Journal of the Korean Statistical Society (1999), **28***: 2, pp 151–165*

Limiting Distributions of Trimmed Least Squares Estimators in Unstable AR(1) Models

Sangyeol Lee[1]

ABSTRACT

This paper considers the trimmed least squares estimator of the autoregression parameter in the unstable AR(1) model: $X_t = \phi X_{t-1} + \varepsilon_t$, where ε_t are iid random variables with mean 0 and variance $\sigma^2 > 0$, and ϕ is the real number with $|\phi| = 1$. The trimmed least squares estimator for ϕ is defined in analogy of that of Welsh (1987). The limiting distribution of the trimmed least squares estimator is derived under certain regularity conditions.

Keywords: An unstable AR(1) model; Robust estimation; Trimmed least squares estimator, Weak convergence.

1. INTRODUCTION

Consider the first order autoregressive model:

$$X_j = \phi X_{j-1} + \varepsilon_j, \tag{1.1}$$

where ε_j are iid random variables with mean 0 and variance $\sigma^2 > 0$. It is well-known that the parameter ϕ determines the characters of $\{X_t\}$. For example, if $|\phi| < 1$, $\{X_t\}$ is stationary; if $|\phi| > 1$, $\{X_t\}$ is explosive provided the initial random variable X_0 is given; if $|\phi| = 1$, the process is unstable. Moreover, the statistical properties of the least squares estimator of ϕ is well-studied in the literature. White (1958) claimed that under the normal assumption on ε_t, the least squares estimator ϕ_{LS} will have a limiting distribution of the random variable that is a functional of a standard Brownian motion, namely,

$$n(\phi_{LS} - \phi) \xrightarrow{d} \frac{1}{2}(B^2(1) - 1) / \int_0^1 B^2(t)dt, \tag{1.2}$$

where B denotes a standard Brownian motion. The above is proved in a more general setting by Chan and Wei (1988). As related literature, we refer to Rao (1978) and Dickey and Fuller (1979).

[1]Department of Statistics, Seoul National University, Seoul, 151-742, Korea

Although the least squares estimator has a variety of decent properties, it is well-known that the estimator has a drawback like the sensitivity to outliers. To overcome such a problem, it is common to consider employing a robust estimator. In stationary time series, a great deal of results exist on robust estimation (cf. Martin and Yohai (1986)). Also, in nonstationary time series, there are papers concerning robust estimation. For example, Knight (1991) and Herce (1996) considered M and LAD (least absolute deviation) estimators in random walk models. However, despite of its popularity, the trimmed least squares estimator did not get a full attention from researchers. This motivates us to study the trimmed least squares estimator for the autoregression parameter ϕ, when $\{X_t\}$ in (1.1) is unstable.

The trimmed mean has long been used as a robust estimator for a location in iid sample since it is easy to compute and understand and works well on real data. Ruppert and Carroll (1980) and Welsh (1987) generalized the trimmed mean of iid sample to the linear regression model. Their idea is to construct the least squares estimator based on the observations whose corresponding residuals, computed based on a preliminary estimator, lie between the $[n\alpha]$th and $[n(1-\beta)]$th largest residuals, where $0 < \alpha < 1/2 < \beta < 1$ are trimming proportions. According to our analysis (cf. Theorem 1.1), the same approach of Welsh remains vaild in our setting. Below, we define the trimmed least squares estimator. Later on, its limiting distribution will be investigated.

Suppose that $X_1, \ldots, X_n$ are available observations. Further assume that $X_0 = 0$. Let ϕ_n be any preliminary estimator of ϕ, such that

$$n(\phi_n - \phi) = O_P(1).$$

A typical example of such an estimator is the least squares estimator (cf. (1.2)). Define the residuals, based on ϕ_n,

$$e_j = X_j - \phi_n X_{j-1}, \quad j = 1, \ldots, n. \tag{1.3}$$

Let $e_{n1}, \ldots, e_{nn}$ denote the ordered random variables of $e_1, \ldots, e_n$. Put

$$\xi_{nq} = \begin{cases} e_{n,nq} & , \ nq \text{ is an integer} \\ e_{n,[nq]+1} & , \ \text{otherwise.} \end{cases}$$

Let α, β be the real numbers such that $0 < \alpha < 1/2 < \beta < 1$, and let

$$\begin{aligned} J_j &= I(e_j \le \xi_{n\alpha}), \\ K_j &= I(\xi_{n\alpha} < e_j \le \xi_{n\beta}), \\ L_j &= I(e_j > \xi_{n\beta}). \end{aligned}$$

Define the α, β trimmed least squares estimator as follows:

$$\hat{\phi}_{n,\alpha,\beta} = \frac{\sum_{j=1}^{n} X_{j-1}\left[\xi_{n\alpha}(J_j - \alpha) + X_j K_j + \xi_{n\beta}(L_j - (1-\beta))\right]}{\sum_{j=1}^{n} X_{j-1}^2 K_j}. \tag{1.4}$$

To obtain the limiting distribution of $\hat{\phi}_{n,\alpha,\beta}$, we assume that

(R): The common distribution F of ε_j has the positve and continuous density f.

Before we state the main theorem, we introduce some notations. Let ξ_q be the number such that $F(\xi_q) = q$, and let

$$\psi_{\alpha,\beta}(x) = \left[\xi_\alpha(I(x \le \xi_\alpha) - \alpha) + xI(\xi_\alpha < x \le \xi_\beta) + \xi_\beta(I(x > \xi_\beta) - (1-\beta))\right].$$

Observe that $\psi_{\alpha,\beta}(x)/(\beta - \alpha) - T(F)$, where

$$T(F) = (\beta - \alpha)^{-1} \int_{\xi_\alpha}^{\xi_\beta} x dF(x),$$

is the influence curve for the trimmed mean in iid setting.

Set $\tau^2 = Var(\psi_{\alpha,\beta}(\varepsilon_j))$, i.e.,

$$\begin{aligned} \tau^2 \ &= \ \xi_\alpha^2 \alpha(1-\alpha) + \int_{\xi_\alpha}^{\xi_\beta} x^2 dF(x) + \xi_\beta^2 \beta(1-\beta) \\ &\quad -2\{\alpha\xi_\alpha + (1-\beta)\xi_\beta\} \int_{\xi_\alpha}^{\xi_\beta} x dF(x) - 2\xi_\alpha \xi_\beta \alpha(1-\beta). \end{aligned} \tag{1.5}$$

Further, let B, B^*, W denote three standard Brownian motions, such that for all $s, t \in [0,1]$,

$$Cov(B(s), W(t)) = \frac{s \wedge t}{\sigma\tau}\{\xi_\alpha \int_{-\infty}^{\xi_\alpha} x dF(x) + \int_{\xi_\alpha}^{\xi_\beta} x^2 dF(x) + \xi_\beta \int_{\xi_\beta}^{\infty} x dF(x)\}. \tag{1.6}$$

and

$$Cov(B(s), B^*(t)) = 0 = Cov(B^*(s), W(t)). \tag{1.7}$$

Actually, we need B, B^*, W because the trimmed least squares estimator is expressed as a functional of

$$B_n(t) = \frac{1}{n^{1/2}\sigma} \sum_{j=1}^{[nt]} \varepsilon_j, \quad B_n^*(t) = \frac{1}{n^{1/2}\sigma} \sum_{j=1}^{[nt]} (-1)^j \varepsilon_j,$$

$$\text{and} \quad W_n(t) = \frac{1}{n^{1/2}\tau} \sum_{j=1}^{[nt]} \{\psi_{\alpha,\beta}(\varepsilon_j) - (\beta - \alpha)T(F)\} \tag{1.8}$$

(see the proof of Theorem 1.1 in Section 3). It is easy to see that (B_n, B_n^*, W_n) converges weakly to (B, B^*, W) in $D^3[0,1]$ space (cf. Chan and Wei, 1988, Theorem 2.2).

The following is the main result of this section.

Theorem 1.1. *Suppose that Condition (R) holds. Then,*

$$n(\hat{\phi}_{n,\alpha,\beta} - \phi) = \frac{n^{-1}\sum_{j=1}^{n} X_{j-1}\psi_{\alpha,\beta}(\varepsilon_j)}{(\beta-\alpha)n^{-2}\sum_{j=1}^{n} X_{j-1}^2} + o_P(1). \tag{1.9}$$

Therefore,

(i) when $\phi = 1$,

$$n(\hat{\phi}_{n,\alpha,\beta} - \phi) \xrightarrow{d} \frac{\sigma\tau\int_0^1 B(t)dW(t) + (\beta-\alpha)T(F)\sigma\int_0^1 B(t)dt}{(\beta-\alpha)\sigma^2\int_0^1 B^2(t)dt}; \tag{1.10}$$

(ii) when $\phi = -1$,

$$n(\hat{\phi}_{n,\alpha,\beta} - \phi) \xrightarrow{d} \frac{\sigma\tau\int_0^1 B^*(t)dW(t) + (\beta-\alpha)T(F)\sigma\int_0^1 B^*(t)dt}{(\beta-\alpha)\sigma^2\int_0^1 (B^*(t))^2dt}. \tag{1.11}$$

Remark 1.1. Theorem 1.1 shows that the limiting distribution of the trimmed least squares estimator does not rely on that of a preliminary estimator for ϕ. If $\alpha = 0, \beta = 1$, $\phi_{n,\alpha,\beta}$ coincides with the ordinary least squares estimator. In this case, when $\phi = 1$, the limiting distribution of the trimmed least squares in (1.10) is the same as in (1.2).

The proof of Theorem 1.1 is provided in Section 3. In proving Theorem 1.1, verifying the following asymptotic result is very crucial: for any $K > 0$,

$$\sup_{|x|\le K} |n^{-1}\sum_{j=1}^{n}\{X_{j-1}[\varepsilon_j I(\varepsilon_j \le x + (\phi_n - \phi)X_{j-1}) - H(x + (\phi_n - \phi)X_{j-1}) + H(x) - \varepsilon_j I(\varepsilon_j \le x)]\}| = o_P(1), \tag{1.12}$$

where $H(x) = E\varepsilon_j I(\varepsilon_j \le x)$. The argument (1.12) is relevant to the oscillation problem in randomly weighted empirical processes. Boldin (1982) and Koul (1992, ch. 7) handled similar problems in stationary autoregressive models. Recently, Lee and Wei (1999) considered the oscillation problem in residual empirical processes from AR(p) unstable processes.

In Section 2, we will verify that (1.12) holds in our setting (cf. Lemmas 2.1 and 2.2). Our lemmas are formulated not only to deal with the unstable AR(1) process but also to cover more general cases.

2. PRELIMINARY RESULTS

In this section, as mentioned earlier, we present some preliminary results needed for proving Theorem 1.1 in Section 1.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Suppose that ε_j are iid random variables with mean 0 and variance σ^2, $\{\mathcal{F}_{nj}; 1 \le j \le n\}$ is a double array of sub σ-fields of $\mathcal{F}$, such that $\mathcal{F}_{nj} \subset \mathcal{F}_{n,j+1}$ for all n, j, and Y_{nj} are $\mathcal{F}_{n,j-1}$ measurable random variables.

Lemma 2.1. *Suppose that*

$$\max_{1\le j\le n} |Y_{nj}| = O_P(a_n), \tag{2.1}$$

$$\sum_{j=1}^{n} Y_{nj}^2 = O_P(b_n), \tag{2.2}$$

where a_n, b_n are positive real numbers. If a sequence of positive real numbers $\{c_n\}$ and a positive real number γ satisfies

$$c_n \to \infty, \quad (nb_n)^{1/2}/c_n = O(n^{\nu_1}) \text{ for some } \nu_1 > 0, \tag{2.3}$$

and

$$\left(\frac{n^{\gamma} c_n^2}{a_n c_n + b_n}\right) / n^{\nu_2} \to \infty \text{ for some } \nu_2 > 0, \tag{2.4}$$

then for any $K, M > 0$,

$$\Gamma_n := \sup_{\{(x,y)\in S_n\}} c_n^{-1} \Big| \sum_{j=1}^{n} \{Y_{nj}[\varepsilon_j I(\varepsilon_j \le x) - H(x) + H(y) - \varepsilon_j I(\varepsilon_j \le y)] \Big| = o_P(1), \tag{2.5}$$

where

$$S_n = \{(x,y) \in R^2; |x-y| \le Kn^{-\gamma}, |x| \le M, |y| \le M\}, \quad H(x) = E\varepsilon_1 I(\varepsilon_1 \le x).$$

Proof: Note that Γ_n is bounded by $\Gamma_n^{++} + \Gamma_n^{+-} + \Gamma_n^{-+} + \Gamma_n^{--}$, where Γ_n^{++} is the random variable that is the same as Γ_n with Y_{nj}, ε_j, H being replaced by $Y_{nj}^+, \varepsilon_j^+, H^+$, where $H^+(x) = E\varepsilon_j^+ I(\varepsilon_j \le x)$, and the other random variables are similarly defined. To prove (2.5) we have to show that each of the four terms is $o_P(1)$. Here, we only provide the proof for $\Gamma_n^{++} = o_P(1)$ since the other cases can be handled similarly.

Partition $[-M, M]$ with the points

$$x_{nr} = -M + 2Mr/N_n, \quad r = 0, \ldots, N_n,$$

where $N_n = n^\lambda$ and λ is a positive integer bigger than γ and $\nu_1 + 1$. Assume that $(x, y) \in S_n$ and in addition $x \in [x_{nr}, x_{n,r+1})$ and $y \in [x_{nl}, x_{n,l+1})$. Note that $|x_{n,l+1} - x_{nr}|$ and $|x_{n,r+1} - x_{n,l}|$ are bounded by $\tilde{K}n^{-\gamma}$ for some $\tilde{K} > 0$, and that for $i = r, r+1$ and $j = l, l+1$, $|H^+(x) - H^+(x_{ni})| \leq K^* N_n^{-1}$ and $|H^+(y) - H^+(x_{nj})| \leq K^* N_n^{-1}$ for some $K^* > 0$. By using the monotonicity property of the indicator function, we can write that $\Gamma_n^{++} \leq \Gamma_n^* + R_n$, where

$$\begin{aligned}\Gamma_n^* &= \sup_{|x_{nl} - x_{nr}| \leq \tilde{K}n^{-\gamma}} c_n^{-1} \Big| \sum_{j=1}^n \{ Y_{nj}^+ [\varepsilon_j^+ I(\varepsilon_j \leq x_{nr}) - H^+(x_{nr}) + H^+(x_{nl}) \\ &\quad - \varepsilon_j^+ I(\varepsilon_j \leq x_{nl})]\} \Big|\end{aligned}$$

and R_n is a random variable which is $O_P((nb_n)^{1/2}/c_n N_n)$. Since R_n is $o_P(1)$ by (2.3), we only have to prove $\Gamma_n^* = o_P(1)$.

In view of (2.1) and (2.2), it suffices to show that for any $A, B > 0$,

$$\Gamma_n^{**} := \Gamma_n^* I(\max_{1 \leq i \leq n} |Y_{ni}| \leq Aa_n, \sum_{i=1}^n Y_{ni}^2 \leq Bb_n) = o_P(1). \tag{2.6}$$

Define

$$\begin{aligned} d_j &= Y_{nj}^+ [\varepsilon_j^+ I(\varepsilon_j \leq x_{nr}) - H^+(x_{nr}) + H^+(x_{nl}) - \varepsilon_j^+ I(\varepsilon_j \leq x_{nl})] \\ &\quad I(\max_{1 \leq i \leq n} |Y_{ni}| \leq Aa_n, \sum_{i=1}^n Y_{ni}^2 \leq Bb_n), \\ d_j' &= d_j I(\max_{1 \leq i \leq j} |Y_{ni}| \leq Aa_n, \sum_{i=1}^j Y_{ni}^2 \leq Bb_n). \end{aligned}$$

Observe that

$$P(d_j \neq d_j' \text{ for some } j = 1, \ldots, n) = 0. \tag{2.7}$$

Further, $\{d_j', \mathcal{F}_j\}$ is a sequence of martingale differences such that $|d_j'| \leq A_1 a_n$ for some $A_1 > 0$, and

$$\sum_{j=1}^n E((d_j')^2 | \mathcal{F}_{j-1}) \leq B_1 b_n n^{-\gamma}, \quad \text{for some } B_1 > 0.$$

Applying Bernstein's inequality for martingales (cf. Shorack and Wellner, 1986, P. 809), we obtain that for $\delta > 0$,

$$P(|\sum_{j=1}^{n} d_j'| \geq c_n\delta) \leq 2\exp\{-c_n^2\delta^2/2(B_1b_nn^{-\gamma}+A_1a_nc_n\delta/3)\} \leq \exp(-\theta n^{\nu_2}),\ \theta > 0,$$

where the last inequality is due to (2.4). Combining this and (2.7), we have that

$$P(\Gamma_n^{**} \geq \delta) \leq 4n^{2\lambda}\exp(-\theta n^{\nu_2}) \to 0.$$

This proves (2.6). □

Lemma 2.2. *Assume that a_n, b_n, c_n satisfy Conditions (2.1)-(2.4). Further, suppose that the sequence of positive real numbers $\{\rho_n\}$ satisfies*

$$\frac{c_n^{-1}(b_n\rho_n^{-1} + (b_nn)^{1/2})}{n^{\zeta_1}} \to 0 \text{ for some } \zeta_1 > 0, \tag{2.8}$$

and

$$\frac{\rho_nc_n^2}{(a_nb_n + a_n^2c_n)n^{\zeta_2}} \to \infty \text{ for some } \zeta_2 > 0. \tag{2.9}$$

Then, for any $L, M > 0$,

$$\begin{aligned}\Lambda_n \ :=\ & \sup_{|x|\leq M, |s|\leq L} |c_n^{-1}\sum_{j=1}^{n}\{Y_{nj}[\varepsilon_jI(\varepsilon_j \leq s\rho_n^{-1}Y_{nj} + x) - H(s\rho_n^{-1}Y_{nj} + x)\\ & +H(x) - \varepsilon_jI(\varepsilon_j \leq x)]\}| = o_P(1).\end{aligned}$$

Proof: As we did in the proof of Lemma 2.1, we only provide the proof for $\Lambda_n^{++} = o_P(1)$, where

$$\begin{aligned}\Lambda_n^{++} \ :=\ & \sup_{|x|\leq M, |s|\leq L} |c_n^{-1}\sum_{j=1}^{n}\{Y_{nj}^+[\varepsilon_j^+I(\varepsilon_j \leq s\rho_n^{-1}Y_{nj} + x) - H^+(s\rho_n^{-1}Y_{nj} + x)\\ & +H^+(x) - \varepsilon_j^+I(\varepsilon_j \leq x)]\}| = o_P(1).\end{aligned} \tag{2.10}$$

Set

$$x_{nr} = -M + 2Mr/N_n, \quad s_{nl} = -L + 2Ll/N_n, \quad r, l = 0, \ldots, N_n,$$

where $N_n = [n^\lambda]$ for some $\lambda > \max\{\zeta_1, \gamma\}$. For $x \in [x_{nr}, x_{n,r+1})$, $s \in I_{nl} = [s_{nl}, s_{n,l+1})$, put

$$\Delta_{nj}^+ = \sup_{s\in I_{nl}} s\rho_n{}^{-1}Y_{nj}, \quad \Delta_{nj}^- = \inf_{s\in I_{nl}} s\rho_n^{-1}Y_{nj}.$$

Note that

$$\begin{aligned}
& c_n^{-1}Y_{nj}^+[\varepsilon_j^+ I(\varepsilon_j \le s\rho_n^{-1}Y_{nj}+x) - H^+(s\rho_n^{-1}Y_{nj}+x) + H^+(x) \\
& -\varepsilon_j^+ I(\varepsilon_j \le x_{n,r+1})] \\
\le\ & c_n^{-1}Y_{nj}^+[\varepsilon_j^+ I(\varepsilon_j \le \Delta_{nj}^+ + x_{n,r+1}) - H^+(\Delta_{nj}^+ + x_{n,r+1}) + H^+(x_{n,r+1}) \\
& -\varepsilon_j^+ I(\varepsilon_j \le x_{n,r+1})] + c_n^{-1}Y_{nj}^+\{H^+(\Delta_{nj}^+ + x_{n,r+1}) - H^+(s\rho_n^{-1}Y_{nj}+x)\} \\
& +c_n^{-1}Y_{nj}^+[\varepsilon_j^+ I(\varepsilon_j \le x_{n,r+1}) - H^+(x_{n,r+1}) + H^+(x) - \varepsilon_j^+ I(\varepsilon_j \le x)].
\end{aligned} \tag{2.11}$$

Similarly, the LHS of (2.11) is bounded from below by the same of the RHS of (2.11) with $\Delta_{nj}^+, x_{n,r+1}$ being replaced by Δ_{nj}^-, x_{nr}. Since the sum of the absolute value of the second term in (2.11) is bounded by

$$\begin{aligned}
& c_n^{-1}\sum_{j=1}^n Y_{nj}^2 O(N_n^{-1}\rho_n^{-1}) + c_n^{-1}\sum_{j=1}^n |Y_{nj}| O(N_n^{-1}) \\
=\ & O_P(c_n^{-1}b_n N_n^{-1}\rho_n^{-1} + c_n^{-1}b_n^{1/2}n^{1/2}N_n^{-1}) = o_P(1),
\end{aligned}$$

where the last equality follows from (2.8), and since the sum of the absolute value of the third term is no more than

$$\sup_{|x-y|\le Kn^{-\lambda}} |c_n^{-1}\sum_{j=1}^n Y_{nj}^+[\varepsilon_j^+ I(\varepsilon_j \le x) - H^+(x) + H^+(y) - \varepsilon_j^+ I(\varepsilon_j \le y)]|,$$

which is $o_P(1)$ due to Lemma 2.1, we can write that $\Lambda_n \le \Lambda_n^{++} + \Lambda_n^{+-} + \Lambda_n^{+-} + R_n$, where

$$\begin{aligned}
\Lambda_n^{++} =\ & \max_{r,s} |c_n^{-1}\sum_{j=1}^n Y_{nj}^+[\varepsilon_j^+ I(\varepsilon_j \le \Delta_{nj}^+ + x_{nr}) - H^+(\Delta_{nj}^+ + x_{nr}) \\
& +H^+(x_{nr}) - \varepsilon_j^+ I(\varepsilon_j \le x_{nr})]|,
\end{aligned}$$

Λ_n^{+-} is the same as Λ_n^{++} with Δ_{nj}^+ being replaced by Λ_{nj}^-, and R_n is a random variable that is $o_P(1)$. Hence, to verify (2.10), it suffices to show $\Lambda_n^{++} = o_P(1)$ and $\Lambda_n^{+-} = o_P(1)$. Here we only prove the former because the latter can be handled similarly.

If we put

$$\begin{aligned}
d_j =\ & Y_{nj}^+[\varepsilon_j^+ I(\varepsilon_j \le \Delta_{nj}^+ + x_{nr}) - H^+(\Delta_{nj}^+ + x_{nr}) + H^+(x_{nr}) - \varepsilon_j^+ I(\varepsilon_j \le x_{nr})] \\
& \times I(\max_{1\le i\le n} |Y_{ni}| \le Aa_n, \sum_{i=1}^n Y_{ni}^2 \le Bb_n), \quad A, B > 0,
\end{aligned}$$

in view of (2.1) and (2.2), we will have $\Lambda_n^{++} = o_P(1)$ as long as $\max_{r,s} |c_n^{-1} \sum_{j=1}^n d_j| = o_P(1)$ for any $A, B > 0$. Here, as we did before in proving Lemma 2.1, consider

$$d_j' = d_j I(\max_{1 \le i \le j} |Y_{ni}| \le A a_n, \sum_{i=1}^j Y_{ni}^2 \le B b_n).$$

It can be seen that $\{d_j', \mathcal{F}_j\}$ forms a sequence of martingale differences with $|d_j'| \le \delta_1 a_n^2 \rho_n^{-1}$ for some $\delta_1 > 0$, and $\sum_{j=1}^n E((d_j')^2 | \mathcal{F}_{j-1}) \le \delta_2 b_n \rho_n^{-1} a_n$ for some $\delta_2 > 0$. Applying Bernstein's inequality for martingales, we obtain that for $z > 0$,

$$P(\max_{r,s} |c_n^{-1} \sum_{j=1}^n d_j'| \ge z) \le (N_n + 1)^2 \exp\{-c_n^2 z^2 / 2(\delta_2 a_n b_n \rho_n^{-1} + a_n^2 c_n \rho_n^{-1} z/3)\},$$

which goes to 0 by (2.9). Since d_j, d_j' satisfy (2.7), we assert (2.10). □

3. PROOFS

We start this section with a Bahadur type representation for the residuals $e_1, \ldots, e_n$ given in (1.3). Throughout, Λ_{nj} denotes $(\phi_n - \phi) X_{j-1}$.

Theorem 3.1. *Let F_n denote the empirical distribution based on $\varepsilon_1, \ldots, \varepsilon_n$. Then for any $\alpha \in (0, 1)$,*

$$n^{1/2}(\xi_{n\alpha} - \xi_\alpha) = n^{1/2}(\alpha - F_n(\xi_\alpha))/f(\xi_\alpha) - n^{-1/2} \sum_{j=1}^n \Lambda_{nj} + o_P(1). \quad (3.1)$$

Particularly,

$$\xi_{n\alpha} - \xi_\alpha = O_P(n^{-1/2}). \quad (3.2)$$

Proof: To establish (3.1), we will adopt the idea of Ghosh (1971). Put

$$\hat{F}_n(x) = n^{-1} \sum_{j=1}^n I(e_j \le x), \hat{G}_n(x) = 1 - \hat{F}_n(x), G_n(x) = 1 - F_n(x), G(x) = 1 - F(x).$$

Let $V_n = n^{1/2}(\xi_{n\alpha} - \xi_\alpha)$. Note that for any t,

$$\begin{aligned} (V_n \le t) &= (\alpha \le \hat{F}_n(n^{-1/2} t + \xi_\alpha)) \\ &= \left(\frac{n^{1/2}(\hat{G}_n(\xi_\alpha + n^{-1/2} t) - G(\xi_\alpha + n^{-1/2} t))}{f(\xi_\alpha)} \le t_n \right), \end{aligned} \quad (3.3)$$

where

$$t_n = n^{1/2}(-\alpha + F(\xi_\alpha + n^{-1/2}t))/f(\xi_\alpha).$$

Obviously, t_n converges to t. Let

$$Z_{t,n} = n^{1/2}(\hat{G}_n(\xi_\alpha + n^{-1/2}t) - G(\xi_\alpha + n^{-1/2}t))/f(\xi_\alpha).$$

Split $Z_{t,n}$ into $I_n + II_n$, where

$$\begin{aligned} I_n &= n^{1/2}(G_n(\xi_\alpha + n^{-1/2}t) - G(\xi_\alpha + n^{-1/2}t))/f(\xi_\alpha), \\ II_n &= n^{1/2}(F_n(\xi_\alpha + n^{-1/2}t) - \hat{F}_n(\xi_\alpha + n^{-1/2}t))/f(\xi_\alpha). \end{aligned}$$

According to Lee and Wei (1996),

$$\sup_x |n^{-1/2} \sum_{j=1}^{n} \{I(e_j \le x) - F(x + \Lambda_{nj}) + F(x) - I(\varepsilon_j \le x)\}| = o_P(1), \quad (3.4)$$

so that

$$\begin{aligned} II_n &= n^{-1/2} \sum_{j=1}^{n} \{F(\xi_\alpha + n^{-1/2}t) - F(\xi_\alpha + n^{-1/2}t + \Lambda_{nj})\}/f(\xi_\alpha) + o_P(1) \\ &= -n^{-1/2} \sum_{j=1}^{n} \Lambda_{nj} + o_P(1), \end{aligned} \quad (3.5)$$

where the last equality is due to the mean value theorem and the fact $\max_{1 \le j \le n} |\Lambda_{nj}| = O_P(n^{-1/2})$. Further, by the arguments of Billingsley (1968, P. 106),

$$|I_n - n^{1/2}(G_n(\xi_\alpha) - G(\xi_\alpha))/f(\xi_\alpha)| = o_P(1). \quad (3.6)$$

Therefore, if we put

$$W_n = n^{1/2}(G_n(\xi_\alpha) - G(\xi_\alpha))/f(\xi_\alpha) - n^{-1/2} \sum_{j=1}^{n} \Lambda_{nj},$$

we have that

$$Z_{t,n} - W_n = o_P(1) \quad (3.7)$$

by (3.5) and (3.6). Particularly, $W_n = O_P(1)$ because

$$n^{-1/2} \sum_{j=1}^{n} \Lambda_{nj} = O_P(1). \quad (3.8)$$

Since, in view of (3.3) and (3.7), V_n and W_n satisfy the Conditions (4) and (5) in Lemma 1 of Ghosh, we obtain (3.1). (3.2) is a direct result of (3.1) and (3.8). □

Lemma 3.1. *Let $\alpha \in (0,1)$. Then,*

$$n^{-1}\sum_{j=1}^{n} X_{j-1}\{I(e_j \le \xi_{n\alpha}) - \alpha\} = n^{-1}\sum_{j=1}^{n} X_{j-1}\{I(\varepsilon_j \le \xi_\alpha) - \alpha\}$$
$$+n^{-1}\sum_{j=1}^{n} X_{j-1}\{\Lambda_{nj} + \xi_{n\alpha} - \xi_\alpha\} f(\xi_\alpha) + o_P(1), \qquad (3.9)$$

$$n^{-1}\sum_{j=1}^{n} X_{j-1}\varepsilon_j I(e_j \le \xi_{n\alpha}) = n^{-1}\sum_{j=1}^{n} X_{j-1}\varepsilon_j I(\varepsilon_j \le \xi_\alpha)$$
$$+n^{-1}\sum_{j=1}^{n} X_{j-1}\{\Lambda_{nj} + \xi_{n\alpha} - \xi_\alpha\}\xi_\alpha f(\xi_\alpha) + o_P(1), \quad (3.10)$$

and

$$n^{-2}\sum_{j=1}^{n} X_{j-1}^2 I(e_j \le \xi_{n\alpha}) = n^{-2}\sum_{j=1}^{n} X_{j-1}^2 I(\varepsilon_j \le \xi_\alpha) + o_P(1). \qquad (3.11)$$

Proof: For brevity, we only provide the proof for (3.10) since the proofs for (3.9) and (3.11) will follow essentially the same lines.

Write

$$n^{-1}\sum_{j=1}^{n} X_{j-1}\varepsilon_j I(e_j \le \xi_{n\alpha}) - n^{-1}\sum_{j=1}^{n} X_{j-1}\varepsilon_j I(\varepsilon_j \le \xi_\alpha) = I_n + II_n + III_n,$$

where

$$\begin{aligned} I_n &= n^{-1}\sum_{j=1}^{n} X_{j-1}\{\varepsilon_j I(\varepsilon_j \le \xi_{n\alpha} + \Lambda_{nj}) - H(\xi_{n\alpha} + \Lambda_{nj}) + H(\xi_{n\alpha}) \\ &\quad -\varepsilon_j I(\varepsilon_j \le \xi_{n\alpha})\}, \\ II_n &= n^{-1}\sum_{j=1}^{n} X_{j-1}\{H(\xi_{n\alpha} + \Lambda_{nj}) - H(\xi_{n\alpha})\}, \\ III_n &= n^{-1}\sum_{j=1}^{n} X_{j-1}\{\varepsilon_j I(\varepsilon_j \le \xi_{n\alpha_n}) - \varepsilon_j I(\varepsilon_j \le \xi_\alpha)\}, \end{aligned}$$

where $H(x) = E\varepsilon_j I(\varepsilon_j \le x)$. First we deal with I_n. Putting

$$Y_{nj} = n^{-1/2}X_{j-1},\ \mathcal{F}_{nj} = \sigma(\varepsilon_i; i \le j),\ a_n = 1,\ b_n = n,\ c_n = n^{1/2},\ \rho_n = n^{1/2},$$
$$\gamma = 1,\ \nu_1 = 1,\ \nu_2 = 2,\ \zeta_1 = 1,\ \zeta_2 = 1/4, \qquad (3.12)$$

one can check that these satisfy the conditions of Lemma 2.2. Hence, $I_n = o_P(1)$.

Second, by the mean value theorem,

$$II_n = n^{-1}\sum_{j=1}^{n} X_{j-1}\Lambda_{nj}H'(c_{nj}),$$

where c_{nj} is a random variable between $\xi_{n\alpha}$ and $\xi_{n\alpha}+\Lambda_{nj}$. Since $\max_{1\le j\le n}|\Lambda_{nj}| = O_P(n^{-1/2})$, in view of Theorem 3.1 we have that $\max_{1\le j\le n}|c_{nj}-\xi_\alpha| = O_P(n^{-1/2})$. This with the fact $n^{-1}\sum_{j=1}^{n}\Lambda_{nj} = O_P(1)$ yields that

$$II_n = n^{-1}\sum_{j=1}^{n} X_{j-1}\Lambda_{nj}\xi_\alpha f(\xi_\alpha) + o_P(1). \tag{3.13}$$

Finally, we deal with III_n. Write

$$III_n = n^{-1}\sum_{j=1}^{n} X_{j-1}\{H(\xi_{n\alpha}) - H(\xi_\alpha)\} + R_n, \tag{3.14}$$

where

$$R_n = n^{-1}\sum_{j=1}^{n} X_{j-1}\{\varepsilon_j I(\varepsilon_j \le \xi_{n\alpha}) - H(\xi_{n\alpha}) + H(\xi_\alpha) - \varepsilon_j I(\varepsilon_j \le \xi_\alpha)\}.$$

By utilizing Theorem 3.1 and applying Lemma 2.1 to the random variables and numbers in (3.12), one can show that $R_n = o_P(1)$. On the other hand, the first term in the RHS of the equality in (3.14) can be rewritten as $n^{-1/2}\sum_{j=1}^{n} X_{j-1}(\xi_{n\alpha} -\xi_\alpha)\xi_\alpha f(\xi_\alpha) + o_P(1)$ by the mean value theorem and Theorem 3.1. Combining this and (3.13)-(3.14), we assert (3.10). □

Proof of Theorem 1.1. Notice that

$$n(\hat{\phi}_{n,\alpha,\beta} - \phi) = \frac{n^{-1}\sum_{j=1}^{n} X_{j-1}\{\xi_{n\alpha}(J_j - \alpha) + \varepsilon_j K_j + \xi_{n\beta}(L_j - (1-\beta))\}}{n^{-2}\sum_{j=1}^{n} X_{j-1}^2 K_j}.$$

¿From Lemma 3.1, it follows that

$$n(\hat{\phi}_{n,\alpha,\beta} - \phi) = (I_n + I_n')/II_n + o_P(1),$$

where

$$I_n = n^{-1}\sum_{j=1}^{n} X_{j-1}\{\psi_{\alpha,\beta}(\varepsilon_j) - (\beta-\alpha)T(F)\},\quad I_n' = n^{-1}\sum_{j=1}^{n} X_{j-1}T(F)(\beta-\alpha),$$

$$II_n = n^{-2} \sum_{j=1}^{n} X_{j-1}^2 I(\xi_\alpha < \varepsilon_j \le \xi_\beta).$$

We first deal with II_n. Put

$$\tilde{X}_j = X_j I(|X_j| \le Bn^{1/2}),\ B > 0.$$

Since for any $B > 0$,

$$n^{-2} \sum_{j=1}^{n} \tilde{X}_{j-1}^2 \{I(\xi_\alpha < \varepsilon_j \le \xi_\beta) - (\beta - \alpha)\} = o_P(1),$$

which can be proved via applying Bernstein's inequality for martingales, we can write that

$$II_n = n^{-2} \sum_{j=1}^{n} X_{j-1}^2 (\beta - \alpha) + o_P(1). \tag{3.15}$$

Assume that $\phi = 1$. According to Theorem 2.4(ii) of Chan and Wei (1988),

$$\Big(\frac{1}{n^{1/2}\sigma} \sum_{j=1}^{[nt]} \varepsilon_j,\ \frac{1}{n^{1/2}\tau} \sum_{j=1}^{[nt]} \{\psi_{\alpha,\beta}(\varepsilon_j) - (\alpha - \beta)T(F)\}, \frac{1}{n\sigma\tau} \sum_{j=1}^{n} X_{j-1}\{\psi_{\alpha,\beta}(\varepsilon_j) - (\beta - \alpha)T(F)\}\Big) \xrightarrow{d} \left(B(t), W(t), \int_0^1 B(t)dW(t)\right), \tag{3.16}$$

where B, W are the Brownian motions in (1.6) and (1.7). Meanwhile, by the continuous mapping theorem,

$$\left(\frac{1}{n^{1/2}\sigma} \sum_{j=1}^{[nt]} \varepsilon_j, \frac{1}{n\sigma} \sum_{j=1}^{n} X_{j-1}, \frac{1}{n^2\sigma^2} \sum_{j=1}^{n} X_{j-1}^2\right) \xrightarrow{d} \left(B(t), \int_0^1 B(t)dt, \int_0^1 B^2(t)dt\right). \tag{3.17}$$

Combining (3.15)-(3.17) and Proposition in the Appendix 3 of Chan and Wei (1988), we obtain (1.10). The proof of (1.11) is quite similar to that of (1.10), and is omitted for brevity. □

References

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Boldin, M. V. (1982). "Estimation of the distribution of noise in an autoregression scheme," *Theorey of Probability and its Applications* **27** 866-871.

Chan, N. H. and Wei, C. Z. (1988). "Limiting distributions of least squares estiamtes of unstable autoregression processes," *The Annals of Statististics* **16** 367-401.

Dickey, D. A. and Fuller, W. A. (1979). "Distribution of the estimators for autoregressive time series with a unit root," *Journal of American Statistical Association* **74** 427-431.

Ghosh, J. K. (1971). "A new proof of the Bahadur representation of quantiles and an application," *The Annals of Mathematical Statistics* **42** 1957-1961.

Herce, M. A. (1996). "Asymptotic theory of LAD estimation in a unit root process with finite variance errors," *Econometric Theory* **12** 129-153.

Knight, K. (1991). "Limit theory for M-estimates in an integrated infinite variance process," *Econometric theory*, **7** 200-212.

Koul, H. L. (1992). *Weighted Empiricals and Linear Models.* IMS Lecture Notes-Monograph Series, Vol. 21. Hayward, California.

Lee, S. and Wei, C. Z. (1999) "On residual empirical processes of stochastic regression models with applications to time series," *To appear in the Annals of Statistics.*

Martin, R. D. and Yohai, V. J. (1986). "Influence functions for time series," *The Annals of Statistics* **14** 781-818.

Rao, M. M. (1978). "Asymptotic distribution of an estimator of the boundary parameter of an unstable process," *The Annals of Statistics* **15** 1667-1682.

Ruppert, D. and Caroll, R. J. (1980). "Trimmed least squares estimation in the linear model," *Journal of American Statistical Association* **75** 828-838.

Shorack, G. and Wellner, J. (1986). *Empirical processes with Applications to Statistics.* Wiley, New York. **75** 828-838.

Welsh, A. H. (1987). "The trimmed mean in the linear model," *The Annals of Statsitics* **15** 20-36.

White, J. S. (1958). "The limiting distribution of the serial correlation coefficient in the explosive case," *The Annals of Mathematical Statistics* **29** 1188-1197.