*The Annals of Probability*
1973, Vol. 1, No. 5, 870–875

# A NOTE ON EMPIRICAL PROCESSES OF STRONG-MIXING SEQUENCES

BY CHANDRAKANT M. DEO

*University of California, Davis*

It is shown that a theorem of Bilingsley ((1968) *Convergence of Probability Measures*, Wiley; Theorem 22.1) about weak convergence of empirical processes of $\varphi$-mixing sequences also holds for a class of strong-mixing sequences.

**1. Introduction and main theorem.** Let $\{\xi_n : -\infty < n < \infty\}$ be a strictly stationary sequence of random variables. Denote by $\mathscr{M}_{-\infty}^{k}$ and $\mathscr{M}_{k}^{\infty}$ the $\sigma$-fields generated by random variables $\{\xi_n : n \leqq k\}$ and $\{\xi_n : n \geqq k\}$ respectively. Let $\varphi_0 = 1$, $\alpha_0 = 1$ and for $n \geqq 1$ define

$$\varphi_n = \sup\{|P(E_2 \mid E_1) - P(E_2)| : E_1 \in \mathscr{M}_{-\infty}^{0},\ E_2 \in \mathscr{M}_{n}^{\infty}\} \tag{1.1}$$

and

$$\alpha_n = \sup\{|P(E_1E_2) - P(E_1)P(E_2)| : E_1 \in \mathscr{M}_{-\infty}^{0},\ E_2 \in \mathscr{M}_{n}^{\infty}\}\,. \tag{1.2}$$

In the definition of $\varphi_n$ we are adopting the convention that $P(E_2 \mid E_1) = 0$ if $P(E_1) = 0$. Clearly $\alpha_n \leqq \varphi_n$. If $\varphi_n \to 0$ the sequence $\{\xi_n\}$ is called $\varphi$-mixing and if $\alpha_n \to 0$, $\{\xi_n\}$ is called strong-mixing.

Suppose now that $0 \leqq \xi_0 \leqq 1$, and $\xi_0$ has a continuous distribution function $F$ on $[0, 1]$. Let $\{F_n(t) : 0 \leqq t \leqq 1\}$ be the empirical process for $\xi_1, \xi_2, \cdots, \xi_n$, i.e., $F_n(t) = n^{-1} \sum_{i=1}^{n} I_{[0,t]}(\xi_i)$ where $I_{[0,t]}(\cdot)$ is the indicator function of the interval $[0, t]$. Normalize $F_n(t)$ as

$$Y_n(t) = n^{\frac{1}{2}}(F_n(t) - F(t))\,, \qquad 0 \leqq t \leqq 1\,. \tag{1.3}$$

Then the stochastic process $\{Y_n(t) : 0 \leqq t \leqq 1\}$ has sample paths in the Skorohod space $D[0, 1]$ of right-continuous functions on $[0, 1]$ with left-limits. For $0 \leqq t \leqq 1$, define the function $g_t$ by

$$g_t(x) = I_{[0,t]}(x) - F(t)\,. \tag{1.4}$$

Then Billingsley (1968), Theorem 22.1, has established the following result.

THEOREM (Billingsley). *Let $\xi_0$ have continuous distribution function $F$ with $F(0) = 0$ and $F(1) = 1$ and suppose further that $\{\xi_n\}$ satisfies the mixing condition*

$$\sum n^2 \varphi_n^{\frac{1}{2}} < \infty\,. \tag{1.5}$$

*Then the sequence $\{Y_n(t) : 0 \leqq t \leqq 1\}$ of normalized empirical processes converges*

Received October 21, 1972; revised January 15, 1973.
*AMS* 1970 *subject classifications.* Primary 60G15, 60F15.
*Key words and phrases.* weak convergence, empirical processes, strong-mixing sequences.

*weakly in* $D[0, 1]$ *to a Gaussian random function* $\{Y(t): 0 \leqq t \leqq 1\}$ *specified by*

$$E(Y(t)) = 0 \tag{1.6}$$

*and*

$$E\{Y(s)Y(t)\} = E\{g_s(\xi_0)g_t(\xi_0)\} + \sum_{k=1}^{\infty} E\{g_s(\xi_0)g_t(\xi_k)\} + \sum_{k=1}^{\infty} E\{g_s(\xi_k)g_t(\xi_0)\}. \tag{1.7}$$

*Furthermore, the series in* (1.7) *converges absolutely and the sample paths of* $Y$ *are continuous with probability one.*

The object of this note is to show that Billingsley's above-mentioned theorem, established for $\varphi$-mixing sequences, remains true for some strong-mixing sequences.

THEOREM 1. *Billingsley's theorem above remains true if the condition* (1.5) *is replaced by*

$$\sum n^2 \alpha_n^{\frac{1}{2}-\tau} < \infty \qquad \text{for some} \quad 0 < \tau < \tfrac{1}{2}. \tag{1.8}$$

Section 2 is devoted to the proof of Theorem 1. In Section 3 we give some examples of stationary sequences which satisfy the conditions of Theorem 1 but which are not even $\varphi$-mixing. Such sequences, therefore, are not covered by Billingsley's theorem nor Sen's (1971) extension of it.

**2. Proof of Theorem** 1. We begin with some lemmas for strong-mixing sequences.

Let $\|\cdot\|_p$ denote the $L_p$-norm of random variables.

LEMMA 1. (Davydov). *Let* $\{\xi_n\}$ *be a strong-mixing stationary sequence. Let* $r_1$, $r_2$, $r_3$ *be positive numbers such that* $r_1^{-1} + r_2^{-1} + r_3^{-1} = 1$. *Suppose that* $X$ *and* $Y$ *are random variables measurable with respect to the* $\sigma$*-fields* $\mathscr{M}_{-\infty}^0$, $\mathscr{M}_n^\infty$ *respectively and assume further that* $\|X\|_{r_1} < \infty$, $\|Y\|_{r_2} < \infty$. *Then*

$$|E(XY) - E(X)E(Y)| \leqq 10\alpha_n^{1/r_3}\|X\|_{r_1}\|Y\|_{r_2}. \tag{2.1}$$

PROOF. This lemma is due to Davydov (1970). The statement on page 492 in [2] however contains a misprint and there is no proof of the lemma given in [2]. We give a proof here for the sake of completeness.

If $\alpha_n = 0$ then $\mathscr{M}_{-\infty}^0$, $\mathscr{M}_n^\infty$ are independent and in that case both sides of (2.1) are zero. We assume, therefore, that $\alpha_n > 0$. Let $M = \alpha_n^{-1/r_1}\|X\|_{r_1}$ and $N = \alpha_n^{-1/r_2}\|Y\|_{r_2}$. Let $X_M$, $Y_N$ denote $X$ and $Y$ truncated at $M$ and $N$ respectively, i.e., $X_M = XI_{[|X| \leqq M]}$ and $Y_N = YI_{[|Y| \leqq N]}$. Also write $\hat{X}_M = X - X_M$ and $\hat{Y}_N = Y - Y_N$. We have, then

$$\begin{aligned} |E(XY) - E(X)E(Y)| &= |E\{(X_M + \hat{X}_M)(Y_N + \hat{Y}_N)\} \\ &\qquad - \{E(X_M) + E(\hat{X}_M)\} \times \{E(Y_N) + E(\hat{Y}_N)\}| \\ &\leqq \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} \end{aligned} \tag{2.2}$$

where

$$\begin{aligned} \mathrm{I} &= |E(X_M Y_N) - E(X_M)E(Y_N)| , \\ \mathrm{II} &= |E(X_M \hat{Y}_N) - E(X_M)E(\hat{Y}_N)| , \\ \mathrm{III} &= |E(\hat{X}_M Y_N) - E(\hat{X}_M)E(Y_N)| , \\ \mathrm{IV} &= |E(\hat{X}_M \hat{Y}_N) - E(\hat{X}_M)E(\hat{Y}_N)| . \end{aligned}$$

Now by Theorem 17.2.1 of [3], we have

$$\mathrm{I} \leqq 4MN\alpha_n = 4\alpha_n^{1/r_3}||X||_{r_1}||Y||_{r_2} . \tag{2.3}$$

Also it is easy to see that

$$\mathrm{II} \leqq 2\alpha_n^{1/r_3}||X||_{r_1}||Y||_{r_2} . \tag{2.4}$$

and

$$\mathrm{III} \leqq 2\alpha_n^{1/r_3}||X||_{r_1}||Y||_{r_2} . \tag{2.5}$$

To handle IV, let us define $r$ by $r_1^{-1} + r_2^{-1} = r^{-1}$. Then,

$$\mathrm{IV} \leqq E|\hat{X}_M \hat{Y}_N| + E|\hat{X}_M|E|\hat{Y}_N| .$$

Now,

$$\begin{aligned} E|\hat{X}_M \hat{Y}_N| &\leqq M^{-r+1}N^{-r+1}E(|X|^r|Y|^r) \\ &\leqq M^{-r+1}N^{-r+1}||X||_{r_1}^r||Y||_{r_2}^r , \quad \text{by Hölder's inequality} \\ &= \alpha_n^{1/r_3}||X||_{r_1}||Y||_{r_2} . \end{aligned}$$

Also it is easy to see that

$$E|\hat{X}_M|E|\hat{Y}_N| \leqq \alpha_n^{1/r_3}||X||_{r_1}||Y||_{r_2} .$$

Thus

$$\mathrm{IV} \leqq 2\alpha_n^{1/r_3}||X||_{r_1}||Y||_{r_2} . \tag{2.6}$$

The proof of the lemma is now complete from (2.2), (2.3), (2.4), (2.5) and (2.6).

LEMMA 2. *Let $\{\xi_n\}$ be a stationary, strong-mixing sequence of random variables satisfying*

$$E(\xi_0) = 0 , \qquad |\xi_0| \leqq C \leqq \infty , \quad \text{a.s.} \quad \textit{and} \quad \sum \alpha_n < \infty . \tag{2.7}$$

*Then the series*

$$E(\xi_0^2) + 2 \sum_{k=1}^{\infty} E(\xi_0\xi_k) \tag{2.8}$$

*converges absolutely.*

*Let $\sigma^2$ denote the sum of the series* (2.8). *Then* $0 \leqq \sigma^2 < \infty$ *and* $n^{-\frac{1}{2}} \sum_{k=1}^{n} \xi_k$ *has limiting normal distribution with mean* 0 *and variance* $\sigma^2$. [*Normal distribution with* $\sigma^2 = 0$ *is understood to be unit mass at* 0.]

PROOF. This is Theorem 18.5.4 of [3]. The case $\sigma^2 = 0$ is handled by noting that if $\sigma^2 = 0$ then $E(S_n^2/n) \to 0$ and so $S_n/n^{\frac{1}{2}} \to 0$ in probability.

LEMMA 3. *Let* $\{\xi_n\} = \{(\xi_n^{(1)}, \xi_n^{(2)}, \cdots, \xi_n^{(d)})\}$ *be a stationary, strong-mixing sequence of d-dimensional random vectors satisfying*

$$(2.9) \qquad E(\xi_0) = 0\,, \qquad |\xi_0^{(j)}| \leqq C < \infty \text{ a.s.}, \quad 1 \leqq j \leqq d\,, \quad \text{and } \sum \alpha_n < \infty\,.$$

*Then* $n^{-\frac{1}{2}} \sum_{k=1}^{n} \xi_k$ *has limiting d-dimensional normal distribution with mean* 0 *and covariances given by*

$$(2.10) \qquad \sigma_{ij} = E\{\xi_0^{(i)}\xi_0^{(j)}\} + \sum_{k=1}^{\infty} E\{\xi_0^{(i)}\xi_k^{(j)}\} + \sum_{k=1}^{\infty} E\{\xi_k^{(i)}\xi_0^{(j)}\}\,, \qquad 1 \leqq i, j \leqq d\,.$$

*The series in* (2.10) *converges absolutely.*

PROOF. See paragraph 3, page 177 of Billingsley (1968). Lemma 3 follows from Lemma 2 by an application of the Cramér–Wold technique.

The next lemma is basic and is comparable to Lemma 1, page 195 of [1].

LEMMA 4. *Let* $\{\xi_n\}$ *be a stationary, strong-mixing sequence of centered, Bernoulli random variables,* i.e., $P\{\xi_0 = 1 - \pi\} = \pi = 1 - P\{\xi_0 = -\pi\}$, $0 < \pi < 1$. *Let* $S_n = \sum_{i=1}^{n} \xi_i$. *Furthermore let* $\sum n^2 \alpha_n^{\frac{1}{2}-\tau} < \infty$ *for some* $0 < \tau < \frac{1}{2}$. *Then*

$$(2.11) \qquad E(S_n^4) \leqq 2400\{n^2\pi^{4/\gamma} + n\pi^{2/\gamma}\}[\sum_{k=0}^{\infty} (k+1)^2 \alpha_k^{\frac{1}{2}-\tau}]^2$$

*where* $\gamma$ *is defined by*

$$(2.12) \qquad \gamma = 2(\tfrac{1}{2} + \tau)^{-1} \quad \text{or equivalently} \quad 2\gamma^{-1} + (\tfrac{1}{2} - \tau) = 1\,.$$

PROOF. Note that $2 < \gamma < 4$ and

$$E|\xi_0|^\gamma = \pi(1-\pi)\{(1-\pi)^{\gamma-1} + \pi^{\gamma-1}\} \leqq \pi(1-\pi) \leqq \pi\,.$$

Now following the proof of Lemma 22.1, page 195 of [1],

$$(2.13) \qquad E(S_n^4) \leqq 4!\, n \sum |E\{\xi_0 \xi_i \xi_{i+j} \xi_{i+j+k}\}|$$

where the indices satisfy

$$(2.14) \qquad i, j, k \geqq 0\,, \qquad i + j + k \leqq n - 1\,.$$

Now by Lemma 1,

$$(2.15) \qquad \begin{aligned} |E[\xi_0(\xi_i \xi_{i+j} \xi_{i+j+k})]| &\leqq 10\alpha_i^{\frac{1}{2}-\tau} ||\xi_0||_\gamma ||\xi_i \xi_{i+j} \xi_{i+j+k}||_\gamma \\ &\leqq 10\alpha_i^{\frac{1}{2}-\tau} ||\xi_0||_\gamma ||\xi_i||_\gamma \\ &\leqq 10\alpha_i^{\frac{1}{2}-\tau} \pi^{2/\gamma}\,. \end{aligned}$$

Similarly

$$(2.16) \qquad |E[(\xi_0 \xi_i \xi_{i+j})\xi_{i+j+k}]| \leqq 10\alpha_k^{\frac{1}{2}-\tau}\pi^{2/\gamma}\,.$$

Using Lemma 1 again.

$$(2.17) \qquad \begin{aligned} |E[(\xi_0\xi_i)(\xi_{i+j}\xi_{i+j+k})| &\leqq |E(\xi_0\xi_i)||E(\xi_0\xi_k)| + 10\alpha_j^{\frac{1}{2}-\tau} ||\xi_0\xi_i||_\gamma ||\xi_0\xi_k||_\gamma \\ &\leqq 100\alpha_i^{\frac{1}{2}-\tau}\alpha_k^{\frac{1}{2}-\tau}\pi^{4/\gamma} + 10\alpha_j^{\frac{1}{2}-\tau}\pi^{2/\gamma}\,. \end{aligned}$$

Hence, by (2.13), (2.15), (2.16) and (2.17), we have

$$(2.18) \qquad E(S_n^4) \leqq 2400n\{\pi^{4/\gamma} \sum_{i,k\leqq j} \alpha_i^{\frac{1}{2}-\tau}\alpha_k^{\frac{1}{2}-\tau} + \pi^{2/\gamma} \sum_{j,k\leqq i} \alpha_i^{\frac{1}{2}-\tau}\}$$

where indices satisfy (2.14).

Now

$$\sum_{i,k\leqq j} \alpha_i^{\frac{1}{2}-\tau}\alpha_k^{\frac{1}{2}-\tau} \leqq \sum_{j=0}^{n-1}[\sum_{i=0}^{\infty}\alpha_i^{\frac{1}{2}-\tau}]^2 = n[\sum_{i=0}^{\infty}\alpha_i^{\frac{1}{2}-\tau}]^2 . \tag{2.19}$$

And

$$\sum_{j,k\leqq i} \alpha_i^{\frac{1}{2}-\tau} \leqq \sum_{i=0}^{n-1}(i+1)^2\alpha_i^{\frac{1}{2}-\tau} \leqq \sum_{i=0}^{\infty}(i+1)^2\alpha_i^{\frac{1}{2}-\tau} . \tag{2.20}$$

Using the bounds (2.19) and (2.20) in (2.18) we get the required result.

Outline of Proof of Theorem 1. We simply indicate the modifications to be made in the well-written proof of Theorem 22.1 of [1]. Same notation is used.

The first part of the proof which shows that we can, without losing generality, assume that $\xi_0$ is uniformly distributed on [0, 1] remains unchanged.

The second part of Billingsley's proof shows that the finite-dimensional distributions of $\{Y_n(t)\}$ converge to those of $\{Y(t)\}$. Using our Lemma 3 the same arguments apply here.

It remains to show that given $\varepsilon > 0$, $\eta > 0$, we can find $\delta > 0$, $0 < \delta < 1$ such that $P\{w(Y_n, \delta) \geqq \varepsilon\} \leqq \eta$ for all sufficiently large $n$.

Applying our Lemma 4 to $\{g_t(\xi_n) - g_s(\xi_n)\}$ we get

$$E\{|\sum_{i=1}^{n}(g_t(\xi_i) - g_s(\xi_i))|^4\} \leqq K_1(n^2|t-s|^{4/\gamma} + n|t-s|^{2/\gamma})$$

where $K_1$ depends on the $\alpha$-sequence only.

Therefore if $\varepsilon < 1$, and $\varepsilon/n \leqq |t-s|^{2/\gamma}$ we get

$$E\{|Y_n(t) - Y_n(s)|^4\} \leqq \frac{2K_1}{\varepsilon}|t-s|^{4/\gamma} . \tag{2.12}$$

Our (2.21) replaces (22.15) of [1].

Now by our (2.21) and Theorem 12.2 of [1] we get

$$P\{\max_{i\leqq m}|Y_n(s+ip) - Y_n(s)| \geqq \lambda\} \leqq \frac{K_2}{\varepsilon\lambda^4} m^{4/\gamma}p^{4/\gamma} . \tag{2.22}$$

where $K_2$ depends only on the $\alpha$-sequence. This replace (22.16) of [1]. From here to (22.19) of [1] everything remains unchanged. We need to change (22.19) of [1] to

$$\left(\frac{\varepsilon}{n}\right)^{\gamma/2} \leqq p < \frac{\varepsilon}{n^{\frac{1}{2}}} \tag{2.23}$$

and (22.20) of [1] then is changed to

$$P\{\sup_{s\leqq t\leqq s+mp}|Y_n(t) - Y_n(s)| \geqq 4\varepsilon\} \leqq \frac{K_2}{\varepsilon^5} m^{4/\gamma}p^{4/\gamma} . \tag{2.24}$$

Now choose $1 > \delta > 0$ so that $(K_2/\varepsilon^{-5})\delta^{4/\gamma-1} < \eta$. Such a choice of $\delta$ is possible because $4/\gamma - 1 > 0$. It then follows that

$$P\{\sup_{s\leqq t\leqq s+\delta}|Y_n(t) - Y_n(s)| \geqq 4\varepsilon\} < \eta\delta , \tag{2.25}$$

provided there exists a $p$ and an integer $m$ such that (2.23) holds and $mp = \delta$.

But this is equivalent to requiring the existence of an integer $m$ such that $(\delta/\varepsilon)n^{\frac{1}{2}} < m \leqq (\delta/\varepsilon^{\gamma/2})n^{\gamma/2}$. Since $\gamma > 2$ this is clearly possible for all sufficiently large $n$. Rest of the proof is same as in [1] and Theorem 1 is completely proven.

**3. Examples.** In this section we give some examples of stationary sequences to which Theorem 1 applies but which are not $\varphi$-mixing.

It is known (see e.g. Theorem 17.3.2 of [3]) that a stationary Gaussian sequence is $\varphi$-mixing iff it is $m$-dependent for some $m$ and this is the case iff its spectral density $f(\lambda)$ is of the form

$$f(\lambda) = |P(e^{i\lambda})|^2\,, \qquad -\pi \leqq \lambda \leqq \pi \tag{3.1}$$

where $P(z)$ is a polynomial of a complex variable $z$.

Let now $\{\xi_n\}$ be a stationary Gaussian sequence with spectral density $f(\lambda)$ such that (i) $f(\lambda) \neq 0$, $-\pi \leqq \lambda \leqq \pi$ and (ii) $f(\lambda)$ has bounded seventh derivative or more generally has a sixth derivative which satisfies Hölder condition of some order $\beta > 0$. It then follows from Lemma 10.6 of [5] or Theorem 8, page 253 of [4] that the $\alpha$-sequence defined for $\{\xi_n\}$ as in (1.2) satisfies the condition (1.8). If now, moreover, $f(\lambda)$ is not of the form (3.1) then $\{\xi_n\}$ is strong-mixing satisfying (1.8) but is not $\varphi$-mixing. Let now $\Phi$ be the cumulative distribution function of the standard normal distribution and let $\{\zeta_n\}$ be a stationary Gaussian sequence which is not $\varphi$-mixing but which satisfies (1.8). Define

$$\xi_n = \Phi(\zeta_n)\,, \qquad -\infty < n < \infty\,.$$

Then $\{\xi_n\}$ is a stationary sequence of random variables which satisfies all the conditions of our Theorem 1 but which is not $\varphi$-mixing. This is clear in view of the fact that $\Phi$ is one-one so that both $\{\zeta_n\}$ and $\{\xi_n\}$ have the same $\alpha$-sequence and the same $\varphi$-sequence. As a specific example let $\{\zeta_n\}$ be a stationary, Markovian, Gaussian sequence with spectral density

$$f(\lambda) = (1-p^2)(1-2p\cos\lambda + p^2)^{-1}\,, \qquad -\pi \leqq \lambda \leqq \pi \ \text{ and } \ 0 < p < 1\,.$$

Then $\{\Phi(\zeta_n)\}$ is a stationary sequence of uniformly distributed random variables which is not $\varphi$-mixing but to which our Theorem 1 applies.

## REFERENCES


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

[2] Davydov, Yu. A. (1970). The invariance principle for stationary processes. *Theor. Probability Appl.* **15** 487–498.

[3] Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

[4] Ibragimov, I. A. and Rozanov, Yu. A. (1970). *Gaussian Stochastic Processes.* Izdatelstvo Nauka, Moscow (in Russian).

[5] Rozanov, Yu. A. (1967). *Stationary Random Processes.* Holden-Day, San Francisco.

[6] Sen, P. K. (1971). A note on weak convergence of empirical processes for $\phi$-mixing random variables. *Ann. Math. Statist.* **42** 2131–2134.


Departmemt of Mathematics
University of California
Davis, California 95616