The Annals of Probability
1991, Vol. 19, No. 1, 401–422

NONLINEAR RENEWAL THEORY FOR CONDITIONAL RANDOM WALKS

BY INCHI HU

University of Pennsylvania

Herein boundary crossing behavior of conditional random walks is studied. Asymptotic distributions of the exit time and the excess over the boundary are derived. In the course of derivation, two results of independent interest are also obtained: Lemma 4.1 shows that a conditional random walk behaves like an unconditional one locally in a very strong sense. Theorem B.1 describes a class of distributions over which the renewal theorem holds uniformly. Applications are given for modified repeated significance tests and change-point problems.

1. Introduction. Let $X, X_1, X_2, \dots$ be i.i.d. random vectors whose common density (with respect to Lebesgue measure over R^d) belongs to the d-dimensional exponential family

$$f_\theta(x) = \exp\{\theta \cdot x - \psi(\theta)\} f(x), \tag{1.1}$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_d)$, $x = (x_1, x_2, \dots, x_d)$ and $\cdot$ denotes the usual inner product. It is well known that $E_\theta X = \nabla\psi(\theta) = \mu(\theta)$ and $\mathrm{Cov}_\theta\, X = \nabla^2\psi(\theta) = \Sigma_\theta$. Here and in the sequel, we use the notation $\nabla\psi$ to denote the gradient vector $(\partial\psi/\partial\theta_1, \dots, \partial\psi/\partial\theta_d)$ and $\nabla^2\psi$ to denote the Hessian matrix $(\partial^2\psi/\partial\theta_i\,\partial\theta_j)_{1 \le i \le d, 1 \le j \le d}$. Let

$$\Theta = \mathrm{int}\left\{\theta\colon \int \exp(\theta \cdot x) f(x)\, dx < \infty\right\}$$

be the interior (Int) of the natural parameter space. We assume that Θ contains a neighborhood of 0. Clearly, $\psi(0) = 0$. Noting that $\nabla\psi$ is a diffeomorphism on Θ, we can write $\theta(\mu)$ to indicate that θ is a function of μ. Similarly, write $\Sigma(\mu)$ for $\Sigma_{\theta(\mu)}$. Let $\Gamma = \nabla\psi(\Theta)$ be the set of possible expectations. We shall assume that, for each θ, there exists a constant M such that

$$\sup_x f_{\theta, n}(x) \le M, \tag{1.2}$$

where $f_{\theta, n}$ denotes the n-fold convolution of f_θ. Moreover (1.2) holds uniformly for θ in compact subsets of Γ.

The following notation will be used throughout this paper. For any vector $v \in R^d$, let $v^{(1)}, v^{(2)}$ be two vectors such that $v = (v^{(1)}, v^{(2)})$, where $v^{(1)} \in R^{d_1}$, $v^{(2)} \in R^{d_2}$ and $d_1 + d_2 = d$. We do not exclude the possibility of $d_1 = 0$. Let the common density $X, X_1, X_2, \dots$ be f_ϑ with $\vartheta = (0^{(1)}, \theta^{(2)})$. Then it is easy to see that $X^{(2)}, X_1^{(2)}, X_2^{(2)}, \dots$ are i.i.d. random vectors with common density

Received August 1988; revised June 1989.
AMS 1980 *subject classifications*. Primary 60K05, 60K40; secondary 62J15.
Key words and phrases. Nonlinear renewal theory, renewal theorem, boundary crossing probabilities, conditional random walks, exponential family.

belongs to the d_2-dimensional exponential family

$$(1.3)\qquad f_\theta^{[2]}(x^{(2)}) = \int f_\vartheta(x)\,dx^{(1)} = \exp\{\theta^{(2)}\cdot x^{(2)} - \psi_2(\theta^{(2)})\} f^{[2]}(x^{(2)}),$$

where $\psi_2(\theta^{(2)}) = \psi(\vartheta)$ and $f^{[2]}(\cdot)$ denotes the density of $X^{(2)}$ when $\theta = 0$.

Use Γ_2 to denote the set of possible expectations of $X^{(2)}$. Put $S_m = \sum_{i=1}^m X_i$. For $\xi_0 \in \Gamma$ ($\xi_0^{(2)} \in \Gamma_2$) and $\xi = m\xi_0$ ($\xi^{(2)} = m\xi_0^{(2)}$) let $P_\xi^{(m)}$ ($P_{\xi^{(2)}}^{(m)}$) denote the conditional probability law given $S_m = \xi$ ($S_m^{(2)} = \xi^{(2)}$, respectively). That is,

$$P_\xi^{(m)}(A) = P_0(A \mid S_m = \xi), \qquad P_{\xi^{(2)}}^{(m)}(A) = P_0(A \mid S_m^{(2)} = \xi^{(2)}).$$

The last assumption on the distributions is that for any positive ε,

$$(1.4)\qquad \lim_{m\to\infty} P_{\xi^{(2)}}^{(m)}(\|S_m/m - \mu_0\| \le \varepsilon) = 1,$$

where μ_0 is given by (3.2).

REMARK. (1.4) is true when $S_m^{(1)}$ and $S_m^{(2)}$ are independent, and it is also true for other interesting cases [see, e.g., Section 2.3 of Hu (1988)].

Define

$$(1.5)\qquad T_m = \inf\{n \ge m_0 : mH(S_n/m, n/m) > 0\},$$

where H is a "smooth" function from R^{d+1} to R. The goal of this paper is to identify the marginal and joint asymptotic distributions of the exit time T_m and the excess over the boundary $mH(S_{T_m}/m, T_m/m)$ under the probability law $P_{\xi^{(2)}}^{(m)}$ as $m \to \infty$.

In the literature, the asymptotic distribution of T_m and $mH(S_{Tm}/m, T_m/m)$ for unconditional random walks have been studied extensively under the name of nonlinear renewal theory; cf. Lai and Siegmund (1977), Woodroofe (1982), Hogan (1984) and Zhang (1988). Nonlinear renewal theory is essential to the approximation of various boundary crossing probabilities which arise from several statistical problems [see, e.g., Siegmund (1986)]. A method which has been quite successful in developing approximations for boundary crossing probabilities can be briefly described as follows.

First, find a measure Q whose likelihood ratio with respect to the probability measure under consideration is a simple function of the excess over the boundary and the exit time. Then use Wald's likelihood ratio identity to convert the boundary crossing probability into the expectation of the likelihood ratio at the exit time. Finally, nonlinear renewal theory can be employed to identify the asymptotic distribution of the likelihood ratio and, hence, deliver an approximation to the boundary crossing probability.

Recently, Hu (1988), James, James and Siegmund (1988) and Siegmund (1988) showed that for some problems it is necessary to take Q to be a conditional probability measure. Thus the knowledge of conditional nonlinear renewal theory is essential to the solutions of these problems.

In Section 2, some examples are presented from which the problem of identifying asymptotic distributions of the exit time and the excess over the boundary for conditional random walks naturally arise. The main results are stated in Section 3. Sections 4 and 5 contains the proofs of these results.

At the heart of the proofs is a restarting method developed by Lai and Siegmund (1977) [also see Hogan (1984) and Siegmund (1985, 1986)]. However, when applying this method to conditional random walks, new difficulties arise: Unlike that of unconditional random walks, the probability law which governs the restarted random walk is not the same as the original one. It turns out that some uniformity results are required. Two uniformity results (Lemma 4.1 and Theorem B.1), which may be useful elsewhere, are obtained in this paper.

2. Examples.

2.1. *Modified repeated significance tests.* Suppose, as in Section 1, $X, X_1, X_2, \dots$ are i.i.d. observations with common density (1.1). Based on these observations, it is desired to test the null hypothesis $\theta^{(1)} = 0^{(1)}$ against the alternative $\theta^{(1)} \neq 0^{(1)}$, where $\theta = (\theta^{(1)}, \theta^{(2)})$. So $\theta^{(2)}$ plays the role of nuisance parameters. Let $l_n(\theta)$ be the log likelihood after observing $X_1, \dots, X_n$. Also, let

$$\phi(x) = \sup_{\theta \in \Theta} [\theta \cdot x - \psi(\theta)], \qquad \phi_2(x) = \sup_{\theta \in \Theta_0} [\theta \cdot x - \psi(\theta)], \tag{2.1}$$

where $\Theta_0 = \{\theta \in \Theta : \theta^{(1)} = 0^{(1)}\}$. Then the generalized likelihood ratio statistic for this testing problem is

$$n\Lambda(S_n/n) = n\phi(S_n/n) - n\phi_2(S_n/n).$$

A repeated significance test (RST) is a sequential test defined in terms of the stopping rule

$$T = \inf\{n \geq m_0 : n\Lambda(S_n/n) > a\}.$$

Sampling stops at $\min(T, m)$ and the null hypothesis is rejected when $T \leq m$. If $T > m$, the null hypothesis should not be rejected. The significance levels of an RST are given by

$$\sup_{\theta \in \Theta_0} P_\theta\{T \leq m\}.$$

The RST, like most sequential tests, is designed to save samples. The price of smaller expected sample size of an RST is a loss of power compared to a fixed sample test of sample size m. The loss of power is a serious disadvantage if, as is often the case in clinical trials, one anticipates a small treatment difference which even a fixed sample test may have insufficient power to detect. Hence it is desirable to modify the RST to increase their power without completely losing their expected sample size advantage in case a large treatment effect does exist. The modified repeated significance test (MRST) is introduced for this purpose. See Hu (1988) and Siegmund (1985) for details. An MRST rejects the null hypothesis when either $T \leq m$ or $T > m$ and $m\Lambda(S_m/m) > c$ for some $c < a$. The significance levels of an MRST are given by

$$\begin{aligned} &\sup_{\theta \in \Theta_0} [P_\theta\{T \leq m\} + P_\theta\{T > m, m\Lambda(S_m/m) > c\}] \\ &\quad = \sup_{\theta \in \Theta_0} [P_\theta\{m\Lambda(S_m/m) > c\} + P_\theta\{T < m, m\Lambda(S_m/m) \leq c\}]. \end{aligned} \tag{2.2}$$

The first term on the RHS of (2.2) can be handled by the usual asymptotic theory. It is the second term which presents new difficulties.

Approximations for significance levels of RST in an exponential family have been provided by Woodroofe (1978, 1979) and Lalley (1983), but their methods seem to be difficult to apply in approximating the significance levels of MRST. This problem is solved by Hu (1988), who shows that [see the proof of Theorem 2 of Hu (1988)]

THEOREM 2.1. *Let*

$$H_1(x,t) = (1-t)\Lambda[(\xi_0 - x)/(1-t)] - a_0$$

and

$$\tau = \inf\{n: mH_1(S_n/m, n/m) > 0\}.$$

Define

$$t_0 = \inf\{t > 0: H_1(t\mu_0, t) = 0\},$$

where $\mu_0 = \mu(\theta_0)$, *with* $\theta_0 = (0^{(1)}, \theta^{(2)}(\xi_0^{(2)}))$. *Suppose* $a \to \infty$, $m \to \infty$ *and* $m_0 \to \infty$ *in such a way that* $a/m = a_0 > 0$ *and* $1 - m_0/m = a_1 > 0$. *Then for each* $\xi_0 = \xi/m$ *such that* $t_0 < 1 - m_0/m$ *and* $\Lambda(\xi_0) < a_0$, *we have*

$$\begin{aligned} P_\xi^{(m)}\{T \le m\} \sim \exp\{-m[a_0 - \Lambda(\xi_0)]\}(1-t_0)^{-d_1/2} & \\ \times|\Sigma[(\xi_0 - t_0\mu_0)/(1-t_0)]|^{-1/2} & \\ \times|\Sigma(\xi_0)|^{1/2} E_{\xi^{(2)}}^{(m)}\{\exp[-mH_1(S_\tau/m, \tau/m)]\}. & \end{aligned} \tag{2.3}$$

In the previous formula, we have used $|M|$ to denote the determinant of a matrix M. Observe that we need to identify the limiting distributions of $mH_1(S_\tau/m, \tau/m)$. Then approximations of significance levels of MRST can be obtained by integrating the RHS of (2.3) with respect to $P\{S_m \in d\xi\}$. Theorem 2.1 can also be used to approximate the p values and power of MRST. See Hu (1988) for details.

2.2. *Change point problems.* Assume that $X_1, X_2, \dots, X_m$ are independent normally distributed random variables and that X_i has mean μ_i and unit variance. Suppose we are interested in testing the null hypothesis $\mu_1 = \cdots = \mu_m$ against the alternative there exist k, $1 < k < m-1$, such that $\mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \mu_{k+2} = \cdots = \mu_m$. Let

$$\begin{aligned} \Lambda_{k,m} &= k(1-k/m)\left[\sum_{i=1}^{k} X_i/k - \sum_{i=k+1}^{m} X_i/(m-k)\right]^2 \\ &= (S_k - kS_m/m)^2/[k(1-k/m)]. \end{aligned}$$

Then the log likelihood ratio statistic is found to be

$$\max_{1 \le n \le m-1} \Lambda_{n,m} \tag{2.4}$$

and the significance level of the likelihood ratio test is the probability, under the null hypothesis, that the random variable (2.4) exceeds some constant c. It is easy to see that under the null hypothesis the random variables

$$S_n - nS_m/m, \qquad n = 1, 2, \dots, m-1,$$

have the same joint distribution as $S_1, S_2, \dots, S_{m-1}$ given $S_m = 0$. The significance level is given by

$$P_0^{(m)}\{T \le m - 1\},$$

where $T = \inf\{n: |S_n| \ge b[n(1 - n/m)]^{1/2}\}$, $b = \sqrt{c}$. The following argument provides a way to approximate

$$P_0^{(m)}\{m_0 \le T \le m_1\}$$

as $b \to \infty$, $m_0 \to \infty$ and $m \to \infty$ in such a way that $m_i/m = t_i (i = 0, 1)$ and $b/\sqrt{m} = \mu_1$ for some $0 < t_0 < t_1 < 1$ and $\mu_1 > 0$.

Let

$$Q^{(m)}(A) = \int_{-\infty}^{\infty} P_\xi^{(m)}(A)\, d\xi/\sqrt{2\pi}.$$

An easy calculation shows that the likelihood ratio of $X_1, \dots, X_n$ under $Q^{(m)}$ relative to $P_0^{(m)}$ is

$$[m(m-n)/n]^{1/2} \exp\{S_n^2/[2n(1 - n/m)]\}$$

from which it follows from Wald's likelihood ratio identity that

$$\begin{aligned}
P_0^{(m)}&\{m_0 \le T \le m_1\} \\
&= m^{-1} \int_{-\infty}^{\infty} E_\xi^{(m)}\left\{\left[\frac{T}{1 - T/m}\right]^{1/2} \right. \\
&\qquad\qquad \left. \times \exp\left\{\frac{-S_T^2}{2T(1 - T/m)}\right\}; m_0 \le T \le m_1\right\} d\xi\Big/\sqrt{2\pi} \\
&= \int_{-\infty}^{\infty} E_{m\xi}^{(m)}\left\{\left[\frac{T}{1 - T/m}\right]^{1/2} \right. \\
&\qquad\qquad \left. \times \exp\left\{\frac{-S_T^2}{2T(1 - T/m)}\right\}; m_0 \le T \le m_1\right\} d\xi\Big/\sqrt{2\pi} \\
&= e^{-c/2} \int_{-\infty}^{\infty} E_{m\xi}^{(m)}\left\{\left[\frac{T}{1 - T/m}\right]^{1/2} \right. \\
&\qquad\qquad \left. \times \exp\left\{-H_2\left(S_T, \frac{T}{m}\right)\right\}; m_0 \le T \le m_1\right\} d\xi\Big/\sqrt{2\pi},
\end{aligned}$$

where

$$H_2(x,t) = \frac{1}{2}\left[\frac{x^2}{t(1-t)} - \mu_1^2\right].$$

Again, we need to know the asymptotic distributions of T/m and $H_2(S_T, T/m)$ under the conditional probability law $P_{m\xi}^{(m)}$ to obtain an approximation.

The example above is taken from Siegmund (1985). James, James and Siegmund (1988) considers the case of unknown variance. Siegmund (1988) obtains confidence sets for the change point when observations are taken from an exponential family. In both articles, the results of conditional renewal theory are also required [see the proof of the Theorem, James, James and Siegmund (1988) and the proof of Theorem B.1 of Siegmund (1988)].

3. Main results. Let $X, X_1, X_2, \dots$ be i.i.d. d-dimensional vectors with common density (1.1) and T_m be the stopping time defined by (1.5). Also let $\tilde{S}_n = \sum_{i=1}^n \tilde{X}_i$, where $\tilde{X}_i$ denotes the $(d+1)$-dimensional vector $(X_i, 1)$. Then T_m can be rewritten as

$$T_m = \inf\{n \geq m_0;\ mH(\tilde{S}_n/m) > 0\}.$$

In the sequel, we shall assume that $m_0/m = r > 0$ as $m \to \infty$. Let

$$\theta_0 = \left(0^{(1)}, \theta^{(2)}(\xi_0^{(2)})\right), \tag{3.1}$$

where $0^{(1)}$ denotes the zero vector of R^{d_1}. Also, let

$$\mu_0 = \mu(\theta_0). \tag{3.2}$$

For $\delta, a, b > 0$, let

$$\mathscr{C}(\delta, a, b) = \{(tv, t): a < t < b, \|v - \mu_0\| < \delta\} \tag{3.3}$$

denote a segment of a cone. Define

$$\mathscr{N} = \{(x,t): H(x,t) < 0\}.$$

Consider the following conditions on H:

1. $\mathscr{N} \supset \mathscr{C}(\delta, r_1, r_2)$, for some $r_1 < r < r_2$.
2. H is continuous in $\mathscr{C}(\delta, 0, 1)$.
3. $t_0 = \inf\{t > 0: H(t\tilde{\mu}) > 0\}$ exists and $r < t_0 < 1$, where

$$\tilde{\mu} = (\mu_0, 1). \tag{3.4}$$

4. H has continuous second partial derivatives in a neighborhood of $t_0\tilde{\mu}$ and $\gamma \cdot \tilde{\mu} > 0$, where

$$\gamma = \nabla H(t_0\tilde{\mu}). \tag{3.5}$$

Condition 1 insures that $\tilde{S}_n/m \in \mathscr{C}(\delta, r_1, r_2)$ for $n = r_1 m, \dots, r_2 m$ (which is true with high probability when m large) implies $H(\tilde{S}_n/m) < 0$. Only then it is meaningful to talk about the first $n \geq m_0$ such that $H(\tilde{S}_n/m)$ is positive. Condition 3 guarantees that (with high probability) the boundary crossing takes place before the condition time m, so that the problem is a genuine conditional boundary crossing problem. If condition 4 is true, then, by implicit function theorem, there exists a continuously differentiable function $h\colon R^d \mapsto R$ such that $h(t_0\mu_1, \dots, t_0\mu_d) = t_0$ and

$$H(x_1 \dots x_d, h(x_1, \dots, x_d)) = 0$$

for all $(x_1, \dots, x_d) \in \mathscr{O}$, where $\mathscr{O}$ is some neighborhood of $(t_0\mu_1, \dots, t_0\mu_d)$. Let

$$\mathscr{F} = \{(x_1, \dots, x_{d+1}) \colon h(x_1, \dots, x_d) = x_{d+1}\}$$

be a d-dimensional surface which is well defined over $\mathscr{O}$ and extended in some smooth way to R^{d+1}. Note that we may choose δ so small that $\mathscr{C}(\delta, 0, 1) \cap \mathscr{F} \subset \mathscr{O}$. In the sequel, we shall assume this. Let

$$T_c = \inf\{n \colon \tilde{S}_n \cdot \gamma > c\},$$

where γ is given by (3.5). The renewal theorem asserts that $P_{\theta_0}\{\tilde{S}_{T_c} \cdot \gamma - c \le x\}$ converges to a distribution function, say, $G_{\theta_0}(x)$, as $c \to \infty$.

THEOREM 3.1. *If conditions* 1–4 *hold, then*

$$\lim_{m \to \infty} P^{(m)}_{\xi^{(2)}}\left\{mH\left(\tilde{S}_{T_m}\right) \le x\right\} = G_{\theta_0}(x)$$

uniformly for $\xi_0^{(2)}$ *in compact subsets of* Γ_2.

THEOREM 3.2. *Let*

$$\mu_\gamma = \gamma \cdot \tilde{\mu}. \tag{3.6}$$

Suppose conditions 1–4 *are true. Then*

$$\lim_{m \to \infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T_m - t_0 m}{\sqrt{mt_0(1 - t_0)\sigma_0^2\mu_\gamma^{-2}}} \le y\right\} = \Phi(y), \tag{3.7}$$

where $\sigma_0^2 = \mathrm{Var}_{\theta_0}(\gamma \cdot \tilde{X})$ *and* $\Phi(y)$ *is the standard normal distribution function. Moreover* (3.7) *holds uniformly for* $\xi_0^{(2)}$ *in compact subsets of* Γ_2.

THEOREM 3.3. *Under condition* 1–4, *we have*

$$\lim_{m \to \infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T_m - t_0 m}{\sqrt{mt_0(1 - t_0)\sigma_0^2\mu_\gamma^{-2}}} \le y, mH\left(\tilde{S}_{T_m}\right) \le x\right\} = \Phi(y)G_{\theta_0}(x) \tag{3.8}$$

uniformly for $\xi_0^{(2)}$ *in compact subsets of* Γ_2.

4. Proof of Theorem 3.1. We begin by mentioning the following useful fact: From Lemma 4.5 and the Kolmogorov inequality applied to the $P_\xi^{(m)}$-martingale $(S_n - n\xi/m)/(1 - n/m)$, if follows that

$$\lim_{m \to \infty} P_\xi^{(m)}\{\|S_n - nS_m/m\| \ge \varepsilon \text{ for some } m_0 \le n \le m(1 - \varepsilon)\} = 0 \tag{4.1}$$

uniformly for ξ_0 in compact subsets of Γ. In view of (1.4), integrating (4.1) with respect to $P^{(m)}_{\xi^{(2)}}\{S^{(1)}_m \in d\xi^{(1)}\}$ gives

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{\|S_n - n\mu_0/m\| \ge n\varepsilon \text{ for some } m_0 \le n \le m(1-\varepsilon)\} = 0, \tag{4.2}$$

where μ_0 is given by (3.2).

Although the proof of Theorem 3.1 is very complicated and technical, the basic strategy is not difficult to explain.

Let

$$T^*_m = \inf\{n: \tilde{S}_n \text{ crosses } m\mathscr{F}\}.$$

It is easy to see that $T_m = T^*_m$ with high probability. In fact, if

$$\tilde{S}_n/m \in \{(tv, t): t > 0, \|v - \mu_0\| < \delta\}$$

for $n = m_0, \ldots, m$, then $T_m = T^*_m$. In what follows, we shall construct a surface $m\mathscr{F}^{[m^{1/3}]}$ which is inside and $O(m^{1/3})$ away from $m\mathscr{F}$. The idea is to condition on the time when the conditional random walk first crosses $m\mathscr{F}^{[m^{1/3}]}$ and then restart the conditional random walk from there. Lemma 4.3 shows that the restarted random walk will cross $m\mathscr{F}$ in a time of $O(m^{1/3})$. Then we can simplify the problem in two aspects. First, by Lemma 4.1, within time of $o(m^{1/2})$, a conditional random walk behaves like an unconditional one. Second, the part of $m\mathscr{F}$ where the boundary crossing takes place is, with high probability, a set of diameter $o(m^{1/2})$, which is well approximated by a hyperplane. Thus the asymptotic distribution of the excess over the boundary is the same as the excess over a hyperplane by an unconditional random walk. The proof of Theorem 3.1 is preceded by five lemmas.

LEMMA 4.1. *Suppose $n = o(m^{1/2})$. Let $\mathscr{E}_n$ denote the σ-field generated by $X_1, \ldots, X_n$. Then*

$$\lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} \sup_{A\in\mathscr{E}_n} |P^{(m)}_{\xi^{(2)}}\{A\} - P_{\theta_0}\{A\}| = 0,$$

where $K \subset \Gamma_2$ is a compact subset of R^{d_2} and θ_0 is given by (3.1).

PROOF. We need Lemma 4.2, whose proof will be given in Appendix A.

LEMMA 4.2. *Suppose $n = o(m^{1/2})$. Let $L_n(\theta_0)$ denote the likelihood ratio*

$$\frac{dP^{(m)}_{\xi^{(2)}}}{dP_{\theta_0}}(S_1, \ldots, S_n).$$

Then, on the set $B = \{S_n = o(m^{1/2})\}$,

$$\lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} |L_n(\theta_0) - 1| = 0. \tag{4.3}$$

Consequently,

$$L_n(\theta_0) \to 1 \quad a.e.\ P_{\theta_0}.$$

It follows from Wald's likelihood ratio identity that

$$
\begin{aligned}
&\lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} \sup_{A\in\mathscr{E}_n} |P_{\xi^{(2)}}^{(m)}\{A\} - P_{\theta_0}\{A\}| \\
&\quad = \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} \sup_{A\in\mathscr{E}_n} |E_{\theta_0}\{L_n(\theta_0); A\} - P_{\theta_0}\{A\}| \\
&\quad \le \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} \sup_{A\in\mathscr{E}_n} E_{\theta_0}\{|1 - L_n(\theta_0)|; A\} \\
&\quad \le \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} E_{\theta_0}\{|1 - L_n(\theta_0)|\} \qquad (4.4)\\
&\quad = \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} 2E_{\theta_0}\{[1 - L_n(\theta_0)]^+\} \\
&\quad = \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} 2\Big[E_{\theta_0}\{[1 - L_n(\theta_0)]^+; B\} + E_{\theta_0}\{[1 - L_n(\theta_0)]^+; B^c\}\Big] \\
&\quad \le \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} 2E_{\theta_0}\{[1 - L_n(\theta_0)]^+; B\} + \lim_{m\to\infty} \sup_{\xi^{(2)}\in mK} 2P_{\theta_0}\{B^c\}.
\end{aligned}
$$

In previous equalities, we have used $[1 - L_n(\theta_0)]^+$ to denote $\max\{[1 - L_n(\theta_0)], 0\}$. The first term on the RHS of (4.4) equals zero by (4.3). By SLLN, the second term can be shown to be zero. This completes the proof. □

REMARK. Although Lemma 4.1 is not difficult to prove, its importance can not be overemphasized. It shows that a conditional random walk behaves like an unconditional one locally as the condition time becomes remote. As we shall see later, the only time segment of a conditional random walk which is important in determining the excess over the boundary is $o(m^{1/2})$. Thus, by Lemma 4.1, the conditional random walk crosses the boundary as if it were an unconditional one. That is why the asymptotic distribution of the excess over the boundary is the same as an unconditional boundary crossing problem.

It will be convenient to use the notation

$$
\begin{aligned}
mD &= \{mx \colon x \in D\}, \\
\operatorname{dis}(C, D) &= \inf\{\|x - y\| \colon x \in C, y \in D\}, \\
m\mathscr{F}^{[k]} &= m\mathscr{F} - k\tilde{\mu}/\|\tilde{\mu}\|.
\end{aligned}
$$

For simplicity, we shall use $\mathscr{C}$ to denote $\mathscr{C}(\delta, 0, 1)$ [for definition, see (3.3)].

LEMMA 4.3. *Define*

$$
\tau = \inf\{n \colon \tilde{S}_n \text{ crosses } m\mathscr{F}^{[m^{1/3}]}\}.
$$

For any given $\varepsilon > 0$, *there exist* m_ε, $\alpha_1 > 0$ *and* $\mathscr{C}$ *such that*

$$
P_{\xi^{(2)}}^{(m)}\{m^{1/3} \ge \operatorname{dis}(\tilde{S}_\tau, m\mathscr{F} \cap m\mathscr{C}) \ge \alpha_1 m^{1/3}\} \ge 1 - \varepsilon,
$$

for all $m > m_\varepsilon$.

PROOF. Lemma 4.3 is a simple consequence of Lemmas 4.4 and 4.5, whose proofs will be given in Appendix A.

LEMMA 4.4. *There exist* $1 > \alpha_1 > 0$ *and* $\mathscr{C}$ *such that*

$$k \geq \operatorname{dis}(m\mathscr{F} \cap m\mathscr{C}, m\mathscr{F}^{[k]} \cap m\mathscr{C}) \geq \alpha_1 k. \tag{4.5}$$

LEMMA 4.5. *For any positive* k,

$$\lim_{m \to \infty} E_{\xi^{(2)}}^{(m)}\{\|X_1\|^k\} = E_{\theta_0}\{\|X_1\|^k\}$$

uniformly for $\xi_0^{(2)}$ *in compact subsets of* Γ_2.

Since with $P_{\xi^{(2)}}^{(m)}$ probability close to 1 the conditional random walk $\tilde{S}_n$ crosses $m\mathscr{F}$ within the cone $m\mathscr{C}$ as $m \to \infty$, it follows from Lemma 4.4 that

$$P_{\xi^{(2)}}^{(m)}\{m^{1/3} \geq \operatorname{dis}(\tilde{S}_\tau, m\mathscr{F} \cap m\mathscr{C})\} > 1 - \varepsilon/2 \tag{4.6}$$

for m sufficiently large. For the same reason,

$$\begin{aligned} &P_{\xi^{(2)}}^{(m)}\{\operatorname{dis}(\tilde{S}_\tau, m\mathscr{F} \cap m\mathscr{C}) \geq \alpha_1 m^{1/3} - \|\tilde{X}_\tau\|\} \\ &\quad \geq P_{\xi^{(2)}}^{(m)}\{\operatorname{dis}(\tilde{S}_\tau, m\mathscr{F} \cap m\mathscr{C}) \\ &\qquad\quad \geq \operatorname{dis}(m\mathscr{F} \cap m\mathscr{C}, m\mathscr{F}^{[m^{1/3}]} \cap m\mathscr{C}) - \|\tilde{X}_\tau\|\} \\ &\quad > 1 - \varepsilon/4. \end{aligned} \tag{4.7}$$

Since there exists a constant β such that $\tau \leq m\beta$ with $P_{\xi^{(2)}}^{(m)}$ probability close to 1 as $m \to \infty$, we have

$$\left| P_{\xi^{(2)}}^{(m)}\{\|\tilde{X}_\tau\| > m^\varepsilon\} - P_{\xi^{(2)}}^{(m)}\{\|\tilde{X}_\tau\| > m^\varepsilon; \tau \leq m\beta\} \right| < \varepsilon/4. \tag{4.8}$$

It is easy to see that

$$\begin{aligned} P_{\xi^{(2)}}^{(m)}\{\|\tilde{X}_\tau\| > m^\varepsilon; \tau \leq m\beta\} &\leq P_{\xi^{(2)}}^{(m)}\left\{\max_{1 \leq i \leq \beta m} \|\tilde{X}_i\| > m^\varepsilon\right\} \\ &\leq \beta m P_{\xi^{(2)}}^{(m)}\{\|\tilde{X}_1\| > m^\varepsilon\}. \end{aligned} \tag{4.9}$$

The last term of (4.9) tends to zero by Lemma 4.5 and Chebyshev's inequality. Hence, by (4.8),

$$P_{\xi^{(2)}}^{(m)}\{\|\tilde{X}_\tau\| > m^\varepsilon\} < \varepsilon/2. \tag{4.10}$$

Choose $\varepsilon < 1/3$, and in view of (4.6), (4.7) and (4.10), the proof is complete. □

PROOF OF THEOREM 3. Let

$$A_1 = \left\{mH\left(\tilde{S}_{T_m}\right) \leq x\right\}.$$

Denote by A_τ the event that the excess over the surface $m\mathscr{F}$ by the conditional random walk starting from $\tilde{S}_\tau$ is no greater than x. That is,

$$A_\tau = \left\{ mH\left(\frac{S'_{\tau_1} + \tilde{S}_\tau}{m} \right) \le x \right\},$$

where

$$S'_n = \tilde{S}_{\tau+n} - \tilde{S}_\tau, \qquad \tau_1 = \inf\{n > 0: S'_n + \tilde{S}_\tau \text{ crosses } m\mathscr{F}\}.$$

It is easily seen that

$$(4.11) \qquad \begin{aligned} \lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{A_1\} &= \lim_{m\to\infty} E^{(m)}_{\xi^{(2)}}\{P^{(m)}_{\xi^{(2)}}[A_1 | \mathscr{E}_\tau]\} \\ &= \lim_{m\to\infty} E^{(m)}_{\xi^{(2)}}\{E^{(m)}_{\xi^{(2)}}[P^{(m-\tau)}_{\zeta^{(2)}}(A_\tau) | \mathscr{E}_\tau]\}, \end{aligned}$$

where $\mathscr{E}_\tau$ is the σ-field generated by events prior to τ, $\zeta^{(2)} = \xi^{(2)} - S^{(2)}_\tau$. Set $\zeta^{(2)}_0 = \zeta^{(2)}/(m-\tau)$. It follows from (4.2) and Lemma 4.3 that, for $t_0 < t < 1$ and $\delta > 0$, the event

$$B_\delta = \{|\zeta^{(2)}_0 - \xi^{(2)}_0| \le \delta, |\tilde{S}_\tau/m - t_0\tilde{\mu}| \le \delta, \tau \le mt\}$$
$$\cap \{\operatorname{dis}(\tilde{S}_\tau, m\mathscr{F} \cap m\mathscr{C}) = O(m^{1/3})\}$$

has $P^{(m)}_{\xi^{(2)}}$ probability tending to 1 as $m \to \infty$. Hence the last term of (4.11) equals

$$(4.12) \qquad \lim_{m\to\infty} E^{(m)}_{\xi^{(2)}}\{P^{(m-\tau)}_{\zeta^{(2)}}[A_\tau]; B_\delta\}.$$

Let $\tilde{R}_m$ be such that $S'_{\tau_1} + \tilde{S}_\tau - \tilde{R}_m \in m\mathscr{F}$ and $\|\tilde{R}_m\| = \operatorname{dis}(S'_{\tau_1} + \tilde{S}_\tau, m\mathscr{F})$. A Taylor expansion shows that

$$(4.13) \qquad mH\left(\frac{\tilde{S}_{\tau+\tau_1}}{m}\right) = mH\left(\frac{\tilde{S}_{\tau+\tau_1} - \tilde{R}_m}{m}\right) + \nabla H\left(\frac{\tilde{v}_m}{m}\right) \cdot \tilde{R}_m,$$

where $\tilde{v}_m$ is a point on the line segment joining $S'_{\tau_1} + \tilde{S}_\tau$ and $S'_{\tau_1} + \tilde{S}_\tau - \tilde{R}_m$. The first term on the RHS of (4.13) vanishes by the definition of $m\mathscr{F}$. Thus the excess over the boundary equals the second term, which is the excess in the normal direction.

Define $B_1 = \{\tau_1 = O(m^{1/3})\}$. By Lemma 4.1, we have

$$(4.14) \qquad \lim_{m\to\infty} \sup_{|\zeta^{(2)}_0 - \xi^{(2)}_0| \le \delta} |P^{(m-\tau)}_{\zeta^{(2)}}\{B_1\} - 1| = 0 \quad \text{on } B_\delta.$$

Hence the subset of $m\mathscr{F}$ on which the boundary crossing takes place is of diameter $O(m^{1/3})$ with high $P^{(m-\tau)}_{\zeta^{(2)}}$ probability. Let

$$\theta_\zeta = \left(0^{(1)}, \theta^{(2)}(\zeta^{(2)}_0)\right).$$

Let $\tilde{u} \in m\mathscr{F}$ be the point on the ray $\{\tilde{S}_\tau + a\tilde{\mu}_\zeta: a > 0\}$, where $\tilde{\mu}_\zeta = (\mu(\theta_\zeta), 1)$.

Let u denote the vector in R^d with the same components as the first d components of $\tilde{u}$. Now we can compare $m\mathscr{F}$ with its tangent plant at $\tilde{u}$ in a neighborhood of diameter $O(m^{1/3})$ via a Taylor expansion

$$(4.15)\quad mh[(x+u)/m] = mh(u/m) + \nabla h(u/m)\cdot x + O(\|x\|^2/m),$$

where $\|x\| = O(m^{1/3})$. Hence the random walk $S_n' + \tilde{S}_\tau$ crosses $m\mathscr{F}$ as if it were a hyperplane which is $O(m^{1/3})$ away,

$$\mathscr{H} = \{\tilde{x} \in R^{d+1}: \tilde{x}\cdot\gamma_\tau = c\},$$

where

$$(4.16)\quad \begin{aligned} \gamma_\tau &= \nabla H(\tilde{u}/m),\\ c &= (\tilde{u} - \tilde{S}_\tau)\cdot\gamma_\tau = a\tilde{\mu}_\zeta\cdot\gamma_\tau = O_p(m^{1/3}). \end{aligned}$$

S_n' crosses the hyperplane $\mathscr{H}$ when the one-dimensional random walk

$$(4.17)\quad U_n = S_n'\cdot\gamma_\tau$$

crosses the level c. We have previously shown that the excess of interest is in the normal direction. In case of a hyperplane, it equals to $U_{\tau_c} - c$, where

$$\tau_c = \inf\{n: U_n > c\}.$$

Clearly, $A_\tau \cap B_1$ belongs to the σ-field generated by $S_1', \dots, S_n'$ with $n = o(m^{1/2})$. By Lemma 4.1, (4.14) and arguments above, (4.12) becomes

$$(4.18)\quad \begin{aligned} \lim_{m\to\infty} E_{\xi(2)}^{(m)}\{P_{\zeta(2)}^{(m-\tau)}[A_\tau]; B_\delta\} &= \lim_{m\to\infty} E_{\xi(2)}^{(m)}\{P_{\zeta(2)}^{(m-\tau)}[A_\tau \cap B_1]; B_\delta\}\\ &= \lim_{m\to\infty} E_{\xi(2)}^{(m)}\{P_{\theta_\zeta}[A_c]; B_\delta\}, \end{aligned}$$

where $A_c = \{U_{\tau_c} - c \le x\}$.

Standard renewal theory asserts that $U_{\tau_c} - c$ has asymptotic distribution with distribution function, say, $G_{\theta_\zeta}(x)$ as $c \to \infty$. Hence (however see the remark below), the last term of (4.18) becomes

$$(4.19)\quad \lim_{m\to\infty} E_{\xi(2)}^{(m)}\{P_{\theta_\zeta}[A_c]; B_\delta\} = \lim_{m\to\infty} E_{\xi(2)}^{(m)}\{G_{\theta_\zeta}(x); B_\delta\}.$$

The RHS of (4.19) can be made arbitrarily close to $G_{\theta_0}(x)$ by letting $\delta \to 0$. This completes the proof. □

REMARK. Although the arguments which lead to (4.19) are intuitively clear, there are two technical difficulties one has to overcome in order to obtain a rigorous proof.

LEMMA 4.6. *The excess over the surface* $m\mathscr{F}$ *by the random walk* $\tilde{S}_{\tau+n}$ *is asymptotically the same as over the hyperplane* $\mathscr{H}$ *under* P_{θ_ζ}.

LEMMA 4.7. *$P_{\theta_\zeta}[A_c]$ converges uniformly (in $\zeta^{(2)}$ and γ_τ) to $G_{\theta_\zeta}(x)$.*

These two issues will be settled in Appendix B.

5. Proof of Theorems 3.2 and 3.3.

PROOF OF THEOREM 3.2. We shall first treat the case $\xi^{(2)} = \xi$, or equivalently $d_1 = 0$. The proof is based on the following result.

PROPOSITION 5.1 [Borisov (1978)]. *Let Y_i, $i = 1, 2, \dots$, be a sequence of i.i.d. absolute continuous random variables such that $EY_1 = 0$, $\mathrm{Var}(Y_1) = 1$ and the moment generating function $E(e^{tY_1})$ is finite in a neighborhood of zero. Also let $V_n = \sum_{i=1}^n Y_i$ and $W_n(t)$ be the random polygonal curve with vertices at $(i/n, V_n/\sqrt{n})$, that is,*

$$W_n(t) = V_i/\sqrt{n} + \sqrt{n}\,(t - i/n)Y_{i+1}, \quad \text{for } i/n \le t \le (i+1)/n.$$

Let $W_n^0(t)$ be the random process with the same (finite-dimensional) distribution as $W_n(t)$ conditioning on $W_n(1) = a$. Let $W^0(t)$ denote the Brownian bridge tied down at $W^0(1) = a$. Then there exits a representation of the process $W_n^0(t)$ on the same probability space as the process $W^0(t)$ such that

$$P\left(\sup_{0 \le t \le 1} |W_n^0(t) - W^0(t)| > \frac{C \log n + x}{\sqrt{n}} \right) \le Ke^{-\lambda x},$$

where C, K and λ are positive constants depending only on the common distribution of Y_i. Moreover, the result holds uniformly in a neighborhood of a.

Define

$$\mathscr{H}^{[m]} = \left\{x \in R^{d+1} \colon x \cdot \gamma = \left(mt_0 - m^{1/3}\right)\mu_\gamma\right\}$$

and

$$T' = \inf\{n \colon U_n' \ge c\},$$

where $c = (mt_0 - m^{1/3})\mu_\gamma$ and $U_n' = \tilde{S}_n \cdot \mu_\gamma$ [μ_γ is given by (3.6)].

From Proposition 5.1, it follows that, with high $P_\xi^{(m)}$ probability, the deviation of the conditional random walk $\tilde{S}_n$ from the line of drift is $O(m^{1/2})$. Hence, with high probability, the subset $\mathscr{A}$ of $m\mathscr{F}$, where the boundary crossing takes place, is a $O(m^{1/2})$ neighborhood of $mt_0\tilde{\mu}$. By (4.15), with high probability, U_n' crosses $\mathscr{H}^{[m]}$ before $\tilde{S}_n$ crosses $m\mathscr{F}$. On the other hand, the distance between $\mathscr{A}$ and $\mathscr{H}^{[m]}$ is $O(m^{1/3})$. Thus,

$$|T' - T_m| = O_p(m^{1/3}) = o_p(m^{1/2}). \tag{5.1}$$

In view of (5.1), we can replace T_m by T' in (3.7) without changing the limit.

It is obvious that

$$\frac{\mu_\gamma T' - c}{\sqrt{mt_0(1-t_0)\sigma_0^2}} = \frac{\mu_\gamma T' - U'_{T'}}{\sqrt{mt_0(1-t_0)\sigma_0^2}} + \frac{U'_{T'} - c}{\sqrt{mt_0(1-t_0)\sigma_0^2}}.$$

Note that $U'_{T'} - c$ is the excess over the boundary at the stopping time T', which is $O_p(1)$ by Theorem 3.1. This shows that it suffices to prove

$$\lim_{m\to\infty} P_\xi^{(m)}\left\{\frac{\mu_\gamma T' - U'_{T'}}{\sqrt{mt_0(1-t_0)\sigma_0^2}} \le x\right\} = \Phi(x).$$

Let

$$W_n = (U'_n - n\mu_\gamma)/\sqrt{m}.$$

Then

$$P_\xi^{(m)}\left\{\frac{\mu_\gamma T' - U'_{T'}}{\sqrt{mt_0(1-t_0)\sigma_0^2}} \le x\right\} = P\left\{\frac{-W_{T'}}{\sqrt{t_0(1-t_0)\sigma_0^2}} \le x | W_1 = 0\right\}.$$

Put $t_1 = (1-\delta)t_0 m, t_2 = (1+\delta)t_0 m$. Observe that

$$(5.2)\quad \{|W_{T'} - W_{t_0 m}| < \varepsilon\} \supset \{T' \in [t_1, t_2]\} \cap \left\{\sup_{t_1 \le n \le t_2} |W_n - W_{mt_0}| < \varepsilon\right\}.$$

It is easy to see that the conditional probability of the first set on the RHS of (5.2) tends to 1, while that of the second set converges to

$$P\left\{\sup_{(1-\delta)t_0 \le t \le (1+\delta)t_0} |W^0(t) - W^0(t_0)| \le \varepsilon\right\},$$

which tends to 1 as $\delta \to 0$. ε being arbitrary, it follows that $W_{T'}$ has the same asymptotic distribution as $W_{t_0 m}$, by Proposition 5.1, which is asymptotically distributed as $N(0, t_0(1-t_0)\sigma_0^2)$. This completes the proof of Theorem 3.2 when $\xi^{(2)} = \xi$, or equivalently $d_1 = 0$.

We now turn to the case $d_1 \ne 0$. Replace ξ_0 by S_m/m everywhere in defining the event

$$A = \left\{\frac{T_m - t_0 m}{\sqrt{mt_0(1-t_0)\sigma_0^2\mu_\gamma^{-2}}} \le y\right\}.$$

Thus the event A remains the same while we integrate out $\xi^{(1)}$ with respect to $P_{\xi^{(2)}}^{(m)}\{S_m^{(1)} \in d\xi^{(1)}\}$. Theorem 3.2 now follows from (1.4). □

PROOF OF THEOREM 3.3. The proof of Theorem 3.3 will only be sketched, since all relevant techniques have appeared in previous proofs. By the same argument as in the proof of Theorem 3.2, it suffices to prove (3.8) with T_m replaced by T'. Restarting the random walk from $\tilde{S}_{T'}$, it only takes the restarted random walk $\tilde{S}_{T'+n} O_p(m^{1/3})$ extra steps to cross $m\mathscr{F}$. Theorem 3.1 shows that the boundary crossing behavior is asymptotically the same as an unconditional random walk.

Let

$$A = \left\{ \frac{T' - t_0 m}{\sqrt{m t_0 (1 - t_0) \sigma_0^2 \mu_\gamma^{-2}}} \le y,\ mH\left(\tilde{S}_{T_m}\right) \le x \right\},$$

$$A_1 = \left\{ \frac{T' - t_0 m}{\sqrt{m t_0 (1 - t_0) \sigma_0^2 \mu_\gamma^{-2}}} \le y \right\} \cap \{T' < T_m\},$$

$$A_{T'} = \left\{ mH\left(\frac{S'_{T''} + \tilde{S}_{T'}}{m} \right) \le x \right\},$$

where

$$T'' = \inf\left\{ n > 0 \colon S'_n + \tilde{S}_{T'} \text{ crosses } m\mathscr{F} \right\}.$$

Then

$$\begin{aligned}
\lim_{m \to \infty} P_{\xi^{(2)}}^{(m)}\{A\} &= \lim_{m \to \infty} E_{\xi^{(2)}}^{(m)}\left\{ P_{\xi^{(2)}}^{(m)}\{A | \mathscr{E}_{T'}\} \right\} \\
&= \lim_{m \to \infty} E_{\xi^{(2)}}^{(m)}\left\{ E_{\xi^{(2)}}^{(m)}\left[P_{\zeta^{(2)}}^{(m - T')}(A_{T'}) | \mathscr{E}_{T'} \right]; A_1 \right\} \\
&= \lim_{m \to \infty} E_{\xi^{(2)}}^{(m)}\{G_{\theta_0}(x); A_1\} \\
&= \Phi(y) G_{\theta_0}(x).
\end{aligned} \tag{5.3}$$

□

APPENDIX A

Proof of Lemmas 4.2, 4.4 and 4.5.

PROOF OF LEMMA 4.2. We need the following result.

PROPOSITION A.1 [Borovkov and Rogozin (1965)]. *Let f_n be the density of S_n under P_0. Then*

$$f_n(nx) = (2\pi n)^{-d/2} |\Sigma(x)|^{-1/2} \exp[-n\phi(x)] \left(1 + O\left(\frac{1}{n} \right) \right), \tag{A.1}$$

where ϕ is given by (2.1). *Moreover the limit above is attained uniformly over compact sets of Γ.*

Let $f^{[2]}$ be as in (1.3). Then

$$\frac{dP^{(m)}_{\xi^{(2)}}}{dP_{\theta_0}}(S_1,\dots,S_n)$$

$$\text{(A.2)}\qquad = \frac{f(X_1)f(X_2)\cdots f(X_n)f^{[2]}_{m-n}(\xi^{(2)}-S^{(2)}_n)/f^{[2]}_m(\xi^{(2)})}{f_{\theta_0}(X_1)f_{\theta_0}(X_2)\cdots f_{\theta_0}(X_n)}$$

$$= \frac{f^{[2]}_{m-n}(\xi^{(2)}-S^{(2)}_n)}{f^{[2]}_m(\xi^{(2)})\exp\{S^{(2)}_n\cdot\theta^{(2)}_0 - n\psi_2(\theta^{(2)}_0)\}}.$$

By (A.1), the last term of (A.2) tends to

$$\text{(A.3)}\qquad \begin{aligned}&[m/(m-n)]^{d_2/2}|\Sigma_2(\xi^{(2)}/m)|^{1/2}|\Sigma_2[(\xi^{(2)}-S^{(2)}_n)/(m-n)]|^{-1/2}\\ &\times\exp\{-(m-n)\phi_2[(\xi^{(2)}-S^{(2)}_n)/(m-n)]\}\\ &\times\exp[m\phi_2(\xi^{(2)}/m)]\exp[-S^{(2)}_n\cdot\theta^{(2)}_0+n\psi_2(\theta^{(2)}_0)],\end{aligned}$$

where Σ_2 denotes the covariance matrix of $X^{(2)}$. A Taylor expansion, for $S_n \in B$ and $\xi^{(2)} \in mK$, gives

$$\begin{aligned}&\phi_2[(\xi^{(2)}-S^{(2)}_n)/(m-n)]\\ &\quad=\phi_2[\xi^{(2)}/m]+[(n\xi^{(2)}/m-S^{(2)}_n)/(m-n)]\nabla\phi_2(\xi^{(2)}/m)\\ &\qquad+o(m^{-1})\quad\text{a.s. } P_{\theta_0}.\end{aligned}$$

Observing that $\nabla\phi_2(\xi^{(2)}/m) = \theta^{(2)}_0$ we have

$$\text{(A.4)}\qquad \begin{aligned}&(m-n)\phi_2[(\xi^{(2)}-S^{(2)}_n)/(m-n)]\\ &\quad=(m-n)\phi_2[\xi^{(2)}/m]+[(n\xi^{(2)}/m-S^{(2)}_n)]\cdot\theta^{(2)}_0+o(1)\quad\text{a.s. } P_{\theta_0}\\ &\quad=m\phi_2[\xi^{(2)}/m]-[S^{(2)}_n\cdot\theta^{(2)}_0-n\psi_2(\theta^{(2)}_0)]+o(1)\quad\text{a.s. } P_{\theta_0}.\end{aligned}$$

Substituting (A.4) into (A.3) completes the proof. □

PROOF OF LEMMA 4.4. The first inequality of (4.5) follows trivially from the definition of $m\mathscr{F}^{[k]}$. We proceed to prove the second inequality. Since the inequality is invariant under rotation, we may choose the coordinate system to be such that $\tilde{\mu} = (0,\dots,0,\|\tilde{\mu}\|)$. Let $M = \sup_{z\in\mathscr{O}}\|\nabla h(z)\|$ (for definition of $\mathscr{O}$, see Section 3). If $M = 0$, then h is constant over $\mathscr{O}$ and Lemma 4.4 holds trivially. For $M \neq 0$, let $\varepsilon = (2M)^{-1}$. Clearly,

$$\begin{aligned}&[\operatorname{dis}(m\mathscr{F}\cap m\mathscr{C}, m\mathscr{F}^{[k]}\cap m\mathscr{C})]^2\\ &\quad\geq \inf_{x,y\in\mathscr{F}\cap\mathscr{C}}[\|m(x-y)-k\mu/\|\mu\|\,\|^2]\\ &\quad= \inf_{x,y\in\mathscr{F}\cap\mathscr{C}}[m^2\|x-y\|^2-2mk(x-y)\cdot(\mu/\|\mu\|)+k^2].\end{aligned}$$

Note that

$$
\text{(A.5)}\quad
\begin{aligned}
&m^2\|x-y\| - 2mk(x-y)\cdot(\mu/\|\mu\|) + k^2\\
&\quad= m^2\Big[\|(x_1,\dots,x_d)-(y_1,\dots,y_d)\|^2\\
&\qquad\quad + |h(x_1,\dots,x_d)-h(y_1,\dots,y_d)|^2\Big]\\
&\qquad - 2mk\big[h(x_1,\dots,x_d)-h(y_1,\dots,y_d)\big] + k^2\\
&\quad= m^2\|(x_1,\dots,x_d)-(y_1,\dots,y_d)\|^2\\
&\qquad + \big\{m\big[h(x_1,\dots,x_d)-h(y_1,\dots,y_d)\big]-k\big\}^2,
\end{aligned}
$$

where the first equality follows from $x\cdot\mu/\|\mu\| = h(x_1,\dots,x_d)$. Let

$$
\begin{aligned}
\varepsilon_1 = \inf\{&\|m(x-y)-k\mu/\|\mu\|\,\|: x,y\in\mathscr{F}\cap\mathscr{C},\\
&\|(x_1,\dots,x_d)-(y_1,\dots,y_d)\| > \varepsilon k/m\},\\
\varepsilon_2 = \inf\{&\|m(x-y)-k\mu/\|\mu\|\,\|: x,y\in\mathscr{F}\cap\mathscr{C},\\
&\|(x_1,\dots,x_d)-(y_1,\dots,y_d)\| \le \varepsilon k/m\}.
\end{aligned}
$$

By (A.5), $\varepsilon_1 \ge \varepsilon k$. By the mean value theorem,

$$|h(x_1,\dots,x_d)-h(y_1,\dots,y_d)| \le M\|(x_1,\dots,x_d)-(y_1,\dots,y_d)\|.$$

Hence $\varepsilon_2 \ge k - \varepsilon Mk = k/2$. It follows that $\operatorname{dis}(m\mathscr{F}\cap m\mathscr{C}, m\mathscr{F}^{[k]}\cap m\mathscr{C}) \ge \min(\varepsilon,\frac{1}{2})k$. This completes the proof. □

PROOF OF LEMMA 4.5. By Proposition A.1 and arguments similar to that of Lemma 4.1, one can show that the conditional density of X_1 given $S_m^{(2)} = \xi^{(2)}$ is

$$\text{(A.6)}\qquad \frac{f(x)\,f_{m-1}^{[2]}(\xi^{(2)}-x^{(2)})}{f_m^{[2]}(\xi^{(2)})} = \left(1 + \sum_{i=0}^{2}\|x\|^i O\left(\frac{1}{m}\right)\right) f_{\theta_0}(x)$$

uniformly for $x = o(m^{1/2})$.

Put $a_m = m^{1/3}$. For any positive integer k,

$$
\begin{aligned}
E_{\xi^{(2)}}^{(m)}\big(\|X_1\|^k\big) &= \int \|x\|^2 \frac{f(x)\,f_{m-1}^{[2]}(\xi^{(2)}-x^{(2)})}{f_m^{[2]}(\xi^{(2)})}\,dx\\
&= \int_{\|x\|\le a_m} + \int_{\|x\|>a_m}\\
&= I_1 + I_2.
\end{aligned}
$$

From (A.6),

$$\left| I_1 - \int \|x_1\|^k f_{\theta_0}(x)\,dx \right| \le \int_{\|x_1\| > a_m} \|x_1\|^k f_{\theta_0}(x)\,dx + O(1/m) \int_{\|x\| \le a_m} \sum_{i=0}^{2} \|x\|^i f_{\theta_0}(x)\,dx. \tag{A.7}$$

By the assumptions made in Section 1, $E_{\theta_0}(\|X_1\|^k) < \infty$. Hence the RHS of (A.7) tends to 0 as $m \to \infty$. We turn to the estimates of I_2. We may write the conditional density of X_1 given $S_m^{(2)} = \xi^{(2)}$ as

$$\frac{f_{\theta_0}(x)\, f_{\theta_0, m-1}^{[2]}(\xi^{(2)} - x^{(2)})}{f_{\theta_0, m}^{[2]}(\xi^{(2)})},$$

where $f_{\theta,n}^{[2]}$ denotes the n-fold convolution of the density $f_\theta^{[2]}$ with itself. By (1.2), there exist a constant M such that $\sup_x f_{\theta_0, m-1}^{[2]}(x) < M$, for m sufficiently large. From Proposition A.1, it follows that

$$f_{\theta_0, m}^{[2]}(\xi^{(2)}) \sim (2\pi m)^{-1/2} |\Sigma_2(\xi/m)|^{-1/2}.$$

Hence

$$\begin{aligned} I_2 &= \int_{\|x\| > a_n} \|x\|^k \frac{f_{\theta_0}(x)\, f_{\theta_0, m-1}^{[2]}(\xi^{(2)} - x^{(2)})}{f_{\theta_0, m}^{[2]}(\xi^{(2)})}\,dx \\ &\le 2M[2\pi m |\Sigma(\xi/m)|]^{1/2} \int_{\|x\| > a_m} \|x\|^k f_{\theta_0}(x)\,dx \\ &\le 2M[2\pi m |\Sigma(\xi/m)|]^{1/2} \exp(-t a_m/2)\, E_{\theta_0}\{\exp(t\|X_1\|)\} \end{aligned} \tag{A.8}$$

for some $t > 0$ and m sufficiently large. By the assumptions made in Section 1, $E_{\theta_0}\{\exp(t\|X_1\|)\} < \infty$. So the RHS of (A.8) tends to 0 as $m \to \infty$. From (A.7) and (A.8), it follows that $E_{\xi^{(2)}}^{(m)}(\|X\|^k)$ converges to $E_{\theta_0}(\|X\|^k)$ uniformly for $\xi_0^{(2)}$ in compact subsets of $\Gamma^{(2)}$. This completes the proof. □

APPENDIX B

Proof of Lemmas 4.6 and 4.7.

PROOF OF LEMMA 4.6. Let

$$\tau_\varepsilon = \inf\{n > 0 : \tilde{S}_{\tau+n} - \tilde{S}_\tau \text{ crosses } \mathscr{H}^{[\varepsilon]}\},$$

$$\tau_{-\varepsilon} = \inf\{n > 0 : \tilde{S}_{\tau+n} - \tilde{S}_\tau \text{ crosses } \mathscr{H}^{[-\varepsilon]}\},$$

where $\mathscr{H}^{[\pm\varepsilon]} = \mathscr{H} \pm \varepsilon\gamma_\tau/\|\gamma_\tau\|$, with γ_τ given by (4.16). It follows from (4.15) that, with high P_{θ_ζ} probability, the part of $m\mathscr{F} - \tilde{S}_\tau$ where the boundary crossing takes place is caught between $\mathscr{H}^{[-\varepsilon]}$ and $\mathscr{H}^{[+\varepsilon]}$, for given $\varepsilon > 0$ and m sufficiently large. The proof now proceeds like that of Lai and Siegmund

(1977) [also see Siegmund (1985) Section 9.2]. We have

$$\begin{aligned} &P_{\theta_\zeta}\{U_{\tau_\varepsilon} - (c+\varepsilon) \le x - \varepsilon\} - \varepsilon \\ &\quad \le P_{\theta_\zeta}\{A_\tau\} \le P_{\theta_\zeta}\{U_{\tau_{-\varepsilon}} - (c-\varepsilon) \le x + \varepsilon\} + \varepsilon, \end{aligned} \tag{B.1}$$

on B_δ, for m sufficiently large. The proof is completed by letting $\varepsilon \to 0$. □

Proof of Lemma 4.7. To establish the uniform covergence of $P_{\theta_\zeta}\{A_\tau\}$, it is sufficient, from (B.1), to show the uniform covergence of

$$\lim_{c\to\infty} P_{\theta_\zeta}\{U_{\tau_c} - c \le x\}. \tag{B.2}$$

Recall that the convergence of (B.2) follows from the renewal theorem, which says that the renewal measure is asymptotically the same as Lebesgue measure times a constant. Thus the uniform convergence of (B.2) can be obtained from the uniformity of the renewal theorem. We now proceed to show the latter. The starting point is Stone (1965). Let F be a nonlattice probability distribution function with finite third moment such that

$$\mu = \int_{-\infty}^{\infty} x\,dF(x) > 0, \qquad \nu = \int_{-\infty}^{\infty} x^2\,dF(x), \qquad \kappa = \int_{-\infty}^{\infty} |x|^3\,dF(x).$$

Denote by $F^{(n)}$ the n-fold convolution of F with itself. Let f be the characteristic function of F. Also, let $K(x)$ be a symmetric probability density function, with characteristic function $k(\theta)$, such that $K(x)$ has finite third moment and $k(\theta) = 0$ for $|\theta| \ge 1$. Define

$$U(x,h) = \sum_{n=0}^{\infty} \left[F^{(n)}(x+h/2) - F^{(n)}(x-h/2)\right]$$

and

$$V(x,h,a) = \int_{-\infty}^{\infty} a^{-1}K(a^{-1}y)U(x-y,h)\,dy.$$

Let $h > 0$ be fixed and let $N < \infty$ be an upper bound to $U(x,h)$. Choose $\varepsilon(0 < \varepsilon \le \frac{1}{2})$ and a_0 such that

$$\int_{|y|\ge \varepsilon h} a_0^{-1}K(a_0^{-1}y)\,dy \le \varepsilon.$$

Stone (1965) shows that

$$V(x,(t-2\varepsilon)h,a_0) - N\varepsilon \le U(x,h) \le (1-\varepsilon)^{-1}V(x,(1+\varepsilon)h,a_0) \tag{B.3}$$

and

$$\begin{aligned} V(x,h,a) = \frac{h}{\mu} + \frac{h}{2\pi}\int_{-a^{-1}}^{a^{-1}} \Re\Bigg\{ e^{-ix\theta}\left(\frac{\sin h\theta/2}{h\theta/2}\right)k(a\theta) \\ \times\left(\frac{1}{1-f(\theta)} - \frac{i}{\mu\theta}\right)\Bigg\}\,d\theta + \frac{o_x(1)}{\mu x^2}, \end{aligned} \tag{B.4}$$

where $o_x(1) \to 0$ as $x \to \infty$ uniformly for a and h in bounded sets and does not depend on f; $\Re(z)$ denotes the real part of a complex number z. The renewal theorem says that

$$\lim_{x\to\infty} U(x,h) = h/\mu. \tag{B.5}$$

It follows from (B.3) and (B.4) that, to show (B.5) holds uniformly over a class $\mathscr{L}$ of distributions, it is sufficient to show that

$$\int_{-a^{-1}}^{a^{-1}} e^{-ix\theta}\left(\frac{\sin h\theta/2}{h\theta/2}\right)k(a\theta)\left(\frac{1}{1-f(\theta)} - \frac{i}{\mu\theta}\right)d\theta \to 0 \quad \text{uniformly over } \mathscr{L} \tag{B.6}$$

as $x \to \infty$. The following theorem is all we need to finish the proof.

THEOREM B.1. *Let $\mathscr{L} = \{F_\alpha\}_{\alpha\in C}$, where C is some index set, be a class of distribution functions such that the following conditions hold.*

(i) *Each F_α is nonlattice and has finite third moment.*
(ii) *The first moment μ_α and the third moment κ_α of F_α satisfy*

$$\inf_{\alpha\in C}\mu_\alpha > 0, \qquad \sup_{\alpha\in C}\kappa_\alpha < \infty.$$

(iii) *For any sequence of distribution functions $\{F_{\alpha_i}\} \in \mathscr{L}$, there exists a subsequence $\{F_{\alpha_i'}\}$ that converges in distribution to a nonlattice distribution function.*

Then

$$\lim_{x\to\infty}\sup_{\alpha\in C}\left|U_\alpha(x,h) - \frac{h}{\mu_\alpha}\right| = 0.$$

PROOF OF THEOREM B.1. The proof of Theorem B.1 will only be sketched. Integrating the LHS of (B.6) by parts gives

$$\begin{aligned}&\frac{e^{-ix\theta}}{-ix}\left(\frac{\sin h\theta/2}{h\theta/2}\right)k(a\theta)\left(\frac{1}{1-f_\alpha(\theta)} - \frac{i}{\mu_\alpha\theta}\right)\Bigg|_{a^{-1}}^{a^{-1}}\\ &+\int_{-a^{-1}}^{a^{-1}}\frac{e^{-ix\theta}}{ix}\frac{d}{d\theta}\left[\left(\frac{\sin h\theta/2}{h\theta/2}\right)k(a\theta)\left(\frac{1}{1-f_\alpha(\theta)} - \frac{i}{\mu_\alpha\theta}\right)\right]d\theta.\end{aligned} \tag{B.7}$$

From (B.7), it is clear that to show (B.6) it is sufficient to prove

$$\sup_{\alpha\in C}\sup_{|\theta|\le a^{-1}}\left|\frac{1}{1-f_\alpha(\theta)} - \frac{i}{\mu_\alpha\theta}\right| < \infty \tag{B.8}$$

and

$$\sup_{\alpha \in C} \sup_{|\theta| \le a^{-1}} \left| \frac{d}{d\theta} \left[\frac{1}{1 - f_\alpha(\theta)} - \frac{i}{\mu_\alpha \theta} \right] \right| < \infty. \tag{B.9}$$

We sketch the proof of (B.9), that of (B.8) being similar and easier. Using (iii), one can show that, for some $\delta > 0$,

$$\sup_{\alpha \in C} \sup_{\delta \le |\theta| \le a^{-1}} \left| \frac{d}{d\theta} \left[\frac{1}{1 - f_\alpha(\theta)} - \frac{i}{\mu_\alpha \theta} \right] \right|$$
$$= \sup_{\alpha \in C} \sup_{\delta \le |\theta| \le a^{-1}} \left| \frac{f_\alpha'(\theta)}{[1 - f_\alpha(\theta)]^2} + \frac{i}{\mu_\alpha \theta^2} \right| < \infty.$$

Hence it suffices to show that

$$\sup_{\alpha \in C} \sup_{|\theta| \le \delta} \left| \frac{f_\alpha'(\theta)}{[1 - f_\alpha(\theta)]^2} + \frac{i}{\mu_\alpha \theta^2} \right| < \infty,$$

for some $\delta > 0$. By standard arguments, one can show that

$$\sup_{\alpha \in C} \sup_{|\theta| \le \delta} \left| \frac{f_\alpha'(\theta)}{[1 - f_\alpha(\theta)]^2} + \frac{i}{\mu_\alpha \theta^2} \right| < \sup_{\alpha \in C} \frac{4k_\alpha}{\mu_\alpha^2}, \tag{B.10}$$

where $\delta = \inf_{\alpha \in C} \mu_\alpha / \nu_\alpha > 0$ whose existence is guaranteed by condition (ii). By condition (ii) again, the RHS of (B.13) is less than infinity. This concludes the proof of Theorem B.1. □

REMARK. Kartashov (1980) also obtains a uniform renewal theorem along with convergence rate. However, his assumptions are stronger than that of Theorem B.1.

The idea now is to apply Theorem B.1 to the distributions of the first ladder height [for definition, see, e.g., Siegmund (1985), page 168] of the random walk U_n, which is indexed by $\alpha = (\zeta_0^{(2)}, \gamma_\tau)$. Only condition (iii) of Theorem B.1 needs justification. The index set

$$C = \{\alpha: |\zeta_0^{(2)} - \xi_0^{(2)}| \le \delta, |\gamma_\tau - \gamma| \le \delta\}$$

is a compact subset of R^{d_2+d+1}. By compactness, for any sequence $\{\alpha_i\} \in C$, there exist a convergent subsequence with limit $\alpha \in C$. But this implies the sequence of distributions indexed by $\{\alpha_i\}$ converges in distribution to the distribution indexed by α, which is obviously nonlattice. This completes the proof of Lemma 4.7. □

Acknowledgments. This work is based on a portion of the author's doctoral dissertation, which was written at Stanford University under the supervision of Professor David Siegmund. His guidance and encouragement are gratefully acknowledged. The author would also like to thank a referee for his helpful suggestions.

REFERENCES

BORISOV, I. S. (1978). On the rate of convergence in the "conditional" invariance principle. *Theory Probab. Appl.* **23** 63–76.

BOROVKOV, A. A. and ROGOZIN, B. A. (1965). On the multidimensional central limit theorem. *Theory Probab. Appl.* **10** 55–62.

HOGAN, M. (1984). Problems in boundary crossing for random walks. Ph.D. dissertation, Stanford Univ.

HU, I. (1988). Repeated significance tests for exponential families. *Ann. Statist.* **16** 1643–1666.

JAMES, B., JAMES, K. L. and SIEGMUND, D. (1988). Conditional boundary crossing probabilities with applications to change-point problems. *Ann. Probab.* **16** 825–839.

KARTASHOV, N. V. (1980). A uniform asymptotic renewal theorem. *Theory Probab. Appl.* **25** 589–592.

LAI, T. L. and SIEGMUND, D. (1977). A nonlinear renewal theory with application to sequential analysis. I. *Ann. Statist.* **5** 946–954.

LALLEY, S. (1983). Repeated likelihood ratio test for curved exponential families. *Z. Wahrsch. Verw. Gebiete* **62** 293–321.

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

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

SIEGMUND, D. (1988). Confidence sets in change-point problems. *Internat. Statist. Rev.* **56** 31–48.

STONE, C. J. (1965). Characteristic functions and renewal theory. *Trans. Amer. Math. Soc.* **120** 327–342.

WOODROOFE, M. (1978). Large deviations of likelihood ratio statistics with application to sequential testing. *Ann. Statist.* **6** 72–84.

WOODROOFE, M. (1979). Repeated likelihood ratio tests. *Biometrika* **66** 453–464.

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

ZHANG, C. (1988). A nonlinear renewal theory. *Ann. Probab.* **16** 793–824.

DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104-6302