Journal of the Korean
Statistical Society
Vol. 26, No. 3, 1997

Bootstrap of LAD Estimate in Infinite Variance AR(1) Processes

Hee-Jeong Kang [1]

Abstract

This paper proves that the standard bootstrap approximation for the least absolute deviation (LAD) estimate of β in AR(1) processes with infinite variance error terms is asymptotically valid in probability when the bootstrap resample size is much smaller than the original sample size. The theoretical validity results are supported by simulation studies.

Key Words : Bootstrap; LAD estimator; Stationary autoregression; Regular variation.

1. INTRODUCTION

Let $\{X_t\}$ be the first-order autoregressive process defined by

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

[1]Department of Statistics, Chonbuk National University, Chonju, Chonbuk 561-756, Korea.

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 satisfies the following conditions

$$P[\,|\varepsilon_1| > x\,] = x^{-\alpha}L(x)\,, \qquad \lim_{x\to\infty}\frac{P[\,\varepsilon_1 > x\,]}{P[\,|\varepsilon_1| > x\,]} = p \tag{1.2}$$

where $L(x)$ is a slowly varying function at ∞, $1 < \alpha < 2$ and $0 \le p \le 1$.

The AR(1) process under the above conditions is asymptotically stationary with an infinite variance and the distribution of the error terms $\{\varepsilon_t\}$ is in the domain of attraction of a stable law (written as $\varepsilon_1 \in D(\alpha)$). See Feller (1971) for more details on the domain of attraction of a stable law. Without loss of generality, we assume that ε_1 has a continuous distribution with median 0. The main focus of this paper is to consider estimates of the parameter β of the process and their distributional properties under the above conditions in (1.2).

Davis *et al.* (1992) have shown that under the conditions in (1.2), the LAD estimator $\hat\beta_{LAD}$ which satisfies the following

$$\sum_{t=1}^{n}\left|X_t - \hat\beta_{LAD}X_{t-1}\right| = \inf_{|\beta|<1}\sum_{t=1}^{n}|X_t - \beta X_{t-1}| \tag{1.3}$$

has a limiting distribution when $\hat\beta_{LAD}$ (denoted as $\hat\beta_n$) has been suitably normalized. However, it is only existential result and computationally intractable.

It is a well known fact that the bootstrap method can provide an alternative procedure for studying the distributional properties of various statistics of interest. In this paper, it is shown that if the bootstrap sample size m is much smaller than the original sample size n, such as $m \to \infty$ and $m < n^{\frac{2}{1+\alpha}-\epsilon}$ for any $\epsilon > 0$, then the bootstrap distribution of the LAD estimator will approximate the true sampling distribution of $\hat\beta_n$. In the finite variance case, Bose (1988) succeeded in proving that the bootstrap method is superior to the normal approximation technique in approximating the distribution of the least squares estimator for the stationary autoregressions when the bootstrap resample size m is equal to the original sample size n. See Theorem 3.9 of Bose (1988) for the details.

In the next section, we present the main results of this paper on the weak consistency of the bootstrap distribution of the LAD estimator. The works of Davis *et al.* (1992) are especially relevant. Section 3 includes some technical results needed in the proof of the main results and Section 4 contains the proof of the validity of the bootstrap procedure. The theoretical results are supported by Monte Carlo simulations in Section 5.

2. MAIN RESULTS

Let $X_1^*, \ldots, X_m^*$ be the bootstrap sample constructed by the recursive formula

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

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 (1.3). Let $\hat{\beta}_m^*$ be the LAD estimator which is obtained by the following equation:

$$\sum_{t=1}^{m} \left| X_t^* - \hat{\beta}_m^* X_{t-1}^* \right| = \inf_{|\phi|<1} \sum_{t=1}^{m} \left| X_t^* - \phi X_{t-1}^* \right| . \tag{2.1}$$

The main aim of this study is to approximate the sampling distribution of $(\hat{\beta}_n - \beta)$ (suitably normalized) by that of $(\hat{\beta}_m^* - \hat{\beta}_n)$ (suitably normalized), given $X_1, \ldots, X_n$. Using the weak consistency of some stochastic processes and the convexity argument, Davis *et al.* (1992) have proved the existence of the limit distribution of the LAD estimator $\hat{\beta}_n$. The limiting stochastic process they used is

$$W(u) = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \left[\left| \varepsilon_{k,i} - u\beta^{i-1} \delta_k \Gamma_k^{-1/\alpha} \right| - |\varepsilon_{k,i}| \right] \tag{2.2}$$

where $\{\varepsilon_{k,i}\}$, $\{\delta_k\}$ and $\{\Gamma_k\}$ are independent sequences of random variables, $\varepsilon_{k,i} \stackrel{d}{=} \varepsilon_1$ *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.

In order to establish the weak consistency of the bootstrap distribution of the LAD estimate, let $W_m^*(u)$ be a stochastic process which is defined by

$$W_m^*(u) = \sum_{t=2}^{m} \left[\left| \varepsilon_t^* - a_m^{-1} u X_{t-1}^* \right| - | \varepsilon_t^* | \right] \tag{2.3}$$

where a_m is

$$a_m = \inf \{ x \; ; \; P [|\varepsilon_1| > x] \leq m^{-1} \}$$

and u is defined as $a_m(\phi - \hat{\beta}_n)$, $\phi \in R$ with $\hat{\beta}_n$ in (1.3). Note that the minimizer of $W_m^*(u)$ is $a_m(\hat{\beta}_m^* - \hat{\beta}_n)$ where $\hat{\beta}_m^*$ is a LAD estimator which is based on a bootstrap sample $X_1^*, \ldots, X_m^*$.

The following theorem is the main result of this paper on the weak consistency of the bootstrap distribution of the LAD estimator. The proofs of the following theorems and lemma will be in Section 4. Here, we assume that $m = m(n) \uparrow \infty$ as $n \uparrow \infty$ throughout the paper.

Theorem 2.1. Let $\{X_t\}$ be an AR(1) process satisfying the conditions in (1.2). Let $\xi_{m,n}^* = a_m(\hat{\beta}_m^* - \hat{\beta}_n)$. If $1 < \alpha < 2$ and $E(|\varepsilon_1|^\gamma) < \infty$ for some $\gamma < 1 - \alpha$, Then, $\sup_x \left| P^*(\xi_{m,n}^* \le x) - P(\xi \le x) \right| \xrightarrow{\mathrm{P}} 0$.
i.e., for any $\delta > 0$ and $\epsilon > 0$,

$$P\{\, \omega \,;\, \sup_x \left| P^*(\xi_{m,n}^* \le x) - P(\xi \le x) \right| > \delta \,\} \longrightarrow 0 \tag{2.4}$$

provided,

$$m < n^{\frac{2}{1+\alpha} - \epsilon}. \tag{2.5}$$

Remark 1. The boundedness of the density function f of ε_1 in a neighborhood of 0 is sufficient for the moment condition. See Davis *et al.* (1992).

One of the difficulties of the above theorem is that the normalizing factor a_m of the bootstrap distribution of the LAD estimator is not known. However, the following lemma suggests an alternative scaling factor, which is known. Thus, we can say that the true sampling distribution of the LAD estimator can be estimated without the knowledge of the underlying distribution or the limiting distribution.

Lemma 2.1. Let $\{\varepsilon_t\}$ have a distribution in $D(\alpha)$. Let $|\hat{\varepsilon}|_{(1)}, \ldots, |\hat{\varepsilon}|_{(n)}$ be the order statistics of $|\hat{\varepsilon}_1|, \ldots, |\hat{\varepsilon}_n|$ where $\{\hat{\varepsilon}_t\}$'s are residuals. Then,

$$\frac{|\hat{\varepsilon}|_{(n-[\frac{n}{m}]+1)}}{a_m} \xrightarrow{\mathrm{P}} 1$$

as m , $n \to \infty$ in such a way that $m/n \to 0$ where $[x]$ is the largest integer function.

Theorem 2.2. Let $\xi_{m,n,1}^* = |\hat{\varepsilon}|_{(n-[\frac{n}{m}]+1)}(\hat{\beta}_m^* - \hat{\beta}_n)$. Then, under the conditions of Theorem 2.1, $\sup_x \left| P^*(\xi_{m,n,1}^* \le x) - P(\xi \le x) \right| \xrightarrow{\mathrm{P}} 0$.
i.e., for any $\delta > 0$ and $\epsilon > 0$,

$$P\{\, \omega \,;\, \sup_x \left| P^*(\xi_{m,n,1}^* \le x) - P(\xi \le x) \right| > \delta \,\} \longrightarrow 0$$

provided,

$$m < n^{\frac{2}{1+\alpha} - \epsilon} .$$

The proofs of the main results of this study rely on point process techniques as can be found in Davis and Resnick (1985). We essentially adopt the ideas of that paper to derive some results on the weak convergence of the point processes based on the bootstrap random variables $\{X_t^*\}$. See Kang (1997) for more details. For further background on point processes, see Resnick (1987).

3. SOME PRELIMINARIES

In this paper, we want to show the validity of the bootstrap method by proving that if $m < n^{\frac{2}{1+\alpha} - \epsilon}$ for any $\epsilon > 0$, then

$$a_m(\hat{\beta}_m^* - \hat{\beta}_n) \xrightarrow{d} \xi \quad \textit{in probability} \tag{3.1}$$

where ξ is the minimum of the stochastic process $W(u)$ in (2.2) and $\hat{\beta}_n$ and $\hat{\beta}_m^*$ are the LAD estimators defined in (1.3) and (2.1), respectively. However, due to Lemma 2.2 in Davis *et al.* (1992), we can easily get the above result (3.1) if we are able to prove that

$$W_m^*(\cdot) \xrightarrow{d} W(\cdot) \quad \textit{in probability} \tag{3.2}$$

on $C(R)$ under the condition (2.5). For this goal, with the following eqivalence

$$\begin{aligned} |x-y| - |x| &= y\,(\,I(x<0) - I(x>0)\,) \\ &\quad + 2(y-x)\,(\,I(y>x>0) - I(y<x<0)\,), \end{aligned} \tag{3.3}$$

we can write the stochastic process $W_m^*(u)$ in (2.3) as follows:

$$W_m^*(u) \overset{\text{let}}{=} W_{m1}^* + W_{m2}^* \tag{3.4}$$

where
$$\begin{aligned} W_{m1}^* &= \sum_{t=2}^{m} \{\, Y_{mt}^*\,(\,I(\varepsilon_t^* < 0) - I(\varepsilon_t^* > 0)\,)\,\}, \\ W_{m2}^* &= 2\sum_{t=2}^{m} \{\,(Y_{mt}^* - \varepsilon_t^*)\,(\,I(Y_{mt}^* > \varepsilon_t^* > 0) - I(Y_{mt}^* < \varepsilon_t^* < 0)\,)\,\}, \\ Y_{mt}^* &= Y_{mt}^*(u) = a_m^{-1} u X_{t-1}^* . \end{aligned}$$

Also by (3.3), we can get

$$W(u) \stackrel{\text{let}}{=} W_1 + W_2 \qquad (\text{ let } c_i = u\beta^{i-1})$$

where $W_1 = \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \left\{ c_i\delta_k\Gamma_k^{-1/\alpha} \; (I(\varepsilon_{k,i} < 0) - I(\varepsilon_{k,i} > 0)) \right\}$,

$$W_2 = 2\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \left\{ \left(c_i\delta_k\Gamma_k^{-1/\alpha} - \varepsilon_{k,i}\right) \left(I(c_i\delta_k\Gamma_k^{-1/\alpha} > \varepsilon_{k,i} > 0) - I(c_i\delta_k\Gamma_k^{-1/\alpha} < \varepsilon_{k,i} < 0)\right)\right\}.$$

Now, in order to follow the idea of Davis *et al.* (1992), for any $M > 0$ and $\delta > 0$, we can express the first term W^*_{m1} in (3.4) as

$$\begin{aligned} W^*_{m1} &= W^*_{m1}\, I(|\varepsilon^*_t| \le M)\, I(|Y^*_{mt}| > \delta) \\ &+ W^*_{m1}\, I(|\varepsilon^*_t| > M)\, I(|Y^*_{mt}| > \delta) + W^*_{m1}\, I(|Y^*_{mt}| \le \delta) \\ &\stackrel{\text{let}}{=} W^*_{m11} + W^*_{m12} + W^*_{m13} \end{aligned}$$

Then by Corollary 3.6 of Kang (1997), we can see that

$$W^*_{m11} \xrightarrow{\text{d}} W_{1\delta M} \quad \textit{in probability} \quad (\text{ as } m \to \infty) \tag{3.5}$$

where $W_{1\delta M} = W_1\; I(|\varepsilon_{k,i}| \le M)\; I(|c_i\delta_k\Gamma_k^{-1/\alpha}| > \delta)$. Therefore, with the result of Proposition A.3 in Davis *et al.* (1992), it is not too difficult to prove that

$$W^*_{m11} \xrightarrow{\text{d}} W_1 \quad \textit{in probability} \tag{3.6}$$

as $m \to \infty$, $M \to \infty$ and $\delta \to 0$.

As for the second term W^*_{m2} in (3.4), we have that for any $\delta > 0$,

$$\begin{aligned} W^*_{m2} &= W^*_{m2}\, I(|Y^*_{mt}| > \delta) + W^*_{m2}\, I(|Y^*_{mt}| \le \delta) \\ &\stackrel{\text{let}}{=} W^*_{m21} + W^*_{m22}\,. \end{aligned}$$

Here, by the same arguments in (3.5) and (3.6), it is easy to show that as $m \to \infty$ and $\delta \to 0$,

$$W^*_{m21} \xrightarrow{\text{d}} W_2 \quad \textit{in probability}\,.$$

In the next lemma, we will state a condition that is sufficient to get the convergence in (3.2).

Lemma 3.1. Suppose that $\{X^*_{nu}\}$ is a sequence of real-valued bootstrap stochastic processes such that for each u, $X^*_{nu} \xrightarrow{\text{d}} X_u$ *in probability* as

$n \to \infty$. Suppose further that X_u and X are real-valued stochastic processes such that $X_u \xrightarrow{\text{P}} X$ as $u \to \infty$ and the distribution of X is continuous. If for every $\eta > 0$, and $\epsilon > 0$,

$$\lim_{u \to \infty} \lim_{n \to \infty} P(\, P^*(\, |X_n^* \, - \, X_{nu}^*| > \eta \,) \, > \, \epsilon \,) \, = \, 0 \tag{3.7}$$

Then, $X_n^* \xrightarrow{\text{d}} X$ *in probability.*

Proof. First notice that for all $x \in R$ and for every $\eta > 0$,

$$\begin{aligned} & P^*(\, X_n^* \le x \,) \, - \, P(\, X \le x \,) \\ & \quad \le \, P^*(\, X_{nu}^* \le x + \eta \,) \, + \, P^*(\, |X_{nu}^* - X_n^*| > \eta \,) \, - \, P(\, X \le x \,) \\ & \quad \le \, P^*(\, X_{nu}^* \le x + \eta \,) \, - \, P(\, X_u \le x + \eta \,) \, + \, P^*(\, |X_{nu}^* - X_n^*| > \eta \,) \\ & \qquad + P(\, X \le x + 2\eta \,) \, - \, P(\, X \le x \,) \, + \, P(\, |X - X_u| > \eta \,) \,. \end{aligned} \tag{3.8}$$

Similarly,

$$\begin{aligned} & P^*(\, X_n^* \le x \,) \, - \, P(\, X \le x \,) \\ & \quad \ge \, P^*(\, X_{nu}^* \le x - \eta \,) \, - \, P(\, X_u \le x - \eta \,) \, - \, P^*(\, |X_{nu}^* - X_n^*| > \eta \,) \\ & \qquad + P(\, X \le x - 2\eta \,) \, - \, P(\, X \le x \,) \, - \, P(\, |X - X_u| > \eta \,) \,. \end{aligned} \tag{3.9}$$

Thus from (3.8) and (3.9),

$$\begin{aligned} & \sup_{x \in R} |\, P^*(\, X_n^* \le x \,) \, - \, P(\, X \le x \,) \,| \\ & \quad \le \, \sup_{y \in R} |\, P^*(\, X_{nu}^* \le y \,) \, - \, P(\, X_u \le y \,) \,| \, + \, P^*(\, |X_{nu}^* - X_n^*| > \eta \,) \\ & \qquad + \, P(\, |X_u - X| > \eta \,) \, + \, \sup_{x \in R} |\, P(\, x - 2\eta \le X \le x + 2\eta \,) \,| \\ & \quad \overset{\text{let}}{=} \, \Delta_{nu} \, + \, \Gamma_{nu} \, + \, H_u \, + V_\eta \,. \end{aligned}$$

Since for each u, $X_{nu}^* \xrightarrow{\text{d}} X_u$ *in probability* as $n \to \infty$, $\Delta_{nu} \xrightarrow{\text{P}} 0$. Also, as $u \to \infty$, by (3.7) and $X_u \xrightarrow{\text{P}} X$, we have $\Gamma_{nu} + H_u \xrightarrow{\text{P}} 0$. Therefore, the result follows by letting $\eta \to 0$ because the distribution of X is continuous.

Hence, in order to prove the convergence in (3.2), it suffices to show that by Lemma 3.1, for every $\eta > 0$, and $\epsilon > 0$,

$$\lim_{\delta \to 0} \lim_{M \to \infty} \lim_{m \to \infty} P(\, P^*(\, |W_{m12}^*| > \eta \,) \, > \, \epsilon \,) \, = \, 0 \,, \tag{3.10}$$

$$\lim_{\delta \to 0} \lim_{m \to \infty} P(\, P^*(\, |W_{m13}^*| > \eta \,) \, > \, \epsilon \,) \, = \, 0 \,, \tag{3.11}$$

$$\lim_{\delta \to 0} \lim_{m \to \infty} P(\, P^*(\, |W_{m22}^*| > \eta \,) \, > \, \epsilon \,) \, = \, 0 \,. \tag{3.12}$$

In the next section, we shall concentrate our efforts on proving the convergences in (3.10), (3.11) and (3.12).

4. PROOFS OF THE MAIN RESULTS

First, we will prove Theorem 2.1.

Proof of Theorem 2.1. In order to obtain the desired result (2.4), as mentioned in the previous section, it suffices to show that (3.10), (3.11) and (3.12) hold under the condition (2.5). Further, to guarantee the moment condition met, we assume that the distribution of ε_1 has a density f which is bounded in a neighborhood of 0 as stated in Remark 1 of Section 2.
To prove (3.10), let $V_t^* = (\ I(\varepsilon_t^* < 0) - I(\varepsilon_t^* > 0)\ )\ I(|\varepsilon_t^*| > M)$ and $Y_{mt}^* = a_m^{-1} u X_{t-1}^*$. Then, we can have

$$
\begin{aligned}
P^* & \left(\left| \sum_{t=2}^{m} \{ Y_{mt}^* \ (I(\varepsilon_t^* < 0) - I(\varepsilon_t^* > 0)) \} \ I(|\varepsilon_t^*| > M) \ I(|Y_{mt}^*| > \delta) \right| > \eta \right) \\
& \le P^* \left(\bigcup_{t=2}^{m} \left\{ (|X_{t-1}^*| > a_m \delta) \bigcap (|V_t^*| > 0) \right\} \right) \\
& \le P^* \left(\bigcup_{t=2}^{m} \left\{ \left(\left| \sum_{j=0}^{r-1} \varepsilon_{t-1-j}^* \hat{\beta}_n^j \right| > \frac{a_m \delta}{2} \right) \bigcap (|V_t^*| > 0) \right\} \right) \\
& \quad + P^* \left(\bigcup_{t=2}^{m} \left\{ \left| X_{t-1}^* - \sum_{j=0}^{r-1} \varepsilon_{t-1-j}^* \hat{\beta}_n^j \right| > \frac{a_m \delta}{2} \right\} \right) \\
& \le m \ P^* \left(\left| \sum_{j=0}^{r-1} \varepsilon_{j+1}^* \hat{\beta}_n^j \right| > \frac{a_m \delta}{2} \right) P^* (\ |V_1^*| > 0 \ ) \\
& \quad + P^* \left(\max_{1 \le j \le m} |\varepsilon_j^*| > \frac{a_m \delta}{2} \frac{1 - |\hat{\beta}_n|}{|\hat{\beta}_n|^r} \right) \\
& \le \sum_{j=0}^{r-1} m \ P^* \left(\left| \varepsilon_{j+1}^* \right| > \left| \hat{\beta}_n \right|^{-j} b_j \frac{a_m \delta}{2} \right) P^* (\ |V_1^*| > 0 \ ) \\
& \quad + m \ P^* \left(|\varepsilon_1^*| > \frac{a_m \delta}{2} \frac{1 - |\hat{\beta}_n|}{|\hat{\beta}_n|^r} \right)
\end{aligned} \tag{4.1}
$$

where $b_j = a\,|\beta|^{j/2}$, $\sum_{j=0}^{r-1} b_j = 1$ and $a = (\ (1 - |\beta|^{1/2}) \ / \ (1 - |\beta|^{r/2})\ )$. Here, by the result of Proposition 3.3 in Kang (1997) and the fact that $\hat{\beta}_n \xrightarrow{\text{a.s.}} \beta$, we can show that as $m,\ n \to \infty$ and $m/n \to 0$,

$$m\ P^*\left(\left| \varepsilon^*_{j+1} \right| > \left|\hat{\beta}_n\right|^{-j}\ b_j\ \frac{a_m \delta}{2} \right) \ \xrightarrow{\text{P}}\ \left(|\beta|^{-j}\ b_j\ \frac{\delta}{2} \right)^{-\alpha},$$

$$P^*\left(|V_1^*| > 0 \right) = P^*\left(|\varepsilon_1^*| > M \right) \ \xrightarrow{\text{a.s.}}\ P\left(|\varepsilon_1| > M \right),$$

$$m\ P^*\left(|\varepsilon_1^*| > \frac{a_m \delta}{2}\ \frac{1-|\hat{\beta}_n|}{|\hat{\beta}_n|^r} \right) \ \xrightarrow{\text{P}}\ \left(\frac{\delta}{2}\ \frac{1-|\beta|}{|\beta|^r} \right)^{-\alpha}. \tag{4.2}$$

Thus by (4.1) and (4.2), we can prove that

$$\limsup_{m\to\infty}\ P^*\left(\left| \sum_{t=2}^{m} \{\ Y^*_{mt}\ (\ I(\varepsilon^*_t < 0) - I(\varepsilon^*_t > 0)\ )\ \} I(|\varepsilon^*_t| > M) I(|Y^*_{mt}| > \delta) \right| > \eta \right)$$
$$\le\ \sum_{j=0}^{r-1} \left(|\beta|^{-j}\ b_j\ \frac{\delta}{2} \right)^{-\alpha}\ P\left(|\varepsilon_1| > M \right) + \left(\frac{\delta}{2}\ \frac{1-|\beta|}{|\beta|^r} \right)^{-\alpha} \quad \textit{in probability}$$

Therefore, the result follows by letting $r \to \infty$ and $M \to \infty$ because by the fact that $|\beta| < 1$, we can see that $|\beta|^r \to 0$, $\sum_{j=0}^{r-1} (\ |\beta|^{-j}\ b_j\ \delta/2\ )^{-\alpha} \le Const \times (\ 1 - |\beta|^{r/2}\ )^{\alpha} \ \to\ Const$ as $r \to \infty$ and $P(\ |\varepsilon_1| > M\ ) \to 0$ as $M \to \infty$.

As for the equation in (3.11), we have that for all $\eta > 0$,

$$P^*\left(\left| \sum_{t=2}^{m} \{\ Y^*_{mt}\ (\ I(\varepsilon^*_t < 0) - I(\varepsilon^*_t > 0)\ )\ \}\ I(\ |Y^*_{mt}| \le \delta\ ) \right| > \eta \right)$$
$$\le\ P^*\left(\left| \sum_{t=2}^{m} (\ Y^*_{mt}\ I(\ |Y^*_{mt}| \le \delta\ )\ )\ \{\ (\ I(\varepsilon^*_t < 0) - I(\varepsilon^*_t > 0)\ ) \right.\right.$$
$$\left.\left. -\ (\ 1 - 2P^*(\varepsilon^*_t > 0)\ )\ \}\ \right| > \frac{\eta}{2} \right)$$
$$+\ P^*\left(\left| (\ 1 - 2P^*(\varepsilon^*_t > 0)\ ) \sum_{t=2}^{m} (\ Y^*_{mt}\ I(\ |Y^*_{mt}| \le \delta\ )\ ) \right| > \frac{\eta}{2} \right)$$
$$\stackrel{\text{let}}{=}\ H_1\ +\ H_2$$

So, it suffices to show that $H_1 \xrightarrow{\text{P}} 0$ and $H_2 \xrightarrow{\text{P}} 0$ as $m \to \infty$, $\delta \to 0$ and under the condition (2.5). However, with the condition (2.5) and the boundedness of the density function f of error terms in a neighborhood of 0 and $E(|\varepsilon_1|^r) < \infty$ for any r such that $1 < r < \alpha < 2$, it is not too difficult to show that $H_1 \xrightarrow{\text{P}} 0$ and $H_2 \xrightarrow{\text{P}} 0$. Hence, the proof is omitted for the sake of clarity.

Now, we need to prove (3.12). First, we notice that for every $\eta > 0$,

$$P^*\left(\left|\sum_{t=2}^{m}\{(Y^*_{mt}-\varepsilon^*_t)\,(\,I(Y^*_{mt}>\varepsilon^*_t>0)-I(Y^*_{mt}<\varepsilon^*_t<0)\,)\}\,I(|Y^*_{mt}|\le\delta)\right|>\eta\right)$$
$$= P^*\left(\sum_{t=2}^{m}(Y^*_{mt}-\varepsilon^*_t)\;I(\,0<\varepsilon^*_t<Y^*_{mt}\le\delta\,)\;>\;\eta\right)$$
$$+\;P^*\left(\sum_{t=2}^{m}(Y^*_{mt}-\varepsilon^*_t)\;I(\,-\delta\le Y^*_{mt}<\varepsilon^*_t<0\,)\;<\;-\eta\right)$$
$$\overset{\text{let}}{=}\;\Delta_1\;+\;\Delta_2$$

Again, with the condition (2.5) and the fact that the density function f of the distribution of error trems is bounded in a neighborhood of 0, it is also not too difficult to prove that $\Delta_1 \xrightarrow{P} 0$ and $\Delta_2 \xrightarrow{P} 0$. Thus, we shall omit the proof for the brevity of this paper and this completes the proof of Theorem 2.1.

Before proving Lemma 2.1, we need the following lemma.

Lemma 4.1. Let $U_1,\ldots,U_n$ and $V_1,\ldots,V_n$ be sequences of real numbers such that $V_i - h \le U_i \le V_i + h$ for all $1 \le i \le n$ and for some $h > 0$. Suppose $U_{(1)},\ldots,U_{(n)}$ and $V_{(1)},\ldots,V_{(n)}$ be the ordered values of $U_1,\ldots,U_n$ and $V_1,\ldots,V_n$, respectively. Then, for all $1 \le j \le n$,

$$V_{(j)} - h \;\le\; U_{(j)} \;\le\; V_{(j)} + h\,.$$

Proof. Suppose $V_{(j)} = V_i$ for some i and $V_{i_1},\ldots,V_{i_j}$ be such that

$$V_{i_k} \;\le\; V_i \qquad \text{for} \quad k = 1,\ldots,j\,.$$

Then,
$$\begin{aligned} V_{(j)} + h &= \max\left\{V_{i_1},\ldots,V_{i_j}\right\} + h \\ &= \max\left\{V_{i_1}+h,\ldots,V_{i_j}+h\right\} \\ &\ge \max\left\{U_{i_1},\ldots,U_{i_j}\right\} = U_{(j)}\,. \end{aligned} \tag{4.3}$$

Similarly, we have

$$V_{(j)} - h \;\le\; \max\left\{U_{i_1},\ldots,U_{i_j}\right\} = U_{(j)}\,. \tag{4.4}$$

Therefore, Lemma 4.1 follows from (4.3) and (4.4).

Proof of Lemma 2.1. Let $|\varepsilon|_{(1)}, \ldots, |\varepsilon|_{(n)}$ be the order statistics of $|\varepsilon_1|, \ldots, |\varepsilon_n|$. Then, notice that for all $1 \leq t \leq n$,

$$
\begin{aligned}
|\hat{\varepsilon}_t| &\leq |\varepsilon_t| + \left|\hat{\beta}_n - \beta\right| \max_{1\leq i\leq n} |\varepsilon_i| \frac{1}{1-|\beta|}, \\
|\hat{\varepsilon}_t| &\geq |\varepsilon_t| - \left|\hat{\beta}_n - \beta\right| \max_{1\leq i\leq n} |\varepsilon_i| \frac{1}{1-|\beta|}.
\end{aligned}
$$

Thus, we can get by Lemma 4.1,

$$
\begin{aligned}
\frac{|\hat{\varepsilon}|_{(n-[\frac{n}{m}]+1)}}{a_m} &\leq \frac{|\varepsilon|_{(n-[\frac{n}{m}]+1)}}{a_m} + \frac{1}{a_m}\left|\hat{\beta}_n - \beta\right| \max_{1\leq i\leq n} |\varepsilon_i| \frac{1}{1-|\beta|}, \\
\frac{|\hat{\varepsilon}|_{(n-[\frac{n}{m}]+1)}}{a_m} &\geq \frac{|\varepsilon|_{(n-[\frac{n}{m}]+1)}}{a_m} - \frac{1}{a_m}\left|\hat{\beta}_n - \beta\right| \max_{1\leq i\leq n} |\varepsilon_i| \frac{1}{1-|\beta|}.
\end{aligned}
$$

Here, by the fact that $a_n(\hat{\beta}_n - \beta) = O_p(1)$ and $\max_{1\leq i\leq n} |\varepsilon_i| / a_n = O_p(1)$ and $a_m \to \infty$ as $m \to \infty$, it suffices to show that

$$
\frac{|\varepsilon|_{(n-[\frac{n}{m}]+1)}}{a_m} \xrightarrow{P} 1 \tag{4.5}
$$

as $m, n \to \infty$ in such a way that $m/n \to 0$. To prove (4.5), let $x > 0$ be a real number. Then, we have

$$
\begin{aligned}
P\left(\frac{1}{a_m}|\varepsilon|_{(n-[\frac{n}{m}]+1)} > x\right) &= P\left(\sum_{t=1}^{n} I(|\varepsilon_t| > a_m x) \geq \frac{n}{m}\right) \\
&= P\left(\frac{m}{n}\sum_{t=1}^{n} I(|\varepsilon_t| > a_m x) - mP_m + mP_m \geq 1\right)
\end{aligned}
$$

where $P_m = P(|\varepsilon_1| > a_m x)$. Here, since it is easy to show that

$$
\frac{m}{n}\sum_{t=1}^{n} I(|\varepsilon_t| > a_m x) - mP_m \xrightarrow{P} 0
$$

as $m, n \to \infty$ in such a way that $m/n \to 0$, and from the definition of a_m,

$$
mP_m \longrightarrow x^{-\alpha} \quad \text{for all } x > 0 \quad \text{as } m \to \infty,
$$

we can easily obtain that for $1 < \alpha < 2$,

$$
\begin{aligned}
&P\left(\frac{1}{a_m}|\varepsilon|_{(n-[\frac{n}{m}]+1)} > x\right) \\
&= P\left(\frac{m}{n}\sum_{t=1}^{n} I(|\varepsilon_t| > a_m x) - mP_m + mP_m \geq 1\right) \to \begin{cases} 1 & \text{if } x \leq 1 \\ 0 & \text{if } x > 1. \end{cases}
\end{aligned}
$$

Hence, we proved (4.5) and this completes the proof of Lemma 2.1.

Proof of Theorem 2.2. This follows from Theorem 2.1 and Lemma 2.1.

5. SIMULATION STUDIES

To verify the validity of the bootstrap method, we simulated the AR(1) process with $\beta = .7$ and $n = 200$ in (1.1). For generating the stable random error terms with index $1 < \alpha < 2$, see Gross and Steiger (1979). We also refer to Bloomfield and Steiger (1980) for the algorithm of calculating the LAD estimator. To support the theoretical results, we present Monte Carlo results that calculate the empirical coverage probability of 95% confidence interval for the parameter of the AR(1) process with infinite variance. The bootstrap distribution of the LAD estimator was used for the upper and lower bound of the confidence interval so that the theoretical validity of the bootstrap method would be verified by comparing the empirical coverage probabilitiy with the nominal 95% confidence level because the limiting distribution of the LAD estimator is not known. Table 1 shows the simulation results using

Table 1. Empirical coverage probabilities of the 95% confidence interval for the parameter β with various α.

	$n = 200$		
	$m = 20$	$m = 40$	$m = 200$
$\alpha = 1.8$	0.917	0.921	0.925
	0.921	0.923	0.934
	0.925	0.927	0.936
	0.929	0.929	0.937
	0.932	0.931	0.938
$\alpha = 1.5$	0.941	0.942	0.952
	0.943	0.946	0.955
	0.946	0.947	0.957
	0.947	0.948	0.961
	0.953	0.950	0.966
$\alpha = 1.2$	0.952	0.954	0.969
	0.955	0.956	0.970
	0.957	0.959	0.973
	0.958	0.960	0.976
	0.961	0.962	0.977

1000 bootstrap replications and 1000 iterations for the empirical coverage probability. As expected, the bootstrap method is good enough to reach the nominal 95% confidence level whenever the bootstrap sample size m is much smaller than the original sample size n.

REFERENCES

(1) Bloomfield, P. and Steiger, W. L. (1980). Least Absolute Deviations Curve-Fitting, *SIAM Journal on Scientific and Statistical Computing*, **1**, 290-301.

(2) Bose, A. (1988). Edgeworth correction by bootstrap in autoregressions, *The Annals of Statistics*, **16**, No. 4, 1709-1722.

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

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

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

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

(7) Kang, H. J. (1997). On Asymptotic Properties of Bootstrap for Autoregressive Processes with Regularly Varying Tail Probabilities, *Journal of the Korean Statistical Society*, **26**, No. 1, 31-45.

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