The Annals of Applied Probability
1992, Vol. 2, No. 3, 714–738

ON MOMENTS OF THE FIRST LADDER HEIGHT OF RANDOM WALKS WITH SMALL DRIFT

BY JOSEPH T. CHANG

Yale University

This paper presents some results that are useful in the study of asymptotic approximations of boundary crossing probabilities for random walks. The main result is a refinement of an asymptotic expansion of Siegmund concerning moments of the first ladder height of random walks having small positive drift. An analysis of the covariance between the first passage time and the overshoot of a random walk over a horizontal boundary contributes to the development of the main result and is of independent interest as well. An application of these results to a "moderate deviations" approximation for the probability distribution of the time to false alarm in the cusum procedure is briefly described.

1. Introduction. The study of boundary crossing probabilities for random walks is a field in which asymptotic approximations have been particularly successful and often remarkably accurate. The "corrected diffusion approximations" of Siegmund (1979) are a prime example, having applications in such fields as sequential analysis, queueing theory and insurance risk theory; see Siegmund (1985a, b) and Asmussen (1987), for example. One of the fundamental tools in the development of those approximations consists of two results, reproduced for convenience as Theorems 1.1 and 1.2, which describe the asymptotic behavior of the moments of the first ladder height of a random walk as the drift of the random walk tends to 0. The main purpose of this paper is to develop refinements of those results. This is done in Theorems 4.1 and 4.2. Along the way, we will develop some results, which are also of some independent interest, concerning the covariance between the first passage time of a random walk over a horizontal boundary and the amount by which the random walk overshoots that boundary.

In this introductory section we first introduce some notation and assumptions related to exponential families and random walks. Then we give some background and motivation for the problems treated in the paper, including a discussion of some ideas involved in their solution. This is followed by a brief description of an application of the results of this paper to a problem about the cusum procedure [Page (1954) and Lorden (1971)], which is a sequential method used in quality control for detecting a change in a probability distribution. There we state a theorem containing a new asymptotic approximation for the time until the cusum procedure gives a false alarm. In lieu of the full

Received November 1990; revised September 1991.

AMS 1980 *subject classifications*. Primary 60J15; secondary 60F99, 62L10.

Key words and phrases. Random walk, exponential family, uniform renewal theorem, first ladder height, first passage time, overshoot, boundary crossing probability, cusum procedure, corrected diffusion approximation, moderate deviations.

development that will appear elsewhere, we give an indication of the role played by the results of this paper in that development.

1.1. *Assumptions and notation.* The following terminology and notation will be used throughout the paper. We shall be concerned with one-parameter exponential families of distributions of the form $\{F_\theta: \theta \in \Theta\}$, where

$$F_\theta(dx) = e^{\theta x - \psi(\theta)} F_0(dx). \tag{1.1}$$

Whenever we consider such an exponential family, we shall assume that Θ contains an interval about 0, and also that the family has been standardized so that the distribution F_0 corresponding to the parameter value $\theta = 0$ has mean 0 and variance 1—that is, $\int x F_0(dx) = 0$ and $\int x^2 F_0(dx) = 1$. For convenience, let us call a one-parameter exponential family satisfying the conditions of the previous sentence a *standard exponential family*. Usually, we will also assume that the distribution F_0 is *strongly nonlattice*, that is,

$$\limsup_{|\lambda| \to \infty} |E_0 \exp(i\lambda X_1)| < 1.$$

We will generally assume that $X_1, X_2, \dots$ are independent and identically distributed, having a distribution F_θ that is a member of a standard exponential family. In this situation, the notation P_θ and E_θ will denote probability and expectation.

The function ψ appearing in (1.1), the cumulant generating function of the $\theta = 0$ distribution, has several familiar properties. For any $\theta \in \Theta$, the derivative $\psi'(\theta) =: \mu$ is the mean of the distribution F_θ. The second derivative $\psi''(\theta)$ is the variance of F_θ. Thus, our assumptions for "standardizing" the family amount to assuming $\psi'(0) = 0$ and $\psi''(0) = 1$. The Taylor expansion for ψ about $\theta = 0$ is

$$\psi(\theta) = \frac{1}{2}\theta^2 + \frac{\gamma}{6}\theta^3 + \frac{\kappa}{24}\theta^4 + \cdots \tag{1.2}$$

where $\gamma = E_0 X_1^3$ and $\kappa = E_0 X_1^4 - 3$ are the third and fourth cumulants of F_0.

A convenient convention that will be used consistently below is to let θ_0 and θ_1 denote elements of Θ satisfying

$$\theta_0 < 0 < \theta_1 \quad \text{and} \quad \psi(\theta_0) = \psi(\theta_1).$$

Other notation in this setting includes $\Delta := \theta_1 - \theta_0$, $\mu_0 := \psi'(\theta_0)$ and $\mu_1 := \psi'(\theta_1)$.

A particularly simple example of a standard exponential family that is also particularly important is the normal family, in which F_θ is the $N(\theta, 1)$ distribution. In this case, $\psi(\theta) = \theta^2/2$ and $\mu = \theta$.

With $X_1, X_2, \dots$ as above, define the *random walk* $\{S_n: n \geq 0\}$ by $S_0 = 0$ and $S_n = X_1 + \cdots + X_n$ for $n > 0$. The *first ladder epoch* τ_+ and the *first descending ladder epoch* τ_- are defined by $\tau_+ = \inf\{n: S_n > 0\}$ and $\tau_- = \inf\{n > 0: S_n \leq 0\}$, and the *first ladder height* is S_{τ_+}. More general *first* *passage times* are defined by

$$\tau_b \equiv \tau(b) = \inf\{n: S_n > b\}$$

for $b \geq 0$, and the *residual* at b is defined to be

$$R_b \equiv R(b) = S_{\tau(b)} - b. \tag{1.3}$$

Suppose F_0 is strongly nonlattice. Under any distribution P_θ for $\theta \geq 0$, R_b approaches a limiting distribution as $b \to \infty$. Let R_∞ denote a random variable whose distribution is this limiting distribution. The moments of R_∞ are given by

$$\rho^{(a)}(\theta) := E_\theta(R_\infty^a) = \frac{E_\theta\left(S_{\tau_+}^{a+1}\right)}{(a+1)E_\theta S_{\tau_+}} \tag{1.4}$$

for $a > 0$. For convenience let $\rho(\theta) \equiv \rho^{(1)}(\theta)$, $\rho_+^{(a)} \equiv \rho^{(a)}(0)$ and $\rho_+ \equiv \rho^{(1)}(0)$. For positive integers a, the quantities $\rho_-^{(a)}$ and ρ_- are defined analogously, using τ_- in place of τ_+; for example, $\rho_- = E_0(S_{\tau_-}^2)/(2E_0 S_{\tau_-})$.

A fundamental tool is *Wald's likelihood ratio identity*, which we will use in the following form. Let $X_1, X_2, \ldots$ be as above, let $\mathscr{F}_n$ be the sigma field generated by $X_1, \ldots, X_n$, let τ be a stopping time with respect to $\{\mathscr{F}_n\}$, and suppose Y is measurable with respect to $\mathscr{F}_\tau$. Then for $\theta, \theta' \in \Theta$,

$$E_\theta(Y; \tau < \infty) = E_{\theta'}\{Ye^{(\theta-\theta')S_\tau - \tau(\psi(\theta)-\psi(\theta'))}; \tau < \infty\}.$$

We will also use the term *Wald's equation* to refer to the familiar fact that in the above setting gives $E_\theta S_\tau = E_\theta X_1 E_\theta \tau$, for example.

Two relations that will be used below are

$$(E_0 S_{\tau_-})(E_0 S_{\tau_+}) = -\tfrac{1}{2} \tag{1.5}$$

and

$$\tfrac{1}{2}\rho_+^{(2)} + \tfrac{1}{2}\rho_-^{(2)} + \rho_+\rho_- = \tfrac{1}{12}(\kappa + 3), \tag{1.6}$$

which hold in any standard exponential family. These may be derived by differentiating the Wiener–Hopf factorization

$$\left(1 - E_0\{e^{i\lambda S_{\tau_-}}\}\right)\left(1 - E_0\{e^{i\lambda S_{\tau_+}}\}\right) = 1 - E_0\{e^{i\lambda X_1}\}$$

two and four times, then setting $\lambda = 0$.

1.2. *Covariance between first passage time and overshoot.* We retain the assumptions that $X_1, X_2, \ldots$ are iid with distribution F_θ belonging to a standard exponential family $\{F_\theta: \theta \in \Theta\}$ and F_0 is strongly nonlattice. Let $\theta > 0$ be fixed for now. Stam (1968) showed that as $b \to \infty$, τ_b and R_b are asymptotically independent in the sense that

$$P_\theta\left\{\frac{\tau_b - b/\mu}{\left(b\sigma^2/\mu^3\right)^{1/2}} \leq x,\ R_b \leq y\right\} \to \Phi(x) P_\theta\{R_\infty \leq y\}.$$

From this it is plausible that $\mathrm{Cov}_\theta(\tau_b, R_b) = o(b^{1/2})$. In fact, it turns out that actually $\mathrm{Cov}_\theta(\tau_b, R_b)$ approaches a constant as $b \to \infty$: Lai and Siegmund (1979) showed that

$$\lim_{b\to\infty} \mathrm{Cov}_\theta(\tau_b, R_b) = \frac{1}{\mu}\int_{[0,\infty)} (E_\theta R_x - E_\theta R_\infty) P_\theta\{M > -x\}\, dx \tag{1.7}$$
$$=: C(\theta)$$

where $M := \min_{n\ge 0} S_n$.

Now we want to drop the assumption that θ is fixed; applications in diffusion normalizations and "moderate deviations" normalizations such as Theorem 1.3 below involve the behavior of $\mathrm{Cov}_\theta(\tau_b, R_b)$ as $\theta \downarrow 0$ and $b \to \infty$ simultaneously. To begin to get a feeling for the behavior of the function $C(b,\theta) := \mathrm{Cov}_\theta(\tau_b, R_b)$, it is instructive to consider its limiting behavior along various lines in the plane $\{(b,\theta)\colon 0 \le b \le \infty,\ \theta \ge 0\}$. Limiting behavior of $C(b,\theta)$ as $b \to \infty$ along lines of constant θ was considered in (1.7). What happens if b is constant and $\theta \downarrow 0$? First suppose $b = \infty$; from (1.7) it makes sense to define $C(\infty,\theta) := C(\theta)$. It is interesting that as $\theta \downarrow 0$, $C(\theta)$ approaches a constant C_0; this may be shown using results from Section 4 about the asymptotic expansion of moments of S_{τ_+} for small θ. Next suppose b is a fixed finite number. It can be shown that in general $C(b,\theta)$ will either approach $+\infty$ or approach $-\infty$ as $\theta \downarrow 0$. For example, using Theorem 1.1, one can show that for $b = 0$ in the normal case we have

$$\lim_{\theta\downarrow 0} \mu\, \mathrm{Cov}_\theta(\tau_0, R_0) = 2^{-1/2}(\rho - 2^{-1/2}) < 0$$

($\rho \doteq 0.583$), so that here $\mathrm{Cov}_\theta(\tau_0, R_0) \to -\infty$ as $\theta \downarrow 0$.

Thus, it appears that while $\mathrm{Cov}_\theta(\tau_b, R_b)$ behaves "nicely" if we first let $b \to \infty$ and then $\theta \downarrow 0$, it behaves "badly" if we first let $\theta \downarrow 0$ and then $b \to \infty$. What sort of behavior is exhibited if we let $\theta \downarrow 0$ and $b \to \infty$ together, say in accordance with a diffusion or moderate deviations normalization? Section 3.1 will show that the behavior is "nice."

Section 3.2 begins by establishing a representation of $\mathrm{Cov}_\theta(\tau_b, R_b^a)$ in terms of the derivative $\dot{E}_\theta(R_b^a)$ of $E_\theta(R_b^a)$ with respect to θ. (The superscripts a are powers here.) This representation is combined with a result from Section 3.1 to obtain a result about $\dot{E}_\theta R_b^a$ that is used in the analysis of $E_\theta R_b^a$ in Section 4.

1.3. *Moments of the first ladder height.* We begin by stating Siegmund's results.

THEOREM 1.1 [Siegmund (1979)]. *Let $\{F_\theta\colon \theta \in \Theta\}$ be a standard exponential family of distributions, with F_0 nonlattice. For any $a > 0$,*

$$\lim_{\theta\downarrow 0} \mu E_\theta(\tau_+ S_{\tau_+}^a) = E_0 S_{\tau_+}^{a+1}/(a+1). \tag{1.8}$$

Theorem 1.1 provides the main ingredient in the proof of the following closely related result.

THEOREM 1.2 [Siegmund (1979)]. *Under the conditions of Theorem* 1.1, *for any* $a > 0$ *we have*

$$E_\theta S_{\tau_+}^a = E_0 S_{\tau_+}^a + \frac{a}{a+1}\left(E_0 S_{\tau_+}^{a+1}\right)\theta + o(\theta) \tag{1.9}$$

as $\theta \downarrow 0$.

It is Theorem 1.2 that has been used so extensively in deriving approximations. For example, these results are an important part of the theory behind the well-known "correction" of adding ρ_+ ($\doteq 0.583$ in the normal case) to the boundary in corrected diffusion approximations. Perhaps the simplest example is the result $P_{\theta_0}\{\tau_b < \infty\} = e^{-\Delta b} + o(\Delta^2)$, whose proof relies on Theorem 1.2. However, as will be indicated below, the results are not strong enough to obtain the approximation related to the cusum procedure described in Section 1.4. Refining Theorem 1.2 is the main problem treated in this paper.

What sorts of refinements would we like? It is natural to ask whether the "$o(\theta)$" that appears in Theorem 1.2 may be strengthened to read "$O(\theta^2)$." At the cost of strengthening the hypotheses on F_0 slightly from "nonlattice" to "strongly nonlattice," we will show that the answer to this question is "yes." The next natural question becomes that of determining the coefficient of θ^2. Having done this, one would ask whether the remainder $o(\theta^2)$ is in fact $O(\theta^3)$, and so on.

To sketch some of the ideas involved, for convenience let

$$h(\theta) := E_\theta S_{\tau_+}^a.$$

In analyzing the behavior of h near $\theta = 0$, the first property to check is continuity, that is, whether

$$h(\theta) = h(0) + o(1) \quad \text{as } \theta \downarrow 0.$$

This is shown in Siegmund (1979). To find an asymptotic expansion, the next item of interest is the derivative $h'(0)$, if it exists. As Siegmund (1979) shows using Wald's likelihood ratio identity, for $\theta > 0$,

$$h'(\theta) = E_\theta\left\{S_{\tau_+}^a\left(S_{\tau_+} - \mu\tau_+\right)\right\}. \tag{1.10}$$

Note that $h'(0)$ cannot be obtained simply by substituting $\theta = 0$ in the right-hand side of (1.10), since the second term then becomes of the form "$0 \times \infty$." It is necessary to take the less direct approach of finding the limit of $h'(\theta)$ as $\theta \downarrow 0$, the nontrivial part of which is finding

$$\lim_{\theta \downarrow 0} \mu E_\theta\left(\tau_+ S_{\tau_+}^a\right). \tag{1.11}$$

It is intuitively plausible that (1.11) should be finite, which would imply that $\lim_{\theta \downarrow 0} h'(\theta)$ is finite, at which point we could conclude that

$$h(\theta) = h(0) + O(\theta) \quad \text{as } \theta \downarrow 0.$$

A harder problem is finding a nice expression for the limit, which was done by Siegmund in Theorem 1.1. Having identified $h'(0) = \lim_{\theta \downarrow 0} h'(\theta)$, we may write

$$h(\theta) = h(0) + h'(0)\theta + o(\theta), \tag{1.12}$$

which is the content of Theorem 1.2. If we want to go further and find $h''(0)$, analogously to what we did above in the case of $h'(0)$, we find using Wald's likelihood ratio identity that

$$h''(\theta) = E_\theta\Big\{-\mu'(\theta)\tau_+ S^a_{\tau_+} + S^a_{\tau_+}(S_{\tau_+} - \mu\tau_+)^2\Big\},$$

and attempt to take the limit as $\theta \downarrow 0$. Given Theorem 1.1, the nontrivial part is

$$\lim_{\theta\downarrow 0}\big\{-\mu'(\theta)\tau_+ S^a_{\tau_+} + \mu^2\tau_+^2 S^a_{\tau_+}\big\}. \tag{1.13}$$

In this case, it is not at all apparent that the limit is even finite; in fact, as $\theta \downarrow 0$, (1.13) is of the form "$-\infty + \infty$."

Showing that (1.13) is finite would enable us to strengthen (1.12) to

$$h(\theta) = h(0) + h'(0)\theta + O(\theta^2) \quad \text{as } \theta \downarrow 0.$$

Identifying the limit and denoting it by $h''(0)$ would give

$$h(\theta) = h(0) + h'(0)\theta + \tfrac{1}{2}h''(0)\theta^2 + o(\theta^2) \quad \text{as } \theta \downarrow 0.$$

In Theorem 4.2 below, we will go one step further and obtain an expansion of the form

$$h(\theta) = h(0) + h'(0)\theta + \tfrac{1}{2}h''(0)\theta^2 + O(\theta^3) \quad \text{as } \theta \downarrow 0.$$

1.4. *Application to a "moderate deviations" result for the cusum procedure.* Suppose $\theta_0 < 0$. Then under P_{θ_0} the random walk $\{S_n: n \geq 0\}$ drifts downward with drift $E_{\theta_0}X_1 < 0$. Define the *reflected random walk* $\{W_n: n \geq 0\}$ with reflecting barrier at 0 by

$$W_n = S_n - \min_{0\leq k\leq n} S_k.$$

Taking $b > 0$ and defining $t_b = \inf\{n: W_n > b\}$, our focus of attention is the probability $P_{\theta_0}\{t_b \leq m\}$. In terms of the cusum procedure, we are interested in the probability distribution of the time until the procedure gives a "false alarm." The following result of Chang (1989) gives an approximation.

THEOREM 1.3. *Suppose $\{F_\theta: \theta \in \Theta\}$ is a standard exponential family such that the distribution F_0 is strongly nonlattice. Let*

$$c := 1 - 2(\rho_+^{(2)} - \rho_+^2) - (\gamma^2/9) + (\kappa/3).$$

Assume $\theta_0 \uparrow 0$, $b \to \infty$ and $m \to \infty$ in such a way that for some $\delta > 0$ and some k we have

$$|\theta_0|^{1+\delta} b \to \infty, \tag{1.14}$$

$$|\theta_0|^k m \to 0$$

and

$$m \geq \frac{b}{\mu_1}(1 + \delta). \tag{1.15}$$

Then

$$\begin{aligned} P_{\theta_0}\{t_b \leq m\} = e^{-\Delta(b+\rho_+-\rho_-)}\big[\Delta|\mu_0|\{m + c - (b + \rho_+ - \rho_-)/\mu_1\} \\ +3 - (2/3)\gamma\Delta + O(\Delta^5 m)\big]. \end{aligned} \tag{1.16}$$

A novel feature of Theorem 1.3 is the use of "moderate deviations" normalizations, which are in a sense intermediate between normalizations usually called "large deviations," in which the drift of the random walk is constant, and diffusion normalizations, in which the drift of the random walk approaches 0 fast enough to make the probability in question approach a positive constant. The starting point for Theorem 1.3 was in the work of Siegmund (1988), who developed a large deviations approximation for $P_{\theta_0}\{t_b \leq m\}$. His approximation unfortunately contained constants complicated enough to make numerical evaluation impractical. However, Siegmund (1988) was able to use his large deviations approximation to derive in a heuristic manner an approximation whose numerical evaluation is extremely simple. The heuristic nature of Siegmund's formula stems from the fact that it was derived by algebraically combining results of different theorems that assumed different normalizations that were not consistent with each other. This precluded the formulation and proof of a bona fide asymptotic expansion in a single, consistent normalization with an error term of specified order of magnitude, and so on. Theorem 1.3 provides this sort of mathematical foundation for Siegmund's proposed approximation.

Theorem 1.3 also carries the required calculations out to a higher order of accuracy than Siegmund did, resulting in the extra term "c" that seems to improve the accuracy of the approximation. To give an indication of the effect of this extra term, Table 1 compares (1.16) with Siegmund's approximation—which is (1.16) with the "c" removed—on the same example Siegmund (1988) used to illustrate his approximation: The exponential family is the $\{N(\theta, 1)\}$ family, $\theta_0 = -0.5$, $b = 3$, and m takes the values displayed in the table. The entries in the "true" column are the actual probabilities $P_{-0.5}\{t_3 \leq m\}$, computed numerically by Waldmann (1986). The entries in the "with c" and

TABLE 1

m	True	Without c	With c
9	0.0542	0.0517	0.0556
12	0.0786	0.0750	0.0789
15	0.1024	0.0983	0.1021
18	0.1257	0.1215	0.1254

"without c" columns are the results of evaluating (1.16) and Siegmund's approximation, respectively.

The main results of this paper, Theorems 4.1 and 4.2, are important in the development of Theorem 1.3. To see why, suppose Theorems 4.1 and 4.2 were not available, and we were working with Theorems 1.1 and 1.2 instead. Then it turns out that the "$O(\Delta^5 m)$" in (1.16) would have to be replaced by "$o(\Delta^3 m)$." However, from assumptions (1.14) and (1.15) it follows that $\Delta^2 m \to \infty$. Therefore, all terms of order Δ or less—including the "c," "$\rho_+ - \rho_-$" and "$(2/3)\gamma\Delta$" in square brackets in (1.16)—are of order $o(\Delta^3 m)$. Thus, being of smaller order of magnitude than the error term, these terms would not be "justified" for inclusion in the approximation. It is interesting that the seemingly innocuous advance of showing that the $o(\theta)$ in (1.9) may be replaced by $O(\theta^2)$ is enough to justify the "$(2/3)\gamma\Delta$" and the "$\rho_+ - \rho_-$" in square brackets in (1.16). The full force of Theorems 4.1 and 4.2 is needed to justify the "c."

The results of Section 3, apart from their contributing to the development of the results in Section 4, are of some interest in their own right. In fact, they are also needed in the proof of Theorem 1.3. Define $T = \inf\{n > 0\colon S_n \notin (0, b]\}$. The quantity $E_{\theta_0}\{T;\, S_T > b\}$ is an important ingredient in the desired probability $P_{\theta_0}\{t_b \le m\}$; see Lemma 8 of Siegmund (1988) for the connection. It turns out that we need to obtain an expansion for $E_{\theta_0}\{T;\, S_T > b\}$ up to order $\Delta^2 e^{-\Delta b}$, which requires showing that

$$\operatorname{Cov}_{\theta_1}\!\left(\tau_b, e^{-\Delta R_b}\right) - \int_{(-\infty,0]} \operatorname{Cov}_{\theta_1}\!\left(\tau_{b-s}, e^{-\Delta R_{b-s}}\right) P_{\theta_1}\{S_T \in ds\} = O(\Delta^2).$$

This is an easy consequence of Theorem 3.2.

2. A uniform renewal theorem and consequences.

A fundamental renewal theoretic development that provides an important tool for the present paper is the work of Stone (1965). A uniform version due to Siegmund (1979) of a theorem of Stone (1965) is stated in this section, and some simple consequences of that result are developed here and used in later sections.

Let $\{F_\theta\colon \theta \in \Theta\}$ be a standard exponential family, suppose that the distribution F_0 is strongly nonlattice and let $X_1, X_2, \ldots$ be iid with distribution F_θ for some nonnegative θ in the interior of Θ. Define $\tau_+^{(0)} := 0$, and for $b > 0$ define the nth ladder epoch recursively by

$$\tau_+^{(n)} := \inf\{k\colon S_k > S_{\tau_+^{(n-1)}}\}.$$

Note that for convenience we are letting $\tau_+ \equiv \tau_+^{(1)}$. Letting

$$U_\theta^+(x) := \sum_{n=0}^{\infty} P_\theta\{S_{\tau_+^{(n)}} \le x\}, \tag{2.1}$$

it can be shown, using a theorem of Stone (1965), that for some $r > 0$,

$$U_\theta^+(x) = \frac{x}{E_\theta S_{\tau_+}} + \frac{E_\theta S_{\tau_+}^2}{2(E_\theta S_{\tau_+})^2} + O(e^{-rx}) \quad \text{as } x \to \infty. \tag{2.2}$$

Thus, given an exponential family, for any fixed $\theta \geq 0$ Stone's theorem guarantees an exponential rate of decrease of the "error term" in (2.2) as $x \to \infty$. However, in many applications, such as corrected diffusion approximations and the approximation presented in Theorem 1.3, we are letting $\theta \downarrow 0$ simultaneously with $x \to \infty$. Consequently, we must contemplate the unpleasant possibility that as $\theta \downarrow 0$ the rate of convergence to 0 of the error term in (2.2) might conceivably become slower and slower. The next result guarantees that this cannot happen; that is, there is a certain exponential rate of convergence that applies uniformly to all θ in some neighborhood of 0.

THEOREM 2.1 [Siegmund (1979)]. *Suppose that $\{F_\theta: \theta \in \Theta\}$ is a standard exponential family and that the distribution F_0 is strongly nonlattice. Then there exist $r > 0$, $\theta^* > 0$, and C such that*

$$\left| U_\theta^+(x) - \frac{x}{E_\theta S_{\tau_+}} - \frac{E_\theta S_{\tau_+}^2}{2(E_\theta S_{\tau_+})^2} \right| \leq Ce^{-rx}$$

for all nonnegative x and all $\theta \in [0, \theta^*]$.

The basic reason why Theorem 2.1 will be useful is as follows. By the "renewal argument," many functions of interest satisfy renewal equations of the form

$$Z_\theta(x) = z_\theta(x) + \int_0^x Z_\theta(x-t) P_\theta\{S_{\tau_+} \in dt\}. \tag{2.3}$$

For example, with R_x defined as in (1.3),

$$Z_\theta(x) = E_\theta R_x^a \tag{2.4}$$

and

$$Z_{\theta_1}(x) = E_{\theta_1}(e^{-\Delta R_x}) \tag{2.5}$$

are two such functions. An elementary consequence of the relation (2.3) is that the function Z_θ may be expressed as the convolution of z_θ with the renewal measure U_θ^+, that is,

$$Z_\theta(x) = \int_0^x z_\theta(x-t) U_\theta^+(dt).$$

Accordingly, one would expect that the statement made by Theorem 2.1 about U_θ^+ could have implications for any function that satisfies a renewal equation of the form (2.3).

Let us start with a result from which other examples such as (2.4) and (2.5) may be treated easily.

THEOREM 2.2. *Under the assumptions of Theorem* 2.1, *there exist* $r > 0$, $\theta^* > 0$ *and* C *such that*

$$|P_\theta\{R_x \le y\} - P_\theta\{R_\infty \le y\}| \le Ce^{-r(x+y)} \tag{2.6}$$

for all $x \ge 0$, $y \ge 0$ *and* $\theta \in [0, \theta^*]$.

PROOF. Since

$$P_\theta\{R_x > y\} = \int_{[0,x]} P_\theta\{S_{\tau_+} > x + y - t\} U_\theta^+(dt)$$

by a renewal argument, and since

$$P_\theta\{R_\infty > y\} = \int_y^\infty P_\theta\{S_{\tau_+} > t\} \frac{dt}{E_\theta S_{\tau_+}}$$
$$= \int_{-\infty}^x P_\theta\{S_{\tau_+} > x + y - t\} \frac{dt}{E_\theta S_{\tau_+}},$$

we have

$$\begin{aligned} P_\theta\{R_x \le y\} - P_\theta\{R_\infty \le y\} &= \int_{[0,x]} P_\theta\{S_{\tau_+} > x + y - t\} \left\{ U_\theta^+(dt) - \frac{dt}{E_\theta S_{\tau_+}} \right\} \\ &\quad - \int_{-\infty}^0 P_\theta\{S_{\tau_+} > x + y - t\} \frac{dt}{E_\theta S_{\tau_+}} \\ &=: J_1 - J_2. \end{aligned} \tag{2.7}$$

Define

$$\varepsilon_\theta(x) := U_\theta^+(x) - \frac{x}{E_\theta S_{\tau_+}} - \frac{E_\theta S_{\tau_+}^2}{2(E_\theta S_{\tau_+})^2},$$

so that

$$J_1 = \int_{[0,x]} P_\theta\{S_{\tau_+} > x + y - t\} \varepsilon_\theta(dt).$$

Then integration by parts gives

$$\begin{aligned} J_1 &= P_\theta\{S_{\tau_+} \ge y\}\varepsilon_\theta(x) + \frac{E_\theta S_{\tau_+}^2}{2(E_\theta S_{\tau_+})^2} P_\theta\{S_{\tau_+} > x + y\} \\ &\quad - \int_{[0,x]} \varepsilon_\theta(t) P_\theta\{S_{\tau_+} \in x + y - dt\}. \end{aligned} \tag{2.8}$$

To bound $|J_1|$, note that Theorem 2.1 says that

$$|\varepsilon_\theta(x)| \le C_1 e^{-r_1 x} \tag{2.9}$$

for all x and all $\theta \in [0, \theta_1^*]$. Also, using the requirement that Θ contain an interval about 0, the well-known fact that for any λ

$$E_0(e^{\lambda X_1}) < \infty \quad \text{implies} \quad E_0(e^{\lambda S_{\tau_+}}) < \infty$$

[for which a reference is Siegmund's (1985) Problem 8.9], and Wald's likelihood ratio identity, one can show that there exist C_2, $r_2 > 0$ and $\theta_2^* > 0$ such that

$$E_\theta(e^{r_2 S_{\tau_+}}) \le C_2 \qquad \forall\ \theta \in [0, \theta_2^*],$$

so that

$$P_\theta\{S_{\tau_+} \ge z\} \le C_2 e^{-r_2 z} \qquad \forall\ \theta \in [0, \theta_2^*], \forall\ z \ge 0. \tag{2.10}$$

Therefore, by (2.9) and (2.10), letting $r_3 = r_1 \wedge r_2$ and $\theta_3^* = \theta_1^* \wedge \theta_2^*$ for example, clearly there is a C_3 such that

$$\left| P_\theta\{S_{\tau_+} \ge y\} \varepsilon_\theta(x) + \frac{E_\theta S_{\tau_+}^2}{2(E_\theta S_{\tau_+})^2} P_\theta\{S_{\tau_+} > x + y\} \right| \le C_3 e^{-r_3(x+y)} \tag{2.11}$$

for all $x \ge 0$, $y \ge 0$ and $\theta \in [0, \theta_3^*]$. For the integral in (2.8), since

$$\left| \int_{[0, x/2]} \varepsilon_\theta(t) P_\theta\{S_{\tau_+} \in x + y - dt\} \right| \le C_1 P_\theta\left\{ S_{\tau_+} \ge y + \frac{x}{2} \right\} \le C_1 C_2 e^{-r_2(y + x/2)}$$

and

$$\left| \int_{(x/2, x]} \varepsilon_\theta(t) P_\theta\{S_{\tau_+} \in x + y - dt\} \right| \le C_1 e^{-r_1 x/2} P_\theta\{S_{\tau_+} \ge y\} \le C_1 C_2 e^{-(r_1 x/2 + r_2 y)},$$

there is an $r_4 > 0$ and a C_4 such that the integral over $[0, x]$ is bounded by $C_4 e^{-r_4(x+y)}$ for all $x \ge 0$, $y \ge 0$ and $\theta \in [0, \theta_3^*]$. Thus, by (2.8) there exist $r_5 > 0$ and C_5 such that

$$|J_1| \le C_5 e^{-r_5(x+y)} \tag{2.12}$$

for all $x \ge 0$, $y \ge 0$ and $\theta \in [0, \theta_3^*]$.

For J_2, since $E_\theta S_{\tau_+}$ is bounded below for $\theta \ge 0$, we have

$$|J_2| \le \int_{-\infty}^0 C_2 e^{-r_2(x+y-t)} \frac{dt}{C_6} = C_7 e^{-r_2(x+y)}. \tag{2.13}$$

This proof is completed by combining (2.7), (2.12) and (2.13). □

The following result gives bounds related to (2.4) and (2.5) that will be used below.

COROLLARY 2.3. *Assume the conditions of Theorem* 2.1. *Then for any $a > 0$ there exist $r > 0$, $\theta^* > 0$ and C such that*

$$|E_\theta R_x^a - E_\theta R_\infty^a| \le C e^{-rx} \tag{2.14}$$

for all nonnegative x and all $\theta \in [0, \theta^*]$. Also, there exist $r > 0$, $\theta^* > 0$ and C

such that

$$\left|E_{\theta_1}(e^{-\Delta R_x}) - E_{\theta_1}(e^{-\Delta R_\infty})\right| \le C\Delta e^{-rx} \tag{2.15}$$

for all nonnegative x and all $\theta_1 \in [0, \theta^*]$.

PROOF. Since

$$E_\theta R_x^a = E_\theta \int_0^{R_x} a y^{a-1}\, dy = a\int_0^\infty y^{a-1} P_\theta\{R_x > y\}\, dy$$

and similarly

$$E_\theta R_\infty^a = a\int_0^\infty y^{a-1} P_\theta\{R_\infty > y\}\, dy,$$

using Theorem 2.2 we see that there exist $r > 0$, $\theta^* > 0$ and C_i such that

$$\begin{aligned} |E_\theta R_x^a - E_\theta R_\infty^a| &\le a\int_0^\infty y^{a-1} |P_\theta\{R_x > y\} - P_\theta\{R_\infty > y\}|\, dy \\ &\le C_1 \int_0^\infty y^{a-1} e^{-r(x+y)}\, dy \\ &= C_2 e^{-rx} \end{aligned}$$

for all $x \ge 0$ and all $\theta \in [0, \theta^*]$. This proves (2.14); (2.15) is proved in the same way. □

3. Covariance between first passage time and overshoot.

3.1. *Rate of convergence.* For later use, we consider the more general problem of analyzing the behavior of $\mathrm{Cov}_\theta(\tau_b, R_b^a)$ for $a > 0$.

THEOREM 3.1. *Let $\{F_\theta\colon \theta \in \Theta\}$ be a standard exponential family, with F_0 strongly nonlattice. Let $a > 0$ and define*

$$C^{(a)}(\theta) := \frac{1}{\mu}\int_{[0,\infty)} (E_\theta R_x^a - E_\theta R_\infty^a) P_\theta\{M > -x\}\, dx.$$

Then there exist A, $r > 0$ and $\theta^* > 0$ such that

$$\left|\mathrm{Cov}_\theta(\tau_b, R_b^a) - C^{(a)}(\theta)\right| \le \mu^{-1} A e^{-rb}$$

for all b and for all $\theta \in (0, \theta^*]$.

PROOF. First observe that it suffices to show that there exist A_1, A_2, $r > 0$ and $\theta^* > 0$ such that

$$\left|\mathrm{Cov}_\theta(\tau_b, R_b^a) - C^{(a)}(\theta)\right| \le \mu^{-1}(A_1 b + A_2) e^{-rb} \tag{3.1}$$

for all b and for all $\theta \in (0, \theta^*]$. Next recall three results of Lai and Siegmund (1979); they consider the case $a = 1$, but the same proofs also work for general positive a. Assume $\theta > 0$ is fixed. The first result is that

$$\operatorname{Cov}_\theta(\tau_b, R_b^a) = \int_{[0,\infty)} (E_\theta R_x^a - E_\theta R_b^a) Q_\theta(dx; b), \tag{3.2}$$

where

$$Q_\theta(dx; b) := \sum_{n=0}^{\infty} P_\theta\{\tau_b > n, S_n \in b - dx\}. \tag{3.3}$$

Second, Lai and Siegmund show in the proof of their Theorem 4 that as $b \to \infty$,

$$Q_\theta(dx; b) \to \mu^{-1} P_\theta\{M > -x\}\, dx =: Q_\theta(dx; \infty) \tag{3.4}$$

in the sense that

$$Q_\theta(x; b) := \int_{[0,x)} Q_\theta(dy; b) \to \int_{[0,x)} Q_\theta(dy; \infty) =: Q_\theta(x; \infty) \tag{3.5}$$

for all $x \geq 0$. Lastly, as mentioned above, they also show [see (60) on their page 71 for the case $a = 1$] that

$$\operatorname{Cov}_\theta(\tau_b, R_b^a) \to \int_{[0,\infty)} (E_\theta R_x^a - E_\theta R_\infty^a) Q_\theta(dx; \infty) \triangleq C^{(a)}(\theta) \tag{3.6}$$

as $b \to \infty$. From (3.2) and (3.6), noting that $\int_{[0,\infty)} Q_\theta(dx; b) = E_\theta \tau_b$, we obtain

$$\begin{aligned} &\operatorname{Cov}_\theta(\tau_b, R_b^a) - C^{(a)}(\theta) \\ &\quad = \int_{[0,\infty)} (E_\theta R_x^a - E_\theta R_\infty^a)[Q_\theta(dx; b) - Q_\theta(dx; \infty)] \\ &\qquad + (E_\theta R_\infty^a - E_\theta R_b^a) E_\theta \tau_b. \end{aligned} \tag{3.7}$$

For convenience, let us commit the linguistic abuse of saying that a function of θ and b "satisfies (3.1)" if there exist A_1, A_2, $r > 0$ and $\theta^* > 0$ such that the function is bounded in absolute value by the right-hand side of (3.1) for all b and for all $\theta \in (0, \theta^*]$. Note that if two functions both satisfy (3.1), then so does their sum. We shall prove the theorem by showing that both of the two terms on the right-hand side of (3.7) satisfy (3.1).

Let $B > 0$, $s > 0$ and $\theta^* > 0$ be chosen such that for all $x \geq 0$ and $0 \leq \theta \leq \theta^*$ we have

$$E_\theta S_{\tau_+} \geq B^{-1}, \tag{3.8}$$

$$E_\theta R_x^p \leq B \quad \text{for } p \in \{1, a\}, \tag{3.9}$$

$$|E_\theta R_x^a - E_\theta R_\infty^a| \leq B e^{-sx} \tag{3.10}$$

and

$$\left| U_\theta^+(x) - (x + E_\theta R_\infty)/E_\theta S_{\tau_+} \right| \leq B e^{-sx}, \tag{3.11}$$

where $U_\theta^+(x)$ is defined in (2.1). Assurance that we may fulfill conditions (3.8)–(3.11) is provided by the continuity of $E_\theta S_{\tau_+}$ in θ for (3.8), Theorem 3 of

Lorden (1970) for (3.9), Corollary 2.3 for (3.10), and Theorem 2.1 for (3.11). Then from Wald's equation, (3.9) and (3.10), clearly

$$|(E_\theta R_b^a - E_\theta R_\infty^a) E_\theta \tau_b| \le Be^{-sb}\left(\frac{b+B}{\mu}\right)$$

for all $b \ge 0$ and $\theta \in (0, \theta^*)$. This takes care of the last term in (3.7).

It remains to show that the integral in (3.7) satisfies (3.1). Since we will be reexpressing the integral in (3.7) by integration by parts, we will be concerned with the size of differences of the form $Q_\theta(x, b) - Q_\theta(x, \infty)$. In fact, we claim that

$$|Q_\theta(x;b) - Q_\theta(x,\infty)| \le 4B^2\mu^{-1}e^{-s(b-x)} \tag{3.12}$$

for all $b > 0$, $0 \le x \le b$ and $\theta \in (0, \theta^*]$.

To prove (3.12), start with the relation

$$Q_\theta(x;b) = (E_\theta\tau_+)\int_{[0,x)} P_\theta\{M > -y\} U_\theta^+(b - x + dy)$$

of Lai and Siegmund [(1979), pages 65 and 66]. Also, by Wald's equation and the definitions in (3.4) and (3.5),

$$Q_\theta(x;\infty) = (E_\theta\tau_+)\int_{[0,x)} P_\theta\{M > -y\}\frac{dy}{E_\theta S_{\tau_+}}.$$

Therefore,

$$\begin{aligned}
&\frac{Q_\theta(x;b) - Q_\theta(x;\infty)}{E_\theta\tau_+}\\
&\quad= \int_{[0,x)} P_\theta\{M > -y\}\, d\left[U_\theta^+(b-x+y) - U_\theta^+(b-x) - \frac{y}{E_\theta S_{\tau_+}}\right]\\
&\quad= P_\theta\{M > -x\}\left[U_\theta^+(b-) - U_\theta^+(b-x) - \frac{x}{E_\theta S_{\tau_+}}\right]\\
&\qquad-\int_{[0,x)}\left[U_\theta^+(b-x+y) - U_\theta^+(b-x) - \frac{y}{E_\theta S_{\tau_+}}\right] d(P_\theta\{M > -y\}),
\end{aligned}$$

so that, using (3.11), we obtain

$$\frac{|Q_\theta(x;b) - Q_\theta(x;\infty)|}{E_\theta\tau_+} \le 4Be^{-s(b-x)}. \tag{3.13}$$

Thus, since (3.9) together with Wald's equation implies that $E_\theta\tau_+ \le B/\mu$ for all $\theta \in (0, \theta^*]$, the claim (3.12) is proved by (3.13).

We now return to the task of showing that the integral in (3.7) satisfies (3.1). We want to show that the integral is small for large b. The idea is this: For x large, $E_\theta R_x^a$ is close to $E_\theta R_\infty^a$, while for x small (compared to b, which is large), "$Q_\theta(dx; b)$ is close to $Q_\theta(dx; \infty)$." This motivates splitting the range of integration $[0,\infty)$ into two subintervals; $[0, b/2)$ and $[b/2,\infty)$ will do. The integral over $[0, b/2)$ is

$$\begin{aligned}\int_{[0,b/2)} &(E_\theta R_x^a - E_\theta R_\infty^a)[Q_\theta(dx;b) - Q_\theta(dx;\infty)]\\ &= (E_\theta R_{b/2}^a - E_\theta R_\infty^a)[Q_\theta(b/2;b) - Q_\theta(b/2;\infty)]\\ &\quad - \int_{[0,b/2)}[Q_\theta(x;b) - Q_\theta(x;\infty)]\,d(E_\theta R_x^a),\end{aligned} \tag{3.14}$$

where we have used $Q_\theta(0-;b) = 0 = Q_\theta(0-;\infty)$. However, by (3.10) and (3.12),

$$\left|(E_\theta R_{b/2}^a - E_\theta R_\infty^a)[Q_\theta(b/2;b) - Q_\theta(b/2;\infty)]\right| \le 4B^3\mu^{-1}e^{-sb}. \tag{3.15}$$

Furthermore, since the reasoning of Lorden (1970) that gave his equation (1) also gives

$$\int_0^x aE_\theta R_y^{a-1}\,dy = (E_\theta S_{\tau_+}^a)U_\theta^+(x) - E_\theta R_x^a, \tag{3.16}$$

we have

$$\begin{aligned}&\left|\int_{[0,b/2)}[Q_\theta(x;b) - Q_\theta(x;\infty)]\,d(E_\theta R_x^a)\right|\\ &\quad\le \Big(\sup_{0\le x<b/2}|Q_\theta(x;b) - Q_\theta(x;\infty)|\Big)\\ &\quad\quad\times\left\{(E_\theta S_{\tau_+}^a)U_\theta^+(b/2) + \int_0^{b/2} aE_\theta R_x^{a-1}\,dx\right\}\\ &\quad\le (4B^2\mu^{-1}e^{-sb/2})\big[2(E_\theta S_{\tau_+}^a)U_\theta^+(b/2)\big]\\ &\quad\le 8\mu^{-1}B^4[b/2 + B]e^{-sb/2}.\end{aligned} \tag{3.17}$$

The first two inequalities in (3.17) use (3.16), and the last inequality uses (3.8) and (3.9) to say that

$$U_\theta^+(b/2) = (E_\theta S_{\tau_+})^{-1}(b/2 + E_\theta R_{b/2}) \le B(b/2 + B).$$

Combining (3.15) and (3.17) shows that (3.14) satisfies (3.1).

Finally, for the integral over $[b/2,\infty)$, use (3.10) to write

$$\begin{aligned}&\left|\int_{[b/2,\infty)}(E_\theta R_x^a - E_\theta R_\infty^a)[Q_\theta(dx;b) - Q_\theta(dx;\infty)]\right|\\ &\quad\le \int_{[b/2,\infty)} Be^{-sx}Q_\theta(dx;b) + \int_{[b/2,\infty)} Be^{-sx}Q_\theta(dx;\infty).\end{aligned} \tag{3.18}$$

It is easy to see that the last two integrals satisfy (3.1), using

$$\int_{[0,\infty)} Q_\theta(dx;b) = E_\theta \tau_b \le \mu^{-1}(b+B)$$

and definition (3.4). This completes the proof. □

The next result was mentioned in Section 1.4. To state it, define

$$\tilde{C}(\theta_1) := \lim_{b\to\infty} \operatorname{Cov}_{\theta_1}(\tau_b, e^{-\Delta R_b}),$$

which can be shown to exist and identified to be

$$\tilde{C}(\theta_1) = \int_{[0,\infty)} \left[E_{\theta_1}(e^{-\Delta R_x}) - E_{\theta_1}(e^{-\Delta R_\infty})\right] Q_{\theta_1}(dx;\infty)$$

in the same manner in which Lai and Siegmund (1979) treated $C(\theta)$.

THEOREM 3.2. *Under the conditions of Theorem* 3.1, *there exist* A, $r > 0$ *and* $\theta^* > 0$ *such that*

$$\left|\operatorname{Cov}_{\theta_1}(\tau_b, e^{-\Delta R_b}) - \tilde{C}(\theta_1)\right| \le Ae^{-rb}$$

for all b *and for all* $\theta_1 \in (0, \theta^*]$.

PROOF. By the same reasoning that gave (3.7), we may obtain

$$\begin{aligned}\operatorname{Cov}_{\theta_1}&(\tau_b, e^{-\Delta R_b}) - \tilde{C}(\theta_1)\\ &= \int_{[0,\infty)} \left[E_\theta(e^{-\Delta R_x}) - E_\theta(e^{-\Delta R_\infty})\right]\left[Q_\theta(dx;b) - Q_\theta(dx;\infty)\right]\\ &\quad + \left[E_\theta(e^{-\Delta R_\infty}) - E_\theta(e^{-\Delta R_b})\right] E_\theta \tau_b.\end{aligned}$$

From here we proceed completely analogously to the proof of Theorem 3.1, with the bound (3.10) replaced by the bound of (2.15). This provides an extra factor of Δ all of the way through the proof. □

3.2. *Another representation and a consequence.* In this section, we will present a representation of $\operatorname{Cov}_\theta(\tau_b, R_b^a)$ that could form the starting point for an alternative approach to analyzing this quantity. From the representation we will also derive a consequence that will be useful in the next section. For notational convenience, throughout this section we will use a dot to denote differentiation with respect to θ.

THEOREM 3.3. *Assuming the conditions of Theorem* 3.1, *for all positive* θ *in the interior of* Θ *we have*

$$\operatorname{Cov}_\theta(\tau_b, R_b^a) = \frac{1}{\mu}\left[E_\theta R_b^{a+1} - (E_\theta R_b)(E_\theta R_b^a) - \dot{E}_\theta R_b^a\right]. \tag{3.19}$$

PROOF. For $\theta > 0$, by Wald's likelihood ratio identity,

$$E_\theta R_b^a = E_0\{R_b^a e^{\theta S_{\tau_b} - \tau_b \psi(\theta)}\}.$$

Using the inequality

$$P_0\{R_b > x\} \le U_0^+(1) \int_{x-2}^{\infty} P_0\{S_{\tau_+} > y\}\, dy,$$

for which a reference is Siegmund's (1985) Problem 8.8, it is not difficult to justify the use of dominated convergence to interchange differentiation with expectation to obtain

$$\begin{aligned} \dot{E}_\theta R_b^a &= E_0\Big\{R_b^a e^{\theta S_{\tau_b} - \tau_b \psi(\theta)}\Big(S_{\tau_b} - \tau_b \dot{\psi}(\theta)\Big)\Big\} \\ &= E_\theta\big\{R_b^a(S_{\tau_b} - \mu\tau_b)\big\}, \end{aligned}$$

where we have set $\mu = \mu(\theta) = \dot{\psi}(\theta)$. From this, making the substitution $S_{\tau_b} = b + R_b$ and rearranging give

$$E_\theta(\tau_b R_b^a) = \frac{1}{\mu}\left[bE_\theta R_b^a + E_\theta R_b^{a+1} - \dot{E}_\theta R_b^a\right].$$

However, clearly

$$(E_\theta \tau_b)(E_\theta R_b^a) = \frac{1}{\mu}(b + E_\theta R_b)(E_\theta R_b^a).$$

The desired result is obtained by subtracting the last two displays. □

In the case $a = 1$ the previous result takes the particularly neat form

$$\operatorname{Cov}_\theta(\tau_b, R_b) = \frac{1}{\mu}\left[\operatorname{Var}_\theta(R_b) - \dot{E}_\theta R_b\right].$$

To state the next result, which will be used in the proof of Lemma 4.5, define $\rho^{(a)}(\theta) := E_\theta R_\infty^a$.

COROLLARY 3.4. *Under the conditions of Theorem* 3.1, *there exist* A, $r > 0$ *and* $\theta^* > 0$ *such that*

$$\left|\dot{E}_\theta R_b^a - \dot{\rho}^{(a)}(\theta)\right| \le Ae^{-rb}$$

for all b *and for all* $\theta \in (0, \theta^*]$.

PROOF. Rearranging (3.19) gives

$$\dot{E}_\theta R_b^a = E_\theta R_b^{a+1} - (E_\theta R_b)(E_\theta R_b^a) - \mu \operatorname{Cov}_\theta(\tau_b, R_b^a).$$

Define

$$g(\theta) := \rho^{(a+1)}(\theta) - \rho(\theta)\rho^{(a)}(\theta) - \mu C^{(a)}(\theta).$$

Then using Corollary 2.3 and Theorem 3.1, it is easy to see that there exist A, $r > 0$ and $\theta^* > 0$ such that

$$\left|\dot{E}_\theta R_b^a - g(\theta)\right| \le Ae^{-rb} \tag{3.20}$$

for all $b \ge 0$ and all $\theta \in (0, \theta^*)$. Thus, letting $f_b(\theta) := E_\theta R_b^a$ and $f(\theta) := \rho^{(a)}(\theta)$, the situation we have here is that as $b \to \infty$, $f_b(\theta) \to f(\theta)$ for all θ and $\dot{f}_b(\theta) \to g(\theta)$ uniformly in $\theta \in (0, \theta^*)$. Under such circumstances, a theorem of elementary analysis [see, e.g., Apostol (1974), Theorem 9.13] implies that f is differentiable in $(0, \theta^*)$ and $\dot{f}(\theta) = g(\theta)$ there. Substituting $\dot{f}(\theta)$ for $g(\theta)$ in (3.20) gives the desired result. □

4. Moments of the first ladder height.

4.1. *Results*. Let $\tau_-^{(0)} := 0$, and for $n > 0$ define the nth weakly descending ladder epoch recursively by

$$\tau_-^{(n)} := \inf\left\{k > \tau_-^{(n-1)}\colon S_k \le S_{\tau_-^{(n-1)}}\right\}.$$

Define

$$\alpha^{(a)} := \int_{[0,\infty)} (E_0 R_x^a - E_0 R_\infty^a) U_0^-(dx), \tag{4.1}$$

where

$$U_0^-(x) \triangleq \sum_{n=0}^{\infty} P_0\left\{-S_{\tau_-^{(n)}} \le x\right\}$$

is the renewal function corresponding to the renewal process $\{-S_{\tau_-^{(n)}}\colon n = 0, 1, \dots\}$. To see that the right-hand side of (4.1) does indeed define a finite number, observe that by (2.14) and integration by parts, for any $a > 0$ there exist $r > 0$ and C such that

$$\int_{[0,\infty)} |E_0 R_x^a - E_0 R_\infty^a| U_0^-(dx) \le C\int_{[0,\infty)} e^{-rx} U_0^-(dx) = rC\int_{[0,\infty)} U_0^-(x) e^{-rx}\, dx.$$

However, since $U_0^-(\cdot)$ increases at a linear rate, clearly the last integral is finite.

With these definitions, now we can state our refinement of Theorem 1.1.

THEOREM 4.1. *Suppose that $\{F_\theta\colon \theta \in \Theta\}$ is a standard exponential family, with F_0 strongly nonlattice. Let $\alpha^{(a)}$ be as defined in* (4.1). *Then for any $a > 0$, as $\theta \downarrow 0$ we have*

$$\mu E_\theta\left(\tau_+ S_{\tau_+}^a\right) = \frac{1}{a+1} E_0 S_{\tau_+}^{a+1} + \left(\frac{1}{a+2} E_0 S_{\tau_+}^{a+2} + \alpha^{(a)}\right)\theta + O(\theta^2). \tag{4.2}$$

A proof of Theorem 4.1 will be given below. Next, just as Theorem 1.1 provided the main ingredient in the proof of Theorem 1.2, here it will also be easy to obtain from Theorem 4.1 the following refinement of Theorem 1.2.

THEOREM 4.2. *Under the conditions of Theorem* 4.1, *for any* $a > 0$, *as* $\theta \downarrow 0$ *we have*

$$(4.3)\qquad \begin{aligned} E_\theta S_{\tau_+}^a &= E_0 S_{\tau_+}^a + \frac{a}{a+1}\left(E_0 S_{\tau_+}^{a+1}\right)\theta \\ &\quad + \frac{1}{2}\left(\frac{a}{a+2} E_0 S_{\tau_+}^{a+2} - \alpha^{(a)}\right)\theta^2 + O(\theta^3). \end{aligned}$$

Theorem 4.2 certainly settles the "$o(\theta)$ versus $O(\theta^2)$" question discussed in Sections 1.3 and 1.4; in fact, it provides an expression for the coefficient of θ^2 and also settles the analogous "$o(\theta^2)$ versus $O(\theta^3)$" question raised by the existence of that coefficient. To mention a question that remains open, note that the coefficient of θ in Theorem 1.2 is "explicit," in the sense that it is a simple function of moments of S_{τ_+} under the $\theta = 0$ distribution. In contrast, as of yet no such explicit expression is available for the number $\alpha^{(a)}$ that appears in (4.3).

There is a special case in which we can give such an explicit expression for the coefficient of θ^2.

COROLLARY 4.3. *In addition to the conditions of Theorem* 4.1, *suppose the distribution* F_0 *is continuous and symmetric about* 0, *so that* $F_0(-x) = 1 - F_0(x)$. *Then*

$$(4.4)\qquad E_\theta S_{\tau_+} = \frac{1}{\sqrt{2}}\left[1 + \theta\rho_+ + \frac{\theta^2}{2}\left(\rho_+^2 + \frac{\kappa}{6}\right) + O(\theta^3)\right]$$

as $\theta \downarrow 0$.

For example, in the case of the normal family $\{N(\theta, 1)\}$, for which $\kappa = 0$, the result is particularly simple:

$$E_\theta S_{\tau_+} = \frac{1}{\sqrt{2}} e^{\theta\rho_+} + O(\theta^3).$$

4.2. *Proofs.* We start with a proof of Theorem 4.2 assuming the truth of Theorem 4.1, and then give a proof of Theorem 4.1.

PROOF THAT THEOREM 4.1 IMPLIES THEOREM 4.2. Let $a > 0$ and retain the notation $h(\theta) \triangleq E_\theta S_{\tau_+}^a$ from above. As Siegmund (1979) shows, for some $\varepsilon > 0$, h is continuously differentiable in $(0, \varepsilon)$ and continuous on $[0, \varepsilon]$. Therefore, for small θ_1 we may write

$$(4.5)\qquad h(\theta_1) = h(0) + \int_0^{\theta_1} h'(\theta)\, d\theta.$$

Combining (1.10), Theorem 1.2 with a replaced by $a+1$, and Theorem 4.1 gives

$$(4.6)\quad \begin{aligned} h'(\theta) &= E_\theta S_{\tau_+}^{a+1} - \mu E_\theta\left(\tau_+ S_{\tau_+}^a\right) \\ &= E_0 S_{\tau_+}^{a+1} + \frac{a+1}{a+2}\left(E_0 S_{\tau_+}^{a+2}\right)\theta + o(\theta) \\ &\quad - \frac{1}{a+1}E_0 S_{\tau_+}^{a+1} - \left(\frac{1}{a+2}E_0 S_{\tau_+}^{a+2} + \alpha^{(a)}\right)\theta + O(\theta^2) \\ &= \frac{a}{a+1}E_0 S_{\tau_+}^{a+1} + \left(\frac{a}{a+2}E_0 S_{\tau_+}^{a+2} - \alpha^{(a)}\right)\theta + o(\theta). \end{aligned}$$

Therefore, by substituting (4.6) into (4.5) and some elementary analysis,

$$(4.7)\quad \begin{aligned} E_{\theta_1} S_{\tau_+}^a = h(\theta_1) &= E_0 S_{\tau_+}^a + \frac{a}{a+1}\left(F_0 S_{\tau_+}^{a+1}\right)\theta_1 \\ &\quad + \frac{1}{2}\left(\frac{a}{a+2}E_0 S_{\tau_+}^{a+2} - \alpha^{(a)}\right)\theta_1^2 + o\left(\theta_1^2\right) \\ &= E_0 S_{\tau_+}^a + \frac{a}{a+1}\left(E_0 S_{\tau_+}^{a+1}\right)\theta_1 + O\left(\theta_1^2\right). \end{aligned}$$

Now change a to $a+1$ and θ_1 to θ in (4.7), and go back and substitute the result into the first line of (4.6). Then repeating the calculation in (4.6) leads to

$$h'(\theta) = \frac{a}{a+1}E_0 S_{\tau_+}^{a+1} + \left(\frac{a}{a+2}E_0 S_{\tau_+}^{a+2} - \alpha^{(a)}\right)\theta + O(\theta^2),$$

from which (4.5) and more elementary analysis gives the desired result (4.3). □

To begin the proof of Theorem 4.1, we have the following simple but useful lemma.

LEMMA 4.4. *Assume the conditions of Theorem* 4.1. *For $\theta \in \Theta$ define the measure U_θ^- by*

$$U_\theta^-(B) := \sum_{n=0}^{\infty} P_\theta\left\{\tau_-^{(n)} < \infty,\ -S_{\tau_-^{(n)}} \in B\right\}$$

for Borel subsets $B \subset [0,\infty)$. Then for all such subsets B,

$$U_\theta^-(B) \uparrow U_0^-(B) \quad \text{as } \theta \downarrow 0.$$

PROOF. Fix a Borel set $B \subset [0,\infty)$. By Wald's likelihood ratio identity, for any $n \geq 0$ we have

$$P_\theta\left\{\tau_-^{(n)} < \infty,\ -S_{\tau_-^{(n)}} \in B\right\} = E_0\left[\exp\left\{\theta S_{\tau_-^{(n)}} - \tau_-^{(n)}\psi(\theta)\right\};\ -S_{\tau_-^{(n)}} \in B\right].$$

However, as $\theta \downarrow 0$,

$$\exp\left\{\theta S_{\tau_-^{(n)}} - \tau_-^{(n)}\psi(\theta)\right\} \uparrow 1.$$

Therefore, by the monotone convergence theorem,

$$P_\theta\left\{\tau_-^{(n)} < \infty,\ -S_{\tau_-^{(n)}} \in B\right\} \uparrow P_0\left\{-S_{\tau_-^{(n)}} \in B\right\} \quad \text{as } \theta \downarrow 0. \tag{4.8}$$

From (4.8), the desired result is obtained by summing over $n \geq 0$ and applying the monotone convergence theorem once again. □

LEMMA 4.5. *Under the conditions of Theorem* 4.1,

$$\int_{[0,\infty)} (E_\theta R_x^a - E_\theta R_\infty^a) U_\theta^-(dx) = \alpha^{(a)} + O(\theta)$$

as $\theta \downarrow 0$, *where* $\alpha^{(a)}$ *is defined by* (4.1).

PROOF. Define

$$f_\theta(x) := E_\theta R_x^a - E_\theta R_\infty^a$$

and write the difference

$$\int_{[0,\infty)} (E_\theta R_x^a - E_\theta R_\infty^a) U_\theta^-(dx) - \alpha^{(a)}$$

as

$$\begin{aligned}
&\int_{[0,\infty)} f_\theta(x) U_\theta^-(dx) - \int_{[0,\infty)} f_0(x) U_0^-(dx) \\
&\quad = \int_{[0,\infty)} [f_\theta(x) - f_0(x)] U_\theta^-(dx) - \int_{[0,\infty)} f_0(x)[U_0^-(dx) - U_\theta^-(dx)] \\
&\quad =: J_1 - J_2.
\end{aligned}$$

To complete the proof, we will show that J_1 and J_2 are $O(\theta)$ as $\theta \downarrow 0$.

For J_1, write

$$f_\theta(x) - f_0(x) = \int_0^\theta \dot{f}_\eta(x)\, d\eta,$$

where the dot denotes differentiation (here with respect to η) as in Section 3. By Corollary 3.4, there exist A, $r > 0$ and $\theta^* > 0$ such that $|\dot{f}_\eta(x)| \leq Ae^{-rx}$ for all $x \geq 0$ and all $\eta \in (0, \theta^*]$. Therefore, letting $\theta \in [0, \theta^*]$, we have

$$|f_\theta(x) - f_0(x)| \leq \theta A e^{-rx}$$

for all $x \geq 0$, so that

$$|J_1| \leq \theta A \int_{[0,\infty)} e^{-rx} U_\theta^-(dx) \leq \theta A \int_{[0,\infty)} e^{-rx} U_0^-(dx),$$

where the second inequality follows from the monotonicity in Lemma 4.4. Thus, since the last integral is clearly finite, $J_1 = O(\theta)$.

For J_2, use Lemma 4.4 again to observe that $U_0^-(dx) - U_\theta^-(dx)$ is a nonnegative measure for each $\theta > 0$. From this and the bound $|f_0(x)| \leq Ae^{-rx}$, say, we obtain

$$|J_2| \leq A\int_{[0,\infty)} e^{-rx}\{U_0^-(dx) - U_\theta^-(dx)\}.$$

However,

$$\begin{aligned}\int_{[0,\infty)} e^{-rx}U_\theta^-(dx) &= \sum_{n=0}^{\infty} \int_{[0,\infty)} e^{-rx}P_\theta\left\{\tau_-^{(n)} < \infty,\ -S_{\tau_-^{(n)}} \in dx\right\}\\ &= \sum_{n=0}^{\infty} E_\theta\left\{\exp\left(rS_{\tau_-^{(n)}}\right);\ \tau_-^{(n)} < \infty\right\}\\ &= \sum_{n=0}^{\infty} \left[E_\theta\left(e^{rS_{\tau_-}};\ \tau_- < \infty\right)\right]^n\\ &= \left[1 - E_\theta\left(e^{rS_{\tau_-}};\ \tau - < \infty\right)\right]^{-1},\end{aligned}$$

so that

$$\begin{aligned}\int_{[0,\infty)} e^{-rx}\{U_0^-(dx) - U_\theta^-(dx)\} &= \frac{E_0\left(e^{rS_{\tau_-}}\right) - E_\theta\left(e^{rS_{\tau_-}};\ \tau_- < \infty\right)}{\left[1 - E_0\left(e^{rS_{\tau_-}}\right)\right]\left[1 - E_\theta\left(e^{rS_{\tau_-}};\ \tau_- < \infty\right)\right]}\\ &\leq \frac{E_0\left[e^{rS_{\tau_-}}\left(1 - e^{\theta S_{\tau_-} - \tau_-\psi(\theta)}\right)\right]}{\left[1 - E_0\left(e^{rS_{\tau_-}}\right)\right]^2},\end{aligned}$$

where the last inequality follows from the relations

$$E_\theta\left(e^{rS_{\tau_-}};\ \tau_- < \infty\right) = E_0\left(e^{rS_{\tau_-}}e^{\theta S_{\tau_-} - \tau_-\psi(\theta)}\right) \leq E_0\left(e^{rS_{\tau_-}}\right).$$

Therefore, since

$$\begin{aligned}E_0\left[e^{rS_{\tau_-}}\left(1 - e^{\theta S_{\tau_-} - \tau_-\psi(\theta)}\right)\right] &\leq E_0\left(1 - e^{\theta S_{\tau_-} - \tau_-\psi(\theta)}\right) = 1 - P_\theta\{\tau_- < \infty\}\\ &= P_\theta\{\tau_- = \infty\} = 1/(E_\theta\tau_+) = \mu/\left(E_\theta S_{\tau_+}\right),\end{aligned}$$

which is clearly $O(\theta)$, we have $J_2 = O(\theta)$. This completes the proof. □

Proof of Theorem 4.1. Recalling the definition (3.3) of Q_θ and applying duality, we obtain

$$Q_\theta(dx;0) \triangleq \sum_{n=0}^{\infty} P_\theta\{\tau_+ > n,\ -S_n \in dx\} = U_\theta^-(dx).$$

Therefore, by (3.2) with $b = 0$,

$$\begin{aligned}\operatorname{Cov}_\theta\left(\tau_+, S_{\tau_+}^a\right) &= \int_{[0,\infty)} \left(E_\theta R_x^a - E_\theta S_{\tau_+}^a\right)U_\theta^-(dx)\\ &= \int_{[0,\infty)} \left(E_\theta R_x^a\right)U_\theta^-(dx) - \left(E_\theta S_{\tau_+}^a\right)E_\theta\tau_+,\end{aligned}$$

so that

$$\mu E_\theta(\tau_+ S^a_{\tau_+}) = \mu \int_{[0,\infty)} (E_\theta R^a_x) U^-_\theta(dx)$$

$$(4.9) \qquad = \mu(E_\theta R^a_\infty) E_\theta \tau_+ + \mu \int_{[0,\infty)} (E_\theta R^a_x - E_\theta R^a_\infty) U^-_\theta(dx)$$

$$= \frac{1}{a+1} E_\theta S^{a+1}_{\tau_+} + \alpha^{(a)}\theta + O(\theta^2),$$

where the last equality uses (1.4), Lemma 4.5 and the fact that $\mu = \theta + O(\theta^2)$. Thus, the proof reduces to establishing an expansion for $E_\theta S^{a+1}_{\tau_+}$ up to $O(\theta^2)$. For the derivative we have

$$\dot{E}_\theta S^{a+1}_{\tau_+} = E_\theta S^{a+2}_{\tau_+} - \mu E_\theta(\tau_+ S^{a+1}_{\tau_+})$$

$$= E_\theta S^{a+2}_{\tau_+} - \left[\frac{1}{a+2} E_\theta S^{a+2}_{\tau_+} + O(\theta)\right]$$

$$= \frac{a+1}{a+2} E_0 S^{a+2}_{\tau_+} + O(\theta),$$

where the first equality is familiar from the proof that Theorem 4.1 implies Theorem 4.2, and the second and third equalities are simple consequences of (4.9) and Theorem 1.2, respectively. Integrating the last display gives

$$E_\theta S^{a+1}_{\tau_+} = E_0 S^{a+1}_{\tau_+} + \left[\frac{a+1}{a+2} E_0 S^{a+2}_{\tau_+}\right]\theta + O(\theta^2),$$

which, when combined with (4.9), gives the desired result (4.2). □

Finally, Corollary 4.3 may be proved by showing that (4.3) reduces to (4.4) when $a = 1$ and the assumed conditions hold. The fact that $E_0 S_{\tau_+} = 1/\sqrt{2}$ follows from (1.5). Using this, the coefficient of θ is immediate from the definition of ρ_+. The calculation that yields the coefficient of θ^2 proceeds as follows. By definition (4.1) and the given assumptions,

$$\alpha^{(1)} = \int_{[0,\infty)} (E_0 R_x - \rho_+) U(dx),$$

where $U(dx) := U^+_0(dx) = U^-_0(dx)$. Wald's equation gives

$$(4.10) \qquad U(x) = \sqrt{2}(E_0 R_x + x)$$

for $x \geq 0$, and for this discussion let us define $E_0 R_x = -x$ for $x < 0$, so that (4.10) holds for all real x. From this,

$$\int_{[0,c]} (E_0 R_x - \rho_+) U(dx) = \sqrt{2} \int_{[0,c]} (E_0 R_x)\, d(E_0 R_x)$$

$$+ \sqrt{2} \int_{[0,c]} E_0 R_x \, dx - \rho_+ U(c).$$

However,

$$\int_{[0,c]} (E_0 R_x)\, d(E_0 R_x) = \int_{(0,c]} (E_0 R_x)\, d(E_0 R_x) + \int_{\{0\}} (E_0 R_x)\, d(E_0 R_x)$$
$$= \tfrac{1}{2}\left[(E_0 R_c)^2 - (E_0 R_0)^2\right] + (E_0 R_0)^2$$
$$= \tfrac{1}{2}(E_0 R_c)^2 + \tfrac{1}{4},$$

and (3.16) gives

$$\int_{[0,c]} E_0 R_x\, dx = \frac{1}{2}\left[E_0 S_{\tau_+}^2 U(c) - E_0 R_c^2\right] = \frac{1}{\sqrt{2}} \rho_+ U(c) - \frac{1}{2} E_0 R_c^2.$$

Thus,

$$\int_{[0,c]} (E_0 R_x - \rho_+) U(dx) = \frac{1}{\sqrt{2}}\left[\frac{1}{2} + (E_0 R_c)^2 - E_0 R_c^2\right],$$

so that

$$\alpha^{(1)} = \frac{1}{\sqrt{2}}\left(\frac{1}{2} + \rho_+^2 - \rho_+^{(2)}\right).$$

Returning to (4.3), additional calculation and an invocation of (1.6), which here takes the form $\rho_+^{(2)} - \rho_+^2 = (\kappa + 3)/12$, show the coefficient of θ^2 to be $(\rho_+^2 - \kappa/6)/(2\sqrt{2})$.

Acknowledgments. This work is a revision of a portion of my Ph.D. dissertation, written at Stanford under the supervision of David Siegmund. All of the problems discussed in this paper build upon earlier work of Professor Siegmund, and I would like to express my sincere appreciation for his generous encouragement and extensive guidance. I benefitted from an enjoyable period of struggling together with Jim Fill on the main problem of this paper. I also enjoyed many stimulating conversations with T. L. Lai and Larry Shepp. Finally, I am very grateful to an anonymous referee for insightful comments and helpful suggestions that improved the presentation.

REFERENCES

APOSTOL, T. M. (1974). *Mathematical Analysis*. Addison-Wesley, Reading, Mass.

ASMUSSEN, S. (1987). *Applied Probability and Queues*. Wiley, New York.

CHANG, J. T. (1989). Random walks, moderate deviations, and the cusum procedure. Technical Report 2, Stanford Univ.

LAI, T. L. and SIEGMUND, D. (1979). A nonlinear renewal theory with applications to sequential analysis. II. *Ann. Statist.* **7** 60–76.

LORDEN, G. (1970). On excess over the boundary. *Ann. Math. Statist.* **41** 520–527.

LORDEN, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.* **42** 1897–1908.

PAGE, E. S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115.

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

SIEGMUND, D. (1985a). Corrected diffusion approximations and their applications. *Proc. Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. Le Cam and R. Olshen, eds.). Wadsworth, Belmont, Calif.

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

SIEGMUND, D. (1988). Approximate tail probabilities for the maxima of some random fields. *Ann. Probab.* **16** 487–501.

STAM, A. J. (1968). Two theorems in r-dimensional renewal theory. *Z. Wahrsch. Verw. Gebiete* **10** 81–86.

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

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.

DEPARTMENT OF STATISTICS
YALE UNIVERSITY
BOX 2179 YALE STATION
NEW HAVEN, CONNECTICUT 06520