J. Korean Math. Soc. **33** (1996), No. 4, pp. 909–928

AN EMPIRICAL LIL FOR STATIONARY MARTINGALE DIFFERENCES: AN INVARIANCE PRINCIPLE APPROACH

JONGSIG BAE

1. Introduction

In this paper, an empirical law of the iterated logarithm is investigated in the context of a stationary and ergodic martingale differences whose values taking in a measurable space.

We assume the integrability condition on the metric entropy with bracketing and derive an empirical law of the iterated logarithm from an invariance principle of an empirical central limit theorem for stationary martingale differences whose proof depends on a chaining argument with stratification (see Bae, J. and Levental, S. (1995)).

We restate the invariance principle of an empirical central limit theorem for stationary martingale differences as in the form of Dudley and Philipp (see Dudley, R. M. and Philipp, W. (1983)) and derive an empirical law of the iterated logarithm for stationary martingale differences via a method adapted from Ossiander's empirical law of the iterated logarithm for iid random variables (see Ossiander, M. (1987)).

The paper generalizes the empirical law of the iterated logarithm of iid sequences of Ossiander (1987). Examples on Markov chains and the Baker's transformation are provided.

Let $X_1, X_2, \dots$ be a stationary and ergodic process taking values in a measurable space $(S, \mathcal{B})$. From stationarity we may assume that the process $X = (X_i)$ is double-sided (see, for example, Breiman, L. (1968)). Define for each $i \in Z$, $\mathcal{F}_i$ to be the σ-field generated by

---

Received December 27, 1995.

1991 AMS Subject Classification: Primary 60F17; secondary 60F05.

Key words: Stationary martingale differences, empirical LIL, invariance principle, empirical CLT, Markov chains, Baker's transformation.

$\{X_j : j \leq i\}$. Assume that for each $f \in \mathcal{F}$, $\{f(X_i), \mathcal{F}_i\}$ is a martingale differences. I. e. $E(f(X_i)|\mathcal{F}_{i-1}) = 0$ for each i. Define

$$S_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(X_i) \quad \text{for each } f \in \mathcal{F}. \tag{1}$$

We will prove that, under the assumption of the integrability of the metric entropy with bracketing, $\mathcal{F}$ satisfies an empirical law of the iterated logarithm of Strassen type (see, for example, Kuelbs, J. and Dudley, R. M. (1980)).

2. A Setup and Main Results

We use the following setup to state the problem in a concrete fashion. Let $(S, \mathcal{B})$ be a measurable space. Choose $(\Omega = S^Z, \mathcal{T} = \mathcal{B}^Z, P)$ as our basic probability space. Let $T : \Omega \to \Omega$ be the left shift transformation. Assume that T is ergodic. Denote by $X = ..., X_{-1}, X_0, X_1, ...$ the coordinate maps on Ω. From our assumption it follows that $\{X_i\}_{i \in Z}$ is a stationary and ergodic process. Next we define for each $i \in Z$ a σ-field $M_i = \sigma(X_j : j \leq i)$ and

$$H_i = \{f : \Omega \to R : f \text{ is } M_i \text{ measurable and } f \in L^2(\Omega)\}.$$

For each $f \in L^2(\Omega)$ we simply denote $E_{i-1}(f)$ to mean $E(f|M_{i-1})$ and

$$H_0 \ominus H_{-1} = \{f \in H_0 : E(f \cdot g) = 0 \text{ for each } g \in H_{-1}\}.$$

Finally on $L^2(\Omega)$ we define a metric d by $d(f,g) = [E(f-g)^2]^{1/2}$, for n$f$, $g \in L^2(\Omega)$.

Let $\mathcal{F} \subseteq H_0 \ominus H_{-1}$. From our setup it follows that for every $f \in \mathcal{F}$, $\{f(T^i(X)), M_i\}$ is a stationary martingale difference sequence. We write

$$V_i = T^i(X),$$

and

$$V = T^0(X)(= X).$$

For every $f \in \mathcal{F}$ we define

$$S_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(V_i). \tag{2}$$

Next we define the metric entropy with bracketing (see, for example, Dudley, R. M. (1984)).

DEFINITION 1. Let $\delta > 0$. For the metric space $(\mathcal{F}, d)$ we define the covering number with bracketing $\nu^B(\delta, \mathcal{F}, d)$, or $\nu^B(\delta)$ if there is no risk of ambiguity, as the smallest n for which there exists $\{f^l_{0,\delta}, f^u_{0,\delta}, \cdots, f^l_{n,\delta}, f^u_{n,\delta}\} \subseteq H_0$ so that for every $f \in \mathcal{F}$ there exist some $0 \le i \le n$ satisfying $f^l_{i,\delta} \le f \le f^u_{i,\delta}$ and $d(f^l_{i,\delta}, f^u_{i,\delta}) < \delta$. We also define the metric entropy with bracketing to be $\ln \nu^B(\delta, \mathcal{F}, d)$.

We will use the following notations: Let $\mathcal{F}_\delta = \{(f^l_{i,\delta}, f^u_{i,\delta}) : 0 \le i \le \nu^B(\delta)\}$ for $0 < \delta \le 1$ and $\mathcal{F}_0 = \mathcal{F} \times \mathcal{F}$. Let $\tilde{\mathcal{F}} = \cup_{0 \le \delta \le 1} \mathcal{F}_\delta$. For a function $\varphi : \mathcal{F} \to R$, we let $||\varphi||_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |\varphi(f)|$ denote the sup of $|\varphi|$ over $\mathcal{F}$. We write $||\cdot||$ in stead of $||\cdot||_{\mathcal{F}}$ when there is no risk of ambiguity. We define

$$\mathcal{M} = \{f \in L^2(\Omega) : E(f) = 0\}.$$

It is easy to see that $\mathcal{M}$ is a closed subspace of the Hilbert space $L^2(\Omega)$, and hence $\mathcal{M}$ is also a Hilbert space. Let $\mathcal{U}$ be the unit ball of $\mathcal{M}$:

$$\mathcal{U} = \{f \in \mathcal{M} : ||f||^2 = E|f|^2 \le 1\}.$$

Then $\mathcal{U}$ defines a set $\mathcal{U}(\mathcal{F})$ of functions on $\mathcal{F}$:

$$\mathcal{U}(\mathcal{F}) = \{f \to E(f \cdot g) : f \in \mathcal{F}, g \in \mathcal{U}\}.$$

We are now ready to state our main result of this paper.

THEOREM 1. (An Empirical LIL for Stationary Martingale Differences) *Assume that*

(a) $$\int_0^1 [\ln \nu^B(u, \mathcal{F}, d)]^{\frac{1}{2}} du < \infty$$

and

(b) *there exists a constant $D > 0$ such that*

$$P^* \left\{ \sup_{(f,g) \in \tilde{\mathcal{F}}} \sum_{i=1}^n \frac{E_{i-1}[f(V_i) - g(V_i)]^2}{n d^2(f,g)} \ge D \right\} \to 0.$$

Then

$$\left\{ \frac{S_n(f)}{\sqrt{2 \log \log n}} : f \in \mathcal{F}, n \ge 3 \right\}$$

is relatively compact with respect to $||\cdot||_{\mathcal{F}}$ a.s., and the set of its limit points is $\mathcal{U}(\mathcal{F})$.

REMARK. Notice that the integrability condition of the metric entropy with bracketing of (a) implies the total boundedness of the metric space $(\mathcal{F}, d)$.

In order to see that our result generalizes that of iid problem, let ξ be a random variable on a probability space $(S, \mathcal{B}, P_0)$, and let $\{\xi_i, i \geq 1\}$ be a sequence of independent copies of ξ. Let $\mathcal{F}$ be a class of real valued functions defined on S. Suppose that

$$E_0 f(\xi) = 0 \text{ for all } f \in \mathcal{F}$$

and

$$E_0 f^2(\xi) < \infty \text{ for all } f \in \mathcal{F}.$$

The following restatement of Theorem 4.2 of Ossiander (1987) will be a special case of Theorem 1.

THEOREM 2. (An ELIL for IID Random Variables)
If $\int_0^1 [\ln \nu^B(u, \mathcal{F}, L^2(P_0))]^{\frac{1}{2}} du < \infty$, *then*

$$\left\{ \frac{\sum_{i=1}^n f(\xi_i)}{\sqrt{2n \log\log n}} : f \in \mathcal{F}, n \geq 3 \right\}$$

is relatively compact with respect to $||\cdot||_{\mathcal{F}}$ *a.s., and the set of its limit points is* $\mathcal{U}(\mathcal{F})$ *where*

$$\mathcal{U}(\mathcal{F}) = \{f \to E_0 f(\xi) g(\xi) : f \in \mathcal{F}, g \in \mathcal{U}\}$$

with

$$\mathcal{U} = \{g \in L^2(S, \mathcal{B}, P_0) : E_0 g^2(\xi) \leq 1\}$$

Proof of Theorem 2. Consider $P = (P_0)^Z$ so that (X_i) are iid. In this case we see that

$$E_{i-1}[f(\xi_i) - g(\xi_i)]^2 = E_0[f(\xi) - g(\xi)]^2 = d^2(f, g).$$

So the condition (b) of the Theorem 1 holds with $D = 2$. This implies that Theorem 2 is a special case of Theorem 1. $\square$

3. An Empirical CLT and Its Invariance Principle for Stationary Martingale Differences

In this section we consider an invariance principle of an Empirical CLT for stationary martingale differences. Define

$$(3) \qquad S_n(f,t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} f(V_i), \quad \text{for } f \in \mathcal{F} \text{ and for } t \in [0,1],$$

where $[x]$ denotes the integer part of x. We write $\mathcal{S} = \mathcal{F} \otimes [0,1]$. Define a pseudo-metric ρ on $\mathcal{S}$ by

$$\rho((f,t),(g,s)) = \max\{d(f,g), |t-s|\}.$$

It is well known that $B(\mathcal{S})$ is complete in the sup-norm, so that $(B(\mathcal{S}), \|\cdot\|_{\mathcal{S}})$ forms a Banach space. We use the following definition of weak convergence due to Hoffmann-Jϕrgensen (see Hoffmann-Jϕrgensen, J. (1991)).

DEFINITION 2. A sequence of $B(\mathcal{S})$-valued random functions $\{Y_n : n \geq 1\}$ converges in law to a $B(\mathcal{S})$-valued Borel measurable random function Y, denoted $Y_n \Rightarrow Y$, if

$$Eg(Y) = \lim_{n \to \infty} E^* g(Y_n), \forall g \in C(B(\mathcal{S}), \|\cdot\|_{\mathcal{S}}),$$

where $C(B(\mathcal{S}), \|\cdot\|_{\mathcal{S}})$ is the set of all bounded, continuous functions from $(B(\mathcal{S}), \|\cdot\|_{\mathcal{S}})$ into R. Here E^* denotes upper expectation.

The following Proposition 1 whose proof depends on a chaining argument with stratification appears in Bae and Levental (1995).

PROPOSITION 1. *Assume that* (a) $\int_0^1 [\ln \nu^B(u, \mathcal{F}, d)]^{\frac{1}{2}} du < \infty$ *and* (b) *there exists a constant* $D > 0$ *such that*

$$P^* \left\{ \sup_{(f,g) \in \tilde{\mathcal{F}}} \sum_{i=1}^{n} \frac{E_{i-1}[f(X_i) - g(X_i)]^2}{n d^2(f,g)} \geq D \right\} \to 0.$$

Then there exists a Gaussian process $\{Z(f,t) : (f,t) \in \mathcal{F} \otimes [0,1]\}$ *with bounded and continuous sample paths such that*

$$S_n \Rightarrow Z,$$

as random elements of $B(\mathcal{F} \otimes [0,1])$. The Gaussian process has the mean $EZ(f,t) = 0$ and the covariance structure $EZ(f,t)Z(g,s) = (t \wedge s)E(f \cdot g)$.

4. Proof of Theorem 1

We begin with

DEFINITION 3. A sequence of $B(\mathcal{S})$-valued random functions $\{Y_n : n \geq 1\}$ converges in probability to 0, denoted $Y_n \to^P 0$, if

$$\lim_{n\to\infty} P^*\{|Y_n| > \epsilon\} = 0, \text{ for every } \epsilon > 0.$$

We will use the following restatement of Proposition 1 in the proof of Theorem 1. See Theorem 1.3 of Dudley and Philipp (1983). See also Theorem 4.1 of Ossiander (1987).

PROPOSITION 2. *Assume that*

(a) $\int_0^1 [\ln \nu^B(u,\mathcal{F},d)]^{\frac{1}{2}} du < \infty$

and

(b) *there exists a constant $D > 0$ such that*

$$P^*\left\{\sup_{(f,g)\in\tilde{\mathcal{F}}} \sum_{i=1}^{n} \frac{E_{i-1}[f(V_i) - g(V_i)]^2}{nd^2(f,g)} \geq D\right\} \to 0.$$

Then there exists a sequence of i.i.d. copies of a Gaussian process $\{Z(f) : f \in \mathcal{F}\}$, defined on $(\Omega, \mathcal{T}, P)$, with bounded and continuous sample paths on $\mathcal{F}$ with the mean $EZ(f) = 0$ and the covariance structure $EZ(f)Z(g) = E(f \cdot g)$ such that

$$\frac{1}{\sqrt{n}} \max_{1\leq k\leq n} \sup_{f\in\mathcal{F}} |\sum_{i=1}^{k}(f(V_i) - Z_i(f))| \to^P 0 \quad , \text{as } n \to \infty.$$

The Z_i's can also be chosen such that, with probability 1 for some measurable U_n

$$\sup_{f\in\mathcal{F}} |\frac{1}{\sqrt{n}} \sum_{i=1}^{n}(f(V_i) - Z_i(f))| \leq U_n = o(\sqrt{\log\log n}).$$

The following corollary is easily follows from the above proposition.

COROLLARY 1. *Assume that*

(a) $\int_0^1 [\ln \nu^B(u, \mathcal{F}, d)]^{\frac{1}{2}} du < \infty$

and

(b) *there exists a constant $D > 0$ such that*

$$P^* \left\{ \sup_{(f,g)\in \tilde{\mathcal{F}}} \sum_{i=1}^{n} \frac{E_{i-1}[f(V_i) - g(V_i)]^2}{nd^2(f, g)} \geq D \right\} \to 0.$$

Then there exist a sequence $\{\tilde{Z}_n : n \geq 1\}$, with bounded and continuous sample paths, of copies of a Gaussian process $\{Z(f) : f \in \mathcal{F}\}$ defined on $(\Omega, \mathcal{T}, P)$ such that

$$||S_n - \tilde{Z}_n|| \to^P 0, \text{ as } n \to \infty.$$

The Z_i's can also be chosen such that, with probability 1 for some measurable U_n

$$||S_n - \tilde{Z}_n|| \leq U_n = o(\sqrt{\log\log n}).$$

Proof. Let $\{Z_i\}$ be as in Proposition 2. Set $\tilde{Z}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i$. Observe that

$$||S_n - \tilde{Z}_n|| \leq \sup_{f\in\mathcal{F}} |\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f(V_i) - Z_i(f))| + \sup_{f\in\mathcal{F}} |\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(f) - \tilde{Z}_n(f)|.$$

Since the Gaussian processes Z and $\tilde{Z}$ have the same mean and the same covariance structure, they have the same distribution. That is,

$$P\{||S_n - \tilde{Z}_n|| = 0\} = 1.$$

Proposition 1 implies the results.

PROPOSITION 3. (Theorem 4.3 of Pisier(1975))

Suppose that $\int_0^1 [\ln \nu^B(u, \mathcal{F}, d)]^{\frac{1}{2}} du < \infty$. *Let* $\{Z_i : i \geq 1\}$ *be a sequence of i.i.d. copies of a Gaussian process* $\{Z(f) : f \in \mathcal{F}\}$ *defined on* $(\Omega, \mathcal{T}, P)$. *Suppose* $\{Z(f) : f \in \mathcal{F}\}$ *has bounded and continuous sample paths with* $EZ(f) = 0$ *and* $E||Z||^2 < \infty$. *Then* Z *satisfies the empirical law of the iterated logarithm. That is,*

$$\left\{ \frac{\sum_{i=1}^n Z_i(f)}{\sqrt{2n \log \log n}} : f \in \mathcal{F}, n \geq 3 \right\}$$

is relatively compact with respect to $|| \cdot ||$ a.s., *and the set of its limit points is*

$$\mathcal{U}(\mathcal{F}) = \{f \to EZ(f)Z(g) : f \in \mathcal{F}, g \in \mathcal{U}\}$$

where $\mathcal{U} = \{g \in L^2(\Omega, \mathcal{T}, P) : EZ^2(g) \leq 1\}$.

REMARK:. Z takes values in $C(\mathcal{F})$, the bounded and continuous functions from $\mathcal{F}$ to R, forms a separable Banach space with the sup-norm $|| \cdot ||_{\mathcal{F}}$.

We are now ready to prove the main result of the paper.

Proof of Theorem 1:. By Proposition 1, there exists a Gaussian process $\{Z(f) : f \in \mathcal{F}\}$ with bounded and continuous sample paths whose mean is zero and covariance structure is

$$EZ(f)Z(g) = E(f \cdot g). \tag{4}$$

Apply Proposition 2 to choose a sequence $\{Z_i : i \geq 1\}$ of i.i.d. copies of $\{Z(f) : f \in \mathcal{F}\}$ such that

$$|| \frac{1}{\sqrt{2 \log \log n}} (S_n - \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i) || \leq Y_n = o(1), \tag{5}$$

with probability 1 for some sequence of measurable $Y_n's$. By Proposition 3, Z satisfies the empirical law of the logarithm. That is,

$$\left\{ \frac{\sum_{i=1}^n Z_i(f)}{\sqrt{2n \log \log n}} : f \in \mathcal{F}, n \geq 3 \right\}$$

is relatively compact with respect to $||\cdot||$ a.s., and the set of its limit points is

$$\mathcal{U}(\mathcal{F}) = \{f \to EZ(f)Z(g) : f \in \mathcal{F}, g \in \mathcal{U}\}$$

where $\mathcal{U} = \{g \in L^2(\Omega, \mathcal{T}, P) : EZ^2(g) \leq 1\}$. This, together with Eq.(4) and Eq.(5), completes theproof of Theorem 1. $\square$

5. An Empirical LIL for Stationary Markov Chains

Let $X_0, X_1, X_2, \cdots$ be a strictly stationary and ergodic Markov chain taking values in a measurable space $(S, \mathcal{B})$ with transition mechanism $P(x, dy)$ and initial distribution α. We assume there exists $b \geq 1$ so that

$$P(x, dy) = p(x, y)\alpha(dy), x, y \in S \tag{6}$$

and

$$0 \leq p(x, y) \leq b < \infty, x, y \in S. \tag{7}$$

Let $P : L^2(\alpha) \to L^2(\alpha)$ be defined by

$$P(g)(x) = \int_S g(y)P(x, dy),\ g \in L^2(\alpha). \tag{8}$$

Denote $||f||$ to be the $L^2(\alpha)$-norm of f. Then from Jensen inequality we note that

$$\sup_{x \in S} |Pg(x)| \leq b||g||,\ g \in L^2(\alpha). \tag{9}$$

For every $f \in \mathcal{F} \subseteq L^2(\alpha)$, we define

$$S_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{f(X_i) - E_\alpha f(X_0)\}. \tag{10}$$

REMARK:. In order to derive an empirical law of iterated logarithm for stationary Markov chains we will use the basic idea in Bae and Levental (1995) where the problem is translated into that of a stationary martingale differences as was originally done by Gordin and Lifsic (1978). We give the proofs of arguments as needed for the completeness of the proof of Theorem 3.

Let $\mathcal{M}_\alpha = \{f \in L^2(\alpha) : E_\alpha(f) = 0\}$. Notice that $\mathcal{M}_\alpha$ is a closed subspace of the Hilbert space $L^2(\alpha)$ and $(I-P)\mathcal{M}_\alpha \subseteq \mathcal{M}_\alpha$.

LEMMA 1. *Assume* (6) *and* (7). *Then*
(a) $I-P$ *is a one to one and onto operator on* $\mathcal{M}_\alpha$
(b) $(I-P)^{-1} = \sum_{i=0}^{\infty} P^i$ *on* $\mathcal{M}_\alpha$, *where the convergence is in operator norm.*

Proof of Lemma 1:. (a) From the ergodicity it follows that $I-P$ is a one to one operator on $\mathcal{M}_\alpha$, in other words, 1 is not an eigenvalue of P. Note that P is a compact operator on $\mathcal{M}_\alpha$ (see Yosida, K. (1965), Example 2, p.277). This implies that every nonzero element in the spectrum of P is an eigenvalue (see Yosida, K. (1965), Proposition 1, p.284). So 1 belongs to the resolvent set of P and $I-P$ is also onto. (b) $\{X_i\}$ is a φ-mixing Markov chain: namely, $\varphi(n)$ converges to zero exponentially fast, where $\varphi(n) = \sup_{A\in\mathcal{B},B\in\mathcal{B},\alpha(A)>0} |P(X_n \in B|X_0 \in A) - \alpha(B)|$ (see Ibragimov, I. A. and Linnik, J. V. (1971), p.367-368). Since $||P^n|| \le 2\sqrt{\varphi(n)}$ (Ibragimov. I. A. and Linnik, J. V. (1971), Theorem 17.2.3), it follows that $\limsup_n (||P^n||)^{1/n} < 1$. Therefore we have the representation $(I-P)^{-1} = \sum_{i=0}^{\infty} P^i$ on $\mathcal{M}_\alpha$.

From stationarity of the Markov chain, using the Kolmogorov consistency theorem, we may assume that the process (X_i) is double sided. Also each $f : S \to R$ will be considered as $f : S^Z \to R$ by putting $f((X_i)) = f(X_0)$.

Let $\mathcal{G} = \{(I-P)^{-1}[f - E_\alpha f] : f \in \mathcal{F}\}$. Observe that $f - E_\alpha f \in \mathcal{M}_\alpha$. We assume without loss of generality that $E_\alpha f = 0$, $f \in \mathcal{F}$. Then $\mathcal{G}$ can be rewritten as $\{(I-P)^{-1} f : f \in \mathcal{F}\}$. For every $g \in \mathcal{G}$ we define $\tilde{g} : S^Z \to R$ by $\tilde{g}(X) = g(X_0) - Pg(X_{-1})$. Observe that $Pg(X_{-1}) = E\{g(X_0)|X_{-1}\}$. This implies that $\{\tilde{g}(T^i X))\}$ is a martingale difference for each $g \in \mathcal{G}$.

We are ready to state an empirical law of the iterated logarithm for stationary Markov chains.

THEOREM 3. *Let $\mathcal{F}$ be a class of real valued functions defined on S. Assume that (6) and (7) hold. If $\int_0^1 [\ln \nu^B(u, \mathcal{F}, L^2(\alpha))]^{1/2} du < \infty$, then*

$$\left\{ \frac{S_n(f)}{\sqrt{2 \log \log n}} : f \in \mathcal{F}, n \geq 3 \right\}$$

is relatively compact with respect to $\|\cdot\|_{\mathcal{F}}$ a.s., and the set of its limit points is

$$\mathcal{L} = \{g(X_0) - Pg(X_{-1}) \to E_\alpha(g(X_0) - Pg(X_{-1}))(h(X_0) - Ph(X_{-1})) \\ : g = (I-P)^{-1}(f - E_\alpha f), f \in \mathcal{F}, h \in \mathcal{U}\}$$

where

$$\mathcal{U} = \{h \in \mathcal{M}_\alpha : E_\alpha(h(X_0) - Ph(X_{-1}))^2 \leq 1, \\ h = (I-P)^{-1}(f - E_\alpha f), f \in \mathcal{F}\}.$$

LEMMA 2. *If $\int_0^1 [\ln \nu^B(u, \mathcal{F}, L^2(\alpha))]^{1/2} du < \infty$, then $\int_0^1 [\ln \nu^B(u, \{\tilde{g}\}, d)]^{\frac{1}{2}} du < \infty$.*

Proof of Lemma 2:. Let $u > 0$. Write $\nu^B(u) = \nu^B(u, \mathcal{F}, L^2(\alpha))$. By definition of $\nu^B(u)$, there is $\{f^l_{0,u}, f^u_{0,u}, \cdots, f^l_{\nu^B(u),u}, f^u_{\nu^B(u),u}\}$ so that for every $f \in \mathcal{F}$ there exists $0 \leq i \leq \nu^B(u)$ satisfying $f^l_{i,u} \leq f \leq f^u_{i,u}$ and $\|f^u_{i,u} - f^l_{i,u}\| < u$. Let $M = \sup_{f \in \mathcal{F}} \|f\|$. Since $\|P^N\| \to 0$ as $N \to \infty$, we can choose N so that $\|P^N\| \leq \frac{u}{2bM\|(I-P)^{-1}\|}$. From Eq.(9) we see that

$$\sup_{x \in S} \left| \sum_{n=N+1}^{\infty} P^n f(x) \right| = \sup_{x \in S} \left| P \sum_{n=N}^{\infty} P^n f(x) \right| \leq \frac{u}{2}. \tag{11}$$

We also see that

$$g = \sum_{n=0}^{N} P^n f + \sum_{n=N+1}^{\infty} P^n f \leq \sum_{n=0}^{N} P^n f^u + \sum_{n=N+1}^{\infty} P^n f, \tag{12}$$

Define

$$g^l_{j,u} := \sum_{n=0}^{N} P^n f^l_{j,u} - \frac{u}{2} \text{ and } g^u_{j,u} := \sum_{n=0}^{N} P^n f^u_{j,u} + \frac{u}{2},$$

for $j = 0, \cdots, \nu^B(u)$. Then since $g = (I-P)^{-1} f$ and $f^l_{i,u} \le f \le f^u_{i,u}$, using Eq.(11) and Eq.(12), we get $g^l_{i,u} \le g \le g^u_{i,u}$. We also have

$$d(g^l_{i,u}, g^u_{i,u}) = \|g^u_{i,u} - g^l_{i,u}\| \le u + \|\sum_{n=0}^{N} P^n(f^u_{i,u} - f^l_{iu})\| \le C \cdot u,$$

where $C = 1 + \sup_N \|\sum_{n=0}^{N} P^n\|$.

We now define the brackets $\{\tilde{g}^l_{0,u}, \tilde{g}^u_{0,u}, \cdots, \tilde{g}^l_{\nu^B(u),u}, \tilde{g}^u_{\nu^B(u),u}\}$ for $\tilde{g}$ by the following equations

$$\tilde{g}^l_{j,u} := g^l_{j,u}(X_0) - P g^u_{j,u}(X_{-1}), \quad \text{and} \quad \tilde{g}^u_{j,u} := g^u_{j,u}(X_0) - P g^l_{j,u}(X_{-1}),$$

for $j = 0, \cdots, \nu^B(u)$. From the inequalities

$$g^l_{i,u}(X_0) \le g(X_0) \le g^u_{i,u}(X_0), \quad \text{and}$$
$$P g^l_{i,u}(X_{-1}) \le P g(X_{-1}) \le P g^u_{i,u}(X_{-1})$$

we have

$$\tilde{g}^l_{i,u} \le \tilde{g} \le \tilde{g}^u_{i,u}$$

and

$$d(\tilde{g}^l_{i,u}, \tilde{g}^u_{i,u}) \le \|g^u_{i,u} - g^l_{i,u}\| + \|P(g^u_{i,u} - g^l_{i,u})\| \le 2\|(g^u_{i,v} - g^l_{i,u})\| \le 2C \cdot u.$$

We conclude that

$$\begin{aligned}\int_0^1 [\ln \nu^B(u, \{\tilde{g}\}, d)]^{\frac{1}{2}} du \le & \int_0^1 [\ln \nu^B(\frac{u}{2C}, \mathcal{F}, L^2(\alpha))]^{\frac{1}{2}} du \\ & = 2C \int_0^{\frac{1}{2C}} [\ln \nu^B(u, \mathcal{F}, L^2(\alpha))]^{\frac{1}{2}} du \\ & < \infty.\end{aligned}$$

The proof of Lemma 2 is completed. □

LEMMA 3.

$$\sup_{f\in\mathcal{F}} \Big|\sum_{i=0}^{n-1}\{f(X_i)-\tilde{g}[T^i(X)]\}\Big| = o_P(\sqrt{n\log\log n}). \tag{13}$$

Proof of Lemma 3:. Since $(I-P)g=f$ we see that

$$\sup_{f\in\mathcal{F}} \Big|\sum_{i=0}^{n-1}\{f(X_i)-\tilde{g}[T^i(X)]\}\Big| = \sup_{g\in\mathcal{G}}|g(X_0)-g(X_n)| \leq 2G(X_n),$$

where $G(\cdot)=\sup_{g\in\mathcal{G}}|g(\cdot)|$. Notice that $G\in L^2(\Omega)$ as follows from the proof of Lemma 2. Now Eq.(13) follows from Markov inequality.

LEMMA 4. *There exists a constant $D>0$ such that*

$$P^*\left\{\sup_{(h_1,h_2)\in\{\tilde{g}\}}\sum_{i=1}^{n}\frac{E_{i-1}[h_1(T^i(X))-h_2(T^i(X))]^2}{nd^2(h_1,h_2)}\geq D\right\}=0.$$

Proof of Lemma 4:. Let us first assume that $||P||<1$. We choose $D=\frac{4b}{1-||P||^2}+1>0$. Note that

$$P(g^2)(x)=\int_S g^2(y)p(x,y)\alpha(da)\leq b\int_S g^2(y)\alpha(dy)=b||g||^2, \tag{14}$$

and

$$||g||^2-||Pg||^2\geq ||g||^2(1-||P||^2). \tag{15}$$

So for $(\tilde{g_1},\tilde{g_2})\in\{\tilde{g}\}\times\{\tilde{g}\}$

$$\begin{aligned}
&E_{-1}[(\tilde{g_1}-\tilde{g_2})(X)]^2\\
&=P(g_1-g_2)^2(X_0)-[P(g_1-g_2)(X_{-1})]^2\\
&\leq P(g_1-g_2)^2(X_0)\\
&\leq b||g_1-g_2||^2\\
&\leq \frac{b}{1-||P||^2}(||g_1-g_2||^2-||P(g_1-g_2)||^2)\\
&=\frac{b}{1-||P||^2}\cdot d^2(\tilde{g_1},\tilde{g_2}).
\end{aligned}$$

Similarly for $(\tilde{g}^l, \tilde{g}^u) \in \bigcup_{0<\delta\le 1}\{\tilde{g}\}_\delta$

$$\begin{aligned}
&E_{-1}[(\tilde{g}^u - \tilde{g}^l)(X)]^2 \\
&\le 2E_{-1}[(g^u - g^l)^2(X_0) + \{P(g^u - g^l)(X_{-1})\}^2] \\
&= 4P(g^u - g^l)^2(X_{-1}) \\
&\le 4b\|g^u - g^l\|^2 \\
&= 4bd^2(\tilde{g}^l, \tilde{g}^u).
\end{aligned}$$

Therefore we have

$$P^*\left\{\sup_{(h_1,h_2)\in\{\bar{g}\}} \sum_{i=1}^n \frac{E_{i-1}[h_1((X_i)) - h_2(X_i)]^2}{nd^2(h_1,h_2)} \ge D\right\} = 0.$$

If $\|P\| = 1$ then there exists N so that $\|P^N\| < 1$. We will work with N Markov chains $(X_{k+Nl})_{l=0}^\infty$, $k = 0, 1, \cdots, N-1$. For each Markov chain the appropriate Markov operator is P^N. Choose D_k for each $k = 0, ..., N-1$ so that

$$P^*\left\{\sup_{(h_1,h_2)\in\{\bar{g}\}} \sum_{i=1}^n \frac{E_{i-1}[h_1(X_{k+Ni}) - h_2(X_{k+Ni})]^2}{nd^2(h_1,h_2)} \ge D_k\right\} = 0.$$

Write $D = \max\{D_0, ..., D_{N-1}\}$. Then

$$\begin{aligned}
&P^*\left\{\sup_{(h_1,h_2)\in\{\bar{g}\}} \sum_{i=1}^n \frac{E_{i-1}[h_1(T^i(X)) - h_2(T^i(X))]^2}{nd^2(h_1,h_2)} \ge D\right\} \\
&\le P^*\left\{\sup_{(h_1,h_2)\in\{\bar{g}\}} \sum_{k=0}^{N-1}\sum_{i=1}^n \frac{E_{i-1}[h_1(X_{k+Ni}) - h_2(X_{k+Ni})]^2}{nd^2(h_1,h_2)} \ge D_k\right\} \\
&\le \sum_{k=0}^{N-1} P^*\left\{\sup_{(h_1,h_2)\in\{\bar{g}\}} \sum_{i=1}^n \frac{E_{i-1}[h_1(X_{k+Ni}) - h_2(X_{k+Ni})]^2}{nd^2(h_1,h_2)} \ge D_k\right\} \\
&= 0.
\end{aligned}$$

The proof of Lemma 4 is completed.

Proof of Theorem 3:. This follows from Lemma 2, Lemma 3, and Lemma 4 apply to $\{\tilde{g}\}$ by noticing that the set of the limit points in this case is

$$\mathcal{L} = \{g(X_0) - Pg(X_{-1}) \to E_\alpha(g(X_0) - Pg(X_{-1}))(h(X_0) - Ph(X_{-1})) \\ : g = (I-P)^{-1}(f - E_\alpha f), f \in \mathcal{F}, h \in \mathcal{U}\}$$

where

$$\mathcal{U} = \{h \in \mathcal{M}_\alpha : E_\alpha(h(X_0) - Ph(X_{-1}))^2 \\ \leq 1, h = (I-P)^{-1}(f - E_\alpha f), f \in \mathcal{F}\}.$$

The proof Theorem 3 is completed.

6. An Empirical LIL for the Baker's Transformation

Let $\Omega = [0,1) \times [0,1)$ be the sample space, $\mathcal{T}$ be the Borel sets and P be the Lebesgue measure. Define the Baker's transformation $\phi : [0,1) \times [0,1) \to [0,1) \times [0,1)$ by

$$\phi(x,y) = \begin{cases} (2x, \frac{y}{2}) & \text{if } 0 \leq x < \frac{1}{2}, \\ (2x-1, \frac{y+1}{2}) & \text{if } \frac{1}{2} \leq x < 1. \end{cases}$$

We can think about $(\ldots, x_{-1}, x_0, x_1, \ldots) \in \{0,1\}^Z$ as a point (x,y) in the half open unit square $[0,1) \times [0,1)$ by putting $x = \sum_{i=0}^\infty \frac{x_i}{2^i+1}$ and $y = \sum_{i=1}^\infty \frac{x_{-i}}{2^i}$. It is known that the transformation is ergodic (See Durrett, R. (1991)). For $t \in [0,1]$, we define

$$f_t(x,y) = \begin{cases} 1_{[0,t]}(y) & \text{if } 0 \leq x < \frac{1}{2} \\ -1_{[0,t]}(y) & \text{if } \frac{1}{2} \leq x < 1. \end{cases}$$

Consider the class of functions $\mathcal{F} = \{f_t(x,y)\}_{0 \leq t \leq 1}$. We denote $\phi^i(x,y) = (x_i, y_i)$ for $(x,y) \in [0,1) \times [0,1)$, $i = 0, 1, \ldots$. Define

$$Z_n(t) = Z_n(f_t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n f_t(x_i, y_i), \; f_t \in \mathcal{F}. \tag{16}$$

Notice that there is a one to one correspondence between $\mathcal{F} = \{f_t(x,y) : 0 \leq t \leq 1\}$ and $[0,1]$. The metric d and the sup-norm $||\cdot||_{\mathcal{F}}$ in this case is given by $d^2(f_t, f_s) = |t-s|$ and $||\varphi||_{\mathcal{F}} = \sup_{0 \leq t \leq 1} |\varphi(t)|$ respectively.

The following theorem is an empirical law of the iterated logarithm for the Baker's transformation.

THEOREM 4.

$$\left\{\frac{\sum_{i=1}^n f_t(x_i,y_i)}{\sqrt{2n\log\log n}} : t\in[0,1], n\geq 3\right\}$$

is relatively compact with respect to the sup-norm a.s., and the set of its limit points is

$$\mathcal{U}([0,1]) = \left\{t\to\int_0^{\frac{1}{2}}\int_0^t g(x,y)dadx \right.$$
$$\left. -\int_{\frac{1}{2}}^1\int_0^t g(x,y)dydx : t\in[0,1], g\in\mathcal{U}\right\}$$

where

$$\mathcal{U} = \left\{g\in L^2([0,1)\times[0,1)) : \int_0^1\int_0^1 g^2(x,y)dadx \right.$$
$$\left. -(\int_0^1\int_0^1 g(x,y)dydx)^2\leq 1\right\}.$$

The following lemma appears in Pollard (1984).

LEMMA 5. (Bernstein's Inequality) *Let $Y_1,...,Y_n$ be independent random variables with zero means and bounded ranges : $|Y_i|\leq M$. Write σ_i^2 for the variance of Y_i. Suppose $V\geq\sigma_1^2+\cdots+\sigma_n^2$. Then for each $\eta>0$,*

$$P\left\{|\sum_{i=1}^n Y_i|>\eta\right\}\leq 2\exp\left\{-\frac{\frac{\eta^2}{2}}{V+\frac{M\eta}{3}}\right\}$$

We observe that $E(f_t(x,y)|y)=0$ for each $0\leq t\leq 1$, which means that $\{f_t(x_i,y_i)\}$ is a martingale differences for each t. In the proof of Theorem 5 we will apply Theorem 1 to the class of functions $\mathcal{F} = \{f_t(x,y) : 0\leq t\leq 1\}$.

LEMMA 6. *There exists a $D > 0$ such that*

$$P\left\{\sup_{t\in[0,1]}\sum_{i=1}^{n}\frac{1_{[0,t]}(y_i)}{nt} > D\right\} \to 0, \quad \text{as } n \to \infty.$$

Proof of Lemma 6:. It will be enough to show that

$$P\left\{\sup_{0\le k\le[\sqrt{n}]-1}\sum_{i=1}^{n}1_{[\frac{k}{\sqrt{n}},\frac{k+1}{\sqrt{n}}]}(y_i) > 8\sqrt{n}\right\} \to 0, \text{as } n \to \infty. \tag{17}$$

Claim:

$$\sum_{i=2km+1}^{(2k+1)m} 2^m P\left\{\sum_{i=1}^{2^{2m}}(1_{[0,\frac{1}{2^m}]}(y_i) - \frac{1}{2^m}) > 2^m\right\} \to 0, \quad \text{as } m \to \infty.$$

To prove the claim we observe that

$$Y_k = (1_{[0,\frac{1}{2^m}]}(y_i) - \frac{1}{2^m}),\ k = 0,\ 1,\ 2, \ldots$$

is an i.i.d. sequence since the Y_k depends on disjoint subsets of the sequence of $0's$ and $1's$. Note that $EY_k = 0$, and $Var(Y_k) \le \frac{m^2}{2^m}$ as follows from

$$\begin{aligned} Var(Y_k) &= Var\left(\sum_{i=2km+1}^{(2k+1)m} 1_{[0,\frac{1}{2^m}]}(y_i)\right) \\ &\le E\left(m\sum_{i=2km+1}^{(2k+1)m} 1_{[0,\frac{1}{2^m}]}(y_i)\right) = \frac{m^2}{2^m}. \end{aligned}$$

Apply Lemma 5 with $V = m^2 2^m$, an upper bound of $\sum_{k=0}^{[\frac{2^{2m}}{2m}]} Y_k$, $M = m$, and $\eta = 2^{m-1}$ to have

$$2^m P\left\{\sum_{k=0}^{[\frac{2^{2m}}{2m}]} Y_k > 2^{m-1}\right\} \to 0, \text{ as } m \to \infty, \tag{18}$$

as follows from the inequality

$$2^m P\left\{\sum_{k=0}^{[\frac{2^{2m}}{2m}]} Y_k > 2^{m-1}\right\} \leq 2^{m+1}\cdot \exp\left\{-\frac{2^m}{8(m^2+\frac{m}{6})}\right\}.$$

Similarly for the Z_k defined by

$$Z_k = \sum_{i=(2k+1)m+1}^{(2k+2)m} (1_{[0,\frac{1}{2^m}]}(y_i) - \frac{1}{2^m}),\ k=0,1,2,...,$$

we have

$$2^m P\left\{\sum_{k=0}^{[\frac{2^{2m}}{2m}]} Z_k > 2^{m-1}\right\} \to 0,\ \text{as } m\to\infty. \tag{19}$$

The claim follows from Eq.(18) and Eq.(19). Using the representation of $0's$ and $1's$ we see that the distribution of

$$\left\{1_{[\frac{k}{\sqrt{n}},\frac{k+1}{\sqrt{n}}]}(y_i):\ i=1,...,2^{2m}\right\}$$

does not depend on k, so that of

$$\sum_{i=1}^{2^{2m}} 1_{[\frac{k}{\sqrt{n}},\frac{k+1}{\sqrt{n}}]}(y_i)$$

does not depend on k either. Therefore from the claim we get

$$P\left\{\sup_{0\leq k\leq 2^m-1}\sum_{i=1}^{2^{2m}} 1_{[\frac{k}{\sqrt{n}},\frac{k+1}{\sqrt{n}}]}(y_i) > 2\cdot 2^m\right\}\to 0,\ \text{as}\ \ m\to 0.$$

We have proved Eq.(17) for $n = 2^{2m}$, $m = 1, 2, \ldots$. For $2^{2m} < n < 2^{2(m+1)}$ we have

$$\begin{aligned}
P&\left\{\sup_{0\le k\le[\sqrt{n}]-1}\sum_{i=1}^{n}1_{[\frac{k}{\sqrt{n}},\frac{k+1}{\sqrt{n}}]}(y_i) > 8\sqrt{n}\right\}\\
&\le \sum_{k=0}^{[\sqrt{n}]-1} P\left\{\sum_{i=1}^{n}1_{[\frac{k}{\sqrt{n}},\frac{k+1}{\sqrt{n}}]}(y_i) > 8\sqrt{n}\right\}\\
&= [\sqrt{n}]P\left\{\sum_{i=1}^{n}1_{[0,\frac{1}{\sqrt{n}}]}(y_i) > 8\cdot\sqrt{n}\right\}\\
&\le 2^{(m+1)}P\left\{\sum_{i=1}^{2^{2(m+1)}}1_{[0,\frac{1}{2^m}]}(y_i) > 8\cdot 2^m\right\}\\
&\le 4\cdot 2^{(m+1)}P\left\{\sum_{i=1}^{2^{2m}}1_{[0,\frac{1}{2^m}]}(y_i) > 2\cdot 2^m\right\}\\
&\to 0, \text{ as } m\to\infty,
\end{aligned}$$

as follows from the claim. The proof of Lemma (6) is completed.

Proof of Theorem 4:. We verify the conditions of Theorem 1. Since there is a one to one correspondence between $\mathcal{F} = \{f_t(x,y) : 0 \le t \le 1\}$ and $[0,1]$, the integrability condition of the covering number with bracketing is obvious. Recall that the metric d in this case is given by $d^2(f_t, f_s) = |t-s|$. Note that

$$E_{i-1}(f_t - f_s)^2(x_i, y_i) = 1_{[t,s]}(y_i) \text{ for } 0 \le t < s \le 1.$$

Condition (b) follows from Lemma 6. Finally observe that the set of its limit points $\mathcal{U}([0,1])$ in this case is given by

$$\left\{t \to \int_0^{\frac{1}{2}}\int_0^t g(x,y)dadx - \int_{\frac{1}{2}}^1\int_0^t g(x,y)dydx : t\in[0,1], g\in\mathcal{U}\right\}$$

where

$$\mathcal{U} = \left\{g \in L^2([0,1)\times[0,1)) : \int_0^1\int_0^1 g^2(x,y)dadx \right.$$
$$\left. -(\int_0^1\int_0^1 g(x,y)dydx)^2 \le 1\right\}.$$

This completes the proof of Theorem 4. □

References

1. Bae, J. and Levental, S., *Uniform CLT for Markov Chains and Its Invariance Principle: A Martingale Approach.*, J. Theoret. Probab. **8** (1995), 549-570.
2. Breiman, L., *Probability.*, Addison-Wesley, Reading, MA, 1968.
3. Brown, B. M., *Martingale Central Limit Theorems*, Ann. Math. Statist **42** (1971), 59-66.
4. Dudley, R. M., *A Coruse on Empirical Processes*, Lecture notes in Math. 1097 Springer-Verlag, New York, 1984.
5. Dudley, R. M. and Philipp, W., *Invariance principles for sums of Banach space valued random elements and empirical processes*, Z. Wahrsch. verw. Gebiete **62** (1983), 509-552.
6. Durrett, R., *Probability: Theory and Examples*, Wadsworth, Belmont, CA, 1991.
7. Hoffmann-Jϕrgensen, J., *Stochastic processes on Polish spaces*, Aarhus Universitet. Matematisk Institut, 1991.
8. Ibragimov, I. A. and Linnik, J. V., *Independent and Stationary Sequence of Random Variables*, "Nauka". Moscow, 1965; English transl. Noordhoff. Groningen, 1971.
9. Kuelbs, J. and Dudley, R. M., *Log Log law for Empirical Measures*, Ann. Probab. **8** (1980), 405-418.
10. Levental, S., *A Uniform CLT for Uniformly Bounded Families of Martingale Differences*, J. Theoret. Probab **2** (1989), 271-287.
11. Ossiander, M., *A Central Limit Theorem under Metric Entropy with L_2 Bracketing*, Ann. Probab **15** (1987), 897-919
12. Pisier, G., *Le theoreme de la limite centrale et laloi du logarithme itere dans les espaces de Banach*, Seminaire Maurey-Schwarz 1975-1976 exposes Nos. 3 et 4, 1975/1976.
13. Pollard, D., *Convergence of Stochastic Processes. Springer series in Statistics*, Springer-Verlag, New York, 1984.
14. Yosida, K., *Functional Analysis Springer-Verlag*, New York, 1965.

Department of Mathematics
Sung Kyun Kwan University
Suwon 440-740, Korea