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

$$M = \sup_{0 \leq k < \inf} S'_k,$$

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 $\hat{\nu} > \nu$ or $\hat{\nu} < \nu$, we can write the bias as

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

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 $E_{\theta_0 \theta}[.]$ the expectation when both P_{θ_0} and P_θ are involved.

Pollak and Siegmund (1986) shows that the quasistationary distribution of T_ν converges to the stationary distribution of T_ν as $d \to \infty$. That means,

$$\begin{aligned}\lim_{d\to\infty}\lim_{\nu\to\infty} P_{\theta_0}(T_\nu < y|N > \nu) &= \lim_{d\to\infty}\lim_{\nu\to\infty} P_{\theta_0}(T_\nu < y)\\ &= P_{\theta_0}(M < y),\end{aligned}$$

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

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 goes below to zero with initial starting point M. Given $\hat{\nu} > \nu$, the bias $\hat{\nu} - \nu$ is asymptotically equal to τ_{-M} plus the length, say γ_m for a CUSUM process T_n starting from zero until the last zero point time under $P_\theta(.)$. Define

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

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

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

We can see that γ_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 ν, 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 σ_M. Thus, as $d, \nu \to \infty$, we have

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

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 $\nu, d \to \infty$,*

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

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 θ_0 and θ 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 τ_- if the random walk goes below zero.

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

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

where β_1 is given in Lemma 1.1.

Proof: By using Wald's likelihood ratio identity by changing the measure $P_\theta(.)$ to $P_{\tilde{\theta}}(.)$ and Lemma 1.1, we have

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

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 $M = 0$ or $M > 0$, we have

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

From Lemma 1.2, we have

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

For the second term in (2.1), by using the Wald's likelihood ratio identity by changing parameters θ to $\tilde{\theta}$ and θ_0 to θ_1, we have

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

and

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

From Corollary 1.1, we know

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

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

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

The first term in (2.2) is

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

The third term in (2.2) is approximately equal to

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

The fourth term in (2.2) is

$$\Delta\Delta_0 \int_0^\infty E_0(R_{-x} - \rho_-) dE_0(R_x - \rho_+) + o(\theta^2).$$

The second term in (2.2), by integrating by parts, can be approximated as

$$e^{\Delta\rho_-}\left(P_{\theta_0}(\tau_+<\infty)-e^{-\Delta_0\rho_+}\right)+\Delta\Delta_0\int_0^\infty(E_0R_x-\rho_+)dx+o(\theta^2).$$

Combining the above approximations, we have

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

$$\begin{aligned}P_{\theta_0\theta}(\tau_{-M}<\infty) &= \frac{\Delta_0}{\Delta+\Delta_0}e^{\Delta\rho_- - \Delta_0\rho_+}+e^{\Delta\rho_-}\left(1-e^{-\Delta_0\rho_+}\right)\\ &\quad + \Delta\Delta_0(-\frac{1}{2}-\rho_-E_0S_{\tau_-}+\int_0^\infty(E_0R_{-x}-\rho_-)dx\\ &\quad + \int_0^\infty(E_0R_{-x}-\rho_-)d(E_0R_x-\rho_+)+\int_0^\infty(E_0R_x-\rho_+)dx)\\ &\quad + o(\theta^2).\end{aligned}$$

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,*

$$\begin{aligned}E_{\theta_0\theta}[\tau_{-M};\tau_{-M}<\infty] &= -\frac{\Delta_0}{\tilde{\mu}}\left(\frac{1}{(\Delta+\Delta_0)^2}-\frac{\rho_-}{\Delta+\Delta_0}\right)\\ &\quad -\frac{\Delta_0}{\tilde{\mu}}(\frac{1}{2}+\rho_-(E_0S_{\tau_+}-\rho_+)-\int_0^\infty(E_0R_{-x}-\rho_-)dx\\ &\quad -\int_0^\infty(E_0R_{-x}-\rho_-)d(E_0R_x-\rho_+)\\ &\quad -\int_0^\infty(E_0R_x-\rho_+)dx)+o(1).\end{aligned}$$

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

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

The first term in (2.3) can be approximated by using Lemmas 2.2 and 1.2 as

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

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

$$
\begin{aligned}
-\textstyle\int_0^\infty & E_\theta[\tau_{-x};\tau_{-x}<\infty]dP_{\theta_0}(\tau_x<\infty) \\
&= -\int_0^\infty E_{\tilde{\theta}}\left[\tau_{-x}e^{-\Delta(x-R_{-x})}\right]dE_{\theta_1}e^{-\Delta_0(x+R_x)} \\
&= -\int_0^\infty E_{\tilde{\theta}}(\tau_{-x})e^{-\Delta(x-\rho_-)}dE_{\theta_1}e^{-\Delta_0(x+R_x)}+o(1) \\
&= -\frac{1}{\tilde{\mu}}[\int_0^\infty(-x+E_0R_{-x})e^{-\Delta(x-\rho_-)}de^{-\Delta_0(x+\rho_+)} \\
&\quad -\Delta_0\int_0^\infty(-x+E_0R_{-x})d(E_0R_x-\rho_+)]+o(1).
\end{aligned}
\tag{2.4}
$$

The first term in (2.4) is approximately equal to

$$
-\frac{\Delta_0}{\tilde{\mu}}\left[e^{\Delta\rho_--\Delta_0\rho_+}\left(\frac{1}{(\Delta+\Delta_0)^2}-\frac{\rho_-}{\Delta+\Delta_0}\right)-\int_0^\infty(E_0R_{-x}-\rho_-)dx\right]+o(1).
$$

The second term in (2.4) is equal to

$$
\begin{aligned}
& -\frac{\Delta_0}{\tilde{\mu}}\left[\int_0^\infty(x-\rho_-)d(E_0R_x-\rho_+)-\int_0^\infty(E_0R_{-x}-\rho_-)d(E_0R-x-\rho_+)\right] \\
= \quad & -\frac{\Delta_0}{\tilde{\mu}}[\rho_-(E_0S_{\tau_+}-\rho_+)-\int_0^\infty(E_0R-x-\rho_+)dx \\
& -\int_0^\infty(E_0R_{-x}-\rho_-)d(E_0R_x-\rho_+)].
\end{aligned}
$$

Combining the above results, we complete the proof.

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

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

For the second term in (2.5), we write

$$
\begin{aligned}
E_{\theta_0}\left[\sigma_M E_{\tilde{\theta}}e^{\Delta(-M+R_{-M})}\right] &= E_{\theta_0}\left[\sigma_M e^{-\Delta M}\right]e^{-\Delta\rho_-} \\
&\quad + E_{\theta_0}\left[\sigma_M e^{-\Delta M}\left(E_{\tilde{\theta}}e^{\Delta R_{-M}}-e^{\Delta\rho_-}\right)\right].
\end{aligned}
$$

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

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

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

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

Note that K is a geometric random variable with

$$
P(K=k)=p^k(1-p)
$$

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 (τ_+, S_{τ_+}).

Thus,

$$\begin{aligned}
E_{\theta_0}\left[\sigma_M e^{-\Delta M}\right] &= E_{\theta_0}\left[\tau_+^{(K)} \exp(-\Delta S_{\tau_+^{(K)}})\right] \\
&= \sum_{k=1}^{\infty} E_{\theta_0}\left[\tau_+^{(k)} \exp(-\Delta S_{\tau_+^{(k)}}); K = k\right] \\
&= \sum_{k=1}^{\infty} k E_{\theta_0}\left[\tau_+ \exp(-\Delta S_{\tau_+}); \tau_+ < \infty\right] \\
&\quad \times \left(E_{\theta_0}\left[\exp(-\Delta S_{\tau_+}); \tau_+ < \infty\right]\right)^{k-1} P_{\theta_0}(\tau_+ = \infty) \\
&= \frac{E_{\theta_0}\left[\tau_+ \exp(-\Delta S_{\tau_+}); \tau_+ < \infty\right]}{\left(1 - E_{\theta_0}\left[\exp(-\Delta S_{\tau_+}); \tau_+ < \infty\right]\right)^2} P_{\theta_0}(\tau_+ = \infty).
\end{aligned}$$

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

Lemma 2.5: *As $\theta_0, \theta \to$ at the same order,*

$$\begin{aligned}
&1 - E_{\theta_0}\left[\exp(-\Delta S_{\tau_+}); \tau_+ < \infty\right] \\
&= (\Delta + \Delta_0) E_0 S_{\tau_+} \exp\left(-(\Delta - \theta_0)\rho_+ + \frac{1}{2}(\Delta - \theta_0)^2(\rho_+^{(2)} - \rho_+^2) - \frac{\theta_0^2}{2}\frac{\alpha_1}{E_0 S_{\tau_+}}\right) \\
&\qquad \times (1 + o(\theta^2)).
\end{aligned}$$

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

$$P_{\theta_0}(\tau_+ = \infty) = \Delta_0 E_0 S_{\tau_+} \exp\left(\theta_0 \rho_+ + \frac{1}{2}\theta_0^2\left(\rho_+^{(2)} - \rho_+^2 - \frac{\alpha}{E_0 S_{\tau_+}}\right)\right)(1 + o(\theta_0^2)).$$

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

Lemma 2.6: *As $\theta_0, \theta \to$ at the same order,*

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

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

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

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

$$E_{\theta_0}\left[\sigma_M e^{-\Delta M}\left(E_{\tilde{\theta}} e^{\Delta R_{-M}} - e^{\Delta \rho_-}\right)\right] \tag{2.6}$$

$$\begin{aligned} &= \Delta E_{\theta_0}[\sigma_M(E_0 R_{-M} - \rho_-)](1 + o(1)) \\ &= -\Delta \int_0^\infty E_{\theta_0}[\sigma_x | M = x](E_0 R_{-x} - \rho_-) dP_{\theta_0}(M > x) \\ &= \Delta\Delta_0 \int_0^\infty E_{\theta_0}[\sigma_x | M = x](E_0 R_{-x} - \rho_-) d(x + E_0 R_x)(1 + o(1)). \end{aligned}$$

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

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

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 $\theta_0 \to 0$,*

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

Finally, we have the following result:

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

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

$$-\frac{\Delta_0}{\mu_1(\Delta+\Delta_0)^2} e^{\gamma\Delta/3-2\theta(\theta-\theta_0)(\rho_+^{(2)}-\rho_+^2-\alpha_1/E_0S_{\tau_+})-2(\theta-\theta_0)^2(\alpha_1/E_0S_{\tau_+})}(1+o(\theta^2)).$$

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*

$$\lim_{d\to\infty}\lim_{\nu\to\infty} E^{\nu}[\hat{\nu}-\nu|N>\nu] = -\frac{1}{\tilde{\mu}\Delta}\left(\frac{\Delta}{\Delta+\Delta_0}e^{\Delta\rho_- -\Delta_0\rho_+} + e^{\Delta\rho_-}\left(1-e^{-\Delta_0\rho_+}\right)\right)$$

$$\begin{aligned}
&-\frac{\Delta}{\tilde{\mu}}\left(\frac{1}{(\Delta+\Delta_0)^2}-\frac{\rho_-}{\Delta+\Delta_0}\right)e^{\Delta\rho_- -\Delta_0\rho_+}\\
&+\frac{\theta_0}{\theta-\theta_0}\frac{\beta_1}{E_0S_{\tau_+}}+\frac{2\theta_0}{\theta}\rho_-\rho_+\\
&-\frac{\theta}{\theta-\theta_0}\left(\rho_+^{(2)}-\rho_+^2-\frac{\alpha_1}{E_0S\tau_+}\right)+o(1).
\end{aligned}$$

A similar result can be obtained for the absolute bias $E^{\nu}[|\hat{\nu}-\nu||N>\nu]$. In particular, when $\theta=\theta_1$, we have the following result.

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

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

$$\begin{aligned}
&= -\frac{3}{4\Delta_0}\left(\frac{1}{\mu_0}+\frac{1}{\mu_1}\right)-\frac{\gamma}{12}\left(\frac{1}{\mu_0}+\frac{1}{\mu}\right)-\frac{1}{2}\frac{\beta_1}{E_0S_{\tau_-}}\\
&\quad -\frac{1}{2}\left(\rho_+^{(2)}-\rho_+^2-\frac{\alpha_1}{E_0S_{\tau_+}}\right)+o(1)\\
&= -\frac{\gamma}{4\theta_0}+\frac{17\gamma^2}{288}-\frac{\kappa}{16}-\frac{1}{2}\frac{\beta_1}{E_0S_{\tau_-}}-\frac{1}{2}\left(\rho_+^{(2)}-\rho_+^2-\frac{\alpha_1}{E_0S_{\tau_+}}\right)+o(1).
\end{aligned}$$

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

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

Some tedious simplifications give the expected results.

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 1.2, we have

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

and thus,

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

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

$$\begin{aligned}
E_\theta(\gamma_m) &= \frac{1}{2\theta^2} - \frac{1}{4} + o(1), \\
P_{\theta_0\theta}(\tau_{-M} < \infty) &= -\frac{\theta_0}{\theta - \theta_0} e^{-\theta(\theta-\theta_0)} + o(\theta^2), \\
E_{\theta_0\theta}[\tau_{-M}; \tau_{-M} < \infty] &= -\frac{\theta_0}{2\theta(\theta-\theta_0)^2} e^{(\theta-\theta_0)^2} + o(1), \\
E_{\theta_0\theta}[\sigma_M; \tau_{-M} < \infty] &= \frac{1}{2\theta_0^2} - \frac{1}{2(\theta-\theta_0)^2} + o(1).
\end{aligned}$$

Summarizing the above results, we have the following corollary:

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

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

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

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

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 the sequential sampling plan case, which gives

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

which is zero at $\theta = -\theta_0$ and at $\theta = -\theta_0$,

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

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 θ. 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 ν is very little. However, as the local bias is at the order $O(1/\theta_0^2)$ at $\theta = -\theta_0$, which approaches ∞ 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 worths 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

ν	θ_0	θ	$E[\hat{\nu}-\nu\|N>\nu]$	$E[\|\hat{\nu}-\nu\|\|N>\nu]$
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(2.149)
		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)

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

At $\theta=-\theta_0$,

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

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

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

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

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 qq-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

$$\theta_0 = -1, \quad \theta_1 = 1, \quad and \quad \sigma^2 = 1.$$

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:

[1] 0.00 0.00 0.10 0.00 0.00 0.00 1.30 0.00 0.00 0.00 0.00 0.32 0.00 0.00 0.00
[16] 0.12 0.00 1.41 1.54 0.22 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 1.61 2.69
[31] 3.50 5.74 6.02 7.16 9.35 9.82 12.09 11.73 11.13 11.18 12.33 14.32 18.47 19.68 21.86
[46] 20.70 19.70 20.85 22.54 23.77 25.42 26.46 27.35 28.26 30.47 31.51 33.36 34.79 34.27 36.00
[61] 37.55 38.43 39.47 39.72 39.65 40.27 41.50 41.22 42.85 45.24 47.85 48.88 50.18 52.05 53.44
[76] 52.92 53.84 54.65 55.66 56.34 58.19 60.00 61.10 60.50 60.95 60.86 62.29 62.70 62.70 63.98
[91] 63.62 64.18 64.77 63.21 63.71 65.54 65.99 68.05 70.14 72.02

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 -0.982; 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

378	286	871	66
36	114	48	291
15	108	123	4
31	188	457	369
215	233	498	338
11	28	49	336
137	22	131	19
4	61	182	329
15	78	255	330
72	99	195	312
96	326	224	171
124	275	566	145
50	54	390	75
120	217	72	364
203	113	228	37
176	32	271	19
55	23	208	156
93	151	517	47
59	361	1613	129
315	312	54	1630
59	354	326	29
61	58	1312	217
1	275	348	7
13	78	745	18
189	17	217	1357
345	1205	120	
20	644	275	
81	467	20	

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 $i \le \nu$ 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 $i = 1, 2, ...$

Denote

$$f_0(x) = e^{-(x+1)},$$

for $x \le -1$, 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 $x_i = 0.005 y_i - 1$ for $i = 1, ...109$, we can fit the data into the standardized model with

$$\theta_0 = -0.550 \quad and \quad \theta_1 = 0.403.$$

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

[1] 0.88812755 0.06794922 0.00000000 0.00000000 0.07393498
[6] 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
[11] 0.00000000 0.00000000 0.00000000 0.00000000 0.01399443
[16] 0.00000000 0.00000000 0.00000000 0.00000000 0.57343963
[21] 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
[26] 0.72329102 0.00000000 0.00000000 0.42858328 0.00000000

```
[31] 0.00000000 0.00000000 0.16384582 0.00000000 0.00000000
[36] 0.00000000 0.00000000 0.00000000 0.62838514 1.00202291
[41] 0.27175541 0.35568049 0.00000000 0.00000000 0.00000000
[46] 0.00000000 0.80321176 1.36166625 2.12991269 1.41962538
[51] 1.79326315 1.18287677 0.26779256 5.28682351 7.50363341
[56] 8.83632010 12.18700554 11.42676777 11.04115848 12.32389470
[61] 13.81142782 13.05618510 12.71053618 12.61963463 12.89337147
[66] 12.86740552 12.98629592 14.81349220 15.76156031 15.12120366
[71] 15.26007424 15.61373183 15.65270148 17.23514049 24.29215038
[76] 23.56188289 24.19026803 29.74376895 30.48204510 33.20335470
[81] 33.28727977 32.88668534 33.26032311 32.36022404 31.68989711
[86] 32.14345562 31.16343580 32.00660794 32.69493363 33.37326923
[91] 32.46817511 33.11154539 33.75991071 34.31836520 34.17251814
[96] 33.89679987 33.27142836 34.08962526 33.27444197 32.36934786
[101] 32.14857510 31.38334228 31.02770327 38.16962896 37.31448530
[106] 37.39841038 36.43337570 35.52328654 41.30156455
```

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 $i \le \nu$ and $N(0, \sigma^2)$ for $i > \nu$.

For a reference value σ_1^2 for σ^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 χ^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

$$\theta_0 = \frac{1}{\sqrt{2}}\left(1 - \frac{1/\sigma_0^2}{\lambda^*}\right) \quad \text{and} \quad \theta_1 = \frac{1}{\sqrt{2}}\left(1 - \frac{1/\sigma_1^2}{\lambda^*}\right),$$

and generally

$$\theta = \frac{1}{\sqrt{2}}\left(1 - \frac{1/\sigma^2}{\lambda^*}\right).$$

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 $(1+\epsilon_0)^2\chi_p^2$ for $\imath \le \nu$ and $(1+\epsilon_1)^2\chi_p^2$ for $i > \nu$, where χ_p^2 is the standard chi-square random variable with degree of freedom p. Then, by using Example 1.3, we define

$$\lambda^* = \left(\frac{1}{(1+\epsilon_0)^2} - \frac{1}{(1+\epsilon_1)^2}\right) / \ln[(1+\epsilon_1)^2/(1+\epsilon_0)^2],$$

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 θ_0 and θ_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 σ_1^2 is the variance of the last 143 modified observations.

Now, we calculate λ^* as

$$\lambda^* = \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) / \ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right) = (1 - 1/7.239)/\log(7.239) = 0.4354.$$

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

460	457	452	459	462	459	463	479
493	490	492	498	499	497	496	490
489	478	487	491	487	482	479	478
479	477	479	475	479	476	476	478
479	477	476	475	475	473	474	474
474	465	466	467	471	471	467	473
481	488	490	489	489	485	491	492
494	499	498	500	497	494	495	500
504	513	511	514	510	509	515	519
523	531	547	551	547	551	547	541
545	549	545	549	547	543	540	539
532	517	527	540	542	538	541	541
547	553	559	557	557	560	571	571
569	575	580	584	585	590	599	603
599	596	585	587	587	581	583	592
596	596	595	598	598	595	595	592
588	582	576	578	589	585	580	579
584	581	581	577	577	578	580	586
583	581	576	571	575	575	573	577
582	584	579	572	577	571	560	549
556	557	563	564	567	561	559	553
553	553	547	550	544	541	532	525
542	555	558	551	551	552	553	557
557	548	547	545	545	539	539	535
537	535	536	537	543	548	546	547
548	549	553	552	551	550	553	554
551	551	545	547	547	537	539	538
533	525	513	510	521	521	521	523
516	511	518	517	520	519	519	519
518	513	499	485	454	462	473	482
486	475	459	451	453	446	455	452
457	449	450	435	415	398	399	361
383	393	385	360	364	365	370	374
359	335	323	306	333	330	336	328
316	320	332	320	333	344	339	350
351	350	345	350	359	375	379	376
382	370	365	367	372	373	363	371
369	376	387	387	376	385	385	380
373	382	377	376	379	386	387	386
389	394	393	409	411	409	408	393
391	388	396	387	383	388	382	384
382	383	383	388	395	392	386	383
377	364	369	355	350	353	340	350
349	358	360	360	366	359	356	355
367	357	361	355	348	343	330	340
339	331	345	352	346	352	357	

and

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

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

```
[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 0.20
[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 0.00 0.51 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 0.05 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 0.44
[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 0.20 0.06 2.62 3.72 3.11 2.91 2.21
[181] 1.51 0.81 0.27 0.00 0.15 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
```

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.

Springer

http://www.springer.com/978-0-387-22927-0

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