The Annals of Probability
1988, Vol. 16, No. 2, 487–501

APPROXIMATE TAIL PROBABILITIES FOR THE MAXIMA OF SOME RANDOM FIELDS[1]

By David Siegmund

Stanford University

For random walks $\{S_n\}$ whose distribution can be embedded in an exponential family, large-deviation approximations are obtained for the probability that $\max_{0 \le i < j \le m}(S_j - S_i) \ge b$ (i) conditionally given S_m and (ii) unconditionally. The method used in the conditional case seems applicable to maxima of a reasonably large class of random fields. For the unconditional probability a more special argument is used, and more precise results obtained.

1. Introduction. Hogan and Siegmund (1986) adapt the method developed by Pickands (1969), Qualls and Watanabe (1973) and Bickel and Rosenblatt (1973) to obtain explicit large-deviation approximations for the maxima of several Gaussian random fields arising in statistics. Using a special argument for one particular case, they suggest a heuristic second-order approximation for that case; and they show by a Monte Carlo experiment that the second-order approximation frequently gives considerably better numerical results.

The purpose of this paper is to show that the method developed by Woodroofe (1976, 1982) for problems in one-dimensional time can be adapted to study maxima of random fields. Overall, it involves simpler computations than the previous method and consequently seems potentially capable of delivering a genuine second-order approximation should one seem desirable. See Woodroofe and Takahashi (1982) for an example in one-dimensional time.

Let $x_1, x_2, \ldots$ be independent, identically distributed random variables, and put $S_n = x_1 + \cdots + x_n$, $n = 0, 1, \ldots$. For $b > 0$ define

$$t = t(b) = \inf\left\{n: \max_{0 \le k \le n} (S_n - S_k) \ge b\right\}.$$

Theorem 1 below gives a large-deviation approximation to the conditional probability

$$P\{t \le m | S_m = \xi\}, \qquad \xi < b, \tag{1}$$

when the distribution of the x's can be embedded in an exponential family. Although we discuss only this concrete case, it will be apparent that the method is reasonably general. See Hogan and Siegmund (1986) and Section 4 for additional examples. For applications of (1) see Levin and Kline (1985) and Adler and Brown (1986).

Received August 1986; revised December 1986.

[1]Research supported by Office of Naval Research Contract N0014-77-C-0306 (NR-042-373) and by National Science Foundation Grant DMS-80-00235.

AMS 1980 *subject classifications*. Primary 60F10, 60G60; secondary 60K05, 62N10.

Key words and phrases. Large deviations, random field, CUSUM test.

The unconditional probability

$$P\{t \le m\} \tag{2}$$

gives the distribution of the run length of a CUSUM test [e.g., van Dobben de Bruyn (1968)] and the probability that at least one among the first m customers in a single server queue has a waiting time exceeding b. Several recent papers have discussed its numerical evaluation [e.g., Woodhall (1983) and Waldmann (1986)].

Although the method of Theorem 1 can also be applied to the unconditional probability (2), one can use a special, considerably simpler argument and obtain a more precise approximation, which in this case provides justification for the Hogan–Siegmund heuristic and indicates what one can expect to gain from second-order approximations in related problems. This line of reasoning is developed in Section 3.

Section 4 contains additional examples and miscellaneous remarks.

2. Conditional probability. In order to facilitate comparisons between Theorem 1 below and the related Theorem 8.72 of Siegmund (1985), we use the notation of that result, which was proved by a method that does not seem to adapt to multidimensional indexing sets. Here we modify a method developed by Woodroofe (1976, 1982) in one-dimensional time.

The following discussion omits some technical details which occur even in the one-dimensional case and concentrates on issues which only arise because of the multidimensional indexing set.

Let P_μ denote the probability which makes $x_1, x_2, \ldots$ independent, identically distributed random variables with probability distribution of the form

$$P_\mu\{x_n \in dx\} = \exp[\theta x - \psi(\theta)]\, dF(x)$$

relative to some fixed probability distribution F, which without loss of generality, is assumed to have mean 0. The parameters μ and θ have the one-to-one relation $\mu = \psi'(\theta)$ $(= E_\mu x_1)$. In Section 3 it is notationally convenient to standardize F to have unit variance.

Let $S_n = x_1 + \cdots + x_n$, $n = 0, 1, \ldots$.

It is convenient to assume that for all μ there exists an n_0 such that

$$\int_{-\infty}^{\infty} |E_\mu \exp(i\lambda x_1)|^{n_0}\, d\lambda < \infty.$$

This implies that for all $n \ge n_0$ the P_μ distribution of S_n has a continuous, bounded density function, $f_{\mu, n}$, and as $n \to \infty$,

$$f_{\mu, n}(\sigma n^{1/2} y + n\mu)\sigma n^{1/2} \to \varphi(y), \tag{3}$$

uniformly in y, where $\sigma^2 = \psi''(\theta)$ and φ denotes the standard normal density function [cf. Feller (1972), page 516]. Also assume that x_1 has a density function. An alternative technical condition would be that F is an arithmetic distribution; and by using the technique of Lalley (1984), one can perhaps eliminate all such assumptions.

Let

$$P_\xi^{(m)}(A) = P_\mu(A|S_m = \xi),$$

for events A defined in terms of $x_1, \dots, x_m$. By sufficiency the conditional probability does not depend on μ. Also let

$$\tau_+(\tau_-) = \inf\{n: S_n > (\le)0\}.$$

THEOREM 1. *Let $b = m\zeta$ and $\xi = m\xi_0$ for arbitrary fixed $\zeta > 0$ and $\xi_0 < \zeta$. Assume there exist $\mu_2 < 0 < \mu_1$ (necessarily unique) such that*

$$1 = \mu_1^{-1}\zeta + |\mu_2|^{-1}(\zeta - \xi_0) \tag{4}$$

and

$$\psi[\theta(\mu_2)] = \psi[\theta(\mu_1)], \tag{5}$$

where $\theta(\mu)$ denotes the inverse of the function $\mu = \psi'(\theta)$. Let $\theta_i = \theta(\mu_i)$, $i = 1,2$, $\theta_0 = \theta(\xi_0)$, and $\sigma_i^2 = \psi''(\theta_i)$, $i = 0,1,2$. Then as $m \to \infty$,

$$P_\xi^{(m)}\{t \le m\} \sim mC(\zeta, \xi_0)\exp[-mB(\zeta, \xi_0)], \tag{6}$$

where

$$B(\zeta, \xi_0) = (\theta_1 - \theta_2)\zeta - \psi(\theta_2) + \psi(\theta_0) + (\theta_2 - \theta_0)\psi'(\theta_0) \tag{7}$$

and

$$C(\zeta, \xi_0) = \frac{P_{\mu_2}^2\{\tau_+ = \infty\}P_{\mu_1}^2\{\tau_- = \infty\}}{(\theta_1 - \theta_2)\mu_1|\mu_2|} \frac{\sigma_0(\zeta - \xi_0)|\mu_2|^{-1}}{\left(\sigma_1^2\zeta/\mu_1^3 + \sigma_2^2(\zeta - \xi_0)/|\mu_2|^3\right)^{1/2}}. \tag{8}$$

REMARKS. (i) To evaluate $C(\zeta, \xi_0)$ it is usually adequate to use the approximation

$$\begin{aligned} &P_{\mu_2}\{\tau_+ = \infty\}P_{\mu_1}\{\tau_- = \infty\}/[(\theta_1 - \theta_2)\mu_1] \\ &\quad = \exp[-(\theta_1 - \theta_2)\rho_+] + o\left[(\theta_1 - \theta_2)^2\right], \end{aligned} \tag{9}$$

as $\theta_1 - \theta_2 \to 0$, where $\rho_+ = E_0 S_{\tau_+}^2/(2E_0 S_{\tau_+})$ can be calculated numerically [cf. Siegmund (1985), Proposition 10.37 and Theorem 10.55]. A similar approximation holds for $P_{\mu_2}\{\tau_+ = \infty\}P_{\mu_1}\{\tau_- = \infty\}/[(\theta_1 - \theta_2)|\mu_2|]$ in terms of $\rho_- = E_0(S_{\tau_-}^2)/(2E_0 S_{\tau_-})$. (ii) The exponent B defined in (7) also appears in the considerably simpler one-dimensional problem involving $P_\xi^{(m)}\{\max_{0<n<m} S_n \ge b\}$. See Siegmund (1982), Theorem 1, or (1985), Theorem 8.72, for that result and for an explanation of the meaning of (4). In the special case $dF(x) = \varphi(x)\,dx$ it is easy to see that $\mu_1 = -\mu_2 = 2\zeta - \xi_0$ and $B = 2\zeta(\zeta - \xi_0)$, so Theorem 2 of Hogan and Siegmund (1986) is a special case.

PROOF OF THEOREM 1. Let

$$D = \left\{(i_0, j_0): i_0 \ge m^{1/2},\ j_0 \le m - m^{1/2},\ |j_0 - i_0 - m\zeta/\mu_1| < m^{7/12}\right\}.$$

For each j_0 let $D(j_0) = \{i_0: (i_0, j_0) \in D\}$. Also for each $0 \le i_0 < j_0 \le m$ let

$$J = J(i_0, j_0) = \left\{(i, j): 0 \le i < j \le m,\ j < j_0 \text{ or } j = j_0 \text{ and } i < i_0\right\} \cap D.$$

Then for arbitrary $x_0 > 0$,

$$
\begin{aligned}
&P_\xi^{(m)}\{t \le m\} \\
&\quad \le \sum_{(i_0, j_0)\in D} \int_0^{x_0} P_\xi^{(m)}\{S_{j_0} - S_{i_0} \in b + dx\} \\
&\qquad\qquad \times P_\xi^{(m)}\{S_j - S_i < b, \forall (i, j) \in J | S_{j_0} - S_{i_0} = b + x\} \\
&\qquad + \sum_{j_0}\left[P_\xi^{(m)}\left\{\max_{\substack{i_0 \notin D(j_0)\\ i_0 < j_0}} (S_{j_0} - S_{i_0}) \ge b\right\}\right. \\
&\qquad\qquad \left. + P_\xi^{(m)}\left\{\max_{i_0 \in D(j_0)} (S_{j_0} - S_{i_0}) \ge b + x_0\right\}\right];
\end{aligned} \tag{10}
$$

and the first sum on the right-hand side of (10) is a lower bound for $P_\xi^{(m)}\{t \le m\}$. From Lemmas 1–5 below one obtains (6), but with a constant C of the form

$$C = C'[\text{second fraction in (8)}]/\mu_1|\mu_2|. \tag{11}$$

In (11)

$$C' = \int_0^\infty e^{-(\theta_1-\theta_2)x} P_{\mu_2}\left\{\max_{n\ge 1} S_n \le -x\right\} P_{\mu_1}\left\{\min_{n \ge 0} S_n + \min_{n\ge 1} S_n' \ge x\right\} dx, \tag{12}$$

where $\{S_n', n = 0, 1, \dots\}$ is an independent copy of the random walk $\{S_n, n = 0, 1, \dots\}$. The proof of Theorem 1 is completed by the evaluation of C' given below in Lemma 7.

Lemma 1 is an easy, well-known consequence of (3) and the relation

$$f_{0,n}(y) = \exp[-\hat{\theta} y + n\psi(\hat{\theta})] f_{\hat{\mu}, n}(y),$$

where $\hat{\mu} = \psi'(\hat{\theta}) = y/n$. [Actually (3) must be strengthened slightly to provide for the desired uniformity.]

LEMMA 1. *Let* $\mu = \psi'(\theta)$, $\sigma^2 = \psi''(\theta)$, *and assume* $|x| \le \delta_n n^{2/3}$ *for some sequence* $\delta_n \to 0$. *Then uniformly in* x *as* $n \to \infty$,

$$f_{0,n}(n\mu + x) \sim \exp[-\theta(n\mu + x) + n\psi(\theta)] \varphi(x/\sigma n^{1/2})/\sigma n^{1/2}.$$

Lemma 2 follows from Lemma 1 and the identity

$$P_\xi^{(m)}\{S_n \in dy\} = f_{0,n}(y) f_{0,m-n}(\xi - y)\, dy / f_{0,m}(\xi).$$

LEMMA 2. *Suppose* $(i_0, j_0) \in D$ *and put* $n = j_0 - i_0$. *Then uniformly in* (i_0, j_0) *and* $|x| \le m^{1/12}$

$$
\begin{aligned}
P_\xi^{(m)}\{S_{j_0} - S_{i_0} \in m\zeta + dx\} &\sim \exp[-mB(\zeta, \xi_0) - (\theta_1 - \theta_2)x] \\
&\quad \times \frac{(2\pi m)^{1/2}\sigma_0}{\sigma_1\sigma_2 n^{1/2}(m-n)^{1/2}} \varphi\left[\frac{|\mu_2|(n - m\zeta/\mu_1)}{\sigma_2(m-n)^{1/2}}\right] \\
&\quad \times \varphi\left[\frac{\mu_1(n - m\zeta/\mu_1)}{\sigma_1 n^{1/2}}\right] dx.
\end{aligned} \tag{13}
$$

Lemma 3. *As $m \to \infty$,*

$$\sum_{j_0} P_\xi^{(m)}\Big\{\max_{\substack{i_0 \notin D(j_0)\\ i_0 < j_0}} (S_{j_0} - S_{i_0}) \ge b\Big\} = o\big[m \exp\{-mB(\zeta, \xi_0)\}\big]$$

and

$$\sum_{j_0} P_\xi^{(m)}\Big\{\max_{i_0 \in D(j_0)} (S_{j_0} - S_{i_0}) \ge b + x_0\Big\} \le m\delta(x_0)\exp\{-mB(\zeta, \xi_0)\},$$

where $\delta(x_0) \to 0$ as $x_0 \to \infty$.

Proof. Since the proofs of both assertions are similar, we consider the second one. The sum contains $m - 1$ terms, each of which is bounded by

$$P_\xi^{(m)}\Big\{\max_{0<n<m} S_n \ge b + x_0\Big\}.$$

A careful reading of the proof of Siegmund (1982), Theorem 1, or (1985), Theorem 8.72, shows that this probability

$$\sim K(\zeta, \xi_0)\exp\{-mB(\zeta, \xi_0) - (\theta_1 - \theta_2)x_0\}$$

provided $x_0 = o(m^{1/2})$, where $K(\zeta, \xi_0)$ is a constant similar to C defined in (8) which does not depend on m or on x_0. This more than proves the second assertion, and the first is proved similarly. □

Although the following lemma is not difficult to prove, its importance cannot be over emphasized. It shows that the two-dimensional random field under consideration here behaves locally like a superposition of independent one-dimensional random fields, and thus makes possible the explicit evaluation of C.

Lemma 4. *Suppose $i_0 \ge m^{1/2}$, $j_0 \le m - m^{1/2}$ and $j_0 - i_0 \sim m\zeta/\mu_1$. Then uniformly in (i_0, j_0) and x in compact subsets of $[0, \infty)$, as $m \to \infty$,*

$$\begin{aligned} &P_\xi^{(m)}\{S_j - S_i < b, \forall\, (i, j) \in J \mid S_{j_0} - S_{i_0} = b + x\} \\ &\quad \to P_{\mu_2}\Big\{\max_{n \ge 1} S_n \le -x\Big\} P_{\mu_1}\Big\{\min_{n \ge 0} S_n + \min_{n \ge 1} S'_n \ge x\Big\}, \end{aligned} \tag{14}$$

where $\{S'_n,\ n = 0, 1, \dots\}$ is an independent copy of $\{S_n,\ n = 0, 1, \dots\}$.

Proof. Given $S_{j_0} - S_{i_0} = m\zeta + x$, the event on the left-hand side of (14) equals

$$\Big\{S_{j_0} - S_{j_0 - j} + \big(S_{i_0 + i} - S_{i_0}\big) > x, \forall\, (i, j)\colon (i_0 + i, j_0 - j) \in J\Big\}.$$

It is easy to see that in the quantification $\forall\, (i, j)$ such that $(i_0 + i, j_0 - j) \in J$ the indices (i, j) with $i \le -1$ and $j \ge 1$ are redundant, because the required inequality holds for these (i, j) if it holds for $i \le -1$, $j = 0$ and for $i = 0$, $j \ge 1$. Hence the event above equals

$$\begin{aligned} \Big\{&S_{i_0 + i} - S_{i_0} > x, \forall\, i \le -1\colon (i_0 + i, j_0) \in J; \\ &S_{j_0} - S_{j_0 - j} + S_{i_0 + i} - S_{i_0} > x, \forall\, i \ge 0,\ j \ge 1\colon (i_0 + i, j_0 - j) \in J\Big\}, \end{aligned}$$

which is contained in or contains

$$\left\{ \min_{-n_1 < i \le -1} \left(S_{i_0+i} - S_{i_0}\right) > x, \min_{1 \le j < n_2} \left(S_{j_0} - S_{j_0-j}\right) + \min_{0 \le i < n_2} \left(S_{i_0+i} - S_{i_0}\right) > x \right\},$$

according as $n_1 = n_2$ is arbitrary, but fixed, or $n_1 = i_0$, $n_2 = j_0 - i_0$.

It is easy to see that for arbitrary $n = 1, 2, \ldots,$ given $S_m = m\xi_0$ and $S_{j_0} - S_{i_0} = m\zeta + x$, as $m \to \infty$ the joint distribution of

$$S_{j_0} - S_{j_0-j}, \qquad j = 0, 1, \ldots, n,$$

converges to the P_{μ_1} joint distribution of S_j, $j = 0, 1, \ldots, n$; the joint distribution of

$$S_{i_0+i} - S_{i_0}, \qquad i = 0, 1, \ldots, n,$$

converges to the P_{μ_1} joint distribution of S_i, $i = 0, 1, \ldots, n$; the joint distribution of

$$S_{i_0+i} - S_{i_0}, \qquad i = 0, -1, \ldots, -n,$$

converges to the P_{μ_2} joint distribution of $-S_i$, $i = 0, 1, \ldots, n$; and asymptotically these three collections of random variables are stochastically independent.

The proof is completed by letting $m \to \infty$ with n held fixed and the three minima restricted to indices with $|i| < n$ and $j < n$, then letting $n \to \infty$ and showing that the indices $|i| \ge n$ or $j \ge n$ do not contribute in the limit. The details of this final step are similar to the one-dimensional case and are omitted. □

LEMMA 5. *As* $m \to \infty$,

$$\sum_D \int_0^\infty P_\xi^{(m)}\{S_{j_0} - S_{i_0} \in b + dx\} P_\xi^{(m)}\{S_j - S_i < b, \forall\,(i, j) \in J | S_{j_0} - S_{i_0} = b + x\}$$

$\sim$ *right-hand side of* (6),

with C *as given in* (11) *and* (12).

PROOF. To sum the approximations provided by Lemmas 2–4 over D, observe that by (4) there are asymptotically $m(1 - \zeta/\mu_1) = m(\zeta - \xi_0)/|\mu_2|$ terms i_0, and for each i_0 the sum over $j_0 = i_0 + n$ of [from (13)]

$$\frac{(2\pi m)^{1/2}\sigma_0}{\sigma_1\sigma_2 n^{1/2}(m-n)^{1/2}} \varphi\left[\frac{|\mu_2|(n - m\zeta/\mu_1)}{\sigma_2(m-n)^{1/2}}\right] \varphi\left[\frac{\mu_1(n - m\zeta/\mu_1)}{\sigma_1 n^{1/2}}\right]$$

converges to

$$\sigma_0(\mu_1|\mu_2|)^{-1}\{\sigma_1^2\zeta/\mu_1^3 + \sigma_2^2(\zeta - \xi_0)/|\mu_2|^3\}^{-1/2}. \qquad \square$$

To complete the proof of the theorem we must evaluate the constant C' defined in (12). This part of the argument seems substantially more difficult than the analogous result of Hogan and Siegmund (1986), Lemma 3.4. The following lemma is well known [e.g., Woodroofe (1982), page 26].

LEMMA 6. *For* $x > 0$,

$$P_{\mu_1}\Big\{\min_{n\geq 1} S_n \geq x\Big\} = \mu_1 P_{\mu_1}\{S_{\tau_+} \geq x\}/E_{\mu_1}(S_{\tau_+})$$

and

$$P_{\mu_2}\Big\{\max_{n\geq 1} S_n \leq -x\Big\} = \mu_2 P_{\mu_2}\{S_{\tau_-} \leq -x\}/E_{\mu_2} S_{\tau_-}.$$

LEMMA 7. *Let* $\{S_n', n = 0, 1, \dots\}$ *be an independent copy of* $\{S_n, n = 0, 1, \dots\}$. *Then*

$$\int_0^\infty \exp[-(\theta_1 - \theta_2)x] P_{\mu_2}\Big\{\max_{n\geq 1} S_n \leq -x\Big\} P_{\mu_1}\Big\{\min_{n\geq 0} S_n + \min_{n\geq 1} S_n' \geq x\Big\}\, dx \tag{15}$$
$$= (\theta_1 - \theta_2)^{-1} P_{\mu_2}^2\{\tau_+ = \infty\} P_{\mu_1}^2\{\tau_- = \infty\}.$$

PROOF. For $y \leq 0$ let $\tau(y) = \inf\{n: S_n \leq y\}$, and observe that by Wald's likelihood ratio identity and (5)

$$P_{\mu_1}\{\tau(y) < \infty\} = E_{\mu_2} \exp\big[(\theta_1 - \theta_2) S_{\tau(y)}\big]. \tag{16}$$

Also

$$P_{\mu_1}\Big\{\min_{n\geq 0} S_n \leq y\Big\} = P_{\mu_1}\{\tau(y) < \infty\}, \qquad y \leq 0. \tag{17}$$

If one writes the convolution appearing in the integrand in (15) as an integral, and uses Lemma 6, (16) and (17), after some manipulation one obtains

$$\int_0^\infty \exp[-(\theta_1 - \theta_2)x] P_{\mu_2}\Big\{\max_{n\geq 1} S_n \leq -x\Big\} P_{\mu_1}\Big\{\min_{n\geq 0} S_n + \min_{n\geq 1} S_n' \geq x\Big\}\, dx$$
$$= \frac{\mu_1|\mu_2|}{E_{\mu_1} S_{\tau_+} |E_{\mu_2} S_{\tau_-}|} \int_0^\infty P_{\mu_2}\{S_{\tau_-} \leq -x\}\, e^{-(\theta_1 - \theta_2)x}$$
$$\times \int_x^\infty P_{\mu_1}\{S_{\tau_+} \in d\eta\}\big[1 - E_{\mu_2} \exp\{(\theta_1 - \theta_2) S_{\tau(x-\eta)}\}\big]\, dx$$
$$= \frac{\mu_1|\mu_2|}{E_{\mu_1} S_{\tau_+}} \int_0^\infty P_{\mu_1}\{S_{\tau_+} \in d\eta\}\, e^{-(\theta_1 - \theta_2)\eta}$$
$$\times \int_0^\eta \frac{P_{\mu_2}\{S_{\tau_-} \leq -x\}}{|E_{\mu_2} S_{\tau_-}|} \big\{\exp[(\theta_1 - \theta_2)(\eta - x)]$$
$$- E_{\mu_2} \exp\big[(\theta_1 - \theta_2)\big(S_{\tau(x-\eta)} + \eta - x\big)\big]\big\}\, dx.$$

Consider the inner integral to be of the form [in the notation of Feller (1971), Chapter XI] $F_0 * Z(\eta)$, where $dF_0(x) = P_{\mu_2}\{|S_{\tau_-}| \leq x\}\, dx/|E_{\mu_2} S_{\tau_-}|$ is the stationary distribution for the renewal process determined by $F(x) = P_{\mu_2}\{|S_{\tau_-}| \leq x\}$ and Z is a solution of the renewal equation $Z = z + F * Z$, or equivalently in terms of the renewal measure U, $Z = U * z$. It is known that Z determines z uniquely; and since F_0 is the stationary distribution, $F_0 * U$ is proportional to

Lebesgue measure and

$$F_0 * Z(\eta) = (F_0 * U) * z(\eta) = \int_0^\eta z(x)\,dx / |E_{\mu_2} S_{\tau_-}|.$$

It is easy to see in the present case that

$$z(x) = \exp[(\theta_1 - \theta_2)x]\left[1 - E_{\mu_2}\exp\{(\theta_1 - \theta_2)S_{\tau_-}\}\right],$$

and hence the left-hand side of (15) equals

$$\frac{\mu_1\mu_2\left[1 - E_{\mu_2}\exp\{(\theta_1 - \theta_2)S_{\tau_-}\}\right]}{E_{\mu_1}S_{\tau_+}|E_{\mu_2}S_{\tau_-}|}\int_0^\infty P_{\mu_1}\{S_{\tau_+} \in d\eta\}\, e^{-(\theta_1-\theta_2)\eta}\int_0^\eta e^{(\theta_1-\theta_2)x}\,dx$$

$$= \frac{\mu_1|\mu_2|\left[1 - E_{\mu_2}\exp\{(\theta_1 - \theta_2)S_{\tau_-}\}\right]\left[1 - E_{\mu_1}\exp\{-(\theta_1 - \theta_2)S_{\tau_+}\}\right]}{(\theta_1 - \theta_2)E_{\mu_1}S_{\tau_+}|E_{\mu_2}S_{\tau_-}|}$$

$$= P_{\mu_1}^2\{\tau_- = \infty\}P_{\mu_2}^2\{\tau_+ = \infty\}/(\theta_1 - \theta_2),$$

where the last equality is a consequence of Wald's identities and the relations $P_{\mu_1}\{\tau_- = \infty\}E_{\mu_1}(\tau_+) = 1$, $P_{\mu_2}\{\tau_+ = \infty\}E_{\mu_2}(\tau_-) = 1$. □

3. Unconditional probability. We continue to use the notation of Section 2 and consider the unconditional probability (2) with $P = P_\mu$ for some $\mu < 0$. In principle, one can obtain a first-order large-deviation approximation for $P_\mu\{t \le m\}$ by integrating the approximation of Theorem 1 with respect to the distribution of S_m. Here we consider a different approach and obtain a second-order approximation to (2). This method was mentioned briefly in Siegmund (1986), and for a related continuous-time problem it was used by Hogan and Siegmund (1986).

The proofs of the following results are for the most part modifications of arguments given in Siegmund (1975, 1979)—see also Siegmund (1985), Chapters VIII and X. Consequently, the steps of the argument are given in a sequence of lemmas, but most details are omitted.

THEOREM 2. *Assume that the P_μ distributions of x_1 are strongly nonarithmetic in the sense that*

$$\limsup_{|\lambda|\to\infty} |E_\mu \exp(i\lambda x_1)| < 1.$$

Let $\mu_2 < 0$ and assume there exists $\mu_1 > 0$ such that (5) *holds. Let $\Delta = \theta_1 - \theta_2$, where $\theta_i = \theta(\mu_i)$, $i = 1, 2$. Let $b > 0$ and assume that for some $\delta > 0$,*

$$(18)\qquad m\mu_1/b \ge 1 + \delta, \qquad m^2\exp(-\Delta b) \to 0.$$

Then as $m, b \to \infty$,

$$(19)\qquad P_{\mu_2}\{t \le m\} = \Delta|\mu_2|\nu_+\nu_-\exp(-\Delta b)\{m - b/\mu_1 + D + o(1)\},$$

where

$$(20)\qquad \nu_+ = P_{\mu_2}\{\tau_+ = \infty\}P_{\mu_1}\{\tau_- = \infty\}/(\mu_1\Delta), \qquad \nu_- = \mu_1\nu_+/|\mu_2|,$$

and

$$D = 1 - \mu_1^{-1}E_{\mu_1}S_{\tau_+}^2/(2E_{\mu_1}S_{\tau_+}) - \mu_1^{-1}E_{\mu_1}(S_{\tau_-};\ \tau_- < \infty)E_{\mu_1}(\tau_+)$$

$$+E_{\mu_1}(\tau_-;\ \tau_- < \infty)E_{\mu_1}(\tau_+) + E_{\mu_2}(\tau_+;\ \tau_+ < \infty)E_{\mu_2}(\tau_-) \tag{21}$$

$$-(\mu_1\nu_+)^{-1}\int_0^\infty \left\{E_{\mu_1}\exp[-\Delta(S_{\tau(x)} - x)] - \nu_+\right\}P_{\mu_1}\left\{\min_{n\geq 0} S_n > -x\right\}dx,$$

with $\tau(x) = \inf\{n\colon S_n \geq x\}$.

REMARK. The constant $\Delta\nu_+\nu_-$ in (19) and (20) is the first factor in the constant C of (6) and (8). The local expansions of remark (i) following Theorem 1 applied to ν_+ and ν_- together with the similar local expansion of D given in Theorem 3 below lead to the much simpler approximation

$$P_{\mu_2}\{t \leq m\} \cong \exp[-\Delta(b + \rho_+ - \rho_-)] \times \left\{\Delta|\mu_2|[m - \mu_1^{-1}(b + \rho_+ - \rho_-)] + 3 - 7\gamma\Delta/6\right\}, \tag{22}$$

where $\gamma = E_0 x_1^3$ and the P_0 distribution of x_1 has been standardized to have unit variance. In the case of the unconditional probability (2), (22) gives precise meaning to the Hogan–Siegmund heuristic approximation (which applies only when $\gamma = 0$).

THEOREM 3. *Assume that the P_0 distribution of x_1 is standardized to have unit variance. For D given by* (21), *as* $\Delta \to 0$,

$$\Delta|\mu_2|D = 3 - \Delta|\mu_2|(\rho_+ - \rho_-)/\mu_1 - 7\Delta\gamma/6 + o(\Delta),$$

where $\rho_\pm = \frac{1}{2}E_0(S_{\tau_\pm}^2)/E_0(S_{\tau_\pm})$ *and* $\gamma = E_0 x_1^3$.

Table 1 contains some values of $p = P_\mu\{t \leq m\}$ computed numerically by Waldmann (1986). For comparison it gives first-order ($\hat{p}_1$) and second-order ($\hat{p}_2$) approximations from (22). The x's are normally distributed with mean $\mu_2 = -0.5$ and $b = 3$. The value of $\rho_+ = |\rho_-| \cong 0.583$ [e.g., Siegmund (1985), page 225]. The approximation $\hat{p}_2$ is quite good, but $\hat{p}_1$ is rather poor. However, the values of b and m are quite small. Hogan and Siegmund (1986) compared similar approximations for (1) in a Monte Carlo experiment involving generally larger sample sizes and found that the first-order approximation was reasonably good when $\xi < 0$

TABLE 1

Approximations to $p = P_\mu\{t \leq m\}$

m	p	$\hat{p}_1$	$\hat{p}_2$
9	0.054	0.023	0.052
12	0.079	0.047	0.076
15	0.102	0.070	0.098
18	0.126	0.093	0.122

but not when $\xi \geq 0$. Their heuristic second-order approximation was good in all cases. Some Monte Carlo experimentation shows that the approximation $\hat{p}_2$ begins to deteriorate when p is about 0.2, and as expected is poor for large values of p.

Theorem 2 is a consequence of Lemmas 8–10.

Let $T = \inf\{n: S_n \notin (0, b)\}$.

LEMMA 8. *For arbitrary* $b > 0$, $\mu < 0$,

$$P_\mu\{t \leq m\} \leq P_\mu\{\tau_+ = \infty\} E_\mu\{(m - T + 1); T < m, S_T \geq b\} + \sum_{n=0}^{m-1} P_\mu\{n < \tau_+ < \infty\} P_\mu\{T \leq m - n, S_T \geq b\}. \tag{23}$$

A lower bound for $P_\mu\{t \leq m\}$ *is given by the right-hand side of* (23) *divided by* $1 + E_\mu\{(m - T + 1); T < m, S_T \geq b\}$.

PROOF. See Siegmund (1986), Proposition 3.23. □

LEMMA 9. *Under the condition* (18)

$$P_{\mu_2}\{t \leq m\} = P_{\mu_2}\{\tau_+ = \infty\} E_{\mu_2}(m - T + 1; T < m, S_T \geq b) + E_{\mu_2}(\tau_+; \tau_+ < \infty) P_{\mu_2}\{S_T \geq b\} + o(e^{-\Delta b}).$$

PROOF. Lemma 9 follows from Lemma 8, the inequality

$$E_{\mu_2}\{m - T + 1; T < m, S_T \geq b\} \leq m P_{\mu_2}\{S_T \geq b\} \leq m e^{-\Delta b},$$

and a related elementary inequality. □

LEMMA 10. *Under the conditions of Theorem* 2

$$P_{\mu_2}\{T \leq m, S_T \geq b\} = P_{\mu_2}\{S_T \geq b\}(1 + o(m^{-1})) = \nu_+ P_{\mu_1}\{\tau_- = \infty\} \exp(-\Delta b)(1 + o(m^{-1}))$$

and

$$\begin{aligned} E_{\mu_2}(T; T < m, S_T \geq b) &= \nu_+ \exp(-\Delta b)\Big\{P_{\mu_1}\{\tau_- = \infty\}[b + \rho_+(\mu_1)]/\mu_1 \\ &\quad + \mu_1^{-1} E_{\mu_1}(S_{\tau_-}; \tau_- < \infty) - E_{\mu_1}(\tau_-; \tau_- < \infty) \\ &\quad + P_{\mu_1}\{\tau_- = \infty\}(\nu_+ \mu_1)^{-1} \int_0^\infty [E_{\mu_1} \exp\{-\Delta(S_{\tau(x)} - x)\} - \nu_+] \\ &\quad \times P_{\mu_1}\Big\{\min_{n \geq 0} S_n > -x\Big\} dx + o(1)\Big\}, \end{aligned}$$

where $\rho_+(\mu_1) = \frac{1}{2} E_{\mu_1}(S_{\tau_+}^2)/E_{\mu_1}(S_{\tau_+})$.

Theorem 2 follows easily by substitution of the approximations of Lemma 10 into Lemma 9 and use of the well-known relation $P_{\mu_1}\{\tau_- = \infty\} = 1/E_{\mu_1}(\tau_+)$ to rewrite the resulting expression.

Theorem 3 follows from Lemmas 11–12.

LEMMA 11. *Assume that the P_0-distribution of x_1 has been standardized to have variance* 1. *Then as $\Delta \to 0$,*

$$E_{\mu_2}(\tau_+;\ \tau_+ < \infty) = \mu_1^{-1} E_0(S_{\tau_+})(1 - \rho_+\mu_1 + o(\Delta)),$$
$$E_{\mu_2}(\tau_-) = \mu_2^{-1} E_0(S_{\tau_-})(1 + \rho_-\mu_2 + o(\Delta))$$

and

$$\mu_1^{-1} E_{\mu_1}(S_{\tau_+}) E_{\mu_1}(S_{\tau_-};\ \tau_- < \infty) = -(2\mu_1)^{-1} + \rho_+ - \tfrac{1}{2}\gamma + o(1),$$

where $\rho_\pm = \frac{1}{2}E_0(S_{\tau_\pm}^2)/E_0(S_{\tau_\pm})$ and $\gamma = E_0(s_1^3)$. Also $\rho_+ + \rho_- = \gamma/3$.

PROOF. The first two approximations have proofs similar to (9) [cf. Siegmund (1985), Proposition 10.37]. For the third, differentiate the Wiener–Hopf factorization of the characteristic functions of S_{τ_+} and S_{τ_-} twice [e.g., Siegmund (1985), Theorem 8.41] to obtain

$$\mu_1^{-1} E_{\mu_1}(S_{\tau_+}) E_{\mu_1}(S_{\tau_-};\ \tau_- < \infty) = \frac{E_{\mu_1}(S_{\tau_+}^2)}{2E_{\mu_1}(S_{\tau_+})} - \frac{E_{\mu_1}(x_1^2)}{2\mu_1},$$

and then let $\mu_1 \to 0$. The identity $\rho_+ + \rho_- = \gamma/3$ follows easily from a threefold differentiation of the Wiener–Hopf factorization. □

LEMMA 12. *Let $\tau(x) = \inf\{n\colon S_n \ge x\}$. Then*

$$\int_0^\infty \{E_{\mu_1}\exp[-\Delta(S_{\tau(x)} - x)] - \nu_+\} P_{\mu_1}\Big\{\min_{n \ge 0} S_n > -x\Big\}\, dx \to 0,$$

as $\mu_1 \to 0$.

PROOF. Since $P_{\mu_1}\{\min_{n \ge 0} S_n > -x\} \to 0$ for each fixed x as $\mu_1 \to 0$, it suffices to consider the integral from x_0 to ∞, with x_0 arbitrarily large. Stone's (1965) renewal theorem with exponentially small remainder can be made uniform in μ_1, as indicated briefly by Siegmund (1979), to show that for some $\delta > 0$,

$$P_{\mu_1}\{S_{\tau(x)} - x > y\} = (E_{\mu_1} S_{\tau_+})^{-1} \int_y^\infty P_{\mu_1}\{S_{\tau_+} > u\}\, du + O(e^{-\delta x}), \tag{24}$$

uniformly in y and μ_1 close to 0. Integration by parts and (24) show that uniformly in μ_1,

$$E_{\mu_1}\exp[-\Delta(S_{\tau(x)} - x)] = \nu_+ + O(e^{-\delta x}),$$

which allows one to complete the proof by letting $\mu_1 \to 0$, then $x_0 \to \infty$. □

4. Discussion. The structure which makes possible the explicit evaluation in Theorem 1 is found in the proof of Lemma 4. Locally the increments to the two-dimensional random field are approximately a superposition of independent one-dimensional random fields. Somewhat more precisely, if $\{W_m(i, j),\ i, j = 0, 1, \dots, m\}$ denotes a sequence of (two-dimensional) random fields, the required property is that for the appropriate (i_0, j_0), which typically are proportional to m, conditional on $W_m(i_0, j_0)$ assuming a large value, the increments $W_m(i_0 + i, j_0 - j) - W_m(i_0, j_0)$, perhaps normalized, converge in law as $m \to \infty$ to a sum of independent random walks of the form $S_{1i} + S_{2j}$, $i = 0, \pm 1, \pm 2, \dots$, $j = 0, 1, 2, \dots$.

Although this property is quite special, there are natural problems which have the required structure. Hogan and Siegmund (1986) discuss the two-dimensional pinned Brownian sheet. Some other examples follow.

(i) Let $W(s, t)$, $0 \le s \le 1$, $0 \le t < \infty$, denote the Kiefer–Müller process, i.e., the Gaussian random field with mean 0 and

$$EW(s_1, t_1)W(s_2, t_2) = 4(s_1 \wedge s_2)(1 - s_1 \vee s_2)(t_1 \wedge t_2).$$

The following result is of interest to a statistician who several times as data accumulate announces Kolmogorov–Smirnov confidence bands for a distribution function and wants to know the overall confidence to attach to the several statements. A related, slightly different result gives an approximation to the asymptotic significance level of a nonparametric test for a change point discussed by Deshayes and Picard (1981) and Picard (1985). For fixed $c > 0$ and $m_0 = mt_0 < m_1 = mt_1$, as $m \to \infty$,

$$P\left\{ \max_{\substack{0 \le i \le m \\ m_0 \le j \le m_1}} j^{-1/2} |W(i/m, j)| \ge cm^{1/2} \right\} \tag{25}$$
$$\sim 2\nu(2c) mc^2 \exp\left(-\tfrac{1}{2} mc^2\right) \int_{ct_1^{-1/2}}^{ct_0^{-1/2}} x^{-1} \nu(x)\, dx,$$

where

$$\nu(x) = 2x^{-2} \exp\left(-2 \sum_1^\infty n^{-1} \Phi\left(-\tfrac{1}{2} x n^{1/2}\right) \right)$$
$$= \exp(-\rho x) + o(x^2), \qquad x \to 0,$$

Φ denotes the standard normal distribution function, and $\rho \cong 0.583$. To obtain the analogous result for a continuous maximization over $s \in [0, 1]$ (resp. $t \in [m_0, m_1]$) one replaces $\nu(2c)$ [resp. $\nu(x)$] by 1 in (25).

(ii) For this example we use the notation of Section 2 but restrict ourselves to the special case that the underlying distribution function F is standard normal. For testing a hypothesis of no change point against an epidemic alternative, as discussed by Levin and Kline (1985), the significance level of (a slight generalization of) the likelihood ratio statistic is

$$P_0^{(m)}\left\{ \max_{\substack{0 < i < j < m \\ \delta m < j - i < (1-\delta)m}} \frac{(S_j - S_i)}{\{(j - i)[1 - (j - i)/m]\}^{1/2}} \ge b \right\}.$$

By a calculation similar to the proof of Theorem 1, one obtains for arbitrary $0 \leq \delta < \frac{1}{2}$ as $m \to \infty$ and $b \to \infty$ with $b/m^{1/2}$ equal a fixed positive constant that this probability

$$\sim \frac{b^3\varphi(b)}{4}\int_{1/2}^{1-\delta}\nu^2\left\{\frac{b}{m^{1/2}[u(1-u)]^{1/2}}\right\}\frac{du}{u^2(1-u)^2},$$

where φ denotes the standard normal density function and ν is as defined in example (i).

(iii) Simple modifications of the proof of Theorem 1 yield a large deviation approximation for the Kuiper (1960) statistic

$$\max_{0<x<y<1}\{y - F_n(y) - [x - F_n(x)]\}, \tag{26}$$

where F_n denotes the empirical distribution for a sample of n independent random variables uniformly distributed on $(0, 1)$. From a standard representation of uniform order statistics by sums of exponentially distributed variables, it follows that the probability that (26) exceeds ζ equals

$$P\left\{\max_{0\leq i<j\leq n}\left[W_j - W_i - (j-i)\right] \geq n\zeta - 1 \mid W_{n+1} - (n+1) = -1\right\},$$

where $W_k = y_1 + \cdots + y_k$ and the y's are independent standard exponential random variables. If one puts $m = n + 1$, $b = (m-1)\zeta - 1$ and $\xi = -1$, this probability is almost in the form required by Theorem 1. Minor modifications in the proof of that result yield

$$P\left\{\max_{0<x<y<1}[y - F_n(y) - \{x - F_n(x)\}] \geq \zeta\right\}$$

$$\sim \frac{n\theta_1(1-\theta_1)\zeta^{1/2}\exp\{-n[(\theta_1-\theta_2)\zeta + \theta_2 + \log(1-\theta_2)]\}}{\{|\theta_2|(1-\theta_2)[1 + \theta_2^3(1-\theta_1)/\{\theta_1^3(1-\theta_2)\}]\}^{1/2}},$$

where $\theta_2 < 0 < \theta_1$ satisfy $\theta_1 - \theta_2 = \log[(1-\theta_2)/(1-\theta_1)]$ and $\theta_1^{-1} + |\theta_2|^{-1} = \zeta^{-1}$.

Siegmund (1982) obtains the analogous approximation for the ordinary Kolmogorov–Smirnov statistic and shows numerically that it provides extraordinarily accurate numerical results, but Hogan and Siegmund's (1986) Monte Carlo experiment for a normal random walk indicates that one cannot expect comparable accuracy in this case, unless the sample size is fairly large. It would be interesting to obtain a second-order approximation along the lines of Theorems 2 and 3.

The approximation of Theorem 2 is concerned with the probability that a CUSUM test for a process which is in control terminates well in advance of its average run length. Although this probability is of particular interest, one would also like to have approximations (i) to the right-hand tail of the distribution of t and approximations which are valid, (ii) when $\mu \geq 0$, (iii) for tests with fast initial response feature [Lucas and Crozier (1982)] and (iv) for two-sided tests. Corrected diffusion approximations [Siegmund (1979, 1985)] seem to provide a unified approach to these problems which takes appropriate advantage of the special structure of the process $S_n - \min_{0 \le k \le n} S_k$, $n = 0, 1, \dots$, but they unfortunately do not seem to apply to more general random fields.

One simple approximation in the special case $\gamma = E_0(x_1^3) = 0$ is as follows: Approximate $P_\mu\{t \le m\}$ for a boundary at b by the analogous probability for a Brownian motion process with boundary at $b' = b + \rho_+ - \rho_-$. The approximating Brownian probability can be evaluated as an infinite series [Sweet and Hardin (1970)] and for values of m which are not too small one needs only a single term of the series to obtain good numerical results. For the case of normal x's considered in Section 3, one obtains when $-\mu b' > 1$, $b' = b + 2\rho_+ \cong b + 1.166$,

$$P_\mu\{t > m\} \cong \frac{2q \sinh(qb') \exp\{\mu b' - \frac{1}{2}(\mu^2 - q^2)m\}}{-b'(\mu^2 - q^2)(1 + \mu \sinh^2(qb')/q^2 b')}, \tag{27}$$

where $q > 0$ satisfies $\tanh(qb') = -q/\mu$.

For the small values of m in Table 1 one expects the approximation provided by (27) to be poor unless one includes more terms of the infinite series. For example, for $m = 12$, (27) yields 0.070, whereas (22) gives 0.076 and the correct value is 0.079. For $m = 82$ and 345, for which according to Waldmann (1986) the exact values of $P_\mu\{t \le m\}$ are 0.50 and 0.95, respectively, the approximation (22) is poor, but (27) yields 0.492 and 0.948.

The corrected diffusion approach to distributional problems associated with CUSUM tests will be discussed systematically in a future paper.

Note added in proof. Results essentially equivalent to Example (ii) of Section 4 have been obtained independently by Qi-wei Yao. Details of his argument will be published.

REFERENCES

ADLER, R. and BROWN, L. (1986). Tail behaviour for suprema of empirical processes. *Ann. Probab.* **14** 1–30.

BICKEL, P. and ROSENBLATT, M. (1973). Two-dimensional random fields. In *Multivariate Analysis III* (P. R. Krishnaiah, ed.) 3–15. Academic, New York.

DESHAYES, J. and PICARD, D. (1981). Convergence de processes à double indice; Application aux tests de rupture dans un modèle. *C. R. Acad. Sci. Paris Sér. A-B* **292** 449–452.

FELLER, W. (1972). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

HOGAN, M. and SIEGMUND, D. (1986). Large deviations for the maxima of some random fields. *Adv. in Appl. Math.* **7** 2–22.

KUIPER, N. H. (1960). Tests concerning random points on a circle. *Proc. Kon. Akad. Wetensch.* **A63** 38–47.

LALLEY, S. (1984). Limit theorems for first passage times in linear and nonlinear renewal theory. *Adv. in Appl. Probab.* **16** 766–803.

LEVIN, B. and KLINE, J. (1985). The cusum test of homogeneity with an application to spontaneous abortion epidemiology. *Statist. in Medicine* **4** 469–488.

LUCAS, J. M. and CROZIER, R. B. (1982). Fast initial response (FIR) for cumulative sum quality control schemes. *Technometrics* **24** 199–205.

PICARD, D. (1985). Testing and estimating change points in time series. *Adv. in Appl. Probab.* **17** 841–867.

PICKANDS, J. (1969). Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145** 51–73.

QUALLS, C. and WATANABE, H. (1973). Asymptotic properties of Gaussian random fields. *Trans. Amer. Math. Soc.* **177** 155–171.

SIEGMUND, D. (1975). Error probabilities and average sample number of the sequential probability ratio test. *J. Roy. Statist. Soc. Ser. B* **37** 394–401.

SIEGMUND, D. (1979). Corrected diffusion approximations in certain random walk problems. *Adv. in Appl. Probab.* **11** 701–719.

SIEGMUND, D. (1982). Large deviations for boundary crossing probabilities. *Ann. Probab.* **10** 581–588.

SIEGMUND, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer, Berlin.

SIEGMUND, D. (1986). Boundary crossing probabilities and statistical applications. *Ann. Statist.* **14** 361–404.

STONE, C. (1965). On moment generating functions and renewal theory. *Ann. Math. Statist.* **36** 1298–1301.

SWEET, A. L. and HARDIN, J. C. (1970). Solutions for some diffusion processes with two barriers. *J. Appl. Probab.* **7** 423–431.

VAN DOBBEN DE BRUYN, C. S. (1968). *Cumulative Sum Tests.* Griffin, London.

WALDMANN, K. H. (1986). Bounds for the distribution of the run length of one-sided and two-sided CUSUM quality control schemes. *Technometrics* **28** 61–67.

WOODHALL, W. H. (1983). The distribution of the run length of one-sided CUSUM procedures for continuous variables. *Technometrics* **25** 295–301.

WOODROOFE, M. (1976). Frequentist properties of Bayesian sequential tests. *Biometrika* **63** 101–110.

WOODROOFE, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* SIAM, Philadelphia.

WOODROOFE, M. and TAKAHASHI, H. (1982). Asymptotic expansions for the error probabilities of some repeated significance tests. *Ann. Statist.* **10** 895–908. Correction (1985) **13** 837.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305