The Annals of Probability
1987, Vol. 15, No. 1, 154–163

THE LAW OF LARGE NUMBERS FOR PARTIAL SUM PROCESSES INDEXED BY SETS[1]

BY EVARIST GINÉ AND JOEL ZINN

Texas A & M University

This note provides necessary and sufficient conditions for the law of large numbers for partial sum processes indexed by sets of $[0,1]^d$ and based on i.i.d. integrable random variables.

1. Introduction. Let X be a real random variable and let $\{X_j\colon j \in \mathbb{N}^d\}$ be a family of independent identically distributed random variables with $\mathscr{L}(X_j) = \mathscr{L}(X)$. The partial sum processes corresponding to $\{X_j\}$ and indexed by subsets A of $[0,1]^d$ are defined as

$$S_n(X, A) := S_n(A) := \sum_{|j| \le n} X_j \delta_{j/n}(A), \qquad A \subset [0,1]^d, \tag{1}$$

where, for $j = (j_1, \dots, j_d) \in \mathbb{N}^d$, we write $|j| = \max_{1 \le k \le d} j_k$. Several central limit theorems and laws of the iterated logarithm have been obtained for the processes $S_n(A)$, uniformly over classes $\mathscr{A}$ of subsets of $[0,1]^d$ under certain metric entropy assumptions (on $\mathscr{A}$). See, e.g., [1], [2], and references therein. Actually these limit theorems are usually proved for "smoothed versions" of S_n (which we define below). It is therefore natural that there should exist a general law of large numbers for S_n and for its smoothed versions so that the CLT and the LIL would be refinements that hold under stronger conditions. The aim of this note is to present such a result.

For references to previous work on the law of large numbers for partial sum processes, see Bass and Pyke [3], which contains the following interesting result. Let $\mathscr{A}$ be a family of Borel measurable sets satisfying

$$\lim_{\delta \to 0} \sup_{A \in \mathscr{A}} \lambda(A(\delta)) = 0, \tag{2}$$

where λ is Lebesgue measure and, for $A \in \mathscr{A}$,

$$A(\delta) := \left\{x \in \mathbb{R}^d\colon \inf_{y \in \partial A} |y - x| < \delta\right\}$$

is the δ annulus about the boundary ∂A of A. Assume also $E|X_j| < \infty$. Then

$$\text{a.s.-} \lim_{n \to \infty} \|S_n(X) - (EX)\lambda\|_{\mathscr{A}} / n^d = 0, \tag{3}$$

Received February 1985.
[1]Work partially supported by National Science Foundation grants DMS-83-18610 and DMS-83-01367.
AMS 1980 *subject classifications*. Primary 60F15; secondary 60G50.
Key words and phrases. Weak and strong laws of large numbers, partial sum processes, metric entropy.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability*.
www.jstor.org

where, if F is a set valued function, we write

(4) $$\|F\|_{\mathscr{A}} := \sup_{A \in \mathscr{A}} |F(A)|.$$

It turns out that condition (2) is appropriate to ensure

(5) $$\left\| \sum_{|j| \le n} \delta_{j/n}/n^d - \lambda \right\|_{\mathscr{A}} \to 0,$$

but that if the X_j are centered (note that $ES_n(X, A)$ need not equal $\lambda(A)EX$), then (3) holds under much weaker conditions which are also necessary. (See Corollary 2, Proposition 1, and Remark 2 below.)

Define, for X, X_j, $j \in \mathbb{N}^d$, i.i.d. random variables, and for finite measures λ_n, $n \in \mathbb{N}$, the processes

(6) $$S_{\lambda_n}(X, A) := \sum_{|j| \le n} X_j \lambda_n(A \cap I_{nj}), \qquad A \in \mathscr{A}, n \in \mathbb{N},$$

where $\mathscr{A}$ is a class of Borel subsets of $[0,1]^d$. Here we assume

(7) $$0 < E|X| < \infty,$$

and, letting

(8) $$I_{nj} := \{(x_1, \dots, x_d) : (j_k - 1)/n < x_k \le j_k/n, 1 \le k \le d\},$$

where $j = (j_1, \dots, j_d)$, we also assume that $\{\lambda_n\}$ is a sequence of positive Borel measures on $[0,1]^d$ satisfying

(9) $$c_1/n^d \le \lambda_n(I_{nj}) \le c_2/n^d \quad \text{for some } c_1 > 0,\ c_2 < \infty, \text{for all } n \in \mathbb{N}.$$

In (1), $\lambda_n = \sum_{|j| \le n} \delta_{j/n}/n^d$ and $S_n/n^d = S_{\lambda_n}$. The usual λ_n in [1] and [2] is $\lambda_n = \lambda$, Lebesgue measure.

One of our two main results states that, under conditions (7) and (9), $\|S_{\lambda_n}(X)\|_{\mathscr{A}}$ converges to zero or not independently of the law of X as long as it is integrable; the other one gives necessary and sufficient conditions for $\|S_{\lambda_n}(X)\|_{\mathscr{A}} \to 0$ (a.s. or in probability) in terms of the size of $\mathscr{A}$ for different metrics associated to $\{\lambda_n\}$ (the size of $\mathscr{A}$ is measured by metric entropy). Let us then recall that if (T, d) is a metric or pseudometric space, the covering number $N(\tau, T, d)$, $0 < \tau \le$ diameter of (T, d), is defined as

$$N(\tau, T, d) := \inf\left\{m : \exists t_1, \dots, t_m \in T \text{ such that } \sup_{t \in T} \min_{r \le m} d(t_r, t) \le \tau\right\}.$$

Then $\ln N(\tau, T, d)$ is the metric entropy of T for the distance d. We will use on $\mathscr{A}$ the distances

(10) $$d_{\lambda_n, p}(A, B) = n^{d(1-1/p)} \left(\sum_{|j| \le n} \left| \lambda_n(A \cap I_{nj}) - \lambda_n(B \cap I_{nj}) \right|^p \right)^{1/p},$$

$$1 \le p < \ A, B \in \mathscr{A},$$

(11) $$d_{\lambda_n, \infty}(A, B) = \max_{|j| \le n} n^d \left| \lambda_n(A \cap I_{nj}) - \lambda_n(B \cap I_{nj}) \right|, \qquad A, B \in \mathscr{A},$$

and will write

(12) $$N_{\lambda_n,p}(\tau,\mathscr{A}) := N(\tau,\mathscr{A},d_{\lambda_n,p}), \qquad 1\le p\le\infty,\ \tau>0.$$

Note that for any $1\le p\le\infty$,

$$d_{\lambda_n,1}\le d_{\lambda_n,p}\le d_{\lambda_n,\infty}$$

and therefore

(13) $$N_{\lambda_n,1}(\tau,\mathscr{A})\le N_{\lambda_n,p}(\tau,\mathscr{A})\le N_{\lambda_n,\infty}(\tau,\infty), \qquad \tau>0.$$

If $\lambda_n=\sum_{|j|\le n}\delta_{j/n}/n^d$ then we denote $d_{\lambda_n,p}$ simply by $d_{n,p}$ and $N_{\lambda_n,p}$ by $N_{n,p}$, and we have:

(14) $$\begin{aligned} d_{n,p}(A,B) &= \Big(\sum_{|j|\le n}\delta_{j/n}(A\,\triangle\,B)/n^d\Big)^{1/p}, \qquad 1\le p<\infty,\ A,B\in\mathscr{A},\\ d_{n,\infty}(A,B) &= \max_{|j|\le n}\delta_{j/n}(A\,\triangle\,B).\end{aligned}$$

So, for $1\le p<\infty$, $d_{n,p}(A,B)$ is the $L_p(\sum_{|j|\le n}\delta_{j/n}/n^d)$ distance between the indicator functions I_A and I_B. Note also that $d_{n,\infty}(A,B)=0$ if and only if $A\cap\{j/n\colon |j|\le n\}=B\cap\{j/n\colon |j|\le n\}$, and $d_{n,\infty}(A,B)=1$ otherwise. Hence, if

(15) $\Delta_n^{\mathscr{A}}:=$ number of different subsets $A\cap\{j/n\colon |j|\le n\}$, $\qquad A\in\mathscr{A}$,

we have that, for all $\tau<1$,

(16) $$N(\tau,\mathscr{A},d_{n,\infty})=\Delta_n^{\mathscr{A}}.$$

The quantities $\Delta_n^{\mathscr{A}}$ were first used in connection with the law of large numbers by Vapnik and Červonenkis [6].

Finally, here is some more notation. $\{\varepsilon_j\colon j\in\mathbb{N}^d\}$ denotes always, in what follows, a Rademacher family of random variables (i.e., the ε_j are i.i.d. and $P\{\varepsilon_j=1\}=P\{\varepsilon_j=-1\}=\frac12$) independent of any other set of random variables that appear in the argument where they are used. Also, we will write $S_{\lambda_n}(\varepsilon,A):=\sum_{|j|\le n}\varepsilon_j\lambda_n(A\cap I_{nj})$, $A\subset[0,1]^d$. $\{X_j'\colon j\in\mathbb{N}^d\}$ denotes always an independent copy of $\{X_j\colon j\in\mathbb{N}^d\}$ so that $\{X_j-X_j'\colon j\in\mathbb{N}^d\}$ is a set of independent symmetric random variables that symmetrizes $\{X_j\}$. Note also that $S_n(X,A)$ and $S_{\lambda_n}(X,A)$ are written as $S_n(X)$, $S_{\lambda_n}(X)$ and $S_n(A)$, $S_{\lambda_n}(A)$ when no confusion is possible.

2. Results and proofs. Before stating the laws of large numbers we make the following trivial (but convenient) observation about measurability.

LEMMA 1. *Let $\mathscr{A}$ be any collection of measurable subsets of $[0,1]^d$, let $\{\lambda_n\}_{n=1}^\infty$ be a family of finite positive Borel measures on $[0,1]^d$, and let $\{X_j\colon j\in\mathbb{N}^d\}$ be i.i.d real random variables with the law of X. Then for all $n\in\mathbb{N}$, $\|S_{\lambda_n}(X)\|_{\mathscr{A}}$ is a Borel measurable function of the $\mathbb{R}^{n^d}$-valued random vector $X_n=\{X_j\colon |j|\le n\}$. If λ_n is discrete, the conclusion holds even if the subsets $A\in\mathscr{A}$ are not measurable.*

PROOF. Let $C_{n,\mathscr{A}} \subset [0, \lambda_n([0,1]^d)]^{n^d}$ be defined as

$$C_{n,\mathscr{A}} = \left\{ \left(\lambda_n \left(A \cap I_{n_j} \right) \right)_{|j| \le n} : A \in \mathscr{A} \right\}.$$

Let $D_{n,\mathscr{A}}$ be a countable dense subset of $\overline{C}_{n,\mathscr{A}}$. Then

$$\left\| S_{\lambda_n}(X) \right\|_{\mathscr{A}} = \sup_{y \in C_{n,\mathscr{A}}} |\langle X, y \rangle| = \sup_{y \in D_{n,\mathscr{A}}} |\langle X, y \rangle|. \qquad \square$$

Now we prove that, under conditions (7) and (9), the sequence $\|S_{\lambda_n}(X)\|_{\mathscr{A}}$ converges to zero or not, a.s. or in probability, independently of the law of X as long as it is integrable. This is in contrast to the law of large numbers for empirical measures ([4], [6], [7]).

THEOREM 1. *Let $\{X_j : j \in \mathbb{N}^d\}$ be i.i.d. real random variables with the law of X such that $0 < E|X| < \infty$ and $EX = 0$. Let $\{\lambda_n\}_{n=1}^{\infty}$ be a sequence of positive Borel measures on $[0,1]^d$ satisfying condition* (9), *and let $\mathscr{A}$ be a class of Borel subsets of $[0,1]^d$ (if the measures λ_n are discrete then the sets in $\mathscr{A}$ need not be Borel). Then the following are equivalent*:

(i) $\lim_{n\to\infty} \|S_{\lambda_n}(X)\|_{\mathscr{A}} = 0$ *a.s.* (*respectively, in probability*);
(ii) $\lim_{n\to\infty} \|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} = 0$ *a.s.* (*respectively, in probability*).

PROOF. Only the case of a.s. convergence is proved since convergence in probability can be treated similarly. For $0 < M < \infty$ and $j \in \mathbb{N}^d$, let $X_j^M = X_j I(|X| > M)$, and let $X_{j,M} = X_j - X_j^M$. Then, since $EX = 0$ and $E|X| < \infty$, the strong law of large numbers in $\mathbb{R}$ gives, as in [4], Theorem 8.3, that

$$\begin{aligned} &\text{a.s.-} \lim_{M\to\infty} \limsup_{n\to\infty} \left\| S_{\lambda_n}(X^M - EX^M) \right\|_{\mathscr{A}} \\ &\quad \le \text{a.s.-} \lim_{M\to\infty} \limsup_{n\to\infty} \sum_{|j|\le n} c_2 |X_j^M - EX_j^M| / n^d \\ &\quad = c_2 \lim_{M\to\infty} E|X^M - EX^M| = 0. \end{aligned} \tag{17}$$

Therefore, by considering $(X_{j,M} - EX_{j,M})/2M$ instead of X_j, $j \in \mathbb{N}^d$, we may (and do) assume that the random variables X_j in (i) are centered and bounded by 1.

Suppose (i) holds. Then

$$\text{a.s.-} \lim_{n\to\infty} \left\| S_{\lambda_n}(X - X') \right\|_{\mathscr{A}} = 0.$$

Hence,

$$\lim_{N\to\infty} \sup_{n\ge N} \left\| S_{\lambda_n}(X - X') \right\|_{\mathscr{A}} = 0 \quad \text{in probability.}$$

By Lemma 1, for all $N \in \mathbb{N}$,

$$\mathscr{L}\left(\sup_{n\ge N} \|S_{\lambda_n}(X - X')\|_{\mathscr{A}} \right) = \mathscr{L}\left(\sup_{n\ge N} \|S_{\lambda_n}(\varepsilon|X - X'|)\|_{\mathscr{A}} \right).$$

So

$$\lim_{N\to\infty} \sup_{n\geq N} \left\| S_{\lambda_n}(\varepsilon|X - X'|) \right\|_{\mathscr{A}} = 0 \quad \text{in probability.}$$

By applying Hoffmann–Jørgensen's inequality ([5], pages 164–165; e.g., [4], Lemma 2.8) to the random variables

$$Y_{j,N} = \left\{ \varepsilon_j |X_j - X_j'| \lambda_n(A \cap I_{nj}) \colon n \geq j \vee N, A \in \mathscr{A} \right\} \in l^\infty(\mathbb{N} \times \mathscr{A}),$$

we obtain

$$\lim_{N\to\infty} E \sup_{n\geq N} \left\| S_{\lambda_n}(\varepsilon|X_j - X_j'|) \right\| = \lim_{N\to\infty} E \left\| \sum_j Y_{j,N} \right\| = 0. \tag{18}$$

(Note that, although $l^\infty(\mathbb{N} \times \mathscr{A})$ is not separable, $\|Y_j\|$ is a random variable by Lemma 1.) But, by Fubini's theorem and Jensen's inequality,

$$E \sup_{n\geq N} \left\| S_{\lambda_n}(\varepsilon|X - X'|) \right\|_{\mathscr{A}} \geq E|X - X'| E \sup_{n\geq N} \left\| S_{\lambda_n}(\varepsilon) \right\|_{\mathscr{A}}.$$

This, together with (18), proves (ii).

Assume now that (ii) holds. Using convexity and the contraction principle in [5], Corollary 4.2, applied to the variables $Y_{j,N}$, it follows that

$$E \sup_{n\geq N} \left\| S_{\lambda_n}(X) \right\|_{\mathscr{A}} \leq E \sup_{n\geq N} \left\| S_{\lambda_n}(X - X') \right\|_{\mathscr{A}}$$

$$= E \sup_{n\geq N} \left\| S_{\lambda_n}(\varepsilon|X - X'|) \right\|_{\mathscr{A}} \leq 2E \sup_{n\geq N} \left\| S_{\lambda_n}(\varepsilon) \right\|_{\mathscr{A}}.$$

But if (ii) holds then this last expectation tends to zero as $N \to \infty$ by Hoffmann–Jørgensen's inequality. Hence $\|S_{\lambda_n}(X)\|_{\mathscr{A}} \to 0$ a.s. □

The next result gives concrete conditions on the class $\mathscr{A}$ which are necessary and sufficient for convergence of $\{\|S_{\lambda_n}\|_{\mathscr{A}}\}$.

Theorem 2. *Let X, $\{X_j\colon j \in \mathbb{N}^d\}$, $\{\lambda_n\}$, and $\mathscr{A}$ be as in Theorem 1. Then the following are equivalent:*

(i) $\lim_{n\to\infty} \|S_{\lambda_n}(X)\|_{\mathscr{A}} = 0$ *a.s.*;
(ii) $\lim_{n\to\infty} \|S_{\lambda_n}(X)\|_{\mathscr{A}} = 0$ *in probability*;
(iii) $\lim_{n\to\infty} [\ln N_{\lambda_n, p}(\tau, \mathscr{A})]/n^d = 0$ *for some* $p \in [1, \infty]$ *and all* $\tau > 0$;
(iv) $\lim_{n\to\infty} [\ln N_{\lambda_n, p}(\tau, \mathscr{A})]/n^d = 0$ *for every* $p \in [1, \infty]$ *and all* $\tau > 0$.

Proof. By Theorem 1 and (13), it is enough to prove:

(I) $[\ln N_{\lambda_n, 1}(\tau, \mathscr{A})]/n^d \to 0$ for all $\tau > 0 \Rightarrow \|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} \to 0$ a.s. and
(II) $\|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} \to 0$ in probability $\Rightarrow [\ln N_{\lambda_n, \infty}(\tau, \mathscr{A})]/n^d \to 0$ for all $\tau > 0$.

Proof of statement (I): Given $\tau > 0$ let $\mathscr{A}_{\tau/2} \subset \mathscr{A}$ be the family of centers of a minimal covering of $\mathscr{A}$ by $d_{\lambda_n, 1}$-balls of radius not larger than $\tau/2$ and center in $\mathscr{A}$. Then $\#\mathscr{A}_{\tau/2} = N_{\lambda_n, 1}(\tau/2, \mathscr{A})$, and, by hypothesis, for all $\tau > 0$ there exists N_τ such that if $n \geq N_\tau$, then

$$N_{\lambda,1}(\tau/2, \mathscr{A}) \leq \exp\{\tau^2 n^d/16c_2^2\}.$$

Then, since for each $A \in \mathscr{A}$ there is $B \in \mathscr{A}_{\tau/2}$ such that $|S_{\lambda_n}(\varepsilon, A) - S_{\lambda_n}(\varepsilon, B)| \leq d_{\lambda_n,1}(A, B) \leq \tau/2$, the standard sub-Gaussian estimate ([4], inequality (2.17)) gives

$$P\{\|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} > \tau\} \leq P\{\|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}_{\tau/2}} > \tau/2\}$$

$$\leq N_{\lambda_n,1}(\tau/2, \mathscr{A}) \sup_{A \in \mathscr{A}} P\{|S_{\lambda_n}(\varepsilon, A)| > \tau/2\}$$

$$\leq 2N_{\lambda_n,1}(\tau/2, \mathscr{A})\exp\{-\tau^2 n^d/8c_2^2\} \leq 2\exp\{-\tau^2 n^d/16c_2^2\}.$$

Therefore $\sum_n P\{\|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} > \tau\} < \infty$ for all $\tau > 0$ and statement (I) is proved.

Proof of statement (II): As shown in the proof of Theorem 1, if $\|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} \to 0$ in probability, then

$$\lim_{n\to\infty} E\|S_{\lambda_n}(\varepsilon)\|_{\mathscr{A}} = 0. \tag{19}$$

By dividing by c_2 if necessary, we may assume $n^d\lambda_n(I_{nj}) \leq 1$. Then $n^d C_{n,\mathscr{A}}$ is a subset of $[0,1]^d$ ($C_{n,\mathscr{A}}$ is defined in the proof of Lemma 1 above). Let $L_{n,\mathscr{A}} \subset [0,1]^{n^d}$ be the convex hull of $n^d C_{n,\mathscr{A}}$; that is,

$$L_{n,\mathscr{A}} = \text{convex hull of } \Big\{x \in [0,1]^{n^d}: x = (x_j)_{|j| \leq n},$$

$$x_j = n^d\lambda_n(A \cap I_{nj}), A \in \mathscr{A}\Big\}.$$

So, (19) is just

$$\lim_{n\to\infty} E \sup_{x \in L_{n,\mathscr{A}}} \Big|\sum_{|j|\leq n} \varepsilon_j x_j\Big| \Big/ n^d = 0. \tag{19′}$$

Let $N_{n,\infty}(\tau, L_{n,\mathscr{A}})$ be the covering number of $L_{n,\mathscr{A}}$ for the distance $d(x, y) = \max_{|j|\leq n}|x_j - y_j|$, $x, y \in [0,1]^{n^d}$. The proof of Lemma 4 in [7] shows that there exists $t(\tau) < \infty$, independent of n, such that, if

$$N_{n,\infty}(3\tau/2, L_{n,\mathscr{A}}) > \exp\{2n^d \ln(1+\tau)\}, \tag{20}$$

then

$$E \sup_{x \in L_{n,\mathscr{A}}} \Big|\sum_{|j|\leq n} \varepsilon_j x_j\Big| \Big/ n^d \geq \tau\big[(1+\tau)^{1/3} - 1\big]\big(t(\tau) - n^{-d}\big)/2. \tag{21}$$

Since $N_{n,\infty}(\tau, L_{n,\mathscr{A}}) \geq N_{\lambda_n,\infty}(\tau, \mathscr{A})$ for all $\tau > 0$, if (iii) with $p = \infty$ does not hold, then for some $\tau > 0$ there exists N_τ' such that (20) holds for $n \geq N_\tau'$. Therefore (21) also holds for these values of n, in contradiction with (19′). □

Remark 1. Here is a different proof of Theorem 2 for $p = 2$. By Theorem 1 the law of large numbers is independent of X. So we can take $x_j = g_j$ i.i.d.

standard normal random variables, and then $S_{\lambda_n}(g, A) = \sum_{|j| \le n} g_j \lambda_n(A \cap I_{nj})$ is a Gaussian process. Using Dudley's majorization and Sudakov's minorization as in [4], Theorem 8.3, one obtains Theorem 2 for $p = 2$. Actually, Sudakov's minorization gives that the condition

$$\sup_{\tau > 0} \tau^2 \ln N_{\lambda_n, 2}(n^{d/2}\tau, \mathscr{A}) \to 0 \quad \text{as } n \to \infty$$

is necessary for $\|S_{\lambda_n}\|_{\mathscr{A}} \to 0$. This is a stronger statement than necessity in Theorem 2 for $p = 2$.

Recall the definition of $S_n(X, A)$ from (1) as well as those of $d_{n,p}$, $d_{n,\infty}$, and $\Delta_n^{\mathscr{A}}$ ((14)–(16)) for the case $\lambda_n = \sum_{|j| \le n} \delta_{j/n}/n^d$. In this case, Theorem 2 gives:

Corollary 1. *Let X and X_j, $j \in \mathbb{N}^d$, be as in Theorem 1, and let $\mathscr{A}$ be a class of subsets of $[0,1]^d$. Then the following are equivalent:*

(i) $\|S_n(X)/n^d\|_{\mathscr{A}} \to 0$ *a.s.;*
(ii) $\|S_n(X)/n^d\|_{\mathscr{A}} \to 0$ *in probability;*
(iii) *for all $\tau > 0$ and for some (all) $p \in [1, \infty)$,* $[\ln N(\tau, \mathscr{A}, d_{n,p})]/n^d \to 0$;
(iv) $[\ln \Delta_n^{\mathscr{A}}]/n^d \to 0$.

Next we consider the situation studied in [3]. Here λ is Lebesgue measure.

Corollary 2. *Let X, X_j, $j \in \mathbb{N}^d$, be i.i.d. real random variables such that $0 < E|X| < \infty$ and $EX \ne 0$. Let $\mathscr{A}$ be a family of Borel subsets of $[0,1]^d$. Then*

$$\lim_{n \to \infty} \|S_n(X)/n^d - (EX)\lambda\|_{\mathscr{A}} = 0 \tag{22}$$

a.s. (or equivalently in probability) if and only if both,

(i) *any of the conditions* (i)–(iv) *in Corollary* 1 *hold for $X - EX$ and*
(ii) $\lim_{n \to \infty} \|\lambda - \sum_{|j| \le n} \delta_{j/n}/n^d\|_{\mathscr{A}} = 0$.

Proof. Since

$$\begin{aligned} &S_n(X, A)/n^d - (EX)\lambda(A) \\ &\quad = S_n(X - EX, A)/n^d + (EX)\left(\textstyle\sum_{|j| \le n} \delta_{j/n}(A)/n^d - \lambda(A)\right), \quad A \in \mathscr{A}, \end{aligned} \tag{23}$$

sufficiency of conditions (i) and (ii) is obvious. Suppose now that (22) holds with convergence in probability. Then (22) also holds with the variables X_j replaced by their symmetrizations $X_j - X_j'$; that is,

$$\lim_{n \to \infty} \|S_n(X - X')/n^d\|_{\mathscr{A}} = 0$$

in probability. Since $E|X_j - X_j'| \ne 0$, condition (iv) in Corollary 1 holds and therefore so do all the others for an i.i.d. set $\{Y_j: j \in \mathbb{N}^d\}$ of integrable, centered random variables. In particular we can take $Y_j = X_j - EX_j$. So,

$$\|S_n(X - EX)/n^d\|_{\mathscr{A}} \to 0$$

in probability and (a.s.). This and (22), (23) give condition (ii) by the triangle inequality. □

The following proposition connects Corollaries 1 and 2 with the law of large numbers of Bass and Pyke [3] described in the Introduction.

Proposition 1. *Let $\mathscr{A}$ be a class of Borel subsets of $[0,1]^d$ such that*

$$\lim_{\delta \to 0} \sup_{A \in \mathscr{A}} \lambda(A(\delta)) = 0. \tag{2}$$

Then

(i) *for all* $\tau > 0$, $\sup_n N_{n,1}(\tau, \mathscr{A}) < \infty$, *and*
(ii) $\lim_{n\to\infty} \|\sum_{|j| \le n} \delta_{j/n}/n^d - \lambda\|_{\mathscr{A}} = 0$.

Proof. Let us denote

$$r_{\mathscr{A}}(\delta) := \sup_{A \in \mathscr{A}} \lambda(A(\delta)), \qquad \delta > 0.$$

Given $\tau > 0$, let m_τ be such that

$$r_{\mathscr{A}}\left(d^{1/2}/m_\tau\right) < \tau/8, \tag{24}$$

and let N_τ be such that for all $n \ge N_\tau$,

$$\left\| \sum_{|j| \le n} \delta_{j/n}/n^d - \lambda \right\|_{\mathscr{R}_{m_\tau}} < \tau/8, \tag{25}$$

where, for each $m \in \mathbb{N}$,

$$\mathscr{R}_m = \left\{B\colon B \text{ is a union of squares } I_{mj},\ |j| \le m\right\}.$$

(2) ensures $m_\tau < \infty$. N_τ exists because $\sum_{|j| \le n}\delta_{j/n}/n^d \to_{w^*} \lambda$ and $\mathscr{R}_m$ consists of a finite number of λ-continuity sets.

Given $A \in \mathscr{A}$, let $A_m := \bigcup\{I_{mj}\colon |j| \le m, I_{mj} \subseteq A\}$ and $A^m := \bigcup\{I_{mj}\colon |j| \le m, A \cap I_{mj} \ne \varnothing\}$. Then $A^m \setminus A_m \in \mathscr{R}_m$ and $A^m \setminus A_m \subseteq A(d^{1/2}/m)$.

So by (24) and (25), if $A \in \mathscr{A}$ and $n \ge N_\tau$,

$$\begin{aligned} \sum_{|j| \le n} \delta_{j/n}(A^{m_\tau} \setminus A)/n^d &\le \sum_{|j| \le n} \delta_{j/n}\left(A^{m_\tau} \setminus A_{m_\tau}\right)/n^d \\ &\le \lambda\left(A^{m_\tau} \setminus A_{m_\tau}\right) + \tau/8 \le \tau/4. \end{aligned} \tag{26}$$

Then (26), the triangle inequality, and (25) give

$$N_{n,1}(\tau, \mathscr{A}) \le N_{n,1}\left(\tau/4, \mathscr{R}_{m_\tau}\right) \le N\left(\tau/8, \mathscr{R}_{m_\tau}, \lambda\right), \qquad n \ge N_\tau,$$

where we denote $N(\tau, \mathscr{R}_m, \lambda) := N(\tau, \mathscr{R}_m, d_\lambda)$ with $d_\lambda(A, B) = \lambda(A \vartriangle B)$. But $N(\tau/6, \mathscr{R}_{m_\tau}, \lambda)$ is independent of n. This proves (i). The proof of (ii) follows along similar lines and is omitted. □

Remark 2. The previous proposition is a source of examples of classes satisfying Corollary 2 (and therefore also Corollary 1). Here are some trivial examples of (a) classes of sets that satisfy Corollary 1 but not (22) and (b) classes that satisfy (22) but not condition (2). Let $B_r \subset [0,1]$, $r \in \mathbb{Q} \cap [0,1]$, be Borel subsets of $[0,1]$ and let $A_r = \{(x, y): x = r,\ y \in B_r\} \subset [0,1]^2$. Define

$$\mathscr{A} = \{A: A \text{ is a union of } A_r\text{'s}\}.$$

Then $\Delta_n^{\mathscr{A}} \leq 2^n$ ($= 2^n$ if $B_r \cap \{k/n, 1 \leq k \leq n\} \neq \phi$ for all n) so that

$$[\ln \Delta_n^{\mathscr{A}}]/n^2 \to 0$$

and Corollary 1 holds. If $B_r = \mathbb{Q} \cap [0,1]$ for all r, then

$$\Big\| \sum_{|j| \leq n} \delta_{j/n}/n^d - \lambda \Big\|_{\mathscr{A}} = 1$$

and (22) does not hold. If $B_r = Q^c \cap [0,1]$ for all r, then

$$\Big\| \sum_{|j| \leq n} \delta_{j/n}/n^d - \lambda \Big\|_{\mathscr{A}} = 0,$$

but condition (2) fails to hold.

If $\lambda_n = \mu$, $n \in \mathbb{N}$, in Theorem 2 (e.g., Lebesgue measure as in [1] and [2]), then

$$\begin{aligned} d_{\lambda_n, 1}(A, B) &= \sum_{|j| \leq n} |\mu(A \cap I_{nj}) - \mu(B \cap I_{nj})| \\ &\leq \sum_{|j| \leq n} \mu((A \cap I_{nj}) \triangle (B \cap I_{nj})) \qquad (27) \\ &= \mu(A \triangle B), \qquad n \in \mathbb{N}. \end{aligned}$$

So, if we let $N(\tau, \mathscr{A}, \mu)$ denote the covering number of $\mathscr{A}$ for the distance $d(A, B) = \mu(A \triangle B)$, then (27) and Theorem 1 give the following:

COROLLARY 3. *Let $\{X_j: j \in \mathbb{N}^d\}$ be i.i.d. with $E|X_j| < \infty$ and $EX_j = 0$. Let μ be a positive measure on $[0,1]^d$ such that for some $c < \infty$ and for all $n \in \mathbb{N}$ and $|j| \leq n$,*

$$\mu(I_{nj}) \leq c/n^d.$$

Let $\mathscr{A}$ be a class of Borel measurable subsets of $[0,1]^d$ such that

$$N(\tau, \mathscr{A}, \mu) < \infty \qquad (28)$$

for all $\tau > 0$. Then

$$\lim_{n \to \infty} \Big\| \sum_{|j| \leq n} X_j \mu(\cdot \cap I_{nj}) \Big\|_{\mathscr{A}} = 0 \quad a.s. \qquad (29)$$

REMARK 3. (i) Condition (28) for $\mathscr{A}$ is equivalent to $\mathscr{A}$ being relatively compact in $L_1(\mu)$ (as a class of indicator functions).

(ii) Arguments similar to those in Proposition 1 show that if $\mathscr{A}$ satisfies condition (2) for μ (instead of λ) then it also satisfies (28) and therefore (29).

(iii) But condition (28) is not necessary for the law of large numbers (29) to hold. To see this let $\mathscr{B}_1 \subset [0,1]$ be the class of all Borel sets of $[0,1]$ and let $\mathscr{A} = \{B \times [0,1]\colon B \in \mathscr{B}_1\} \subset [0,1]^2$. Let $\mu = \mu_n$ = Lebesgue measure on $[0,1]$, and let $\mu^2 = \mu_n^2$ = Lebesgue measure on $[0,1]^2$, $n \in \mathbb{N}$. Then, since $\mathscr{B}_1$ is not totally bounded in $L_1(\mu)$, neither is $\mathscr{A}$ totally bounded in $L_1(\mu^2)$: that is, (28) does not hold for $\mathscr{A}$ and μ^2. However, for $\tau > 0$,

$$N_{\mu_n^2, 1}(\tau, \mathscr{A}) = N_{\mu_n, 1}(\tau, \mathscr{B}) \le N_{\mu_n, \infty}(\tau, \mathscr{B}) \le \left(\left[\tau^{-1}\right] + 1\right)^n,$$

and therefore,

$$\left[\ln N_{\mu_n^2, 1}(\tau, \mathscr{A})\right]/n^2 \to 0,$$

showing that the pair $(\mathscr{A}, \mu^2)$ satisfies the law of large numbers (29).

REFERENCES

[1] ALEXANDER, K. and PYKE, R. (1986). A uniform central limit theorem for set-indexed partial sum processes with finite variance. *Ann. Probab.* **14** 582–597.

[2] BASS, R. F. (1985). Law of the iterated logarithm for set-indexed partial sum processes with finite variance. *Z. Wahrsch. verw. Gebiete* **70** 591–608.

[3] BASS, R. F. and PYKE, R. (1984). A strong law of large numbers for partial-sum processes indexed by sets. *Ann. Probab.* **12** 268–271.

[4] GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–989.

[5] HOFFMANN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–182.

[6] VAPNIK, V. N. and ČERVONENKIS, A. JA. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.* **16** 164–180.

[7] VAPNIK, V. N. and ČERVONENKIS, A. JA. (1981). Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory Probab. Appl.* **26** 532–553.

DEPARTMENT OF MATHEMATICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843