Bull. Korean Math. Soc. **51** (2014), No. 2, pp. 317–328
http://dx.doi.org/10.4134/BKMS.2014.51.2.317

THE SECOND CENTRAL LIMIT THEOREM FOR MARTINGALE DIFFERENCE ARRAYS

Jongsig Bae, Doobae Jun, and Shlomo Levental

Abstract. In Bae *et al.* [2], we have considered the uniform CLT for the martingale difference arrays under the uniformly integrable entropy. In this paper, we prove the same problem under the bracketing entropy condition. The proofs are based on Freedman inequality combined with a chaining argument that utilizes majorizing measures. The results of present paper generalize those for a sequence of stationary martingale differences. The results also generalize independent problems.

1. Introduction and the main results

In Bae *et al.* [2], we have obtained the uniform CLT for the martingale difference arrays under the uniformly integrable entropy condition.

In view of the history of developing the theory of empirical process, it is natural to prove the same problem under the bracketing entropy condition. The proofs will be based on Freedman inequality combined with a chaining argument that utilizes majorizing measures. The results of present paper generalize those for a sequence of stationary martingale differences of Bae and Levental [1]. The results also generalize independent problems such as in Van der Vaart and Wellner [9].

Consider the setting of a martingale differences array of Bae *et al.* [2]. In measuring the size of the class $\mathcal{F}$ we are going to use the concepts of bracketing number. We introduce the following definition in Van der Vaart and Wellner [9]. See also Dudley [3].

Definition 1. Let ρ be a pseudo metric on $\mathcal{F}$. Given two functions l and u, the bracket $[l, u]$ is the set of all functions f with $l \leq f \leq u$. An ϵ-bracket is a bracket $[l, u]$ with $\rho(l, u) < \epsilon$. The bracketing number $N_{[\,]}(\epsilon, \mathcal{F}, \rho)$ is the minimum number of ϵ-brackets needed to cover $\mathcal{F}$.

Received August 12, 2011; Revised November 28, 2012.

2010 *Mathematics Subject Classification.* Primary 60F17; Secondary 60F05.

Key words and phrases. central limit theorem, martingale difference array, bracketing entropy, majorizing measure, eventual uniform equicontinuity.

Define

$$\rho_n^2(f,g) := \sum_{j=1}^{j(n)} E\left[(V_{nj}(f) - V_{nj}(g)\right]^2 \text{ for } f, g \in \mathcal{F}.$$

We use the following definition of bracketing entropy.

Definition 2. For $\epsilon > 0$, $j \leq j(n)$, $n \in \mathbb{N}$, we define the covering number with bracketing $N_{[\ ]}(\epsilon, \mathcal{F}, \rho_n)$ as the minimal number of sets N_ϵ^n in a partition $\mathcal{F} = \cup_{k=1}^{N_\epsilon} \mathcal{F}_{\epsilon k}^n$ of $\mathcal{F}$ into sets $\mathcal{F}_{\epsilon k}^n$ such that, for every partitioning set $\mathcal{F}_{\epsilon k}^n$

$$\left(\sum_{j=1}^{j(n)} E^* \sup_{f,g \in \mathcal{F}_{\epsilon k}^n} |V_{nj}(f) - V_{nj}(g)|^2\right)^{1/2} \leq \epsilon.$$

We are ready to state a result on an eventual uniform equicontinuity for martingale difference arrays.

Theorem 1. *Let* $\{V_{nj}(f) : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$ *be a martingale differences array of* $\mathcal{L}_2$*-process indexed by a class* $\mathcal{F}$ *of measurable functions on a measurable space* $(\mathbf{X}, \mathcal{X})$*. Suppose that*

$$P^*\left(\sup_{f,g\in\mathcal{F}} \frac{\sigma_n^2(f,g)}{\rho_n^2(f,g)} \geq L\right) \to 0 \text{ as } n \to \infty \text{ for a constant } L. \tag{1}$$

Suppose that

$$\sum_{j=1}^{j(n)} E^*||V_{nj}||_{\mathcal{F}}^2\{||V_{nj}||_{\mathcal{F}} > \eta\} \to 0 \text{ for every } \eta > 0, \tag{2}$$

$$\int_0^{\delta_n} \left[\log N_{[\ ]}(\epsilon, \mathcal{F}, \rho_n)\right]^{1/2} d\epsilon \to 0 \text{ for every } \delta_n \downarrow 0. \tag{3}$$

Then

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} E^*\left(\sup_{\rho_n(f,g)\leq\delta} |S_n(f) - S_n(g)|\right) = 0.$$

Remark 1. It is inevitable, to the best of our knowledge, to use a complicate chaining argument with stratifications in the proof of Theorem 1. We communicated with Professor Ossiander, who originally proved the bracketing CLT using the chaining argument with stratifications, see Ossiander [6], on the possibility of removing stratifications. See Theorem 7.2.1 of Dudley [4]. See also Theorem 5.12 of van de Geer [8] where no stratifications are employed for a uniformly bounded class of functions.

Theorem 2. *Let* $\{V_{nj}(f) : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$ *be a martingale differences array of* $\mathcal{L}_2$*-process indexed by a class* $\mathcal{F}$ *of measurable functions on a* *measurable space* $(\mathbf{X}, \mathcal{X})$. *Suppose that*

$$P^*\left(\sup_{f,g\in\mathcal{F}} \frac{\sigma_n^2(f,g)}{\rho_n^2(f,g)} \geq L\right) \to 0 \text{ as } n\to\infty \text{ for a constant } L. \tag{4}$$

Suppose that

$$\sum_{j=1}^{j(n)} E^*||V_{nj}||_{\mathcal{F}}^2\{||V_{nj}||_{\mathcal{F}} > \eta\} \to 0 \text{ for every } \eta > 0, \tag{5}$$

$$\int_0^{\delta_n} \left[\log N_{[\ ]}(\epsilon, \mathcal{F}, \rho_n)\right]^{1/2} d\epsilon \to 0 \text{ for every } \delta_n \downarrow 0. \tag{6}$$

Suppose also that, as $n\to\infty$, *for each* $f \in \mathcal{F}$

$$\sum_{j=1}^{j(n)} v_{n,j}(f) \to^P \sigma^2(f), \tag{7}$$

where each $\sigma^2(f)$ *is a positive constant. Suppose there exists a Gaussian process* Z *such that finite dimensional distributions of* S_n *converges to those of* Z. *Then*

$$S_n \Rightarrow Z \text{ as random elements of } B(\mathcal{F}).$$

The limiting process $Z = \{Z(f) : f \in \mathcal{F}\}$ *is mean zero Gaussian process with covariance function* $EZ(f)Z(g)$ *and the sample paths of* Z *belong to* $U_B(\mathcal{F}, \rho_n)$.

Proof of Theorem 2. Note that $(\mathcal{F},\rho_n)$ is a totally bounded pseudometric space. By Theorem 1 and Markov inequality for outer expectation, we have the eventual uniform equicontinuity of the process $(S_n(f) : f \in \mathcal{F})$:

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P^*\left(\sup_{\rho_n(f,g)\leq\delta} |S_n(f) - S_n(g)| > \epsilon\right) = 0.$$

This result, together with the assumptions of finite dimensional distribution convergence, see for example Pollard [7], complete the proof. □

We will use the following Proposition 1 in the proof of Theorem 1.

Proposition 1. *Let* $\{V_{nj}^0(f)) : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$ *be a martingale differences array of* $\mathcal{L}_2$*-process indexed by a class* $\mathcal{F}$ *of measurable functions* $(\mathbf{X}, \mathcal{X})$ *bounded by a sequence* η_n *with* $\eta_n \downarrow 0$. *Let* τ_n *be a finite stopping time relative to the* σ*-fields* $\{\mathcal{E}_{nj} : 0 \leq j \leq j(n), n \in \mathbb{N}\}$ *that satisfies almost surely* $\sigma_{\tau_n}^2(f,g) \leq L\rho_{\tau_n}^2(f,g)$ *for* $f, g \in \mathcal{F}$ *and for a constant* L. *Suppose that*

$$\int_0^{\delta_n} \left[\log N_{[\ ]}(\epsilon, \mathcal{F}, \rho_n)\right]^{1/2} d\epsilon \to 0 \quad \text{for every } \delta_n \downarrow 0.$$

Then,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} E^*\left(\sup_{\rho_n(f,g)\leq\delta}\left|\sum_{j=1}^{\tau_n}\left[V_{nj}^0(f) - V_{nj}^0(g)\right]\right|\right) = 0.$$

2. Proof of Theorem 1

We introduce the following truncation argument. For $a > 0$, let

$$\psi(a,x) = \begin{cases} a & \text{if } a < x, \\ x & \text{if } |x| \le a, \\ -a & \text{if } x < -a. \end{cases}$$

For $\delta > 0$, $j \le j(n)$, $n \in \mathbb{N}$, and $f \in \mathcal{F}$, let

$$V_{nj}^{\delta}(f) = \psi(\delta, V_{nj}(f)),$$

so that $V_{nj}^{\delta}(f)$ is a truncation of $V_{nj}(f)$ at the level δ. We simplify the notation by writing $\overline{V_{nj}^{\delta}(f)} := V_{nj}^{\delta}(f) - E_{n,j-1}V_{nj}^{\delta}(f)$, and define

$$S_n^{(\delta)}(f) = \sum_{j=1}^{j(n)} \overline{V_{nj}^{\delta}(f)} \text{ for } f \in \mathcal{F}.$$

Write for a stopping time τ_n,

$$S_{\tau_n}^{(\delta)}(f) = \sum_{j \le \tau_n} \overline{V_{nj}^{\delta}(f)} \text{ for } f \in \mathcal{F}.$$

Since $\{V_{nj}(f), \mathcal{E}_{nj}\}$ is a martingale difference, for any $\delta > 0$, using the identity

$$|E_{n,j-1}(V_{nj}(f)\{|V_{nj}(f)| > \delta\})| = |E_{n,j-1}(V_{nj}(f)\{|V_{nj}(f)| \le \delta\})|,$$

we have

$$\begin{aligned} ||S_n - S_n^{(\delta)}||_{\mathcal{F}} \le {} & \frac{1}{\delta}\sum_{j=1}^{j(n)} ||V_{nj}||_{\mathcal{F}}^2\{||V_{nj}||_{\mathcal{F}} > \delta\} \\ & + \frac{2}{\delta}\sum_{j=1}^{j(n)} E_{n,j-1}||V_{nj}||_{\mathcal{F}}^2\{||V_{nj}||_{\mathcal{F}} > \delta\}. \end{aligned}$$

Therefore, using (2), we obtain $E^*||S_n - S_n^{(\delta)}||_{\mathcal{F}} = o(1)$. For any $\delta > 0$, and $\eta > 0$

$$\begin{aligned} & \sup_{\rho_n(f,g) \le \eta} |S_n(f) - S_n(g)| \\ \le {} & \sup_{\rho_n(f,g) \le \eta} |S_n^{(\delta)}(f) - S_n^{(\delta)}(g)| + 2||S_n(f) - S_n^{(\delta)}(f)||_{\mathcal{F}}. \end{aligned}$$

Therefore, we have

$$E^*\left(\sup_{\rho_n(f,g) \le \eta} |S_n(f) - S_n(g)|\right) \le E^*\left(\sup_{\rho_n(f,g) \le \eta} |S_n^{(\delta)}(f) - S_n^{(\delta)}(g)|\right) + o(1).$$

Define a stopping time τ_n by, for $n \ge 1$

$$\tau_n := n \wedge \max\left\{k \ge 0 : \sup_{f,g \in \mathcal{F}} \frac{\sigma_k^2(f,g)}{\rho_k^2(f,g)} < L\right\}.$$

Since the random variables $\sigma_k^2(f,g)$ predictable, we see that τ_n is a stopping time. Observe that

$$P^*\left(\sup_{f,g\in\mathcal{F}} \frac{\sigma^2_{\tau_n}(f,g)}{\rho^2_{\tau_n}(f,g)} \geq L\right) = 0.$$

Since $P^*(\tau_n < n) \to 0$ as $n\to\infty$, it is enough to prove that for every $\delta > 0$

$$\lim_{\eta\downarrow 0}\limsup_{n\to\infty} E^*\left(\sup_{\rho_{\tau_n}(f,g)\leq\eta} |S^{(\delta)}_{\tau_n}(f) - S^{(\delta)}_{\tau_n}(g)|\right) = 0.$$

Let $\delta > 0$. Choose $\eta_n \downarrow 0$ such that $|\overline{V^\delta_{nj}(f)}| \leq \eta_n$, apply Proposition 1 with

$$V^0_{nj}(f) := \overline{V^\delta_{nj}(f)}$$

and we conclude that

$$\lim_{\eta\downarrow 0}\limsup_{n\to\infty} E^*\left(\sup_{\rho_{\tau_n}(f,g)\leq\eta} \left|S^{(\delta)}_{\tau_n}(f) - S^{(\delta)}_{\tau_n}(g)\right|\right) = 0.$$

The proof of Theorem 1 is completed.

3. Proof of Proposition 1

In the proof of Proposition 1, we will use the Freedman inequality and an argument of chaining that utilizes majorizing measures. For a random variable ξ, we use the notation $||\xi||_\infty$ to denote the essential supremum of $|\xi|$. We also use the notation $\preceq$ to mean the left hand side is bounded by a constant times the right hand side.

Lemma 1. *Let $(d_j)_{1\leq j\leq n}$ be a martingale difference with respect to an increasing σ-fields $(\mathcal{E}_j)_{0\leq j\leq n}$. Suppose that $||d_j||_\infty \leq M$ for a constant $M<\infty$, $j = 1,\ldots,n$. Let $\tau_n \leq n$ be a stopping time relative to $(\mathcal{E}_i)$ that satisfies $||\sum_{j=1}^{\tau_n} E(d_j^2|\mathcal{E}_{j-1})||_\infty \leq V$ for a constant V. Then, for every $y > 0$*

$$P\left(\left|\sum_{j=1}^{\tau_n} d_j\right| > y\right) \leq 2\cdot\exp\left[-\frac{1}{2}\frac{y^2}{V+My}\right].$$

Proof. See Proposition 2.1 in Freedman [5]. □

Lemma 2. *Let X be a random variable such that*

$$P(|X|>y) \leq 2\cdot\exp\left[-\frac{1}{2}\frac{y^2}{V+My}\right] \text{ for every } y>0.$$

Then,

$$E(|X|;A) \preceq \left(M\log\frac{1}{\mu} + \sqrt{V}\sqrt{\log\frac{1}{\mu}}\right)(\mu + P(A))$$

for every measurable set A and every constant $0<\mu<e^{-1}$.

Proof. See Lemma 2.11.17 in Van der Vaart and Wellner [9]. □

Lemma 3. *Under the assumptions of Proposition 1, there exists for every n a sequence of nested partitions $\mathcal{F} = \cup_k \mathcal{F}_{qk}^n$ and discrete subprobability measure μ_n on $\mathcal{F}$ such that, for every k and n,*

$$\lim_{q_0\to\infty} \limsup_{n\to\infty} \sup_f \sum_{q=q_0+1}^{\infty} 2^{-q} \sqrt{\log \frac{1}{\mu_n(\mathcal{F}_q^n f)}} = 0, \tag{8}$$

$$\sum_{j=1}^{\tau_n} E \sup_{f,g\in\mathcal{F}_{qk}^n} (V_{nj}^0(f) - V_{nj}^0(g))^2 \le 2^{-2q}. \tag{9}$$

Here $\mathcal{F}_q^n f$ is the set in the q-th partition to which f belongs.

Proof. For every $q \in \mathbb{N}$, let N_q^n be the number of sets in the q-th partition. Define $\mu_n := \sum_{q=1}^{\infty} 2^{-q}\mu_{n,q}$, where $\mu_{n,q}(\mathcal{F}_q^n f) = (N_q^n)^{-1}$ for every f and q. The assumptions of Proposition 1 justify the existence of the discrete subprobability measure and the displayed properties. The proof is completed. □

In the proof of Proposition 1, we will modify the chaining argument in Van der Vaart and Wellner [9] to a martingale difference context using the Freedman inequality.

Proof of Proposition 1. First note that $|V_{nj}^0| \le \eta_n$ for every j and n. Let $\mathcal{F} = \cup_k \mathcal{F}_{qk}^n$ be a sequence of nested partitions as in Lemma 3. We may assume without loss of generality that $\mu_n(\mathcal{F}_q^n f) \le 1/4$ for every q and f. Most of the argument is carried out for a fixed n and this index will be suppressed in the notation. Choose an element f_{qk} from each set $\mathcal{F}_{qk}$ and define

$$\begin{aligned}
\pi_q f &= f_{qk}, \\
(\Delta_q f)_{nj} &= \sup_{f,g\in\mathcal{F}_{qk}} |V_{nj}^0(f) - V_{nj}^0(g)|, \text{ if } f \in \mathcal{F}_{qk}, \\
a_q f &= 2^{-q} \Big/ \sqrt{\log \frac{1}{\mu(\mathcal{F}_{q+1} f)}}.
\end{aligned}$$

For $q > q_0$, define

$$\begin{aligned}
(A_{q-1}f)_{nj} &= \{(\Delta_{q_0} f)_{nj} \le a_{q_0} f, \ldots, (\Delta_{q-1} f)_{nj} \le a_{q-1} f\}, \\
(B_q f)_{nj} &= \{(\Delta_{q_0} f)_{nj} \le a_{q_0} f, \ldots, (\Delta_{q-1} f)_{nj} \le a_{q-1} f, (\Delta_q f)_{nj} > a_q f\}, \\
(B_{q_0} f)_{nj} &= \{(\Delta_{q_0} f)_{nj} > a_{q_0} f\}.
\end{aligned}$$

Now, decompose

$$\begin{aligned}
V_{nj}^0(f) - V_{nj}^0(\pi_{q_0} f) = {} & [V_{nj}^0(f) - V_{nj}^0(\pi_{q_0} f)](B_{q_0} f)_{nj} \\
& + \sum_{q=q_0+1}^{\infty} [V_{nj}^0(f) - V_{nj}^0(\pi_{q_0} f)](B_q f)_{nj} \\
& + \sum_{q=q_0+1}^{\infty} [V_{nj}^0(\pi_q f) - V_{nj}^0(\pi_{q-1} f)](A_{q-1} f)_{nj}.
\end{aligned}$$

Claim 1.

$$\lim_{q_0\to\infty}\limsup_{n\to\infty} E^*\left(\left\|\sum_{j=1}^{\tau_n}[V^0_{nj}(f)-V^0_{nj}(\pi^n_{q_0}f)]\right\|_{\mathcal{F}}\right)=0.$$

Proof. Since $(\Delta_q f)_{nj} \le 2\eta_n$, the first term in the decomposition is zero for every fixed q_0 and for large n. Notice that $(\Delta_q f)_{nj} \le (\Delta_{q-1} f)_{nj} \le a_{q-1}f$ on $(B_q(f))_{nj}$. Therefore

$$[V^0_{nj}(f)-V^0_{nj}(\pi_{q_0}f)](B_qf)_{nj} \le (\Delta_q f)_{nj}(B_q(f))_{nj} \le a_{q-1}f.$$

Consider the martingale differences

$$d_{n,j}(f) := (\Delta_q f)_{nj}(B_q(f))_{nj} - E_{n,j-1}(\Delta_q f)_{nj}(B_q(f))_{nj}.$$

Then $|d_{n,j}(f)| \le 2a_{q-1}f := M$. We observe that

$$\sum_{j=1}^{\tau_n} E_{n,j-1}[(\Delta_q f)_{nj}(B_q(f))_{nj}]\{(\Delta_q f)_{nj} > a_q f\} \le L\cdot\frac{2^{-2q}}{a_q f} \tag{10}$$

as follows from

$$\begin{aligned}
&P\left(\sum_{j=1}^{\tau_n} E_{n,j-1}[(\Delta_q f)_{nj}(B_q(f))_{nj}]\{(\Delta_q f)_{nj} > a_q f\} > L\cdot\frac{2^{-2q}}{a_q f}\right)\\
&= P\left(a_q f\sum_{j=1}^{\tau_n} E_{n,j-1}[(\Delta_q f)_{nj}(B_q(f))_{nj}]\{(\Delta_q f)_{nj} > a_q f\} > L\cdot 2^{-2q}\right)\\
&\le P\left(\sum_{j=1}^{\tau_n} E_{n,j-1}\left[(\Delta_q f)_{nj}\right]^2 > L\cdot 2^{-2q}\right)\\
&\le P\left(\sum_{j=1}^{\tau_n} E_{n,j-1}\left[\sup_{f,g\in\mathcal{F}^n_{qk}}|V^0_{nj}(f)-V^0_{nj}(g)|\right]^2 > L\cdot 2^{-2q}\right)\\
&\le \sum_{k=1}^{N^n_q} P\left(\sum_{j=1}^{\tau_n} E_{n,j-1}|V^0_{nj}(f)-V^0_{nj}(g)|^2 > L\cdot 2^{-2q}\right)\\
&\le N^n_q\cdot P\left(\rho^2_{\tau_n}(f,g) > 2^{-2q}\right) = 0.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{j=1}^{\tau_n} E_{n,j-1}[d_{nj}(f)]^2\\
&\le 4\sum_{j=1}^{\tau_n} E_{n,j-1}[(\Delta_q f)_{nj}(B_q(f))_{nj}]^2
\end{aligned}$$

$$
\begin{aligned}
&\le 4\sum_{j=1}^{\tau_n} a_{q-1}fE_{n,j-1}[(\Delta_q f)_{nj}(B_q(f))_{nj}]\{(\Delta_q f)_{nj} > a_q f\}\\
&\le 4\frac{a_{q-1}f}{a_q f}\sum_{j=1}^{\tau_n} E_{n,j-1}[(\Delta_q f)_{nj}(B_q(f))_{nj}]^2\{(\Delta_q f)_{nj} > a_q f\}\\
&\le 4L\cdot\frac{a_{q-1}f}{a_q f}2^{-2q} \text{ by (10)}\\
&:= V.
\end{aligned}
$$

By Freedman inequality of Lemma 1, for every $y > 0$

$$
P\left(\left|\sum_{j=1}^{\tau_n} d_{n,j}(f)\right| > y\right) \le 2\cdot\exp\left[-\frac{1}{2}\frac{y^2}{V+My}\right].
$$

By Lemma 2 for every f and measurable set A,

$$
\begin{aligned}
&E(|\sum_{j=1}^{\tau_n} d_{n,j}(f)|;A)\\
&\preceq \left(a_{q-1}f\log\frac{1}{\mu(\mathcal{F}_q f)} + \sqrt{\frac{a_{q-1}}{a_q}}2^{-q}\sqrt{\log\frac{1}{\mu(\mathcal{F}_q f)}}\right)(\mu(\mathcal{F}_q f)+P(A))\\
&\preceq 2^{-q}\sqrt{\log\frac{1}{\mu(\mathcal{F}_q f)}}(\mu(\mathcal{F}_q f)+P(A))
\end{aligned}
$$

since $\mu(\mathcal{F}_{q+1}f) \le \mu(\mathcal{F}_q f)$. For each q, let $\Omega = \cup_k \Omega_{qk}$ be a partition such that the maximum $\left|\left|\sum_{j=1}^{\tau_n} d_{n,j}(f)\right|\right|_{\mathcal{F}}$ is achieved at f_{qk} on the set Ω_{qk}. For every q, there are as many sets Ω_{qk} as $\mathcal{F}_{qk}$ in the q-th partition. Then

$$
\begin{aligned}
&E\Big(||\sum_{q=q_0+1}^{\infty}\sum_{j=1}^{\tau_n} d_{n,j}(f)||_{\mathcal{F}}\Big)\\
&\le \sum_{q=q_0+1}^{\infty}\sum_k E\Big(|\sum_{j=1}^{\tau_n} d_{n,j}(f_{qk})|;\Omega_{qk}\Big)\\
&\le \sum_k \sup_f \sum_{q=q_0+1}^{\infty} 2^{-q}\sqrt{\log\frac{1}{\mu(\mathcal{F}_q f)}}(\mu(\mathcal{F}_q f_{qk})+P(\Omega_{qk}))\\
&\le 2\sup_f \sum_{q=q_0+1}^{\infty} 2^{-q}\sqrt{\log\frac{1}{\mu(\mathcal{F}_q f)}}.
\end{aligned}
$$

Next, note that

$$
\sum_{q=q_0+1}^{\infty}\sum_{j=1}^{\tau_n} E_{n,j-1}(\Delta_q f)_{nj}(B_q(f))_{nj}
$$

$$\le \sum_{q=q_0+1}^{\infty} \sum_{j=1}^{\tau_n} E_{n,j-1}(\Delta_q f)_{nj}(B_q(f))_{nj}\{(\Delta_q f)_{nj} > a_q f\}$$

$$\le L \cdot \sup_f \sum_{q=q_0+1}^{\infty} \frac{2^{-2q}}{a_q f} \preceq \sup_f \sum_{q=q_0+1}^{\infty} 2^{-q}\sqrt{\log \frac{1}{\mu(\mathcal{F}_q f)}}.$$

Triangular inequality yields

$$E\left(\left\|\sum_{q=q_0+1}^{\infty}\sum_{j=1}^{\tau_n}(\Delta_q f)_{nj}(B_q(f))_{nj}\right\|_{\mathcal{F}}\right) \preceq \sup_f \sum_{q=q_0+1}^{\infty} 2^{-q}\sqrt{\log \frac{1}{\mu(\mathcal{F}_q f)}}.$$

Therefore

$$\lim_{q_0\to\infty}\limsup_{n\to\infty} E\left(\left\|\sum_{q=q_0+1}^{\infty}\sum_{j=1}^{\tau_n}(\Delta_q f)_{nj}(B_q(f))_{nj}\right\|_{\mathcal{F}}\right) = 0.$$

This proves that

$$\lim_{q_0\to\infty}\limsup_{n\to\infty} E\left(\left\|\sum_{q=q_0+1}^{\infty}\sum_{j=1}^{\tau_n}[V_{nj}^0(f) - V_{nj}^0(\pi_q f)](B_q f)_{nj}\right\|_{\mathcal{F}}\right) = 0.$$

By considering the martingale differences

$$(\Delta_q f)_{nj}(A_{q-1}(f))_{nj} - E_{n,j-1}(\Delta_q f)_{nj}(A_{q-1}(f))_{nj}$$

and

$$\sum_{j=1}^{\tau_n} E_{n,j-1}[(\Delta_q f)_{nj}(A_{q-1}(f))_{nj}]^2 \le L \cdot 2^{-2(q-1)},$$

the third term can be handled in a similar way

$$\lim_{q_0\to\infty}\limsup_{n\to\infty} E\left(\left\|\sum_{q=q_0+1}^{\infty}\sum_{j=1}^{\tau_n}[V_{nj}^0(\pi_q f) - V_{nj}^0(\pi_{q-1} f)](A_{q-1} f)_{nj}\right\|_{\mathcal{F}}\right) = 0.$$

The proof of the claim is completed. □

Finally if the q_0-th partition consists of $N_{q_0}^n$ sets, then Freedman inequality and Lemma 2 in Bae *et al.* [2] yield

$$E \sup_{\rho_{\tau_n}(f,g)<\delta_n} \left|\sum_{j=1}^{\tau_n}[V_{nj}^0(\pi_{q_0}^n f) - V_{nj}^0(\pi_{q_0}^n g)]\right| \preceq \log N_{q_0}^n \eta_n + \sqrt{\log N_{q_0}^n}(2^{-q_0} + \delta_n).$$

The entropy condition implies that

$$\lim_{q_0\to\infty}\limsup_{n\to\infty} E \sup_{\rho_{\tau_n}(f,g)<\delta_n} \left|\sum_{j=1}^{\tau_n}\left[V_{nj}^0(\pi_{q_0}^n f) - V_{nj}^0(\pi_{q_0}^n g)\right]\right| = 0.$$

Combining this with Claim (1) we see that

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} E^* \sup_{\rho_{\tau_n}(f,g)<\delta} \left| \sum_{j=1}^{\tau_n} \left[V_{nj}^0(f) - V_{nj}^0(g) \right] \right| = 0.$$

The proof of Proposition 1 is completed. □

4. Applications to stationary martingale differences

In this section we introduce a setup of a stationary martingale-difference and consider applications.

Let S be a set and $\mathcal{B}$ be a σ-field on S. We consider $(\Omega = S^{\mathbb{Z}}, \mathcal{T} = \mathcal{B}^{\mathbb{Z}}, P)$ as the basic probability space. We denote by T the left shift on Ω. We assume that P is invariant under T, i.e., $PT^{-1} = P$, and that T is ergodic. We denote by $X = \dots, X_{-1}, X_0, X_1, \dots$ the coordinate maps on Ω. From our assumptions it follows that $(X_j)_{j\in\mathbb{Z}}$ is a stationary and ergodic process. Next we define for $j \in \mathbb{Z}$ a σ-field $M_j := \sigma(X_i : i \le j)$ and $H_j := \{f : \Omega \to R : f \in M_j \text{ and } f \in L^2(\Omega)\}$. We denote for $f \in L^2(\Omega)$, $E_{j-1}(f) := E(f|M_{j-1})$, and $H_0 \ominus H_{-1} := \{f \in H_0 : E(f \cdot g) = 0 \text{ for } g \in H_{-1}\}$.

As a first application, we will regain the uniform CLT for the stationary martingale-difference sequence. See Bae and Levental [1]. This justifies, in a sense, that those in previous sections can be considered as a generalization of stationary martingale difference sequence to the non-stationary martingale difference arrays. For every $f,\ g \in L^2(\Omega)$ we put $\rho(f,g) := \left[E(f-g)^2\right]^{1/2}$. Consider $\mathcal{F} \subseteq H_0 \ominus H_{-1}$ with an envelope F satisfying $EF^2 < \infty$. From our setup, it follows that, for every $f \in \mathcal{F}$, $\{f(T^j(X)), M_j\}$ is a stationary martingale-difference sequence. Consider the process $\{Z_n(f) : f \in \mathcal{F}\}$ for the stationary martingale-difference defined by

$$Z_n(f) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} f(\xi_j) \text{ for } f \in \mathcal{F}, \tag{11}$$

where $\xi_j := T^j(X)$, $\xi_0 := T^0(X)(= X)$.

We use the following restatement of Definition 1 for the class $\mathcal{F}$.

Definition 3. For $\epsilon > 0$, we define the covering number with bracketing $N_{[\,]}(\epsilon, \mathcal{F}, \rho)$ as the smallest m for which there exists $\{f_{0,\epsilon}^l, f_{0,\epsilon}^u, \dots, f_{m,\epsilon}^l, f_{m,\epsilon}^u\}$ so that for every $f \in \mathcal{F}$ there exist some $0 \le i \le m$ satisfying $f_{i,\epsilon}^l \le f \le f_{i,\epsilon}^u$ and $\rho(f_{i,\epsilon}^l, f_{i,\epsilon}^u) < \epsilon$.

We equip the space $\mathcal{F}$ with the pseudometric ρ so that $(\mathcal{F}, \rho)$ is totally bounded. The following CLT for the sequence of stationary martingale differences generalizes that of IID result, see for example Ossiander [6], to the stationary and ergodic martingale-difference.

Theorem 3. *Suppose that $\mathcal{F}$ has the bracketing entropy with respect to L_2-norm:*

$$\int_0^\infty [\ln N_{[\ ]}(\epsilon, \mathcal{F}, \rho)]^{1/2} d\epsilon < \infty. \tag{12}$$

Suppose that

$$E^* \sup_{f,g \in \mathcal{F}} \frac{E_{-1}\left[f(\xi_0) - g(\xi_0)\right]^2}{\rho^2(f,g)} < \infty. \tag{13}$$

Then

$$Z_n(\cdot) \Rightarrow Z(\cdot) \text{ as random elements of } B(\mathcal{F}).$$

The limiting process $Z = (Z(f) : f \in \mathcal{F}) \in U_B(\mathcal{F}, \rho)$ is mean zero Gaussian with covariance function $EZ(f)Z(g) = Ef(X)g(X)$.

Proof. To apply Theorem 2, we consider the martingale difference arrays

$$\{n^{-1/2} f(\xi_j) : n = 1, 2, \ldots, f \in \mathcal{F}\}.$$

In this case the bracketing numbers in Theorem 2 reduce to $N_{[\ ]}(\epsilon, \mathcal{F}, \rho)$. Therefore the assumption (12) implies that

$$\int_0^{\delta_n} \left[\log N_{[\ ]}(\epsilon, \mathcal{F}, \rho)\right]^{1/2} d\epsilon \to 0 \quad \text{for every } \delta_n \downarrow 0.$$

The assumption (13) is sufficient to the existence of a constant L satisfying

$$P^*\left(\sup_{f,g \in \mathcal{F}} \sum_{j=1}^n \frac{E_{j-1}\left[f(\xi_j) - g(\xi_j)\right]^2}{n\rho^2(f,g)} \geq L\right) \to 0.$$

Since $N_{[\ ]}(1, \mathcal{F}, \rho) < \infty$, we see that

$$||f(\cdot)||_{\mathcal{F}} \leq \sum_{i=0}^{N_{[\ ]}(1,\mathcal{F},\rho)} \left(|f_{i,1}^l(\cdot)| + |f_{i,1}^u(\cdot)|\right) \in \mathcal{L}_2.$$

Hence, using stationarity and the dominate convergence, we obtain for every $\eta > 0$

$$\begin{aligned} &\frac{1}{n}\sum_{j=1}^n E^* ||f(V_j)||_{\mathcal{F}}^2 \{||f(V_j)||_{\mathcal{F}} > \sqrt{n}\eta\} \\ &= E^* ||f(V_0)||_{\mathcal{F}}^2 \{||f(V_0)||_{\mathcal{F}} > \sqrt{n}\eta\} \to 0. \end{aligned}$$

The condition (7) follows from our setting. Finally, the finite dimensional distribution convergence follows from the one dimensional CLT. This completes the proof. □

Remark 2. The result of Theorem 3 of Bae *et al.* [2] is still valid if we replace the uniform integrable entropy condition by the bracketing entropy condition: Suppose that

$$\int_0^{\delta_n} [\ln N_{[\,]}(\epsilon ||F_n||_{P,2}, \mathcal{F}_n, L_2(P))]^{1/2} d\epsilon \to 0 \quad \text{for every } \delta_n \downarrow 0.$$

Remark 3. Our result generalizes that of IID problem. See Theorem 2.11.23 of Van der Vaart and Wellner [9].

References

[1] J. Bae and S. Levental, *Uniform CLT for Markov chains and its invariance principle: A Martingale approach*, J. Theoret. Probab. **8** (1995), no. 3, 549–570.

[2] J. Bae, D. Jun, and S. Levental, *The uniform CLT for Martingale difference arrays under the uniformly integrable entropy*, Bull. Korean Math. Soc. **47** (2010), no. 1, 39–51.

[3] R. M. Dudley, *A Course on Empirical Processes*, Lecture notes in Math. 1097, Springer-Verlag, New York. 1984.

[4] ———, *Uniform Central Limit Theorems*, Cambridge Studies in Advanced Mathematics 63, Cambridge University Press, Cambridge, 1999.

[5] D. Freedman, *On tail probabilities for Martingales*, Ann. Probab. **3** (1975), 100–118.

[6] M. Ossiander, *A central limit theorem under metric entropy with $\mathcal{L}_2$ bracketing*, Ann. Probab. **15** (1987), no. 3, 897–919.

[7] D. Pollard, *Empirical Processes: Theory and Applications*, Regional conference series in Probability and Statistics 2, Inst. Math. Statist., Hayward, CA. 1990.

[8] S. van der Geer, *Empirical Processes in M-Estimation*, Cambridge Series in Statistical and Probabilistic Mathematics. 2000.

[9] A. W. Van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes with Applications to Statistics*, Springer series in Statistics, Springer-Verlag, New York, 1996.

Jongsig Bae
Department of Mathematics
Institute of Basic Science
Sungkyunkwan University
Suwon 440-746, Korea
E-mail address: jsbae@skku.edu

Doobae Jun
Department of Mathematics(and RING)
Gyeongsang National University
Jinju 660-701, Korea
E-mail address: dbjun@gnu.ac.kr

Shlomo Levental
Department of Statistics and Probability
Michigan State University
East Lansing, Michigan 48824, USA
E-mail address: levental@stt.msu.edu