YOKOHAMA MATHEMATICAL
JOURNAL VOL. 32, 1984

A NOTE ON THE ALMOST SURE APPROXIMATION OF THE EMPIRICAL PROCESS OF WEAKLY DEPENDENT RANDOM VECTORS

By

SOMPONG DHOMPONGSA*

(Received December 8, 1983)

ABSTRACT. In this paper we improve upon results on the almost sure approximation of the empirical process of weakly dependent random vectors, recently obtained by Berkes and Philipp (1977) and Philipp and Pinzur (1980). For strongly mixing sequences we improve the bounds on the mixing rates, and for absolutely regular sequences we improve the error term. We also extend these results to random vectors which are functions of the given sequence.

1. Introduction and statement of results

Let $\{\xi_n, n\geqq 1\}$ be a strictly stationary sequence of random vectors with values in $\boldsymbol{R}^e$ and let $\mathscr{M}_a^b$ denote the σ-field generated by $\{\xi_n, a\leqq n\leqq b\}$. The sequence is called absolutely regular if for some $\beta(n)\downarrow 0$

(1.1) $$E\sup_{A\in\mathscr{M}_{k+n}^{\infty}}|P(A\mid\mathscr{M}_1^k)-P(A)|\leqq\beta(n)$$

for all $k, n\geqq 1$. Th sequence $\{\xi_n, n\geqq 1\}$ is said to satisfy a strong mixing condition if for some $\rho(n)\downarrow 0$

(1.2) $$|P(AB)-P(A)P(B)|\leqq\rho(n)$$

for all $A\in\mathscr{M}_{k+n}^{\infty}$, $B\in\mathscr{M}_1^k$ and all $k, n\geqq 1$. Since (1.2) is equivalent with

$$\sup_{A\in\mathscr{M}_{k+n}^{\infty}}E\,|\,P(A\mid\mathscr{M}_1^k)-P(A)|\leqq\rho^*(n)$$

with $\rho(n)\leqq\rho^*(n)\leqq 2\rho(n)$ we see that every absolutely regular sequence satisfies a strong mixing condition. The converse is not true; there are examples of strongly mixing sequences that are not absolutely regular. The well-known ϕ-mixing condition which is still more restrictive than (1.1) will not be needed in this paper.

---

* Supported by the Thai Government.

AMS Subject Classification. Primary: 60F15

Running Head. Empirical processes.

Key words and phrases. Invariance principles, empirical processes, random vectors, Kiefer processes, mixing random variables.

Let f be a measurable mapping from the space of infinite sequences of vectors in $\boldsymbol{R}^e$ into $\boldsymbol{R}^q$. Define

(1.3) $$\eta_n=f(\xi_n, \xi_{n+1}, \cdots) \quad n\geqq 1$$

and let

$$\eta_{nm}=E(\eta_n \mid \mathscr{M}_n^{n+m}) \quad n\geqq 1,\ m\geqq 0 .$$

We shall assume that η_n can be closely approximated by η_{nm} in the form

(1.4) $$E\,|\eta_n-\eta_{nm}|\leqq\psi(m) \quad n\geqq 1,\ m\geqq 0 ;$$

Let F denote the common distribution function of η_n. The empirical process of $\{\eta_n, n\geqq 1\}$ is defined as

$$R(s, t)=\sum_{n\leqq t}(1\{\eta_n\leqq s\}-F(s)) , \quad s\in\boldsymbol{R}^q , \quad t\geqq 0 .$$

(For two points $u=(u_1, \cdots, u_q)$ and $v=(v_1, \cdots, v_q)$ in $\boldsymbol{R}^q$, we write $u\leqq v$ to mean that $u_i\leqq v_i$ for all $1\leqq i\leqq q$.)

In recent papers Berkes and Philipp (1977) and Philipp and Pinzur (1980) proved approximation theorems for $R(s, t)$ by a Kiefer process $K(s, t)$ with various degrees of accuracy. The formal definition of a Kiefer process is given in these papers, but it also can be viewed as a $C(\boldsymbol{R}^q)$-valued Brownian motion with covariance structure given below. The purpose of this note is to improve upon these results by improving the bounds on the mixing rates and the error terms in some of them.

Write

(1.5) $$g_n(s)=1\{\eta_n\leqq s\}-F(s) , \quad s\in\boldsymbol{R}^q .$$

Then the two series defining the covariance function

(1.6) $$\Gamma(s, s')=E\{g_1(s)g_1(s')\}+\sum_{n\geqq 2} E\{g_1(s)g_n(s')\} + \sum_{n\geqq 2} E\{g_n(s)g_1(s')\}$$

converge absolutely.

Theorem 1. *Let $\{\xi_n, n\geqq 1\}$ be a strictly stationary, absolutely regular sequence of random vectors with values in $\boldsymbol{R}^q$ and common continuous distribution function F. Suppose that the mixing rate $\beta(n)$ satisfies*[1]

(1.7) $$\beta(n)\ll n^{-(1+\theta)q-2}$$

[1] For two sequences of real numbers (x_n) and (y_n), $x_n\ll y_n$ means there exists a constant $c\geqq 0$ such that $|x_n|\leqq c\,|y_n|$ for all n.

for some $\theta>0$. *Then the series in* (1.6) (*with* η_n *replaced by* ξ_n *in* (1.5)) *converge absolutely. Moreover, without changing its distribution we can redefine the empirical process* $\{R(s, t),\ s\in \boldsymbol{R}^q,\ t\geqq 0\}$ *of* $\{\xi_n, n\geqq 1\}$ *on a richer probability space on which there exists a Kiefer process* $\{K(s, t),\ s\in \boldsymbol{R}^q,\ t\geqq 0\}$ *with covariance function* $E(K(s, t)K(s', t'))=\Gamma(s, s')\min(t, t')$ *such that with probability* 1

$$\sup_{t\leqq T}\sup_{s\in \boldsymbol{R}^q}|R(s, t)-K(s, t)|\ll T^{(1/2)-\lambda} \tag{1.8}$$

where $\lambda>0$ *only depends on* q *and* θ.

For the analogous result on the empirical process of $\{\eta_n, n\geqq 1\}$ we have to put restrictions on the common distribution function F of η_n.

Theorem 2. *Let* $\{\xi_n, n\geqq 1\}$ *be a strictly stationary, absolutely regular sequence of random vectors with values in* $\boldsymbol{R}^e$. *Suppose that the random vectors* $\eta_n\in \boldsymbol{R}^q$ *are defined by* (1.3) *and satisfy* (1.4) *with*

$$\psi(m)\ll m^{-(8+\delta)} \tag{1.9}$$

for some $\delta>0$. *Suppose that the mixing rate* $\beta(n)$ *of the sequence* $\{\xi_n, n\geqq 1\}$ *in* (1.1) *satisfies* (1.7). *Assume further that all the marginals* $F_p(1\leqq p\leqq q)$ *of the common distribution function* F *of* η_n *concentrate on* $[0, 1]$ *and satisfy a Lipschitz condition of the form*

$$|F_p(x)-F_p(y)|\leqq C\cdot|x-y| \quad x, y\in[0, 1]\,. \tag{1.10}$$

Here $C=C_F$ *is a constant. Then the conclutions of Theorem* 1 *remain valid for the empirical process* $\{R(s, t),\ s\in[0, 1]^q,\ t\geqq 0\}$ *of* $\{\eta_n, n\geqq 1\}$.

For strongly mixing random variables we obtain similar results.

Theorem 3. *Let* $\{\eta_n, n\geqq 1\}$ *be a strictly stationary sequence of random vectors in* $\boldsymbol{R}^q$ *with common, but arbitrary distribution function and satisfying a strong mixing condition with mixing rate*

$$\rho(n)\ll n^{-(1+\theta)q-2}\,.$$

Then the conclusions of Theorem 1 *remain valid with the error term in* (1.8) *replaced by* $T^{1/2}(\log T)^{-\lambda}$.

Similarly a result analogous to Theorem 2 can be formulated and proved for strongly mixing sequences. Recently, Yoshihara (1979) proved that the mixing rate in Theorem 1 of Berkes and Philipp (1977) could be relaxed to $\rho(n)\ll n^{-3-\theta}$. However, there appears to be a misprint in his paper since one term in the exponential bound in Lemma 2 is $\exp(-A^2C_0^2l^\rho\log\log N)$ when it must be

$\exp(-A^2C_0^2l^{-\rho}\log\log N)$ to make the proofs of his Lemmas 3 and 4 work.

2. Sketch of proofs

Since the time the papers [3] and [11] were written several of their ingredients were improved. We now employ these improvements and add a few new ideas.

2.1 As has been shown in sections 4 and 5 of [11] there is no loss of generality to assume that all the marginals of the common distribution function F in Theorem 1 have uniform distribution over $[0, 1]$. In fact, by the continuity of F, we can show that the sequence of random vectors $\{\eta_n, n\geqq 1\}$ in $\boldsymbol{R}^q$ constructed in section 5 of [11] whose components are uniformly distributed over $[0, 1]$ such that[2] $(\xi_n, n\geqq 1)=(F^{-1}\eta_n, n\geqq 1)$ in distribution satisfies an absolutely regular condition (1.1) with the rate in (1.7). Now, under the new assumption, Theorem 1 (case of uniform marginals) is a part of Theorem 2 where $C_F=1$ and $\psi(m)=0$ for all m. Hence, we prove Theorem 2 first. We basically follow the proof of Theorem 1 of [3]. We assume without loss of generality that $\theta\leqq 1/3$. Define

$$t_k=[k^{21/\varepsilon\theta+72q/\alpha\theta}]$$

where

$$\varepsilon=\frac{3+\theta}{3(2+\theta)} \tag{2.1}$$

$$\alpha=\frac{\theta}{8(2+\theta)(4+3q)}$$

$$r_k=\left[\frac{4}{\alpha}\log k/\log 2\right]$$

and

$$d_k=2^{qr_k}\,.$$

In the proof of Proposition 1 of [11] we replace the application of Yurinskii's (1977) theorem by the following partial generalization of Dehling (1980).

[2] For a nondecreasing function G on $\boldsymbol{R}$, define the inverse G^{-1} of G on the smallest interval containing the range of G by

$$G^{-1}(y)=\inf\{x\colon G(x)\geqq y\}\,.$$

Let $F_p(1\leqq p\leqq q)$ be the marginals of F. Define a mapping

$$F^{-1}\colon [0, 1]^q\to\boldsymbol{R}^q$$

by

$$F^{-1}(u_1, \cdots, u_q)=(F_1^{-1}(u_1), \cdots, F_q^{-1}(u_q))\,, \qquad u=(u_1, \cdots, u_q)\in[0, 1]^q\,.$$

Theorem A. *Let $\{x_n, n\geqq 1\}$ be a weakly stationary, absolutely regular sequence of random variables with values in $\boldsymbol{R}^d$, centered at expectations and with $(2+\delta)$-th moments uniformly bounded by $\rho_{2+\delta}$, where $0<\delta\leqq 1$. Suppose that the mixing rate $\beta(n)$ satisfies*

$$\beta(k)\ll k^{-(1+\varepsilon)(1+2/\delta)} \tag{2.2}$$

for some $0<\varepsilon\leqq 1$. Then the two series defining the covariance function T of the sequence $\{x_n, n\geqq 1\}$ defined by

$$T(f, g)=E\{f(x_1)g(x_1)\}+\sum_{n\geqq 2} E\{f(x_1)g(x_n)\}+\sum_{n\geqq 2} E\{f(x_n)g(x_1)\}$$

converge absolutely. Moreover, there exist constants $\lambda>0$ and C depending only on ε, δ and the constant implied by $\ll$ in (2.2) such that the Proholov distance

$$\pi(\mathscr{L}(n^{-1/2}\sum_{j\leqq n} x_j);\ N(0, T))\leqq Cn^{-\lambda}d^{1/2}(1+\rho_{2+\delta}^{1/3})\ .$$

Here, for any random variable X, $\mathscr{L}(X)$ stands for the distribution of X.

For the proof of the oscillation estimates in section 3.3 of [3] we replace the application of the classical exponential bound Lemma 2 of [11] by the following one.

Proposition 1. *Let $\{\xi_n, n\geqq 1\}$ and $\{\eta_n, n\geqq 1\}$ be as in Theorem 2. Using the same ordering of elements in $\boldsymbol{R}^q$ as in [11] we put*

$$x_n=x_n(s, s')=g_n(s')-g_n(s)$$

by (1.5) and

$$D_n=D_n(s, s')=F(s')-F(s)$$

for $s<s'$ in $[0, 1]^q$. Then there exist constants $A\geqq 1$ and $C\geqq 1$ depending only on q, θ and the constant implied by $\ll$ in (1.9) such that

$$P\left\{\left|\sum_{n=H+1}^{H+2^N} x_n\right|>ARD^{\alpha}2^{(1/2)N}(\log N)^{1/2}\right\} \tag{2.3}$$
$$\leqq C(\exp(-8RD^{-\alpha}\log N)+R^{-2}2^{(-1/2)qN(1+(1/2)\theta)})$$

for all $H\geqq 0$, $N, R\geqq 1$ and all $s, s'\in[0, 1]^q$ with $s<s'$. Here α is defined in (2.1).

The proof of Proposition 1 is a modification of the proof of Proposition 3.3.1 of [10]. We will give a sketch in section 3.

In section 3.3 of [11] it was shown that if the t_k's have polynomial growth, then Lemmas 3 and 4 of [11] continue to hold. Note that in their proofs the lengths of all the sums are powers of 2. Hence for the proof of Theorem 2 it remains to check the details of section 3.5 of [11]. We replace the application of Theorem A by a stronger variant, due to Dehling and Philipp (1980). This variant can be simply formulated as follows. If in Theorem A of [11] the condition corresponding to the ϕ-mixing condition is replaced by[3]

$$E \sup_{C \in L_k} |P(C \mid \bigvee_{j<k} L_j) - P(C)| \leqq \beta_k$$

then the conclusion of Theorem A remains valid with ϕ_k replaced by $\beta_k^{1/2}$. There are no more changes necessary. This concludes the proof of Theorem 2.

2.2 The proof of Theorem 1 can be proved using Theorem 2 by the technique employing in section 4 of [11].

2.3 The proof of Theorem 3 is the same as the proof of Theorem 2 of [11] except that we apply Proposition 1 of the present paper instead of Proposition 1 of [11]. We also apply Proposition 2.1 of [9] to obtain a similar result as in Proposition 4.1 of [3].

3. Sketch of proof of Proposition 1.

Except for a few minor changes the proof of Proposition 1 is almost identical with the proof of Proposition 3.3.1 in [10]. We can assume without loss of generality that $H=0$. We note that

$$E(\sum_{n \in N} x_n)^2 = N\sigma^2 + 0(D^{1/2})$$

where

$$\sigma^2 = Ex_1^2 + 2\sum_{n=2}^{\infty} E(x_1 x_n) \ll D^{3/4} .$$

The main observation being that for fixed m the random variables η_{nm} also are absolutely regular. The details can be worked out in the same way as in [3] and [11]. However, Lemma 3.2.1 in [10] has to be replaced by the following one. The proof is the same.

Lemma 1. *Let X and Y be random vectors in* $[0, 1]^q$ *with* $E|X-Y| \leqq \varepsilon$. *Suppose that the marginals F_p of the distrition of X satisfy a Lipschitz condition* (1.10). *Then for all* $t \in [0, 1]^q$

$$E\,|1\{X \leqq t\} - 1\{Y \leqq t\}| \leqq 2(C_F+1)\varepsilon^{1/2} .$$

[3] For a family of σ-fields $\{L_\lambda : \lambda \in \Lambda\}$ of a given space, we write $\bigvee_{\lambda \in \Lambda} L_\lambda$ for the smallest σ-field of the space containing the set $\bigcup_{\lambda \in \Lambda} L_\lambda$.

Thus for the proof of Proposition 1 we can assume $D \geqq 2^{-(2+\theta)qN}$ since otherwise we apply Chebyshev's inequality. As on p. 329 of [10] we define blocks H_j and I_j inductively, each of length 2^t where

$$t=\left[\left(2\alpha(2+\theta)q+1+\frac{2+\theta}{4}q\right)N/((1+\theta)q+2)\right]+1\,,$$

and the random variables y_j and z_j are also defined in the same fashion where here we put

$$m=[\{\max(2^{(2\alpha(2+\theta)q+1+(2+\theta/4)q)N},\, 2^t D^{-1/2},\, 2^{((1+\theta)q+2/2)t})\}^{2/(8+\delta)}]+1\,.$$

We replace $R^{-3}N^{-1.1}$ in Lemma 3.3.1 of [10] by $R^{-2}2^{-(1/2)(1+(\theta/2))qN}$. We use the estimate $\|v_j\|_2 \ll 2^{-(2+\theta/2)qt}$ instead of the estimate of $\|v_j\|_4$ in Lemma 3.3.4. Thus $R^{-4}N^{-3/2}$ in Lemma 3.3.5 is changed to $R^{-2}2^{-(1/2)(1+(\theta/2))qN}$. We define $k=2^{N-t-1}$ and replace Lemma 3.3.3 by

$$\sum_{j\leqq k} Ey_j^2 \ll D^{3\alpha}2^N\,.$$

Hence Lemma 3.3.6 becomes

$$P\{\sum_{j\leqq k} E(y_j^2 \mid \mathscr{L}_{j-1}) \geqq 2RBD^{3\alpha}2^N\} \ll R^{-2}2^{-(1/2)(1+(\theta/2))qN}\,. \tag{3.1}$$

Finally we prove the analogue of Lemma 3.3.7 of [10], that is we prove

$$P\{|\sum_{j\leqq k} y_j| \geqq 8RBD^{\alpha}2^{N/2}(\log N)^{1/2}\} \ll \exp(-8RD^{-\alpha}\log N) + R^{-2}2^{-(1/2)(1+(\theta/2))qN}\,. \tag{3.2}$$

Here we define

$$\begin{aligned} U_n &= \sum_{j\leqq n} y_j\,, \quad n\leqq k \\ &= U_k \quad\ , \quad n>k \\ U_0 &= 0 \\ S_n^2 &= \sum_{j\leqq n} E(y_j^2 \mid \mathscr{L}_{j-1})\,, \quad n\leqq k \\ &= S_k^2 \quad\qquad\qquad , \quad n>k\,. \end{aligned}$$

We choose $C=C_q>1$ so that

$$8(\log N)^{1/2} \leqq C2^{\alpha(2+\theta)qN}\,, \quad N\geqq 1\,.$$

Then we take B in (3.1) so large that $B\geqq C^2$. We put

$$\lambda=\frac{2}{C}D^{-2\alpha}(\log N)^{1/2}2^{-N/2}\,,$$

$$K=4RBD^{3\alpha}2^N\,,$$

$$c=\frac{1}{\lambda},$$

and

$$T_n=\exp(\lambda U_n-\lambda^2 S_n^2), \quad n\geqq 1$$
$$T_0=1 .$$

We finish the proof of (3.2) as on p. 332 of [10].

The proof of relation (2.3) can now be completed in the same way as on p. 332–333 of [10].

Acknowledgement

I am very grateful to Professor Walter Philipp for his constant encouragement, helpful advice, patience and assistance throughout the preparation of this paper.

References

[1] R. B. Ash, *Real analysis and probability*. Academic Press, New York (1972).
[2] I. Berkes and W. Philipp, *Approximation theorems for independent and weakly dependent random vectors*. Ann. Probab. **7** (1979), 29–54.
[3] I. Berkes and W. Philipp, *An almost sure invariance principle for the empirical distribution function of mixing random variables*. Z. Wahrscheinlichkeitstheorie verw. Gebiete. **41** (1977), 115–137.
[4] H. Dehling, *Limit theorems for sums of weakly dependent Banach space valued random variables*. Preprint (1980).
[5] H. Dehling and W. Philipp, *Almost sure invariance principles for weakly dependent vector-valued random variables*. Preprint (1980).
[6] J. L. Doob, *Stochastic processes*. Wiley, New York (1953).
[7] A. Dvoretzky, *Asymptotic normality for sums of dependent random variables*. Proc. 6th Berkeley Sympos. Math. Statist. Probab. Vol. II, (1970), 513–535.
[8] I. A. Ibragimov, *Some limit theorems for stationary processes*. Theory Prob. Appl., **7** (1962), 349–382.
[9] J. Kuelbs and W. Philipp, *Almost sure invariance principles for partial sums of mixing B-valued random variables*. Ann. Probab., **8** (1980).
[10] W. Philipp, *A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables*. Ann. Probab., **5**(3) (1977), 319–350.
[11] W. Philipp and L. Pinzur, *Almost sure approximation theorems for the multivariate empirical process*. Z. Wahrscheinlichkeitstheorie verw. Gebiete. **54** (1980), 1–13,
[12] W. F. Stout, *Almost sure convergence*. Academic Press, New York (1974).
[13] V. A. Volkonokii and Yu. A. Rozanov *Some limit theorems for random functions*. Theory Probab. Appl., **4** (1959), 178–197.
[14] M. Wichura, *Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments*. Ann. Prob., **1** (1973), 272–296.
[15] K. Yoshihara, *Note on an almost sure invariance principle for some empirical processes*. Yokohama Math. J., **27** (1979), 105–110.

[16] V. V. Yurinskii, *On the error of the Gaussian approximation for convolutions*. Theory Prob. Appl., **22** (1977), 236–247.

University of Illinois,
Urbana Il. 61801,
U.S.A.