Journal of the Korean Statistical Society (1999), **28**: *4, pp 435–441*

Rate of Convergence of Empirical Distributions and Quantiles in Linear Processes with Applications to Trimmed Mean†

Sangyeol Lee[1]

ABSTRACT

A 'convergence in probability' rate of the empirical distributions and quantiles of linear processes is obtained. As an application of the limit theorems, a trimmed mean for the location of the linear process is considered. It is shown that the trimmed mean is asymptotically normal. A consistent estimator for the asymptotic variance of the trimmed mean is provided.

Keywords: Linear Process; Empirical Distributions; Empirical Quantiles; Rate of Convergence; Trimmed Mean.

1. INTRODUCTION AND THE MAIN RESULTS

Consider the linear process

$$X_j = \mu + \sum_{i=0}^{\infty} a_i \varepsilon_{j-i}, \tag{1.1}$$

where ε_t are iid random variables with $E|\varepsilon_t|^\alpha < \infty$ for some $\alpha > 0$ and $\{a_j\}$ is a sequence of real numbers with $|a_i| \leq ci^{-q}$ for some $c, q > 0$, where $q\alpha > 1$ or $q > 1$ according as $0 < \alpha < 1$ and $\alpha \geq 1$. The process in (1.1) covers a broad class of time series models including the most popular ARMA models. In the literature, there has been an attempt to study the rate of the empirical distributions and quantiles. For example, Hannan and Hesse (1988) and Hesse (1990) derived a rate of their almost sure convergence in linear processes. However, for example, Hesse (1990, Theorem 2) only covers the process with geometrically decaying coefficients. Therefore, there is a need to extend to more general processes. Since the almost sure convergence result is often beyond what one requires for statistical

†This work was supported by S.N.U. Posco Research Fund (1998).

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

analysis, we here study a rate of the convergence in probability of the empirical distributions and quantiles. Actually, de Wet and Venter (1974) and Lee (1995) derived an asymptotic expansion of the trimmed mean to establish a central limit theorem for i.i.d. r.v.'s and linear processes using such a convergence result. Since a strong mixing condition imposed by Lee for using the results by Babu and Singh (1978) does not necessarily hold for linear processes (cf. Bradley, 1986), we here adopt the mixingale approach of McLeish (1975).

Let $F_\mu(\cdot) = F(\cdot - \mu)$ and $f_\mu = f(\cdot - \mu)$ denote the marginal distribution and density of X_1. For $u \in (0,1)$, let ξ_u be the number with $F(\xi_u) = u$. Let $\tilde{X}_j = X_j - \mu$, and $F_n(x) = n^{-1/2}\sum_{j=1}^{n} I(\tilde{X}_j \leq x)$, and $\xi_{nu} = F_n^{-1}(u) = \inf\{y; F_n(y) \geq u\}$. Finally, let $\eta = (q\tilde{\alpha} - 1)/2$, where $\tilde{\alpha} = \min\{1, \alpha\}$. Here is the main result of this article.

Theorem 1.1. *Suppose that* $\sup_x f(x) < \infty$ *and* $\eta > 0$. *Then*

$$n^{\lambda} \sup_{x} |F_n(x) - F(x)| = o_P(1) \quad \text{for all} \quad \lambda \in (0, 1/(2 + 2/\eta)). \tag{1.2}$$

Theorem 1.1 particularly shows that when $\{a_j\}$ is exponentially decaying to 0, the deviations of the empirical and true distributions go to 0 in probability with rate $n^{-\lambda}$ for any λ arbitrarily close to 1/2.

The following is a subsidiary result of Theorem 1.1.

Corollary 1.1. *Suppose that* $\sup_x f(x) < \infty$, $K_\delta = \inf_{\delta \leq x \leq 1-\delta} f(x) > 0$ *for all* $0 < \delta < 1/2$, *and* $\eta > 0$. *Then for every* $0 < \delta < 1/2$,

$$n^{\lambda} \sup_{\delta \leq u \leq 1-\delta} |\xi_{nu} - \xi_u| = o_P(1) \quad \text{for all} \quad \lambda \in (0, 1/(2 + 2/\eta)). \tag{1.3}$$

Proof: Given $\varepsilon > 0$, put $S_n(\varepsilon) = \{n^{\lambda} \sup_x |F_n(x) - F(x)| < \varepsilon\}$. Assume $n^{-\lambda}\varepsilon \leq \delta/2$. On $S_n(\varepsilon)$, we have that $\xi_{u-n^{-\lambda}\varepsilon} \leq \xi_{nu} \leq \xi_{u+n^{-\lambda}\varepsilon}$ for all $u \in [\delta, 1-\delta]$, and thus $P(\{n^{\lambda} \sup_{\delta \leq u \leq 1-\delta} |\xi_{nu} - \xi_u| > \varepsilon/K_{\delta/2}\} \cap S_n(\varepsilon)) = 0$. Then (1.3) is yielded by this and the fact $P(S_n(\varepsilon)) \to 1$. □

We need the following lemma to prove Theorem 1.1.

Lemma 1.1. *Let* Y, Z *be independent r.v.'s and* $X = Y + Z$. *If* F *denotes the distribution function of* X, *and if* $E|Z|^{\alpha} \leq 1$ *for some* $0 < \alpha \leq 1$ *and* $\sup_x |F'(x)| < \infty$, *then* $\sup_x E|I(X \leq x) - I(Y \leq x)| = O((E|Z|^{\alpha})^{1/2})$.

Proof: For any $b > 0$, we have that

$$E|I(X \le x) - I(Y \le x)| \le P(|Z| > b) + \max\{|F(x+b) - F(x)|, |F(x-b) - F(x)|\}.$$

Putting $b = (E|Z|^\alpha)^{1/2\alpha}$, we obtain the lemma. □

Proof of Theorem 1.1. Let

$$X_{mj} = \sum_{i=0}^{m-1} a_i \varepsilon_{j-i}, \quad F^{(m)}(x) = P(X_{mj} \le x) \text{ and } F_n^{(m)}(x) = n^{-1} \sum_{i=1}^{n} I(X_{mj} \le x).$$

Let $\zeta, \theta > 0$ be the real numbers such that $\zeta < 1 - 2\lambda, \theta > \lambda, \lambda + \theta - \eta\zeta < 0$. Putting $m = n^\zeta$ and applying Lemma 1.1, we have that

$$\sup_x E|I(X_j \le x) - I(X_{mj} \le x)| \le C(E|X_j - X_{mj}|^{\tilde{\alpha}})^{1/2} = O(n^{-\eta\zeta}), \tag{1.4}$$

and thus

$$\sup_x |F(x) - F^{(m)}(x)| = O(n^{-\eta\zeta}). \tag{1.5}$$

We first show that

$$n^\lambda \sup_x |F_n^{(m)}(x) - F(x)| = o_P(1). \tag{1.6}$$

Let $-\infty = x_0 < x_1 < \cdots < x_n = \infty$ be the numbers such that $F(x_j) = j/n^\theta$, $j = 0, \ldots, n^\theta$. By the monotonicity of $F^{(m)}$ and $F_n^{(m)}$, we can write that $\sup_x |F_n^{(m)}(x) - F^{(m)}(x)| \le I_1 + I_2$, where

$$I_1 = \max_{0 \le j \le n^\theta} |F_n^{(m)}(x_j) - F^{(m)}(x_j)|, \quad I_2 = \max_{0 \le j \le n^\theta} (F^{(m)}(x_{j+1}) - F^{(m)}(x_j)).$$

Notice that by (1.5), $n^\lambda \max_{0 \le j \le n^\theta} |F^{(m)}(x_j) - F(x_j)| = O(n^{-(\eta\zeta - \lambda)})$, and so

$$n^\lambda I_2 \le O(n^{-(\eta\zeta - \lambda)}) + O(n^{\lambda - \theta}) = o(1). \tag{1.7}$$

To deal with I_1, put $d_{ij} = I(X_{mi} \le x_j) - F^{(m)}(x_j)$. For each n, we can find u, v such that $n = mu + v$, $0 \le v < m$. For simplicity, we assume that $n = mu$. Write $\sum_{i=1}^n d_{ij} = D_1 + \cdots + D_m$, where $D_k = \sum_{l=1}^u d_{(l-1)m+k,j}, k = 1, \ldots, m$. Note that $|d_{ij}| \le 1$ and $Var(D_k) \le cu$ for some $c > 0$. Applying Bernstein's inequality, we have for any $\delta > 0$,

$$P(n^\lambda |n^{-1} \sum_{k=1}^m D_k| > \delta) \le 2m \exp\left\{ -\frac{\delta^2 n^{2-2\lambda-2\zeta}}{2cu + (1/3)n^{1-\lambda-\zeta}} \right\},$$

which yields $n^\lambda I_1 = o_P(1)$. This together with (1.7) implies (1.6).

Now, in view of (1.6), it suffices to show that

$$n^\lambda \sup_x |F_n(x) - F(x) - F_n^{(m)}(x) + F^{(m)}(x)| = o_P(1).$$

Due to (1.5), we only have to verify that

$$II = \sup_x |F_n(x) - F_n^{(m)}(x)| = o_P(n^{-\lambda}). \tag{1.8}$$

Write $II \le II_1 + II_2$, where

$$II_1 = \max_{0\le j\le n^\theta} |F_n(x_j) - F_n^{(m)}(x_j)|, \quad II_2 = \max_{0\le j\le n^\theta} |F_n^{(m)}(x_{j+1}) - F_n^{(m)}(x_j)|.$$

From (1.4), we have that $En^\lambda II_1 O(n^{\theta+\lambda-\eta\zeta}) = o(1)$. Since $n^\lambda II_2 \le n^\lambda(2I_1 + I_2) = o_P(1)$, (1.8) is proved. □

2. TRIMMED MEAN

Suppose that the observations $X_1, \dots, X_n$ are given. Let $X_{n1}, \dots, X_{nn}$ be the corresponding order statistics. Let a and b be positive real numbers with $a + b < 1$. For $u \in (0,1)$, define $q_n(u) = [nu]$ or $[nu]+1$ corresponding as nu is an integer or not. Put $r = q_n(a)$, $s = q_n(1-b)$ and $\gamma = 1 - a - b$, respectively. Define the a, b-trimmed mean as follows:

$$\overline{X}_{a,b} = n^{-1}\gamma^{-1}\sum_{i=r}^{s} X_{ni}. \tag{2.1}$$

Let

$$W_i = \gamma^{-1}\left\{\xi_a^*(I(\tilde{X}_i < \xi_a) - a) + X_i I_i - EX_i I_i + \xi_{1-b}^*(I(\tilde{X}_i > \xi_{1-b}) - b)\right\},$$

where $I_i = I(\xi_a \le \tilde{X}_i \le \xi_{1-b})$ and $\xi_u^* = \xi_u + \mu$ for $0 < u < 1$. In addition, set

$$\theta = \mu + \gamma^{-1}\int_{\xi_a}^{\xi_{1-b}} x dF(x). \tag{2.2}$$

Theorem 2.1. *Suppose that f satisfies the conditions in Corollary 1.1 and $\eta > 1$. Then,*

$$\overline{X}_{a,b} - \theta = n^{-1}\sum_{i=1}^{n} W_i + o_P(n^{-2\lambda}) \tag{2.3}$$

for any $\lambda \in (0, 1/(2+2/\eta))$, *and* $\tau^2(a,b) = EW_1^2 + 2\sum_{i=1}^{\infty} EW_1 W_{i+1}$ *exists. If in addition* $\tau^2(a,b) > 0$, *it holds that as* $n \to \infty$,

$$n^{1/2}(\overline{X}_{a,b} - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2(a,b)). \tag{2.4}$$

We need two lemmas to prove Theorem 2.1. Throughout, $||\cdot||_2$ denotes the L^2 norm, $\mathcal{F}_j$ denotes the σ-field generated by ε_i, $i \le j$, and $S_n = \sum_{i=1}^{n} W_i$. The following lemma is a direct result of Lemma 1.1.

Lemma 2.1. *Let* X, Y, Z *be the same r.v.'s in Lemma 1.1, and let* $I = (\zeta_1 \le X \le \zeta_2)$ *and* $J = (\zeta_1 \le Y \le \zeta_2)$. *Suppose that* $\mathcal{F}$ *is a* σ*-field,* Z *is* $\mathcal{F}$*-measurable,* $E|Z|^\alpha \le 1$ *for some* $0 < \alpha \le 1$, *and* Y *is independent of* $\mathcal{F}$. *Then under the same condition of Lemma 1.1, for any* $A > 0$,

$$E[|XI - YJ|\ |\mathcal{F}] \le K\{I(|Z| > A) + (E|Z|^\alpha)^{1/2} + |Z|^\alpha\}, \quad K > 0.$$

Lemma 2.2. *Under the same conditions of Theorem 2.1,*

(1) $||E(W_{j+k}|\mathcal{F}_j)||_2 \le \psi_k$, *where* $\psi_k = O(k^{-\eta})$.

(2) $\{W_j, \mathcal{F}_j\}$ *is a mixingale with* ψ_k *of size -1/2 in the sense of McLeish.*

(3) $\tau^2(a,b) = \sum_{h=-\infty}^{\infty} \gamma_w(h)$ *is absolutely summable, where* $\gamma_w(h) = EW_1 W_{1+h}$.

(4) $E|n^{-1}E[(S_{m+n} - S_m)^2|\mathcal{F}_0] - \tau^2(a,b)| \to 0$ *as* $m, n \to \infty$,

Proof: Applying Lemmas 1.1 and 2.1 with $Y = \sum_{l=-\infty}^{k-1} a_l \varepsilon_{j+k-l}$ and $Z = \sum_{l=k}^{\infty} a_l \varepsilon_{j+k-l}$, we attain (1). Since $k^{-\eta}$ satisfies the conditions in Definition 2.4 of McLeish, (2) is proved. (3) is an immediate result of (1). (4) is proved by the fact that $E|E(W_i W_j|\mathcal{F}_0) - EW_i W_j| = O(i^{-\eta} + j^{-\eta})$, which is due to Lemmas 1.1 and 2.1. □

Proof of Theorem 2.1. Write that $\sum_{i=r}^{s} X_{ni} = \sum_{i=1}^{n} X_i I(\xi_a^* \le X_i \le \xi_{1-b}^*) + U_{n1} + U_{n2}$, where $U_{n1} = B_n(I(X_{nr} < \xi_a^*) - I(X_{nr} > \xi_a^*))$. Here, B_n denotes the sum of X_i's between ξ_a^* and X_{nr_n}. The term U_{n2} is analogously defined. By Theorem 1.1 and Corollary 1.1, there exists $\lambda > 1/4$, such that

$$nF_n(\xi_a^*) - (r-1) = o_P(n^{1-\lambda}), \quad X_{nr} - \xi_a^* = o_P(n^{-\lambda}). \tag{2.5}$$

Thus $U_{n1} = (nF_n(\xi_a^*) - (r-1))\xi_a^* + o_P(n^{1-2\lambda})$. Since a similar expression holds for U_{n2}, (2.3) follows. The rest of the theorem follows from Lemma 2.2 and Theorem 2.6 of McLeish (1975). □

In the remainder of this section, we estimate the asymptotic variance of the trimed mean. Assume that $a = b$ and the density of ε_t is symmetric about zero. In that case, θ becomes μ due to (2.2) and W_i is rewritten as follows:

$$W_i = \gamma^{-1}\left\{\xi_a^*(I(\tilde{X}_i < \xi_a) - a) + \tilde{X}_i I_i + \mu(I_i - \gamma) + \xi_{1-a}^*(I(\tilde{X}_i > \xi_{1-a}) - a)\right\}$$

with $I_i = I(\xi_a \le \tilde{X}_i \le \xi_{1-a})$. Futher, from Theorem 2.1, we can write

$$\overline{X}_a := \overline{X}_{a,1-a} = \mu + n^{-1}\sum_{i=1}^{n} W_i + o_P(n^{-1/2}). \tag{2.6}$$

As an estimate of $\tau^2(a) = \sum_{h=-\infty}^{\infty} EW_1W_{1+h}$, one may consider $\tau_n^2(a) = \sum_{|h|\le h_n}\gamma_n(h)$, where $\gamma_n(h) = n^{-1}\sum_{i=1}^{n-|h|} W_iW_{i+h}$, and $\{h_n\}$ is a sequence of positive integers diverging to ∞ with $h_n = O(n^\rho)$, $\rho \in (0, 1/3]$. However, since $\gamma_n(h)$ are unobservable, replacing $\gamma_n(h)$ by $\hat{\gamma}_n(h) = n^{-1}\sum_{i=1}^{n-|h|}\hat{W}_i\hat{W}_{i+|h|}$;

$$\begin{aligned}\hat{W}_i \;&=\; \gamma^{-1}\Big\{X_{nr}(I(X_i < X_{nr}) - a) + (X_i - \overline{X}_a)I(X_{nr} \le X_i \le X_{ns}) \\ &\quad + (I(X_{nr} \le X_i \le X_{ns}) - \gamma)\bar{X}_a + X_{ns}(I(X_i > X_{ns}) - a)\Big\},\end{aligned}$$

we employ $\hat{\tau}_n^2(a) = \sum_{|h|\le h_n}\hat{\gamma}_n(h)$.

Proof: Suppose that f satisfies the conditions in Corollary 1.1, and $\eta > 2$. Then, $\hat{\tau}_n^2(a) \xrightarrow{P} \tau^2(a)$. □

The following is useful to prove Theorem 2.2.

Lemma 2.3. *Under the condition of Theorem 2.2, $\tau_n^2(a) \xrightarrow{P} \tau^2(a)$.*

Proof: Let $x_n = \sum_{|h|\le h_n}(1 - \frac{|h|}{n})EW_1W_{1+|h|}$. Since by Lemma 2.2 (3), $\sum_{h=-\infty}^{\infty}$ $|EW_1W_{1+h}| < \infty$, the lemma will follow if $\tau_n^2(a) - x_n \xrightarrow{P} 0$. To verify this, we set $U_{i,h} = W_iW_{i+h} - EW_iW_{i+h}$ for $h \ge 0$. Put $\xi_{l,h} = (1 - l/(n-h))EU_{1,h}U_{1+l,h}$. Then by stationarity, $E(\gamma_n(h) - y_n(h))^2 = I_n(h) + II_n(h)$, where $I_n(h) = n^{-2}(n-h)(EU_{1,h}^2 + 2\sum_{l=1}^{h}\xi_{l,h})$ and $II_n(h) = 2n^{-2}(n-h)\sum_{l=h+1}^{n-h}\xi_{l,h}$. Since $E|U_{1,h}U_{1+l,h}|$ $= O((h-l)^{-\eta})$ for $h > l$, due to (2.5), we have that $\max_{0\le h\le h_n}|I_n(h)| = O(n^{-1}h_n)$. Similarly, for some $K > 0$, $\max_{0\le h\le h_n}|II_n(h)| \le Kn^{-1}\sum_{k=1}^{\infty}k^{-\eta} = O(n^{-1})$. Hence, we obtain $\max_{0\le h\le h_n}E(\gamma_n(h) - y_n(h))^2 = O(n^{-1}h_n)$. Thus, $\{E(\tau_n^2(a) - x_n)^2\}^{1/2} = O(n^{-1/2}h_n) \to 0$ as $n \to \infty$. □

Proof of Theorem 2.2. Without loss of generality, assume that $\mu = 0$. Put $\hat{W}_i = W_i + R_i$. It suffices to show that $n^{-1}h_n \sum_{i=1}^n \{R_i^2 + |R_i|\} = o_P(1)$. Define $H_i = I(X_i < \xi_a) - a, K_i = I(X_i > \xi_{1-a}) - a, \hat{H}_i = I(X_i < X_{nr}) - a, \hat{I}_i = I(X_{nr} \le X_i \le X_{ns})$, $\hat{K}_i = I(X_i > X_{ns}) - a$. Now, split R_i into the four terms; $R_i = R_{i1} + R_{i2} + R_{i3} + R_{i4}$, where $R_{i1} = X_{nr}\hat{H}_i - \xi_a H_i$, $R_{i2} = (X_i - \bar{X}_a)\hat{I}_i - X_i I_i$, $R_{i3} = X_{ns}\hat{K}_i - \xi_{1-a}K_i$ and $R_{i4} = (I_i - \gamma)\mu - (\hat{I}_i - \gamma)\bar{X}_a$.

First, notice that for some $\lambda > 1/3$, $\sum_{i=1}^n R_{i1}^2 = o_P(n^{1-\lambda})$, where we have used (2.5). This in turn implies that $n^{-1}h_n \sum_{i=1}^n R_{i1}^2 = o_P(n^{\rho-\lambda}) = o_P(1)$. Similarly, it can be shown that $n^{-1}h_n \sum_{i=1}^n R_{i3}^2 = o_P(1)$. Next, observe that by (2.5) and (2.6), $\sum_{i=1}^n R_{i2}^2 = O_P(n^{1-\lambda})$, which asserts $n^{-1}h_n \sum_{i=1}^n R_{i2}^2 = o_P(1)$. Finally, it is easy to see $\sum_{i=1}^n R_{i4}^2 = O_P(1)$. Thus $n^{-1}h_n \sum_{i=1}^n R_i^2 = o_P(1)$. In a similar fashion, we can see that $n^{-1}h_n \sum_{i=1}^n |R_i| = o_P(1)$. This completes the proof. □

References

Babu, G. J. and Singh, K. (1978). "On deviations between empirical and quantile processes for mixing random variables," *Journal of Multivariate Analysis.* **8**, 532-549.

Bradley, R. C. (1986). "Basic properties of strong mixing conditions," In *Dependence in Probability and Statistics* E. Eberein and M. S. Taqqu ed. 165-192.

de Wet, T. and Venter, J. H. (1974). "An asymptotic representation of trimmed means with application," *South African Statististics Journal* **8**, 127- 134.

Hannan, E. J. and Hesse, C. H. (1988). "Rate of convergence for the quantile function of a linear process," *Australian Journal of Statististics* **30A** 283-295.

Hesse, C. H. (1990). "Rates of convergence for the empirical distribution function and the empirical characteristic function of a broad class of linear processes," *Journal of Multivariate Analysis* **35** 186-202.

Lee, S. (1995). "A trimmed mean of location of an AR(∞) stationary process," *Journal of Statistical Planning and Inferences* **48**, 131-140.

McLeish, D. L. (1975). "Invariance principles for dependent variables," *Zeit. Wahrsch. Verw. Gebiete* **30** 168-178.