TAIL APPROXIMATIONS FOR MAXIMA OF RANDOM FIELDS

BY

D. SIEGMUND

TECHNICAL REPORT NO. 333
SEPTEMBER 1989

PREPARED UNDER THE AUSPICIES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS89-05874

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
TAIL APPROXIMATIONS FOR MAXIMA OF RANDOM FIELDS

BY

D. SIEGMUND

TECHNICAL REPORT NO. 333
SEPTEMBER 1989

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS89-05874

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Tail Approximations for Maxima of Random Fields

D. Siegmund
Stanford University

Abstract

The method for approximating the tail of the distribution of the maximum of a random field first developed by Woodroofe (1976, 1978) and extended to the multidimensional case by Siegmund (1988) is applied to approximate the significance level of the likelihood ratio test for no change-point against the alternative of at most two changes in the mean of a sequence of normal observations with common variance.
1. Introduction.

Although the number of processes $X(t), t \in T$, for which one can evaluate exactly $\Pr\{\max X(t) > b\}$ is very small, the last twenty years have seen a variety of methods developed to obtain approximations to such probabilities, especially asymptotic approximations as $b \to \infty$ and hence the probability converges to zero.

The number of methods is quite large when the indexing set $T$ is one-dimensional although admittedly even there different methods are more or less successful in dealing with different problems. When $T$ is several dimensional, the number of available methods shrinks considerably. Historically the first of these is the asymptotic inclusion-exclusion method developed by Pickands (1969) in one dimension and extended to $d > 1$ dimensions independently by Qualls and Watanabe (1973) and Bickel and Rosenblatt (1973). However, the resulting approximation involves the maximum of a derived random field and in general cannot be explicitly evaluated. Hogan and Siegmund (1986) observed that this method does yield explicit, easily computed approximations in many special cases. More recently, a beautiful heuristic method for obtaining approximations in a number of problems has been developed by Aldous (1989). The structure of the underlying random field which is important for obtaining explicit approximations is that locally it behaves like a superposition of independent random walks. In Aldous' (1989) terminology this is the "uncorrelated orthogonal increments property."

The subject of this paper is a third method, which seems to work particularly well in a variety of problems arising in statistics. The method was pioneered by Woodroofe (1976, 1978) in one dimensional problems of sequential analysis. Siegmund (1988) discussed a number of examples where the method can be adapted for random fields, particularly those arising in change-point problems. (See also Kim and Siegmund (1989).) This paper is concerned with a related example requiring some new techniques, which seem related to what Aldous (1989) calls the marked clump technique.
Section 2 contains a description of a statistical problem and a statement of the main results. A proof is discussed informally in Section 3. Some supporting renewal theoretic lemmas appear in Section 4. Several technical calculations which appear in all boundary crossing problems have been omitted.

2. Tests to Detect Change-points.

Let \( x_n \) \((n = 1, 2, \ldots, m)\) be independent and normally distributed. Assume there exist \( 1 \leq j_1 < j_2 \leq m \) such that \( x_n \) is \( N(\mu_1, \sigma^2) \) for \( n \leq j_1 \), \( N(\mu_2, \sigma^2) \) for \( j_1 < n \leq j_2 \), and \( N(\mu_3, \sigma^2) \) for \( j_2 < n \leq m \). The mean values \( \mu_1, \mu_2, \mu_3 \) and the change-points \( j_1, j_2 \) are all unknown. Usually \( \sigma^2 \) is also unknown, but for simplicity it is assumed known and without loss of generality equal to one for most of this paper. We shall be concerned with testing the hypothesis of no change-point, \( H_0 : \mu_1 = \mu_2 = \mu_3 \), against the alternative that at least two of the \( \mu_i \) are different. There is a substantial literature on the simpler model of at most one change-point, the special case in which it is known that \( \mu_2 = \mu_3 \) or equivalently \( j_2 = m \). See, for example, James, James, and Siegmund (1987).

The likelihood ratio statistic for testing \( H_0 \) is as follows. Let \( S_n = x_1 + \cdots + x_n \) and for \( 0 \leq i \leq j \leq m \) put \( \bar{x}_{i,j} = (S_j - S_i)/(j - i) \) \((\bar{x}_{i,i} = 0)\). Let

\[
Q_{i,j} = i(\bar{x}_{0,i} - \bar{x}_{0,m})^2 + (j - i)(\bar{x}_{i,j} - \bar{x}_{0,m})^2 + (m - j)(\bar{x}_{j,m} - \bar{x}_{0,m})^2.
\]

The (log) likelihood ratio statistic for testing \( H_0 \) is

\[
Q = \max_{1 \leq i < j \leq m} Q_{i,j}.
\]

If the only possible change-points were \( j_1 = i, j_2 = j \), we would be faced with the standard problem of testing the equality of means of three normal populations, and \( Q_{i,j} \) would be the (log) likelihood ratio statistic. Under \( H_0 \) the distribution of \( Q_{i,j} \) is \( \chi^2 \) with two degrees of freedom. Since in fact \( j_1 \) and \( j_2 \) are unknown parameters, the (log) likelihood ratio statistic is \( Q = \max Q_{i,j} \). The significance level of the likelihood ratio test is

\[
P \{Q^{1/2} > b\},
\]
where $P$ denotes probability under $H_0$. The principal results of this paper are approxima-
tions for (3) and for the significance level of the likelihood ratio test when $\sigma^2$ is unknown,
which are given below in Theorems (9) and (10).

It will be convenient to introduce the following notation: $\varphi(x) = (2\pi)^{-1/2} \exp(-\frac{1}{2}x^2)$,
$\Phi(x) = \int_{-\infty}^{x} \varphi(z)dz$, and

$$\nu(x) = 2x^{-2} \exp \left\{ -2 \sum_{1}^{\infty} n^{-1} \Phi(-\frac{1}{2}xn^{1/2}) \right\} \quad (x > 0).$$

(4)

A useful approximation for the function $\nu$ is

$$\nu(x) \approx \exp(-\rho x) + o(x^2) \quad (x \to 0),$$

(5)

where $\rho \cong 0.583$ (cf. Siegmund (1985), Proposition 10.37).

For the simpler alternative of at most one change the likelihood ratio test statistic is

$$\max_{1 \leq k < m} \left\{ (S_k - kS_m/m)^2 / [k(1 - k/m)] \right\}.$$

It is known (Siegmund (1985), Theorem 11.30) that if $m \to \infty$, $b \to \infty$, and
$b/m^{1/2} \to b_0 > 0$, then under $H_0$ for any $0 \leq t_0 < t_1 \leq 1$,

$$P \left\{ \max_{m_0 < k < m_1} \left[ |S_k - kS_m/m| / \{ k(1 - k/m) \}^{1/2} \right] > b \right\}$$

$$\sim b_0 \varphi(b) \int_{t_0}^{t_1} \left[ t(1 - t) \right]^{-1} \nu \left\{ b_0 / \{ t(1 - t) \}^{1/2} \right\} dt.$$

(6)

The special function $\nu$ defined in (4) appears on the right hand side of (6) because the
process $S_k - kS_m/m$ jumps over the curved boundary $b[k(1 - k/m)]^{1/2}$ when crossing it.
If $S_k - kS_m/m$ were replaced by its continuous time analogue, viz. a Brownian Bridge
$B_m(t)$ for $0 \leq t \leq m$, the function $\nu$ on the right hand side of (6) would be replaced by 1,
and this approximating expression would simplify to

$$b_0 \varphi(b) \log \left[ (1 - t_1)/t_0 \right] \left[ (1 - t_1)/t_0 \right],$$

provided in this case that $0 < t_0 < t_1 < 1.$

As indicated in Section 1, (6), which involves a one dimensional indexing set, can be
proved by a much larger class of methods than are available for studying the distribution
of (2). More closely related to the results of this paper is the approximation obtained independently by Siegmund (1988) and Yao (1989) to the significance level of the likelihood ratio test of $H_0$ against the alternative that $\mu_1 = \mu_3 \neq \mu_2$. The test statistic is

$$\max_{1 \leq i < j < m} \left\{ \frac{(S_j - S_i - (j - i)S_m/m)^2}{[(j - i)(1 - (j - i)/m)]} \right\}.$$ 

In this case, if $b \to \infty$, $m \to \infty$, while $b/m^{1/2} \to b_0$, then for any $0 \leq \delta < \frac{1}{2}$ when $H_0$ holds,

$$P \left\{ \max_{0 < i < j < m \atop \delta m < j-i < (1-\delta)m} \frac{|S_j - S_i - (j-i)S_m/m|}{[(j-i)(1-(j-i)/m)^{1/2}] > b} \right\} \sim \frac{1}{2} b^3 \varphi(b) \int_{1/2}^{1-\delta} \nu^2 \left\{ \frac{b_0}{t(t-1)^{1/2}} \right\} \frac{dt}{t^2(1-t)^2}.$$

In order to state the principle results of this paper, the following additional notation is helpful. For $0 < s < t < 1$, $0 \leq \omega < 2\pi$ let

$$(\xi_1 = \xi_1(s,t,\omega) = \frac{1}{2} \cos^2 \omega / [s(1-s/t)],$$

$$\xi_2 = \frac{1}{2} \left\{ \frac{\cos \omega}{[t(1-t)]^{1/2}} - \sin \omega \left[ \frac{s}{t(1-t)} \right]^{1/2} \right\}^2.$$ 

(9) **Theorem.** Suppose $b \to \infty$, $m \to \infty$, and $b/m^{1/2} \to b_0 > 0$. For any $0 \leq \delta_i < \frac{1}{4}$ ($i = 1, 2, 3$), when $H_0$ is true,

$$P \left\{ \max_{\delta_1 m < i < j < (1-\delta_2)m \atop j-i > \delta_3 m} Q_{ij}^{1/2} > b \right\}$$

$$\sim (2\pi)^{-1} b^4 \exp(-\frac{1}{2} b^2) \int_{\xi_1 \leq \xi < \xi_1 + \frac{1-\delta_2}{2}} \int_{\xi_1 \leq \xi \leq \xi_2} \int_{0 < \omega < \delta_1} \xi_1 \xi_2 \nu \left[ b_0 \left( \frac{2\xi_1}{b} \right)^{1/2} \right] \nu \left[ b_0 \left( \frac{2\xi_2}{b} \right)^{1/2} \right] d\omega d\xi_1 d\xi_2 dt.$$

The following result is appropriate when $\sigma^2 = \text{var}(x_n)$ is unknown and must be estimated from the data. Its proof will not be discussed in this paper. See James, James, and Siegmund (1988) for an analogous result in the simpler case of at most one change.
(10) **Theorem.** Assume the conditions of (9) and that \(0 < b_0 < 1\). Then when \(H_0\) is true

\[
P \left\{ \max_{j-i > \delta_0} \max_{0 < \epsilon < 1-\epsilon_2} \frac{Q_{i,j}}{b^2 m^{-1}} > b^2 m^{-1} \sum_{k=1}^{m} (x_k - \bar{x}_{i,j})^2 \right\}
\]

\[
\sim (2\pi)^{-1} b^4 (1-b_0^2)^{(m-5)/2} \int_0^{2\pi} \int_{-\epsilon_2}^{\epsilon_2} \int_{-\epsilon_2}^{\epsilon_2} \xi_1 \xi_2 \nu \left[ b_0 \left\{ 2\xi_1/(1-b_0^2) \right\}^{1/2} \right] \nu \left[ b_0 \left\{ 2\xi_2/(1-b_0^2) \right\}^{1/2} \right] \, dw \, ds \, dt.
\]

**Remark.** Numerical evaluation of the integrals in (9) and (10) is somewhat onerous, but becomes considerably easier if one uses the approximation (5) for the function \(\nu\) defined in (4). In the simpler cases already studied this approximation is satisfactory provided \(m\) is not too small, say \(m \geq 25\); and hence one expects it will also prove useful in the present case. For example, if in (7) \(b = 3.5\), \(m = 40\), and \(\delta = 0.1\), the right hand side equals 0.0504. If the approximation (5) for \(\nu\) is used the result is 0.0494. On the other hand, if \(\nu\) is replaced by 1, which would be appropriate if \(S_n - nS_m/m\) were in fact continuous Brownian Bridge, \(B_m(t)\) for \(0 \leq t \leq m\), the integral in (7) can be evaluated explicitly but the approximation is a very poor 0.25.

3. **Proof of (9).**

In order to prove (9) it is helpful to rewrite the definition (1) in a way which clearly exhibits each \(Q_{i,j}\) as a random variable whose distribution under \(H_0\) is \(\chi^2\) with two degrees of freedom. Let \(S_n^{(m)} = S_n - nS_m/m\). Note that under \(H_0\) \(S_n^{(m)} (n = 0, 1, \ldots, m)\) is a discrete time Brownian Bridge process, with mean zero and covariance function \(n_1(1-n_2/m)\) for \(n_1 \leq n_2\). It is easily verified that

\[
Q_{i,j} = Z_1^2 + Z_2^2,
\]

where

\[
Z_1 = (S_i - iS_j/j) / [i(1-i/j)]^{1/2} = (S_i^{(m)} - iS_j^{(m)}/j) / [i(1-i(j))]^{1/2}
\]

and

\[
Z_2 = S_j^{(m)} / [j(1-j/m)]^{1/2}
\]
are independent standard normal random variables under $H_0$. Although the dependence of $Z_1$ and $Z_2$ on $i$ and $j$ has been suppressed, no confusion should result since $Z_1$ and $Z_2$ will only appear in the company of other variables which indicate the appropriate indices.

From $Q_{i,j}^{1/2} = \sup_{0 \leq \omega < 2\pi} [Z_1 \cos \omega + Z_2 \sin \omega]$ it follows that $Q^{1/2}$ defined in (2) is the maximum of a three dimensional Gaussian random field. Kim and Siegmund (1989) have used such a representation in a similar problem, and it would undoubtedly work here. However, a somewhat different approach seems more natural and simpler computationally. This approach would presumably also work for Kim and Siegmund.

Let $D = \{(i_0, j_0) : \delta_1 m < i_0 < j_0 < (1 - \delta_2)m, \ j_0 - i_0 > \delta_3 m\}$, and for each $(i_0, j_0) \in D$ let $J_0 = J_0(i_0, j_0) = \{(i, j) : (i, j) \in D, \ i > i_0 \text{ and } j = j_0, \text{ or } j > j_0\}$. If $Q_{i,j} > b$ for some $(i, j) \in D$, there must be a “last” value $(i_0, j_0)$, where “last” means that $Q_{i,j} < b$ for all $i > i_0$ and $j = j_0$ and for all $j > j_0$. Hence

\[
P \left\{ \max_{(i,j) \in D} Q_{i,j}^{1/2} > b \right\} = \sum_{(i_0,j_0) \in D} \int_0^\infty \int_0^{2\pi} P \left\{ Q_{i_0,j_0}^{1/2} \in b + m^{-1/2} dx, \ \arg(Z_1, Z_2) \in d\omega \right\} \times P \left\{ \max_{(i,j) \in J_0} Q_{i,j}^{1/2} \leq b | Q_{i_0,j_0}^{1/2} = b + m^{-1/2}x, \ \arg(Z_1, Z_2) = \omega \right\}.
\]

Obviously

\[
P \left\{ Q_{i_0,j_0}^{1/2} \in b + m^{-1/2} dx, \ \arg(Z_1, Z_2) \in d\omega \right\} = (b + m^{-1/2}x) \exp \left[ -\frac{1}{2} (b + m^{-1/2}x)^2 \right] m^{-1/2} dx (2\pi)^{-1} d\omega \sim b_0 \exp \left[ -\frac{1}{2} b^2 - b_0 x \right] dx (2\pi)^{-1} d\omega
\]

uniformly for bounded $x$. Some technical analysis along the lines given in Siegmund (1988) shows that it is permissible to ignore terms $(i_0, j_0)$ in (11) near the boundaries of $D$, and in analyzing the conditional probability to consider only values of $(i, j)$ within distance $0(m^{1/8})$ of $(i_0, j_0)$ and arbitrary fixed values of $x$.

The conditional probability in (11) equals

\[
P \left\{ \max_{(i,j) \in J_0} m^{1/2} (Q_{i,j}^{1/2} - Q_{i_0,j_0}^{1/2}) \leq -x | Q_{i_0,j_0}^{1/2} \right. \]

\[
= b + m^{-1/2}x, \ Z_1 = (b + m^{-1/2}x) \cos \omega \}.
\]
Let \( f(y_1, y_2, s, t) = \left( (y_1 - sy_2/t)^2 / s(1-s/t) + y_2^2 / t(1-t) \right)^{1/2}, \) so

\[
Q_{i,j}^{1/2} = m^{1/2} f \left[ m^{-1} S_i^{(m)}, m^{-1} S_j^{(m)}, m^{-1} i, m^{-1} j \right].
\]

Moreover, a Taylor series expansion shows that if \( k = i - i_0 \) and \( \ell = j - j_0 \) are \( O(m^{1/8}) \), then

\[
(14) \quad m^{1/2}(Q_{i,j}^{1/2} - Q_{i_0,j_0}^{1/2}) = (S_{i_0+k}^{(m)} - S_{i_0}^{(m)}) \frac{\partial f}{\partial y_1} + (S_{j_0+\ell}^{(m)} - S_{j_0}^{(m)}) \frac{\partial f}{\partial y_2} + k \frac{\partial f}{\partial s} + \ell \frac{\partial f}{\partial t} + O_p(m^{-3/4}),
\]

where the partial derivatives are evaluated at \( m^{-1}(S_{i_0}^{(m)}, S_{j_0}^{(m)}, i_0, j_0) \).

Suppose now that \( m \to \infty \) with \( m^{-1}(i_0, j_0) \to (s, t) \) for some \( 0 < s < t < 1 \). A long messy calculation of means and covariances shows that under the conditioning indicated in (13) the right hand side of (14) converges in law to the random field

\[
W_k + W_{\ell}' \quad (k, \ell = 0, \pm 1, \pm 2, \ldots),
\]

where \( W_k \) and \( W_{\ell}' \) are mutually independent random walks with \( W_0 = W_0' = 0 \). The increments of \( W_k \) are \( N(-b_0\xi_1, 2\xi_1) \) and the increments of \( W_{\ell}' \) are \( N(-b_0\xi_2, 2\xi_2) \), where \( \xi_1 \) and \( \xi_2 \) are defined in (8). It follows from the definition of \( J_0 \) that the conditional probability in (13) converges to

\[
(15) \quad P \left\{ \max_{k \geq 0} W_k + \max_{\ell \geq 0} W_{\ell}' \leq -x, \quad \max_{k \leq 0} W_k + \max_{\ell \geq 1} W_{\ell}' \leq -x \right\}.
\]

It is necessary to remember that in (15), as well as in (16) and (17) below, the distribution of the random walks \( \{W_k\} \) and \( \{W_{\ell}'\} \) depend upon \( \xi_1 \) and \( \xi_2 \), and hence ultimately on \( (s, t, \omega) \).

Since \( \max_{k \leq 0} W_k \geq 0, \max_{k \leq 0} W_k + \max_{\ell \geq 1} W_{\ell}' \leq -x \) implies \( \max_{\ell \geq 1} W_{\ell}' \leq -x \) and hence \( \max_{\ell \geq 0} W_{\ell}' = 0 \). It follows that the probability (15) equals

\[
(16) \quad P \left\{ \max_{k \geq 1} W_k \leq -x, \max_{k \leq 0} W_k + \max_{\ell \geq 1} W_{\ell}' \leq -x \right\} = P \left\{ \max_{k \geq 1} W_k \leq -x \right\} P \left\{ \max_{k \leq 0} W_k + \max_{\ell \geq 1} W_{\ell}' \leq -x \right\}.
\]
Substituting (12) and (16) into (11) and approximating the sum over $D$ by a double integral, one obtains

\begin{equation}
(17) \quad P \left\{ \max_{(i,j) \in D} Q_{i,j}^{1/2} > b \right\} \sim (2\pi)^{-1/2} b_0 m^2 \exp(-\frac{1}{2} b^2) 
\times \int \int \int \int_0^\infty \exp(-b_0 x) P \left\{ \max_{k \geq 1} W_k \leq -x \right\} P \left\{ \max_{\ell \geq 1} W'_\ell \leq -x \right\} dx \omega ds dt,
\end{equation}

where the outer three integrals are over the domain indicated in the statement of the theorem.

The proof is concluded by appealing to Lemma (21) in Section 4 in order to evaluate the inner integral on the right hand side of (17) in terms of the function $\nu$ defined in (4).

Remarks. (i) The preceding calculation breaks down in a neighborhood of points $(s, t, \omega)$ where $\xi_1$ or $\xi_2$ vanishes, e.g. along $\omega = \pi/2$. Since the totality of such points has measure 0, it is easy to see that they cause no real difficulty, although they present one more technical obstacle to a completely rigorous exposition.

(ii) In Siegmund (1988) it was necessary to evaluate an integral similar to the inner integral in (17). In fact the same method works here. Since that renewal theoretic argument is a particularly interesting feature of the method of this paper, it is given in detail in the following section. An extension to deal with higher dimensional random fields is also given.

(iii) The structure of the present problem which makes possible an evaluation of the inner integral in (17) is that (a) locally the random field behaves like a superposition of independent random walks, and (b) the constant $b_0$ in the exponential function is twice the ratio of the mean to the variance for the two random walks. Although (a) is at first surprising, examination of the variety of examples considered by Hogan and Siegmund (1986), Siegmund (1988), and Aldous (1989) leads one to view it as a natural consequence of the structure of the original random field. On the other hand (b) appears as the outcome of an elaborate calculation, and after comparison with previously studied examples it still seems rather mysterious.

In this section random walk and renewal theory are used to evaluate the inner integral in (17) in terms of the function $\nu$ defined in (4). The method comes from Siegmund (1988). An extension appropriate for higher dimensional random fields is also given. The notation is not necessarily consistent with that of Sections 1–3.

Let $W_n(n = 1, 2, \ldots, W_0 = 0)$ be a random walk with $\mu = E(W_1) > 0$. Let $\tau_+ = \inf\{n : W_n > 0\}$.

(18) Lemma. For all $x \geq 0$

$$\mu^{-1}P \left\{ \min_{n \geq 1} W_n > x \right\} = [E(W_{\tau_+})]^{-1}P\{S_{\tau_+} > x\}.$$ 

Proof. This result is well-known. See, for example, Woodroofe (1982, Theorem 2.7) or Siegmund (1985, Problem 8.13).

(19) Lemma. If $W_1$ is $N(\mu, \sigma^2)$, then

$$\int_0^\infty \exp(-2\mu x/\sigma^2)P \left\{ \min_{n \geq 1} W_n > x \right\} dx = \mu \nu(2\mu/\sigma) = \frac{1}{2}\sigma^2 P \left\{ \min_{n \geq 0} S_n = 0 \right\} / E(W_{\tau_+}),$$

where $\nu$ is defined in (4).

Proof. We write $P_{\mu,\sigma^2}$ to denote dependence of probabilities on the parameters $\mu, \sigma^2$. Let $\zeta = \mu/\sigma$. By the change of variables $y = x\sigma$ and Lemma (18) one obtains

$$\int_0^\infty \exp(-2\mu x/\sigma^2)P_{\mu,\sigma^2} \left\{ \min_{n \geq 1} W_n > x \right\} dx$$

$$= \sigma \int_0^\infty \exp(-2\zeta y)P_{\zeta,1} \left\{ \min_{n \geq 1} W_n > y \right\} dy$$

$$= \mu [E_{\zeta,1}(W_{\tau_+})]^{-1} \int_0^\infty \exp(-2\zeta y)P_{\zeta,1} \{W_{\tau_+} > y\} dy$$

$$= \mu [2\zeta E_{\zeta,1}(W_{\tau_+})]^{-1}[1 - E_{\zeta,1} \exp(-2\zeta W_{\tau_+})]$$

$$= \mu [2\zeta E_{\zeta,1}(W_{\tau_+})]^{-1}P_{-\zeta,1} \{\tau_+ = \infty\},$$
where the last equality is a consequence of Wald’s likelihood ratio identity. That the last three expressions all equal \( \mu \nu (2 \zeta) \) is well known (e.g., Siegmund (1985, pp. 175-176)). That the last one also equals

\[
\frac{1}{2} \sigma^2 P_{\mu, \sigma^2} \left\{ \min_{n \geq 0} S_n = 0 \right\} / E_{\mu, \sigma^2}(W_{\tau_+})
\]

is obvious.

Now let \( W'_n (n = 0, 1, \ldots) \) be a second random walk, independent of the first, with \( \mu' = E(W'_1) > 0 \). Let \( M = \min_{n \geq 1} W_n, M^- = \min(M, 0) = \min_{n \geq 0} W_n \), and let \( M' = \min_{n \geq 1} W'_n \).

(20) Lemma. For any \( \lambda > 0 \)

\[
\int_0^\infty \exp(-\lambda x) P\{M > x\} P\{M^- + M' > x\} dx
= \int_0^\infty \exp(-\lambda y) P\{M' \in dy\} \int_y^\infty \exp [\lambda(y - x)] P\{M^- > -(y - x)\} dx.
\]

Proof. The result follows at once by writing

\[
P\{M^- + M' > x\} = \int_x^\infty P\{M' \in dy\} P\{M^- > x - y\}
\]

and inverting the order of integration.

(21) Lemma. Suppose \( W_1 \) is \( N(\mu, \sigma^2) \), \( W'_1 \) is \( N(\mu', (\sigma')^2) \), and that \( \alpha = \mu / \sigma^2 = \mu' / (\sigma')^2 \). Then

\[
\int_0^\infty \exp(-2\alpha x) P\{M > x\} P\{M^- + M' > x\} dx = 2\alpha \mu \mu' \nu(2\mu / \sigma) \nu(2\mu' / \sigma')
\]

Proof. It suffices to evaluate the right hand side of (20) for \( \lambda = 2\alpha \). By (18) the inner integral on the right hand side of (20) (with \( \lambda = 2\alpha \)) is

\[
\mu \int_0^y P\{W_{\tau_+} > x\} \exp \{2\alpha(y - x)\} P\{M^- > -(y - x)\} dx
= \mu \int_0^y F_0(dx) Z(y - x),
\]
where $F_0(dx) = P\{W_{r_+} > x\}dx/EW_{r_+}$ is the stationary measure for the renewal process determined by $F(dx) = P\{W_{r_+} \in dx\}$ and, as will be shown in Lemma (23) below,

$$Z(x) = \exp(2\alpha x)P\{M^- > -x\}$$

is a solution of the renewal equation

$$Z(y) = z(y) + F \ast Z(y)$$

associated with $z(y) = \exp(2\alpha y)P\{M^- = 0\}$. From standard renewal theory (e.g., Feller (1972, Chapter XI)) one knows that $Z = U \ast z$, where $U = \sum_{n=0}^{\infty} F^{n*}$ is the renewal measure associated with $F$. Also $F_0 \ast U(dx) = dx/EW_{r_+}$ and hence

$$F_0 \ast Z(y) = F_0 \ast U \ast z(y) = \int_0^y ds \exp \left[2\alpha(y - x)\right] P\{M^- = 0\}/EW_{r_+}$$

$$= (2\alpha EW_{r_+})^{-1} P\{M^- = 0\} \left[\exp(2\alpha y) - 1\right].$$

Substituting this into (22), one sees from (20) that the left hand side of (21) equals

$$\mu(2\alpha EW_{r_+})^{-1} P\{M^- = 0\} \int_0^\infty [1 - \exp(-2\alpha y)] P\{M' \in dy\}$$

$$= \mu(EW_{r_+})^{-1} P\{M^- = 0\} \int_0^\infty \exp(-2\alpha y)P\{M' > y\}dy$$

$$= 2\alpha \mu \nu(2\mu/\sigma)\nu(2\mu'/\sigma'),$$

where the final equality is a consequence of (19).

The following result has been used above in the proof of (21).

**Lemma.** Suppose $W$ is $N(\mu, \sigma^2)$ and put $\alpha = \mu/\sigma^2$. Let $F(dx) = P\{W_{r_+} \in dx\}$. The function $Z(x) = \exp(2\alpha x)P\{M^- > -x\}$ satisfies the renewal equation $Z = z + F \ast Z$ with $z(x) = \exp(2\alpha x)P\{M^- = 0\}$.

**Proof.** As above we write $P_{\mu, \sigma^2}$ to denote dependence of probabilities on $\mu, \sigma^2$. Let $G(x) = P_{\mu, \sigma^2}\{M^- > -x\} = P_{-\mu, \sigma^2}\{\max_{n \geq 0} W_n < x\}$ $(x > 0)$, so $Z(x) = \exp(2\alpha x)G(x)$. Then

$$G(x) = P_{-\mu, \sigma^2}\{\tau_+ = \infty\} + \int_{(0,x)} P_{-\mu, \sigma^2}\{\tau_+ < \infty, W_{r_+} \in dy\} G(x - y).$$
Multiplying by \( \exp(2\alpha x) \) yields
\[
Z(x) = \exp(2\alpha x) P_{\mu, \sigma^2} \{ M^- = 0 \} \\
+ \int_{(0,x)} P_{-\mu, \sigma^2} \{ \tau_+ < \infty, W_{\tau_+} \in dy \} \exp(2\alpha y) Z(x - y).
\]

By Wald's likelihood ratio identity
\[
P_{-\mu, \sigma^2} \{ \tau_+ < \infty, W_{\tau_+} \in dy \} \exp(2\alpha y) = P_{\mu, \sigma^2} \{ W_{\tau_+} \in dy \},
\]
which completes the proof.

The following result indicates how (21) can be extended to deal with higher dimensional random fields exhibiting the same structure as that encountered above. Let \( W''_n, n = 0, 1, \ldots \) be a third random walk, independent of \( \{ W_n \} \) and \( \{ W'_n \} \), with \( W''_1 \) having a \( N(\mu'', \sigma'')^2 \) distribution.

(24) **Proposition.** Assume \( \alpha = \mu/\sigma^2 = \mu'//(\sigma')^2 = \mu''//(\sigma'')^2 \). Let \( M_1, \ldots, M_5 \) be independent random variables. Suppose \( M_1 \) and \( M_2 \) have the distribution of \( \min_{n \geq 1} W_n, M_3 \) and \( M_4 \) have the distribution of \( \min_{n \geq 1} W'_n \), and \( M_5 \) has the distribution of \( \min_{n \geq 1} W''_n \). Let \( a^- \) denote \( \min(a, 0) \). Then
\[
\int_0^\infty \exp(-2\alpha x) P\{ M_1 > x \} P\{ M_2^- + M_3 > x, M_2^- + M_4^- + M_5 > x \} dx \\
= (2\alpha)^2 \mu \mu' \nu(2\mu/\sigma) \nu(2\mu'/\sigma') \nu(2\mu''/\sigma'').
\]

**Proof.** Since
\[
P\{ M_2^- + M_3 > x, M_2^- + M_4^- + M_5 > x \}
= \int_x^\infty P\{ \min(M_3, M_4^- + M_5) \in dy \} P\{ M_2^- > -(y - x) \},
\]
by (18) the integral on the left hand side of (24) equals
\[
\mu \int_0^\infty P\{ \min(M_3, M_4^- + M_5) \in dy \} \exp(-2\alpha y) \\
\int_0^y \frac{P\{ W_{\tau_+} > x \}}{EW_{\tau_+}} \exp[2\alpha(y - x)] P\{ M_2^- > -(y - x) \} dx.
\]

12
By the proof of (21) the inner integral equals \((2\alpha EW_{\tau_1})^{-1} P\{M_1^- = 0\}[\exp(2\alpha y) - 1]\) and hence the entire expression equals

\[
\mu(2\alpha EW_{\tau_1})^{-1} P\{M_1^- = 0\} \int_0^\infty P\{\min(M_3, M_4^- + M_5) \in dy\} [1 - \exp(-2\alpha y)]
\]

\[
= (2\alpha)\mu\nu(2\mu/\sigma) \int_0^\infty \exp(-2\alpha x) P\{M_3 > x\} P\{M_4^- + M_5 > x\} dx
\]

by (19) and an integration by parts. This reduces the present case to the one previously considered, so (24) follows from (21).
References


Woodroofe, M. (1978). Large deviations of the likelihood ratio statistic with applications to sequential testing, Ann. Statist. 6, 72-84.