The Cusum of Squares Test for Scale Changes in Infinite Order Moving Average Processes

SANGYEOL LEE and SIYUN PARK
Seoul National University

ABSTRACT. In this paper we consider the problem of testing for a scale change in the infinite order moving average process

\[ X_j = \sum_{i=0}^{\infty} a_i \varepsilon_{j-i}, \]

where \( \varepsilon_j \) are i.i.d. r.v.s with \( E|\varepsilon_1|^{a} < \infty \) for some \( a > 0 \). In performing the test, a cusum of squares test statistic analogous to Inclán & Tiao’s (1994) statistic is considered. It is well-known from the literature that outliers affect test procedures leading to false conclusions. In order to remedy this, a cusum of squares test based on trimmed observations is considered. It is demonstrated that this test is robust against outliers and is valid for infinite variance processes as well. Simulation results are given for illustration.

Key words: cusum of squares test, infinite order moving average processes, infinite variance processes, mixingale central limit theorem, robust test, test for a scale change

1. Introduction

Since economic time series are frequently affected by events such as changes in fiscal or monetary policy, the problem of testing the parameter constancy of a time series has received considerable attention from researchers; see, for example, Bagshaw & Johnson (1977), Picard (1985), Kramer et al. (1988), Tang & MacNeil (1993) and papers cited therein. The problem of testing the change in variance of a time series was first tackled by Hsu et al. (1974). Their work was directed toward the modelling of stock returns by using, instead of the Pareto distribution, a normal probability model with a non-stationary variance subject to step changes at irregular time intervals. As an example of related work, we refer to Wichern et al. (1976), which considers a detection procedure for a sudden variance change in first-order autoregressive models based on a likelihood method; see also Baufays & Rasson (1985).

As in the case of time series, there is a vast number of papers aimed at testing the parameter constancy in independent samples. See Hinkley (1971), Zacks (1983), Csörgő & Horváth (1988) and Krishnaiah & Miao (1988) for reviews of earlier works. We refer to Hsu (1977) for a test of variance changes in an i.i.d. sample. Recently, Inclán & Tiao (1994) proposed a cusum of squares test for testing a variance change in i.i.d. normal r.v.s, based on the earlier work by Brown et al. (1975) that deals with the problem of testing the constancy of the regression coefficients in regression models. In Inclán & Tiao (1994), provided the observations \( \varepsilon_t \), \( t = 1, \ldots, n \), are given, a variance change is detected based on the test statistic:

\[
IT_n = \max_{1 \leq k \leq n} \left( \frac{n}{2} \right)^{1/2} \left[ \frac{1}{n} \sum_{j=1}^{n} \varepsilon_j^2 \right]^{1/2} - \frac{k}{n} \right].
\]

A large value of \( IT_n \) indicates the existence of a variance change. Critical values are obtainable asymptotically, since \( IT_n \) has the same limiting distribution as \( \sup_{0 \leq t \leq 1} |B^t(t)| \), where \( B^t \) is a standard Brownian bridge.
In deriving the limiting distribution of $IT_n$, the key theorem was Donsker’s invariance principle. Actually the normality assumption is not necessarily required for obtaining the limiting distribution. Only a slight modification of $IT_n$ gives the same asymptotic distribution in non-normal r.v.s. Since the invariance principle holds in weakly dependent processes (cf. Billingsley, 1968), Inclán and Tiao’s method can be extended to autocorrelated observations.

In this article, we concentrate on testing for an abrupt change of the variance in MA($\infty$) (infinite order moving average) processes. We adopt the MA($\infty$) process as our basic process because it is a quite general process in time series and covers the well-known ARMA process. In some articles, such as Wichern et al. (1976) and Davis (1979), the time series under consideration is assumed to follow an ARMA($p$, $q$) model, where $p$, $q$ are known orders. However, assuming an ARMA model with known orders might be unrealistic because a sudden change of variance could make it hard to identify the correct order of the model. This is another reason for us to employ the MA($\infty$) process. Later, we will see that the cusum of squares test defined analogously to $IT_n$ (cf. theorem 1) performs well in MA($\infty$) processes.

The objective of this paper is not only to extend Inclán and Tiao’s test to MA($\infty$) processes but also to derive a test which possesses a robust character against outliers. We do so because, in general, outliers affect test procedures and can lead to false conclusions, and this phenomenon can occur in the cusum of squares test. As an example of related work on robust tests, we refer to Davis (1979), which considers a robust method for testing the variance change in ARMA models.

For a robust test, our approach uses the cusum of squares of trimmed observations. The basic idea is to use the observations between the $[nu]$th and $[nv]$th largest observations instead of the whole set of observations, where $u$, $v$ are numbers in $(0, 1)$, and to construct a test statistic similar to (1) (cf. theorem 2). The underlying reasoning is the same as for the use of the trimmed mean for estimating the location of an i.i.d. sample. A merit of the trimming test is that a scale change can also be detected in infinite variance processes, while an ordinary test is of no use in such cases; see Brockwell & Davis (1990, p. 535) for the definition of infinite variance processes. Our simulation results demonstrate that the trimming method produces a robust test.

This article is organized as follows. In section 2, we introduce the ordinary and trimming cusum of squares tests and derive their limiting distributions. In section 3, we perform a local power study of the trimming cusum of squares test and show that it has a non-trivial local power under a certain sequence of contiguous alternatives. In section 4, we present simulation results in order to evaluate the performance of the cusum of squares tests. All the theorems in sections 2 and 3 (theorems 1, 2, 3 and 4) are proved in the appendix.

2. Cusum of squares test

In this section we introduce the cusum of squares test and give the related asymptotic results. Let $\{X_j\}$ be an MA($\infty$) process such that

$$X_j = \sum_{i=0}^{\infty} a_i \varepsilon_{j-i},$$

where $\varepsilon_i$ are i.i.d. r.v.s. In obtaining the limiting distribution of the test statistic, Donsker’s invariance principle for dependent processes plays an important role as we mentioned earlier. Here, we do not assume any mixing conditions for $\{X_j\}$ since it is well-known that $\{X_j\}$ is not necessarily strongly mixing (cf. Bradly, 1986). Instead, we adopt the mixingale approach introduced by McLeish (1975).

First, we extend Inclán and Tiao’s test to the process in (2). To this end, we assume that
The cusum of squares test for scale changes

The cusum of squares test based on trimmed observations in order to handle the case where the data is contaminated by outliers such as large errors, or has the characteristics of a marginal distribution with infinite variance. Towards this end, we consider the model to accommodate outliers as follows:

\[ U_j = (1 - p_j) X_j + p_j V_j, \]

where \( p_j \) are i.i.d. r.v.s with \( 0 \leq p_j \leq 1 \), the contaminating process \( \{ V_j \} \) is a sequence of i.i.d. r.v.s with \( E V_j = 0 \) and \( E V_j^2 < \infty \), and \( \{ p_j \} \), \( \{ V_j \} \) and \( \{ X_j \} \) are all independent. In this case, unlike before, we do not apply the eighth moment condition to \( e_1 \). Instead, we assume that \( E|e_1|^{\alpha} < \infty \) for some \( \alpha > 0 \). Also, we assume that \( \sum_{i=0}^{\infty} |a_i|^{\alpha} < \infty \) and \( \sum_{i=0}^{\infty} |a_i| < \infty \) according to whether \( \alpha < 2 \) or \( \alpha \geq 1 \). Here, we consider the case of \( \alpha < 2 \) to accommodate infinite

The following theorem shows that the test statistic \( T_n \) in (5) below, constructed by analogy to \( IT_n \) in (1), has the same limiting distribution as \( IT_n \) under certain regularity conditions.

**Theorem 1**

Suppose that \( E|e_1|^8 < \infty \) and \( |a_i| \leq c i^{-q} \) for some \( c > 0 \) and \( q > 5/2 \). Assume that a sequence of positive integers \( \{ h_n \} \) satisfies

\[ H: h_n \to \infty \quad \text{and} \quad h_n = O(n^\rho) \quad \text{for some } \rho \in (0, 1/2). \]

Then, as \( n \to \infty \),

\[ \hat{\phi}^2 := \sum_{|h| \leq h_n} \hat{\gamma}(h) E \to \phi^2, \]

and

\[ T_n := \max_{1 \leq k \leq n} \frac{n^{1/2} \sum_{i=1}^{k} X_i^2}{\phi} - \frac{k}{n} \to \sup_{0 < r < 1} |B^r(t)|, \]

where \( B^\rho \) denotes a standard Brownian bridge.

**Remark 1.** The model (2) satisfying the conditions in theorem 1 covers ARMA models, since in ARMA models, the coefficients are geometrically bounded. The existence of \( \phi^2 \) is guaranteed by Phillips & Solo (1992, th. 3.8) or Brockwell & Davis (1990, prop. 7.3.4.). Indeed, one can check that \( \sum_{|h| = \infty} |\hat{\gamma}(h)| < \infty \).

In practice, a large value of \( T_n \) implies a variance change. Given any significance level, the corresponding critical value is obtainable from an existing table. For example, at the level \( \alpha = 0.05 \) we reject the null hypothesis, under which no variance change is assumed to occur, if \( T_n > 1.358 \) (cf. Inclán & Tiao, 1994). In section 4, empirical sizes and powers are simulated for a Gaussian AR(1) process. The results demonstrate the validity of our test.

Although the extension of Inclán and Tiao’s test to MA(\( \infty \)) processes is of much interest, it is not our main objective as mentioned earlier. In the remainder of this section, we develop the cusum of squares test based on trimmed observations in order to handle the case where the data is contaminated by outliers such as large errors, or has the characteristics of a marginal distribution with infinite variance. Towards this end, we consider the model to accommodate outliers as follows:

\[ U_j = (1 - p_j) X_j + p_j V_j, \]
variance processes. Note that (6) becomes an I.O. (innovation outlier) model when \( p_j = 0 \) for all \( j \) and the errors \( \varepsilon_j \) follow a heavy-tailed non-Gaussian distribution, and it becomes an A.O. (additive outlier) model when \( \alpha \geq 2 \) and \( p_j = 1/2 \) for all \( j \) and \( V_j \) has an appropriate distribution (cf. Fox, 1972; Denby & Martin, 1979). It also indicates an S.O. (substitutive outliers) model when \( \alpha \geq 2 \) and \( p_j \) are Bernoulli r.v.s (cf. Bustos, 1982).

Let \( F \) denote the distribution of \( U_1 \) and assume that the density \( f = F' \) satisfies

\[
R: f(x) > 0 \quad \text{for all } x \quad \text{and} \quad \sup_x f(x) < \infty.
\]

For \( u \in (0, 1) \), let \( \xi_u \) be a number such that \( F(\xi_u) = u \). Provided \( U_1, \ldots, U_n \) are given, set

\[
\tilde{\xi}_{nu} = \begin{cases} 
U_{(n[x])}, & nu \text{ is an integer} \\
U_{(n[nu]+1)}, & nu \text{ is not an integer},
\end{cases}
\]

where \( U_{(n1)}, \ldots, U_{(nn)} \) denote the ordered r.v.s of \( U_1, \ldots, U_n \), and \( [x] \) is the largest integer not exceeding \( x \). Let \( u < v \) be numbers in \( (0, 1) \). We denote \( \Psi_j = U_j^2 I(\tilde{\xi}_{nu} < U_j \leq \tilde{\xi}_{nv}) \),

\[
\mu^* = n^{-1} \sum_{j=1}^n \Psi_j^2, \quad \text{and} \quad (\tau^*)^2 = \sum_{|h| \leq h_n} \gamma^*(h),
\]

where

\[
\gamma^*(h) = n^{-1} \sum_{i=1}^{n-|h|} (\Psi_i^2 - \mu^*) \left( \frac{\Psi_i^2}{\Psi_i^2} - \mu^* \right) \quad \text{for } |h| < n,
\]

and \( \{h_n\} \) is a sequence of positive integers satisfying (3).

Throughout the following, \( a \wedge b \) denotes the minimum of \( a \) and \( b \). We introduce below the trimming cusum of squares test.

**Theorem 2**

Suppose that \( E|\varepsilon_1|^\alpha < \infty, |a_i| \leq ci^{-q} \) for some \( \alpha, q > 0 \) with \( (\alpha \wedge 1)q > 7 \), and \( h_n \) satisfies (3) with \( \rho \in (0, 3/8) \). Let

\[
T_n^* = \frac{n^{1/2} \mu^*}{\tau^*} \max_{1 \leq k \leq n} \left| \frac{\sum_{j=1}^k \Psi_j^2}{n} - \frac{k}{n} \right|
\]

Then, under condition \( R \),

\[
T_n^* \to^d \sup_{0 \leq t \leq 1} \left| B^\rho(t) \right| \quad \text{as } n \to \infty.
\]

**Remark 2.** If \( \{a_i\} \) is geometrically bounded, \( \rho \) can take any number in \( (0, 1/2) \). See the proof of lemma 7 in the appendix. If we analyse the data in a real context, the level of \( U_j \) would be unknown and thus the above test may not be directly applicable. In such a case, we could consider a test based on mean adjusted r.v.s. Namely, the test statistic can be constructed based on \( U_j - \tilde{\delta}_U \), where \( \tilde{\delta}_U \) is an estimate of the level \( \delta_U \), such as the trimmed mean. See Lee (1995, 1999) for using the trimmed mean as an estimator of the location parameter in linear processes. Although we do not deal with this case in detail, we can see that the same limiting distribution as in theorem 2 is obtainable. A similar discussion can be made for the test in theorem 1. As with \( T_n \) in (5), a large value of \( T_n^* \) indicates a scale change. The simulation results in section 4 demonstrate that the test based on trimmed observations is robust against outliers while \( T_n \) performs poorly in the presence of outliers.
3. Local power study for $T_n^*$

In this section, we study the local power of the test based on $T_n^*$ in theorem 2 under a series of alternative hypotheses. Here we restrict ourselves to the case that $U_j \equiv X_j$ and $X_j$ are i.i.d. r.v.s for the sake of technical convenience. However, the results presented below may be extended to MA($\infty$) processes under certain regularity conditions. We will show that the test based on $T_n^*$ has a non-trivial local power (cf. theorem 5).

Assume that $X_1, X_2, \ldots$ are i.i.d. r.v.s with the distribution $F$ and density $f$ satisfying (7) and in addition,

$$\sup_x |x^i f''(x)| < \infty \quad \text{and} \quad \sup_x |x^i f^n(x)| < \infty, \quad i = 0, 1, 2. \quad (11)$$

Let $\Theta_j$ be i.i.d. r.v.s, independent of $\{X_j\}$, with $E\Theta_1 = 0$, $E\Theta_1^2 = \sigma^2_\Theta$ and $E|\Theta_1|^3 < \infty$. Set $R_j = X_j + n^{-\frac{1}{4}}\lambda \Theta_j$, $j = 1, \ldots, n$, where $\lambda$ is a real number. To construct alternative hypotheses, we introduce an array of r.v.s $\{X_{nj}; j = 1, \ldots, n\}$, $n \geq 1$, where $X_{nj}$ is identical to $X_j$ under the null hypothesis $H_0$. Consider the following sequence of contiguous alternatives:

$\{H_n: X_{nj} = X_j, \quad j = 1, \ldots, [n\theta]\}$

$\{X_{nj} = R_j, \quad j = [n\theta] + 1, \ldots, n, \quad 0 < \theta < 1\}$

For a real number $x$, we define

$$F_n(x) = n^{-1} \sum_{j=1}^n I(X_{nj} \leq x),$$

and for $u \in (0, 1)$, we set

$$\xi_{nu} = \inf \{x: F_n(x) \geq u\},$$

which is identical to the expression defined in (8). The following is an analogue of Ghosh (1971, th. 1), which is useful for our analysis presented below.

**Theorem 3**

Under $\{H_n\}$,

$$n^{\frac{1}{2}}(\xi_{nu} - \xi_u) = n^{\frac{1}{2}}(u - F_n(\xi_u))/f(\xi_u) + o_P(1), \quad (12)$$

and

$$\xi_{nu} - \xi_u = O_P(n^{-1/2}). \quad (13)$$

Since we are dealing with an i.i.d. sample, we do not need $\tau^*$ in $T_n^*$. Instead, we use $\hat{\tau}$, where

$$\hat{\tau}^2 = n^{-1} \sum_{j=1}^n X_{nj}^4 I(\xi_{nu} \leq X_{nj} \leq \xi_{nu}) - (n^{-1} \sum_{j=1}^n X_{nj}^2 I(\xi_{nu} \leq X_{nj} \leq \xi_{nu}))^2.$$

For $u \in (0, 1)$ and $l = 2, 4$,

$$\left| n^{-1} \sum_{j=1}^n X_{nj}^4 I(X_{nj} \leq \xi_u) - I(X_{nj} \leq \xi_{nu}) \right| \leq \max(\xi_{nu}, \xi_u)|F_n(\xi_u) - u|$$

$$= O_P(n^{-1/2}),$$

which is due to theorem 3, we have that under $\{H_n\}$,

$$\hat{\tau}^2 \stackrel{p}{\rightarrow} \tau^2 := \text{var}(X_1^2 I(\xi_u \leq X_1 \leq \xi_u)) \quad \text{as} \quad n \to \infty.$$
Hence, to derive the limiting distribution of $T_n^*$, it suffices to investigate the asymptotic behaviour of $\hat{H}_n(t) := \hat{H}_n(t) - t\hat{H}_n(1)$, where

$$\hat{H}_n(t) = n^{-1/2} \sum_{j=1}^{[nt]} \{ X_j^2 I(\xi_u \leq X_j \leq \xi_v) - \mu \} \quad \text{and} \quad \mu = EX_1^2 I(\xi_u \leq X_1 \leq \xi_v).$$

We put

$$\hat{H}_n(t) = n^{-1/2} \sum_{j=1}^{[nt]} \{ X_j^2 I(\xi_u \leq X_j \leq \xi_v) - \mu \}$$

and

$$\hat{L}_n(t) = \hat{H}_n(t) - t\hat{H}_n(1).$$

Using theorem 3 and Bernstein’s inequality (cf. Pollard, 1984, p. 193), and following essentially the same argument as in the proof of lemma 6, one can verify that

$$\sup_{0 \leq |t| \leq 1} |\hat{L}_n(t) - \hat{L}_n(t)| = o_P(1),$$

which indicates that the limiting distribution of $\hat{L}_n$ will determine that of $T_n^*$.

Setting

$$\hat{H}_n^*(t) = n^{-1/2} \sum_{j=1}^{[nt]} \{ X_j^2 I(\xi_u \leq X_j \leq \xi_v) - \mu \}$$

and

$$\mu_R = ER_1^2 I(\xi_u \leq R_1 \leq \xi_v),$$

we can write

$$\hat{H}_n(t) = \begin{cases} \hat{H}_n^*(t), & t \leq \theta \\ \hat{H}_n^*(t) + n^{-1/2}([nt] - [n\theta])(\mu_R - \mu) + \omega_n(t), & t \geq \theta, \end{cases}$$

where

$$\omega_n(t) = n^{-1/2} \sum_{j=[nt]+1}^{[nt]} \{ R_j^2 I(\xi_u \leq R_j \leq \xi_v) - \mu_R + \mu - X_j^2 I(\xi_u \leq X_j \leq \xi_v) \}.$$

Notice that

$$\text{var}\left\{ R_j^2 I(\xi_u \leq R_j \leq \xi_v) - X_j^2 I(\xi_u \leq X_j \leq \xi_v) \right\}$$

$$\leq B_1 \{ E[I(\xi_u \leq R_j \leq \xi_v) - I(\xi_u \leq X_j \leq \xi_v)] + 2n^{-1/4}h|E[\Theta_2]| \}$$

$$\leq B_2 n^{-1/4}$$

for $B_1 > 0$ and $B_2 > 0$.

Now, it remains for us to handle $\mu_R - \mu$ in order to obtain the limiting distribution of $\hat{H}_n$.

The following explanation is concerned with this task.

**Theorem 4**

As $n \to \infty$,

$$n^{1/2}(\mu_R - \mu) \to d := (1/2)\lambda^2 \sigma^2 \left\{ (2v - 2u - 2\xi_v f(\xi_v) + 2\xi_u f(\xi_u) + \xi_v^2 f' (\xi_v) - \xi_u^2 f' (\xi_u)) \right\}.$$
Combining this with (15), we can see that
\[ \tau^{-1} \mathcal{H}_n(t) \xrightarrow{d} B(t) + d(t - \theta)I(t \geq \theta), \tag{16} \]
where \( B \) denotes a standard Brownian motion. The following is an immediate result of (14) and (16).

**Theorem 5**
Under \( \{ H_n \} \), \( \mathcal{H}_n \) converges weakly to a Gaussian process \( \mathcal{L} \) as \( n \) tends to \( \infty \), where
\[
\mathcal{L}(t) = \tau D(t) + \tau B'(t),
\]
\[
D(t) = -d(1 - \theta)tI(t < \theta) - d\theta(1 - t)I(t \geq \theta).
\]
Hence, as \( n \to \infty \),
\[ T_n^* \xrightarrow{d} \sup_{0 \leq t \leq 1} |D(t) + B'(t)|. \]

Theorem 5 shows that \( T_n^* \) possesses a non-trivial local power. The result appeals to our intuition because \( |D(t)| \) has its maximum value \( |d\theta(1 - \theta)| \) at \( t = \theta \) and this has a maximum at \( \theta = 1/2 \).

### 4. Simulation results and discussion

In this section, we evaluate the performance of the test statistics \( T_n \) and \( T_n^* \) in theorems 1 and 2 through a simulation study. The empirical sizes and powers are calculated at a nominal level of 0.05, and \( h_n = n^{0.333} \) are used for \( \hat{\phi}^2 \) and \( (\tau^*)^2 \) (cf. theorems 1 and 2). In each simulation 100 initial observations are discarded to remove initialization effects.

In order to examine the performance of \( T_n \), we consider the first order autoregressive process \( X_j = \phi X_{j-1} + \varepsilon_j \), where \( \varepsilon_j \) are i.i.d. standard normal r.v.s and \( X_0 = 0 \). The empirical sizes are calculated with sets of 200, 300, and 500 observations generated from the AR(1) model with \( \phi = 0.1, 0.5, 0.8 \). The figures in the “size” column of Table 1 indicate the proportion of the number of rejections of the null hypothesis \( H_0 \), under which no variance changes are assumed to occur, out of 2000 repetitions. They indicate that size distortions become larger as \( \phi \) tends to 1. This phenomenon is due to the fact that the correlation of the series becomes stronger as \( \phi \) tends to 1. To examine the power, we consider the alternative hypothesis

| Table 1. Empirical sizes and powers of \( T_n \) for the Gaussian AR(1) process |
|---|---|---|---|
| \( n \) | \( \phi \) | 0.1 | 0.5 | 0.8 |
| size | 200 | 0.041 | 0.033 | 0.013 |
| | 300 | 0.045 | 0.032 | 0.025 |
| | 500 | 0.042 | 0.038 | 0.021 |
| power: \( \Delta = 4 \) | 200 | 0.989 | 0.913 | 0.478 |
| | 300 | 1.000 | 0.993 | 0.746 |
| | 500 | 1.000 | 1.000 | 0.963 |
\[ H_1: \epsilon_j \sim \mathcal{N}(0, 1), \quad j = 1, \ldots, [n/2], \]
\[ \epsilon_j \sim \mathcal{N}(0, \Delta), \quad j = [n/2] + 1, \ldots, n, \]
where \( n \) denotes the sample size and \([n/2]\) is the point where the variance change occurs. For \( \Delta = 4, n = 200, 300, 500, \) and the values of \( \phi \) mentioned above, the number of rejections of the null hypothesis are calculated out of 2000 repetitions. The results are summarized in Table 1. There, as might be expected, it can be observed that the power increases as \( n \) increases and \( \phi \) decreases to 0. This enables us to conclude that the test performs well for data with moderate sample size unless it is strongly correlated.

Table 2 summarizes the empirical sizes and powers of \( T_n \) when outliers are involved in the data. Here we assume that \( \{X_j\} \) is contaminated by the outliers \( V_j \), which are i.i.d. \( \mathcal{N}(0, \sigma^2_V) \), so that the observed data follow the model \( U_j = (1 - p_j)X_j + p_jV_j \), where \( p_j \) follow a Bernoulli distribution with success probability \( p = 0.1 \). It is assumed that \( \{p_j\}, \{\epsilon_j\} \), and \( \{V_j\} \) are all independent. The empirical sizes and powers based on the \( U_j \)s are calculated out of 2000 repetitions for \( n = 200, 300, 500, \phi = 0.1, 0.5, 0.8, \sigma^2_V = 25, 100 \) and \( \Delta = 4 \). Table 2 shows that the test has very severe size distortions and very low powers. These results confirm that outliers strongly affect the test in a negative manner.

To examine whether \( T^*_n \) in theorem 2 can remedy the bad effects caused by outliers, the empirical sizes and powers are calculated in the same setting as above using the trimming proportions \( u = 0.05 \) and \( v = 0.95 \). The results in Table 3 show that the size distortions are significantly lessened and the powers are remarkably improved. This demonstrates that the trimming method yields a robust test against outliers.

In order to see how \( T^*_n \) works for an infinite variance process, its performance is evaluated for the AR(1) process with Cauchy-distributed errors. More precisely, we consider the model \( X_j = \phi X_{j-1} + \delta^{1/2}\epsilon_j \), where \( \epsilon_j \) are i.i.d. Cauchy (0,1). Under the null hypothesis \( H_0 \), we assume that \( \delta = 1 \). The alternative under consideration is as follows:

\[ H_1: \delta = 1, \quad j = 1, \ldots, [n/2], \]
\[ \delta = 9, \quad j = [n/2] + 1, \ldots, n. \]

The sizes and powers are calculated for \( n = 200, 300, 500, 1000 \) and \( \phi = 0.1, 0.5, 0.8 \). The results, presented in Table 4, show that the test has a similar pattern in size and power as

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \sigma^2_V )</th>
<th>( \phi )</th>
<th>( 0.1 )</th>
<th>( 0.5 )</th>
<th>( 0.8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>25</td>
<td>0.020</td>
<td>0.024</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.021</td>
<td>0.020</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>25</td>
<td>0.025</td>
<td>0.030</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.031</td>
<td>0.022</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>25</td>
<td>0.024</td>
<td>0.028</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.027</td>
<td>0.030</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td><strong>power: ( \Delta = 4 )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>25</td>
<td>0.034</td>
<td>0.290</td>
<td>0.324</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.211</td>
<td>0.134</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>25</td>
<td>0.306</td>
<td>0.418</td>
<td>0.543</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.041</td>
<td>0.051</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>25</td>
<td>0.481</td>
<td>0.638</td>
<td>0.841</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.053</td>
<td>0.079</td>
<td>0.226</td>
<td></td>
</tr>
</tbody>
</table>

observed in Table 3. The validity of $T_{n}^{*}$ is apparent when it is compared to $T_{n}$; see Table 5, which shows the complete failure of $T_{n}$. For infinite variance processes, a larger sample size and a larger magnitude of change $|\delta|$ are required to constitute a good test.

Although we do not report them here, we also produced simulation results in the case where change point is located at either the $[n/4]$ or $[3n/4]$th position of the series. As might be expected, the powers diminished compared to the case in which the change point is at the middle of the series. In addition, we also produced simulation results when $\sigma_{V}^{2}$ is rather small, say, 5 and 10. Compared to the values reported in Table 3, the powers tend to decrease. However, the power loss was not significant when $n = 300, 500$.

Table 3. Empirical sizes and powers of $T_{n}^{*}$ with trimming proportions $u = 0.05$ and $v = 0.95$ for the Gaussian AR(1) process with outliers

<table>
<thead>
<tr>
<th>n</th>
<th>$\sigma_{V}^{2}$</th>
<th>0.1</th>
<th>0.5</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>200</td>
<td>25</td>
<td>0.031</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>0.031</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>25</td>
<td>0.053</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>0.033</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>25</td>
<td>0.047</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>0.046</td>
<td>0.038</td>
</tr>
<tr>
<td>power: $A = 4$</td>
<td>200</td>
<td>25</td>
<td>0.789</td>
<td>0.675</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>0.826</td>
<td>0.756</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>25</td>
<td>0.945</td>
<td>0.890</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>0.978</td>
<td>0.955</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>25</td>
<td>0.998</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>0.999</td>
<td>0.998</td>
</tr>
</tbody>
</table>

Table 4. Empirical sizes and powers of $T_{n}^{*}$ with trimming proportions $u = 0.05$ and $v = 0.95$ for the AR(1) process with Cauchy errors

<table>
<thead>
<tr>
<th>n</th>
<th>$\sigma_{V}^{2}$</th>
<th>0.1</th>
<th>0.5</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>200</td>
<td>0.034</td>
<td>0.030</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.043</td>
<td>0.032</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.044</td>
<td>0.037</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.041</td>
<td>0.042</td>
<td>0.036</td>
</tr>
<tr>
<td>power: $\delta = 9$</td>
<td>200</td>
<td>0.420</td>
<td>0.302</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.674</td>
<td>0.481</td>
<td>0.197</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.933</td>
<td>0.785</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.998</td>
<td>0.993</td>
<td>0.703</td>
</tr>
</tbody>
</table>

Table 5. Empirical sizes and powers of $T_{n}$ for the AR(1) process with Cauchy errors

<table>
<thead>
<tr>
<th>n</th>
<th>$\sigma_{V}^{2}$</th>
<th>0.1</th>
<th>0.5</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>1000</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>power: $\delta = 9$</td>
<td>1000</td>
<td>0.019</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Despite the fact that the trimming method performs well against outliers, there is a need to check if it performs adequately when it is applied to data without outliers. In order to examine this, the performance of $T_n^*$ is investigated for the Gaussian AR(1) process $\{X_t\}$ and is compared to $T_n$ for $\phi = 0.1, 0.5, 0.8$, $\Delta = 4$ and $n = 200, 300, 500$. Table 6 shows that the sizes are not severely distorted but the powers are not as good as those in Table 1. This result tells us that the trimming method might cause power losses when applied to data with no outliers. However, such a defect will not be very significant when the sample size is fairly large.

In practice, optimal selection of trimming proportions is an important issue. We have seen through a local power study (cf. theorems 4 and 5) that the trimming proportions influence the power of $T_n^*$. The formula in theorem 4 shows that the power depends upon $u$ and $v$. It also indicates difficulty in choosing optimal trimming proportions because the power depends upon the unknown density function of the data. Since it is not feasible to identify the probabilistic structure of outliers in actual practice, it would seem difficult to develop a method to select optimal values of $u$ and $v$ which work appropriately in every situation. Here, we suggest that values of $u = 0.05$, $v = 0.95$ be used since they consistently produced reasonable results in our simulation study.

So far, we have seen that the trimming test works well and has a robust character against outliers. Of course, it is not proper to say that our test always performs well in every situation, particularly when a small sample is provided. However, we believe that our test provides a functional tool to detect a scale change when data is suspected to be contaminated by outliers.

**Acknowledgements**

The authors wish to acknowledge the financial support of the Korea Research Foundation made in the programme year of 1998. We would like to thank an associate editor and two referees for their helpful comments and suggestions which led to a significant improvement in the presentation of the paper.

**References**


Received January 1999, in final form October 2000

Sangyeol Lee and Siyun Park, Department of Statistics, Seoul National University, Seoul, 151-742, Korea.

Appendix

Let $\mathcal{F}_i$ denote the $\sigma$-field generated by $\varepsilon_i$, $i \leq j$ and let $\| \cdot \|_p$ denote the $\mathcal{L}_p$ norm. Throughout the following, $C$ denotes a universal constant.

**Lemma 1**

Under the conditions of theorem 1,

$$
\left\| E(X_i^2 | \mathcal{F}_0) - \nu \right\|_2 = O(i^{-2(q-1)}),
$$

and

\[ E[EX_i^2X_j^2|\mathcal{F}_0] - EX_i^2X_j^2 = O(i^{-(q-1)/2} + j^{-(q-1)/2}). \] (19)

**Proof.** Note that
\[ E(X_i^2|\mathcal{F}_0) - \nu = -\sum_{l=1}^{\infty} a_l^2\sigma_{\varepsilon}^2 + \sum_{l=1}^{\infty} a_l^2\varepsilon_{i-l}^2 + \sum_{l,m=1,\ell \neq \ell}^{\infty} a_l a_m\varepsilon_{i-l}\varepsilon_{i-m}. \]

Using this equation, together with Minkowski’s inequality and the monotone convergence theorem, we get
\[
\|E(X_i^2|\mathcal{F}_0) - \nu\|_2 \leq 2 \sum_{l=1}^{\infty} a_l^4\sigma_{\varepsilon}^2 + \left( \sum_{l=1}^{\infty} |a_l| \right)^2 \sigma_{\varepsilon}^2 = O(i^{-2(q-1)}),
\]
which proves (18). In order to show (19), we introduce \( Y_i = \sum_{l=0}^{i-1} a_l\varepsilon_{i-l} \) and \( Z_i = \sum_{l=1}^{\infty} a_l\varepsilon_{i-l}. \)

Note that \( Y \) and \( Z \) are \( \sigma(\varepsilon; \ell \geq 1) \) and \( \sigma(\varepsilon; \ell \leq 0) \)-measurable, respectively, and \( X_i = Y_i + Z_i. \) Since \( E[E(X_i^2X_j^2|\mathcal{F}_0) - EX_i^2X_j^2] \leq C(\sum_{l=1,\ell} E[Z_l] + EZ_i^2) + \sum_{\ell=1,2} E[Z_lZ_j], \) we obtain (19).

**Proof of theorem 1.** We first prove (4). It can be readily seen that
\[ \hat{\gamma}(h) = n^{-1} \sum_{i=1}^{n} (X_i^2 - \nu)(X_{i+h}^2 - \nu) + R_n(h), \]
where \( \sum_{h=0}^{hn}|R_n(h)| = o_P(1). \) Hence, it suffices to show that
\[ \sum_{|h| \leq h_n} \left\{ n^{-1} \sum_{i=1}^{n} (X_i^2 - \nu)(X_{i+h}^2 - \nu) - \gamma(h) \right\} = o_P(1), \]
which will hold if
\[ n^{-1/2} \sum_{h=0}^{h_n} A_{nh} \rightarrow 0, \] (20)
where \( A_{nh} = E^{1/2}\left\{ n^{-1/2} \sum_{i=1}^{n} (X_i^2 X_{i+h}^2 - EX_i^2 X_{i+h}^2) \right\}^2. \) Put
\[ V_{hl} = EX_i^2 X_{i+h}^2 X_{i+1}^2 X_{i+1+h}^2 - (EX_i^2 X_{i+1+h}^2)^2. \]

Using the stationary property, we can write
\[ A_{nh}^2 = V_{h0} + 2 \sum_{j=1}^{n-1} (1 - l/n)V_{hl}. \]

For \( I > h, |V_{hl}| \) is bounded by
\[
|EX_i^2 X_{i+h}^2 [E(X_{i+1}^2 X_{i+1+h}^2|\mathcal{F}_0) - EX_i^2 X_{i+h}^2]| \\
\leq (EX_i^2 X_{i+h}^2)^{1/2} [E[E(X_{i-h}^2 X_i^2|\mathcal{F}_0) - EX_{i-h}^2 X_i^2]^2]^{1/2} \\
= O((l - h)^{-(q-1)/2}),
\]
where the last equality can be proved in a similar fashion to (19). For \( 0 < l < h, \) we split \( V_{hl} \) into \( I + II, \) where
\[ I = EX_i^2 X_{i+1}^2 X_{i+1+h}^2 X_{i+1+h+l} - (EX_i^2 X_{i+1})^2, \]
\[ II = (EX_i^2 X_{i+1}^2)^2 - (EX_i^2 X_{i+1+h})^2. \]
Similarly to (21), we can see that 
\[ |I| = O(h - l)^{(q-1/2)}. \]
On the other hand, since
\[ II = (EX_1^2X_{1+l}^2 + EX_1^2X_{1+h}^2)(EX_1^2X_{1+l}^2 - EX_1^2X_{1+h}^2), \]
and since the first term of the right hand side in the above equality is bounded by a constant, we have that
\[ |II| \leq C \sum_{j=h,l} |EX_1^2((EX_1^2|F_1) - \nu)|. \]

This and (18) yield
\[ |II| = O(h^{-2(q-1)} + l^{-2(q-1)}). \]
Hence,
\[ |V_{hl}| = O((h - l)^{-2(q-1/2)} + h^{-2(q-1)} + l^{-2(q-1)}), \quad 0 < l < h. \]

Since \( V_{hl} \) is bounded and \( A_{n0} = O(1) \), it follows from (21), (22) and (3) that
\[
\begin{align*}
n^{-1/2} \sum_{h=0}^{h_n} A_{nh} & \leq O(n^{-1/2}) + Cn^{-1/2} \sum_{h=0}^{h_n} \left( \sum_{h < l < n} |V_{hl}|^{1/2} + \sum_{0 < l < h} |V_{hl}|^{1/2} + \sum_{l=0,h} |V_{hl}|^{1/2} \right) \\
& = O(h_n/n^{1/2}) \rightarrow 0.
\end{align*}
\]

This proves (20) and thus (4).

Now we are ready to prove (5). Consider the stochastic process:
\[ S_n(t) = n^{-1/2} \varphi^{-1} \left[ \sum_{j=1}^{[nt]} (X_j - \nu) \right]. \]

According to th. 3.8 of Phillips & Solo (1992), \( S_n \) converges weakly to a standard Brownian motion. Note that \( T_n \) is a function of \( S_n \), viz.,
\[ T_n = (\varphi / \hat{\varphi}) \sup_{0 \leq t \leq 1} |S_n(t) - tS_n(1) + (t - [nt]/n)S_n(1)|. \]

Hence, (5) follows from (4) and the continuous mapping theorem.

Now we prove theorem 2 based on the mixingale central limit theorem of McLeish (1975). Here we only provide the proof for the case that \( p_j \) in (6) are identically 0 so that \( U_j \equiv X_j \), since the other cases can be proven in a similar manner. Throughout the following, we denote \( \tilde{\alpha} = 1 \wedge \alpha \) and \( \beta = (\tilde{\alpha} - 1)/2 \). Here is the definition of a sequence of mixingales introduced by McLeish.

**Definition 1**

Let \( \{ \mathcal{F}_j \} \) be a sequence of non-decreasing sub \( \sigma \)-fields and let \( \{ Y_j \} \) be a sequence of r.v.s adapted to \( \{ \mathcal{F}_j \} \). \( \{ (Y_j, F_j) \} \) is a mixingale if there exists a positive sequence \( \psi_k \rightarrow 0 \) as \( k \rightarrow \infty \), such that \( \| E(Y_{j+k}|F_j) \|_2 \leq \psi_k \). The sequence \( \{ \psi_k \} \) is considered to be of size \(-p\) if there exists a positive sequence \( \{ L(k) \} \) such that

(a) \( \sum_{k=1}^{\infty} (kL(k))^{-1} < \infty \);
(b) \( L(k) - L(k - 1) = O(L(k)/k) \);
(c) \( L(k) \) is eventually non-decreasing;
(d) \( \psi_k = o((k^{1/2}L(k))^{-2p}) \).

Remark 3. A typical example of $L(k)$ satisfying (a)–(c) is $k^d$ for some $d > 0$. According to McLeish (1975, th. 2.5 and 2.6), if $\{Y_j, \mathcal{T}_j\}$ is a mixingale with $\psi_k$ of size $-1/2$, $\{Y^2_j\}$ is uniformly integrable, and

$$|E\{n^{-1}(S_{k+n} - S_k)^2|\mathcal{T}_{k-m}\} - \sigma^2| \to 0$$

as $m, k, n \to \infty$,

where $S_n = Y_1 + \cdots + Y_n$, then $W_n$, defined by $W_n(t) = 1/(\sqrt{n}\alpha)S_{[nt]}$, $0 \leq t \leq 1$, converges weakly to a standard Brownian motion.

Lemma 2

Suppose that the density $f$ of $X_1$ satisfies $\sup_x |f(x)| < \infty$, $E|\epsilon_1| < \infty$ and $|a_i| \leq c_i^{-d}$ for some $c_i > 0$ and $q > 2/\alpha$. Let $W_j = X_j^2 I(\xi_1 \leq X_j \leq \xi_2) - \mu$, where $\xi_1 < \xi_2$ are real numbers and $\mu = EX_1^2 I(\xi_1 \leq X \leq \xi_2)$. Then $\{W_j, \mathcal{T}_j\}$ is a mixingale with the sequence $\psi_k$ of size $-1/2$.

Proof. Split $X_{j+k}$ into $Y + Z$, where $Y = \sum_{i=0}^{k-1} a_i \epsilon_j + k - i$ and $Z = \sum_{i=k}^{\infty} a_i \epsilon_j + i - j$. Note that $Y$ and $Z$ are $\sigma(\epsilon_j; i \geq j + 1)$ and $\sigma(\epsilon_i; i \leq j)$-measurable r.vs, respectively. Fix $A > 0$. Write $E(W_{j+k}|\mathcal{T}_j) = A_1 + A_2$, where

$$A_1 = E(W_{j+k}|\mathcal{T}_j)I(|Z| \leq A) \quad \text{and} \quad A_2 = E(W_{j+k}|\mathcal{T}_j)I(|Z| > A).$$

Set $I_{j+k} = I(\xi_1 \leq X_{j+k} \leq \xi_2)$ and $J = I(\xi_1 \leq Y \leq \xi_2)$. Without loss of generality, we assume that $E|Z|^{\alpha} < 1$. First, we deal with $A_2$. Observe that $W_{j+k}$ is bounded and then

$$\|A_2\|_2 \leq C(E|Z|^{\alpha})^{1/2}. \quad (23)$$

Next, split $A_1$ into $A_3 + A_4$, where

$$A_3 = E(X_{j+k}^2|J_{j+k} - J)|\mathcal{T}_j)I(|Z| \leq A) \quad \text{and} \quad A_4 = E(X_{j+k}^2|J - \mu|\mathcal{T}_j)I(|Z| \leq A).$$

Since $X_{j+k}^2$ is bounded on $(I_{j+k} = 1, J = 0)$ and $(I_{j+k} = 0, J = 1, |Z| \leq A)$, $|A_3|$ is no more than a positive constant times $E(|I_{j+k} - J|\mathcal{T}_j)I(|Z| \leq A)$, which equals

$$E(|I_{j+k} - J| |Z| I(|Z| \leq A)) \leq \sum_{i=1}^{2} |F_Y(\xi_i - Z) - F_Y(\xi_i)|I(|Z| \leq A)$$

$$\leq 2 \sup_x |F(x) - F_Y(x)| + \sum_{i=1}^{2} |F(\xi_i - Z) - F(\xi_i)|I(|Z| \leq A),$$

where $F_Y$ denotes the distribution of $Y$ and the first inequality is due to the independence of $Y$ and $Z$. For any $b > 0$, we have

$$|F(x) - F_Y(x)| \leq E\{I(X_{j+k} \leq x) - I(X_{j+k} \leq x + Z)|I(|Z| \leq b)\} + P(|Z| > b)$$

$$\leq E|Z|^{\alpha} / b^{\alpha} + \max\{|F(x + b) - F(x)|, |F(x) - F(x)|\}.$$

Putting $b = (E|Z|^{\alpha})^{1/2\alpha}$ and using the mean value theorem, we get

$$\sup_{-\infty < x < \infty} |F(x) - F_Y(x)| \leq C(E|Z|^{\alpha})^{1/2},$$

which immediately yields

$$E(|I_{j+k} - J|\mathcal{T}_j)I(|Z| \leq A) \leq C(E|Z|^{\alpha})^{1/2} + |Z|^{\alpha}. \quad (24)$$

Hence, we have

$$\|A_3\|_2 \leq C(E|Z|^{\alpha})^{1/2} + (E|Z|^{2\alpha})^{1/2}. \quad (25)$$

Now, we deal with $A_4$. Setting
$A_5 = E\{X_{j+k}^2 J_i \mid \mathcal{T}_j \} I(\mid Z \mid \leq A) - EX_{j+k}^2 J_i I(\mid Z \mid \leq A)$,

we can write

$A_4 = A_5 - EA_3 - EX_{j+k}^2 I_j I(\mid Z \mid > A) + \mu I(\mid Z \mid > A)$.

Substituting $Y + Z$ for $X_{j+k}$ in $A_5$, one can readily see that

$\|A_5\|_2^2 \leq C \sum_{i=0}^{2} \text{var}(Z^i I(\mid Z \mid \leq A)) \leq CE\mid Z \mid^\alpha$.

Using this and (25), we can show that $\|A_4\|_2$ has the same upper bound as presented in (25). Combining this, (23) and (25), we have

$\|E(W_{j+k} \mid \mathcal{T}_j)\|_2 = O(k^{-\beta})$. \hspace{1cm} (26)

Since $\beta - (1/2 + d) > 0$ for some $d > 0$, $k^{-\beta}$ satisfies (a)–(d) in definition 1, and thus the lemma is established.

**Lemma 3**

Suppose that $\varepsilon_1$ satisfies the conditions in lemma 2 except $q > 3/\alpha$. Then if $\gamma_w(h)$ denotes $EW_1W_{1+h} \mid \gamma_w(h) = O(h^{-\beta})$ and therefore $\sum_{h=-\infty}^{\infty} \mid \gamma_w(h) \mid < \infty$.

**Proof.** Owing to (26), we have $\mid \gamma_w(h) \mid \leq C \|E(W_{1+h} \mid \mathcal{T}_1)\|_2 = O(h^{-\beta})$. Therefore, $\sum_{h=-\infty}^{\infty} \mid \gamma_w(h) \mid < \infty$.

The following is needed to establish a mixingale invariance principle for $\{W_j, \mathcal{T}_j\}$ (cf. remark 1).

**Lemma 4**

Let $\tau^2 = \sum_{h=-\infty}^{\infty} \gamma_w(h)$ and $S_n = W_1 + \ldots + W_n$. Under the conditions of lemma 3,

$E[n^{-1} E(\{S_{m+n} \mid \mathcal{T}_0\} - \tau^2)] \to 0 \hspace{1cm} \text{as} \hspace{1cm} n \to \infty$. \hspace{1cm} (27)

**Proof.** Since by lemma 3 and the dominated convergence theorem,

$n^{-1} \sum_{i,j=m+1}^{m+n} EW_iW_j - \tau^2 = \sum_{|h| < n} (1 - |h|/n) \gamma_w(h) - \tau^2 \to 0 \hspace{1cm} \text{as} \hspace{1cm} n \to \infty$,

(27) follows if

$n^{-1} \sum_{i,j=m+1}^{m+n} E[EW_iW_j \mid \mathcal{T}_0] - EW_iW_j \to 0 \hspace{1cm} \text{as} \hspace{1cm} n \to \infty$. \hspace{1cm} (28)

Putting $V_{ij} = E(\tilde{W}_i \tilde{W}_j \mid \mathcal{T}_0) - E\tilde{W}_i \tilde{W}_j$, where $\tilde{W}_i = W_i + \mu$, we have that

$n^{-1} \sum_{i,j=m+1}^{m+n} E[EW_iW_j \mid \mathcal{T}_0] - EW_iW_j$

$\leq n^{-1} \sum_{i,j=m+1}^{m+n} (E[V_{ij}] + \mu \|E(W_i \mid \mathcal{T}_0)\|_2 + \mu \|E(W_i \mid \mathcal{T}_0)\|_2)$

$\leq n^{-1} \sum_{i,j=m+1}^{m+n} E[V_{ij}] + n^{-1} \sum_{i,j=m+1}^{m+n} O(i^{-\beta} + j^{-\beta})$, 

[© Board of the Foundation of the Scandinavian Journal of Statistics 2001.]
where the last inequality follows from (26). Since \( \beta > 1 \), the second term on the right hand side of the last inequality goes to 0 as \( m, n \to \infty \). Thus, to verify (28) it suffices to show that

\[
E[V_{ij}^2] = O(i^{-\beta} + j^{-\beta}).
\]

(29)

We put \( Y_i = \sum_{j=1}^{i-1} a_i \mathbf{e}_{i-j}, \quad Z_i = \sum_{t=i}^{\infty} a_i \mathbf{e}_{t-i}, \quad J_i = I(\xi_1 \leq Y_i \leq \xi_2) \) and \( S_y = I(|Z_i| \leq A, |Y_i| \leq A) \). Recall that \( I_1 = I(\xi_1 \leq Y_i \leq \xi_2) \). We then split \( V_{ij} \) into \( V_{ij1} + V_{ij2} \), where \( V_{ij1} = V_y S_y \) and \( V_{ij2} = V_y S_y \) where \( S_y^c \) is the complement of \( S_y \). Since \( |V_{ij}| \) is bounded, we have \( E[V_{ij2}] = O(i^{-\beta} + j^{-\beta}) \) by Markov's inequality. Hence, we only have to deal with \( V_{ij1} \). We write \( V_{ij1} = A_1 + A_2, \) where \( A_1 = E(W_i(W_j - X_{i,j}^2)X_{j}J_{j} | \mathcal{F}_0)S_y \) and \( A_2 = E(X_{i,j}^2J_{i}X_{j}J_{j} | \mathcal{F}_0) - E(W_iW_j)S_y \). First, we deal with \( A_1 \). By (24), we have

\[
E[A_1] \leq CE[E[|I_1|J_i - J_{j} | \mathcal{F}_0)S_y] = O(i^{-\beta} + j^{-\beta}).
\]

(30)

Next, we deal with \( A_2 \). The following is easy to check:

\[
E(X_{i,j}^2J_{i}X_{j}J_{j} | \mathcal{F}_0)S_y = (EY_{i,j}^2J_{i}X_{j}J_{j})S_y + \gamma_{ij}, \quad \text{where } E[\gamma_{ij}] = O(i^{-\beta} + j^{-\beta}).
\]

This together with (30) yields \( E(W_iW_j)S_y = (EY_{i,j}^2J_{i}X_{j}J_{j})S_y + \rho_{ij}, \) where \( E[\rho_{ij}] = O(i^{-\beta} + j^{-\beta}) \). Therefore, \( E[A_2] = O(i^{-\beta} + j^{-\beta}) \), which completes the proof.

**Proposition 1**

*Under the conditions of lemma 3, \( W_n \), defined by \( W_n(t) = n^{-1/2} \tau^{-1} S_{\lceil nt \rceil} \), converges weakly to a standard Brownian motion.*

**Proof.** The theorem follows from lemmas 2, 4 and th. 2.5 and 2.6 of McLeish (1975).

**Lemma 5**

*Suppose that the conditions of lemma 3 with \( q > 5/\alpha \) are satisfied. Recall that \( \tilde{W}_i = W_i + \mu \). We put \( \tilde{W} = n^{-1} \sum_{i=1}^{n} \tilde{W}_i, \) and for \( \|h\| < n, \) define

\[
\tilde{\gamma}_n(h) = n^{-1} \sum_{i=1}^{n-\|h\|} (\tilde{W}_i - \tilde{W})(\tilde{W}_{i+\|h\|} - \tilde{W}).
\]

Then if \( \{h_n\} \) is a sequence of positive integers satisfying (3) with \( 0 < \rho \leq 3/8 \),

\[
\tilde{\gamma}^2 := \sum_{\|h\| \leq h_n} \tilde{\gamma}_n(h) \to \tau^2 \quad \text{in probability.}
\]

(31)

**Proof.** Note that (31) holds if

\[
n^{-1/2} \sum_{h=0}^{h_n} B_{nh} \to 0,
\]

(32)

where \( B_{nh}^2 = E\{n^{-1/2} \sum_{i=1}^{n} (\tilde{W}_i \tilde{W}_{i+h} - E\tilde{W}_i \tilde{W}_{i+h})\}^2 \), which equals \( Q_{h0} + 2 \sum_{l=i}^{n} (1 - l/n) Q_{hl} \), with

\[
Q_{hl} = E\tilde{W}_i \tilde{W}_{i+h} \tilde{W}_{i+l} \tilde{W}_{i+h+l} - (E\tilde{W}_i \tilde{W}_{i+h})^2, \quad h, l \geq 0.
\]

Here, following the same arguments as in (21) and (22), and utilizing the arguments in (26) and (29), we can show (32). We omit the details for brevity.
Lemma 6
Suppose that $E[|\varepsilon_1|^a] < \infty$, $|a_i| \leq ci^{-q}$ for some $\alpha$, $q > 0$, and let $\lambda$ be a positive real number satisfying $\lambda + \eta > 1$, $\theta < \lambda < \theta + \beta$, $\eta - 0\beta + 1/2 < 0$ for some $\theta \in (0, 1/2)$ and $\eta \in (0, 1)$. Then if $\xi_n - \xi_u = O_p(n^{-2})$,
\[
\Gamma_n := \sup_{0 \leq \tau \leq 1} \left| \sum_{i=1}^{[n\tau]} \left( X_i^2 I(\xi_u \leq X_i \leq \xi_{nu}) - H(\xi_u, \xi_{nu}) \right) I(\xi_{nu} \geq \xi_u) \right|^p \rightarrow 0, \tag{33}
\]
where $H(\xi_1, \xi_2) = EX_1^2 I(\xi_1 \leq X_1 \leq \xi_2)$ for $\xi_1 \leq \xi_2$. Similarly,
\[
\sup_{0 \leq \tau \leq 1} \left| \sum_{i=1}^{[n\tau]} \left( X_i^2 I(\xi_{nu} \leq X_i \leq \xi_u) - H(\xi_{nu}, \xi_u) \right) I(\xi_{nu} \leq \xi_u) \right|^p \rightarrow 0. \tag{34}
\]

Proof. We only prove (33) because (34) can be proved in a similar fashion. Since $\xi_{nu} - \xi_u = O_p(n^{-2})$, $\Gamma_n$ will go to 0 in probability if
\[
\Gamma_{n1} := \sup_{\xi_1, \xi_u : 0 < \xi_1 \leq \xi_u + K n^{-2}} \left| \sum_{i=1}^{[n\tau]} \left( X_i^2 I(\xi_u \leq X_i \leq \xi) - H(\xi_u, \xi) \right) \right|^p \rightarrow 0, \tag{35}
\]
where $S_n = \{ \xi_1, \xi_u \leq \xi \leq \xi_u + K n^{-2} \}$ and $K$ is any positive real number. We partition the interval $S_n$ by the points $\xi_i = \xi_u + K n^{-2} i / [n\eta]$, $i = 0, \ldots, [n\eta]$. If $\xi \in S_n$, $\xi$ belongs to certain $[\xi_j, \xi_{j+1}]$. In this case,
\[
X_i^2 I(\xi_u \leq X_i \leq \xi) \leq X_i^2 I(\xi_u \leq X_i \leq \xi_{j+1}) \leq X_i^2 I(\xi_u \leq X_i \leq \xi_{j+1}) \leq X_i^2 I(\xi_u \leq X_i \leq \xi_{j+1})
\]
and
\[
H(\xi_{nu}, \xi) \leq H(\xi_{nu}, \xi_{j+1}).
\]
Using the above and the fact: $|H(\xi_{nu}, \xi_{j+1}) - H(\xi_{nu}, \xi)| \leq C n^{-(\lambda + \eta)}$ for $l = j, j + 1$, where $\lambda + \eta > 1$ by our assumption, $\Gamma_{n1}$ is bounded by
\[
\Gamma_{n2} := \max_{0 \leq \tau \leq 1} \max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} d_{ij} \right|^p,
\]
where
\[
d_{ij} = X_i^2 I(\xi_u \leq X_i \leq \xi) - H(\xi_u, \xi).
\]
Here, to show $\Gamma_{n2} = o_P(1)$, we introduce $Y_{mi} = \sum_{j=0}^{m-1} a_j \varepsilon_{i-j}$, where $m = [n\eta]$, and $H_m(\xi_1, \xi_2) = E Y_{mi}^2 I(\xi_1 \leq Y_{mi} \leq \xi_2)$. Using the arguments in (24) and (25), one can show that for any $A < B$,
\[
\sup_{A \leq \xi_1 \leq \xi_2 \leq B} E |X_i^2 I(\xi_1 \leq X_i \leq \xi_2) - Y_{mi}^2 I(\xi_1 \leq Y_{mi} \leq \xi_2)| = O(n^{-0\beta}),
\]
\[
\sup_{A \leq \xi_1 \leq \xi_2 \leq B} |H(\xi_1, \xi_2) - H_m(\xi_1, \xi_2)| = O(n^{0\beta}).
\]
Therefore, since $\eta + 1/2 - 0\beta < 0$, it suffices to verify that
\[
\Gamma_{n3} := \max_{0 \leq \tau \leq 1} \max_{[n\eta], 1 \leq k \leq n} \left| \sum_{i=1}^{k} d_{ij}^{*} \right|^p = o_P(1), \tag{36}
\]
where $d_{ij}^{*} = Y_{mi}^2 I(\xi_u \leq Y_{mi} \leq \xi) - H_m(\xi_u, \xi)$ and $\omega$ is any real number in $(0, 1/2)$. We assume that $k = mu + v$, where $v = 1, \ldots, m - 1$ and write $\sum_{i=1}^{k} d_{ij}^{*} = D_1 + \cdots + D_m$, where $D_i = \sum_{i=1}^{\nu} d_{ij}^{*} i = 1, \ldots, v$ and $D_i = \sum_{i=1}^{\nu} d_{ij}^{*} i = v + 1, \ldots, m$. Note that
each $D_i$ is a sum of independent r.v.s. Since $|d_{ij}^\theta| \leq C$, $E d_{ij}^\theta = 0$, $\text{var}(d_{ij}^\theta) \leq C n^{-\lambda}$ and $u + 1 \leq C n^{1-\theta}$, applying Bernstein’s inequality, we have that for $i = 1, \ldots, m$, and any $\delta > 0$,

$$P(|D_i| > n^{1/2} \delta m^{-1}) \leq 2 \exp\{-2^{-1} n \delta^2 / C n^{-\lambda} + C n^{1/2+\theta} \delta / 3\} \leq 2 \exp(-Bn^{\kappa}) \text{ for some } B, \kappa > 0$$

This in turn implies that

$$P(T_n > \delta) \leq 2(n^\theta + 1)n^{\theta+1} \exp(-Bn^{\kappa}) \to 0,$$

which proves (36).

**Remark 4.** According to Lee (1999), it holds that for every $u \in (0, 1)$,

$$\xi_{nu} - \xi_u = O_P(n^{-\lambda}) \quad (37)$$

for any $\lambda \in (0, 1/(2 + 2/\beta))$. Thus, provided $q \lambda > 7$ (or equivalently $\beta > 3$), one can find real numbers $\eta$ and $\theta$ required in lemma 6. Actually, in a similar fashion to the proof of lemma 6, one can also show that

$$\sup_{0 < u \leq 1} \left| n^{-1/2} \sum_{i=1}^{\lceil n \rceil} \left[ I(\xi_u \leq X_i \leq \xi_{nu}) - (F(\xi_u) - F(\xi_{nu})) \right] I(\xi_{nu} \geq \xi_u) \right| = o_P(1), \quad (38)$$

$$\sup_{0 < u \leq 1} \left| n^{-1/2} \sum_{i=1}^{\lceil n \rceil} \left[ I(\xi_{nu} \leq X_i \leq \xi_u) - (F(\xi_u) - F(\xi_{nu})) \right] I(\xi_{nu} < \xi_u) \right| = o_P(1). \quad (39)$$

The above is useful for establishing the following lemma.

**Lemma 7**

*Under the conditions of theorem 2, as $n \to \infty$,*

$$(\tau^*)^2 \to \tau^2 \text{ in probability.}$$

**Proof.** In view of lemma 5, it suffices to show that $(\tau^*)^2 - \tau^2 \stackrel{p}{\to} 0$. Write $\sum_{|h| \leq h_n} \{\hat{\gamma}(h) - \gamma^*(h)\} = \sum_{|h| \leq h_n} (A_1(h) + A_2(h))$, where

$$A_1(h) = n^{-1} \sum_{i=1}^{n-|h|} (\tilde{W}_i \tilde{W}_{i+|h|} - \mu^2 \psi^2_{i+|h|}),$$

$$A_2(h) = n^{-1} \sum_{i=1}^{n-|h|} \left[ (\mu^* \psi^2_{i+|h|}) - \mu^* \psi^2_{i+|h|} \right] + (\tilde{W}_i \tilde{W}_{i+|h|}) + (\tilde{W}_i^2 - (\mu^*)^2).$$

Due to (39), we have that for all $|h| < n$,

$$|A_1(h)| \leq O_P(1) \left\{ n^{-1} \sum_{i=1}^{n} |I(X_i \leq \xi_u) - I(X_i \leq \xi_{nu})| + n^{-1} \sum_{i=1}^{n} |I(X_i \leq \xi_u) - I(X_i \leq \xi_{nu})| \right\}.$$

In view of (37)–(39), we can write that for some $\lambda > 3/8$,

$$\sum_{|h| \leq h_n} |A_1(h)| \leq O_P(1) h_n \left\{ |F(\xi_{nu}) - F(\xi_u)| + |F(\xi_{nu}) - F(\xi_u)| \right\} + o_P(1) = O_P(n^{\lambda-\lambda}) + o_P(1),$$

where $0 < \rho \leq 3/8$ is the number in (3). Hence, $\sum_{|h| \leq h_0} |A_1(h)| = o_P(1)$. In a similar fashion, one can show that $\sum_{|h| \leq h_0} |A_2(h)| = o_P(1)$, whose proof is omitted for brevity. This establishes the lemma.

**Proof of theorem 2.** Let $W_j = X_j^2 I(\xi_u \leq X_j \leq \xi_v) - \mu, \mu = EX_j^2 I(\xi_u \leq X_j \leq \xi_v)$ and $W_j = \hat{W}_j - \mu$. Since $(x^*)^2 \rightarrow t^2$ by lemma 7, (10) holds only if $B_n^*$, defined by

$$B_n^*(t) = n^{-1/2} \sum_{j=1}^{[n]} (\Psi^2_j - \mu) - tn^{-1/2} \sum_{j=1}^{n} (\Psi^2_j - \mu), \quad 0 \leq t \leq 1,$$

converges weakly to $B^o$. Putting

$$B_n(t) = n^{-1/2} \sum_{j=1}^{[n]} W_j - tn^{-1/2} \sum_{j=1}^{n} W_j, \quad 0 \leq t \leq 1,$$

we can write

$$|B_n(t) - B_n^*(t)| = \sum_{j=1}^{[n]} \left\{ A_n(s_j, \xi_u, \xi_v) + A_n(s_j, \xi_l, \xi_r) \right\} + A_n(t),$$

where

$$A_n(s, a, b) = \left| n^{-1/2} \sum_{j=1}^{[n]} \{ X_j^2 I(a \leq X_j \leq b) - H(a, b) \} \right| I(a \leq b),$$

$H(\cdot, \cdot)$ is the function in lemma 6, and $\sup_{0 \leq t \leq 1} |A_n(t)| = o_P(1)$. From lemma 6, it can be shown that $\sup_{0 \leq t \leq 1} |B_n(t) - B_n^*(t)| = o_P(1)$. Since $B_n$ converges weakly to $B^o$ by proposition 1, the theorem is established.

**Proof of theorem 3.** Let $F_R$ denote the distribution of $R_1$. Recall that the density $f$ of $X_1$ satisfies (7) and (11). By a Taylor series expansion, we have

$$F_R(x) - F(x) = n^{-1/2} \lambda^2 \sigma_0^2 f'(x)/2 + r_n(x), \quad (40)$$

where $\sup_{x} |r_n(x)| = O(n^{-3/4})$. Below, we prove the theorem following the idea of Ghosh (1971).

Let $t$ be any real number. We put

$$V_n = n^{1/2}(\xi_u - \xi_u), \quad W_n = n^{1/2}(F(\xi_u) - F_n(\xi_u))/f(\xi_u)$$

and

$$Z_n = n^{1/2}\{F(\xi_u + tn^{-1/2}) - F_n(\xi_u + tn^{-1/2})\}/f(\xi_u).$$

We now write

$$-f(\xi_u)W_n = n^{-1/2} \sum_{j=1}^{[n\theta]} \{ I(X_j \leq \xi_u) - F(\xi_u) \} + n^{-1/2} \sum_{j=[n\theta]+1}^{n} \{ I(R_j \leq \xi_u) - F_n(\xi_u) \}

+ n^{-1/2}(n - [n\theta])(F_R(\xi_u) - F(\xi_u)).$$

Clearly, the first two terms of the right hand side of the above equality are $O_P(1)$. Meanwhile, the third term is $O(1)$ due to (40). Hence, we have

$$W_n = o_P(1). \quad (41)$$

Next, we prove
We split $f(\xi_n)(W_n - Z_m)$ into $A_1 + A_2 + A_3$, where

$$A_1 = n^{-1/2} \sum_{j=1}^{[n\theta]} \{ I(X_j \leq \xi_n + tn^{-1/2}) - F(\xi_n + tn^{-1/2}) + F(\xi_n) - I(X_j \leq \xi_n) \},$$

$$A_2 = n^{-1/2} \sum_{j=[n\theta]+1}^{n} \{ I(R_j \leq \xi_n + tn^{-1/2}) - F_R(\xi_n + tn^{-1/2}) + F_R(\xi_n) - I(R_j \leq \xi_n) \},$$

$$A_3 = n^{-1/2} \sum_{j=[n\theta]+1}^{n} \{(F_R(\xi_n + tn^{-1/2}) - F(\xi_n + tn^{-1/2})) - (F_R(\xi_n) - F(\xi_n)) \}.$$

Obviously, $A_1$ and $A_2$ are $o_P(1)$. Since $A_3 = o_P(1)$ in view of (40), (42) is proved.

Now, notice that

$$(V_n \leq t) = \{ u \leq F_n(\xi_n + tn^{-1/2}) \} = (Z_m \leq t_n),$$

where $t_n = n^{-1/2}\{-u + F(\xi_n + tn^{-1/2})\}/f(\xi_n) \overset{p}{\to} t$. From this, (42) and (41), one can see that (5) and (4) in lem. 1 of Ghosh are satisfied. Henceforth $V_n - W_n = o_P(1)$, which entails (12). (13) is a direct result of (12).

**Proof of theorem 4.** Write $n^{-1/2}(\mu_R - \mu) = A_1 + A_2$, where

$$A_1 = n^{1/2} E(R_1^2 - X_1^2) I(\xi_n \leq X_1 \leq \xi_0).$$

$$A_2 = n^{1/2} E R_1^2 J_{RX}, \quad \text{where} \quad J_{RX} = I(\xi_n \leq R_1 \leq \xi_0) - I(\xi_n \leq X_1 \leq \xi_0).$$

One can easily see that

$$A_1 = \lambda^2 \sigma_0^2 (u - u).$$

(43)

To deal with $A_2$, we split $A_2$ into $A_{21} + A_{22} + A_{23}$, where

$$A_{21} = n^{1/2} E X_1^2 J_{RX}, \quad A_{22} = 2n^{1/4} \lambda EX_1 \Theta_1 J_{RX} \quad \text{and} \quad A_{23} = \lambda^2 E \Theta_1^2 J_{RX}.$$

First, note that $A_{23} = o_P(1)$ by the dominated convergence theorem. Second, we observe that

$$A_{21} = n^{1/2} E \{G_u(-n^{-1/4} \lambda \Theta_1) - G_u(0) - (G_u(-n^{-1/4} \lambda \Theta_1)) - G_u(0) \},$$

where $G_u(x) = \int_{\xi_u}^{\xi_u + x} y^2 dF(y)$ for $u \in (0, 1)$. By a Taylor series expansion and (11),

$$A_{21} \rightarrow \frac{1}{2} \lambda^2 \sigma_0^2 \{ 2 \xi_u f(\xi_u) + \xi_u^2 f'(\xi_u) - 2 \xi_u f(\xi_u) - \xi_u^2 f'(\xi_u) \}.$$  

(44)

Finally, we deal with $A_{22}$. Putting $\tilde{G}_u(x) = \int_{\xi_u}^{\xi_u + x} y dF(y)$ for $u \in (0, 1)$, we can write that

$$A_{22} = 2n^{1/4} \lambda E \{ \Theta_1 [\tilde{G}_u(-n^{-1/4} \lambda \Theta_1) - \tilde{G}_u(0) - (\tilde{G}_u(-n^{-1/4} \lambda \Theta_1)) - \tilde{G}_u(0)] \}.$$  

By a Taylor series expansion and (11), we have

$$A_{22} \rightarrow -2 \lambda^2 \sigma_0^2 (\xi_u f(\xi_u) - \xi_u f(\xi_u)).$$

Combining this and (43) and (44), we establish the theorem.