Journal of the Korean
Statistical Society
Vol. 26, No. 1, 1997

On Asymptotic Properties of Bootstrap for Autoregressive Processes with Regularly Varying Tail Probabilities†

Hee-Jeong Kang[1]

Abstract

Let $X_t = \beta X_{t-1} + \varepsilon_t$ be an autoregressive process where $|\beta| < 1$ and $\{\varepsilon_t\}$ is independent and identically distributed with regularly varying tail probabilities. This process is called the asymptotically stationary first-order autoregressive process (AR(1)) with infinite variance. In this paper, we obtain a host of weak convergences of some point processes based on bootstrapping of $\{X_t\}$. These kinds of results can be generalized under the infinite variance assumption to ensure the asymptotic validity of the bootstrap method for various functionals of $\{X_t\}$ such as partial sums, sample covariance and sample correlation functions, etc.

Key Words : Bootstrap; Autoregressive process; Regular variation; Point processes; LAD estimator.

---

†This paper was supported (in part) by research funds of Chonbuk National University.
[1]Dept. of Statistics, Chonbuk National University, Chonju, Chonbuk 561-756, Korea.

1. INTRODUCTION

Recently, there has been increasing interest in developing stochastic processes to model heavy-tailed data which usually can be found in economic phenomena. One of assumptions in these models is the error terms have regularly varying tail probabilities. In such cases, asymptotic behaviors of various functionals of observations can be obtained by point process convergence results as in Davis and Resnick (1985) and Davis, Knight and Liu (1992). In this paper, we investigate the bootstrap of point processes which are related to the distribution of LAD (Least Absolute Deviation) estimators and prove that the bootstrap procedure is asymptotically valid when the bootstrap resample size m satisfies $m \longrightarrow \infty$ and $m/n \longrightarrow 0$ as the original sample size n goes to infinity. For this purpose, let $\{X_t\}$ be observations from the first-order autoregressive process (AR(1)) satisfying the difference equations

$$X_t = \beta X_{t-1} + \varepsilon_t \ , \qquad X_0 = 0 \ , \qquad t = 1, \ldots, n \tag{1.1}$$

where β is a parameter of the process with $|\beta| < 1$ and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random errors whose distribution is in the domain of attraction of a stable distribution with index $1 < \alpha < 2$ (written as $\varepsilon_1 \in D(\alpha)$). In other words, the distribution of $\{\varepsilon_t\}$ satisfies the following conditions

$$P[\ |\varepsilon_1| > x\ ] = x^{-\alpha} L(x) \tag{1.2}$$

and

$$\lim_{x \to \infty} \frac{P[\ \varepsilon_1 > x\ ]}{P[\ |\varepsilon_1| > x\ ]} = p \tag{1.3}$$

where $L(x)$ is a slowly varying function at ∞ and $0 \leq p \leq 1$. See Feller (1971) for more details on the domain of attraction of a stable law. The AR(1) process under the above conditions is asymptotically stationary and has a finite mean and an infinite variance. Without loss of generality, we assume that ε_1 has a continuous distribution. In fact, the latter two conditions (1.2) and (1.3) imply that

$$nP(\ |\varepsilon_1|\ >\ a_n x\ )\ \longrightarrow\ x^{-\alpha} \quad \text{for all} \quad x > 0 \tag{1.4}$$

where $\{a_n\}$ is a sequence of positive constants such that

$$a_n\ =\ \inf\{\ x\ ;\ P\ [\ |\varepsilon_1| > x\ ] \leq n^{-1}\ \}. \tag{1.5}$$

Under the above infinite variance conditions, Davis, Knight and Liu (1992) showed that the LAD estimator $\hat{\beta}_n$ which satisfies the following

$$\sum_{t=1}^{n} \left| X_t - \hat{\beta}_n X_{t-1} \right| = \inf_{|\beta|<1} \sum_{t=1}^{n} |X_t - \beta X_{t-1}| \tag{1.6}$$

converges to the unknown parameter β in (1.1) faster than the least squares estimator and the rate of convergence is getting fast as α, the index of law, decreases to 1. Also, they proved by the point process techniques that the LAD estimator $\hat{\beta}_n$ in (1.6) has a limiting distribution when $\hat{\beta}_n$ has been normalized with the sequences of $\{a_n\}$ in (1.5).

The purpose of this study is to show that these point process results are approximated well by the bootstrap method and these kinds of results can be generalized to approximate the weak limit behavior of various functionals of observations under these infinite variance assumptions.

The implementation of the bootstrap procedure can be done as following; given the original sample $X_1, \ldots, X_n$, construct the bootstrap sample $X_1^*, \ldots, X_m^*$ by the recursive formula

$$X_t^* = \hat{\beta}_n X_{t-1}^* + \varepsilon_t^* , \qquad X_0^* = 0 , \qquad t = 1, \ldots, m. \tag{1.7}$$

where $\{\varepsilon_t^*\}$'s are i.i.d. random variables from $\hat{F}_n$, the empirical distribution function of the residuals, $\hat{\varepsilon}_t = X_t - \hat{\beta}_n X_{t-1}$, $t = 1, \ldots, n$ with $\hat{\beta}_n$ given by (1.6). In general, the presence of (*) will denote that we are dealing with the bootstrap quantity.

The limit behaviors of various functionals of observations in the present context heavily depend on the assumption in (1.4). In order to reproduce the same asymptotic results with the above restriction on the bootstrap resample size m, we need to show that the similar result holds for the bootstrap random variables, namely that for all $x > 0$,

$$mP^*(|\varepsilon_1^*| > a_m x) \longrightarrow x^{-\alpha} \qquad \textit{in probability} \tag{1.8}$$

where $\{a_m\}$ is defined in (1.5).

The remainder of this paper is organized as follows. In Section 2, we consider the properties of the bootstrap distribution of random variables $\{\varepsilon_t^*\}$. Section 3 contains the extention of these results to the point processes based on the bootstarp random variables $\{X_t^*\}$ in (1.7). We essentially follow the ideas of Davis and Resnick (1985) to derive the same asymptotic results for the point processes based on $\{X_t^*\}$. The applications of these point process convergence results under the infinite variance settings are considered in Section 4. For further background on point processes, see Davis and Resnick (1985) and Resnick (1987).

2. PRELIMINARY RESULTS

Let $\{\varepsilon_t\}$ be an i.i.d. sequence having regularly varying tail probabilities as defined in (1.2) and (1.3). Further let $\{\varepsilon_{k,i}\}$, $\{\delta_k\}$ and $\{\Gamma_k\}$ be independent sequences of random variables such that $\varepsilon_{k,i} \stackrel{d}{=} \varepsilon_1$ *i.i.d.*, $\{\delta_k\}$ are *i.i.d.* with $P[\delta_k = 1] = p = 1 - P[\delta_k = -1]$, and $\Gamma_k = E_1 + \cdots + E_k$, where $\{E_i\}$'s are *i.i.d.* exponential r.v.'s with mean 1. Also denote $\{\varepsilon_t^*\}$ be a sequence of bootstrap random variables from $\hat{F}_n$, the empirical distribution of residuals $\{\hat{\varepsilon}_t\}$. For a fixed integer $r > 1$, let

$$J_m^* = \sum_{k=1}^{m} I_{(\varepsilon_k^*,\, a_m^{-1} Z_k^{(r)})} \qquad \text{and} \qquad J = \sum_{k=1}^{\infty}\sum_{i=1}^{r} I_{(\varepsilon_{k,i},\, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i)} \tag{2.1}$$

where $I_x(B)$ is defined as an indicator function of x for $B \in R^{r+1}$, a_m is in (1.5), $Z_k^{(r)} = (\varepsilon_{k-1}^*, \varepsilon_{k-2}^*, \ldots, \varepsilon_{k-r}^*)$ and $\mathbf{e}_i \in R^r$ is the basis element with ith component equal to one and the rest zero. See Davis and Resnick (1985) for more details on the relevant state space and relevant measure for the processes $\{J_m^*\}$ and J.

Define the measure $\mu(dt, dx) = dt \times \lambda(dx)$ on the space $R \times (\bar{R}\backslash\{0\})$ where $\lambda(dx) = \alpha p x^{-\alpha-1} 1_{(0,\infty)}(x)dx + \alpha(1-p)(-x)^{-\alpha-1} 1_{(-\infty,0)}(x)dx$ and let S be the collection of all sets B of the form

$$B = (b_0, c_0] \times (b_1, c_1] \times \cdots \times (b_r, c_r] \tag{2.2}$$

where the r-dimensional rectangle $(b_1, c_1] \times \cdots \times (b_r, c_r]$ is bounded away from $(0, 0, \cdots, 0)$ and $b_i < c_i, b_i \neq 0, c_i \neq 0$ for $i = 1, \ldots, r$. Moreover, since $B \in S$ is bounded away from zero, either

(C1) $\quad (b_1, c_1] \times \cdots \times (b_r, c_r] \cap \{y\mathbf{e}_i : y \in R\} = \emptyset \quad$ for $\quad i = 1, \ldots, r$

or

(C2) $\quad (b_1, c_1] \times \cdots \times (b_r, c_r] \cap \{y\mathbf{e}_i : y \in R\} = \begin{cases} (b_{i'}, c_{i'}] & i = i' \\ \emptyset & i \neq i'. \end{cases}$

That is, the rectangle $(b_1, c_1] \times \cdots \times (b_r, c_r]$ either has empty intersection with all of the coordinate axes or intersects exactly one in an interval. Note that in **(C2)**, $b_i < 0 < c_i$ for $i \neq i'$ and $0 \notin (b_{i'}, c_{i'}]$. Also, it is obvious that

$$P(\ J(\partial B) = 0\ ) = 1 \ \text{ for all } \ B \in S.$$

In the following lemmas, we state some preliminary results which are necessary to prove some basic convergences of the point processes $\{J_m^*\}$ and J in (2.1).

Lemma 2.1. With probability 1,

$$\sup_{x \in R} |\ P^*(\ \varepsilon_1^* \le x\ ) - P(\ \varepsilon_1 \le x\ )\ | \longrightarrow 0\ . \tag{2.3}$$

Proof. Let $F(x) = P(\ \varepsilon_1 \le x\ )$. First notice that for all $\delta > 0$, and for all $x \in R$,

$$\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n} I(\hat{\varepsilon}_t \le x) \\
&= \frac{1}{n}\sum_{t=1}^{n} I(\hat{\varepsilon}_t \le x,\ |\hat{\varepsilon}_t - \varepsilon_t| \le \delta) + \frac{1}{n}\sum_{t=1}^{n} I(\hat{\varepsilon}_t \le x, |\hat{\varepsilon}_t - \varepsilon_t| > \delta) \\
&\le \frac{1}{n}\sum_{t=1}^{n} I(\varepsilon_t \le x + \delta) + \frac{1}{\delta}\frac{1}{n}\sum_{t=1}^{n} |\hat{\varepsilon}_t - \varepsilon_t|\ .
\end{aligned} \tag{2.4}$$

Similarly,

$$\frac{1}{n}\sum_{t=1}^{n} I(\hat{\varepsilon}_t \le x) \ \ge\ \frac{1}{n}\sum_{t=1}^{n} I(\varepsilon_t \le x - \delta) - \frac{1}{\delta}\frac{1}{n}\sum_{t=1}^{n} |\hat{\varepsilon}_t - \varepsilon_t|\ . \tag{2.5}$$

Thus, by (2.4) and (2.5), for all $\delta > 0$, and for all $x \in R$,

$$\begin{aligned}
\sup_x |\ P^*(\ \varepsilon_1^* \le x\ ) - P(\ \varepsilon_1 \le x\ )\ | &= \sup_x \left| \frac{1}{n}\sum_{t=1}^{n} I(\hat{\varepsilon}_t \le x) - F(x) \right| \\
&\le \sup_x \left| \frac{1}{n}\sum_{t=1}^{n} I(\varepsilon_t \le x) - F(x) \right| + \sup_x |\ F(x+\delta) - F(x-\delta)\ | \\
&\quad + \frac{1}{\delta}\frac{1}{n}\sum_{t=1}^{n} |\hat{\varepsilon}_t - \varepsilon_t| \\
&\le \sup_x \left| \frac{1}{n}\sum_{t=1}^{n} I(\varepsilon_t \le x) - F(x) \right| + \sup_x |\ F(x+\delta) - F(x-\delta)\ | \\
&\quad + \frac{1}{\delta}|\hat{\beta}_n - \beta| \frac{1}{1 - |\beta|}\frac{1}{n}\sum_{t=1}^{n} |\varepsilon_t|\ .
\end{aligned}$$

So, by the Glivenko-Cantelli theorem and by the result of Gross and Steiger (1979), $\hat{\beta}_n \overset{\text{a.s.}}{\to} \beta$, and $(1/n)\sum_{t=1}^{n}|\varepsilon_t| \overset{\text{a.s.}}{\to} E|\varepsilon_1| < \infty$, for all $\delta > 0$, *almost surely,*

$$\begin{aligned}\limsup_{n\to\infty}\ \sup_{x\in R}&|\, P^*(\, \varepsilon_1^* \le x\,) - P(\,\varepsilon_1 \le x\,)\,| \\ &\le \sup_{x\in R}\ |\, F(x+\delta) - F(x-\delta)\,| \quad \longrightarrow \quad 0\end{aligned}$$

as $\delta \to 0$ since the distribution function of ε_1 is continuous. Therefore, (2.3) is proved.

Lemma 2.2. For any interval $(b, c\,]$ on R with $b \neq 0, c \neq 0, b < c$ and $0 \notin (b, c]$,

$$mP^*(\, a_m^{-1}\varepsilon_1^* \in (b, c\,]\,) \longrightarrow \lambda(\,(b, c\,]\,) \quad \textit{in probability} \tag{2.6}$$

provided,

$$m,\ n \longrightarrow \infty \quad \text{in such a way that} \quad m/n \longrightarrow 0 \tag{2.7}$$

where a_m is defined in (1.5) and the measure λ is defined on $\bar{R}\backslash\{0\}$ by $\lambda(dx) = \alpha p x^{-\alpha-1} 1_{(0,\infty)}(x)dx + \alpha(1-p)(-x)^{-\alpha-1}1_{(-\infty,0)}(x)dx$.

Proof. First of all, under the condition (2.7), we shall prove the following results which are needed to prove the above lemma.

$$\frac{m}{n}\sum_{t=1}^{n} I(\, a_m^{-1}\varepsilon_t \in (b, c\,]\,) - mP(\, a_m^{-1}\varepsilon_1 \in (b, c\,]\,) = o_p(1) \tag{2.8}$$

$$\frac{\max_{1\le t\le n}\ |\,\varepsilon_t\,|}{a_n} = O_p(1) \tag{2.9}$$

$$\frac{\max_{1\le t\le n}\ |\,\hat{\varepsilon}_t - \varepsilon_t\,|}{a_m} = o_p(1)\,. \tag{2.10}$$

First, it is easy to show (2.8) with the Chebyshev's inequality and the condition $m/n \to 0$ and the fact that

$$mP(\, a_m^{-1}\varepsilon_1 \in (b, c\,]\,) \longrightarrow \lambda(b, c\,]\,. \tag{2.11}$$

It is also not difficult to show (2.9) since for every $\delta_1 > 0$,

$$\begin{aligned}P&\left(\frac{\max_{1\le t\le n}\ |\,\varepsilon_t\,|}{a_n} > \delta_1\right) \\ &= 1 - P(\,|\varepsilon_t| \le a_n\delta_1 \quad \text{for all} \quad 1 \le t \le n\,) \\ &= 1 - \left(1 - \frac{1}{n}nP(\,|\varepsilon_1| > a_n\delta_1\,)\right)^n \quad \longrightarrow \quad 1 - \exp(-\delta_1^{-\alpha})\end{aligned}$$

which goes to 0 as $\delta_1 \to \infty$. So, (2.9) is proved. Also, since $a_n(\hat{\beta}_n - \beta) = O_p(1)$ by the result of Davis, Knight and Liu (1992) and $\max_{1\le t\le n} | \hat{\varepsilon}_t - \varepsilon_t | = |\hat{\beta}_n - \beta| \max_{1\le t\le n} |X_{t-1}| \le |\hat{\beta}_n - \beta| \max_{1\le t\le n} |\varepsilon_t| \frac{1}{1-|\beta|}$, we have by (2.9) that

$$\frac{\max_{1\le t\le n} | \hat{\varepsilon}_t - \varepsilon_t |}{a_m} \le \frac{1}{a_m} \frac{\max_{1\le t\le n} | \varepsilon_t |}{a_n} a_n|\hat{\beta}_n - \beta| \frac{1}{1-|\beta|} \longrightarrow 0 \quad \textit{in probability}.$$

Now, to prove (2.6), in view of (2.8) and (2.11), it suffices to show that

$$\begin{aligned} & mP^*(a_m^{-1}\varepsilon_1^* \in (b,c]) - \frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\varepsilon_t \in (b,c]) \\ = & \frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\hat{\varepsilon}_t \in (b,c]) - \frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\varepsilon_t \in (b,c]) \\ \longrightarrow & \ 0 \quad \textit{in probability.} \end{aligned}$$

Notice that

$$\begin{aligned} \frac{m}{n}\sum_{t=1}^n & I(a_m^{-1}\hat{\varepsilon}_t \in (b,c]) - \frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\varepsilon_t \in (b,c]) \\ &= \frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\hat{\varepsilon}_t \in (b,c],\ a_m^{-1}\varepsilon_t \notin (b,c]) \\ &\quad - \frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\hat{\varepsilon}_t \notin (b,c],\ a_m^{-1}\varepsilon_t \in (b,c]). \end{aligned} \tag{2.12}$$

Here, by the argument in Lemma 3.2 of Kang (1995), we can show that

$$\frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\hat{\varepsilon}_t \in (b,c],\ a_m^{-1}\varepsilon_t \notin (b,c]) \xrightarrow{P} 0.$$

Similarly, it is not difficult to prove that for the second term in (2.12),

$$\frac{m}{n}\sum_{t=1}^n I(a_m^{-1}\hat{\varepsilon}_t \notin (b,c],\ a_m^{-1}\varepsilon_t \in (b,c]) \xrightarrow{P} 0.$$

This completes the proof.

Lemma 2.3. Under the conditions of Lemma 2.2, for all $x > 0$,

$$mP^*(a_m^{-1}|\varepsilon_1^*| > x) \longrightarrow x^{-\alpha} \quad \textit{in probability}.$$

Proof. By a similar argument in the proof of Lemma 2.2, we can show that for all $x > 0$,

$$\begin{aligned} mP^*(\, a_m^{-1}\varepsilon_1^* \in (-\infty, -x)\,) &\xrightarrow{\text{p}} \lambda(-\infty, -x) \\ mP^*(\, a_m^{-1}\varepsilon_1^* \in (x, +\infty)\,) &\xrightarrow{\text{p}} \lambda(x, +\infty)\,. \end{aligned}$$

Therefore,

$$\begin{aligned} & mP^*(\, a_m^{-1}|\varepsilon_1^*| > x\,) \\ & = \quad mP^*(\, a_m^{-1}\varepsilon_1^* \in (-\infty, -x)\,) + mP^*(\, a_m^{-1}\varepsilon_1^* \in (x, +\infty)\,) \\ & \xrightarrow{\text{p}} \quad \lambda(-\infty, -x) + \lambda(x, +\infty) = x^{-\alpha}\,. \end{aligned}$$

Remark. When $m = n$, i.e., when the bootstrap sample size is equal to the original sample size, then $mP^*(a_m^{-1}\varepsilon_1^* \in (b, c])$ does not converge to $\lambda(b, c]$ in probability but converges in distribution to a poisson random variable. That is, $mP^*(\, a_m^{-1}\varepsilon_1^* \in (b, c]\,) = nP^*(\, a_n^{-1}\varepsilon_1^* \in (b, c]\,) = \sum_{t=1}^n I(\, a_n^{-1}\hat{\varepsilon}_t \in (b, c]\,) \approx \sum_{t=1}^n I(\, a_n^{-1}\varepsilon_t \in (b, c]\,) \xrightarrow{\text{d}} N$ where N is a poisson random variable with mean $\lambda(b, c]$ because $\sum_{t=1}^n I(\, a_n^{-1}\varepsilon_t \in (b, c]\,)$ is a binomial variable and $nP(\, a_n^{-1}\varepsilon_1 \in (b, c]\,) \to \lambda(b, c]$ as $n \to \infty$.

3. BASIC CONVERGENCES

In this section, we shall prove that under the condition of m, $n \to \infty$ in such a way that $m/n \to 0$, for every nonnegative integer value x, and for all $B \in S$,

$$P^*(\, J_m^*(B) = x\,) - P(\, J(B) = x\,) \longrightarrow 0 \quad \textit{in probability}$$

where J_m^* and J are point processes defined in (2.1) and B is a $(r+1)$-dimensional rectangle in (2.2). An argument involving continuous mapping theorem allows us to extend this result to show that a sequence of point processes based on $\{X_t^*\}$ converges to another limit point process. Here, it is worth noting that the process $\sum_{k=1}^\infty I_{(\delta_k \Gamma_k^{-1/\alpha})}$ is a Poisson process on $\bar{R}\backslash\{0\}$ with intensity measure $\lambda(dx) = \alpha p x^{-\alpha-1} 1_{(0,\infty)}(x)dx + \alpha(1-p)(-x)^{-\alpha-1} 1_{(-\infty,0)}(x)dx$. For more details on Poisson processes, see Resnick (1987). From now on, throughout the paper, we will assume that m, $n \to \infty$ in such a way that $m/n \to 0$, unless stated otherwise. We begin with a lemma which says some properties of the point processes J_m^* and J with respect to the cases **(C1)** and **(C2)** where the cases **(C1)** and **(C2)** are defined in the previous section.

Lemma 3.1. Denote J_m^* and J be the point processes defined in (2.1) and let B in (2.2) satisfy **(C1)**. Then,

$$P(\ J(B)\ =\ 0\ )\ =\ 1\quad and\quad E^*(\ J_m^*(B)\ )\ \xrightarrow{\ \mathrm{P}\ }\ 0\,. \tag{3.1}$$

Also, if B satisfies **(C2)**, then

$$\begin{aligned} P(\ J(B)\ =\ 0\ )\ &=\ exp(\ -\mu((b_0,c_0]\times(b_{i'},c_{i'}])\ )\quad and \\ E^*(\ J_m^*(B)\ )\ &\xrightarrow{\ \mathrm{P}\ }\ \mu((b_0,c_0]\times(b_{i'},c_{i'}])\,. \end{aligned} \tag{3.2}$$

Proof. Consider the case when B satisfies **(C1)**. Then it is obvious that

$$P(\ J(B)\ =\ 0\ )\ =\ 1$$

since $(b_1,c_1]\times\cdots\times(b_r,c_r]$ has empty intersection with all of the coordinate axes and

$$\begin{aligned} E^*(\ J_m^*(B)\ )\ &=\ \sum_{k=1}^{m} P^*(\ \varepsilon_k^*\in(b_0,c_0],a_m^{-1}\varepsilon_{k-1}^*\in(b_1,c_1],\ldots,a_m^{-1}\varepsilon_{k-r}^*\in(b_r,c_r]\ ) \\ &=\ \sum_{k=1}^{m} P^*(\ \varepsilon_k^*\in(b_0,c_0]\ )\prod_{i=1}^{r} P^*(\ a_m^{-1}\varepsilon_{k-i}^*\in(b_i,c_i]\ ) \\ &\le\ \sum_{k=1}^{m} P^*(\ \varepsilon_1^*\in(b_0,c_0]\ )\,(\ P^*(\ a_m^{-1}|\varepsilon_1^*|>d\ )\ )^2 \end{aligned}$$

where d is the minimun of $|b_{i_1}|, |c_{i_1}|, |b_{i_2}|$ and $|c_{i_2}|$ which are end points of the intervals that do not contain zero. Thus,

$$E^*(\ J_m^*(B)\ )\ \le\ m(\ P^*(\ a_m^{-1}|\varepsilon_1^*|>d\ )\ )^2\ =\ \frac{1}{m}\,(\ mP^*(\ a_m^{-1}|\varepsilon_1^*|>d\ )\ )^2\ \xrightarrow{\ \mathrm{P}\ }\ 0$$

by Lemma 2.3. Suppose B satisfies **(C2)**. It is also clear that

$$P(\ J(B)\ =\ 0\ )\ =\ \exp(\ -\mu((b_0,c_0]\times(b_{i'},c_{i'}])\ )$$

since the point process J is a PRM *(Poisson random measure)* with mean measure $\mu((b_0,c_0]\times(b_{i'},c_{i'}])$. As above,

$$\begin{aligned} E^*(\ J_m^*(B)\ )\ &=\ \sum_{k=1}^{m} P^*(\ \varepsilon_k^*\in(b_0,c_0],a_m^{-1}\varepsilon_{k-1}^*\in(b_1,c_1],\ldots,a_m^{-1}\varepsilon_{k-r}^*\in(b_r,c_r]\ ) \\ &=\ \sum_{k=1}^{m} P^*(\ \varepsilon_1^*\in(b_0,c_0]\ )\prod_{i=1}^{r} P^*(\ a_m^{-1}\varepsilon_1^*\in(b_i,c_i]\ ) \end{aligned}$$

since $b_i < 0 < c_i$ for $i \neq i'$ and $a_m \to \infty$, $\prod_{i \neq i'} P^*(a_m^{-1}\varepsilon_1^* \in (b_i, c_i]) \to 1$ a.s. Therefore,

$$\begin{aligned} E^*(J_m^*(B)) &\approx P^*(\varepsilon_1^* \in (b_0, c_0])\, m P^*(a_m^{-1}\varepsilon_1^* \in (b_{i'}, c_{i'}]) \\ &\xrightarrow{\text{P}} P(\varepsilon_1 \in (b_0, c_0])\, \lambda(b_{i'}, c_{i'}] = \mu((b_0, c_0] \times (b_{i'}, c_{i'}]) \end{aligned}$$

by Lemma 2.1 and Lemma 2.2.

Proposition 3.2. Let $\tilde{J}_m^* = \sum_{k=1}^m \sum_{i=1}^r I_{(\varepsilon_{k+i}^*, a_m^{-1}\varepsilon_k^* \mathbf{e}_i)}$. Then for all nonnegative integer value x, and for all $B \in S$,

$$P^*(J_m^*(B) = x) - P^*(\tilde{J}_m^*(B) = x) \longrightarrow 0 \quad \textit{in probability.} \tag{3.3}$$

Proof. Since $J_m^*(B)$ and $\tilde{J}_m^*(B)$ are nonnegative integer valued, it suffices to show that for all $\delta > 0$,

$$P^*(| J_m^*(B) - \tilde{J}_m^*(B) | > \delta) \longrightarrow 0 \quad \textit{in probability.} \tag{3.4}$$

First, consider the case when B satisfies **(C1)**. Then by (3.1), $E^*(J_m^*(B)) \xrightarrow{\text{P}} 0$ and $\tilde{J}_m^*(B) = 0$ a.s. because the points of $\tilde{J}_m^*$ are located on the coordinates axes. So, by the Markov's inequality, (3.4) is proved. Now, suppose that B satisfies **(C2)** in which case $0 \in (b_i, c_i\,], i \neq i'$ and $0 \notin (b_{i'}, c_{i'}\,]$. Note that

$$\begin{aligned} J_m^*(B) &= \sum_{k=1}^{i'-1} I_{(\varepsilon_k^*, a_m^{-1} Z_k^{(r)})}(B) + \sum_{k=i'}^{m} I_{(\varepsilon_k^*, a_m^{-1} Z_k^{(r)})}(B) \\ &\overset{\text{let}}{=} J_{m1}^*(B) + J_{m2}^*(B) \end{aligned}$$

By the same reason as in the proof of Lemma 3.1, we can show that

$$\begin{aligned} E^*(J_{m1}^*(B)) &\leq (i'-1)\, P^*(\varepsilon_1^* \in a_m (b_{i'}, c_{i'}\,]) \\ &\longrightarrow 0 \quad \textit{in probability} \end{aligned} \tag{3.5}$$

because $a_m \to \infty$ and $0 \notin (b_{i'}, c_{i'}\,]$. As for the second term $J_{m2}^*(B)$, by (3.2), we can see that

$$E^*(J_{m2}^*(B)) \xrightarrow{\text{P}} P(\varepsilon_1 \in (b_0, c_0\,])\, \lambda(b_{i'}, c_{i'}\,] = \mu((b_0, c_0\,] \times (b_{i'}, c_{i'}\,]) \tag{3.6}$$

Also, by Lemma 2.1 and 2.2,

$$\begin{aligned} E^*(\tilde{J}_m^*(B)) &= \sum_{k=1}^m P^*(\varepsilon_{k+i'}^* \in (b_0, c_0\,],\ a_m^{-1}\varepsilon_k^* \in (b_{i'}, c_{i'}\,]) \\ &= P^*(\varepsilon_1^* \in (b_0, c_0\,])\, m P^*(a_m^{-1}\varepsilon_1^* \in (b_{i'}, c_{i'}\,]) \\ &\xrightarrow{\text{P}} P(\varepsilon_1 \in (b_0, c_0\,])\, \lambda(b_{i'}, c_{i'}\,] = \mu((b_0, c_0\,] \times (b_{i'}, c_{i'}\,]) \end{aligned} \tag{3.7}$$

Therefore, by (3.5), (3.6) and (3.7),

$$E^*(\ J_m^*(B)\ -\ \tilde{J}_m^*(B)\ )\ \longrightarrow\ 0\quad \textit{in probability}$$

and with the Markov's inequality, this completes the proof.

Proposition 3.3. For each fixed positive integer $r > 1$,

$$J_m^* = \sum_{k=1}^{m} I_{(\varepsilon_k^*,\, a_m^{-1} Z_k^{(r)})} \xrightarrow{\ d\ } J = \sum_{k=1}^{\infty}\sum_{i=1}^{r} I_{(\varepsilon_{k,i},\, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i)} \quad \textit{in probability}$$

in $M_p(R \times (\bar{R}^r \backslash \{(0,0,\ldots,0)\}))$. That is, for every nonnegative integer value x, and for all $B \in S$,

$$P^*(\ J_m*(B)\ =\ x\ )\ -\ P(\ J(B)\ =\ x\ )\ \longrightarrow\ 0\quad \textit{in probability} \tag{3.8}$$

where $Z_k^{(r)} = (\varepsilon_{k-1}^*, \varepsilon_{k-2}^*, \ldots, \varepsilon_{k-r}^*)$.

Proof. In order to prove (3.8), by Proposition 3.2, it suffices to show that

$$\sum_{k=1}^{m}\sum_{i=1}^{r} I_{(\varepsilon_{k+i}^*,\, a_m^{-1}\varepsilon_k^* \mathbf{e}_i)} \xrightarrow{\ d\ } \sum_{k=1}^{\infty}\sum_{i=1}^{r} I_{(\varepsilon_{k,i},\, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i)} \quad \textit{in probability.}$$

However, by the continuous mapping theorem, this is equivalent to show that

$$\sum_{k=1}^{m} I_{(\varepsilon_{k+r}^*, \ldots, \varepsilon_{k+1}^*,\, a_m^{-1}\varepsilon_k^*)} \xrightarrow{\ d\ } \sum_{k=1}^{\infty} I_{(\varepsilon_{k,r}, \ldots, \varepsilon_{k,1},\, \delta_k \Gamma_k^{-1/\alpha})} \tag{3.9}$$

in probability because the composition of the following two continuous mappings,

$$\begin{aligned}
&\sum_{k=1}^{\infty} I_{(\varepsilon_{k,r}, \varepsilon_{k,r-1}, \ldots, \varepsilon_{k,1},\, \delta_k \Gamma_k^{-1/\alpha})} \\
&\quad\longmapsto \left(\sum_{k=1}^{\infty} I_{(\varepsilon_{k,r},\, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_r)}, \ldots, \sum_{k=1}^{\infty} I_{(\varepsilon_{k,1},\, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_1)} \right) \\
&\quad\longmapsto \sum_{k=1}^{\infty}\sum_{i=1}^{r} I_{(\varepsilon_{k,i},\, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i)}
\end{aligned}$$

is itself a continuous mapping from $M_p(R \times R \times \cdots \times R \times \bar{R}\backslash\{0\})$ into $M_p(R \times (\bar{R}^r \backslash \{(0,0,\ldots,0)\}))$. Also, by using Laplace functional of point process, which is defined as $\Psi_\nu(f) = E\{\exp\{-\nu(f)\}\} = E\{\exp\{-\int_B f(x)\nu(dx,\omega)\}\}$

where ν is a point process, $f \in C_K^+(B)$ and $C_K^+(B)$ is the space of non-negative continuous functions $B \to R_+$ with compact support, to show (3.9) is equivalent to prove that

$$\Big| E^* \Big\{ \exp\Big\{ -\sum_{k=1}^{m} f\left(\varepsilon_{k+r}^*, \varepsilon_{k+r-1}^*, \ldots, \varepsilon_{k+1}^*, a_m^{-1}\varepsilon_k^*\right) \Big\} \Big\} - E\Big\{ \exp\Big\{ -\sum_{k=1}^{\infty} f\left(\varepsilon_{k,r}, \varepsilon_{k,r-1}, \ldots, \varepsilon_{k,1}, \delta_k \Gamma_k^{-1/\alpha}\right) \Big\} \Big\} \Big| \longrightarrow 0 \quad \textit{in probability} \tag{3.10}$$

for all $f \in C_K^+(B)$. Here, by the results of Lemma 2.1, Lemma 2.2 and the argument in Proposition 3.6 of Kang (1995), it is not too difficult to prove (3.10). So, we will omit the proof for the sake of brevity and this completes the proof of Proposition 3.3. See Kang (1995) for the details.

We now use the result of Proposition 3.3 to obtain some results on the weak convergence of a point process based on $\{X_t^*\}$. The continuous mapping theorem is exceedingly useful to get this result. We begin with a lemma which parallels Lemma 2.3 in Davis and Resnick (1985).

Lemma 3.4. For any $\delta > 0$,

$$\lim_{r\to\infty} \limsup_{m\to\infty} P^* \left(a_m^{-1} \max_{1\le t\le m} \left| X_t^* - \sum_{j=0}^{r-1} \hat\beta_n^j \varepsilon_{t-j}^* \right| > \delta \right) = 0 \quad \textit{in probability}.$$

Proof. First, note that if $1 \le t \le r$ then,

$$P^* \left(a_m^{-1} \max_{1\le t\le r} \left| X_t^* - \sum_{j=0}^{r-1} \hat\beta_n^j \varepsilon_{t-j}^* \right| > \delta \right) = 0 \quad a.s.$$

because $\left| X_t^* - \sum_{j=0}^{r-1} \hat\beta_n^j \varepsilon_{t-j}^* \right| = 0$ for $1 \le t \le r$ (notice that $\varepsilon_{-j}^* = 0$ for $j \ge 0$). If $r < t \le m$, then $\left| X_t^* - \sum_{j=0}^{r-1} \hat\beta_n^j \varepsilon_{t-j}^* \right| = \left| \hat\beta_n^r \sum_{j=r}^{t-1} \hat\beta_n^{j-r} \varepsilon_{t-j}^* \right|$. So, $\max_{r<t\le m} \left| X_t^* - \sum_{j=0}^{r-1} \hat\beta_n^j \varepsilon_{t-j}^* \right| \le \left(|\hat\beta_n|^r / (1 - |\hat\beta_n|) \right) \max_{1\le t\le m} |\varepsilon_t^*|$. Thus, by the convexity argument, for any $\delta > 0$,

$$P^* \left(a_m^{-1} \max_{1\le t\le m} \left| X_t^* - \sum_{j=0}^{r-1} \hat\beta_n^j \varepsilon_{t-j}^* \right| > \delta \right) \le \frac{m}{n} \sum_{t=1}^{n} I\left(|\hat\varepsilon_t| > a_m \, \delta \, \frac{1 - |\hat\beta_n|}{|\hat\beta_n|^r} \right) - \frac{m}{n} \sum_{t=1}^{n} I\left(|\varepsilon_t| > a_m \, \delta \, \frac{1 - |\beta|}{|\beta|^r} \right)$$

$$+ \frac{m}{n} \sum_{t=1}^{n} I\left(\mid \varepsilon_t \mid \; > \; a_m \, \delta \, \frac{1-|\beta|}{|\beta|^r} \right).$$

Here, by (2.8), (2.11) and using the same argument as in the proof of Lemma 2.2 and the fact that $\hat{\beta}_n \xrightarrow{\text{a.s.}} \beta$, we can obtain that

$$\begin{aligned}\limsup_{m\to\infty} P^* &\left(a_m^{-1} \max_{1\le t\le m} \left| X_t^* - \sum_{j=0}^{r-1} \hat{\beta}_n^j \varepsilon_{t-j}^* \right| > \delta \right) \\ &= \left(\delta \left(\frac{1-|\beta|}{|\beta|^r} \right) \right)^{-\alpha} \quad \textit{in probability}\end{aligned}$$

and

$$\lim_{r\to\infty} \left(\delta \left(\frac{1-|\beta|}{|\beta|^r} \right) \right)^{-\alpha} = 0$$

since $|\beta| < 1$ and $1 < \alpha < 2$. This completes the proof.

The following is the main result of this paper which is about the weak convergence of a point process based on the bootstrap random variables $\{X_t^*\}$.

Theorem 3.5. Suppose $\{X_t^*\}$ is the process given by (1.7). Then,

$$\sum_{k=1}^{m} I_{(\, \varepsilon_k^*, a_m^{-1} X_{k-1}^* \,)} \xrightarrow{\text{d}} \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} I_{(\, \varepsilon_{k,i}, \delta_k \Gamma_k^{-1/\alpha} \beta^{i-1} \,)} \quad \textit{in probability}$$

in $M_p(R \times \bar{R}\backslash\{0\})$.

Proof. From Proposition 3.3, we have that for any fixed positive integer $r > 1$,

$$\sum_{k=1}^{m} I_{(\varepsilon_k^*, a_m^{-1} Z_k^{(r)})} \xrightarrow{\text{d}} \sum_{i=1}^{r} \sum_{k=1}^{\infty} I_{(\varepsilon_{k,i}, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i)} \quad \textit{in probability}$$

in $M_p(R \times (\bar{R}^r\backslash\{(0,0,\ldots,0)\}))$. Here, we note that the map

$$(\, v_{k-1}, v_{k-2}, \ldots, v_{k-r} \,) \rightarrow \sum_{i=1}^{r} c_i v_{k-i}$$

induce a continuous map from $M_p(R \times (\bar{R}^r\backslash\{(0,0,\ldots,0)\})) \rightarrow M_p(R \times \bar{R}\backslash\{0\})$. So, by Theorem 5.5 in Billingsley (1968), we can say that

$$\sum_{k=1}^{m} I_{(\varepsilon_k^*, a_m^{-1} \sum_{i=1}^{r} \hat{\beta}_n^{i-1} \varepsilon_{k-i}^*)} \xrightarrow{\text{d}} \sum_{i=1}^{r} \sum_{k=1}^{\infty} I_{(\varepsilon_{k,i}, \delta_k \Gamma_k^{-1/\alpha} \beta^{i-1})} \quad \textit{in probability} \tag{3.11}$$

in $M_p(R \times \bar{R}\backslash\{0\})$. Also, as $r \to \infty$

$$\sum_{i=1}^{r}\sum_{k=1}^{\infty} I_{(\varepsilon_{k,i},\, \delta_k \Gamma_k^{-1/\alpha} \beta^{i-1})} \longrightarrow \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} I_{(\varepsilon_{k,i},\, \delta_k \Gamma_k^{-1/\alpha} \beta^{i-1})} \tag{3.12}$$

pointwise. Thus by (3.11), (3.12) and the argument in Proposition 3.8 of Kang (1995), it suffices to show that for each $\eta > 0$, and for $f \in C_K^+(R \times \bar{R}\backslash\{0\})$,

$$\lim_{r\to\infty} \limsup_{m\to\infty} P^* \left(\left| \sum_{k=1}^{m} f(\varepsilon_k^*, a_m^{-1} X_{k-1}^*) - \sum_{k=1}^{m} f(\varepsilon_k^*, a_m^{-1} \sum_{i=1}^{r} \hat{\beta}_n^{i-1} \varepsilon_{k-i}^*) \right| > \eta \right) = 0 \quad in\ probability. \tag{3.13}$$

The proof of (3.13) is not too difficult when we follow the idea of Theorem 2.4 in Davis and Resnick (1985) and use the results of Lemma 3.4. Hence, the proof will be omitted for brevity and this completes the proof of Theorem 3.5.

Corollary 3.6. For all non-negative continuous functions f on $R \times \bar{R}\backslash\{0\}$ with compact support,

$$\sum_{k=1}^{m} f(\, \varepsilon_k^*, a_m^{-1} X_{k-1}^* \,) \xrightarrow{\text{d}} \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} f(\, \varepsilon_{k,i}, \delta_k \Gamma_k^{-1/\alpha} \beta^{i-1} \,) \quad in\ probability.$$

Proof. This follows from Theorem 3.5 and Theorem 2.1 in Billingsley (1968).

Remark. As an application of the above corollary, Davis, Knight and Liu (1992) used the following function; $f(x,y) = f(x,y)I(|x| \le M)I(|y| > \delta)$ for any $M > 0$ and $\delta > 0$ to show the existence of the limiting distribution for the LAD estimators of AR(1) processes with infinite variance.

4. CONCLUDING REMARKS

In this study, we obtained a bunch of point process convergence results which are related to the bootstrap of LAD estimators. An exceedingly useful result of these point process techniques is Corollary 3.6 and this corollary holds for any point process results which are related to not only LAD estimators but also various functionals of observations with regularly varying error terms. In fact, by using the point process techniques, Davis and Resnick (1985) obtained the weak limit behavior of various functionals of observations

$\{X_t\ , t \geq 1\}$ such as partial sums, sample covariance functions and sample correlation functions under the conditions that $\{X_t = \sum_{j=0}^{\infty} c_j Z_{t-j}\ ,\ t \geq 1\}$ and as usual $\{Z_j\}$ satisfies (1.2) and (1.3) with $0 < \alpha < 2$ and $\sum_{j=0}^{\infty} |c_j|^{\delta} < \infty$ for some $\delta < \alpha\ ,\ \delta \leq 1$. Here, note that the conditions on a real sequence $\{c_j\ ,\ j \geq 0\}$ are always satisfied for the stationary autoregressive processes. Therefore, since the stationary autoregressive processes can be expressed by moving average processes, it is not difficult to show that the bootstrap procedure is asymptotically valid to approximate the weak limit behavior of various functionals of observations of stationary ARMA processes with infinite variance when we follow the ideas of Davis and Resnick (1985) and apply the arguments used in this paper.

REFERENCES

(1) Billingsley, P. (1968). *Convergence of Probability Measures*, John Wiley & Sons, New York.

(2) Davis, R. A., Knight, K. and Liu, J. (1992). *M*-estimation for autoregressions with infinite variance, *Stochastic Processes and their Applications*, **40**, 145-180.

(3) Davis, R. A. and Resnick, S. I. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities, *The Annals of Probability*, **13**, 179-195.

(4) Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, 2nd ed., Vol. II., John Wiley & Sons, New York.

(5) Gross, S. and Steiger, W. L. (1979). Least absolute deviation estimates in autoregression with infinite variance, *Journal of Applied Probability*, **16**, 104-116.

(6) Kang, H. J. (1995). *Bootstrapping Autoregressive Processes with Infinite Variance*, Ph.D. Thesis, Dept. of Statistics, Penn. State University.

(7) Resnick, S. I. (1986). Point processes, regular variation and weak convergence, *Advances in Applied Probability*, **18**, 66-138.

(8) Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, New York.