of these tests. His students

---

* This work was supported by the Deutsche Forschungsgemeinschaft.
AMS 1980 subject classifications. Primary 62L15; Secondary 60G40, 62C10.
Key words and phrases. Sequential Bayes tests for composite hypotheses, repeated significance tests and tests of power one, first exit distributions, tangent approximation.

Lai, Siegmund and others went further, and also studied the operating characteristics of the truncated RST. Their main tool was an extension of a classical technique of sequential analysis, viz., working with the martingale property of mixtures of likelihood-ratios.

I wish to offer some complementary views of the subject. In the first part I shall discuss optimality properties of tests of power one and repeated significance tests. It will turn out that, if one lets the sampling cost depend on the underlying parameter in a natural way, the theory of these procedures fits naturally in the classical sequential testing framework.

In the second part of the paper I shall discuss the tool with which I found the results originally, namely the tangent approximation to first exit distributions of Brownian motion over curved boundaries. Many of the results of Robbins and his colleagues can be derived with this approximation device; in this way they become better understood from the viewpoint of classical fluctuation theory. Finally this excursion through sequential analysis will lead to some interesting new perspectives of the LIL, in particular a new way of proof, and tail probabilities of first exit distributions of Brownian motion over lower class functions (e.g. $(2t \log \log t)^{1/2}$) as $t \to \infty$.

2. Optimality of sequential tests with parabolic and nearly parabolic boundaries.

Bayes tests of power one. We discuss the testing problem of whether the drift is different from zero. Let W(t) denote Brownian motion with unknown drift $\theta \in \mathbf{R}$ and P_θ the associated measure. Let F be a prior on $\mathbf{R}$ given by $F = \gamma\delta_0 + (1-\gamma) \int_{\cdot} \phi(\sqrt{r}\theta)\sqrt{r}d\theta$ with $0 < \gamma < 1$ and $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, consisting of a point mass at $\{\theta = 0\}$ and a smooth normal part on $\{\theta \neq 0\}$. Let the sampling cost be $c\theta^2$, with $c > 0$ for the observation of W per unit time when the underlying measure is P_θ. We assume a loss function which is equal to 1 if $\theta = 0$ and we decide in favour of "$\theta \neq 0$" and which is identically 0 if $\theta \neq 0$. A statistical test consists of a stopping time T of Brownian motion where stopping means a decision in favour of "$\theta \neq 0$". The Bayes risk for this problem is then given by

$$(1) \qquad R(T) = \gamma P_o(T < \infty) + (1-\gamma)c \int_{-\infty}^{\infty} \theta^2 E_\theta T \phi(\sqrt{r}\theta)\sqrt{r}d\theta .$$

The issue is to determine the stopping rule T_c^* which minimizes R(T). For the cost c sufficiently small, T_c^* is a test of power one for the decision problem $H_o : \theta = 0$ versus $H_1 : \theta \neq 0$ (the definition of a test of power one can be found in the paper of Robbins (1970)).

The statistical meaning of the sampling cost "$c\theta^2$" becomes apparent by the following consideration. Let us consider two testing problems of the simple hypotheses:

1) $H_o : \theta = 0$ versus $H_1 : \theta = \theta_1$,

2) $H_o : \theta = 0$ versus $H_1 : \theta = \theta_2$

with $\theta_i > 0$, i=1,2. Let t_i, i=1,2 denote the sample sizes. Then the level-α Neyman-Pearson tests have the same power if and only if $\theta_1^2 t_1 = \theta_2^2 t_2$. Thus the factor θ^2 standardizes the sample sizes in such a way that the testing problems are equivalent.

There is also a basic mathematical reason for letting the sampling costs depend on θ. The Bayes risk (1) with a constant or $|\theta|$ instead of θ^2 would be infinite for all nontrivial stopping times. For more details see Lerche (1985a).

The problem corresponding to (1) for simple hypotheses has an exact solution. Its Bayes risk is given by

$$(2) \qquad R(T) = \gamma P_o(T < \infty) + (1-\gamma)c\theta^2 E_\theta T.$$

The stopping rule T_c^* which minimizes (2) can be calculated directly by using the following result due to Darling and Robbins (1967). It states that for every stopping rule with $P_o(T < \infty) < 1$,

(3) $$E_\theta T \geq 2 \log b/\theta^2, \quad \text{where} \quad b = P_o(T < \infty)^{-1}.$$

Equality in (3) holds for the stopping rule

(4) $$T^* = \inf\{t > 0 \mid \frac{dP_{\theta,t}}{dP_{o,t}} \geq b\},$$

where $\frac{dP_{\theta,t}}{dP_{o,t}}$ denotes the likelihood ratio of P_θ with respect to P_o given the path $W(\mu)$, $0 \leq \mu \leq t$; it is given by

$$\frac{dP_{\theta,t}}{dP_{o,t}} = \exp(\theta W(t) - \frac{1}{2}\theta^2 t).$$

PROPOSITION 1.

(5) $$\min_T R(T) = R(T_c^*) = 2(1-\gamma)c\ [1 + \log b(c)]$$

where T_c^* is given by (4) with the constant $b(c) = \gamma[2(1-\gamma)c]^{-1}$. The minimum is taken over all stopping times $T \geq 0$.

$R(T_c^*)$ does not depend on θ. This suggests that T_c^* can also be expressed in a form independent of θ. In a Bayes formulation this is in fact possible.

Let $F_{x,t}$ denote the posterior with respect to the prior $F = \gamma\delta_o + (1-\gamma)\delta_\theta$ given that $W(t) = x$. Then $F_{x,t} = \gamma(x,t)\delta_o + (1-\gamma(x,t))\delta_\theta$ with $\gamma(x,t) = F_{x,t}(\{0\}) = [1 + \frac{1-\gamma}{\gamma}\frac{dP_{\theta,t}}{dP_{o,t}}(x)]^{-1}$. A little calculation shows that

$$T_c^* = \inf\{t > 0 \mid F_{W(t),t}(\{0\}) \leq \frac{2c}{1+2c}\}.$$

We call such a stopping rule a simple Bayes rule.

We turn now to the case of composite hypothesis with the risk (1). Since by our choice of the sampling cost the simple testing problem (2) has a solution which does not explicitly depend on θ, one might guess that a simple Bayes rule will also be nearly optimal for the risk (1). In fact this is the next result.

Let $F_{x,t}$ now denote the posterior with respect to the prior $F = \gamma\delta_o + (1-\gamma) \int \phi(\sqrt{r}\theta)\sqrt{r}d\theta$ given that $W(t) = x$. It is given by

$$F_{x,t} = \gamma(x,t)\ \delta_o + (1-\gamma(x,t))G_{x,t} \text{ where}$$

$$G_{x,t} = N\ [\frac{x}{t+r}, \frac{1}{t+r}] \text{ and}$$

$$\gamma(x,t) = F_{x,t}(\{0\}) = [1 + \frac{1-\gamma}{\gamma} \int \frac{dP_{\theta,t}}{dP_{o,t}} \phi(\sqrt{r}\theta)\sqrt{r}d\theta]^{-1}.$$

Let T_c denote the simple Bayes rule

$$T_c = \inf\{t > 0 \mid F_{W(t),t}(\{0\}) < \frac{2c}{1+2c}\} \tag{6}$$

and T_c^* the optimal stopping rule for the risk (1).

THEOREM 1.

$$R(T_c^*) = R(T_c) + o(c) \quad \text{as} \quad c \to 0. \tag{7}$$

$$R(T_c) = 2(1-\gamma)c\ [\log b + 1 + \frac{1}{2} \log(2 \log b) - A + o(1)] \tag{8}$$

as $c \to 0$. Here $b = \gamma[2(1-\gamma)c]^{-1}$ and $A = 2 \int_o^\infty \log x\ \phi(x)\ dx$.

Equation (7) states that the risks of the stopping rules T_c^* and T_c differ only by a o(c)-term. The risk of T_c is given by equation (8) up to a o(c)-term. A comparison of the equations (5) and (8) shows that the term $2(1-\gamma)c\ [1/2 \log(2 \log b) - A + o(1)]$ is the increase of the risk due to the ignorance of the parameter $\theta \neq 0$.

We note that the stopping rule (6) can also be expressed as

$$T_c = \inf\{t > 0 \mid \int \frac{dP_{\theta,t}}{dP_{o,t}} \phi(\sqrt{r}\theta)\sqrt{r}d\theta > b\}$$

$$= \inf\{t > 0 \,\Big|\, |W(t)| > \sqrt{(t+r)(\log(\tfrac{t+r}{r}) + 2 \log b)}\}$$

which is a familiar mixture stopping rule of the work of Robbins et al. We also note that the nearly optimal stopping boundary contains only a single log-term and not an iterated one as one might have expected from the LIL. For more details and for proofs see Lerche (1985 and 1986a).

Repeated significance tests. We consider the problem of testing the sign of the drift of Brownian motion $W(t)$. The parameter sets of H_o and H_1 are given by $\Theta_o = \{\theta < 0\}$ and $\Theta_1 = \{\theta > 0\}$. The observation cost is again $c\theta^2$ where c is a positive constant. On the parameter set $\{\theta \neq 0\}$ we put the normal prior $G(d\theta) = \phi(\sqrt{r}\theta)\sqrt{r}d\theta$. The Bayes risk for a decision procedure (T,δ), consisting of a stopping time T of $W(t)$ and a terminal decision rule δ, is given by

$$R(T,\delta) = \int_{\Theta_o} (P_\theta\{H_o \text{ rejected } (\delta)\} + c\theta^2 E_\theta T)G(d\theta) \tag{9}$$

$$+ \int_{\Theta_1} (P_\theta\{H_1 \text{ rejected } (\delta)\} + c\theta^2 E_\theta T)G(d\theta).$$

The objective is to find a decision procedure (T^*,δ^*) with minimal Bayes risk.

Let $G_{x,t}$ denote the posterior of θ with respect to the prior G given that $W(t) = x$. $G_{x,t} = N(\frac{x}{t+r}, \frac{1}{t+r})$. Let T be an arbitrary stopping time. δ^* denotes the terminal decision rule after stopping at time T, which rejects the hypothesis H_o if and only if $G_{W(T),T}(\Theta_o) < G_{W(T),T}(\Theta_1)$. It is well known that $R(T,\delta^*) \leq R(T,\delta)$ for all decision rules δ.

For $\lambda > 0$ the simple Bayes rule is defined as

$T_\lambda = \inf\{t > 0 \mid \min_{i=0,1} G_{W(t),t}(\Theta_i) \leq \Phi(-\lambda)\}$ where Φ denotes the standard normal distribution function. It can also be expressed as

$T_\lambda = \inf\{t > 0 \mid |W(t)| \geq \lambda\sqrt{t+r}\}$, which is the stopping time of the repeated significance test. The following result states that the repeated significance test is an exact Bayes test for the risk (9).

THEOREM 2. Let $0 < c < \infty$. Let $\lambda(c)$ denote the solution of the equation $\phi(\lambda)/\lambda = 2c$ and let $T^* = T_{\lambda(c)}$. Then $\min_{(T,\delta)} R(T,\delta) = R(T^*,\delta^*)$.

There is a well-known analogous result for simple hypothesis which determines the sequential probability-ratio test (SPRT) as a Bayes test. In our setup - with observation cost "$c\theta^2$" - the repeated significance test turns out as an adaptive version of Wald's SPRT. For more details on this topic and for the proof of Theorem 2 see Lerche (1985 and 1986b).

The following heuristic argument was what originally led us to Theorem 1. First rewrite the Bayes risk (1) as an integral with respect to the measure of Brownian motion without drift: $R(T) = \int h(|W(T)|,T)dP_o$. Since the original problem is symmetric, the stopping times

$$T = \inf\{t > 0 \mid |W(t)| > \psi(t)\} \tag{10}$$

are the relevant competitors. For a fixed T given by (10), let $p_\psi(t)$ denote the density of T under the measure P_o. Then

$$R(T) = \int h(\psi(t),t)\ p_\psi(t)dt.$$

Since explicit formulas for p_ψ are unavailable for every ψ, we try to work instead of p_ψ with a good approximation to p_ψ. As will be explained in the next section the tangent approximation is such a useful approximation. With it and a variational argument over ψ we found that the optimal stopping boundary ψ^* asymptotically grows like $(t \log t)^{1/2}$ as $t \to \infty$. For the details (see Lerche (1985)).

3. The tangent approximation.

Let W(t) denote the standard Brownian motion starting from zero at time zero. Let $\psi(t)$ be a continuously differentiable, positive function on $\mathbf{R}_+ = (0,\infty)$ which belongs to the upper class at zero. Let $T = \inf\{t > 0 \mid W(t) \geq \psi(t)\}$ denote the first exit time of W over ψ. By a result of Strassen (1967) the distribution of T has a continuous density p. However this is explicitly known only for few boundaries. For $\psi(t) = \Lambda + bt$, the Bachelier-Levy formula states that

$$(11) \qquad p(t) = \frac{\Lambda}{t^{3/2}} \phi(\frac{\psi(t)}{\sqrt{t}}).$$

Several authors (Daniels (1974), Lorden (1973), Strassen (1967)) had the idea of the tangent approximation, viz. to approximate p(t) by the first exit density at t of Brownian motion over the tangent to the curve at ψ. The tangent approximation (TA) is given by $\frac{\Lambda(t)}{t^{3/2}} \phi(\frac{\psi(t)}{\sqrt{t}})$, where $\Lambda(t) = \psi(t) - t\psi'(t)$ denotes the intercept of the tangent to the curve ψ at t on the space-axis.

Strassen (1967) proved that the TA holds as $t \to 0$, i.e.,

$$(12) \qquad p(t) = \frac{\Lambda(t)}{t^{3/2}} \phi(\frac{\psi(t)}{\sqrt{t}}) (1+o(1)).$$

He used (12) to give an intuitive geometric proof of the difficult part of the Kolmogorov-Petrovski-Erdös test, which is a generalization of the LIL. (A geometric proof of the easy half is given on p.33 of Ito-McKean (1974)).

We state now some results about the TA as a global approximation device when the boundaries recede to infinity.

Let $\{\psi_a;\ a \in \mathbf{R}_+\}$ denote a set of positive, monotone increasing, continuously differentiable functions on $\mathbf{R}_+$. Let $T_a = \inf\{t > 0 \mid W(t) \geq \psi_a(t)\}$ and p_a the density of the distribution of T_a. Let $\Lambda_a(t) = \psi_a(t) - t\psi_a'(t)$. The following theorem is very similar to Theorem 1 of Jennen-Lerche (1981).

THEOREM 3. Let $0 < t_1 \leq \infty$ and $0 < \alpha < 1$. Assume that

(i) $P(T_a < t_1) \to 0$ as $a \to \infty$,

(ii) $\psi_a(t)/t^\alpha$ is monotone decreasing in t for each a,

(iii) for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all a,

$|\psi_a'(s)/\psi_a'(t) - 1| < \varepsilon$ if $|s/t - 1| < \delta$ for $s,t \in (0,t_1)$.

Then

$$(13)\quad p_a(t) = \frac{\Lambda_a(t)}{t^{3/2}} \phi\left(\frac{\psi_a(t)}{\sqrt{t}}\right)(1 + o(1)) \text{ uniformly on } (0,t_1) \text{ as } a \to \infty.$$

Integration yields the following corollary which resembles a result of Cuzick (1981).

COROLLARY 1.

$$(14)\quad P(T_a < t) = \int_0^t \frac{\Lambda_a(u)}{u^{3/2}} \phi\left(\frac{\psi_a(u)}{\sqrt{u}}\right) du(1 + o(1)) \text{ uniformly on } (0,t_1) \text{ as } a \to \infty.$$

We add several remarks to these results. At first we note that the tangent approximation is a purely local approximation. The quantity $\frac{\Lambda_a(t)}{t^{3/2}} \phi\left(\frac{\psi_a(t)}{\sqrt{t}}\right)$ is usually not a probability density (except for straight lines).

Assumption (i) can easily be checked by using the inequality

$$P(T_a < t_1) \le \int_0^{t_1} \frac{\psi_a(t)}{t^{3/2}} \phi\left(\frac{\psi_a(t)}{\sqrt{t}}\right) dt$$

which holds for monotone functions ψ_a.

The case $t_1 = \infty$ is included. Example (15) (below) is of this type. The other examples satisfy the conditions of Theorem 3 on finite intervals:

$$(15)\qquad \psi_a(t) = \sqrt{(t+1)(2a + \log(t+1))},$$

$$(16)\qquad \psi_a(t) = at^\alpha,\ \alpha < \tfrac{1}{2},$$

(17) $$\psi_a(t) = \sqrt{2(r+at)},\ r > 0,$$

(18) $$\psi_a(t) = \sqrt{a}\psi(t/a),\ \psi \text{ a fixed function.}$$

The last example shows that Strassen's result on the tangent approximation (Theorem 3.5 of Strassen (1967)) is contained as a special case of Theorem 3.

For the boundary (15) the formula (14) yields

(19) $$P(T_a < \infty) = \frac{1}{2} e^{-a}(1 + o(1)) \text{ and}$$

(20) $$P(T_a < at_1) = e^{-d}(1- \Phi\ (\sqrt{2/t_1}))(1 + o(1)).$$

The first result agrees with a famous exact result of Robbins-Siegmund (1970) for the two-sided case, the second one with an asymptotic result of Lai-Siegmund (1977).

For the boundary (17) the formula (14) yields

(21) $$P(T < t_1) = \sqrt{a}\ e^{-a} \frac{1}{\sqrt{\pi}} \int_o^t \frac{1+2\frac{a}{t}}{\sqrt{1+\frac{a}{t}}} e^{-r} \frac{dt}{t} (1 + o(1))$$

$$= \sqrt{a}\ e^{-a} \frac{1}{\sqrt{\pi}} \int_{\sqrt{2/t_1}}^{\infty} e^{-r\theta^2/2} \frac{d\theta}{\theta} (1 + o(1)),$$

which agrees with a well known result of Siegmund (1977) for the RST. Also Siegmund's second order result for the RST can be rederived by a refinement of the TA due to Jennen (1985).

Now we try to explain the crucial idea of the proof of Theorem 3, which is closely related to large deviation theory. First note that assumption (i) implies

(i') $$\inf_{o<t\leq t_1} \psi_a(t)/\sqrt{t} \to \infty \quad \text{as} \quad a \to \infty.$$

This can be seen from the following inequality

$$1 - \Phi(\frac{\psi_a(t)}{\sqrt{t}}) = P(W(t) \geq \psi_a(t)) \leq P(T_a \leq t_1) \to 0 \text{ as } a \to \infty \text{ for all } 0 < t \leq t_1.$$

Now we consider instead of Brownian motion the Brownian bridge with space-time endpoints $(0,0)$ and $(\psi_a(t),t)$. If for every $0 < \varepsilon < 1$

$$(22) \qquad P((1-\varepsilon)t < T_a < t | W(t) = \psi_a(t)) = 1 - o(1)$$

uniformly on $(0,t_1]$ as $a \to \infty$, it is intuitively clear that the tangent approximation will hold. But why does (22) hold? To indicate the answer let $\tilde{W}_o$ denote the Brownian bridge with endpoints $(0,0)$ and $(\psi_a(t),t)$. Then

$$(23) \qquad \sup_{o<v<1} \left| \frac{\tilde{W}_o(vt)}{\psi_a(t)} - 1 \right| \longrightarrow 0 \text{ uniformly on } (0,t_1] \text{ as } a \to \infty.$$

This can be seen as follows. Let $W_o(vt) = \tilde{W}_o(vt) - v\psi_a(t)$. Then by the scaling property of the Brownian bridge

$$\sup_{o<v<1} | \tilde{W}_o(vt)/\psi_a(t) - v| = \sup_{0<v<1} \left| \frac{W_o(vt)/\sqrt{t}}{\psi_a(t)/\sqrt{t}} \right| \overset{\mathcal{L}}{=} \sup_{0<v<1} \left| \frac{W_o(v)}{\psi_a(t)/\sqrt{t}} \right|$$

But this expression converges to zero uniformly on $(0,t_1]$ as $a \to \infty$ by (i').

Thus (23) describes the fact that, if the endpoint is high, the Brownian bridge takes nearly the shortest way between $(0,0)$ and $(\psi_a(t),t)$, which is along the ray $\frac{u}{t}\psi_a(t)$. If the boundary $\psi_a(u)$ for $u < t$ is high relative to the ray $\frac{u}{t}\psi_a(t)$, then one might expect that (22) holds. Here height is measured in units of the standard deviation of the Brownian bridge W_0. In fact by condition (i') and (ii),

$$(24) \qquad \inf_{\varepsilon t<u<(1-\varepsilon)t} \sqrt{\frac{t}{u(t-u)}} (\psi_a(u) - \frac{u}{t}\psi_a(t)) \to \infty$$

since $\psi_a(u) > (\frac{u}{t})^\alpha \psi_a(t)$ by condition (ii).

We now consider whether the TA for lower class boundaries at infinity like (17) holds uniformly on $\mathbb{R}_+$. Formula (21) shows that this cannot be true since the integral on its r.h.s. tends to infinity as $t_1 \to \infty$. Nevertheless also for (17) there exists a uniform approximation of the densities on $\mathbb{R}_+$.

THEOREM 4. Assume that

(i') $\inf_{0<t<\infty} \psi_a(t)/\sqrt{t} \to \infty$ as $a \to \infty$,

(ii') there exists a constant $1/2 < \alpha < 1$ such that $\psi_a(t)/t^\alpha$ is decreasing,

(iii') for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all a $|\psi_a'(s)/\psi_a'(t) - 1| < \varepsilon$ if $|s/t - 1| < \delta$,

(iv') there exists a $\gamma > 0$ such that $P(T_a < \gamma) \to 0$ as $a \to \infty$.

Then

$$p_a(t) = P(T_a > t) \frac{\Lambda_a(t)}{t^{3/2}} \phi\left(\frac{\psi_a(t)}{\sqrt{t}}\right) (1 + o(1)) \text{ uniformly on } \mathbb{R}_+ \text{ as } a \to \infty.$$

Theorem 4 leads to a characterization of the uniform TA.

COROLLARY 2. Let the assumptions (i') - (iv') hold. Let $\{h_a;\ a \in \mathbb{R}_+\}$ denote a function with $\lim_{a\to\infty} h_a = \infty$. Then the TA uniformly holds on $(0,h_a)$ as $a \to \infty$ if and only if $P(T_a < h_a) \to 0$.

For the boundary (17), $P(T_a < h_a) \to 0$ holds for $h_a = \exp(e^a a^{-\alpha})$ with $\alpha > 1/2$.

There is a related result for a different type of asymptotic, where one lets $t \to \infty$ for a fixed boundary ψ. For this case one naturally assumes that $P(T > t) > 0$ for all t.

Under the condition that

$$\psi(t)/\sqrt{t} \to \infty \quad \text{as } t \to \infty \tag{25}$$

and some further assumptions,

$$p(t) = P(T > t) \frac{\Lambda(t)}{t^{3/2}} \phi \left(\frac{\psi(t)}{\sqrt{t}}\right) (1 + o(1)) \tag{26}$$

holds as $t \to \infty$.

Statement (26) leads to a new proof of the Kolmogorov-Petrovski-Erdös test by noting that $(\log P(T > t))' = \frac{p(t)}{P(T>t)}$. For a lower-class boundary which satisfies (25), statement (26) also yields the tail probabilities

$$P(T > t) = \exp(-\int_0^t \frac{\Lambda(s)}{s^{3/2}} \phi(\frac{\psi(s)}{\sqrt{s}}) ds(1 + o(1)) \text{ as } t \to \infty.$$

For more details see Lerche (1984).

REFERENCES

Cuzick, J. (1981). Boundary crossing probabilities for stationary Gaussian processes and Brownian motion. Trans. Amer. Math. Soc. **263** 469-492.

Daniels, H. (1974). The maximum size of a closed epidemic. Adv. Appl. Prob. **6** 607-621.

Darling, D. and Robbins, H. (1967). Iterated logarithm inequalities. Proc. Nat. Acad. Sci. U.S.A. **57** 1188-1192.

Darling, D. and Robbins, H. (1968). Some further remarks on inequalities for partial sums. Proc. Nat. Acad. Sci. U.S.A. **60** 1175-1182.

Ito, K. and McKean, H.P. (1974). Diffusion Processes, Grundlehren der matematischen Wissenschaften, 125. Springer-Verlag, Berlin.

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

Jennen, C. (1985). Second-order approximations to the density, mean and variance of Browian first-exit times. Ann. Prob. **13** 126-144.

Lai, T.L. and Siegmund, D. (1977). A nonlinear renewal theory with applications to sequential analysis I, Ann. Statist. 5 946-954.

Lerche, H.R. (1984). Statistical aspects of the law of the iterated logarithm, Habilitationsschrift at the University of Heidelberg.

Lerche, H.R. (1985). On the optimality of sequential tests with parabolic boundaries. Proc. Conf. in Honor of J. Neyman and J. Kiefer, (ed. L. LeCam, R. Olshen), Wadsworth Inc., Vol. II, 579-597.

Lerche, H.R. (1986a). The shape of Bayes tests of power one. To appear in Ann. Statist.

Lerche, H.R. (1986b). An optimal property of the repeated significance test. To appear in Proc. Nat. Acad. Sci. U.S.A.

Lorden, G. (1973). Open-ended tests for Koopman-Darmois families. Ann. Statist. **1** 633-643.

Robbins, H. (1952). Some aspects of the sequential design of experiments. Bull. Amer. Math. Soc. **58** 527-535.

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

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

Robbins, H. and Siegmund, D. (1973). Statistical tests of power one and the integral representation of solutions of certain partial differential equations. Bull. Acad. Sinica **1** 93-120.

Siegmund, D. (1977). Repeated significance tests for a normal mean. Biometrika **64** 177-189.

Strassen, V. (1967). Almost sure behaviour of sums of independent random variables and martingales. Proc. Fifth Berkeley Symp. Math. Statist. Prob., Univ. of Cal. Press, Vol. III, Part I, 315-343.