Journal of the Korean Statistical Society (2003), **32**: *2, pp 203–211*

A NOTE ON WEAK CONVERGENCE OF EMPIRICAL PROCESSES FOR A STATIONARY PHI-MIXING SEQUENCE†

Tae Yoon Kim[1], Jang Han Kim[1] and Tai Sup Lee[2]

Abstract

A new result of weak convergence of the empirical process is established for a stationary ϕ-mixing sequence of random variables, which relaxes the existing conditions on mixing coefficients. The result is basically obtained from bounds for even moments of sums of ϕ-mixing r.v.'s useful for handling triangular arrays with entries decreasing in size.

AMS 2000 subject classifications. Primary 60F05.
Keywords. ϕ-mixing, weak convergence, empirical process.

1. Introduction and Statement of Main Result

Let $\{\xi_j, j \geq 1\}$ be a strictly stationary sequence of random variables satisfying a ϕ-mixing condition

$$\sup\{|P(B|A) - P(B)| : A \in \mathcal{M}_1^k, B \in \mathcal{M}_{k+n}^\infty\} \leq \phi(n) \to 0 \ \ (n \to \infty)$$

where $\mathcal{M}_a^b$ denotes the σ-field generated by ξ_j $(a \leq j \leq b)$. Denote by $F_n(t)$ the empirical distribution function of the ϕ-mixing sequence $\{\xi_j, j \geq 1\}$ at stage n. For $\kappa \in (0,1)$ let

$$Y_n(t) = \frac{n}{\sigma_n}\{F_n(t) - t\}, \ \kappa \leq t \leq 1-\kappa \tag{1.1}$$

be the corresponding empirical process. Here $\sigma_n^2 = \mathrm{Var}(\sum_{i=1}^n g_t(\xi_i))$, $g_t(\xi) = I_{[0,t]}(\xi) - t$, and $I_A(\cdot)$ is the indicator function of A. Further set for $0 < \kappa \leq s, t \leq 1-\kappa < 1$,

$$\sigma(s,t) = \sigma_n^{-2} E\left[\left\{\sum_{i=1}^n g_s(\xi_i)\right\}\left\{\sum_{j=1}^n g_t(\xi_j)\right\}\right].$$

Received February 2003; accepted March 2003.
†This work was supported by Korea Research Foundation Grant, 2002-070-C00018.
[1]Department of Statistics, Keimyung University, Taegu 704-701, Korea
[2]Department of Statistics, Anyang University, Anyang 430-714, Korea

Define $Y(t)$, $\kappa \le t \le 1-\kappa$, to be the Gaussian random function specified by

$$E\{Y(t)\} = 0, \;\; E\{Y(s)Y(t)\} = \sigma(s,t)$$

for $\kappa \le s \le t \le 1-\kappa$. The weak convergence of usual empirical process Y_n (*i.e.*, σ_n replaced by $\sqrt{n}$) has been established by Theorem 22.1 of Billingsley (1968) under the condition $\sum i^2\phi(i)^{1/2} < \infty$. Then Sen (1971), Yoshihara (1974, 1978) and Yokoyama (1980) obtained further developments along this line (in the last paper, the above condition is relaxed to $\sum \phi(i) < \infty$). We show here that if $\kappa > 0$ the above result remains true under less restrictive condition on mixing coefficients, say, (1.2).

THEOREM 1.1. *Let* $\{\xi_i\}$ *be stationary and* ϕ*-mixing with*

$$\phi(i) = O(i^{-\beta}) \tag{1.2}$$

for $1/2 < \beta$. *If* $\kappa > 0$, *then* $\{Y_n(t) : \kappa \le t \le 1-\kappa\}$ *converges in law (as* $n \to \infty$*) to* $\{Y(t) : \kappa \le t \le 1-\kappa\}$.

REMARK 1.1. Notice that since $\sigma_n = 0$ if $t = 0$ and $t = 1$, a positive κ is essential to define the $Y_n(t)$ in Theorem 1.1. Note also that $\sigma(t,t) = 1$ and $\sigma(s,t)$ may be assumed to be a positive constant $c(s,t)$ depending on t and s. If $\sum \phi(i) < \infty$, then

$$\begin{aligned} c(s,t) = &\left[E\{g_s(\xi_1)g_t(\xi_1)\} + \sum_{k=2}^{\infty} E\{g_s(\xi_1)g_t(\xi_k)\} + \sum_{k=2}^{\infty} E\{g_s(\xi_k)g_t(\xi_1)\}\right] \\ &\times \left[E\{g_t^2(\xi_1)\} + 2\sum_{k=2}^{\infty} E\{g_t(\xi_1)g_t(\xi_k)\}\right]^{-1}. \end{aligned}$$

One of limitations from $\kappa > 0$ lies in its applicability to statistical problems. For example, a testing problem of normal *vs.* heavy tailed distribution usually focuses the behavior of the empirical process at $t = 0, 1$. The main mathematical tool for the proof of Theorem 1.1 is the following moment bounds useful for triangular arrays with entries decreasing in size. Applications of these types of moment bounds are usually found in some settings that arise in nonparametric function estimation. See Doukhan, León and Portal (1984), Lemma 9 of Truong and Stone (1992), Lemma A2 of Yu (1993), and Cox and Kim (1995).

LEMMA 1.1. *Let* $\{\xi_i\}$ *be a centered stationary* ϕ*-mixing sequence with*

$$E(|\xi_i|^2) < \frac{c}{\sqrt{n}}, \;\; E(|\xi_i|) < \frac{c}{\sqrt{n}}. \tag{1.3}$$

Let $S_n = \xi_1 + \cdots + \xi_n$, $n \geq 1$ and l be a positive integer. If (1.2) holds, then for all $n \geq 1$ and for some $0 < \theta < 1/2$

$$E\{(S_n)^{2l}\} \leq an^{(1-\theta)l} \tag{1.4}$$

where a does not depend on n.

2. Proofs

Proof of Theorem 1.1. Throughout this proof we simply indicate the modifications to be made in the well-written proof of Theorem 22.1 of Billingsley (1968). Same notations are used here. The first part of Billingsley's proof remains unchanged which we can, without loss of generality, assume that ξ_0 is uniformly distributed on $[0,1]$. The second part of the proof shows that the finite dimensional distributions of $\{Y_n(t)\}$ converge to those of $\{Y(t)\}$. The same line as in Billingsley (1968, p. 197) applies here. Note that one may use invariance principle instead of Theorem 20.1 of Billingsley (1968) here. Indeed it is known that a strictly stationary centered ϕ-mixing sequence with $E(S_n^2) \to \infty$ and $E(|X_i|^{2+\delta}) < \infty$ for some $\delta > 0$ satisfies invariance principle (Ibragimov, 1975). It remains to show that given $\epsilon > 0$, $\eta > 0$, we can find δ, $0 < \delta < 1$ such that $P\{w(Y_n, \delta) \geq \epsilon\} \leq \eta$ for all sufficiently large n.

Fix $1 > \epsilon > 0$ and $\eta > 0$. Since ξ_0 is uniformly distributed,

$$E\{|g_t(\xi_0) - g_s(\xi_0)|\} \leq |t-s|, \quad E\{|g_t(\xi_0) - g_s(\xi_0)|^2\} \leq |t-s|. \tag{2.1}$$

Assume that

$$\frac{\epsilon}{n^b} < |t-s| < \frac{\epsilon}{\sqrt{n}} \quad \text{for some } \frac{1}{2} < b < 1. \tag{2.2}$$

Since (1.3) holds by (2.1) and (2.2), an application of Lemma 1.1 yields

$$E\left\{\left|\sum_{i=1}^{n}(g_t(\xi_i) - g_s(\xi_i))\right|^{2l}\right\} \leq K_1 n^l n^{-\theta l}$$

for some positive constants K_1 and $0 < \theta < 1/2$. Then we have

$$E\{|Y_n(t) - Y_n(s)|^{2l}\} \leq K_1 \sigma_n^{-2l} n^l n^{-\theta l} = (n^\theta h_n)^{-l}$$

where h_n is a slowly varying function. Note that under the conditions of theorem, it is known that $\sigma_n^2 = nh_n$. Using (2.2) and properties of the slowly varying function it is easy to see that the last expression is bounded by

$$K_1 \cdot \left|\frac{t-s}{\epsilon}\right|^{\theta_0 l/b} = K_1 \cdot \left|\frac{t-s}{\epsilon}\right|^{\alpha} \tag{2.3}$$

for some $\theta_0 < \theta$ and $\alpha = \theta_0 l/b > 1$ since one may choose l sufficiently large. Note that $\theta_0/b < 1$.

Assume now that p is a number satisfying $\epsilon/n^b \leq p$ and consider the random variables

$$Y_n(s+ip) - Y_n(s+(i-1)p),\ i = 1, \dots, m,$$

where m is a positive integer. By (2.3) and Theorem 12.2 of Billingsley (1968),

$$P\Big(\max_{i\leq m} |Y_n(s+ip) - Y_n(s)| \geq \lambda\Big) \leq \frac{K}{\lambda^{2l}} \left(\frac{m}{\epsilon}\right)^{\alpha} p^{\alpha} \tag{2.4}$$

for some positive constant K. Then by (22.18) of Billingsley (1968), we have

$$\sup_{s\leq t\leq s+mp} |Y_n(t) - Y_n(s)| \leq 3 \max_{i\leq m} |Y_n(s+ip) - Y_n(s)| + p\sqrt{n}. \tag{2.5}$$

If

$$\frac{\epsilon}{n^b} \leq p < \frac{\epsilon}{\sqrt{n}}, \tag{2.6}$$

then (2.4) applies, and it follows by (2.5) that

$$P\Big(\sup_{s\leq t\leq s+mp} |Y_n(t) - Y_n(s)| \geq 4\epsilon\Big) \leq \frac{K}{\epsilon^{2l+\alpha}} m^{\alpha} p^{\alpha}.$$

Choose $\delta > 0$ so that $K\delta^{\alpha-1}/\epsilon^{2l+\alpha} < \eta$. The choice of such δ is possible because $\alpha > 1$. It then follows that

$$P\Big(\sup_{s\leq t\leq s+\delta} |Y_n(t) - Y_n(s)| \geq 4\epsilon\Big) < \eta\delta,$$

provided there exists a p and an integer m such that (2.6) holds and $mp = \delta$. But this is equivalent to requiring the existence of an integer m such that

$$\frac{\delta}{\epsilon}\sqrt{n} < m < \frac{\delta}{\epsilon} n^b \quad \text{for some } \frac{1}{2} < b < 1$$

which is true for all sufficiently large n. The rest of the proof is same as in Billingsley and Theorem 1.1 is proven. □

PROOF OF LEMMA 1.1. We start with the well known moment inequality for ϕ-mixing. Let r_1, r_2 be positive numbers such that $r_1^{-1} + r_2^{-1} = 1$. Suppose that X and Y are uniform mixing random variables measurable with respect to the σ-fields $\mathcal{M}_{-\infty}^{0}$, M_n^{∞} respectively and assume further that $||X||_{r_1}$, $||Y||_{r_2} < \infty$. Then

$$|E(XY) - E(X)E(Y)| \leq 2\phi_n^{1/r_1} \parallel X \parallel_{r_1} \parallel Y \parallel_{r_2}. \tag{2.7}$$

See Ibragimov (1962) for its proof. First observe that under the conditions of theorem there exists a constant a such that

$$E(S_n^2) \leq an^{1-\theta}$$

for some $0 < \theta < 1/2$. Indeed

$$\begin{aligned} E(S_n^2) &= \sum_{i=1}^{n} E(\xi_i^2) + \sum_{i \neq j}^{n} \sum^{n} |E(\xi_i \xi_j)| \\ &\leq nE(\xi_i^2) + n \sum_{i=1}^{n} |E(\xi_0 \xi_i)| \\ &\leq nE(\xi_i^2) + 2nE(|\xi|) \sum_{i=1}^{n} i^{-\beta} \\ &\leq c(n^{1/2} + 2n^{1/2} n^{1-\beta}) \\ &\leq an^{1-\theta} \end{aligned}$$

for some $0 < \theta < 1/2$. Note that in the above we used (1.2), (1.3) and (2.7) with $r_1 = 1$ and $r_2 = \infty$. This verifies that (1.4) is true for $l = 1$. It will therefore be sufficient to assume that (1.4) is true if l is an integer $m \geq 1$ and prove that it is then true if $l = m + 1$. We thus assume in the following that the lemma is true if $l = m$, *i.e.*, $E(S_n^{2m}) \leq (\text{const}) \cdot n^{(1-\theta)m}$ for some $0 < \theta < 1/2$.

Let k be a positive integer, to be determined more precisely below, and define T_n, $\hat{S}_n$ and C_n, by

$$T_n = \sum_{n+1}^{n+k} \xi_j, \ \hat{S}_n = \sum_{n+k+1}^{2n+k} \xi_j \text{ and } C_n = E(S_n^{2(m+1)}).$$

Then we are to prove

$$C_n \leq an^{(1-\theta)(m+1)}, \ n = 1, 2, \dots \tag{2.8}$$

for a proper choice of a. In order to prove this we first prove that if $\epsilon_1 > 0$,

$$E\{(S_n + \hat{S}_n)^{2m+2}\} \leq (2 + \epsilon_1)C_n + a_1 n^{(1-\theta)(m+1)}, \ n = 1, 2, \dots \tag{2.9}$$

for a proper choice of a_1 and k. In fact remembering that S_n and $\hat{S}_n$ have the same distribution,

$$E\left\{(S_n+\hat{S}_n)^{2m+2}\right\} = \sum_{j=0}^{2m+2}\binom{2m+2}{j}E(S_n^j\hat{S}_n^{\,2m+2-j})$$

$$= E(S_n^{2m+2}) + E(\hat{S}_n^{\,2m+2}) + \sum_{j=1}^{2m+1}\binom{2m+2}{j}E(S_n^j\hat{S}_n^{\,2m+2-j})$$

$$\leq 2C_n + \sum_{j=1}^{2m+1}\binom{2m+2}{j}\left|E(S_n^j\hat{S}_n^{\,2m+2-j})\right|. \tag{2.10}$$

Now for the second term of (2.10) using (2.7) with

$$X = S_n^j,\ Y = \hat{S}_n^{\,2m+2-j},\ r_1 = \frac{2m+2}{j} \text{ and } r_2 = \frac{2m+2}{2m+2-j},$$

we have

$$\left|E(S_n^j\hat{S}_n^{\,2m+2-j})\right| \leq 2\phi_k^{1/(2m+2)}C_n + \left|E(S_n^j)E(S_n^{2m+2-j})\right| \tag{2.11}$$

for $j = 1, \ldots, 2m+1$. We substitute this in (2.10) to obtain

$$E\left\{(S_n+\hat{S}_n)^{2m+2}\right\}$$

$$\leq 2C_n + \sum_{j=1}^{2m+1}\binom{2m+2}{j}\left\{\left|E(S_n^j)E(S_n^{2m+2-j})\right| + 2\phi_k^{1/(2m+2)}C_n\right\}$$

$$\leq (2+b\phi_k^{1/(2m+2)})C_n + \sum_{j=2}^{2m}\binom{2m+2}{j}\left|E(S_n^j)E(S_n^{2m+2-j})\right| \tag{2.12}$$

for some constant b not involving k. In the last expression, we used $E(S_n) = 0$ for $j = 1$ and $j = 2m+1$. An application of Hölder's inequality to the last term in (2.12) yields that for $j = 2, \ldots, 2m$,

$$E(S_n^j)E(S_n^{2m+2-j}) \leq \left(E|S_n|^{2m}\right)^{j/2m}\left(E|S_n|^{2m}\right)^{(2m+2-j)/2m}$$

$$= \left(E|S_n|^{2m}\right)^{(2m+2)/2m}$$

which is at most (const) $\cdot\, n^{(1-\theta)(m+1)}$ by the induction hypothesis. Substituting this back into (2.12), we find that

$$E\{(S_n+\hat{S}_n)^{2m+2}\} \leq (2+b\phi_k^{1/(m+2)})C_n + a_1 n^{(1-\theta)(m+1)}$$

for some constants a_1, b not involving k. To prove (2.9) we need only to increase k, if necessary, to make the second term in the parenthesis $< \epsilon_1$.

Next we prove that, if $\epsilon > 0$, there is a constant a_2 and a value of k for which

$$C_{2n} \leq (2+\epsilon)C_n + a_2 n^{(1-\theta)(m+1)}, \ n \geq 1. \tag{2.13}$$

In fact, applying Minkowski's inequality and (2.9), we find that

$$\begin{aligned} C_{2n} &= E\left\{\left|S_n + \hat{S}_n + T_n - \sum_{j=2n+1}^{2n+k} \xi_j\right|\right\}^{2m+2} \\ &\leq \left[\left(E|S_n + \hat{S}_n|^{2m+2}\right)^{1/(2m+2)} + \sum_{j=n+1}^{n+k} \left(E|\xi_j|^{2m+2}\right)^{1/(2m+2)} \right. \\ &\quad \left. + \sum_{j=2n+1}^{2n+k} \left(E|\xi_j|^{2m+2}\right)^{1/(2m+2)}\right]^{2m+2} \\ &\leq \left[\left\{(2+\epsilon_1)C_n + a_1 n^{(1-\theta)(m+1)}\right\}^{1/(2m+2)} + 2kC_1^{1/(2m+2)}\right]^{(2m+2)} \\ &\leq \left[(1+\epsilon_1)\left\{(2+\epsilon_1)C_n + a_1 n^{(1-\theta)(m+1)}\right\}^{1/(2m+2)}\right]^{2m+2} \end{aligned}$$

if n is sufficiently large. Then

$$C_{2n} \leq (1+\epsilon_1)^{2m+2}\left\{(2+\epsilon_1)C_n + a_1 n^{(1-\theta)(m+1)}\right\}$$

if n is sufficiently large. If ϵ_1 is so small that $(1+\epsilon_1)^{2m+2}(2+\epsilon_1) \leq 2+\epsilon$, there must be a_2 for which (2.13) is true.

According to (2.13)

$$\begin{aligned} C_{2^r} &\leq (2+\epsilon)^r C_1 + a_2\left\{2^{(r-1)(1-\theta)(m+1)} + (2+\epsilon)2^{(r-2)(1-\theta)(m+1)} \right. \\ &\quad \left. + \cdots + (2+\epsilon)^{(r-1)}\right\} \\ &< (2+\epsilon)^r C_1 + a_2 \cdot 2^{(r-1)(1-\theta)(m+1)}\left(1 - \frac{2+\epsilon}{2^{(1-\theta)(m+1)}}\right)^{-1}, \ r \geq 1 \end{aligned}$$

if ϵ is so small that $2+\epsilon < 2^{(1-\theta)(m+1)}$. Such choice is possible because $(1-\theta)(m+1) > 1$ (*i.e.*, $1-\theta > 1/2$ and $m \geq 1$). Then if ϵ is chosen in this way,

$$C_{2^r} < a_3 \cdot 2^{r(1-\theta)(m+1)}, \ r \geq 0 \tag{2.14}$$

where

$$a_3 = C_1 + a_2 \cdot 2^{-(1-\theta)(m+1)} \left(1 - \frac{2+\epsilon}{2^{(1-\theta)(m+1)}}\right)^{-1}.$$

Finally if n is any positive integer it can be written in the form

$$n = 2^r + \nu_1 2^{r-1} + \cdots + \nu_r \leq 2^r + 2^{r-1} + \cdots + 1$$

where

$$2^r \leq n < 2^{r+1}$$

and each ν_j is either 0 or 1. Then S_n can be written as the sum of $r+1$ groups of sum containing 2^r, $\nu_1 2^{r-1}, \ldots$ terms and using Minkowski's inequality, (2.14) and the fact that the $\{\xi_j\}$ process is stationary,

$$\begin{aligned} C_n &\leq \left\{ \left(E|S_{2^r}|^{2m+2}\right)^{1/(2m+2)} + \left(E|S_{2^{r-1}}|^{2m+2}\right)^{1/(2m+2)} \right. \\ &\qquad \left. + \cdots + \left(E|S_1|^{2m+2}\right)^{1/(2m+2)} \right\}^{2m+2} \\ &\leq a_3 \left\{2^{r(1-\theta)/2} + 2^{(r-1)(1-\theta)/2} + \cdots + 1\right\}^{2m+2} \\ &= a_3 \left\{\frac{2^{(r+1)(1-\theta)/2} - 1}{2^{(1-\theta)/2} - 1}\right\}^{2m+2} \\ &\leq a \cdot n^{(1-\theta)(m+1)} \end{aligned}$$

for some constant a, as was to be proved. □

REFERENCES

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

COX, D. AND KIM, T. Y. (1995). "Moment bounds for mixing random variables useful in nonparametric function estimation", *Stochastic Processes and Their Applications*, **56**, 151–158.

DOUKHAN, P., LEÓN, J. AND PORTAL, F. (1984). "Vittesse de convergence dans le théorème central limite pour des variables aléatorie méalngeantes à valeurs un espace de Hilbert", *Compterendus de l'Academie Sciences de Paris*, **298**, 305–308.

IBRAGIMOV, I. A. (1962). "Some limit theorems for stationary processes", *Theory of Probabilty and Its Applications*, **7**, 361–392.

IBRAGIMOV, I. A. (1975). "A remark on the central limit theorem for dependent random variables", *Theory of Probability and Its Applications*, **20**, 135–140.

SEN, P. K. (1971). "A note on weak convergence of empirical processes for ϕ-mixing random variables", *The Annals of Mathematical Statistics*, **42**, 2131–2133.

TRUONG, Y. K. AND STONE, C. J. (1992). "Nonparametric function estimation involving time series", *The Annals of Statistics*, **20**, 77–97.

YOKOYAMA, R. (1980). "A note on weak convergence of empirical processes for ϕ-mixing variables", *Yokohama Mathematical Journal*, **28**, 37–40.

YOSHIHARA, K. (1974). "Extensions of Billingsley's theorems on weak convergence of empirical processes", *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **29**, 87–92.

YOSHIHARA, K. (1978). "Note on multidimensional empirical processes for ϕ-mixing random vectors", *Journal of Multivariate Analysis*, **8**, 584–588.

YU, B. (1993). "Density estimation in the L^∞ norm for dependent data with applications to the Gibbs sampler", *The Annals of Statistics*, **21**, 711-735.