

A class of CUSUM tests using empirical distributions for tail changes in weakly dependent processes

JunHyeong Kim^a, Eunju Hwang^{1,b}

^aDepartment of Industrial Engineering, Hanyang University, Korea;

^bDepartment of Applied Statistics, Gachon University, Korea

Abstract

We consider a wide class of general weakly-dependent processes, called ψ -weak dependence, which unify almost all weak dependence structures of interest found in statistics under natural conditions on process parameters, such as mixing, association, Bernoulli shifts, and Markovian sequences. For detecting the tail behavior of the weakly dependent processes, change point tests are developed by means of cumulative sum (CUSUM) statistics with the empirical distribution functions of sample extremes. The null limiting distribution is established as a Brownian bridge. Its proof is based on the ψ -weak dependence structure and the existence of the phantom distribution function of stationary weakly-dependent processes. A Monte-Carlo study is conducted to see the performance of sizes and powers of the CUSUM tests in GARCH(1, 1) models; in addition, real data applications are given with log-returns of financial data such as the Korean stock price index.

Keywords: weak dependence, tail index, CUSUM test, limit theory, GARCH process

1. Introduction

There are extreme outliers of the real data found in many financial markets and telecommunication networks that play a key role to provide future forecasts and optimal services. Extremal events such as high flood levels of rivers or extremal values of environmental indicators are also often more informative than central parts. Statistical inference on the tail of a distribution has been one of the most important issues for extreme value analysis. The tail index of a time series, which is used as a measure of the thickness of the tail, gives more information than other volatility measures such as variance. For example, it determines the existence of higher-order moments and the limit distribution of sample maximum and minimum.

There is empirical evidence that the tail behavior of real data has changed at some point during the observation period, as mentioned by Hoga (2017), who discussed change point tests for the tail index of β -mixing random variables. Concerning the change point tests on the tail index, we refer to Quintos *et al.* (2001), Jureckova and Picek (2001), Jureckova *et al.* (2009), Kim and Lee (2009, 2011, 2012, 2016), and Hoga (2017). Quintos *et al.* (2001) explored the tests of tail shape constancy in the 1997 Asian financial crisis. Jureckova and Picek (2001) and Kim and Lee (2009) proposed tests for Pareto-type tails, Jureckova *et al.* (2009) and Kim and Lee (2012) dealt with tests for tail index in autoregressive models, while Kim and Lee (2011) extended change point tests to a large class of dependent sequences and Kim and Lee (2016) to GARCH types. The change point tests by Quintos *et al.* (2001), Kim and Lee (2009, 2012, 2016) and Hoga (2017) were developed by means of

¹ Corresponding author: Department of Applied Statistics, Gachon University, 1342 Seongnamdaero, Sujeong-gu, Seongnam-si, Gyeonggi-do 13120, South Korea. E-mail: ehwang@gachon.ac.kr

Hill's (1975) estimator; however, Jureckova and Picek (2001) and Jureckova *et al.* (2009) employed empirical processes along with some threshold sequences. Hill's estimator has been extensively discussed and modified by many authors in various time series models. Jureckova's approach also has a powerful performance as seen in the numerical illustration of Jureckova and Picek (2001), who have proposed a class of nonparametric tests on the tail index for independent and identically distributed (iid) random variables. The proposed tests adequately well the tails even for moderate samples.

We adopt the idea of Jureckova and Picek (2001), rather than Hill's estimate approach, and propose a simpler test for the tail change in weakly-dependent processes. We consider a wide class of general weakly-dependent time series sequences, called ψ -weak dependence, introduced by Doukhan and Louhichi (1999) to unify almost all weak dependence conditions such as mixing, associations, Gaussian sequences and Bernoulli shifts. The ψ -weakly dependent processes, including all classes of weakly dependent stationary processes of interest in statistics under natural conditions on the process parameters have been actively studied by many researchers: Coulon-Prieur and Doukhan (2000) for a central limit theorem under the weak dependence condition, Doukhan and Louhichi (2001) for functional estimation of the density, Ango Nze *et al.* (2002) for nonparametric regression estimation with the weak dependence, Ango Nze and Doukhan (2004) for the weakly dependent time series models and applications to econometrics, Doukhan and Neumann (2007), Hwang and Shin (2013) and Doukhan *et al.* (2015) for probabilistic results of the processes such as moment inequalities and distribution functions. Also see Hwang and Shin (2011, 2012a, 2012b, 2016a, 2016b) and references.

In the ψ -weakly dependent processes, we discuss the change point tests for the tail index, applying cumulative sum (CUSUM) statistics based on the empirical process and threshold sequence. CUSUM test statistics are constructed using empirical distribution functions with sample extremes by splitting a the set of observations into N subsamples of size n where nN is the sample size. A key idea of the CUSUM tests is that some suitable threshold sequence can be selected so that the empirical distribution functions at the threshold values are neither zero nor one a.s. The threshold sequence is adopted as in Jureckova and Picek (2001), who proposed nonparametric tests (but not CUSUM tests) on the tail index for iid random variables. The reason why the threshold of the same form can be chosen in our case is explained with the existence of the phantom distribution function of the weakly dependent processes. The notion of the phantom distribution function has been introduced by O'Brien (1987) and investigated by Doukhan *et al.* (2015) along with sufficient conditions on its existence in stationary and ψ -weakly dependent processes. The null limiting distribution of the CUSUM test statistics as both $N \rightarrow \infty$ and $n \rightarrow \infty$ is established to be a Brownian bridge. The proof is addressed by the weak dependence structure and the existence of the phantom distribution function of the stationary processes under the null hypothesis. This present work makes it possible to extend Jureckova and Picek (2001) of the iid case to general weakly-dependent time series models. A Monte-Carlo study is conducted to see the performance of powers and sizes of the CUSUM tests in GARCH(1, 1) models with heavy-tailed noises when tail indices of the noises are changed. Real data applications are given with the financial data of Korean stock price index.

The remainder of the paper is organized as follows. Section 2 presents the weak dependence and main results and Section 3 deals with GARCH(1, 1) processes. In Section 4, a Monte Carlo study and real data analysis are conducted. In Section 5 a concluding remark is stated while in Section 6 a proof is given along with the existing proposition.

2. Main result

We consider stationary and weakly dependent sequences of a time series satisfying weak dependence

conditions in the following, proposed by Doukhan and Louhichi (1999) to unify all classes of times series under natural conditions on parameters. The following is summarized in the same way as in Doukhan *et al.* (2015).

For $k \in \mathbb{N}$, let \mathcal{M}^k be the family of all bounded measurable non-constant functions f on \mathbb{R}^k satisfying

$$\sup_{(x_1, x_2, \dots, x_k) \in \mathbb{R}^k} |f(x_1, x_2, \dots, x_k)| \leq 1.$$

Define the Lipschitz coefficients on $\mathcal{M}^k \ni f \mapsto \mathbb{L}ip f \in \mathbb{R}^+ \cup \{+\infty\}$ as

$$\mathbb{L}ip f = \sup_{(y_1, \dots, y_k) \neq (x_1, \dots, x_k)} \frac{|f(y_1, \dots, y_k) - f(x_1, \dots, x_k)|}{|y_1 - x_1| + \dots + |y_k - x_k|}.$$

We also set $\mathcal{M} = \cup_{k \in \mathbb{N}} \mathcal{M}^k$. According to Doukhan and Louhichi (1999) and Dedecker *et al.* (2007), a time series $\{X_t\}$ is said to be $(\epsilon, \mathcal{M}, \psi)$ -weakly dependent, (or simply ψ -weakly dependent), if there exists a mapping

$$\psi_\epsilon : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}^+$$

such that

$$\epsilon(r) = \sup \frac{|\text{Cov}(f(X_{i_1}, \dots, X_{i_k}), g(X_{j_1}, \dots, X_{j_\ell}))|}{\psi_\epsilon(f, g)} \rightarrow 0 \quad \text{as } r \rightarrow \infty, \quad (2.1)$$

where the supremum is taken over all pairs of functions $f \in \mathcal{M}^k$, $g \in \mathcal{M}^\ell$ and sets of indices

$$i_1 \leq i_2 \leq \dots \leq i_k \leq j_1 \leq j_2 \leq \dots \leq j_\ell$$

with a gap of size r ; $j_1 - i_k \geq r$. By selecting a mapping ψ_ϵ we obtain the following various coefficients of dependence:

$$\begin{aligned} \psi_\alpha(f, g) &= 4, & \epsilon(r) &= \alpha(r) & (\alpha - \text{mixing}) \\ \psi_\theta(f, g) &= \ell \mathbb{L}ip g, & \epsilon(r) &= \theta(r) & (\theta - \text{dependence}) \\ \psi_\eta(f, g) &= k \mathbb{L}ip f + \ell \mathbb{L}ip g, & \epsilon(r) &= \eta(r) & (\eta - \text{dependence}) \\ \psi_\kappa(f, g) &= k\ell \mathbb{L}ip f \cdot \mathbb{L}ip g, & \epsilon(r) &= \kappa(r) & (\kappa - \text{dependence}) \\ \psi_\lambda(f, g) &= k \mathbb{L}ip f + \ell \mathbb{L}ip g + k\ell \mathbb{L}ip f \cdot \mathbb{L}ip g, & \epsilon(r) &= \lambda(r) & (\lambda - \text{dependence}). \end{aligned} \quad (2.2)$$

In this work we are interested in change of tail behavior of a stationary weakly dependent process $\{X_t\}$ satisfying (2.1) and (2.2). Let $F(x) = P(X_1 \leq x)$ be the marginal distribution function of the stationary process $\{X_t\}$. Suppose that the cumulative distribution function (cdf) F has tail index m with satisfying

$$\lim_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{m \log x} = 1, \quad (2.3)$$

for some $m > 0$. An exponentially tailed distribution F corresponds to $m = \infty$. For $0 < m < \infty$, F is heavy-tailed or Pareto type. We adopt the idea of Jureckova and Picek (2001), instead of Hill's estimate approach, and propose a simpler test for the change on the tail index. Let m_0 be a threshold

tail index of interest in hypothesis. We wish to test. If the tail index m remains in the same interval $(0, m_0]$ for the whole observations. The null hypothesis is $H_0(= H_{m_0})$: “no change in the tail behavior of the stationary weakly dependent process with tail index $m \in (0, m_0]$, i.e., $x^{m_0}(1 - F(x)) \geq 1$ for sufficiently large x ”. Under the null hypothesis H_0 , assuming that $\{X_t\}$ is stationary, F satisfies (2.3) with unchanged tail index m , vs. under the alternative hypothesis H_A , the tail index m of F changes from a value in $(0, m_0]$ to one in (m_0, ∞) with $\limsup_{x \rightarrow \infty} x^{m_0}(1 - F(x)) < 1$.

In order to construct the test statistics, we apply the empirical distribution functions of the extremes, based on splitting the set of observations into N subsamples of size n , described as in Jureckova and Picek (2001). Suppose that $X_1, X_2, \dots, X_{\bar{N}}$ are observed with sample size $\bar{N} = nN$ so that these observations are divided into N groups of each size n . We assume that $n \rightarrow \infty$, $N \rightarrow \infty$ and $n = o(N)$ as $\bar{N} \rightarrow \infty$. For example, we set $n = [\log \bar{N}]$ and $N = [\bar{N} / \log \bar{N}]$ where $[x]$ is the integer part of x . For $s = 1, 2, \dots, N$, the s -th group consists of $\{X_{(s-1)n+1}, \dots, X_{sn}\}$, and its maximum is denoted by $X_{(n),s}$: $X_{(n),s} = \max\{X_{(s-1)n+i} : 1 \leq i \leq n\}$. The empirical distribution function $\hat{F}_{N,n}$ of the sample maxima $\{X_{(n),s} : s = 1, 2, \dots, N\}$ is given by

$$\hat{F}_{N,n}(x) = \frac{1}{N} \sum_{s=1}^N \mathbb{I}(X_{(n),s} \leq x), \quad (2.4)$$

where $\mathbb{I}(\cdot)$ is the indicator function. For the test statistics we choose a threshold sequence $a = a_{\bar{N}, m_0}$, depending on the sample size and the hypothesis index m_0 . For the case of iid random variables, Jureckova and Picek (2001) have selected $a := (nN^{1-\delta})^{1/m_0}$ for $0 < \delta < 2/3$ as the threshold sequence so that $P(X_{\min} > a) \rightarrow 0$ and $P(X_{\max} < a) \rightarrow 0$ as $\bar{N} \rightarrow \infty$ and thus

$$\lim_{\bar{N} \rightarrow \infty} P(0 < \hat{F}_{N,n}(a) < 1) = 1. \quad (2.5)$$

Here $X_{\min} = \min\{X_t : 1 \leq t \leq nN\}$ and $X_{\max} = \max\{X_t : 1 \leq t \leq nN\}$. In this work of weakly dependent time series case, we choose the same sequence $a := (nN^{1-\delta})^{1/m_0}$. The reason will be explained in Remark 1 below, which is due to the existence of the phantom distribution function of the stationary weakly dependent processes (see also Proposition 1 in Section 6).

A class of CUSUM test statistics $\{\hat{T}_N(\cdot) \equiv \hat{T}_{N, m_0}(\cdot) : m_0 > 0\}$ is constructed based on the empirical distribution functions as follows: For $0 \leq z \leq 1$ and $a = (nN^{1-\delta})^{1/m_0}$ for some $0 < \delta < 2/3$, let

$$\hat{T}_N(z) = \frac{1}{\sqrt{N}} \sum_{s=1}^{[Nz]} \mathbb{I}(X_{(n),s} \leq a) - \frac{z}{\sqrt{N}} \sum_{s=1}^N \mathbb{I}(X_{(n),s} \leq a).$$

We will reject the null hypothesis H_0 if $\sup_{0 \leq z \leq 1} |\hat{T}_N(z)|$ is large. Some simulated results for the statistic $\sup_{0 \leq z \leq 1} |\hat{T}_N(z)|$ are given for GARCH(1, 1) models with two changed noises on Table 1 in Section 4. GARCH(1, 1) processes are discussed in the next section. Note that our class of tests can be used for the null hypothesis H'_0 that tail index $m \in [m_0, \infty)$ is unchanged with $x^{m_0}(1 - F(x)) \leq 1$ for large x , vs. the alternative hypothesis H'_A that there exists tail change from $m \in [m_0, \infty)$ to $m \in (0, m_0)$ with $\limsup_{x \rightarrow \infty} x^{m_0}(1 - F(x)) > 1$, because the statistics account for the supremum of the absolute value over $0 \leq z \leq 1$.

A functional central limit theorem (FCLT) establishes the convergence in distribution:

$$\hat{T}_N(z) \xrightarrow{d} \sigma(B(z) - zB(1)) \equiv \sigma B_0(z) \quad \text{as } N \rightarrow \infty,$$

where $\{B_0(z) := B(z) - zB(1), 0 \leq z \leq 1\}$ is the Brownian bridge with the standard Brownian motion $B(\cdot)$ and

$$\sigma^2 = \lim_{N \rightarrow \infty} \text{Var} \left[\frac{1}{\sqrt{N}} \sum_{s=1}^N \mathbb{I}(X_{(n),s} \leq a) \right]. \quad (2.6)$$

We may use $\hat{\sigma}_{N,q}^2$ as an estimator of σ^2 which is defined by

$$\hat{\sigma}_{N,q}^2 = \frac{1}{N} \sum_{s=1}^N \left(\mathbb{I}(X_{(n),s} \leq a) - \hat{F}_{N,n}(a) \right)^2 + 2 \sum_{s=1}^q \omega_{s,q} \hat{\gamma}_{N,s}(a), \quad q < N, \quad (2.7)$$

where $\omega_{s,q} = 1 - s/(q+1)$ and autocovariance estimator

$$\hat{\gamma}_{N,s}(a) = \frac{1}{N} \sum_{i=1}^{N-s} \left(\mathbb{I}(X_{(n),i} \leq a) - \hat{F}_{N,n}(a) \right) \left(\mathbb{I}(X_{(n),i+s} \leq a) - \hat{F}_{N,n}(a) \right)$$

for $0 \leq s \leq N$ (Lo, 1991, Equation (3.6)). Define the CUSUM test statistic $\hat{B}_N(z)$ by

$$\hat{B}_N(z) := \frac{1}{\hat{\sigma}_{N,q}} \hat{T}_N(z).$$

Our main result for the asymptotic of $\hat{B}_N(z)$ is stated in Theorem 1, for which we make the following assumptions under the null hypothesis:

(A1) $\{X_t\}$ is stationary and satisfies the weakly dependent conditions in (2.1) and (2.2).

(A2) $\{X_t\}$ has continuous marginals satisfying

$$P(X_1 \in (x, x+u]) = F(x+u) - F(x) \leq cu^b, \quad x \in \mathbb{R}, u > 0, \quad (2.8)$$

for some $b > 0$ and $c > 0$.

The concentration assumption in (A2) is not restrictive as mentioned by Doukhan *et al.* (2015) (see Section 4.3). By their Theorems 7–9, under conditions (A1)–(A2), $\{X_t\}$ admits a continuous phantom distribution function, that is, there exists $G(\cdot)$ such that

$$P\left(\max_{1 \leq t \leq n} X_t \leq u_n\right) - G^n(u_n) \rightarrow 0$$

as $n \rightarrow \infty$ for every sequence $\{u_n\} \subset \mathbb{R}$. Thanks to the existence of the phantom distribution function, the threshold sequence $a = (nN^{1-\delta})^{1/m_0}$ satisfying (2.5) can be selected as well as the following theorem is valid. More details on significant roles of the phantom distribution function are explored in Remark 1.

Theorem 1. *We assume (A1) and (A2). Under the null hypothesis H_0 , if the weak dependent coefficient in (2.1) and (2.2) has order $\epsilon(r) = O(r^{-\beta})$ for some $\beta > 0$, then we have, as $N \rightarrow \infty$ and $n \rightarrow \infty$ with $n = o(N)$,*

$$\sup_{0 \leq z \leq 1} |\hat{B}_N(z)| \xrightarrow{d} \sup_{0 \leq z \leq 1} |B_0(z)|,$$

where $\{B_0(z) := B(z) - zB(1), 0 \leq z \leq 1\}$ is the Brownian bridge with the standard Brownian motion $B(\cdot)$.

Its proof is drawn in Section 6. Now the following remark explains that the arguments of Jureckova and Picek (2001) can be applied in the case of weakly dependent sequences due to the existence of the phantom distribution function.

Remark 1. As mentioned above, in the case of iid random variables with cdf $F(x)$, the threshold sequence $a = (nN^{1-\delta})^{1/m_0}$ for $0 < \delta < 2/3$ has been chosen by Jureckova and Picek (2001) so that (2.5) holds in hypothesis $x^{m_0}(1 - F(x)) \geq 1$ for sufficiently large x . In this present work we adopt the same threshold so that (2.5) holds for the following reason: Recall that the empirical version of $P(X_{(n),s} \leq x)$ is given by $\hat{F}_{N,n}(x)$ in (2.4). If random variables are iid then $P(X_{(n),s} \leq x) = F(x)^n$, while if $\{X_t\}$ is a stationary and weakly-dependent sequence in this work, the equality does not hold. However, we can approximate using a phantom distribution function. By Proposition 1 (see Doukhan *et al.*, 2015), a continuous phantom distribution function $G(\cdot)$ exists such that

$$P(X_{(n),s} \leq u_n) - G^n(u_n) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

for every sequence $\{u_n\}$ and thus we have approximation $\hat{F}_{N,n}(u_n) - G(u_n)^n \rightarrow 0$ as $n \rightarrow \infty$, and we may say that $\hat{F}_{N,n}(x)$ is also the empirical version of $G(x)^n$. For just simplicity, we consider $M_n := X_{(n),1}$ and for fixed x , write

$$P(M_n \leq x) = P(X_i \leq x \text{ for all } i = 1, 2, \dots, n) = \prod_{j=1}^n P(A_j | D_{j-1}) =: \prod_{j=1}^n g_j(x),$$

where $A_j = \{X_j \leq x\}$, $D_{j-1} = \{X_i \leq x, i = 1, 2, \dots, j-1\}$ for $j = 1, 2, \dots, n$ with D_0 sample space, and $g_j(x) = P(A_j | D_{j-1})$. For a given sequence u_n , we have $\prod_{j=1}^n g_j(u_n) - G(u_n)^n \rightarrow 0$, $G(u_n) - \prod_{j=1}^n g_j(u_n)^{1/n} \rightarrow 0$ as $n \rightarrow \infty$. Thus we may regard $G(x)$ as asymptotically geometric mean of conditional cdfs $\{g_j(x) : j = 1, 2, \dots, n\}$, and thus (2.3) holds for G replaced in F and we have $x^{m_0}(1 - G(x)) \geq 1$ for sufficiently large x . Therefore, by the same argument as in Jureckova and Picek (2001) replaced by G , which plays a role as F asymptotically on pp. 177–178, in the proof of their Theorem 2.1, (2.5) with the selected threshold $a = (nN^{1-\delta})^{1/m_0}$ follows from the fact that $\hat{F}_{N,n}(x)$ is approximated by $G(x)^n$ for sufficiently large x as $n \rightarrow \infty$.

As a specific model of the ψ -weakly dependent process, we deal with a GARCH process and a simulation study is given in the next two sections.

3. GARCH(1, 1) model

In this section, we consider a GARCH(1, 1) process to present a simulation study for the proposed CUSUM tests. The GARCH process $\{X_t : t \in \mathbb{Z}\}$ is given by

$$X_t = \sqrt{h_t} \varepsilon_t, \quad h_t = \omega + \alpha X_{t-1}^2 + \beta h_{t-1}, \quad (3.1)$$

where $\{\varepsilon_t : t \in \mathbb{Z}\}$ are iid random variables with zero mean and unit variance, $\omega > 0$, $\alpha \geq 0$, and $\beta \geq 0$. Here a sufficient condition is given for the asymptotic result of the CUSUM statistics in the case of GARCH(1, 1) processes: On behalf of the stationarity of GARCH(1, 1) model, we assume $\alpha + \beta < 1$; in addition, the following are added for our main result of Theorem 1 in the GARCH(1, 1) processes:

(A1)' The GARCH parameters satisfy $E \log(\beta + \alpha \varepsilon_1^2) < 0$;

(A2)' The noise process $\{\varepsilon_t : t \in \mathbb{Z}\}$ has the absolutely continuous distribution and its cdf has a bounded density.

Condition (A1)' is for the strict stationarity of the GARCH process under the null hypothesis. The strict stationarity along with the bounded density in (A2)' imply the concentration assumption in (A2). Two conditions (A1)' and (A2)' imply conditions (A1) and (A2) as well as ensure the existence of the phantom distribution function for the GARCH process. We discuss these details as follows. Assume that $h_0 = 1$. Straightforwardly we have the identity:

$$h_t = \omega \cdot \left(1 + \sum_{i=1}^{t-1} \sum_{j=1}^{t-i} (\beta + \alpha \varepsilon_{t-j}^2) \right) + \prod_{i=0}^{t-1} (\beta + \alpha \varepsilon_i^2)$$

and the GARCH(1, 1) model (3.1) can be expressed as $X_t = H(\varepsilon_t, \varepsilon_{t-1}, \dots)$ for some function H .

This implies that the GARCH(1, 1) process is a Bernoulli shift with innovation $\{\varepsilon_t : t \in \mathbb{Z}\}$. For the function H , we can express the following weak dependence property: The sequence $\{d_r : r \in \mathbb{Z}\}$ defined by

$$d_r := E \left| H(\varepsilon_{t-j} : j \in \mathbb{Z}) - H(\varepsilon_{t-j} \mathbb{I}\{|j| \leq r\} : j \in \mathbb{Z}) \right|$$

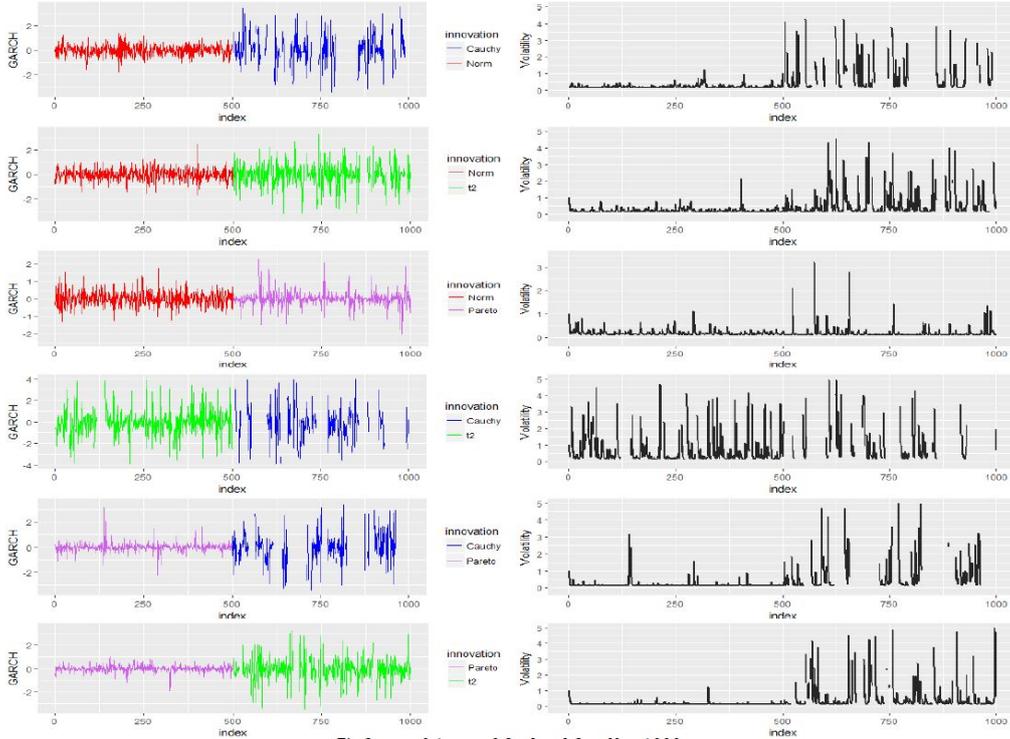
converges to zero as r tends to infinity. According to Lemma 3.1, p. 25, of Dedecker *et al.* (2007), (see Ango Nze and Doukhan (2004)), it follows that the GARCH process is η -weakly dependent with the weak dependence coefficient $\eta(r) \leq 2d_{\lfloor r/2 \rfloor}$. Thus the stationary GARCH(1, 1) process satisfies condition (A1). Also, because it is a Bernoulli shift with iid innovation process $\{\varepsilon_t\}$ which has a bounded density, say f , under conditions (A1)'–(A2)', then concentration assumption (A2) holds with $b = 1$ and $c = \sup_x f(x)$ in (2.8). In conclusion, under conditions (A1)'–(A2)', the GARCH(1, 1) process admits a continuous phantom distribution and Theorem 1 holds for the GARCH model.

4. A Monte-Carlo study

Now we present simulations for the proposed CUSUM tests to see the size and power performances. We consider GARCH(1, 1) models (3.1) with several noises $\{\varepsilon_t\}$. Some pairs of normal, $t(1)$, $t(2)$, Cauchy, and (two-sided) Pareto distributions are employed for the noises in GARCH(1, 1) models. The GARCH (1, 1) models and their volatilities with $h_t = 0.1 + 0.3X_{t-1}^2 + 0.2h_{t-1}$ of size $Nn = 1000$ are generated and depicted in Figure 1, with two changed noise distributions for $t = 1, 2, \dots, 1000$ where noises are changed at time $t = 501$. The left plots of Figure 1 depict the GARCH models and the right depict the volatilities. We see that there exist changes in distributions at the center.

Table 1 gives the numerical illustration for values of the test statistics $\sup_{0 \leq z \leq 1} |\hat{T}_N(z)|$ of simulated sample of the GARCH (1, 1) model with $N = 100, n = 10, \delta = 0.25$ and with $m_0 = 1.00, 1.25, 1.50, \dots, 2.75, 3.00$ and compares those in the GARCH(1, 1) models with two cases of the absence/presence of noise changes. Table 1, the values of $\sup_{0 \leq z \leq 1} |\hat{T}_N(z)|$ are small in most distributions with no changes, whereas those are somewhat larger in cases with two changed innovations.

Finite sample size and power properties of the proposed CUSUM tests $\sup_{0 \leq z \leq 1} |\hat{B}_N(z)|$ in Theorem 1 are investigated in stationary GARCH(1, 1) models with $h_t = 0.1 + 0.05X_{t-1}^2 + 0.9h_{t-1}$ for nominal levels 5%; sample size 800, 1200; $N = 200; n = 4, 6; \delta = 0.25, 0.50; 1000$ replications, as seen in Tables 2–3, respectively, where we see that reasonable sizes and powers are reported.

Figure 1: $GARCH(1, 1)$ models and their volatilities with two changed innovations.Table 1: Simulated test statistics $\sup_{0 \leq z \leq 1} |\hat{T}_N(z)|$ with $h_t = 0.1 + 0.3X_{t-1}^2 + 0.2h_{t-1}$

m_0	Normal	$t(1)$	$t(2)$	Cauchy	Pareto	$N + t(1)$	$N + C$	$N + P$	$t(2) + C$	$t(2) + P$
1.00	0.0998	0.4596	0.0981	1.4273	0.0	1.4669	2.9695	1.7541	1.4378	1.3734
1.25	0.0999	0.1745	0.1945	1.2267	0.0	1.8400	2.3199	1.6881	1.7785	1.0353
1.50	0.0996	0.4235	0.0998	0.8843	0.0	1.5763	2.4109	2.0522	1.6138	1.2572
1.75	0.0995	0.6771	0.1866	1.0368	0.0	1.9319	2.6258	1.4158	1.3595	1.4746
2.00	0.0998	0.4112	0.2407	1.0507	0.0	1.9406	3.1334	1.8647	1.8917	1.5973
2.25	0.0991	0.1749	0.3150	0.7745	0.0	2.1645	3.7539	1.8528	1.7461	1.1337
2.50	0.0997	0.3782	0.2420	0.7692	0.0	1.9278	2.9041	1.6773	2.0532	1.8058
2.75	0.0994	0.3857	0.1525	0.3469	0.0	1.9457	3.3271	1.7049	1.6039	0.3581
3.00	0.0993	0.0574	0.3995	0.8770	0.0	2.0149	3.3641	1.9841	1.7784	1.1438

$N = 100, n = 10, \delta = 0.25; N = \text{Normal}, C = \text{Cauchy}, P = \text{Pareto}.$

As a real data application, the proposed CUSUM tests are applied to log-return sets of the Korean stock price index (KOSPI). We adopt financial data from KOSPI-C (Construction) and KOSPI-F (Finance) between 08-02-2013 and 08-27-2017. Figure 2 provides the plots of the financial real data and their standardized log-returns.

For $\delta = 0.25, 0.5$, the CUSUM test statistics $\sup_{0 \leq z \leq 1} \hat{B}_N(z)$ of KOSPI-C and KOSPI-F are reported in Table 4 with $m_0 = 2.00, 2.25, 2.50, \dots, 4.25, 4.50$. Note that the critical values of the standard Brownian bridge for 1% and 5% are 1.5174 and 1.2238, respectively. Using these values, as seen in Table 4 with $\delta = 0.25$, we reject the null hypotheses for KOSPI-C for $m_0 \in \{2.00, 2.25, \dots, 3.25\}$ for both cases $n = 5, 10$ and for both 1%, 5% level tests; therefore, we can conclude that there exist

Table 2: Size of 5% level tests in GARCH(1, 1) model with $h_t = 0.1 + 0.05X_{t-1}^2 + 0.9h_{t-1}$

Noise distribution	$N * n$	n	δ	m_0								
				1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
$t(1)$	800	4	0.25	0.040	0.043	0.052	0.059	0.046	0.051	0.055	0.050	0.052
			0.50	0.053	0.046	0.046	0.056	0.047	0.047	0.047	0.054	0.050
	1200	6	0.25	0.012	0.009	0.016	0.010	0.008	0.008	0.021	0.011	0.011
			0.50	0.007	0.012	0.016	0.010	0.011	0.013	0.010	0.011	0.014
$t(2)$	800	4	0.25	0.049	0.066	0.061	0.066	0.065	0.064	0.055	0.074	0.056
			0.50	0.052	0.045	0.064	0.066	0.052	0.063	0.071	0.060	0.062
	1200	6	0.25	0.080	0.079	0.082	0.076	0.077	0.081	0.077	0.090	0.086
			0.50	0.084	0.080	0.085	0.083	0.086	0.083	0.078	0.074	0.078
Cauchy	800	4	0.25	0.064	0.066	0.076	0.059	0.073	0.057	0.057	0.061	0.064
			0.50	0.076	0.062	0.065	0.065	0.055	0.067	0.072	0.073	0.076
	1200	6	0.25	0.083	0.094	0.076	0.068	0.092	0.070	0.081	0.092	0.082
			0.50	0.082	0.082	0.073	0.084	0.081	0.071	0.085	0.097	0.080
Pareto	800	4	0.25	0.081	0.050	0.091	0.098	0.081	0.057	0.054	0.064	0.057
			0.50	0.065	0.040	0.046	0.080	0.071	0.054	0.068	0.072	0.057
	1200	6	0.25	0.067	0.076	0.079	0.056	0.075	0.064	0.092	0.047	0.112
			0.50	0.113	0.092	0.117	0.068	0.077	0.085	0.085	0.100	0.093

sample size = 800, 1200; $N = 200$; $n = 4, 6$; $\delta = 0.25, 0.5$; number of replications = 1000.

Table 3: Power of 5% level tests in GARCH(1,1) model with $h_t = 0.1 + 0.05X_{t-1}^2 + 0.9h_{t-1}$

Noise distribution	$N * n$	n	δ	m_0								
				1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
$t(2)+$ Pareto	800	4	0.25	1.000	1.000	0.997	0.989	0.984	0.979	0.968	0.932	0.916
			0.50	0.999	0.994	0.984	0.974	0.957	0.929	0.916	0.897	0.882
	1200	6	0.25	1.000	0.995	0.997	0.997	0.974	0.962	0.938	0.935	0.912
			0.50	0.998	0.993	0.986	0.963	0.952	0.933	0.890	0.866	0.839
$t(2)+$ Cauchy	800	4	0.25	1.000	0.996	0.997	0.994	0.981	0.990	0.952	0.955	0.925
			0.50	0.999	0.993	0.984	0.975	0.957	0.929	0.917	0.897	0.882
	1200	6	0.25	1.000	1.000	0.998	0.993	0.986	0.973	0.958	0.937	0.914
			0.50	1.000	0.987	0.970	0.985	0.943	0.925	0.893	0.867	0.833
Cauchy + Pareto	800	4	0.25	0.852	0.759	0.664	0.600	0.508	0.455	0.432	0.402	0.351
			0.50	0.724	0.633	0.546	0.450	0.429	0.371	0.324	0.275	0.289
	1200	6	0.25	0.893	0.785	0.704	0.592	0.555	0.487	0.489	0.399	0.372
			0.50	0.774	0.669	0.590	0.496	0.425	0.401	0.331	0.296	0.260
Normal + Pareto	800	4	0.25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			0.50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1200	6	0.25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			0.50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

sample size = 800, 1200; $N = 200$; $n = 4, 6$; $\delta = 0.25, 0.5$; number of replications = 1000.

changes of tail behaviors of log-returns of KOSPI-C on those indices m_0 . Also we reject the null hypotheses for KOSPI-F for $m_0 \in \{2.00, 2.25, \dots, 4.00\}$. Thus there exists changes of tail index of log-returns of KOSPI-F from a value in $(0, 4]$ to that in $(4, \infty)$ for both cases $n = 5, 10$ and for both 1%, 5% level tests. However, if $\delta = 0.5$ then different results are reported as $n = 5$ and as $n = 10$. Therefore we need to refine the threshold $a = (nN^{1-\delta})^{1/m_0}$ by finding optimal ratio of n to N or optimal δ . This discussion will remain for a further study along with the estimation problem.

5. Concluding remarks

We have proposed a class of CUSUM tests for tail behavior of the weakly-dependent processes. CUSUM test statistics are constructed using empirical distribution functions with sample extremes

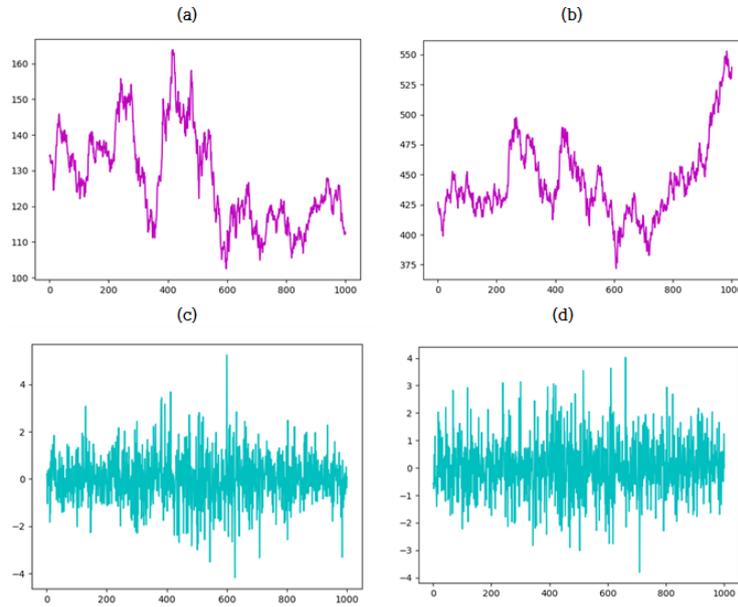


Figure 2: Financial data of KOSPI and their log-returns. (a), (c): KOSPI-Construction data and log-returns; (b), (d): KOSPI-Finance data and log-returns.

Table 4: Computed test statistics $\sup_{0 \leq z \leq 1} |\hat{B}_N(z)|$ for (standardized) log-returns of KOSPI-Construction and KOSPI-Finance

m_0	$\delta = 0.25$				$\delta = 0.5$			
	KOSPI-C		KOSPI-F		KOSPI-C		KOSPI-F	
	$n = 5$	$n = 10$	$n = 5$	$n = 10$	$n = 5$	$n = 10$	$n = 5$	$n = 10$
2.00	2.2357	3.1586	2.2340	3.1590	2.2309	3.1622	2.2356	3.1616
2.25	2.2356	3.1600	2.2356	3.1599	2.2321	3.1572	2.2325	3.1570
2.50	2.2318	3.1586	2.2343	3.1588	2.2318	3.1576	2.2355	3.1617
2.75	2.2347	3.1617	2.2347	3.1605	0.2934	3.1604	2.2350	3.1621
3.00	2.2341	3.1571	2.2355	3.1613	0.3034	0.4094	2.2347	3.1615
3.25	2.2339	3.1601	2.2353	3.1622	0.3017	0.4112	0.2617	3.1568
3.50	0.3057	0.3838	2.2301	3.1618	0.2296	0.4096	0.1799	0.3464
3.75	0.3014	0.3970	2.2327	3.1603	0.1862	0.3057	0.2195	0.2378
4.00	0.2986	0.4042	2.2344	3.1597	0.1961	0.2620	0.1909	0.2320
4.25	0.2956	0.4044	0.2627	0.3609	0.2034	0.2673	0.1827	0.2809
4.50	0.2758	0.3505	0.1774	0.2523	0.1821	0.2755	0.1824	0.2528

sample size = 1000; $n = 5, 10$; $\delta = 0.25, 0.5$.

and its limiting distribution is derived to be a Brownian bridge. Powers and sizes are computed to see the performance of CUSUM tests via simulations. Also real data analysis is given. From the analysis of our CUSUM tests it might be possible to estimate the tail index of the weakly dependent time series models. The estimation needs refinements of the test statistics, to find the optimal m_0 , along with optimal parameters such as δ , q and the optimal ratio of n to N . Future studies should investigate this estimation problem by means of CUSUM test statistics and compare with the existing tail-index estimates such as the Hill estimator. Furthermore, some empirical applications can be obtained when the CUSUM test and estimation are conducted on the tail index of a weakly dependent return data

set. For instance, using a change point detection based on CUSUM tests, both the tail index itself and its change point are estimated; subsequently, the value-at-risk (VaR) or volatility forecasting are then computed. One way to calculate VaR or volatility forecasting in the presence of tail change is to use different weights on periods of changed tail behavior like exponential smoothing. Such methodology on empirical applications will be discussed in a further study.

6. Proof

In order to prove Theorem 1 we need to state the following proposition that is related to the existence of the phantom distribution function of the weakly dependent process with conditions (A1)–(A2) under the null hypothesis. In case of α -mixing, according to Theorem 6 of Doukhan *et al.* (2015), p. 711, the process admits a continuous phantom distribution function. For other cases, their Theorems 7–9 are summarized by the following proposition under conditions (A1)–(A2).

Proposition 1. (Doukhan *et al.*, (2015), Theorems 7–9 and Remark 10 on p. 715) *Let $\{X_t\}$ be a stationary sequence with continuous marginals satisfying (A1) and (A2). For $\psi \in \{\psi_\theta, \psi_\eta, \psi_\kappa, \psi_\lambda\}$ with the corresponding $\epsilon \in \{\theta, \eta, \kappa, \lambda\}$, if $\{X_t\}$ is $(\epsilon, \mathcal{M}, \psi)$ -weakly dependent and fulfills $\epsilon(r) = O(r^{-\beta})$ for some $\beta > 0$, then it admits a continuous phantom distribution function of $G(\cdot)$ such that*

$$P\left(\max_{1 \leq t \leq n} X_t \leq u_n\right) - G^n(u_n) \rightarrow 0$$

as $n \rightarrow \infty$ for every sequence $\{u_n\} \subset \mathbb{R}$.

Proof of Theorem 1:: By the consistency of the estimator in (2.7) it suffices to show that for $0 \leq z \leq 1$, $\hat{T}_N(z) \xrightarrow{d} \sigma B_0(z)$ as $N \rightarrow \infty$ and $n \rightarrow \infty$. For $a = (nN^{1-\delta})^{1/m_0}$, recall $\hat{F}_{N,n}(a) = (1/N) \sum_{s=1}^N \mathbb{I}(X_{(n),s} \leq a)$ and note that $E[\hat{F}_{N,n}(a)] = P(X_{(n),1} \leq a)$ and by (2.6)

$$\lim_{N \rightarrow \infty} N \text{Var}[\hat{F}_{N,n}(a)] = \lim_{N \rightarrow \infty} \text{Var}[\sqrt{N} \hat{F}_{N,n}(a)] = \sigma^2.$$

According to Proposition 1, there exists a function $G(\cdot)$ such that for each $s = 1, 2, \dots, N$, $P(X_{(n),s} \leq a) - G^n(a) \rightarrow 0$ as $n \rightarrow \infty$. The stationarity of the process under the null H_0 ensures that for each s , the phantom distribution functions cannot be too different and must be tail equivalent. Thus we may adopt one equivalent phantom distribution function as G such that the cdf of $X_{(n),s}$ is approximated by $G^n(\cdot)$. Therefore, we see that $\{X_t\}$ has asymptotic independence of maxima (AIM) relative to the sequence $\{a = a_N\}$, (O'Brien, 1987) of $s_1 \neq s_2$, as $n \rightarrow \infty$, $P(X_{(n),s_1} \leq a, X_{(n),s_2} \leq a) - P(X_{(n),s_1} \leq a)P(X_{(n),s_2} \leq a) = G^{2n}(a) - G^n(a)G^n(a) + o(1) \rightarrow 0$, where it is used that two sets $\{X_{(n),s_1} \leq a, X_{(n),s_2} \leq a\}$ and $\{X_{(2n),s_1,s_2} \leq a\}$ are equivalent, with $X_{(2n),s_1,s_2} = \max\{X_{(s_1-1)n+1}, X_{(s_1-1)n+2}, \dots, X_{s_1n}, X_{(s_2-1)n+1}, X_{(s_2-1)n+2}, \dots, X_{s_2n}\}$, and thus $P(X_{(n),s_1} \leq a, X_{(n),s_2} \leq a) - G^{2n}(a) = o(1)$. See Theorem 3.1 and its proof of Jakubowski (1991) for a similar arguments.

Let $U_{(n),s} = \mathbb{I}(X_{(n),s} \leq a) - G^n(a)$ for $s = 1, 2, \dots, N$. Then $\text{Cov}(U_{(n),s_1}, U_{(n),s_2}) \rightarrow 0$ for $s_1 \neq s_2$, and thus by the CLT of weakly dependent sequence it follows that as $N \rightarrow \infty$ and $n \rightarrow \infty$,

$$\frac{1}{\sqrt{N}} \sum_{s=1}^N U_{(n),s} \xrightarrow{d} N(0, \sigma^2),$$

noting that its asymptotic mean and asymptotic variance are zero and σ^2 , respectively. Also, note that $\hat{F}_{N,n}(\cdot)$ is a consistent estimator of $G^n(\cdot)$ with satisfying $\sqrt{N}[\hat{F}_{N,n}(a) - G^n(a)] \xrightarrow{d} N(0, \sigma^2)$ as $N \rightarrow \infty$

and $n \rightarrow \infty$. Thus under the null H_0 we may use the FCLT to show that

$$\frac{1}{\sqrt{N}} \left(\sum_{s=1}^{\lfloor Nz \rfloor} U_{(n),s} - \frac{\lfloor Nz \rfloor}{N} \sum_{s=1}^N U_{(n),s} \right) \xrightarrow{d} \sigma(B(z) - zB(1)) = \sigma B_0(z)$$

for $0 \leq z \leq 1$. The left hand side is

$$\frac{1}{\sqrt{N}} \left(\sum_{s=1}^{\lfloor Nz \rfloor} \{\mathbb{I}(X_{(n),s} \leq a) - G^n(a)\} - \frac{\lfloor Nz \rfloor}{N} \sum_{s=1}^N \{\mathbb{I}(X_{(n),s} \leq a) - G^n(a)\} \right) \equiv \hat{T}_N(z).$$

Therefore the desired result is obtained. \square

Acknowledgements

This work was supported by Research Fund of Gachon University (GCU-2018-0337).

References

- Ango Nze P, Bühlmann P, and Doukhan P (2002). Nonparametric regression estimation under weak dependence beyond mixing and association, *Annals of Statistics*, **30**, 397–430.
- Ango Nze PA and Doukhan P (2004). Weak dependence: models and applications to econometrics, *Econometric Theory*, **20**, 995–1045.
- Coulon-Prieur C and Doukhan P (2000). A triangular central limit theorem under a new weak dependent condition, *Statistics and Probability Letters*, **47**, 61–68.
- Dedecker J, Doukhan P, Lang G, and León JR, Louhichi S, and Prieur C (2007) Weak Dependence: With Examples and Applications. In *Lecture Notes in Statistics* (Vol 190), Springer, New York.
- Doukhan P, Jakubowski A, and Lang G (2015). Phantom distribution functions for some stationary sequences, *Extremes*, **18**, 697–725.
- Doukhan P and Louhichi S (1999). A new weak dependence condition and applications to moment inequalities, *Stochastic Processes and Their Applications*, **84**, 313–342.
- Doukhan P and Louhichi S (2001). Functional estimation of a density under a new weak dependence condition, *Scandinavian Journal of Statistics*, **28**, 325–341.
- Doukhan P and Neumann MH (2007). Probability and moment inequalities for sums of weakly dependent random variables with applications, *Stochastic Processes and Their Applications*, **117**, 878–903.
- Hill B (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, **3**, 1163–1174.
- Hoga Y (2017). Change point tests for the tail index of β -mixing random variables, *Econometric Theory*, **33**, 915–954.
- Hwang E and Shin DW (2011). Semiparametric estimation for partially linear models with ψ -weak dependent errors, *Journal of the Korean Statistical Society*, **40**, 411–424.
- Hwang E and Shin DW (2012a). Stationary bootstrap for kernel density estimators under ψ -weak dependence, *Computational Statistics and Data Analysis*, **56**, 1581–1593.
- Hwang E and Shin DW (2012b). Random central limit theorems for linear processes with weakly dependent innovations, *Journal of the Korean Statistical Society*, **41**, 313–322.
- Hwang E and Shin DW (2013). A study on moment inequalities under a weak dependence, *Journal of the Korean Statistical Society*, **42**, 133–141.

- Hwang E and Shin DW (2016a). Maximal inequalities and an application under a weak dependence, *Journal of Korean Mathematical Society*, **53**, 57–72.
- Hwang E and Shin DW (2016b). Kernel estimators of mode under ψ -weak dependence, *Annals of the Institute of Statistical Mathematics*, **68**, 301–327.
- Jakubowski A (1991). Relative extremal index of two stationary processes, *Stochastic Processes and their Applications*, **37**, 281–297.
- Jureckova J, Koul HL, and Picek J (2009). Testing the tail index in autoregressive models, *Annals of the Institute of Statistical Mathematics*, **61**, 579–598.
- Jureckova J and Picek J (2001). A class of tests on the tail index, *Extremes*, **4**, 165–183.
- Kim M and Lee S (2009). Test for tail index change in stationary time series with Pareto-type marginal distribution, *Bernoulli*, **15**, 325–356.
- Kim M and Lee S (2011). Change point test for tail index for dependent data, *Metrika*, **74**, 297–311.
- Kim M and Lee S (2012). Change point test of tail index for autoregressive processes, *Journal of the Korean Statistical Society*, **41**, 305–312.
- Kim M and Lee S (2016). On the tail index inference for heavy-tailed GARCH-type innovations, *Annals of the Institute of Statistical Mathematics*, **68**, 237–267.
- Lo AW (1991). Long-term memory in stock market prices, *Econometrica*, **59**, 1279–1313.
- O'Brien GL (1987). Extreme values for stationary and Markov sequences, *Annals of Probability*, **15**, 281–291.
- Quintos C, Fan Z, and Phillips PCB (2001). Structural change tests in tail behaviour and the Asian crisis, *The Review of Economics Studies*, **68**, 633–663.

Received July 8, 2019; Revised December 11, 2019; Accepted January 6, 2020