CUSUM TEST
FOR PARAMETER CHANGE
IN TIME SERIES MODELS

Sangyeol Lee
Contents

1. Introduction of the CUSUM test
2. Test for variance change in AR\((p)\) model
3. Test for Parameter Change in Regression Models with ARCH Errors
4. Cusum test for general parameter case
5. Test for Parameter Change based on the MDPDE
Introduction of the CUSUM test

- The cusum test is one of the most frequently used method to detect change points.

- It started in quality control (Page, 1955) and moved to time series analysis since time series data suffer from changes due to the change of governmental policy and critical social events.

- It is easy to understand and implement in actual usage, and can be used for both testing and estimating the locations of changes.

- The cusum test proposed by Inclán and Tiao (1994).

Suppose that \( \{ X_t \} \) is a series of independent r.v.’s with \( (0, \sigma_t^2) \) and one wishes to test

\[
H_0 : \sigma_t^2 \text{ is constant over } X_1, \ldots, X_n \quad \text{vs.} \quad H_1 : \text{not } H_0
\]
• Test statistic:

\[ IT_n := \sqrt{\frac{n}{2}} \max_{1 \leq k \leq n} |D_k|, \text{ where } D_k = \frac{\sum_{t=1}^{k} X_t^2}{\sum_{t=1}^{n} X_t^2} - \frac{k}{n} \]

• If \( X_1, \ldots, X_n \) are i.i.d. \( N(0, \sigma^2) \) under \( H_0 \), then

\[
\sup_{1 \leq k \leq n} \sqrt{\frac{n}{2}}|D_k| \xrightarrow{w} \sup_{0 \leq s \leq 1} |W^o(s)|,
\]

where \( W^o \) is a Brownian bridge: Gaussian process with zero mean and \( Cov(W^o(s), W^o(t)) = s \wedge t \) for all \( s, t \) in \([0,1]\)

• Estimate of the change point: \( k^* = \arg \max_k |D_k| \).

• This procedure can be implemented for detecting multiple changes.
The background for the weak convergence is Donsker’s invariance principle. Note that

\[
\sqrt{n/2} D_k = \frac{1}{\sqrt{n} \sqrt{2}} \left| \sum_{t=1}^{k} X_t^2 - \left( \frac{k}{n} \right) \sum_{t=1}^{n} X_t^2 \right| \times \frac{1}{n^{-1} \sum_{t=1}^{n} X_t^2}.
\]

\[
\left( n^{-1} \sum_{t=1}^{n} X_t^2 \approx \sigma^2 \text{ and } \sqrt{\text{Var}(X_1^2)} = \sqrt{2}\sigma^2 \right)
\]

So,

\[
\sqrt{n/2} D_{[ns]} \to W^o(s), \quad 0 \leq s \leq 1.
\]

\[
\Rightarrow \sup_{1 \leq k \leq n} \sqrt{n/2} |D_k| = \sup_{0 \leq s \leq 1} \sqrt{n/2} |D_{[ns]}| \xrightarrow{d} \sup_{0 \leq s \leq 1} |W^o(s)|.
\]

Reject \( H_0 \) if \( \sup_{1 \leq k \leq n} \sqrt{n/2} |D_k| \) is large.
• For a general i.i.d. case, we can construct the cusum test:

\[ T_n = \frac{1}{\sqrt{n\hat{\tau}^2}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} X_t^2 - \left( \frac{k}{n} \right) \sum_{t=1}^{n} X_t^2 \right|, \]

where

\[ \hat{\tau}^2 = n^{-1} \sum_{t=1}^{n} X_t^4 - \left( n^{-1} \sum_{t=1}^{n} X_t^2 \right)^2 \approx \tau^2 = Var(X_1^2). \]

– This can be extended to stationary time series and the residuals from AR models with unit roots and GARCH models.
Test for variance change in AR($p$) model

- Unstable AR($p$) model:

\[ X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \epsilon_t, \]  

(1)

where $\epsilon_t$ are iid r.v.'s with $E\epsilon_1 = 0$, $E\epsilon_1^2 = \sigma^2$ and $E\epsilon_1^4 < \infty$ and the corresponding characteristic polynomial $\phi(z)$ has a decomposition

\[ \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \]

\[ = (1 - z)^a (1 + z)^b \prod_{m=1}^{l} (1 - 2 \cos \theta_m z + z^2)^{d_m} \psi(z), \]

where $a, b, l, d_m$ are nonnegative integers, $\theta_m$ belongs to $(0, \pi)$ and $\psi(z)$ is the polynomial of order $q = p - (a + b + 2d_1 + \cdots + 2d_l)$ that has no zeros on the unit disk in the complex plane.
• Our goal is to test the following hypotheses:

\[ H_0 : \text{the } \epsilon_t \text{ have the same variance } \sigma^2 \text{ vs. } H_1 : \text{not } H_0. \]

In order to perform a test, we employ the cusum of squares test statistic:

\[
T_n = \frac{1}{\sqrt{n \kappa}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} \epsilon_t^2 - \frac{k}{n} \sum_{t=1}^{n} \epsilon_t^2 \right|, \tag{2}
\]

where \( \kappa^2 = Var(\epsilon_1^2) \).
Since $\epsilon_t$ are unobservable, we use the cusum of squares test $T_{n1}$ based on the residuals:

$$T_{n1} = \frac{1}{\sqrt{n\hat{\kappa}_n}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} \hat{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^{n} \hat{\epsilon}_t^2 \right|,$$

where

$$\hat{\epsilon}_t = X_t - \hat{\phi}_n X_{t-1}, \quad t = 1, \ldots, n,$$

$$X_t = (X_t, \ldots, X_{t-p+1})' \text{ with } X_t = 0, \quad \forall t \leq 0,$$

$$\hat{\phi}_n = \left( \sum_{t=1}^{n} X_{t-1} X_{t-1}' \right)^{-1} \sum_{t=1}^{n} X_{t-1} X_t \quad \text{[LSE of } \phi = (\phi_1, \ldots, \phi_p)'\text{]},$$

$$\hat{\kappa}_n^2 = n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^4 - \left( n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^2 \right)^2 \quad \text{[estimator of } Var(\epsilon_1^2)\text{].}$$
Theorem 1 \textit{Under $H_0$, as $n \to \infty$,}

\[ T_{n1} \xrightarrow{w} \sup_{0 \leq s \leq 1} |W^o(s)|, \quad (4) \]

\textit{where $W^o$ is a Brownian bridge.}

\textit{We reject $H_0$ if $T_{n1}$ is large.}

\begin{itemize}
  \item Unit roots do not affect the limiting distribution of the test unlike the empirical process case (cf. Lee and Wei, 1999).
\end{itemize}
Test for Parameter Change in Regression Models

with ARCH Errors

- Let us consider the model

\[ y_t = \beta' z_t + \epsilon_t, \]  
\[ \epsilon_t = h_t \xi_t, \]
\[ h_t^2 = a(\theta) + \sum_{j=1}^{\infty} b_j(\theta) \epsilon_{t-j}^2, \]

where \( \xi_t \) are iid r.v.’s with zero mean and unit variance, \( \{z_t\} \) is a \( p \)-dim’l strictly stationary process, and \( \theta \rightarrow a(\theta) \) and \( \theta \rightarrow b(\theta) \) are nonnegative continuous real functions defined on a subset \( \mathcal{N} \) in \( \mathbb{R}^d \) with \( a(\theta) > 0 \) and \( \sum_{j=1}^{\infty} b_j(\theta) < \infty \) for all \( \theta \in \mathcal{N} \). We assume that \( y_s, z_s, s < t \) are independent of \( \xi_u, u \geq t \), and \( \{(\epsilon_t, h_t, z_t)\} \) is strong mixing.
– The Model (5) covers a broad class of important models in the financial time series context including GARCH models.

– In particular, it becomes a GARCH(1,1) model if we put $z_t = 0$, $\theta = (\omega, \alpha_1, \alpha_2)$, $\omega, \alpha_1, \alpha_2 > 0$, $\alpha_1 + \alpha_2 < 1$, $a(\theta) = \omega/(1 - \alpha_1)$ and $b_j(\theta) = \alpha_1 \alpha_2^{j-1}$. In this case, \{$(\epsilon_t, h_t, z_t)$\} is geometrically strong mixing (cf. Carrasco and Chen (2002)).
• The objective here is to test the hypotheses

\[ H_0 : \eta = (\beta', \theta')' \text{ remains the same for the whole series } \quad \text{vs.} \quad H_1 : \text{Not } H_0. \]

• For a test, one may construct a cusum test based on \( \{ \hat{e}_t := y_t - \hat{\beta}' z_t \} \) as in Inclán and Tiao (1994) and Kim, Cho and Lee (2000). However, as observed in the simulation study, the test in GARCH(1,1) models is unstable and produces low powers. Thus one has to develop a better test which is not much affected by the GARCH parameters.

• As a candidate, one can naturally consider the cusum test based on \( \{ \xi_t^2 \} \), say,

\[
T_n := \frac{1}{\sqrt{n} \tau} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} \xi_t^2 - \left( \frac{k}{n} \right) \sum_{t=1}^{n} \xi_t^2 \right|, \quad (6)
\]

where \( \tau^2 = Var(\xi_{1}^2) \), since \( T_n \) is free from the GARCH parameters.
• In this case, however, one may speculate whether $T_n$ can detect any changes since $T_n$ itself has no information about the GARCH parameters.

• But since $\xi_t$ are not observable, one should replace $\xi_t^2$’s by the residuals $\hat{\xi}_t^2$, which are obtained via estimating the unknown parameters. Those estimators play an important role to detect changes in the parameters in the presence of changes: the iid property of the true errors still remains when there are no changes. From this reasoning, one can anticipate that the residual cusum test should be more stable and produce better powers.
Now, we construct the residual cusum test. To this end, we assume that

(A1) $E|z_1|^{4+\delta_1} < \infty$, $E|\epsilon_1|^{4+\delta_1} < \infty$ and $E|\xi_1|^{4+\delta_1} < \infty$ for some $\delta_1 > 0$.

(A2) There exists $\delta_2 > 0$ such that

$$\sup_{||\theta - \theta'|| \leq \delta_2, \theta' \in \mathcal{N}} ||\dot{a}(\theta)|| < \infty \quad \text{and} \quad \sum_{j=1}^{\infty} \sup_{||\theta - \theta'|| \leq \delta_2, \theta' \in \mathcal{N}} ||\dot{b}_j(\theta)|| < \infty,$$

where $\dot{a}(\theta)$ and $\dot{b}_j(\theta)$ denote the gradient vectors of $a$ and $b_j$ at $\theta$.

(A3) There exists a sequence of positive integers with $q \to \infty$, $q/\sqrt{n} \to 0$ and $\sqrt{n} \sum_{j=q+1}^{\infty} b_j(\theta) \to 0$ as $n \to \infty$.

(A4) $\{(\epsilon_t, h_t, z_t)\}$ is strong mixing with order $\gamma(h)$ satisfying

$$\sum_{h=1}^{\infty} \gamma(h) \frac{\delta_1}{4+\delta_1} < \infty.$$

• Observe that the last condition in (A3) is satisfied if \( b_j(\theta) \) are geometrically bounded (as in GARCH models), and 
\[ q = [(\log n)]^\zeta, \zeta > 1. \]
Also, if \( z_t \) are identically zero and \( \{y_t\} \) is a GARCH process, \( \{(y_t, h_t)\} \) is geometrically strong mixing (cf. Carrasco and Chen (2002)), so that (A4) is satisfied.

• Now, we construct the residual cusum test. In analogy of \( h_t^2 \), we define

\[
\hat{h}_t^2 = a(\hat{\theta}) + \sum_{j=1}^{q} b_j(\hat{\theta})\hat{\epsilon}_{t-j}^2,
\]
\[
\hat{\epsilon}_t = y_t - \hat{\theta} z_t \quad \text{and} \quad \hat{\xi}_t = \hat{\epsilon}_t / \hat{h}_t,
\]

where \( \hat{\eta} = (\hat{\beta}', \hat{\theta}')' \) is an estimator of \( \eta \) with \( \sqrt{n}(\hat{\eta} - \eta) = O_P(1) \).
Then, we have the following result.

**Theorem 2** Assume that (A1)-(A4) hold. Set

$$\hat{T}_n := \frac{1}{\sqrt{n \hat{\tau}}} \max_{q+1 \leq k \leq n} \left| \sum_{t=q+1}^{k} \hat{\xi}_t^2 - \left( \frac{k}{n} \right) \sum_{t=q+1}^{n} \hat{\xi}_t^2 \right|$$

where $\hat{\tau}^2 = \frac{1}{n-q} \sum_{t=q+1}^{n} \hat{\xi}_t^4 - \left( \frac{1}{n-q} \sum_{t=q+1}^{n} \hat{\xi}_t^2 \right)^2$. Then, under $H_0$, 

$$\hat{T}_n \overset{d}{\rightarrow} \sup_{0 \leq u \leq 1} |B^o(u)|, \quad n \to \infty,$$

where $B^o$ is a Brownian bridge.
• GARCH(1,1) case: we demonstrate that a modification of the test, free from a choice of \( q \), can be constructed for the models with \( h_t^2 \) satisfying a specific equation. Here, considering its extreme popularity in the financial time series context, we concentrate ourselves on the case of GARCH(1,1) errors:

\[
y_t = \beta' z_t + \varepsilon_t, \\
\varepsilon_t = h_t \xi_t, \\
h_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1}^2
\]

with \( \omega > 0, \alpha_1, \alpha_2 \geq 0 \) and \( \alpha_1 + \alpha_2 < 1 \).

• In this case, we can write

\[
h_t^2 = a + \alpha_1 \sum_{j=1}^{\infty} \alpha_2^{j-1} \varepsilon_{t-j}^2
\]
with $a = \omega/(1 - \alpha_1)$, and its estimate is

$$\hat{h}_t^2 = \hat{a} + \alpha_1 \sum_{j=1}^{q} \hat{\alpha}_2^{j-1} \hat{\varepsilon}_{t-j}^2,$$

(9)

- $\hat{\varepsilon}_t = y_t - \hat{\beta}' z_t$, $\hat{\beta}$, $\hat{a}$, $\hat{\alpha}_1$, $\hat{\alpha}_2$ are the estimators for $\beta$, $a$, $\alpha_1$ and $\alpha_2$ satisfying

$$\sqrt{n} (\hat{\beta} - \beta) = O_P (1), \sqrt{n} (\hat{a} - a) = O_P (1),$$

$$\sqrt{n} (\hat{\alpha}_1 - \alpha_1) = O_P (1) \quad \text{and} \quad \sqrt{n} (\hat{\alpha}_2 - \alpha_2) = O_P (1),$$

and $q$ is a sequence of positive integers with $q \to \infty$, $q/\sqrt{n} \to 0$ and $\sqrt{n} \alpha_2^q \to 0$, which ensures (A3).
• Note that the estimate of the conditional variance can be obtained recursively from the equation

$$\tilde{h}_t^2 = \hat{\omega} + \hat{\alpha}_1 \hat{\varepsilon}_{t-1}^2 + \hat{\alpha}_2 \tilde{h}_{t-1}^2,$$

in sofar as initial values $\hat{\varepsilon}_0^2$ and $\tilde{h}_0^2$ are provided. In fact, we can see that for $t \geq 2$, $\tilde{h}_t^2$ can be used instead of $\hat{h}_t^2$.

• Therefore, we have the following.

**Theorem 3** Let $\tilde{h}_t^2$ be the one in (10), and set $\tilde{\xi}_t^2 = \hat{\varepsilon}_t^2 / \tilde{h}_t^2$. Let

$$\tilde{T}_n := \max_{1 \leq t \leq n} T_{n,k} := \frac{1}{\sqrt{n\tilde{\tau}} \max_{1 \leq k \leq n}} \left| \sum_{t=1}^k \tilde{\xi}_t^2 - \left( \frac{k}{n} \right) \sum_{t=1}^n \tilde{\xi}_t^2 \right|,$$

where $\tilde{\tau}^2 = \frac{1}{n} \sum_{t=1}^n \tilde{\xi}_t^4 - \left( \frac{1}{n} \sum_{t=1}^n \tilde{\xi}_t \right)^2$. Then if (A1) holds, under $H_0$, $\tilde{T}_n \overset{d}{\to} \sup_{0 \leq u \leq 1} |B^\alpha(u)|$, $n \to \infty$. 

20
Cusum test for General parameter case

- Suppose that $X_1, X_2, \ldots$ are i.i.d. r.v.'s and one wishes to test for certain parameter change in $\theta \in \mathcal{R}$ based on an estimator $\hat{\theta}_n$.

- Suppose that
\[
\sqrt{n} \left( \hat{\theta}_n - \theta \right) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} l_t(\theta) + \Delta_n,
\]
where $\Delta_n$ is negligible asymptotically and $l_1(\theta), \ldots, l_n(\theta)$ are i.i.d. r.v.'s with
\[
El_t(\theta) = 0, \quad El_t^2(\theta) < \infty.
\]

- If $l_t(\theta)$ are observable, as in the mean and variance case, one may construct the cusum test based on
\[
V_n(s) = \frac{1}{\sqrt{n} \left( El_1^2(\theta) \right)^{1/2}} \left( \sum_{t=1}^{[ns]} l_t(\theta) - \frac{[ns]}{n} \sum_{t=1}^{n} l_t(\theta) \right) \rightarrow W^o(s).
\]
• However, in many situations, those are unobservable. But we can still construct a cusum test as a functional of the estimators.

• Since

\[
\sum_{t=1}^{k} l_t(\theta) = k \left( \hat{\theta}_k - \theta \right) - \sqrt{k} \Delta_k, \quad 1 \leq k \leq n,
\]

one can write

\[
T_n(s) = \frac{[ns]}{\sqrt{n} (El_1^2(\theta))^{1/2}} \left( \hat{\theta}_{[ns]} - \hat{\theta}_n \right)
\]

\[
= V_n(s) + (El_1^2(\theta))^{-1/2} \left( \frac{\sqrt{[ns]}}{\sqrt{n}} \Delta_{[ns]} - \frac{[ns]}{n} \Delta_n \right),
\]

\[
\rightarrow W^o(s) \quad \text{if} \quad \max_{1 \leq k \leq n} \sqrt{\frac{k}{n}} |\Delta_k| = o_P(1).
\]

So,

\[
\sup_{0 \leq s \leq 1} |T_n(s)|^2 \xrightarrow{d} \sup_{0 \leq s \leq 1} |W^o(s)|^2.
\]

• Let \( \{X_t\} \) be stationary time series.

• \( \theta = (\theta_1, \ldots, \theta_p)' \) is the parameter vector which will be examined for constancy.

• \( H_0 : \theta \) does not change for \( X_1, \ldots, X_n \) vs. \( H_1 : \) not \( H_0 \)

• Suppose that \( \hat{\theta}_k \) satisfies the following

\[
\sqrt{k} \left( \hat{\theta}_k - \theta \right) = \frac{1}{\sqrt{k}} \sum_{t=1}^{k} l_t(\theta) + \Delta_k, \]

where \( l_t(\theta) \) forms stationary martingale differences w.r.t. \( \{\mathcal{F}_t\} \) and \( \Gamma = Var(l_t(\theta)) \) is nonsingular.
• If under $H_0$,

$$\frac{1}{\sqrt{n}} \Gamma^{-1/2} \sum_{t=1}^{[ns]} l_t(\theta) \xrightarrow{w} W_p(s)$$

and

$$\max_k \frac{\sqrt{k}}{\sqrt{n}} \| \Delta_k \| = o_P(1), \quad \text{as } n \to \infty,$$

the test statistic

$$T_n := \max_{p \leq k \leq n} \frac{k^2}{n} \left( \hat{\theta}_k - \hat{\theta}_n \right)' \hat{\Gamma}^{-1} \left( \hat{\theta}_k - \hat{\theta}_n \right)$$

where $\hat{\theta}_k = (\hat{\theta}_{k1}, \ldots, \hat{\theta}_{kp})'$ be the estimator of $\theta$ based on $X_1, \ldots, X_k$, $k = 1, \ldots, n$, satisfies

$$\hat{T}_n \xrightarrow{w} \sup_{0 \leq s \leq 1} \sum_{j=1}^p \frac{1}{2} (W_j^o(s))^2$$
**Example 1: Test for RCA(1) model**

- Let \( \{x_t; t = 0, \pm 1, \pm 2, \ldots\} \) be the time series of the RCA(1) model

\[
x_t = (\phi + b_t)x_{t-1} + \epsilon_t, \tag{11}
\]

where

\[
\begin{pmatrix}
    b_t \\
    \epsilon_t
\end{pmatrix}
\iid \begin{pmatrix}
    0 \\
    0
\end{pmatrix}, \begin{pmatrix}
    \omega^2 & 0 \\
    0 & \sigma^2
\end{pmatrix}.
\]

- Nicholls & Quinn (1982) showed that a sufficient condition for the strict stationarity and ergodicity of \( \{x_t\} \) in (11) is \( \phi^2 + \omega^2 < 1 \).

- We assume that \( E\epsilon_t^{2k} < \infty \) and \( E(\phi + b_t)^{2k} < 1 \) for some \( k \) that will be specified later, which immediately yields \( Ex_t^{2k} < \infty \) (cf. Feigin & Tweedie, 1985).
Now, we consider the problem of testing for a change of the parameter vector $\theta = (\phi, \omega^2, \sigma^2)'$ based on a (conditional) LSE $\hat{\theta}$.

Suppose that $x_1, \ldots, x_n$ are a sample from (11) and assume $x_0 = 0$. We intend to test the following hypotheses

$H_0 : \theta = (\phi, \omega^2, \sigma^2)'$ is constant over $x_1, \ldots, x_n.$ vs. $H_1 : \text{not } H_0.$
• In order to construct a cusum test, consider the estimators 
\[ \hat{\theta}_k = (\hat{\phi}_k, \hat{\omega}_k^2, \hat{\sigma}_k^2)' \] 
based on \( x_1, \ldots, x_k \), \( k = 1, \ldots, n \). The estimator \( \hat{\phi}_k \) of \( \phi \) is defined as the minimizer of \( \sum_{t=1}^{k} (x_t - \phi x_{t-1})^2 \), and the estimators \( \hat{\omega}_k^2 \) and \( \hat{\sigma}_k^2 \) are defined as the minimizers of 
\[ \sum_{t=1}^{k} (\hat{u}_{k,t}^2 - \omega^2 x_{t-1}^2 - \sigma^2)^2, \] 
where \( \hat{u}_{k,t} = x_t - \hat{\phi}_k x_{t-1} \), by noticing the equation
\[ E(u_t^2 | F_{t-1}) = \omega^2 x_{t-1}^2 + \sigma^2 \] 
under \( H_0 \), where \( u_t = x_t - \phi x_{t-1} \) and \( F_t = \sigma(\epsilon_s, b_s; s \leq t) \).

• Then, \( \hat{\theta}_k \) is written as
\[
\hat{\theta}_k = \begin{pmatrix}
\hat{\phi}_k \\
\hat{\omega}_k^2 \\
\hat{\sigma}_k^2
\end{pmatrix} = \begin{pmatrix}
\frac{\sum_{t=1}^{k} x_{t-1} x_t}{\sum_{t=1}^{k} x_{t-1}^2} \\
\frac{\sum_{t=1}^{k} x_{t-1}^2}{\sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k})^2} \\
k^{-1} \sum_{t=1}^{k} \hat{u}_{k,t}^2 - \hat{\omega}_k^2 m_{2,k}
\end{pmatrix},
\] 
where \( m_{2,k} = k^{-1} \sum_{t=1}^{k} x_{t-1}^2 \).
Let $\Gamma$ be the $3 \times 3$ symmetric matrix whose entries are

\[
\begin{align*}
\Gamma_{11} &= \frac{\omega^2 E x_1^4 + \sigma^2 E x_1^2}{(E x_1^2)^2}, \\
\Gamma_{22} &= (E x_1^4 - (E x_1^2)^2)^{-2} \left((E b_1^4 - \omega^4)(E x_1^8 - 2 E x_1^2 E x_1^6 + (E x_1^2)^2 E x_1^4)ight. \\
&\quad + 4 \omega^2 \sigma^2 (E x_1^6 - 2 E x_1^2 E x_1^4 + (E x_1^2)^3) + (E \epsilon_1^4 - \sigma^4)(E x_1^4 - (E x_1^2)^2)), \\
\Gamma_{33} &= (E b_1^4 - \omega^4) \left(E x_1^4 - \frac{2 E x_1^2 (E x_1^6 - E x_1^2 E x_1^4)}{E x_1^4 - (E x_1^2)^2}\right) \\
&\quad - 4 \omega^2 \sigma^2 E x_1^2 + E \epsilon_1^4 - \sigma^4 + (E x_1^2)^2 \Gamma_{22}, \\
\Gamma_{12} &= \frac{E b_1^3 E x_1^6 - E b_1^3 E x_1^2 E x_1^4 + E \epsilon_1^3 E x_1^3}{E x_1^2 E x_1^4 - (E x_1^2)^3}, \\
\Gamma_{13} &= \frac{-E b_1^3 E x_1^2 E x_1^6 + E b_1^3 (E x_1^4)^2 - E \epsilon_1^3 E x_1^2 E x_1^3}{E x_1^2 E x_1^4 - (E x_1^2)^3}, \\
\Gamma_{23} &= \frac{(E b_1^4 - \omega^4)(E x_1^6 - E x_1^2 E x_1^4)}{E x_1^4 - (E x_1^2)^2} + 4 \omega^2 \sigma^2 - E x_1^2 \Gamma_{22}.
\end{align*}
\]
Theorem 4  Suppose that $E\epsilon_1^{16} < \infty$ and $E(\phi + b_1)^{16} < 1$. Then under $H_0$, as $n \to \infty$,

$$\frac{[ns]}{\sqrt{n}} \Gamma^{-1/2}(\hat{\theta}_{[ns]} - \hat{\theta}_n) \xrightarrow{w} W_3^o(s).$$

- In view of the result of Theorem 4, we can construct the test statistic

$$T_n = \max_{1 \leq k \leq n} \frac{k^2}{n} \frac{1}{\hat{\Gamma}}(\hat{\theta}_k - \hat{\theta}_n)' \hat{\Gamma}^{-1}(\hat{\theta}_k - \hat{\theta}_n),$$

where $\hat{\Gamma}$ is a consistent estimator of $\Gamma$. We reject $H_0$ at the level $\alpha$ if $T_n \geq C_\alpha$, where $C_\alpha$ is the $(1 - \alpha)$-quantile point of

$$\sup_{0 \leq s \leq 1} \sum_{j=1}^3 (W_j^o(s))^2.$$
Example 2: Test for Autocovariance Function

- Let \( \{x_t; t = 0, \pm 1, \pm 2, \ldots\} \) be a stationary linear process of the form

\[
x_t = \sum_{i=0}^{\infty} a_i \epsilon_{t-i},
\]

(12)

where the real sequence \( \{a_i\} \) satisfies the summability condition \( \sum_{i=0}^{\infty} i |a_i| < \infty \) and \( \epsilon_t \) are iid random variables with mean 0, variance \( \sigma^2_\epsilon \), and \( E|\epsilon_1|^{4\lambda} < \infty \) for some \( \lambda > 1 \). Assume that \( x_1, \ldots, x_n \) are observed, and denote the autocovariance at lag \( h \) by \( \gamma(h) \). As an estimate of \( \gamma(h) \), we use

\[
\hat{\gamma}_n(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_t - \bar{x}_n)(x_{t+h} - \bar{x}_n), \quad \bar{x}_n = \frac{1}{n} \sum_{t=1}^{n} x_t.
\]
Theorem 5  Let

\[ \Lambda_n(s) = \left( \frac{[ns]}{\sqrt{n}} (\hat{\gamma}_{[ns]}(0) - \hat{\gamma}_n(0)), \ldots, \frac{[ns]}{\sqrt{n}} (\hat{\gamma}_{[ns]}(m) - \hat{\gamma}_n(m)) \right)', \]

\[ 0 \leq s \leq 1. \] Then under \( H_0 \), where no changes are assumed to occur in the autocovariance function, we have

\[ \Lambda_n(s)'\Gamma^{-1}\Lambda_n(s) \xrightarrow{w} \sum_{j=0}^{m} (W_j^\circ(s))^2, \]

where \( \Gamma \) is the \((m + 1) \times (m + 1)\) matrix whose \((i, j)\)th entry is

\[ \Gamma_{ij} = \kappa_4 \gamma(i)\gamma(j) + \sum_{r=-\infty}^{\infty} (\gamma(i + r)\gamma(j + r) + \gamma(i - r)\gamma(j + r)), \]

for \( i, j = 0, \ldots, m \), and \( \kappa_4 \) is the kurtosis of \( \epsilon_1 \).
• Since $\Gamma$ is unknown, we should replace it by a consistent estimator $\hat{\Gamma}$.

Now we assume that $\{x_t\}$ in (12) can be rewritten as

$$x_t = \sum_{j=1}^{\infty} \pi_j x_{t-j} + \epsilon_t,$$

where $\pi(z) := 1 - \sum_{j=1}^{\infty} \pi_j z^j$ is analytic on an open set containing the unit disk in the complex plane, and have no zeros in the unit disk.

• Notice that $\{x_t\}$ can be rewritten as in (12), and covers stationary and invertible ARMA processes. We introduce a sequence of positive integers $\{h_n\}$, such that as $n \to \infty$,

$$h_n \to \infty \text{ and } h_n = O(n^\beta) \text{ for some } \beta \in (0, (\lambda - 1)/2\lambda).$$
Then if \( \hat{\kappa}_4 \) is a consistent estimator of \( \kappa_4 \), we have

\[
\hat{\Gamma}_{ij} \xrightarrow{P} \Gamma_{ij},
\]

where

\[
\hat{\Gamma}_{ij} = \hat{\kappa}_4 \hat{\gamma}_n(i) \hat{\gamma}_n(j) + \sum_{r=-h_n}^{h_n} (\hat{\gamma}_n(i + r) \hat{\gamma}_n(j + r) + \hat{\gamma}_n(i - r) \hat{\gamma}_n(j + r)).
\]

Note that a consistent \( \hat{\kappa}_4 \) can be obtained by calculating residuals \( \hat{\epsilon}_t \) via fitting a long AR(\( q \)) model to observations, viz.,

\[
\hat{\kappa}_4 = n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^4 / (n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^2)^2 - 3.
\]

A typical example of \( q \) is \((\log n)^2\).
From Theorem 5 and (14), we obtain the following result.

**Theorem 6** *Under $H_0$,*

$$\Lambda_n(s)'\hat{\Gamma}^{-1}\Lambda_n(s) \xrightarrow{w} \sum_{j=0}^{m} (W_j^\circ(s))^2.$$  

Theorem 6 ensures the Brownian bridge result for the cusum test statistic. The test statistic is defined as

$$\Lambda_n := \max_{m+1 \leq k \leq n} \Lambda_n(k/n)'\hat{\Gamma}^{-1}\Lambda_n(k/n).$$
Test for Parameter Change based on the MDPDE

- Suppose that the time series are contaminated by outliers, and we observe

\[ X_t = (1 - p_t)X_{o,t} + p_tX_{c,t}, \]

where \( p_t \) are iid Bernoulli r.v.’s with success probability \( p \), \( \{X_{o,t}\} \) is a strong mixing process with the marginal density \( f_{\theta} \), the contaminating process \( \{X_{c,t}\} \) is a sequence of iid r.v.’s, and \( \{p_t\} \), \( \{X_{o,t}\} \) and \( \{X_{c,t}\} \) are all independent. Assume that one wishes to test

\[ H_0 : \theta \text{ does not change over } X_1, \ldots, X_n \text{ vs. } H_1 : \text{not } H_0. \]
Minimum Density Power Divergence Estimator (MDPDE)
Basu, Harris, Hjort and Jones (1998)

• Let \( X_1, \ldots, X_n \) be iid r.v.'s with distribution \( G \) having density \( g \)

• We want to choose a model that fits the data in a reasonable way from
the parametric family of models \( \{F_\theta\} \) indexed by the unknown
parameter \( \theta \in \Theta \subset \mathbb{R}^p \), possessing densities \( \{f_\theta\} \).

• Then the best fitting parameter is defined by

\[
\theta_\alpha := T_\alpha(G) = \arg \min_{\theta \in \Theta} d_\alpha(g, f_\theta),
\]

where \( d_\alpha, \alpha \geq 0 \) is a density power divergence defined by

\[
d_\alpha(g, f) := \begin{cases} 
\int \left\{ f^{1+\alpha}(z) - \left(1 + \frac{1}{\alpha}\right) g(z) f^\alpha(z) + \frac{1}{\alpha} g^{1+\alpha}(z) \right\} \, dz, & \alpha > 0 \\
\int g(z) \left(\log g(z) - \log f(z)\right) \, dz, & \alpha = 0.
\end{cases}
\]
• Note that

$$\theta_\alpha = \arg \min_{\theta \in \Theta} d_\alpha (g, f_\theta) = \arg \min_{\theta \in \Theta} E_G V_\alpha (\theta; X),$$

where \(X \sim G\) and

$$V_\alpha (\theta; x) := \begin{cases} \int f_\theta^{1+\alpha} (z) dz - (1 + \frac{1}{\alpha}) f_\theta^\alpha (x) & , \alpha > 0 \\ -\log f_\theta (x) & , \alpha = 0, \end{cases}$$

so the MDPDE \(\hat{\theta}_{\alpha, n}\) of \(\theta_\alpha\) is defined by

$$\hat{\theta}_{\alpha, n} = \arg \min_{\theta \in \Theta} H_{\alpha, n} (\theta), \quad (15)$$

where \(H_{\alpha, n} (\theta) = n^{-1} \sum_{t=1}^{n} V_\alpha (\theta; X_t)\).

• If \(\alpha = 0\), the MDPDE is nothing but MLE.
• Basu et al. (1998) showed that the MDPDE with \( \alpha > 0 \) has robust features against outliers but still possesses the efficiency that the MLE has when the true density belongs to the parametric family \( \{F_\theta\} \).

• Under certain regularity conditions, \( \hat{\theta}_{\alpha,n} \) is weakly consistent for \( \theta_\alpha \) and \( \sqrt{n}(\hat{\theta}_{\alpha,n} - \theta_\alpha) \) is asymptotically normal with mean zero vector.
Main result

• Our objective here is to test the constancy of the unknown parameter $\theta_\alpha$ through the cusum test, based on the MDPDE, introduced by Lee et al. (2003).

• For this task, we set up the null and alternative hypotheses:

$$H_0 : \theta_\alpha \text{ does not change over } X_1, \ldots, X_n \text{ vs. } H_1 : \text{not } H_0.$$ 

• In order to achieve the asymptotic distribution of the cusum test statistic, we need the strong consistency and the functional central limit theorem for the estimator.
(B1) The distribution $F_\theta$ and $G$ have common support, so that the set $\mathcal{X}$ on which the densities are greater than zero is independent of $\theta$.

(B2) There is an open set $\vartheta$ of the parameter space $\Theta$ containing $\theta_\alpha$ such that for all $x \in \mathcal{X}$, and all $\theta \in \vartheta$, the density $f_\theta(x)$ has continuous partial derivatives of order $r(\geq 0)$ with respect to $\theta$ and

$$E \left| \frac{\partial^j f_\theta(X)}{\partial \theta_{i_1} \cdots \partial \theta_{i_j}} \right| < \infty, \quad 0 \leq j \leq r.$$

(B3) The integral $\int f_\theta^{1+\alpha}(z)dz$ can be differentiated $r$-times ($r \geq 0$) with respect to $\theta$, and the derivative can be taken under the integral sign.

(B4) For each $1 \leq i_1, \cdots, i_r \leq p$, there exist functions $M_{\alpha,i_1 \cdots i_r}(x)$ with $EM_{\alpha,i_1 \cdots i_r}(X) < \infty$ such that

$$\left| \frac{\partial^r V_\alpha(\theta; x)}{\partial \theta_{i_1} \cdots \partial \theta_{i_r}} \right| \leq M_{\alpha,i_1 \cdots i_r}(x)$$

for all $\theta \in \vartheta$ and $x \in \mathcal{X}$.
(B5) There exists a nonsingular matrix $J_\alpha$, defined by

$$J_\alpha := \frac{1}{1 + \alpha} E \left( \frac{\partial^2 V_\alpha(\theta_\alpha; X)}{\partial \theta^2} \right).$$

(B6) $\{X_t\}$ is ergodic and strictly stationary.

- In what follows, we assume that $\theta_\alpha$ exists and is unique, and keep the same definition for the estimator $\hat{\theta}_{\alpha,n}$ as in the iid case. In fact, the estimator is obtained by solving the estimating equations

$$U_{\alpha,n}(\theta) = (1 + \alpha)^{-1} \cdot \frac{\partial H_{\alpha,n}(\theta)}{\partial \theta} = 0.$$
Theorem 7 (Strong consistency) Suppose that $H_0$ holds and Conditions (B1) - (B4) and (B6) hold for some nonnegative integer $r$. Then there exists a sequence $\{\hat{\theta}_{\alpha,n}\}$, such that

$$P \left\{ \hat{\theta}_{\alpha,n} \rightarrow \theta_{\alpha}, \text{ as } n \rightarrow \infty \right\} = 1.$$  \hfill (16)

Also, if $V_{\alpha}(\theta; x)$ is differentiable w.r.t. $\theta$, then we have

$$U_{\alpha,n}(\hat{\theta}_{\alpha,n}) = 0$$  \hfill (17)

for sufficiently large $n$. 

Theorem 8 (functional central limit theorem) Assume that

Conditions (B1)-(B6) hold with $r=3$. In addition, suppose that

(C1) \( \{X_t\} \) is strong mixing with mixing order \( \beta(\cdot) \) of size \( -\gamma/(\gamma - 2) \) for \( \gamma > 2 \), i.e., \( \sum_{n=1}^{\infty} \beta(n)^{1-2/\gamma} < \infty \).

(C2) \( E|\partial V_{\alpha}(\theta_{\alpha}; X)/\partial \theta_i|^\gamma < \infty \) for \( i = 1, \ldots, m \).

(C3) \( nK_{\alpha,n} \to K_{\alpha} \) for some positive definite and symmetric matrix \( K_{\alpha} \), where \( K_{\alpha,n} \) is the covariance matrix of \( U_{\alpha,n}(\theta_{\alpha}) \).

Then, under \( H_0 \), we have

\[
\frac{[ns]}{\sqrt{n}} \left( \hat{\theta}_{\alpha,[ns]} - \theta_{\alpha} \right) \xrightarrow{w} J_{\alpha}^{-1} K_{\alpha}^{1/2} W_p(s).
\]
• Utilizing Theorem 8, we can have the following.

**Theorem 9** Define

$$
T_n^0(\alpha) := \max_{p \leq k \leq n} \frac{k^2}{n} \left( \hat{\theta}_{\alpha,k} - \hat{\theta}_{\alpha,n} \right)' \left( J_\alpha K_\alpha^{-1} J_\alpha \right) \left( \hat{\theta}_{\alpha,k} - \hat{\theta}_{\alpha,n} \right). \quad (18)
$$

Suppose that the conditions of Theorem 8 hold. Then, under $H_0$,

$$
T_n^0(\alpha) \xrightarrow{w} \sup_{0 \leq s \leq 1} \left\| W_p^o(s) \right\|^2.
$$

We reject $H_0$ if $T_n^0(\alpha)$ is large.
• Since $J_\alpha$ and $K_\alpha$ are unknown, we should replace them by consistent estimators $\hat{J}_\alpha$ and $\hat{K}_\alpha$.

• Note that

$$J_\alpha = \int u_{\theta_\alpha}(z)u_{\theta_\alpha}(z)' f_{\theta_\alpha}^{1+\alpha}(z) dz$$

$$+ \int (i_{\theta_\alpha}(z) - \alpha u_{\theta_\alpha}(z)u_{\theta_\alpha}(z)') (g(z) - f_{\theta_\alpha}(z)) f_{\theta_\alpha}^\alpha(z) dz,$$

where $u_{\theta}(z) = \partial \log f_{\theta}(z)/\partial \theta$ and $i_{\theta}(z) = -\partial u_{\theta}(z)/\partial \theta$. Therefore, if we put

$$\hat{J}_\alpha = \int \left\{ (1 + \alpha)u_{\hat{\theta}_\alpha,n}(z)u_{\hat{\theta}_\alpha,n}(z)' - I_{\hat{\theta}_\alpha,n}(z) \right\} f_{\hat{\theta}_\alpha,n}^{1+\alpha}(z) dz$$

$$+ \frac{1}{n} \sum_{t=1}^n \left\{ I_{\hat{\theta}_\alpha,n}(X_t) - \alpha u_{\hat{\theta}_\alpha,n}(X_t)u_{\hat{\theta}_\alpha,n}(X_t)' \right\} f_{\hat{\theta}_\alpha,n}^\alpha(X_t),$$

then one can show that $\hat{J}_\alpha$ converges to $J_\alpha$ almost surely.
• Under the assumptions of Theorem 8,

\[ K_\alpha = \sum_{k=-\infty}^{\infty} \frac{1}{(1 + \alpha)^2} \text{Cov} \left( \frac{\partial V_\alpha(\theta_\alpha, X_0)}{\partial \theta}, \frac{\partial V_\alpha(\theta_\alpha, X_k)}{\partial \theta} \right) \]

exists due to Theorem 1.5 in Bosq (1996, page 32).

• Assuming that

   (D1) \( \sum_{n=1}^{\infty} \beta(n)^{1/3} < \infty \);

   (D2) \( E|\partial V_\alpha(\theta_\alpha; X)/\partial \theta_i|^6 < \infty \) for \( i = 1, \ldots, m \);

   (D3) there exists a function \( M_\alpha(x) \) with \( EM_\alpha(X)^2 < \infty \) such that

\[ \sum_{i,j=1}^{p} |\partial^2 V_\alpha(\theta; x)/\partial \theta_i \partial \theta_j| \leq M_\alpha(x) \]

for all \( \theta \in \Theta \) and \( x \in \mathcal{X} \),

one can show that \( \hat{K}_\alpha \rightarrow K_\alpha \) in probability, where

\[ \hat{K}_\alpha = \sum_{k=-h_n}^{h_n} \frac{1}{n(1 + \alpha)^2} \sum_{i=1}^{n-k} \frac{\partial V_\alpha(\hat{\theta}_\alpha,n; X_t)}{\partial \theta} \cdot \frac{\partial V_\alpha(\hat{\theta}_\alpha,n; X_{t+k})'}{\partial \theta} \]

and \{h_n\} is a sequence of positive integers such that

\[ h_n \rightarrow \infty \quad \text{and} \quad h_n/\sqrt{n} \rightarrow 0. \]
Then we have the cusum test.

**Theorem 10** Define the test statistic $T_{n3}(\alpha)$ by

$$T_{n3}(\alpha) := \max_{p \leq k \leq n} \frac{k^2}{n} \left( \hat{\theta}_{\alpha,k} - \hat{\theta}_{\alpha,n} \right)' \hat{J}_\alpha \hat{K}_\alpha^{-1} \hat{J}_\alpha \left( \hat{\theta}_{\alpha,k} - \hat{\theta}_{\alpha,n} \right).$$

Suppose that the conditions of Theorem 8 and (D1)-(D3) hold. Then, under $H_0$,

$$T_{n3}(\alpha) \xrightarrow{w} \sup_{0 \leq s \leq 1} \| W_p^0(s) \|^2.$$

We reject $H_0$ if $T_{n3}(\alpha)$ is large.