2
Change-Point Estimation

2.1 Asymptotic Quasistationary Bias

For notational convenience, we denote by \( \{S'_n\} \), for \( n \geq 0 \), an independent copy of \( \{S_n\} \) and

\[
M = \sup_{0 \leq k < \infty} S'_k,
\]

as the maximum of \( \{S'_n\} \) when \( S'_0 = 0 \) and

\[
\sigma_M = \arg \sup_{0 \leq k < \infty} S'_k
\]

as the corresponding maximum point.

Conditional on \( N > \nu \), depending on whether \( \nu > \nu \) or \( \nu < \nu \), we can write the bias as

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

where \( I_A \) denotes the indicator function for the event \( A \). The notations are consistent with Chapter 1 and, in addition, we shall denote by \( E_{\theta_0}[^\cdot] \) the expectation when both \( P_{\theta_0} \) and \( P_{\theta} \) are involved.
Pollak and Siegmund (1986) shows that the quasistationary distribution of $T_{\nu}$ converges to the stationary distribution of $T_{\nu}$ as $d \to \infty$. That means,

$$\lim_{d \to \infty} \lim_{\nu \to \infty} P_{\theta_0}(T_{\nu} < y|\nu) = \lim_{d \to \infty} \lim_{\nu \to \infty} P_{\theta_0}(T_{\nu} < y) = P_{\theta_0}(M < y),$$

as we note that $T_{\nu}$ is asymptotically equivalent to $M$ in distribution as $\nu \to \infty$.

This implies that when the change occurs, $T_{\nu}$ 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 goes below to zero with initial starting point $M$. Given $\hat{\nu} > \nu$, the bias $\hat{\nu} - \nu$ is asymptotically equal to $\tau_{-M}$ plus the length, say $\gamma_m$ for a CUSUM process $T_n$ starting from zero until the last zero point time under $P_{\theta}(\cdot)$. Define

$$E[X; A] = E[XIA].$$

As $d, \nu \to \infty$, we have

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

We can see that $\gamma_m$ is a geometric summation of i.i.d. 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 $\nu$, we see that $T_{\nu+k}$ behaves like a random walk $\{S'_k\}$ for $k \geq 0$ with maximum value $M$ and thus, $\hat{\nu} - \nu$ is asymptotically distributed as the maximum point $\sigma_M$. Thus, as $d, \nu \to \infty$, we have

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

A similar argument is used in Srivastava and Wu (1999) for the continuous-time analog.

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

**Theorem 2.1 :** As $\nu, d \to \infty$,

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

$$-P_{\theta_0}(\sigma_M; \tau_{-M} = \infty],$$

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

$$+E_{\theta_0}(\sigma_M; \tau_{-M} = \infty].$$
2.2 Second-Order Approximation

In this section, we shall derive the second-order approximation for the asymptotic bias given in Theorem 2.1 in order to investigate the bias numerically and also see some local properties by further assuming $\theta_0$ and $\theta$ approach zero at the same order. The main theoretical tool is the strong renewal theorem given in Section 1.3 and its applications to ladder variables.

There are five terms in Theorem 2.1 which will be evaluated in a sequence of lemmas. Most of the results generalizes the ones given in 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 due to sequential sampling plan.

The first lemma gives the approximation for the expected length $\tau_-$ if the random walk goes below zero.

**Lemma 2.2:** As $\theta \to 0$,

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

where $\beta_1$ is given in Lemma 1.1.

**Proof:** By using Wald’s likelihood ratio identity by changing the measure $P_\theta$ to $\tilde{P}_\theta$ and Lemma 1.1, we have

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

After some algebraic simplifications, we get the result.

**Corollary 2.1:** As $\theta \to 0$,

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

To evaluate $P_{\theta,0}(\tau_- M < \infty)$, we follow a similar technique used in Wu (1999) and only the main steps are provided.

First, by conditioning on whether $M = 0$ or $M > 0$, we have

$$P_{\theta,0}(\tau_- M < \infty) = P_\theta(\tau_- < \infty) P_{\theta,0}(\tau_+ = \infty) + P_{\theta,0}(\tau_- M < \infty; M > 0). \quad (2.1)$$

From Lemma 1.2, we have

$$P_\theta(\tau_- < \infty) P_{\theta,0}(\tau_+ = \infty) = \Delta_0 E_0 S_{\tau_+} e^{\theta_0 \rho_\theta} (1 + \Delta E_0 S_{\tau_+}) + o(\theta^2)$$

$$= \Delta_0 E_0 S_{\tau_+} e^{\theta_0 \rho_\theta} - \frac{\Delta_0}{2} + o(\theta^2).$$
For the second term in (2.1), by using the Wald’s likelihood ratio identity by changing parameters \( \theta \) to \( \tilde{\theta} \) and \( \theta_0 \) to \( \theta_1 \), we have

\[
P_{\theta_0}(\tau_x < \infty) = E_{\tilde{\theta}} e^{\Delta S_{\tau_x}} = e^{-\Delta x} E_{\tilde{\theta}} e^{\Delta R_{\tau_x}},
\]

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 1.1, we know

\[
E_{\theta_0} R_x - \rho_+ = O(e^{-\tau_x}) \quad \text{and} \quad E_{\theta_0} R_{-x} - \rho_- = O(e^{-\tau_x}),
\]

as \( x \to \infty \). Now, we write

\[
P_{\theta_0,\theta}(\tau_- M < \infty, M > 0)
\]

\[
= - \int_0^\infty P_{\theta}(\tau_x < \infty) dP_{\theta_0}(M > x)
\]

\[
= - \int_0^\infty E_{\tilde{\theta}} e^{\Delta S_{\tau_x}} dE_{\tilde{\theta}} e^{-\Delta_0 S_{\tau_x}}
\]

\[
= - \int_0^\infty e^{-\Delta(x-\rho_-)} dE_{\tilde{\theta}} e^{-\Delta_0(x+\rho_+)}
\]

\[
- \int_0^\infty e^{-\Delta(x-\rho_-)} \left( e^{-\Delta_0(x+\rho_+)} \left( E_{\theta_1} e^{-\Delta_0(R_x - \rho_+)} - 1 \right) \right) d\left( e^{-\Delta_0(x+\rho_+)} \left( E_{\theta_1} e^{-\Delta_0(R_x - \rho_+)} - 1 \right) \right)
\]

\[
- \int_0^\infty e^{-\Delta(x-\rho_-)} \left( E_{\theta_0} e^{\Delta(R_x - \rho_-)} - 1 \right) d\left( e^{-\Delta_0(x+\rho_+)} \left( E_{\theta_1} e^{-\Delta_0(R_x - \rho_+)} - 1 \right) \right)
\]

\[
\times d \left( e^{-\Delta_0(x+\rho_+)} \left( E_{\theta_1} e^{-\Delta_0(R_x - \rho_+)} - 1 \right) \right). \tag{2.2}
\]

The first term in (2.2) is

\[
\frac{\Delta_0}{\Delta + \Delta_0} e^{\Delta_0 R_x - \rho_+}.
\]

The third term in (2.2) is approximately equal to

\[
\Delta \Delta_0 \int_0^\infty (E_{\theta_0} R_{-x} - \rho_-) dx + o(\theta^2).
\]

The fourth term in (2.2) is

\[
\Delta \Delta_0 \int_0^\infty E_{\theta_0}(R_{-x} - \rho_-) dE_{\theta_0}(R_x - \rho_+) + o(\theta^2).
\]
The second term in (2.2), by integrating by parts, can be approximated as
\[ e^{\Delta \rho_-} (P_{\theta_0}(\tau_+ < \infty) - e^{-\Delta_0 \rho_+}) + \Delta_0 \int_0^\infty (E_0 R_x - \rho_+) dx + o(\theta^2). \]

Combining the above approximations, we have

**Lemma 2.3:** As \( \theta_0, \theta \to 0 \) at the same order,
\[
P_{\theta_0}(\tau - M < \infty) = \frac{\Delta_0}{\Delta + \Delta_0} e^{\Delta \rho_- - \Delta_0 \rho_+} + e^{\Delta \rho_-} \left( 1 - e^{-\Delta_0 \rho_+} \right)
\]
\[+ \Delta_0 \left( \frac{1}{2} - \rho_- E_0 S_{\tau_-} + \int_0^\infty (E_0 R_x - \rho_-) dx \right)
\]
\[+ \int_0^\infty (E_0 R_x - \rho_-) d(E_0 R_x - \rho_+) + \int_0^\infty (E_0 R_x - \rho_+) dx
\]
\[+ o(1). \]

The evaluation of \( E_{\theta_0, \theta}[\tau - M; \tau - M < \infty] \) is similar.

**Lemma 2.4:** As \( \theta_0, \theta \to 0 \) at the same order,
\[
E_{\theta_0, \theta}[\tau - M; \tau - M < \infty] = -\frac{\Delta_0}{\mu} \left( \frac{1}{\Delta + \Delta_0} - \frac{\rho_-}{\Delta + \Delta_0} \right)
\]
\[- \frac{\Delta_0}{\mu} \left( \frac{1}{2} - \rho_- (E_0 S_{\tau_+} - \rho_+) - \int_0^\infty (E_0 R_x - \rho_-) dx \right)
\]
\[- \int_0^\infty (E_0 R_x - \rho_-) d(E_0 R_x - \rho_+)
\]
\[- \int_0^\infty (E_0 R_x - \rho_+) dx + o(1). \]

**Proof:** Again depending on whether \( \{M = 0\} \) or \( \{M > 0\} \), we have
\[
E_{\theta_0, \theta}[\tau - M; \tau - M < \infty]
\]
\[= E_{\theta}[\tau_+; \tau_- < \infty] P_{\theta_0}(\tau_+ = \infty) - \int_0^\infty E_{\theta}[\tau_+; \tau_- < \infty] dP_{\theta_0}(\tau_x < \infty). \quad (2.3) \]

The first term in (2.3) can be approximated by using Lemmas 2.2 and 1.2 as
\[
E_{\theta}[\tau_+; \tau_- < \infty] P_{\theta_0}(\tau_+ = \infty) = \frac{\mu_0}{\mu} + o(1)
\]
\[= -\frac{\Delta}{2\mu} + o(1). \]

For the second term in (2.3), by conditioning on the value of \( M \) given \( M > 0 \) we use the similar techniques as in Lemma 2.3 and write it as
2. Change-Point Estimation

Combining the above results, we complete the proof.

\[ -\int_0^\infty E_0[\tau-x; \tau-x < \infty]dP_0(\tau_x < \infty) \]
\[ = -\int_0^\infty E_0[\tau-xe^{-\Delta(x-R_x)}]dE_0 e^{-\Delta_0(x+R_x)} \]
\[ = -\int_0^\infty E_0(\tau-x)e^{-\Delta(x-R_x)}dE_0 e^{-\Delta_0(x+R_x)} + o(1) \]
\[ = -\frac{1}{\mu}\left[\int_0^\infty (-x + E_0R_x)e^{-\Delta(x-R_x)}de^{-\Delta_0(x+R_x)} \right. \]
\[ \left. - \Delta_0 \int_0^\infty (-x + E_0R_x)d(E_0R_x - \rho_+) \right] + o(1). \]  \hspace{1cm} (2.4)

The first term in (2.4) is approximately equal to

\[ -\frac{\Delta_0}{\mu} \left[ e^{\Delta\rho - \Delta_0\rho_+} \left( \frac{1}{(\Delta + \Delta_0)^2} - \frac{\rho_+}{\Delta + \Delta_0} \right) - \int_0^\infty (E_0R_x - \rho_+)dx \right] + o(1). \]

The second term in (2.4) is equal to

\[ -\frac{\Delta_0}{\mu} \left[ \int_0^\infty (x - \rho_-)d(E_0R_x - \rho_+) - \int_0^\infty (E_0R_x - \rho_-)d(E_0R - x + \rho_+) \right] \]
\[ = -\frac{\Delta_0}{\mu} [\rho_-(E_0S_{\tau_+} - \rho_+) - \int_0^\infty (E_0R - x + \rho_+)dx] \]
\[ - \int_0^\infty (E_0R_x - \rho_-)d(E_0R_x - \rho_+). \]

Combining the above results, we complete the proof.

Finally, we evaluate \( E_{\theta_0}[\sigma_M; \tau-M < \infty] \). We first write

\[ E_{\theta_0}[\sigma_M; \tau-M < \infty] = E_{\theta_0}\sigma_M - E_{\theta_0} \left[ \sigma_M E_{\theta}e^{\Delta(-M+R-M)} \right]. \]  \hspace{1cm} (2.5)

For the second term in (2.5), we write

\[ E_{\theta_0}\left[ \sigma_M E_{\theta}e^{\Delta(-M+R-M)} \right] = E_{\theta_0} \left[ \sigma_M e^{-\Delta M} \right] e^{-\Delta\rho_-} + E_{\theta_0} \left[ \sigma_M e^{-\Delta M} (E_{\theta}e^{\Delta R-M} - e^{\Delta\rho_-}) \right]. \]

To evaluate \( E_{\theta_0}[\sigma_M e^{-\Delta M}] \), we note that under \( P_0(\cdot) \)

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

where \( =^d \) means equivalence in distribution, \( \tau_{\tau_+}^{(k)} \) is the kth ladder epoch defined in Section 1.3, and

\[ K = \inf\{k > 0 : \tau_{\tau_+}^{(k)} < \infty\}. \]

Note that \( K \) is a geometric random variable with

\[ P(K = k) = p^k(1 - p) \]
2.2. Second-Order Approximation

for \( k \geq 0 \), with terminating probability

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

For given \( K = k \), \((\tau_+^{(k)}, S_{\tau_+^{(k)}})\) is, in distribution, equivalent to the sum of \( k \) i.i.d. random variables distributed as \((\tau_+, S_{\tau_+})\).

Thus,

\[
E_{\theta_0} \left[ \sigma_M e^{-\Delta M} \right] = E_{\theta_0} \left[ \tau_+^{(K)} \exp(-\Delta S_{\tau_+^{(K)}}) \right]
\]

\[
= \sum_{k=1}^{\infty} E_{\theta_0} \left[ \tau_+^{(k)} \exp(-\Delta S_{\tau_+^{(k)}}); K = k \right]
\]

\[
= \sum_{k=1}^{\infty} k E_{\theta_0} \left[ \tau_+ \exp(-\Delta S_{\tau_+}); \tau_+ < \infty \right]
\]

\[
\times \left( E_{\theta_0} \left[ \exp(-\Delta S_{\tau_+}); \tau_+ < \infty \right] \right)^{k-1} P_{\theta_0}(\tau_+ = \infty)
\]

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

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

Lemma 2.5: As \( \theta_0, \theta \to \) at the same order,

\[
1 - E_{\theta_0} \left[ \exp(-\Delta S_{\tau_+}); \tau_+ < \infty \right]
\]

\[
= (\Delta + \Delta_0) E_0 S_{\tau_+} \exp \left( -(\Delta - \theta_0) \rho_+ + \frac{1}{2} (\Delta - \theta_0)^2 (\rho_+^{(2)} - \rho_+^2) - \frac{\theta_0^2}{2 E_0 S_{\tau_+}} \alpha \right)
\]

\[
\times (1 + o(\theta^2)).
\]

Proof: Using Wald’s likelihood ratio identity, we have

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

The Taylor series expansion following the lines of Lemma 1.2 will give the result after some algebraic simplification.

In particular, by letting \( \Delta = 0 \), we have

\[
P_{\theta_0}(\tau_+ = \infty) = E_0 S_{\tau_+} \exp \left( \theta_0 \rho_+ + \frac{1}{2} \theta_0^2 \left( \rho_+^{(2)} - \rho_+^2 - \frac{\alpha}{E_0 S_{\tau_+}} \right) \right) (1 + o(\theta_0^2)).
\]

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,

$$E_{\theta_0} \left[ \tau_+ e^{-\Delta S_{\tau_+}}; \tau_+ < \infty \right] = \frac{E_0 S_{\tau_+}}{\mu_1} \exp(-((\Delta - \theta_0)\rho_+ + \frac{1}{2}(\Delta - \theta_0)^2(\rho_+^{(2)} - \rho_+^2))$$

$$- \frac{\theta_0^2}{2} \frac{\alpha_1}{E_0 S_{\tau_+}} - \theta_1(\Delta + \Delta_0) \frac{\alpha_1}{E_0 S_{\tau_+}} \left( 1 + o(\theta^2) \right).$$

In particular, by letting $\Delta = 0$, we have

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

On the other hand, by conditioning on the value of $M$, we have

$$E_{\theta_0} \left[ \sigma_M e^{-\Delta M} \left( E_{\theta} e^{\Delta R_{-M}} - e^{\Delta \rho_-} \right) \right]$$

$$= \Delta E_{\theta_0} \left[ \sigma_M (E_0 R_{-M} - \rho_-) \right](1 + o(1))$$

$$= -\Delta \int_0^\infty E_{\theta_0} [\sigma_x | M = x] (E_0 R_{-x} - \rho_-) dP_{\theta_0}(M > x)$$

$$= \Delta \Delta_0 \int_0^\infty E_{\theta_0} [\sigma_x | M = x] (E_0 R_{-x} - \rho_-) d(x + E_0 R_{x})(1 + o(1)).$$

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 $O_p(.)$ means at the same order in probability. This implies

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

Thus, (2.6) is at the 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 $\theta_0 \to 0$,

$$E_{\theta_0} [\sigma_M] = \frac{E_{\theta_0} [\tau_+; \tau_+ < \infty]}{P_{\theta_0} (\tau_+ = \infty)}$$

$$= \frac{1}{\Delta_0 \mu_1} e^{-\theta_0^2 (2\alpha_1 / E_0 S_{\tau_+})} \left( 1 + o(\theta_0^2) \right).$$

Finally, we have the following result:

Lemma 2.8: As $\theta_0, \theta \to 0$,

$$E_{\theta_0 \theta} [\sigma_M; \tau_- = \infty] = \frac{1}{\Delta_0 \mu_1} e^{-\theta_0^2 (2\alpha_1 / E_0 S_{\tau_+})}$$
2.2. Second-Order Approximation

\[
- \frac{\Delta_0}{\mu_1(\Delta + \Delta_0)^2} e^{\gamma \Delta/3 - 2\theta(\theta - \theta_0)(\rho_2^2 - \rho_z^2 - \alpha_1/E_0 S_{\tau_+}) - 2(\theta - \theta_0)^2(\alpha_1/E_0 S_{\tau_+})} (1 + o(\theta^2)).
\]

Combining Lemmas 2.1-2.8 and Lemmas 1.1-1.2, we have the following local 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

\[
\lim_{d \to \infty} \lim_{\nu \to \infty} E[\hat{\nu} - \nu] = -\frac{1}{\hat{\mu}} \left( \frac{\theta}{\Delta + \Delta_0} e^{\Delta \rho - \Delta_0 \rho_+} + e^{\Delta \rho} \left( 1 - e^{-\Delta_0 \rho_+} \right) \right)
\]

\[
- \frac{\Delta}{\hat{\mu}} \left( \frac{1}{\Delta + \Delta_0} - \frac{\rho_-}{\Delta + \Delta_0} \right) e^{\Delta \rho - \Delta_0 \rho_+}
\]

\[
+ \frac{\theta_0}{\theta_0 - \theta_0} \frac{\beta_1}{E_0 S_{\tau_+}} + 2\theta_0 \rho - \rho_-
\]

\[
- \frac{\theta}{\theta_0} \left( \rho_2 - \rho_z - \frac{\alpha_1}{E_0 S_{\tau_+}} \right) + o(1).
\]

A similar result can be obtained for the absolute bias \( E[|\hat{\nu} - \nu|] \).

In particular, when \( \theta = \theta_1 \), we have the following result.

**Corollary 2.2:** As \( \theta = \theta_1 \to 0 \),

\[
- \frac{3}{4\Delta_0} \left( \frac{1}{\mu_0} + \frac{1}{\mu_1} \right) - \frac{\gamma}{12} \left( \frac{1}{\mu_0} + \frac{1}{\mu} \right) - \frac{1}{2} \frac{\beta_1}{E_0 S_{\tau_+}}
\]

\[
- \frac{1}{2} \left( \rho_2 - \rho_z - \frac{\alpha_1}{E_0 S_{\tau_+}} \right) + o(1)
\]

\[
= - \frac{\gamma}{4\theta_0} + \frac{17\gamma^2}{288} - \frac{\kappa}{16} \frac{\beta_1}{2 E_0 S_{\tau_+}} - \frac{1}{2} \left( \rho_2 - \rho_z - \frac{\alpha_1}{E_0 S_{\tau_+}} \right) + o(1).
\]

**Proof:** The first equation is a direct simplification of Theorem 2.2. For the second equation, we note that as \( \theta_0 \to 0 \),

\[
\mu_1 = \theta_1 e^{\theta_1(\gamma/2) + \theta_1^2(\gamma^2/8)} (1 + o(\theta_1^2)),
\]

\[
\Delta_0 = 2\theta_1 e^{\theta_1(\gamma/6) + \theta_1^2(\gamma^2/24)} (1 + o(\theta_1^2)),
\]

\[
\theta_0 = -\theta_1 e^{\theta_1(\gamma/3) + \theta_1^2(\gamma^2/18)} (1 + o(\theta_1^2)),
\]

\[
\mu_0 = \theta_1 e^{-\theta_1(\gamma/6) + \theta_1^2(\gamma^2/12)} (1 + o(\theta_1^2)).
\]

Some tedious simplifications give the expected results.
Therefore, the local bias of $\hat{\nu}$ is largely affected by the skewness $\gamma$. If $\gamma > 0$, the local bias becomes positive. If $F_0(x)$ is symmetric, from Corollary 1.2, we have

$$\rho_+^{(2)} - \rho_-^{(2)} = \frac{\alpha_1}{E_0 S_{r_+}} = \frac{\kappa}{6},$$

and thus,

$$E_\nu[\hat{\nu} - \nu | N > \nu] \approx -\frac{7}{48} \frac{\kappa}{\beta_1} + o(1),$$

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

### 2.3 Two Examples

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

#### 2.3.1 Normal Distribution

From Example 1.1, the approximations for the related quantities are simplified as

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

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

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

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

Summarizing the above results, we have the following corollary:

**Corollary 2.3:** As $\theta_0, \theta \to 0$ at the same order,

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

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

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

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

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

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

and at \( \theta = -\theta_0, \)

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

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

\[ \lim_{d \to \infty} \lim_{\nu \to \infty} \mathbb{E}_\nu[\hat{\nu} - \nu|N > \nu] = \frac{1}{2\theta^2} - \frac{1}{2\theta_0^2}, \]

which is zero at \( \theta = -\theta_0 \) and at \( \theta = -\theta_0, \)

\[ \lim_{d \to \infty} \lim_{\nu \to \infty} \mathbb{E}_\nu[|\hat{\nu} - \nu|N > \nu] = \frac{3}{4\theta_0^2}. \]

We see that the sequential sampling plan has local effect at the second-order in the discrete-time case 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. One thousand replications of the CUSUM charts are simulated for several values for \( \theta. \) Only those runs with \( N > \nu \) are used for calculating \( \hat{\nu}. \) Table 2.1 gives the comparison between the simulated results and approximated values. The approximated values from Corollary 2.4 are given in the parentheses. 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 \( \theta_0 = 0.25. \) 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 \( \nu \) is very little. However, as the local bias is at the order \( O(1/\theta_0^2) \) at \( \theta = -\theta_0, \) which approaches \( \infty \) 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 \( \theta d \) approaches a constant, called moderate deviation as considered in Chang (1992), is definitely worth a future study.

### 2.3.2 Exponential Distribution

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

From Example 1.2 and Theorem 2.2, we have the following result:

**Corollary 2.4.** As \( \theta, \theta_0 \to 0 \) at the same order,
Table 2.1: Biases in the Normal Case

\[
\begin{array}{cccc}
\nu & \theta_0 & \theta & E[\hat{\nu} - \nu|N > \nu] \quad E[|\hat{\nu} - \nu||N > \nu] \\
50 & -0.25 & 0.25 & 0.113(-0.125) & 9.737(11.875) \\
& & 0.5 & -4.902(-6.083) & 7.090(8.139) \\
& & 0.75 & -5.682(-7.174) & 6.376(7.826) \\
& & 1.0 & -5.768(-7.550) & 6.188(7.810) \\
-0.5 & 0.5 & 0.268(-0.125) & 3.02(2.875) \\
& & 0.75 & -1.302(-1.211) & 2.338(2.249) \\
& & 1.0 & -1.730(-1.583) & 3.135(1.972) \\
100 & -0.25 & 0.25 & 1.368(-0.125) & 11.728(11.875) \\
& & 0.5 & -5.644(-6.083) & 7.768(8.139) \\
& & 0.75 & -6.181(-7.174) & 6.942(7.826) \\
& & 1.0 & -6.250(-7.55) & 6.520(7.810) \\
-0.5 & 0.5 & -0.223(-0.125) & 3.052(2.875) \\
& & 0.75 & -1.109(-1.211) & 2.208(2.149) \\
& & 1.0 & -1.564(-1.583) & 2.084(1.972) \\
\end{array}
\]

\[
\lim_{d \to \infty} \lim_{\nu \to \infty} E'\hat{\nu} [\hat{\nu} - \nu|N > \nu]
\]

\[
= - \frac{\Delta + 2\Delta_0}{\hat{\mu}(\Delta + \Delta_0)^2} e^{-(1/2)\Delta - \Delta_0} \\
- \frac{1}{\hat{\mu}\Delta} e^{-\Delta/2}(1 - e^{-\Delta_0}) - \frac{\Delta_0}{2\hat{\mu}(\Delta + \Delta_0)} e^{-(1/2)\Delta - \Delta_0}
\]

\[
- \frac{1}{\Delta_0 \mu_1} + \frac{\Delta_0}{\mu_1(\Delta + \Delta_0)^2} e^{(2/3)\Delta} \\
- \frac{\theta}{\theta - \theta_0} - \frac{\theta_0}{\theta} + \frac{7}{18} \theta - \theta_0 + o(1).
\]

At \( \theta = -\theta_0 \),

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

We see that due to the asymmetry of \( F_0(x) \), the local bias becomes positive as \( \theta_0 \) is small.

### 2.4 Case Study

In this section, we conduct three classical case studies to illustrate the applications.
Table 2.2: Nile River Flow from 1871 – 1970

<table>
<thead>
<tr>
<th>Year</th>
<th>Flow</th>
<th>Year</th>
<th>Flow</th>
<th>Year</th>
<th>Flow</th>
<th>Year</th>
<th>Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>1871</td>
<td>1120</td>
<td>1896</td>
<td>1220</td>
<td>1921</td>
<td>768</td>
<td>1946</td>
<td>1040</td>
</tr>
<tr>
<td>1872</td>
<td>1160</td>
<td>1897</td>
<td>1030</td>
<td>1922</td>
<td>845</td>
<td>1947</td>
<td>860</td>
</tr>
<tr>
<td>1873</td>
<td>963</td>
<td>1898</td>
<td>1100</td>
<td>1923</td>
<td>864</td>
<td>1948</td>
<td>874</td>
</tr>
<tr>
<td>1874</td>
<td>1210</td>
<td>1899</td>
<td>774</td>
<td>1924</td>
<td>862</td>
<td>1949</td>
<td>848</td>
</tr>
<tr>
<td>1875</td>
<td>1160</td>
<td>1900</td>
<td>840</td>
<td>1925</td>
<td>698</td>
<td>1950</td>
<td>890</td>
</tr>
<tr>
<td>1876</td>
<td>1160</td>
<td>1901</td>
<td>874</td>
<td>1926</td>
<td>845</td>
<td>1951</td>
<td>744</td>
</tr>
<tr>
<td>1877</td>
<td>813</td>
<td>1902</td>
<td>694</td>
<td>1927</td>
<td>744</td>
<td>1952</td>
<td>749</td>
</tr>
<tr>
<td>1878</td>
<td>1230</td>
<td>1903</td>
<td>940</td>
<td>1928</td>
<td>796</td>
<td>1953</td>
<td>838</td>
</tr>
<tr>
<td>1879</td>
<td>1370</td>
<td>1904</td>
<td>833</td>
<td>1929</td>
<td>1040</td>
<td>1954</td>
<td>1050</td>
</tr>
<tr>
<td>1880</td>
<td>1140</td>
<td>1905</td>
<td>701</td>
<td>1930</td>
<td>759</td>
<td>1955</td>
<td>918</td>
</tr>
<tr>
<td>1881</td>
<td>995</td>
<td>1906</td>
<td>916</td>
<td>1931</td>
<td>781</td>
<td>1956</td>
<td>986</td>
</tr>
<tr>
<td>1882</td>
<td>935</td>
<td>1907</td>
<td>692</td>
<td>1932</td>
<td>865</td>
<td>1957</td>
<td>797</td>
</tr>
<tr>
<td>1883</td>
<td>1110</td>
<td>1908</td>
<td>1020</td>
<td>1933</td>
<td>845</td>
<td>1958</td>
<td>923</td>
</tr>
<tr>
<td>1884</td>
<td>994</td>
<td>1909</td>
<td>1050</td>
<td>1934</td>
<td>944</td>
<td>1959</td>
<td>975</td>
</tr>
<tr>
<td>1885</td>
<td>1020</td>
<td>1910</td>
<td>969</td>
<td>1935</td>
<td>984</td>
<td>1960</td>
<td>815</td>
</tr>
<tr>
<td>1886</td>
<td>960</td>
<td>1911</td>
<td>831</td>
<td>1936</td>
<td>897</td>
<td>1961</td>
<td>1020</td>
</tr>
<tr>
<td>1887</td>
<td>1180</td>
<td>1912</td>
<td>726</td>
<td>1937</td>
<td>822</td>
<td>1962</td>
<td>906</td>
</tr>
<tr>
<td>1888</td>
<td>799</td>
<td>1913</td>
<td>456</td>
<td>1938</td>
<td>1010</td>
<td>1963</td>
<td>901</td>
</tr>
<tr>
<td>1889</td>
<td>958</td>
<td>1914</td>
<td>824</td>
<td>1939</td>
<td>771</td>
<td>1964</td>
<td>1170</td>
</tr>
<tr>
<td>1890</td>
<td>1140</td>
<td>1915</td>
<td>702</td>
<td>1940</td>
<td>676</td>
<td>1965</td>
<td>912</td>
</tr>
<tr>
<td>1891</td>
<td>1100</td>
<td>1916</td>
<td>1120</td>
<td>1941</td>
<td>649</td>
<td>1966</td>
<td>746</td>
</tr>
<tr>
<td>1892</td>
<td>1210</td>
<td>1917</td>
<td>1100</td>
<td>1942</td>
<td>846</td>
<td>1967</td>
<td>919</td>
</tr>
<tr>
<td>1893</td>
<td>1150</td>
<td>1918</td>
<td>832</td>
<td>1943</td>
<td>812</td>
<td>1968</td>
<td>718</td>
</tr>
<tr>
<td>1894</td>
<td>1250</td>
<td>1919</td>
<td>764</td>
<td>1944</td>
<td>742</td>
<td>1969</td>
<td>714</td>
</tr>
<tr>
<td>1895</td>
<td>1260</td>
<td>1920</td>
<td>821</td>
<td>1945</td>
<td>801</td>
<td>1970</td>
<td>740</td>
</tr>
</tbody>
</table>

2.4.1 Nile River Data (Normal Case):

The following Nile River flow data are reproduced from Cobb(1978), and the data are read in columns.

From a scatterplot, we see there is an obvious change around the year 1900. A q-q-normal plot shows the normality is roughly true.

As in Cobb(1978), we assume the independent normality. Assume the pre-change mean is $m_0 = 1100$ and the post-change mean is $m_1 = 850$, with a change magnitude $m_1 - m_0 = 250$ and standard deviation 125.

To apply the CUSUM procedure, we first standardize the data by subtracting all the observations by $m_0 + (m_1 - m_0)/2 = 975$, the average of the pre-change and post-change means, then switch the sign in order to make the change positive, and then divide all the data by 125.

After these transformations, we standardize the observations to $x_i$ for $i =$
1, 2, ..., 100 such that
\[\theta_0 = -1, \quad \theta_1 = 1, \quad \text{and} \quad \sigma^2 = 1.\]

Now we form the CUSUM process by letting \(T_0 = 0\) and
\[T_i = \max(0, T_{i-1} + x_i),\]
for \(i = 1, ..., 100\). The calculated values are reported as follows:

<p>| | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.32</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.12</td>
<td>0.00</td>
<td>1.41</td>
<td>1.54</td>
<td>0.22</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.61</td>
</tr>
<tr>
<td>3.50</td>
<td>5.74</td>
<td>6.02</td>
<td>7.16</td>
<td>9.35</td>
<td>9.82</td>
<td>12.09</td>
<td>11.73</td>
<td>11.13</td>
<td>11.18</td>
<td>12.33</td>
<td>14.32</td>
</tr>
<tr>
<td>19.68</td>
<td>21.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.70</td>
<td>19.70</td>
<td>20.85</td>
<td>22.54</td>
<td>23.77</td>
<td>25.42</td>
<td>26.46</td>
<td>27.35</td>
<td>28.26</td>
<td>30.47</td>
<td>31.51</td>
<td>33.36</td>
</tr>
<tr>
<td>34.79</td>
<td>34.27</td>
<td>36.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.55</td>
<td>38.43</td>
<td>39.47</td>
<td>39.72</td>
<td>39.65</td>
<td>40.27</td>
<td>41.50</td>
<td>41.22</td>
<td>42.85</td>
<td>45.24</td>
<td>47.85</td>
<td>48.88</td>
</tr>
<tr>
<td>50.18</td>
<td>52.05</td>
<td>53.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52.92</td>
<td>53.84</td>
<td>54.65</td>
<td>55.66</td>
<td>56.34</td>
<td>58.19</td>
<td>60.00</td>
<td>61.10</td>
<td>60.50</td>
<td>60.95</td>
<td>60.86</td>
<td>62.29</td>
</tr>
<tr>
<td>62.70</td>
<td>62.70</td>
<td>63.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63.62</td>
<td>64.18</td>
<td>64.77</td>
<td>63.21</td>
<td>63.71</td>
<td>65.24</td>
<td>65.99</td>
<td>68.05</td>
<td>70.14</td>
<td>72.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The reason we report the CUSUM process instead of drawing graphs is to inspect the path directly rather than guessing from the graph. By looking at the CUSUM process, we see that in the fixed sample size case, the maximum likelihood ratio estimator (the last zero point of the CUSUM process) is \(\hat{\nu}_n = 28\) which is the year 1898.

The estimated pre-change mean is the average of the observations from 1970 to 1998, which is \(-0.98\); the estimated post-change mean is 1.00; and the pre-change standard deviation is 1.08 and post-change standard deviation is 1.00. We see that the assumption is roughly correct except for the slight discrepancy from the pre-change standard deviation.

Let us look at the estimator from the sequential sampling plan point of view, which is more natural from the nature of monitoring.

We see that as long as the threshold \(d\) is taken larger than 2.00 and less than 70, a change is always signaled and the maximum likelihood estimator \(\hat{\nu} = 28\) no matter what the value \(d\) is.

Thus, we see that the estimator \(\hat{\nu}\) is stable to the selection of the threshold \(d\).

### 2.4.2 British Coal Mining Disaster (Exponential Case)

The following table gives the intervals in days between successive coal mining disasters in Great Britain for the period 1875–1951. A disaster is defined as involving the death of 10 or more men. The data are taken from Maguire, Pearson, and Wynn(1952) and appeared in many places; the most noticeable is Cox and Lewis (1966). The data are read in columns.
Table 2.3: British Coal Mining Disaster Intervals

<table>
<thead>
<tr>
<th></th>
<th>378</th>
<th>286</th>
<th>871</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>114</td>
<td>48</td>
<td>291</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>108</td>
<td>123</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>188</td>
<td>457</td>
<td>369</td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>233</td>
<td>498</td>
<td>338</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>28</td>
<td>49</td>
<td>336</td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>22</td>
<td>131</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>61</td>
<td>182</td>
<td>329</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>78</td>
<td>255</td>
<td>330</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>99</td>
<td>195</td>
<td>312</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>326</td>
<td>224</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>275</td>
<td>566</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>54</td>
<td>390</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>217</td>
<td>72</td>
<td>364</td>
<td></td>
</tr>
<tr>
<td>203</td>
<td>113</td>
<td>228</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>176</td>
<td>32</td>
<td>271</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>23</td>
<td>208</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>151</td>
<td>517</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>361</td>
<td>1613</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>315</td>
<td>312</td>
<td>54</td>
<td>1630</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>354</td>
<td>326</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>58</td>
<td>1312</td>
<td>217</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>275</td>
<td>348</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>78</td>
<td>745</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>189</td>
<td>17</td>
<td>217</td>
<td>1357</td>
<td></td>
</tr>
<tr>
<td>345</td>
<td>1205</td>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>644</td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>467</td>
<td>20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We first explain the transformation on the data in the exponential case in order to fit them into the frame.

Suppose the original data \( \{Y_i\} \) follow \( \exp(\lambda_0) \) for \( i \leq \nu \) and \( \exp(\lambda_1) \) for \( i > \nu \) where \( \lambda_0 > \lambda_1 \) are the corresponding hazard rates.

Define

\[
\lambda^* = (\lambda_0 - \lambda_1) / \ln(\lambda_0 / \lambda_1),
\]

and make the following data transformation:

\[
X_i = \lambda^* Y_i - 1,
\]

for \( i = 1, 2, \ldots \).

Denote

\[
f_0(x) = e^{-(x+1)},
\]

for \( x \leq -1 \), and then

\[
f_\theta(x) = \exp(\theta x - c(\theta))f_0(x)
\]

satisfies the standardized model with

\[
c(\theta) = - (\theta + \ln(1 - \theta)),
\]

and

\[
\theta_i = 1 - \lambda_i / \lambda^*
\]

for \( i = 0, 1 \) such that

\[
c(\theta_0) = c(\theta_1).
\]

A scatterplot or by looking at the data directly shows that there is a change around the observation 50. So we take the mean from observations 1 to 50 as the pre-change mean and the mean from observations 51 to 109 as the post-change mean, which gives

\[
\lambda_0 = 1/129 \quad \text{and} \quad \lambda_1 = 1/335,
\]

and \( \lambda^* = 0.005 \).

Now, after the data transformation by letting \( x_i = 0.005y_i - 1 \) for \( i = 1, \ldots, 109 \), we can fit the data into the standardized model with

\[
\theta_0 = -0.550 \quad \text{and} \quad \theta_1 = 0.403.
\]

Next, we can formalize the CUSUM process by calculating \( T_n \)'s based on \( x_i \)'s, which are reported as follows:

\[
\begin{align*}
[1] & \quad 0.88812755 \quad 0.06794922 \quad 0.00000000 \quad 0.00000000 \quad 0.07393498 \\
[6] & \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \\
[11] & \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.01399443 \\
[16] & \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.57343963 \\
[21] & \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \quad 0.00000000 \\
[26] & \quad 0.72329102 \quad 0.00000000 \quad 0.00000000 \quad 0.42858328 \quad 0.00000000
\end{align*}
\]
2.4. Case Study

By inspecting the CUSUM process, we see that the estimator $\hat{\nu} = 46$ no matter what the threshold $d$ is (at least 5 and at most 40). Again, we see that the CUSUM procedure is very reliable in terms of the change-point estimator.

2.4.3 IBM Stock Price (Variance Change)

Suppose the original independent observations $\{Y_i\}$ follow $N(0, \sigma_0^2)$ for $i \leq \nu$ and $N(0, \sigma^2)$ for $i > \nu$.

For a reference value $\sigma_1^2$ for $\sigma^2$, we define

$$\lambda^* = \left( \frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) / \ln \left( \frac{\sigma_1^2}{\sigma_0^2} \right)$$

and make the following data transformation:

$$X_i = \frac{1}{\sqrt{2}} (\lambda^* Y_i^2 - 1).$$

Let $f_0(x)$ be the density function of $(\chi^2 - 1)/\sqrt{2}$, where $\chi^2$ is the standard chi-square random variable with degree of freedom 1. Then from Example 1.3, we know

$$c(\theta) = -\frac{\theta}{\sqrt{2}} - \frac{1}{2} \ln(1 - \sqrt{2} \theta).$$

Define

$$\theta_0 = \frac{1}{\sqrt{2}} \left( 1 - \frac{1}{\lambda^*} \right) \quad \text{and} \quad \theta_1 = \frac{1}{\sqrt{2}} \left( 1 - \frac{1}{\lambda^*} \right),$$

and generally

$$\theta = \frac{1}{\sqrt{2}} \left( 1 - \frac{1}{\lambda^*} \right).$$
2. Change-Point Estimation

It can be verified that
\[ c'(\theta) = \frac{1}{\sqrt{2}} \left( \frac{\lambda^*}{1/\sigma^2} - 1 \right) \]
and \( c(\theta_0) = c(\theta_1) \). Thus, the standardized observations \( \{X_i\} \) fit in our model.

**Remark:** Generally, the original independent observations \( \{Y_i\} \) follow \((1 + \epsilon_0)^2 \chi^2_p\) for \( i \leq \nu \) and \((1 + \epsilon_1)^2 \chi^2_p\) for \( i > \nu \), where \( \chi^2_p \) is the standard chi-square random variable with degree of freedom \( p \). Then, by using Example 1.3, we define
\[ \lambda^* = \frac{1}{(1 + \epsilon_0)^2} - \frac{1}{(1 + \epsilon_1)^2} \bigg/ \ln[(1 + \epsilon_1)^2/(1 + \epsilon_0)^2], \]
and make the following transformation:
\[ X_i = \frac{1}{\sqrt{2p}}(\lambda^*Y_i - p). \]

For \( f_0(x) \) defined as in Example 1.3, let
\[ \theta = \sqrt{p/2} \left( 1 - \frac{(1 + \epsilon)^2}{\lambda^*} \right), \]
and define \( \theta_0 \) and \( \theta_1 \) correspondingly. Then we can verify that \( c(\theta_0) = c(\theta_1) \).

The following data set is taken as the IBM stock daily closing prices from May 17 of 1961 to Nov. 2 of 1962, [Box, Jenkins, and Reinsel(1994), pp.542)] for a total of 369 observations. The data are read in rows.

We use the geometric normal random walk model and find that it fits the data quite well with quite small autocorrelation. After taking the difference for the logarithm of the data, the plot of total 368 data shows an obvious increase in the variance roughly around the 225th observation.

The standard deviation for the first 225 observations is found to be 0.00978. So we divide all the data by 0.00978, and denote the modified data as \( \{Y_i\} \)'s, which gives
\[ \sigma_0^2 = 1 \quad \text{and} \quad \sigma_1^2 = 7.239, \]
where \( \sigma_1^2 \) is the variance of the last 143 modified observations.

Now, we calculate \( \lambda^* \) as
\[ \lambda^* = \left( \frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) / \ln \left( \frac{\sigma_1^2}{\sigma_0^2} \right) = \frac{(1 - 1/7.239)}{\log(7.239)} = 0.4354. \]

The transformed data are calculated as
\[ X_i = \frac{1}{\sqrt{2}}(\lambda^*Y_i^2 - 1), \]
and the corresponding conjugate parameters are found to be
\[ \theta_0 = \frac{1}{\sqrt{2}} \left( 1 - \frac{1}{0.4354} \right) = -0.917 \]
Table 2.4: IBM Stock Price

<table>
<thead>
<tr>
<th>Date</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>460</td>
<td>457</td>
</tr>
<tr>
<td>493</td>
<td>490</td>
</tr>
<tr>
<td>489</td>
<td>478</td>
</tr>
<tr>
<td>479</td>
<td>477</td>
</tr>
<tr>
<td>479</td>
<td>477</td>
</tr>
<tr>
<td>474</td>
<td>465</td>
</tr>
<tr>
<td>481</td>
<td>488</td>
</tr>
<tr>
<td>494</td>
<td>499</td>
</tr>
<tr>
<td>504</td>
<td>513</td>
</tr>
<tr>
<td>523</td>
<td>531</td>
</tr>
<tr>
<td>545</td>
<td>549</td>
</tr>
<tr>
<td>532</td>
<td>517</td>
</tr>
<tr>
<td>547</td>
<td>553</td>
</tr>
<tr>
<td>569</td>
<td>575</td>
</tr>
<tr>
<td>599</td>
<td>596</td>
</tr>
<tr>
<td>596</td>
<td>596</td>
</tr>
<tr>
<td>588</td>
<td>582</td>
</tr>
<tr>
<td>584</td>
<td>581</td>
</tr>
<tr>
<td>583</td>
<td>581</td>
</tr>
<tr>
<td>582</td>
<td>584</td>
</tr>
<tr>
<td>556</td>
<td>557</td>
</tr>
<tr>
<td>553</td>
<td>553</td>
</tr>
<tr>
<td>542</td>
<td>548</td>
</tr>
<tr>
<td>537</td>
<td>535</td>
</tr>
<tr>
<td>548</td>
<td>549</td>
</tr>
<tr>
<td>551</td>
<td>551</td>
</tr>
<tr>
<td>533</td>
<td>525</td>
</tr>
<tr>
<td>516</td>
<td>511</td>
</tr>
<tr>
<td>518</td>
<td>513</td>
</tr>
<tr>
<td>486</td>
<td>475</td>
</tr>
<tr>
<td>457</td>
<td>449</td>
</tr>
<tr>
<td>383</td>
<td>393</td>
</tr>
<tr>
<td>359</td>
<td>335</td>
</tr>
<tr>
<td>316</td>
<td>320</td>
</tr>
<tr>
<td>351</td>
<td>350</td>
</tr>
<tr>
<td>382</td>
<td>370</td>
</tr>
<tr>
<td>369</td>
<td>376</td>
</tr>
<tr>
<td>373</td>
<td>382</td>
</tr>
<tr>
<td>389</td>
<td>394</td>
</tr>
<tr>
<td>391</td>
<td>388</td>
</tr>
<tr>
<td>382</td>
<td>383</td>
</tr>
<tr>
<td>377</td>
<td>364</td>
</tr>
<tr>
<td>349</td>
<td>358</td>
</tr>
<tr>
<td>367</td>
<td>357</td>
</tr>
<tr>
<td>339</td>
<td>331</td>
</tr>
</tbody>
</table>
and

$$\theta_1 = \frac{1}{\sqrt{2}} \left( 1 - \frac{1}{7.239 \times 0.4354} \right) = 0.483.$$

The CUSUM process formed by the standardized \( \{X_i\} \)'s are calculated as

\[
\begin{align*}
[1] & 0.00 0.00 0.05 0.00 0.00 0.00 3.01 4.97 4.39 3.73 3.50 2.80 \\
[13] & 2.15 1.45 1.22 0.53 1.49 1.90 1.41 0.92 0.55 0.00 0.00 0.00 \\
[25] & 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[37] & 0.00 0.00 0.00 0.00 0.48 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[49] & 0.16 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[61] & 0.00 0.00 0.00 0.00 0.30 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[73] & 0.00 0.00 0.04 2.16 1.63 1.09 0.78 0.24 0.00 0.00 0.00 0.00 \\
[85] & 0.00 0.00 0.00 0.00 1.93 2.40 3.60 2.94 2.41 1.80 1.10 0.78 \\
[97] & 0.46 0.12 0.00 0.00 0.51 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[109] & 0.00 0.03 0.00 0.00 0.00 0.41 0.00 0.00 0.00 0.00 0.00 0.00 \\
[121] & 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[133] & 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[145] & 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[157] & 0.00 0.51 1.07 0.88 0.18 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[169] & 0.00 0.00 0.00 0.00 0.00 0.00 0.20 0.06 2.62 3.72 3.11 2.91 \\
[181] & 1.51 0.81 0.27 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[193] & 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 \\
[205] & 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.39 0.00 0.00 \\
[217] & 0.00 0.03 1.04 0.45 1.21 0.50 0.00 0.00 0.00 0.00 0.00 0.00 \\
[229] & 0.00 0.00 0.00 0.00 0.00 0.00 1.76 3.66 16.99 17.28 18.34 18.80 \\
[241] & 18.29 19.27 22.34 22.63 21.99 22.06 22.64 22.07 21.76 22.05 21.36 24.35 \\
[253] & 30.78 35.70 35.01 66.55 77.11 78.54 79.19 92.99 92.68 92.00 91.89 91.55 \\
[265] & 96.24 110.94 114.52 123.22 145.53 145.08 145.42 146.58 150.35 150.15 \\
153.80 157.46 \\
[277] & 161.12 164.77 169.17 171.86 171.84 174.42 173.74 173.06 173.02 172.98 \\
187.59 187.24 \\
[289] & 186.74 186.84 189.41 189.30 188.69 188.57 187.89 189.56 190.38 189.77 \\
190.20 192.17 \\
[301] & 191.46 193.43 194.52 193.82 193.66 194.06 195.19 195.04 194.35 193.85 \\
194.22 193.54 \\
[313] & 192.85 192.34 192.15 191.47 195.89 195.26 194.63 193.94 197.75 197.12 \\
196.61 197.24 \\
[325] & 198.24 197.88 197.71 197.78 197.17 196.55 195.86 195.15 194.99 195.31 \\
194.79 194.85 \\
[337] & 194.34 194.43 197.69 197.58 201.69 201.63 201.16 204.98 206.98 206.30 \\
207.68 207.07 \\
[349] & 206.36 206.54 207.30 206.55 205.87 208.72 210.47 210.16 210.36 210.93 \\
210.89 214.99 \\
[361] & 217.15 216.47 217.60 222.42 223.01 223.26 223.50 223.43
\]
2.4. Case Study

We see that no matter what the threshold $d$ will be (at least 5 and at most 223), the change-point estimator is consistently $\hat{\nu} = 234$.

The estimation for the post-change parameter is considered in Chapter 4.
Inference for Change Point and Post Change Means
After a CUSUM Test
Wu, Y.
2005, XIII, 158 p., Softcover
ISBN: 978-0-387-22927-0