Ann. Inst. Statist. Math.
Vol. 56, No. 1, 127–142 (2004)

BIAS OF ESTIMATOR OF CHANGE POINT DETECTED BY A CUSUM PROCEDURE

YANHONG WU

Department of Mathematics, University of the Pacific, Stockton, CA 95211, U.S.A.

(Received December 10, 2001; revised March 14, 2003)

Abstract. For independent observations from a standard one-parameter exponential family, the estimator of change point after being detected by a CUSUM procedure is defined as the last zero point of the CUSUM process before the alarm time. By assuming that the change occurs far away from beginning and the control limit is large, an explicit form for the bias of estimator is derived conditioning on the change being detected. By further assuming that the change magnitude and its reference value approach zero at the same order, the local second order expansion of the bias is obtained for numerical evaluation. It is found that, surprisingly, even in the normal distribution case, the bias is non-zero when the change magnitude equals to its reference value, in contrast to the continuous time analog and the fixed sample size case. Numerical results show that the approximations are quite satisfactory.

Key words and phrases: Change-point estimator, CUSUM procedure, quasi-stationary bias, random walk theory, strong renewal theorem, ladder epoches and ladder heights.

1. Introduction

Let $F_\theta(x)$ belong to a standard one-parameter exponential family of the form

$$dF_\theta(x) = \exp(x\theta - c(\theta))dF_0(x),$$

for $|\theta| \leq K(> 0)$ and $c(0) = c'(0) = 0$, $c''(0) = 1$. Also denote by $\gamma = c^{(3)}(0)$ and $\kappa = c^{(4)}(0)$. Throughout our discussion, we shall assume that $F_0(x)$ is strongly non-arithmetic in the sense that

$$\lim_{|\lambda|\to\infty} \sup \left| \int_{-\infty}^{\infty} e^{i\lambda x} dF_0(x) \right| < 1.$$

This condition implies that $F_\theta(x)$ is also strongly non-arithmetic uniformly for $|\theta| < \theta^*$ for some $\theta^* > 0$ (Siegmund (1979)).

Suppose $\{X_k\}$ are independent random variables which follow distribution $F_{\theta_0}(x)$ for $k \leq \nu$ and $F_\theta(x)$ for $k > \nu$, where $\theta_0 < 0 < \theta$ and ν is the change point. For a pre-selected reference value $\theta_1 > 0$ for θ such that $c(\theta_0) = c(\theta_1)$, the CUSUM procedure (Page (1954)), based on the likelihood ratio test, makes an alarm at the time

$$N = \inf\{n > 0 : T_n = \max(0, T_{n-1} + X_n) > d\}, \quad \text{with} \quad T_0 = 0,$$

where d is the control limit. At $\theta = \theta_1$, the maximum likelihood estimator of ν conditional on the change being detected is given by

$$\hat{\nu} = \max\{k < N : T_k = 0\},$$

which is the last zero point of T_k before the alarm time.

For notational convenience, let $E^\nu[\cdot]$ denote the expectation when the change point is at ν, $E_{\theta_0}[\cdot]$ and $E_\theta[\cdot]$ denote the expectations when no change is assumed, and $E_{\theta_0\theta}[\cdot]$ denote the expectation when both θ_0 and θ are involved.

In this paper, conditioning on the change being detected, we consider the asymptotic quasi-stationary bias and absolute bias of $\hat{\nu}$, defined as

$$\lim_{d\to\infty}\lim_{\nu\to\infty} E^\nu[\hat{\nu} - \nu \mid N > \nu]; \qquad \text{and} \qquad \lim_{d\to\infty}\lim_{\nu\to\infty} E^\nu[|\hat{\nu} - \nu| \mid N > \nu].$$

In Section 2, we first obtain an explicit asymptotic form for the asymptotic bias. Then, by using the strong renewal theorem, we are able to derive the local second order expansion for the bias as θ and θ_0 approach zero. Our main contribution is two-folds. First, we develop a general method for estimating the change point in the exponential family after sequential detection. Second, we show that there are some fundamental differences between the sequential sampling case and fixed sample size case. In Section 3, numerical results in the normal and exponential distribution case are presented by using the approximations. The results show that, in contrast to the fixed sample size case considered in Hinkley (1971) and Wu (1999) for the normal case, the bias is not negligible because of sequential sampling, even when the change magnitude equals the reference value. It is also different from the sequential sampling case in continuous time analog as considered in Srivastava and Wu (1999). Some independent and necessary results on the strong renewal theorem and ladder variables are presented in the Appendix for a complete presentation.

2. Quasi-stationary bias and second order expansion

Let $S_n = \sum_{i=1}^n X_i$ for $n > 0$ with $S_0 = 0$, and

$$\tau_x = \begin{cases} \inf\{n > 0 : S_n \le x\}; & \text{for} \quad x \le 0; \\ \inf\{n > 0 : S_n > x\}; & \text{for} \quad x > 0, \end{cases}$$

denote the boundary crossing time and $R_x = S_{\tau_x} - x$ the overshoot. In special, we denote by $\tau_- = \tau_0$ and $\tau_+ = \lim_{x\to 0+} \tau_x$ as the ladder epoches.

For notational convenience, we denote by $\{S_n'\}$ for $n \ge 0$, an independent copy of $\{S_n\}$ and

$$M = \sup_{0\le k<\infty} S_k',$$

as the maximum of $\{S_n'\}$ and $\sigma_M = \arg\sup_{0\le k<\infty} S_k'$ as the corresponding maximum point.

Conditioning on $N > \nu$, depending on whether $\hat{\nu} > \nu$ or $\hat{\nu} < \nu$, we can write

$$\hat{\nu} - \nu = (\hat{\nu} - \nu) I_{[\hat{\nu}>\nu]} - (\nu - \hat{\nu}) I_{[\hat{\nu}<\nu]},$$

where I_A denotes the indicator function of the event A.

Pollak and Siegmund ((1986), Theorem 2) showed that the quasi-stationary distribution of T_ν converges to the stationary distribution for T_ν as $d \to \infty$. That means,

$$\lim_{d\to\infty}\lim_{\nu\to\infty} P_{\theta_0}(T_\nu < y \mid N > \nu) = \lim_{d\to\infty}\lim_{\nu\to\infty} P_{\theta_0}(T_\nu < y) = P_{\theta_0}(M < y).$$

This implies that when the change occurs, T_ν is asymptotically distributed as M.

Thus, the event $\{\hat{\nu} > \nu\}$ is asymptotically equivalent to the event $\{\tau_{-M} < \infty\}$, i.e. the random walk S_n eventually comes back to zero with initial starting point M. Given $\hat{\nu} > \nu$, the bias $\hat{\nu} - \nu$ is asymptotically τ_{-M} plus the length, say γ_m for a CUSUM process T_n starting from zero until the last zero point time under $P_\theta(\cdot)$. Denote by $E[X; A] = E[XI_A]$. As $d, \nu \to \infty$, we have

$$\begin{aligned} E^\nu[\hat{\nu} - \nu; \hat{\nu} > \nu] &\to E_{\theta_0\theta}[\tau_{-M} + \gamma_m; \tau_{-M} < \infty] \\ &= E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] + E_\theta[\gamma_m] P_{\theta_0\theta}(\tau_{-M} < \infty). \end{aligned}$$

As noted in Wu (1999), γ_m is a geometric summation of iid random variables distributed as $\{\tau_-; \tau_- < \infty\}$ with terminating probability $P_\theta(\tau_- = \infty)$. Thus, we have

LEMMA 2.1.

$$E_\theta[\gamma_m] = \frac{E_\theta[\tau_-; \tau_- < \infty]}{P_\theta(\tau_- = \infty)}.$$

On the other hand, given $\hat{\nu} < \nu$, by looking at T_k backward in time starting from ν, we see that $T_{\nu-k}$ behaves like a random walk $\{S'_k\}$ for $k \geq 0$ with maximum value M and thus, $\nu - \hat{\nu}$ is asymptotically distributed as the maximum point σ_M. Thus, as $\nu, d \to \infty$, we have

$$E^\nu[\nu - \hat{\nu}; \hat{\nu} > \nu] \to E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty] = E_{\theta_0}[\sigma_M P_\theta(\tau_{-M} = \infty)].$$

A similar argument is referred to Srivastava and Wu (1999) for the continuous time analog.

Summarizing the above results, we get the following asymptotic first order result.

THEOREM 2.1. *As $\nu, d \to \infty$,*

$$\begin{aligned} E^\nu[\hat{\nu} - \nu \mid N > \nu] &\to E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] + P_{\theta_0\theta}(\tau_{-M} < \infty)\frac{E_\theta[\tau_-; \tau_- < \infty]}{P_\theta(\tau_- = \infty)} \\ &\quad - E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty]; \\ E^\nu[|\hat{\nu} - \nu| \mid N > \nu] &\to E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] + P_{\theta_0\theta}(\tau_{-M} < \infty)\frac{E_\theta[\tau_-; \tau_- < \infty]}{P_\theta(\tau_- = \infty)} \\ &\quad + E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty]. \end{aligned}$$

In the following, we shall derive the second order expansions for the asymptotic bias in Theorem 2.1 in order to investigate the bias numerically by further assuming both θ_0 and θ approach zero at the same order. The main theoretical tool is the strong renewal theorem and its applications to ladder variables. Following a referee's suggestion, the readers are encouraged to read the related results which are presented in the appendix before coming back for fully understanding the technique. Otherwise, it is better directly going to Section 3 to see the results in the normal case and some numerical demonstration.

There are five terms in Theorem 2.1 which will be evaluated in a sequence of lemmas. Most results generalize the ones of Wu (1999) in the fixed sample size case with normal distribution. However, the technique used here is much more general and can be used for any distribution of exponential family type and also raises more difficulties. An exception to the fixed sample size case is the term $E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty]$ which forms the difference between the fixed sample size case and sequential sampling case and also provides some new technical difficulties.

For notational convenience, we denote by $\Delta = \theta_1 - \theta_0$ and $\tilde{\Delta} = \theta - \tilde{\theta}$ where $c(\theta_0) = c(\theta_1)$ and $c(\theta) = c(\tilde{\theta})$. Also, let $\mu = c'(\theta)$, $\tilde{\mu} = c'(\tilde{\theta})$ and $\mu_i = c'(\theta_i)$ for $i = 0, 1$. Other notations are referred to the appendix.

The first lemma generalizes Lemma 4 of Wu (1999).

LEMMA 2.2. *As* $\theta \to 0$,

$$E_\theta[\tau_-; \tau_- < \infty] = \frac{E_0 S_{\tau_-}}{\tilde{\mu}} e^{\theta\rho_- + \theta^2/2(\rho_-^{(2)} - \rho_-^2 - 5\beta_1/E_0 S_{\tau_-})}(1 + o(\theta^2)),$$

where β_1 *is given in Lemma* A.1.

PROOF. By using Wald's Likelihood Ratio Identity by changing the measure $P_\theta(\cdot)$ to $P_{\tilde{\theta}}(\cdot)$ and Lemma A.1, we have

$$\begin{aligned} E_\theta[\tau_-; \tau_- < \infty] &= E_{\tilde{\theta}}[\tau_- e^{\Delta S_{\tau_-}}] \\ &= \frac{1}{\tilde{\mu}} E_{\tilde{\theta}} S_{\tau_-} + \Delta E_{\tilde{\theta}}(\tau_- S_{\tau_-}) + \frac{\Delta^2}{2} E_{\tilde{\theta}}(\tau_- S_{\tau_-}^2) + o(\Delta). \end{aligned}$$

After some algebraic simplifications, we get the result.

COROLLARY 2.1. *As* $\theta \to 0$,

$$E_\theta[\gamma_m] = -\frac{1}{\Delta\tilde{\mu}} e^{-(2\beta_1/E_0 S_{\tau_-})\theta^2}(1 + o(\theta^2)).$$

To evaluate $P_{\theta_0\theta}(\tau_{-M} < \infty)$, we follow a similar technique used in Wu (1999) and only the main steps are provided.

First, by conditioning on whether $M = 0$ and $M > 0$, we write

$$\begin{aligned} &P_{\theta_0\theta}(\tau_{-M} < \infty) \\ &\quad = P_\theta(\tau_- < \infty) P_{\theta_0}(\tau_+ = \infty) + P_{\theta_0\theta}(\tau_{-M} < \infty; M > 0). \end{aligned} \tag{2.1}$$

From Lemma A.2, we have

$$\begin{aligned} P_\theta(\tau_- < \infty) P_{\theta_0}(\tau_+ = \infty) &= \Delta_0 E_0 S_{\tau_+} e^{\theta_0\rho_+}(1 + \Delta E_0 S_{\tau_-}) + o(\theta^2) \\ &= \Delta_0 E_0 S_{\tau_+} e^{\theta_0\rho_+} - \frac{\Delta\Delta_0}{2} + o(\theta^2). \end{aligned}$$

For the second term of (2.1), by using Wald's Likelihood Ratio Identity, we have

$$P_\theta(\tau_{-x} < \infty) = E_{\tilde{\theta}} e^{\Delta S_{\tau_{-x}}} = e^{-\Delta x} E_{\tilde{\theta}} e^{\Delta R_{-x}},$$

and

$$P_{\theta_0}(M > x) = P_{\theta_0}(\tau_x < \infty) = e^{-\Delta_0 x} E_{\theta_1} e^{\Delta_0 R_x}.$$

From Corollary A.1, we know

$$E_0 R_x - \rho_+ = O(e^{-rx}); \qquad \text{and} \qquad E_0 R_{-x} - \rho_- = O(e^{-rx}),$$

as $x \to \infty$. Now, we write

$$\begin{aligned}
(2.2) \qquad & P_{\theta_0\theta}(\tau_{-M} < \infty, M > 0) \\
&= -\int_0^\infty P_\theta(\tau_{-x} < \infty) dP_{\theta_0}(M > x) \\
&= -\int_0^\infty E_{\tilde{\theta}} e^{\Delta S_{\tau_{-x}}} dE_{\theta_1} e^{-\Delta_0 S_{\tau_x}} \\
&= -\int_0^\infty e^{-\Delta(x-\rho_-)} de^{-\Delta_0(x+\rho_+)} \\
&\quad - \int_0^\infty e^{-\Delta(x-\rho_-)} d(e^{-\Delta_0(x+\rho_+)}(E_{\theta_1} e^{-\Delta_0(R_x-\rho_+)} - 1)) \\
&\quad - \int_0^\infty e^{-\Delta(x-\rho_-)}(E_{\theta_0} e^{\Delta(R_{-x}-\rho_-)} - 1) de^{-\Delta_0(x+\rho_+)} \\
&\quad - \int_0^\infty e^{-\Delta(x-\rho_-)}(E_{\theta_0} e^{\Delta(R_{-x}-\rho_-)} - 1) \\
&\quad \times d(e^{-\Delta_0(x+\rho_+)}(E_{\theta_1} e^{-\Delta_0(R_x-\rho_+)} - 1)).
\end{aligned}$$

The first term of (2.2) is

$$\frac{\Delta_0}{\Delta + \Delta_0} e^{\Delta\rho_- - \Delta_0\rho_+}.$$

The third term of (2.2) is approximated as

$$\Delta\Delta_0 \int_0^\infty (E_0 R_{-x} - \rho_-) dx + o(\theta^2).$$

The fourth term of (2.2) is

$$\Delta\Delta_0 \int_0^\infty E_0(R_{-x} - \rho_-) d(E_0 R_x - \rho_+) + o(\theta^2).$$

The third term of (2.2), by integrating by part, can be approximated as

$$e^{\Delta\rho_-}(P_{\theta_0}(\tau_+ < \infty) - e^{-\Delta_0\rho_+}) + \Delta\Delta_0 \int_0^\infty (E_0 R_x - \rho_+) dx + o(\theta^2).$$

Finally, we have the following approximation, which generalizes Lemma 6 of Wu (1999).

LEMMA 2.3. *As $\theta_0, \theta \to 0$ at the same order,*

$$P_{\theta_0\theta}(\tau_{-M} < \infty) = \frac{\Delta_0}{\Delta + \Delta_0} e^{\Delta\rho_- - \Delta_0\rho_+} + e^{\Delta\rho_-}(1 - e^{-\Delta_0\rho_+}) + \Delta\Delta_0\Big(-\frac{1}{2} - \rho_- E_0 S_{\tau_-} + \int_0^\infty (E_0 R_{-x} - \rho_-)dx + \int_0^\infty E_0(R_{-x} - \rho_-)d(E_0 R_x - \rho_+) + \int_0^\infty (E_0 R_x - \rho_+)dx\Big) + o(\theta^2).$$

The evaluation of $E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty]$ is similar and generalizes Lemma 7 of Wu (1999).

LEMMA 2.4. *As $\theta_0, \theta \to 0$ at the same order,*

$$E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] = -\frac{\Delta_0}{\tilde{\mu}}\left(\frac{1}{(\Delta + \Delta_0)^2} - \frac{\rho_-}{\Delta + \Delta_0}\right) - \frac{\Delta_0}{\tilde{\mu}}\Big(\frac{1}{2} + \rho_-(E_0 S_{\tau_+} - \rho_+) - \int_0^\infty (E_0 R_{-x} - \rho_-)dx - \int_0^\infty E_0(R_{-x} - \rho_-)d(E_0 R_x - \rho_+) - \int_0^\infty (E_0 R_x - \rho_+)dx\Big) + o(1).$$

PROOF. Again depending on whether $\{M = 0\}$ or $\{M > 0\}$, we have

$$E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] = E_\theta[\tau_-; \tau_- < \infty]P_{\theta_0}(\tau_+ = \infty) - \int_0^\infty E_\theta[\tau_{-x}; \tau_{-x} < \infty]dP_{\theta_0}(\tau_x < \infty). \quad (2.3)$$

The first term of (2.3) can be approximated by using Lemma 2.2 and Lemma A.2 as

$$E_\theta[\tau_-; \tau_- < \infty]P_{\theta_0}(\tau_+ = \infty) = \frac{\mu_0}{\tilde{\mu}} + o(1) = -\frac{\Delta_0}{2\tilde{\mu}} + o(1).$$

For the second term of (2.3), we use the similar techniques as in Lemma 2.3 and write

$$\begin{aligned} -\int_0^\infty & E_\theta[\tau_{-x}; \tau_{-x} < \infty]dP_{\theta_0}(\tau_x < \infty) \quad (2.4) \\ &= -\int_0^\infty E_{\tilde{\theta}}(\tau_{-x} e^{-\Delta(x - R_{-x})}) dE_{\theta_1} e^{-\Delta_0(x + R_x)} \\ &= -\int_0^\infty E_{\tilde{\theta}}(\tau_{-x}) e^{-\Delta(x - \rho_-)} dE_{\theta_1} e^{-\Delta_0(x + R_x)} + o(1) \end{aligned}$$

$$= -\frac{1}{\tilde{\mu}}\left[\int_0^\infty (-x + E_0 R_{-x}) e^{-\Delta(x-\rho_-)} de^{-\Delta_0(x+\rho_+)} - \Delta_0 \int_0^\infty (-x + E_0 R_{-x}) d(E_0 R_x - \rho_+)\right] + o(1).$$

The first term of (2.4) is approximated as

$$-\frac{\Delta_0}{\tilde{\mu}}\left[e^{\Delta\rho_- - \Delta_0\rho_+}\left(\frac{1}{(\Delta+\Delta_0)^2} - \frac{\rho_-}{\Delta+\Delta_0}\right) - \int_0^\infty (E_0 R_{-x} - \rho_-) dx\right] + o(1).$$

The second term of (2.4) is equal to

$$-\frac{\Delta_0}{\tilde{\mu}}\left[\int_0^\infty (x - \rho_-) d(E_0 R_x - \rho_+) - \int_0^\infty (E_0 R_{-x} - \rho_-) d(E_0 R_x - \rho_+)\right]$$
$$= -\frac{\Delta_0}{\tilde{\mu}}\left[\rho_-(E_0 S_{\tau_+} - \rho_+) - \int_0^\infty (E_0 R_x - \rho_+) dx - \int_0^\infty (E_0 R_{-x} - \rho_-) d(E_0 R_x - \rho_+)\right].$$

Combining the above results, we complete the proof.

Finally, we evaluate $E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty]$. We first write

$$E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty] = E_{\theta_0}\sigma_M - E_{\theta_0}[\sigma_M E_{\tilde{\theta}} e^{\Delta(-M+R_{-M})}]. \tag{2.5}$$

For the second term of (2.5), we write

$$E_{\theta_0}[\sigma_M E_{\tilde{\theta}} e^{\Delta(-M+R_{-M})}]$$
$$= E_{\theta_0}[\sigma_M e^{-\Delta M}] e^{\Delta\rho_-} + E_{\theta_0}[\sigma_M e^{-\Delta M}(E_{\tilde{\theta}} e^{\Delta R_{-M}} - e^{\Delta\rho_-})].$$

To evaluate $E_{\theta_0}[\sigma_M e^{-\Delta M}]$, we note that under $P_{\theta_0}(\cdot)$

$$(\sigma_M, M) =^d (\tau_+^{(K)}, S_{\tau_+^{(K)}}),$$

where $=^d$ denotes equivalence in distribution, $\tau_+^{(k)}$ is the k-th ladder epoches defined in the Appendix and

$$K = \sup\{k > 0 : \tau_+^{(k)} < \infty\}.$$

Note that K is a geometric random variable $P(K = k) = p^k(1-p)$ for $k \geq 0$, with terminating probability

$$1 - p = P_{\theta_0}(\tau_+ = \infty).$$

For given k, $(\tau_+^{(k)}, S_{\tau_+^{(k)}})$ is, in distribution, equivalent to the sum of k iid random variables distributed as (τ_+, S_{τ_+}).

Thus,

$$E_{\theta_0}[\sigma_M e^{-\Delta M}] = E_{\theta_0}[\tau_+^{(K)} e^{-\Delta S_{\tau_+^{(K)}}}]$$
$$= \sum_{k=1}^\infty E_{\theta_0}[\tau_+^{(k)} e^{-\Delta S_{\tau_+^{(k)}}}; K = k]$$

$$= \sum_{k=1}^{\infty} kE_{\theta_0}[\tau_+ e^{-\Delta S_{\tau+}}; \tau_+ < \infty](E_{\theta_0}(e^{-\Delta S_{\tau+}}; \tau_+ < \infty))^{k-1} P_{\theta_0}(\tau_+ = \infty)$$

$$= \frac{E_{\theta_0}[\tau_+ e^{-\Delta S_{\tau+}}; \tau_+ < \infty]}{(1 - E_{\theta_0}[e^{-\Delta S_{\tau+}}; \tau_+ < \infty])^2} P_{\theta_0}(\tau_+ = \infty).$$

The next two lemmas give the approximations for the related quantities.

LEMMA 2.5. *As $\theta_0, \theta \to 0$ at the same order,*

$$1 - E_{\theta_0}[e^{-\Delta S_{\tau+}}; \tau_+ < \infty]$$
$$= (\Delta + \Delta_0) E_0 S_{\tau_+} e^{-(\Delta-\theta_0)\rho_+ + 1/2(\Delta-\theta_0)^2(\rho_+^{(2)} - \rho_+^2) - (\theta_0^2/2)(\alpha_1/E_0 S_{\tau_+})}(1 + o(\theta^2)).$$

PROOF. Using Wald's Likelihood Ratio Identity, we have

$$1 - E_{\theta_0}[e^{-\Delta S_{\tau+}}; \tau_+ < \infty] = 1 - E_{\theta_1} e^{-(\Delta+\Delta_0)S_{\tau+}}.$$

A Taylor expansion following the lines of Lemma A.2 will give the result after some algebraic simplification.

In particular

$$P_{\theta_0}(\tau_+ = \infty) = \Delta_0 E_0 S_{\tau_+} e^{\theta_0 \rho_+ + 1/2\theta_0^2(\rho_+^{(2)} - \rho_+^2 - \alpha_1/E_0 S_{\tau_+})}(1 + o(\theta_0^2)).$$

The following lemma can be proved similarly as for Lemma 2.2, and its proof is omitted.

LEMMA 2.6. *As $\theta_0, \theta \to 0$ at the same order,*

$$E_{\theta_0}[\tau_+ e^{-\Delta S_{\tau+}}; \tau_+ < \infty]$$
$$= \frac{E_0 S_{\tau_+}}{\mu_1} e^{-(\Delta-\theta_0)\rho_+ + 1/2(\Delta-\theta_0)^2(\rho_+^{(2)} - \rho_+^2) - (\theta_1^2/2)(\alpha_1/E_0 S_{\tau_+}) - \theta_1(\Delta+\Delta_0)(\alpha_1/E_0 S_{\tau_+})}$$
$$\times (1 + o(\theta^2)).$$

In particular,

$$E_{\theta_0}[\tau_+; \tau_+ < \infty] = \frac{E_0 S_{\tau_+}}{\mu_1} e^{\theta_0 \rho_+ + \theta_0^2/2(\rho_+^{(2)} - \rho_+^2 - 5\alpha_1/E_0 S_{\tau_+})}(1 + o(\theta^2)).$$

On the other hand,

$$(2.6) \qquad E_{\theta_0}[\sigma_M e^{-\Delta M}(E_{\tilde{\theta}} e^{\Delta R_{-M}} - e^{\Delta \rho_-})]$$
$$= \Delta E_{\theta_0}[\sigma_M(E_0 R_{-M} - \rho_-)](1 + o(1))$$
$$= -\Delta \int_0^{\infty} E_{\theta_0}[\sigma_x \mid M = x](E_0 R_{-x} - \rho_-) dP_{\theta_0}(M > x)$$
$$= \Delta\Delta_0 \int_0^{\infty} E_{\theta_0}[\sigma_x \mid M = x](R_0 R_{-x} - \rho_-) d(x + E_0 R_x)(1 + o(1)).$$

Since

$$E_{\theta_0}[S_{\tau_+}; \tau_+ < \infty] = E_0 S_{\tau_+}(1 + o(1)),$$

as $\theta_0 \to 0$. Thus, $K = O_p(x)$, where $O_p(\cdot)$ means at the same order in probability. This implies

$$E_{\theta_0}[\sigma_x \mid M = x] = O\left(\frac{x}{\mu_0}\right).$$

Thus, (2.6) is at the order of $O(\theta)$.

By letting $\Delta = 0$, the first term of (2.5) can be evaluated by combining Lemmas 2.5 and 2.6.

LEMMA 2.7. *As* $\theta_0 \to 0$,

$$E_{\theta_0}\sigma_M = \frac{E_{\theta_0}[\tau_+; \tau_+ < \infty]}{P_{\theta_0}(\tau_+ = \infty)} = \frac{1}{\Delta_0 \mu_1} e^{-(2\alpha_1/E_0 S_{\tau_+})\theta_0^2}(1 + o(\theta_0^2)).$$

Finally, we have the following result.

LEMMA 2.8. *As* $\theta_0, \theta \to 0$,

$$\begin{aligned} &E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty] \\ &= \frac{1}{\Delta_0\mu_1} e^{-(2\alpha/E_0 S_{\tau_+})\theta_0^2} \\ &\quad - \frac{\Delta_0}{\mu_1(\Delta + \Delta_0)^2} e^{\gamma\Delta/3 - 2\theta(\theta-\theta_0)(\rho_+^{(2)} - \rho_+^2 - \alpha_1/E_0 S_{\tau_+}) - 2(\theta-\theta_0)^2(\alpha_1/E_0 S_{\tau_+})}(1 + o(\theta^2)). \end{aligned}$$

Combining Lemmas 2.1–2.8 and A.1–A.2, we have the following second order expansion for the asymptotic bias of $\hat{\nu}$.

THEOREM 2.2. *As* $\theta_0, \theta \to 0$ *at the same order, we have*

$$\begin{aligned} \lim_{d\to\infty} \lim_{\nu\to\infty} E^\nu[\hat{\nu} - \nu \mid N > \nu] = &-\frac{1}{\tilde{\mu}\Delta}\left(\frac{\Delta}{\Delta + \Delta_0} e^{\Delta\rho_- - \Delta_0\rho_+} + e^{\Delta\rho_-}(1 - e^{-\Delta_0\rho_+})\right) \\ &- \frac{\Delta}{\tilde{\mu}}\left(\frac{1}{(\Delta + \Delta_0)^2} - \frac{\rho_-}{\Delta + \Delta_0}\right) e^{\Delta\rho_- - \Delta_0\rho_+} \\ &+ \frac{\theta_0}{\theta - \theta_0}\frac{\beta_1}{E_0 S_{\tau_+}} + \frac{2\theta_0}{\theta}\rho_-\rho_+ \\ &- \frac{\theta}{\theta - \theta_0}\left(\rho_+^{(2)} - \rho_+^2 - \frac{\alpha_1}{E_0 S_{\tau_+}}\right) + o(1). \end{aligned}$$

Similar result can be obtained for $E^\nu[|\hat{\nu} - \nu| \mid N > \nu]$ and is omitted.

In special, when $\theta = \theta_1$ we have the following result.

COROLLARY 2.2. *As* $\theta = \theta_1 \to 0$,

$$\lim_{d\to\infty}\lim_{\nu\to\infty} E^{\nu}[\hat{\nu} - \nu \mid N > \nu]$$
$$= -\frac{3}{4\Delta_0}\left(\frac{1}{\mu_0} + \frac{1}{\mu_1}\right) - \frac{\gamma}{12}\left(\frac{1}{\mu_0} + \frac{1}{\mu_1}\right) - \frac{1}{2}\frac{\beta_1}{E_0 S_{\tau_-}}$$
$$- \frac{1}{2}\left(\rho_+^{(2)} - \rho_+^2 - \frac{\alpha_1}{E_0 S_{\tau_+}}\right) + o(1)$$
$$= -\frac{\gamma}{4\theta_0} + \frac{17\gamma^2}{288} - \frac{\kappa}{16} - \frac{1}{2}\frac{\beta_1}{E_0 S_{\tau_-}} - \frac{1}{2}\left(\rho_+^{(2)} - \rho_+^2 - \frac{\alpha_1}{E_0 S_{\tau_+}}\right) + o(1).$$

PROOF. The first equation is a direct simplification of Theorem 2.2. For the second equation, we note that as $\theta_0 \to 0$,

$$\mu_1 = \theta_1 e^{(\gamma/2)\theta_1 + \theta_1^2(\kappa/6 - \gamma^2/8)}(1 + o(\theta_1^2)),$$
$$\Delta_0 = 2\theta_1 e^{(\gamma/6)\theta_1 + (\theta_1^2/24)\gamma^2}(1 + o(\theta_1^2)),$$
$$\theta_0 = -\theta_1 e^{(\gamma/3)\theta_1 + (\gamma^2/18)\theta_1^2}(1 + o(\theta_1^2)),$$
$$\mu_0 = -\theta_1 e^{-(\gamma/6)\theta_1 + \theta_1^2(\kappa/6 - (17/72)\gamma^2)}(1 + o(\theta_1^2)).$$

Some tedious simplifications give the result.

Therefore, the local bias of $\hat{\nu}$ is largely affected by the skewness γ. If $\gamma > 0$, the local bias becomes positive. If $F_0(x)$ is symmetric, from Corollary A.2, we have $\rho_+^{(2)} - \rho_+^2 - \frac{\alpha_1}{E_0 S_{\tau_+}} = \frac{\kappa}{6}$, and thus

$$E^{\nu}[\hat{\nu} - \nu \mid N > \nu] \approx -\frac{7}{48}\kappa - \frac{1}{2}\frac{\beta_1}{E_0 S_{\tau_-}} + o(1),$$

which is surprisingly a non-zero constant, in contrast to the fixed sample size case as given in normal case of next section.

3. Two examples

In this section, we discuss two special cases: normal and exponential distributions.

3.1 *Normal distribution*

Here $c(\theta) = \frac{\theta^2}{2}$ and $F_0(x) = \Phi(x)$ the standard normal distribution, which is symmetric. Thus, $\gamma = \kappa = 0$, and from Corollary A.2

$$\rho_+ = -\rho_- \approx 0.583, \qquad E_0 S_{\tau_+} = -E_0 S_{\tau_-} = \frac{1}{\sqrt{2}},$$
$$\rho_+^{(2)} - \rho_+^2 - \frac{\alpha_1}{E_0 S_{\tau_+}} = \rho_-^{(2)} - \rho_-^2 - \frac{\beta_1}{E_0 S_{\tau_-}} = 0,$$

and

$$\frac{\beta_1}{E_0 S_{\tau_-}} = \rho_+^{(2)} - \rho_+^2 = \frac{1}{4}.$$

The approximations for the related quantities are simplified as

$$E_\theta(\gamma_m) = \frac{1}{2\theta^2} - \frac{1}{4} + o(1);$$

$$P_{\theta_0\theta}(\tau_{-M} < \infty) = -\frac{\theta_0}{\theta - \theta_0} e^{-\theta(\theta-\theta_0)} + o(\theta^2);$$

$$E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] = -\frac{\theta_0}{2\theta(\theta-\theta_0)^2} e^{(\theta-\theta_0)^2} + o(1);$$

$$E_{\theta_0\theta}[\sigma_M; \tau_{-M} = \infty] = \frac{1}{2\theta_0^2} - \frac{1}{2(\theta-\theta_0)^2} + o(1).$$

Summarizing the above results, we have the following corollary.

COROLLARY 3.1. *As $\theta_0, \theta \to 0$ at the same order,*

$$\lim_{d\to\infty} \lim_{\nu\to\infty} E^\nu[\hat{\nu} - \nu \mid N > \nu] = \frac{1}{2\theta^2} - \frac{1}{2\theta_0^2} + \frac{\theta_0}{4(\theta-\theta_0)} + o(1),$$

$$\lim_{d\to\infty} \lim_{\nu\to\infty} E^\nu[|\hat{\nu} - \nu| \mid N > \nu] = \frac{1}{2}\left(\frac{1}{\theta^2} + \frac{1}{\theta_0^2} - \frac{2}{(\theta-\theta_0)^2}\right) + \frac{\theta_0}{4(\theta-\theta_0)} + o(1).$$

At $\theta = \theta_1 = -\theta_0$,

$$\lim_{d\to\infty} \lim_{\nu\to\infty} E^\nu[\hat{\nu} - \nu \mid N > \nu] = -\frac{1}{8} + o(1),$$

$$\lim_{d\to\infty} \lim_{\nu\to\infty} E^\nu[|\hat{\nu} - \nu| \mid N > \nu] = \frac{3}{4\theta_0^2} - \frac{1}{8} + o(1).$$

Remark. Wu (1999) considered the bias of the estimator in the large fixed sample size case, which corresponds to the maximum point of a two-sided random walk, and obtained the following result:

$$E^\nu[\hat{\nu} - \nu] = \frac{1}{2\theta^2} - \frac{1}{2\theta_0^2} + \frac{1}{4}\frac{\theta+\theta_0}{\theta-\theta_0} + o(1);$$

and at $\theta = -\theta_0$,

$$E^\nu[|\hat{\nu} - \nu|] = \frac{3}{4\theta_0^2} - \frac{1}{4} + o(1).$$

Srivastava and Wu (1999) also considered the continuous time analog in sequential sampling case which gives

$$E^\nu[\hat{\nu} - \nu \mid N > \nu] \to \frac{1}{2\theta^2} - \frac{1}{2\theta_0^2},$$

and at $\theta = -\theta_0$,

$$E^\nu[|\hat{\nu} - \nu| \mid N > \nu] \to \frac{3}{4\theta_0^2}.$$

We see that the sequential sampling plan has a local effect at the second order and is negative at $\theta = -\theta_0$.

To show the accuracy of the second order approximations, we conduct a simple simulation study. For $d = 10$ and $\theta_0 = -0.25, -0.5$ we let $\nu = 50$ and 100. 1000

Table 1. Biases in the normal case.

ν	θ_0	θ	$E[\hat{\nu}-\nu \mid N>\nu]$	$E[\lvert\hat{\nu}-\nu\rvert \mid N>\nu]$
50	−0.25	0.25	0.113(−0.125)	9.737(11.875)
		0.5	−4.902(−6.083)	7.090(8.139)
		0.75	−5.682(−7.174)	6.376(7.826)
		1.0	−5.768(−7.55)	6.188(7.81)
	−0.5	0.5	0.268(−0.125)	3.052(2.875)
		0.75	−1.302(−1.211)	2.338(2.149)
		1.0	−1.673(−1.583)	2.135(1.972)
100	−0.25	0.25	1.368(−0.125)	11.728(11.875)
		0.5	−5.644(−6.083)	7.768(8.139)
		0.75	−6.181(−7.174)	6.942(7.826)
		1.0	−6.250(−7.55)	6.520(7.81)
	−0.5	0.5	−0.223(−0.125)	3.052(2.875)
		0.75	−1.109(−1.211)	2.208(2.149)
		1.0	−1.564(−1.583)	2.084(1.972)

replications of the CUSUM charts are simulated for each case. Only those runs with $N > \nu$ are used for calculating $\hat{\nu}$. Table 1 gives the simulated results. The approximated values from Corollary 3.1 are given in the bracket. We see that the approximations are generally good. The case $\nu = 100$ shows quite satisfactory results. Also, we see that approximations for the case $\theta_0 = -0.5$ perform better than those for the case $\theta_0 = -.25$. The reason is that our results are given by first assuming $d, \nu \to \infty$ and then letting $\theta_0, \theta \to 0$. The effect of ν is very little. However, as the local bias is at the order $O(1/\theta_0^2)$ at $\theta = -\theta_0$, which approaches infinity as $\theta_0 \to 0$, there could be an error term at the order, say, $O(1/(d\theta))$ for finitely large d. Thus, the approximation may perform better for $\theta = -\theta_0 = 0.5$. The case when θd approaches a constant, called moderate deviation as considered in Chang (1992), is definitely worthy for a future study.

3.2 *Exponential distribution*

Here, we are interested in quick detection of increment in the mean of an exponential distribution from the initial mean 1.

Let $f_0(x) = e^{-(x+1)}$ for $x \geq -1$. Then $c(\theta) = -\theta - \ln(1-\theta)$ for $|\theta| < 1$. Thus, $c'(\theta) = \theta/(1-\theta)$ and $c^{(k)}(\theta) = \frac{(k-1)!}{(1-\theta)}$ for $k \geq 2$ and $\gamma = 2$ and $\kappa = 6$.

Because of the memoryless property, R_x follows exp(1) for any $x \geq 0$. Also, it is noted that S_{τ_-} follows $U(-1,0)$ (Siegmund (1985), p. 186, Problem 8.10). Thus,

$$E_0 S_{\tau_+} = 1, \quad \rho_+ = 1, \quad \rho_+^{(2)} = 2,$$
$$E_0 S_{\tau_-} = -\frac{1}{2}, \quad \rho_- = -\frac{1}{3}, \quad \rho_-^{(2)} = \frac{1}{6}.$$

Since $\alpha_1 = 0$, we have $\frac{\beta_1}{E_0 S_{\tau_-}} = \frac{7}{18}$ from Corollary A.2.

From Theorem 2.2, we have the following result.

COROLLARY 3.2. *As $\theta, \theta_0 \to 0$ at the same order,*

$$\begin{aligned}\lim_{d\to\infty}\lim_{\nu\to\infty} E^{\nu}[\hat{\nu}-\nu \mid N>\nu] = & -\frac{\Delta+2\Delta_0}{\tilde{\mu}(\Delta+\Delta_0)^2}e^{-(1/2)\Delta-\Delta_0} - \frac{1}{\tilde{\mu}\Delta}e^{-(\Delta/2)}(1-e^{-\Delta_0}) \\ & - \frac{\Delta_0}{2\tilde{\mu}(\Delta+\Delta_0)}e^{-(1/2)\Delta-\Delta_0} \\ & - \frac{1}{\Delta_0\mu_1} + \frac{\Delta_0}{\mu_1(\Delta+\Delta_0)^2}e^{(2/3)\Delta} \\ & - \frac{\theta}{\theta-\theta_0} - \frac{\theta_0}{\theta} + \frac{7}{18}\frac{\theta_0}{\theta-\theta_0} + o(1).\end{aligned}$$

At $\theta = -\theta_0$,

$$\lim_{d\to\infty}\lim_{\nu\to\infty} E^{(\nu)}[\hat{\nu}-\nu \mid N>\nu] = -\frac{1}{2\theta_0} - \frac{17}{24} + o(1).$$

We see that due to the asymmetry of $F_0(x)$, the local bias becomes positive as θ_0 is small.

Acknowledgements

I am grateful to two referees' comments which improved the presentation and to Professor R. Keener for discussions about the results in the normal distribution case. The original version of the paper was written when the author was visiting the Department of Statistics, University of Michigan.

Appendix

Strong renewal theorem and ladder variables

In this appendix, we state the strong renewal theorem and its applications to the approximations for the probabilities and moments associated with the ladder variables. The strong renewal theorem is given in Stone (1965) and developed in Siegmund (1979) for the exponential family.

Let $\tau_-^{(0)} = \tau_+^{(0)} = 0$ and $\tau_+^{(1)} = \tau_+$ and $\tau_-^{(1)} = \tau_-$. For $n > 1$, we define

$$\tau_+^{(n)} = \inf\{k : S_k > S_{\tau_+^{(n-1)}}\};$$
$$\tau_-^{(n)} = \inf\{k : S_k \le S_{\tau_-^{(n-1)}}\},$$

as the k-th ladder epoches. Denote for $x \ge 0$,

$$U_\theta^+(x) = \sum_{n=0}^{\infty} P_\theta(S_{\tau_+^{(n)}} \le x); \qquad U_{\theta_0}^-(x) = \sum_{n=0}^{\infty} P_{\theta_0}(-S_{\tau_-^{(n)}} \le x),$$

as the renewal functions of $\{S_n\}$. The following uniform strong renewal theorem is stated in Chang (1992) which formalizes the results of Siegmund (1979).

UNIFORM STRONG RENEWAL THEOREM. *There exist positive numbers r, $\theta *$ and C such that*

$$\left| U_\theta^+(x) - \frac{x}{E_\theta S_{\tau_+}} - \frac{E_\theta S_{\tau_+}^2}{2(E_\theta S_{\tau_+})^2} \right| \leq Ce^{-rx};$$
$$\left| U_{\theta_0}^-(x) + \frac{x}{E_{\theta_0} S_{\tau_-}} - \frac{E_{\theta_0} S_{\tau_-}^2}{2(E_{\theta_0} S_{\tau_-})^2} \right| \leq Ce^{-rx},$$

uniformly for $-\theta^* \leq \theta_0 \leq 0 \leq \theta \leq \theta^*$ *and* $x \geq 0$.

An immediate consequence of the strong renewal theorem is the exponential convergence rate for the overshoot, which is stated in Chang (1992).

COROLLARY A.1. *There exist positive numbers r, θ^* and C such that*

$$|P_\theta(R_x \leq y) - P_\theta(R_\infty \leq y)| \leq Ce^{-r(x+y)};$$
$$|P_{\theta_0}(R_{-x} > -y) - P_{\theta_0}(R_{-\infty} > -y)| \leq Ce^{-r(x+y)},$$

uniformly for $-\theta^* \leq \theta_0 \leq 0 \leq \theta \leq \theta^*$ *and* $x, y \geq 0$, *where*

$$P_\theta(R_\infty \leq y) = \frac{1}{E_\theta S_{\tau_+}} \int_0^y P_\theta(S_{\tau_+} > x)dx;$$

and

$$P_{\theta_0}(R_{-\infty} > -y) = \frac{1}{-E_{\theta_0} S_{\tau_-}} \int_0^y P_{\theta_0}(S_{\tau_-} < -y)dy.$$

We denote by $\rho_\pm^{(k)} = E_0 R_{\pm\infty}^k$ and $\rho_\pm = \rho_\pm^{(1)}$.

One important application of the strong renewal theorem is to deliver very accurate approximations for the (joint) moments of ladder variables. The following lemma gives two very important approximations which is given in Lemma 10.27 of Siegmund (1985) and extended in Chang (1992).

LEMMA A.1. *As* $0 < \theta \to 0$, *for* $k \geq 1$,

$$E_\theta S_{\tau_+}^k = E_0 S_{\tau_+}^k + \frac{k}{k+1} E_0 S_{\tau_+}^{k+1} \theta + \frac{\theta^2}{2} \left(\frac{k}{k+2} E_0 S_{\tau_+}^{k+2} - \alpha_k \right) + o(\theta^2),$$
$$\mu E_\theta(\tau_+ S_{\tau_+}^k) = \frac{1}{k+1} E_0 S_{\tau_+}^{k+1} + \theta \left(\frac{1}{k+2} E_0 S_{\tau_+}^{k+2} + \alpha_k \right) + o(\theta),$$

where

$$\alpha_k = \int_{0-}^\infty (E_0 R_x^k - \rho_+^{(k)}) U_0^-(dx).$$

Similarly, as $0 > \theta_0 \to 0$,

$$E_{\theta_0} S_{\tau_-}^k = E_0 S_{\tau_-}^k + \frac{k}{k+1} \theta_0 E_0 S_{\tau_-}^{k+1} + \frac{\theta_0^2}{2} \left(\frac{k}{k+2} E_0 S_{\tau_-}^{k+2} - \beta_k \right) + o(\theta_0^2),$$
$$\mu_0 E_{\theta_0}(\tau_- S_{\tau_-}^k) = \frac{1}{k+1} E_0 S_{\tau_-}^{k+1} + \theta_0 \left(\frac{1}{k+2} E_0 S_{\tau_-}^{k+2} + \beta_k \right) + o(\theta_0),$$

where

$$\beta_k = \int_{0+}^{\infty} (E_0 R_{-x}^k - \rho_-^{(k)}) U_0^+(dx).$$

An interesting application of Lemma A.1 is the following two different versions of approximation for the probability $P_\theta(\tau_- = \infty)$.

LEMMA A.2. *As* $\theta \to 0$,

$$\begin{aligned} P_\theta(\tau_- = \infty) &= \frac{\mu}{E_0 S_{\tau_+}} e^{-\rho_+\theta - 1/2(\rho_+^{(2)} - \rho_+^2 - \alpha_1/E_0 S_{\tau_+})\theta^2}(1 + o(\theta^2)), \\ &= -\Delta E_0 S_{\tau_-} e^{\theta\rho_- + 1/2(\rho_-^{(2)} - \rho_-^2 - \beta_1/E_0 S_{\tau_-})\theta^2}(1 + o(\theta^2)). \end{aligned}$$

PROOF. For the first approximation, by using the Wiener-Hopf equation and Lemma A.1, we have

$$\begin{aligned} P_\theta(\tau_- = \infty) &= \frac{1}{E_\theta(\tau_+)} = \frac{\mu}{E_\theta S_{\tau_+}} \\ &= \mu\left(E_0 S_{\tau_+} + \frac{\theta}{2} E_0 S_{\tau_+^2} + \frac{\theta^2}{2}\left(\frac{1}{3} E_0 S_{\tau_+^3} - \alpha_1\right) + o(\theta^2)\right)^{-1} \\ &= \frac{\mu}{E_0 S_{\tau_+}}\left(1 + \theta\rho_+ + \frac{\theta^2}{2}\left(\rho_+^{(2)} - \frac{\alpha_1}{E_0 S_{\tau_+}}\right) + o(\theta^2)\right)^{-1}, \end{aligned}$$

which is equivalent to the required result.

By using the Wald's Likelihood Ratio Identity first and then a Taylor expansion, we have the second approximation:

$$\begin{aligned} P_\theta(\tau_- = \infty) &= 1 - E_{\tilde\theta} e^{\Delta S_{\tau_-}} \\ &= -\left(\Delta E_{\tilde\theta} S_{\tau_-} + \frac{\Delta^2}{2} E_{\tilde\theta} S_{\tau_-}^2 + \frac{\Delta^3}{6} E_{\tilde\theta} S_{\tau_-}^3\right) + o(\Delta^3) \\ &= -\Delta\Bigg(E_0 S_{\tau_-} + \frac{\tilde\theta}{2} E_0 S_{\tau_-}^2 + \frac{\tilde\theta^2}{2}\left(\frac{1}{3} E_0 S_{\tau_-}^3 - \beta_1\right) \\ &\qquad + \frac{\Delta}{2}\left(E_0 S_{\tau_-}^2 + \frac{2}{3}\tilde\theta E_0 S_{\tau_-}^3\right) + \frac{\Delta^2}{6} E_0 S_{\tau_-}^3 + o(\Delta^2)\Bigg) \\ &= -\Delta E_0 S_{\tau_-}\left(1 + \theta\rho_- + \frac{\theta^2}{2}\left(\rho_-^{(2)} - \frac{\beta_1}{E_0 S_{\tau_-}}\right) + o(\theta^2)\right), \end{aligned}$$

which is equivalent to the required result.

An interesting consequence of this lemma is the following link between the moments of X_1 and overshoot under $P_0(\cdot)$.

By matching the two versions of approximation in Lemma A.2, we have

$$-\frac{\Delta}{\mu} E_0 S_{\tau_+} E_0 S_{\tau_-} = e^{-\theta(\rho_+ + \rho_-) - \theta^2/2(\rho_+^{(2)} + \rho_-^{(2)} - \rho_+^2 - \rho_-^2 - \alpha_1/E_0 S_{\tau_+} - \beta_1/E_0 S_{\tau_-})}(1 + o(\theta^2)).$$

A Taylor expansion around zero gives

$$\mu = c'(\theta) = \theta + \frac{\theta^2}{2}\gamma + \frac{\theta^3}{6}\kappa = \theta e^{(\gamma/2)\theta+\theta^2(\kappa/6-\gamma^2/8)}(1+o(\theta^2)),$$

and

$$\Delta = 2\theta + \frac{\gamma}{3}\theta^2 + \frac{\gamma^2}{9}\theta^3 = 2\theta e^{(\gamma/6)\theta+(\theta^2/24)\gamma^2}(1+o(\theta^2)).$$

Thus, we have the following identities:

COROLLARY A.2.

$$E_0 S_{\tau_+} E_0 S_{\tau_-} = -\frac{1}{2};$$

$$\rho_+ + \rho_- = \frac{\gamma}{3};$$

$$\rho_+^{(2)} + \rho_-^{(2)} - \rho_+^2 - \rho_-^2 - \frac{\alpha_1}{E_0 S_{\tau_+}} - \frac{\beta_1}{E_0 S_{\tau_-}} = \frac{1}{3}(\kappa - \gamma^2).$$

REFERENCES

Chang, J. T. (1992). On moments of the first ladder height of random walks with small drift, *The Annals of Applied Probability*, **2**, 714–738.

Hinkley, D. V. (1971). Inference about the change point from cumulative sum tests, *Biometrika*, **58**, 509–523.

Page, E. S. (1954). Continuous inspection schemes, *Biometrika*, **41**, 100–114.

Pollak, M. and Siegmund, D. (1986). Convergence of quasi-stationary to stationary distribution for stochastically monotone Markov processes, *Journal of Applied Probability*, **23**, 215–220.

Siegmund, D. (1979). Corrected diffusion approximations in certain random walk problems, *Advances in Applied Probability*, **11**, 701–719.

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

Srivastava, M. S. and Wu, Y. (1999). Quasi-stationary biases of change point and change magnitude estimation after sequential CUSUM test, *Sequential Analysis*, **18**, 203–216.

Stone, C. (1965). On moment generating functions and renewal theory, *The Annals of Mathematical Statistics*, **36**, 1298–1301.

Wu, Y. (1999). Second order expansions for the moments of minimum point of an unbalanced two-sided normal random walk, *Annals of the Institute of Statistical Mathematics*, **51**, 187–200.