The Annals of Probability
1998, Vol. 26, No. 3, 1370–1383

THE p-VARIATION OF PARTIAL SUM PROCESSES AND THE EMPIRICAL PROCESS

BY JINGHUA QIAN

Bucknell University

The p-variation of a function f is the supremum of the sums of the pth powers of absolute increments of f over nonoverlapping intervals. Let F be a continuous probability distribution function. Dudley has shown that the p-variation of the empirical process is bounded in probability as $n \to \infty$ if and only if $p > 2$, and for $1 \le p \le 2$, the p-variation of the empirical process is at least $n^{1-p/2}$ and is at most of the order $n^{1-p/2}(\log\log n)^{p/2}$ in probability. In this paper, we prove that the exact order of the 2-variation of the empirical process is $\log\log n$ in probability, and for $1 \le p < 2$, the p-variation of the empirical process is of exact order $n^{1-p/2}$ in expectation and almost surely.

Let $S_j := X_1 + X_2 + \cdots + X_j$. Then the p-variation of the partial sum process for $\{X_1, X_2, \dots, X_n\}$ is defined as that of f on $(0, n]$, where $f(t) = S_j$ for $j-1 < t \le j$, $j = 1, 2, \dots, n$. Bretagnolle has shown that the expectation of the p-variation for independent centered random variables X_i with bounded pth moments is of order n for $1 \le p < 2$. We prove that for $p = 2$, the 2-variation of the partial sum process of i.i.d. centered nonconstant random variables with finite $2+\delta$ moment for some $\delta > 0$ is of exact order $n\log\log n$ in probability.

1. Introduction. Wiener first defined p-variation in 1924 [19]. He mainly focused on the case $p = 2$, the 2-variation. For p-variations with $p \ne 2$, the first major work was done by Young [20], partly with Love [12]. More recent applications in probability theory include Bretagnolle [2], which proves an important result for sums of independent mean zero random variables; Lépingle [11], which treats the p-variation for semimartingales; and Pisier and Xu [14], which treats martingales and Banach space interpolation. The p-variation for empirical processes seems to have been first addressed explicitly in a sequence of papers by Dudley [4, 6, 7].

The definitions of the empirical distribution function, the empirical process and the p-variation of a function f are given as follows.

DEFINITION 1.1. Let $X_1, X_2, \dots, X_n$ be i.i.d. (independent, identically distributed) random variables with d.f. (distribution function) F. Let $F_n(x) := n^{-1}\sum_{i=1}^n 1_{(-\infty,x]}(X_i)$ and $\alpha_n := \sqrt{n}(F_n - F)$. Then F_n is called the empirical d.f. of F and α_n is the empirical process.

To introduce p-variation, we define the more general ψ-variation first.

Received February 1998.

AMS 1991 *subject classifications.* Primary 62G30, 62G20, 60G50; secondary: 26A45, 26A48.

Key words and phrases. p-variation norm, ψ-variation, the empirical process, partial sum processes.

DEFINITION 1.2. Let ψ be a Young–Orlicz modulus. That is, ψ is a strictly increasing, convex function from $[0, \infty)$ onto itself. Let f be a function from an interval $J \subset \mathrm{R}$ into R. The *ψ-variation* of f is defined by

$$v_\psi(f) := \sup\left\{\sum_{i=1}^{m} \psi(|f(x_i) - f(x_{i-1})|)\colon x_0 \in J,\ x_0 < x_1 < \cdots < x_m \in J,\ m = 1, 2, \ldots\right\}.$$

If $\psi(x) \equiv x^p$, where $p \geq 1$, then the *p-variation* of f is $v_p(f) := v_\psi(f)$. For a sequence of random variables $X_1, X_2, \ldots, X_n$ and $S_j := \sum_{i=1}^{j} X_i$, the p-variation for partial sums is defined by

$$v_p(\{X_i\}_{i=1}^n) := \max\left\{\sum_{i=1}^{k} |S_{j_i} - S_{j_{i-1}}|^p\colon 0 = j_0 < j_1 < \cdots < j_k = n,\ k = 1, 2, \ldots, n\right\}.$$

REMARKS. For $p = 1$, the p-variation is the usual total variation. For $p = 2$, the 2-variation we are studying here is actually different from the widely studied quadratic variation of a stochastic process, which is defined by a limit in probability as the mesh of partitions goes to 0, or an almost sure limit under further restrictions.

For $p \geq 1$ and $-\infty \leq a < b \leq +\infty$, let $W_p := W_p[a, b]$ be the class of all real-valued functions f on $[a, b]$ with finite p-variation $v_p(f) = v_p(f, [a, b]) < \infty$. For any $f \in W_p$, define $\|f\|_{(p)} := v_p(f)^{1/p}$. Then $\|f\|_{(p)}$ is a seminorm which is 0 only for constants. If we let $\|f\|_{[p]} := \|f\|_{(p)} + \|f\|_\infty$, where $\|f\|_\infty := \sup_x |f(x)|$, then $\|\cdot\|_{[p]}$ is a norm on W_p. It can be shown that W_p is complete for $\|\cdot\|_{[p]}$ and so W_p is a Banach space.

In [4], [6] and [7], Dudley suggests using the p-variation norm to replace the supremum norm when considering the differentiability of some statistical functionals. The reasons are severalfold: first, Fréchet derivatives often exist for p-variation norms while they do not for the supremum norm. Second, uniformly over all possible norms on spaces containing empirical distribution functions, the p-variation norm gives remainder bounds in the differentiation which are of the smallest possible order in a range of cases. Third, the p-variation norm retains the good property of the supremum norm of being invariant under all strictly increasing, continuous transformations of R onto itself. Fourth, the central limit theorem for the ordinary empirical process in the p-variation norms holds for $p > 2$. These as well as other results given in [4], [6] and [7] show the importance and usefulness of the p-variation spaces and norms.

For $p > 2$, the central limit theorem for empirical processes in p-variation holds [4] and there exist versions α_n^* and B_n^* of the empirical process α_n and Brownian bridge B, respectively, on some probability space, such that $\mathrm{E}\|\alpha_n^* -$

$B_n^*\|_{[p]} \le C(p)n^{(1/p)-1/2}$, where $C(p)$ is a constant depending only on p and the bound is optimal up to the constant $C(p)$ [9]. For $p \le 2$, the central limit theorem for empirical processes cannot hold since the p-variation norm of the empirical process is not bounded in probability. Dudley ([7], Theorem 2.2) shows the following:

THEOREM A (Dudley). *Let F be any d.f. and F_n be an empirical d.f. for F. Then for $1 \le p \le 2$, $v_p(n^{1/2}(F_n - F)) = O_p(n^{1-p/2}(\mathsf{LL}\, n)^{p/2})$ as $n \to \infty$. Conversely, if F is continuous, then almost surely for all n, $v_p(n^{1/2}(F_n - F)) \ge n^{1-p/2}$.*

A Young–Orlicz modulus ψ as in Definition 1.2 is said to satisfy the Δ_2 condition if, for some constant $L < \infty$, $\psi(2x) \le L\psi(x)$ for all $x > 0$. Taylor [18] proved the following:

THEOREM B (Taylor). *For the Brownian motion process x_t on $[0, M]$, where $0 < M < \infty$, almost surely $v_{\psi_1}(x_t) < \infty$, where $\mathsf{L}\, x := \max(1, \log x)$, $\psi_1(x) := x^2/\mathsf{LL}(1/x)$ for $0 < x \le e^{-e}$ and $\psi_1(x)$ is defined for $x > e^{-e}$ in any way such that $\psi_1(x)$ is a Young–Orlicz modulus satisfying Δ_2. Moreover, if $\psi_1(x) = o(\psi(x))$ as x decreases to 0, then $v_\psi(x) = +\infty$ almost surely.*

In this paper, we will establish the exact order of $v_p(\alpha_n)$ in probability for $p = 2$, and both in expectation and almost surely for $1 \le p < 2$. In particular, we show that for $p = 2$, the upper bound in Theorem A is sharp up to a constant, while for $1 \le p < 2$, the lower bound is sharp up to a constant. We also show that for $1 \le p < 2$, the almost sure bound is of the same order as the bound in expectation, thus also is of the same order as the bound in probability.

For the p-variation of partial sums, Bretagnolle ([2], Theorem 2) proved:

THEOREM C (Bretagnolle). *Let $X_1, X_2, \ldots, X_n$ be independent random variables with $\mathsf{E}\, X_i = 0$ for $i = 1, 2, \ldots, n$. Then for $p \in (0, 2)$, there exists a constant c_p, depending only on p, such that $\mathsf{E}\, v_p(\{X_i\}_{i=1}^n) \le c_p \sum_{i=1}^n \mathsf{E}|X_i|^p$.*

The paper is organized as follows: In Section 2, we study the 2-variation of partial sum processes. In Section 3, we use the connection between the p-variation of the partial sum process and that of the empirical process to obtain results for the empirical process. In Section 4, we study the almost sure asymptotic behavior of the p-variation of $F_n - F$ for a continuous d.f. F and its empirical d.f. F_n with $1 < p < 2$.

2. The 2-variation of partial sum processes.

In this section, we show that the 2-variation of the partial sum process of i.i.d. centered random variables Y_i with finite $2 + \delta$ moment for some $\delta > 0$ is of order $n\, \mathsf{LL}\, n$ in probability.

To get a lower bound for the 2-variation of the partial sum process of Y_i, we begin with the simple case in which X_i are i.i.d. $N(0,1)$ random variables. It is easily seen that the following holds for some $c > 0$ in place of 1/12. For a proof of the given statement, see Qian ([15], Lemma 2.1).

LEMMA 2.1. *Let $\Phi(x)$ be the standard normal distribution function. Then $\Phi(-x) = 1 - \Phi(x) \geq (1/12)\exp(-3x^2/4)$ for $x \geq 1$.*

Now, let $\{X_i\}_{i\geq 1}$ be i.i.d. $N(0,1)$, $S_k := \sum_{i=1}^k X_i$ for $k \geq 1$ and $S_0 := 0$. Let $K := 25$. Given n, let

$$l := l(n) := \left\lfloor \frac{\log n}{4 \log K} \right\rfloor \quad \text{and} \quad m := m(n) := \left\lfloor \frac{\log n}{2 \log K} \right\rfloor, \tag{1}$$

where $\lfloor x \rfloor$ is the largest integer $\leq x$. Let $\mathrm{L}\,x := \max(1, \log x)$ and let

$$E_n := E_n(\omega) := \left\{ t \in \{0, 1, \dots, n - \lfloor \sqrt{n} \rfloor\} : \omega \in \bigcup_{j=l}^{m} \left\{ |S_{t+K^j}(\omega) - S_t(\omega)| \geq \sqrt{K^j \,\mathrm{LL}\, n}/2 \right\} \right\}.$$

LEMMA 2.2. *For n large enough ($n \geq n_1$, where n_1 is an absolute constant) and for any $t \in \{1, 2, \dots, n - \lfloor \sqrt{n} \rfloor\}$, $\mathrm{P}(\omega : t \in E_n(\omega)) > 1 - p_n$, where $p_n := \exp(-(\log n)^{1/5}/(52 \log K))$.*

PROOF. For the given n and t, let

$$E_n^j := E_n^j(t) := \left\{ \omega : |S_{t+K^j}(\omega) - S_{t+K^{j-1}}(\omega)| \geq \sqrt{K^j \,\mathrm{LL}\, n} \right\},$$

for $j = l+1, l+2, \dots, m$. Then by Lemma 2.1,

$$\begin{aligned} \mathrm{P}(E_n^j) &\geq \frac{1}{12} \exp\left(-\frac{3}{4} \frac{K^j}{K^j - K^{j-1}} \mathrm{LL}\, n \right) \geq \frac{1}{12} \exp\left(-\frac{4}{5} \mathrm{LL}\, n \right) \\ &= \frac{1}{12} (\log n)^{-4/5} \quad \text{for } n \geq \lfloor e^e \rfloor + 1 = 16. \end{aligned} \tag{2}$$

Let $E_n' := E_n'(t) := \bigcup_{j=l+1}^m E_n^j(t)$. Note that for different j's, $E_n^j(t)$ depend on disjoint sums of i.i.d. random variables, so $E_n^j(t)$ are independent, and for $n \geq 16$,

$$\begin{aligned} \mathrm{P}(E_n') &= \mathrm{P}\left(\bigcup_{j=l+1}^m E_n^j \right) = 1 - \mathrm{P}\left(\bigcap_{j=l+1}^m (E_n^j)^c \right) = 1 - \prod_{j=l+1}^m \mathrm{P}((E_n^j)^c) \\ &\geq 1 - \prod_{j=l+1}^m \left(1 - (1/12)(\log n)^{-4/5} \right) \\ &\geq 1 - \exp\left(-(1/12)(m-l)(\log n)^{-4/5} \right) \\ &\geq 1 - \exp\left(-(1/12)(\log n/(4 \log K) - 1)(\log n)^{-4/5} \right) \\ &> 1 - \exp\left(-(\log n)^{1/5}/(52 \log K) \right) =: 1 - p_n, \end{aligned}$$

where the last inequality holds for $n \geq n_1$ for an absolute constant $n_1 = 5^{104} < 5 \times 10^{72}$.

So with probability greater than $1 - p_n$, there exists some $j = l+1, l+2, \ldots, m$, such that $|S_{t+K^j}(\omega) - S_{t+K^{j-1}}(\omega)| \geq \sqrt{K^j \,\mathrm{LL}\, n}$. Then either $|S_{t+K^j} - S_t| \geq \sqrt{K^j \,\mathrm{LL}\, n}/2$ or $|S_{t+K^{j-1}} - S_t| \geq \sqrt{K^j \,\mathrm{LL}\, n}/2 \geq \sqrt{K^{j-1} \,\mathrm{LL}\, n}/2$. This implies that for $\omega \in E_n'(t)$, we have $t \in E_n(\omega)$; therefore, $\mathrm{P}(\omega\colon t \in E_n(\omega)) \geq \mathrm{P}(E_n'(t)) > 1 - p_n$ holds for $n \geq n_1$. □

Lemma 2.2 holds for individual t. It can be extended by a Fubini argument, as in Taylor's [18] proof of Theorem B, to a large portion of points in $\{0, 1, \ldots, n - \lfloor\sqrt{n}\rfloor\}$ simultaneously. Let U_n be the discrete uniform distribution on $\{0, 1, \ldots, n - \lfloor\sqrt{n}\rfloor\}$.

LEMMA 2.3. *For any $n \geq n_1$, $\mathrm{P}\{\omega\colon U_n(E_n(\omega)) > 1 - \sqrt{p_n}\} > 1 - \sqrt{p_n}$.*

PROOF. By Lemma 2.2, for each t, $\int_\Omega 1_{E_n(\omega)}(t)\, d\,\mathrm{P}(\omega) = \mathrm{P}(\omega\colon t \in E_n(\omega)) > 1 - p_n$. Then, interchanging an integral and a finite sum gives

$$\int_\Omega U_n(E_n(\omega))\, d\,\mathrm{P}(\omega) = \int \int_\Omega 1_{E_n(\omega)}(t)\, d\,\mathrm{P}(\omega)\, dU_n(t) > 1 - p_n.$$

If the lemma fails, a contradiction follows easily, so Lemma 2.3 is proved. □

The proof of Theorem 2.1 below will, again as in Taylor's [18] proof of Theorem B, use the following lemma of Vitali type:

LEMMA 2.4 (e.g., Lemma 7.2.2 in [5]). *Let $\lambda(A)$ denote the Lebesgue measure of a set A in R. Let $\mathscr{U}$ be a collection of open intervals in R with bounded union W. Then for any $t < \lambda(W)$, there is a finite, disjoint subcollection $\{V_1, V_2, \ldots, V_q\} \subset \mathscr{U}$, such that $\sum_{i=1}^q \lambda(V_i) \geq t/3$.*

THEOREM 2.1. *Let $\{X_i\}_{i\geq 1}$ be i.i.d. $N(0,1)$ random variables and $v_2(n) := v_2(\{X_i\}_{i=1}^n)$. Then for any $c \in (0, 1/12)$, $P(v_2(n) \geq cn \,\mathrm{LL}\, n) \to 1$ as $n \to \infty$.*

PROOF. By Lemma 2.3, we know that for $n \geq n_1$ with probability greater than $1 - \sqrt{p_n}$, for more than $M := M(n) := \lfloor(1 - \sqrt{p_n})(n - \lfloor\sqrt{n}\rfloor + 1)\rfloor$ integers in $\{0, 1, \ldots, n - \lfloor\sqrt{n}\rfloor\}$, say, $t_1, t_2, \ldots, t_M$, we can find corresponding integers $j(1), j(2), \ldots, j(M)$, such that $\{t_i, t_i + K^{j(i)}\}_{i=1}^M \subset \{1, 2, \ldots, n\}$ (for, $K^{j(i)} \leq K^m \leq \lfloor\sqrt{n}\rfloor$) and $|S_{t_i + K^{j(i)}} - S_{t_i}| \geq \sqrt{K^{j(i)} \,\mathrm{LL}\, n}/2$ for $i = 1, 2, \ldots, M$.

Consider the open intervals with integer endpoints t_i and $t_i + K^{j(i)}$ for $i = 1, 2, \ldots, M$, that is, the open intervals $\{(t_i, t_i + K^{j(i)})\}_{i=1}^M$. Note that $K^{j(i)} \geq K > 1$, $\{t_i, t_i + K^{j(i)}\}_{i=1}^M \subset \{0, 1, \ldots, n\}$ and for n large enough, $M = \lfloor(1 - \sqrt{p_n})(n - \lfloor\sqrt{n}\rfloor + 1)\rfloor > c_1 n$ for any fixed $c_1 \in (0,1)$. Therefore, the union of these open intervals is an open set in $(0, n]$ with Lebesgue measure b, where $b := b(n) \geq M > c_1 n$.

Applying Lemma 2.4, for $c_1 n < b$, yields a finite and disjoint subcollection of the above open intervals, say, $\{(t_i, t_i + K^{j(i)})\}_{i=1}^q$, such that $\sum_{i=1}^q K^{j(i)} \geq c_1 n/3$.

For the open disjoint intervals $\{(t_i, t_i + K^{j(i)})\}_{i=1}^q$, the closures are nonoverlapping except for possible intersections at endpoints. If we renumber the endpoints, we can get $0 \le t_1 < t_1 + K^{j(1)} \le t_2 < t_2 + K^{j(2)} \le \cdots \le t_q < t_q + K^{j(q)} \le n$. Thus, $v_2(n) \ge \sum_{i=1}^q (S_{t_i + K^{j(i)}} - S_{t_i})^2 \ge (1/4)\sum_{i=1}^q K^{j(i)} \operatorname{LL} n \ge (c_1/12) n \operatorname{LL} n$. Let $c := c_1/12$. Then $\mathrm{P}(v_2(n) \ge cn \operatorname{LL} n) \ge 1 - \sqrt{p_n} \to 1$ as $n \to \infty$. □

Theorem 2.1 gives a lower bound for the 2-variation of partial sums of i.i.d. $N(0,1)$ random variables X_i. Now we prove a more general result for i.i.d. random variables Y_i with finite $2+\delta$ moment for some $\delta > 0$.

THEOREM 2.2. *Let $\{Y_i\}_{i \ge 1}$ be i.i.d. with $\mathrm{E}Y_1 = 0$, $\mathrm{E}Y_1^2 = 1$ and $\mathrm{E}|Y_1|^{2+\delta} < \infty$ for some $\delta \in (0,1]$. Let $v_2^Y(n) := v_2(\{Y_i\}_{i=1}^n)$. Then for any $c \in (0, 1/12)$, $\mathrm{P}(v_2^Y(n) \ge cn \operatorname{LL} n) \to 1$ as $n \to \infty$.*

PROOF. Let $T_k := \sum_{i=1}^k Y_i$ for $k = 1, 2, \ldots, n$, $T_0 := 0$ and $L := \mathrm{E}|Y_1|^{2+\delta} < \infty$. Then by an extension of the Berry–Esséen theorem (e.g., [13], page 115), there exists an absolute constant $A > 0$, such that

$$\forall t, \qquad \left|\mathrm{P}\left(T_n/\sqrt{n} \le t\right) - \mathrm{P}\left(S_n/\sqrt{n} \le t\right)\right| \le A \cdot L/n^{\delta/2} =: L'/n^{\delta/2},$$

where $S_n := X_1 + X_2 + \cdots + X_n$ and the X_i's are i.i.d. $N(0,1)$ as in Theorem 2.1.

Let $P_Y^j := \mathrm{P}(T_{t+K^j} - T_{t+K^{j-1}} \le -\sqrt{K^j \operatorname{LL} n})$ and $P_X^j := \mathrm{P}(S_{t+K^j} - S_{t+K^{j-1}} \le -\sqrt{K^j \operatorname{LL} n})$ for $j = l+1, l+2, \ldots, m$, recalling $K := 25$ and for l, m defined as in (1). Then, $|P_Y^j - P_X^j| \le L'/(K^j - K^{j-1})^{\delta/2}$ for $j = l+1, l+2, \ldots, m$. Note that $K^j - K^{j-1} \ge K^{j-1} \ge K^l \ge n^{1/4}/K$ for $j = l+1, l+2, \ldots, m$. Therefore, $|P_Y^j - P_X^j| \le L' \cdot K^{\delta/2}/n^{\delta/8}$.

Recall the definition of E_n^j (as in the proof of Lemma 2.2) and (2) which says that $\mathrm{P}(E_n^j) \ge (\log n)^{-4/5}/12$ for $j = l+1,\ l+2,\ \ldots,\ m$. So $P_X^j = \mathrm{P}(E_n^j)/2 \ge (\log n)^{-4/5}/24 \gg L' \cdot K^{\delta/2}/n^{\delta/8}$ for n large enough. Then for large n, we have $P_Y^j > P_X^j/2 \ge (\log n)^{-4/5}/48$. Let $\widetilde{E}_n^j := \{\omega \colon |T_{t+K^j} - T_{t+K^{j-1}}| \ge \sqrt{K^j \operatorname{LL} n}\}$, $j = l+1, \ldots, m$. Then, $\mathrm{P}(\widetilde{E}_n^j) \ge P_Y^j \ge (\log n)^{-4/5}/48$. One can continue the proof just as in Theorem 2.1 since the rest of the proof only requires the random variables be i.i.d. and not standard normal. □

Next we will establish an upper bound for the 2-variation of partial sums.

THEOREM 2.3. *Let $\{X_i\}_{i \ge 1}$ be i.i.d. with mean 0 and variance 1. Then $v_2(n) := v_2(\{X_i\}_{i=1}^n) = O_p(n \operatorname{LL} n)$ as $n \to \infty$.*

PROOF. Let $\{x_t\}_{t \ge 0}$ be a Brownian motion defined on the same probability space as $\{X_i\}_{i \ge 1}$ such that they are independent. By taking a product space, we can assume such an $\{x_t\}_{t \ge 0}$ exists.

By Skorohod imbedding of sums (e.g., [1], Theorem 13.6), there is a sequence of i.i.d. nonnegative random variables, $\tau_1, \tau_2, \ldots$, defined on the same space, such that the process $\{x_{T(i)}\}_{i=1}^n$ and the process $\{S_i/\sqrt{n}\}_{i=1}^n$ have the same distribution, where $T(i) := \tau_1 + \tau_2 + \cdots + \tau_i$ and $S_i := X_1 + X_2 + \cdots + X_i$ for $i = 1, 2, \ldots, n$. Furthermore, $\mathsf{E}\tau_1 = \mathsf{E}(X_1/\sqrt{n})^2 = 1/n$. Therefore, for $T(1) < T(2) < \cdots < T(n)$, $\mathsf{E}T(i) = i\mathsf{E}\tau_1 = i/n$, $i = 1, 2, \ldots, n$. In particular $\mathsf{E}T(n) = 1$. Then for any given $\varepsilon \in (0, 1)$, if $M > 1/\varepsilon$, by Markov's inequality, for all n, $\mathsf{P}(E_n) \le 1/M < \varepsilon$, where $E_n := \{\omega\colon T(n) > M\}$. Let $\psi_1(0) := 0$ and

$$\psi_1(x) := \begin{cases} x^2/\log\log(1/x), & \text{if } 0 < x \le e^{-e}, \\ (x^2 + (e^e + e^{-e})x - 1)/2, & \text{if } x > e^{-e}. \end{cases}$$

Then $\psi_1(x)$ is a Young–Orlicz modulus satisfying the Δ_2 condition.

Now we will get an upper bound for $S := \sup_\pi \sum_{i=1}^k (x_{T(j_i)} - x_{T(j_{i-1})})^2$ on the set $\Omega \setminus E_n$, where the supremum is taken over all finite partitions π: $0 = j_0 < j_1 < \cdots < j_k = n$. Let

$$\mathsf{I} := \left\{ i \in \{1, 2, \ldots, k\}\colon |x_{T(j_i)} - x_{T(j_{i-1})}| \le 1/\sqrt{n} \right\}$$

and

$$\mathsf{II} := \left\{ i \in \{1, 2, \ldots, k\}\colon |x_{T(j_i)} - x_{T(j_{i-1})}| > 1/\sqrt{n} \right\}.$$

Then

$$\begin{aligned} S &= \sup_\pi \sum_{i \in \mathsf{I}} (x_{T(j_i)} - x_{T(j_{i-1})})^2 \\ &\quad + \sup_\pi \sum_{i \in \mathsf{II}} \psi_1(|x_{T(j_i)} - x_{T(j_{i-1})}|) \cdot \frac{(x_{T(j_i)} - x_{T(j_{i-1})})^2}{\psi_1(|x_{T(j_i)} - x_{T(j_{i-1})}|)} \\ &=: S_{\mathsf{I}} + S_{\mathsf{II}}. \end{aligned}$$

It is clear that $S_{\mathsf{I}} \le \sum_{i \in \mathsf{I}} 1/n \le k/n \le 1$.

For an upper bound for S_{II}, let

$$\phi_i := (x_{T(j_i)} - x_{T(j_{i-1})})^2 / \psi_1(|x_{T(j_i)} - x_{T(j_{i-1})}|), \qquad i \in \mathsf{II}.$$

Then

$$\phi_i = \begin{cases} \log\log \dfrac{1}{|x_{T(j_i)} - x_{T(j_{i-1})}|}, & \text{if } \dfrac{1}{\sqrt{n}} < |x_{T(j_i)} - x_{T(j_{i-1})}| \le e^{-e}; \\ \dfrac{2(x_{T(j_i)} - x_{T(j_{i-1})})^2}{(x_{T(j_i)} - x_{T(j_{i-1})})^2 + (e^e + e^{-e})|x_{T(j_i)} - x_{T(j_{i-1})}| - 1}, \\ \hfill \text{if } |x_{T(j_i)} - x_{T(j_{i-1})}| > e^{-e}. \end{cases}$$

Let

$$\phi_i^{(1)} := \phi_i 1\left\{1/\sqrt{n} < |x_{T(j_i)} - x_{T(j_{i-1})}| \le e^{-e}\right\}$$

and

$$\phi_i^{(2)} := \phi_i 1\left\{|x_{T(j_i)} - x_{T(j_{i-1})}| > e^{-e}\right\},$$

where $1\{\cdots\} := 1_{\{\cdots\}}$. Then for any $i \in \mathrm{II}$, $\phi_i^{(1)} \le \log\log\sqrt{n} \le \log\log n$. Also, since $B := (e^e + e^{-e})|x_{T(j_i)} - x_{T(j_{i-1})}| - 1 > 0$ if $\phi_i^{(2)} > 0$, then $\phi_i^{(2)} = 2A/(A+B) < 2$ for $A := (x_{T(j_i)} - x_{T(j_{i-1})})^2 > 0$ and $B > 0$. Therefore, for any $i \in \mathrm{II}$, $\phi_i \le \max(2, \log\log n) = \log\log n$ for n large enough. Then $S_{\mathrm{II}} \le \log\log n \sup_\pi \sum_{i\in \mathrm{II}} \psi_1(|x_{T(j_i)} - x_{T(j_{i-1})}|)$.

In $\Omega \setminus E_n$, $T(n) \le M$, so $\sup_\pi \sum_{i\in \mathrm{II}} \psi_1(|x_{T(j_i)} - x_{T(j_{i-1})}|) \le v_{\psi_1}(x_t, [0, M]) < \infty$ almost surely by Theorem B. Recall that $\mathrm{P}(\Omega \setminus E_n) \ge 1 - \varepsilon$ for all n, where $\varepsilon > 0$ is arbitrary. Letting $\varepsilon \to 0$, we get $S = S_{\mathrm{I}} + S_{\mathrm{II}} = O_p(\mathrm{LL}\, n)$, as $n \to \infty$. Since $\sum_{i=1}^k (S_{j_i} - S_{j_{i-1}})^2 =_{(\mathrm{d})} n \sum_{i=1}^k (x_{T(j_i)} - x_{T(j_{i-1})})^2$, it follows that $v_2(n) = O_p(n\, \mathrm{LL}\, n)$ as $n \to \infty$. $\square$

Combining Theorems 2.2 and 2.3 gives the exact order of the 2-variation of partial sums in probability up to some constant.

THEOREM 2.4. *Let $\{X_i\}_{i\ge 1}$ be i.i.d. with $\mathrm{E}X_1 = 0$, $\mathrm{E}X_1^2 = 1$ and $\mathrm{E}|X_1|^{2+\delta} < \infty$ for some $\delta > 0$. Then the 2-variation of partial sums of $\{X_i\}_{i=1}^n$ is of exact order $n\,\mathrm{LL}\, n$ in probability. That is, it is $O_p(n\,\mathrm{LL}\, n)$ and not $o_p(n\,\mathrm{LL}\, n)$.*

3. The p-variation of the empirical process.

Recall the definition of the p-variation of a function f in Section 1. We will treat the cases $p = 2$ and $p \in [1, 2)$ separately.

Let F be any nondecreasing function from an interval J into an interval L. Then for any real-valued function g on L, let $(g \circ F)(x) := g(F(x))$. It is easy to see that for $0 < p < \infty$, $v_p(g \circ F) \le v_p(g)$. If F is onto L, or onto its interior and g is continuous at the endpoints, then $v_p(g \circ F) = v_p(g)$.

Let F be any d.f. and let F_n be an empirical d.f. for it. Let U be the uniform $[0, 1]$ d.f. and let U_n be its empirical d.f. Then it is easy to see and known that we can write $F_n \equiv U_n \circ F$ and, of course, $F \equiv U \circ F$. Thus $F_n - F \equiv (U_n - U) \circ F$.

The lower bound for the 2-variation of partial sums will imply the following result for the 2-variation of the empirical process, which is sharp up to the constant c_0 since by Theorem A, $v_2(\alpha_n) = O_p(\mathrm{LL}\, n)$ as $n \to \infty$.

THEOREM 3.1. *Let F be any continuous d.f. on $\mathbb{R}$, let F_n be its empirical d.f. and let $\alpha_n := \sqrt{n}(F_n - F)$. Then $\mathrm{P}(v_2(\alpha_n) \ge c_0\, \mathrm{LL}\, n) \to 1$ as $n \to \infty$, where c_0 is any constant in $(0, 1/12)$.*

PROOF. We can assume as just noted that F is the $U[0, 1]$ d.f. U, since U_n is continuous at 0 and 1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. observations with distribution F. Since F is continuous, a.s. they are distinct, say, $0 =: X_{(0)} < X_{(1)} < \cdots < X_{(n)} < X_{(n+1)} := 1$. Denote the uniform spacings as $s_i := X_{(i)} - X_{(i-1)}$, $i = 1, 2, \ldots, n+1$. Let $\tilde{r}_i := (F_n - F)(X_{(i)}) - (F_n - F)(X_{(i-1)})$ for $i = 1, 2, \ldots, n+1$. Then $\tilde{r}_i = 1/n - s_i$ for $i = 1, 2, \ldots, n$ and $\tilde{r}_{n+1} = -s_{n+1}$. Recalling Definition 1.2, it is clear that

$$v_2(\alpha_n) \ge n\, v_2(\{\tilde{r}_i\}_{i=1}^{n+1}). \tag{3}$$

For the uniform spacings s_i, there is a well known representation theorem as follows:

LEMMA 3.1 (e.g., [16], page 335). *Let $\{\gamma_i\}_{i=1}^{n+1}$ be independent standard exponentially distributed random variables and $\Gamma_{n+1} := \sum_{i=1}^{n+1} \gamma_i$. Then $\{s_j\}_{j=1}^{n+1}$ is distributed as $\{\gamma_j/\Gamma_{n+1}\}_{j=1}^{n+1}$, and Γ_{n+1} is independent of $\{\gamma_j/\Gamma_{n+1}\}_{j=1}^{n+1}$.*

Lemma 3.1 allows $\tilde{r}_i = 1/n - s_i$ to be replaced by $r_i := 1/n - \gamma_i/\Gamma_{n+1}$ for $i = 1, 2, \dots, n$, $\tilde{r}_{n+1} = -s_{n+1}$ by $r_{n+1} := -\gamma_{n+1}/\Gamma_{n+1}$ and

$$v_2(\{r_i\}_{i=1}^{n+1}) =_{(\mathrm{d})} v_2(\{\tilde{r}_i\}_{i=1}^{n+1}). \tag{4}$$

Now, let $y_i = (1 - \gamma_i)/\Gamma_{n+1}$ for $1 \le i \le n+1$; $z_i = 1/n - 1/\Gamma_{n+1}$ for $1 \le i \le n+1$; $w_i = 0$ for $1 \le i \le n$; and $w_{n+1} = -1/n$. For any $u := \{u_i\}_{i=1}^{n+1} \in \mathrm{R}^{n+1}$, let $\|u\| := \|u\|_{(2)} = (v_2(\{u_i\}_{i=1}^{n+1}))^{1/2}$. Recall that the so defined $\|\cdot\|$ is a seminorm. As $r = y + z + w$, we have $|\|r\| - \|y\|| \le \|z\| + \|w\|$. It is clear that $\|w\| = 1/n$, $\|z\| = |(n+1)/n - (n+1)/\Gamma_{n+1}|$, thus

$$|\|r\| - \|y\|| = O_p(n^{-1/2}) \tag{5}$$

holds by the central limit theorem and the law of large numbers. On the other hand, $\Gamma_{n+1}\|y\| = (v_2(\{1-\gamma_i\}_{i=1}^{n+1}))^{1/2}$, and by Theorem 2.2, $\mathrm{P}(v_2(\{1-\gamma_i\}_{i=1}^{n+1}) \ge c(n+1)\mathrm{LL}(n+1)) \to 1$ as $n \to \infty$. Note that $\Gamma_{n+1}/(n+1) \to 1$ a.s., so $\mathrm{P}(\|y\| \ge c\sqrt{\mathrm{LL}(n+1)/(n+1)}) \to 1$ as $n \to \infty$. Combining this with (3), (4) and (5) completes the proof for the uniform case, and so the general case of Theorem 3.1. □

For $1 \le p < 2$ and any continuous distribution function F, by just considering the jumps of height $n^{-1/2}$ of α_n at n distinct points, clearly $v_p(\alpha_n) \ge n^{1-p/2}$. Next we will establish an upper bound on $\mathrm{E}v_p(\alpha_n)$ of the same order as this lower bound, which implies that the exact order of $v_p(\alpha_n)$ for $p \in [1, 2)$ is $n^{1-p/2}$. The proof will be based on Theorem C Bretagnolle of [2].

THEOREM 3.2. *Let F be any d.f. on R, let F_n be an empirical d.f. for F and $\alpha_n := n^{1/2}(F_n - F)$. Then for $1 \le p < 2$, $\mathrm{E}v_p(\alpha_n) \le C_p n^{1-p/2}$, where C_p is a constant depending only on p.*

PROOF. Again, we can assume F is the uniform $[0, 1]$ d.f. U. Let $X_1, X_2, \dots, X_n$ be the n observations on which F_n is based. Almost surely they are distinct. Let the order statistics be $0 =: X_{(0)} < X_{(1)} < \cdots < X_{(n)} < X_{(n+1)} := 1$.

Let $\alpha_n(t-) := \lim_{s\uparrow t} \alpha_n(s)$ for $0 < t \le 1$. Let $u_{2i} := X_{(i)}$ and $u_{2i-1} := X_{(i)}-$ for $i = 1, \dots, n$, $u_0 := 0$, $u_{2n+1} := 1$. It is easy to see that α_n are piecewise monotone functions. By properties of such functions, it can be shown that $v_p(\alpha_n) = v_p(\{\alpha_n(u_i) - \alpha_n(u_{i-1})\}_{i=1}^{2n+1})$ [15]. Let $s_i := X_{(i)} - X_{(i-1)}$, $i = 1, 2, \dots, n+1$, be the uniform spacings. Then $\alpha_n(u_{2i-1}) - \alpha_n(u_{2i-2}) = -\sqrt{n}s_i$ for $i = 1, 2, \dots, n+1$ and $\alpha_n(u_{2i}) - \alpha_n(u_{2i-1}) = 1/\sqrt{n}$ for $i = 1, 2, \dots, n$.

Let γ_i and Γ_{n+1} be the same as in Lemma 3.1 and let $\widetilde{y}_{2i-1} := -s_i =_{(\mathrm{d})} -\gamma_i/\Gamma_{n+1}$ for $i = 1, 2, \ldots, n+1$ and $\widetilde{y}_{2i} := 1/n$ for $i = 1, 2, \ldots, n$. Let $a := \{a_i\}_{i=1}^{2n+1}$ for any a_i. It follows that

$$v_p(\alpha_n) = n^{p/2} v_p(\widetilde{y}). \tag{6}$$

Let $G(n) := \{\Gamma_{n+1} \geq n/2\}$. By independence of $\{s_j\}_{j=1}^{n+1}$ and Γ_{n+1} (Lemma 3.1),

$$\mathsf{E} v_p(\widetilde{y}) = \mathsf{E}(v_p(\widetilde{y}) | G(n)) = \mathsf{E}(v_p(\widetilde{y}) 1_{G(n)})/\mathsf{P}(G(n)).$$

For all n, $\mathsf{P}(G(n)) \geq 1/4$ by Chebyshev's inequality, so $\mathsf{E} v_p(\widetilde{y}) \leq 4\mathsf{E}(v_p(\widetilde{y}) 1_{G(n)})$. Let $y_{2i-1} := -\gamma_i/\Gamma_{n+1}$ for $i = 1, 2, \ldots, n+1$, $y_{2i} := 1/\Gamma_{n+1}$ for $i = 1, 2, \ldots, n$ and let $x_{2i-1} := 0$ for $i = 1, 2, \ldots, n+1$, $x_{2i} := 1/n - 1/\Gamma_{n+1}$ for $i = 1, 2, \ldots, n$. For any $a, b \geq 0$ we have by Jensen's inequality $(a+b)^p \leq 2^{p-1}(a^p + b^p)$. As $\widetilde{y} = y + x$, we have $v_p(\widetilde{y}) \leq 2^{p-1}(v_p(y) + v_p(x))$. So

$$\mathsf{E} v_p(\widetilde{y}) \leq 4 \cdot 2^{p-1} \mathsf{E}[(v_p(y) + v_p(x)) 1_{G(n)}]. \tag{7}$$

For $n < 2p/(2-p)$, the conclusion holds since $v_p(\alpha_n)$ is uniformly bounded for such n, so assume $n \geq 2p/(2-p)$. It is clear that $v_p(x) = |1 - n/\Gamma_{n+1}|^p$, and by Hölder's inequality,

$$\begin{aligned} \mathsf{E}|1 - n/\Gamma_{n+1}|^p &\leq \left[\mathsf{E}(|\Gamma_{n+1} - n|)^{p\cdot(2/p)}\right]^{p/2} \cdot \left[\mathsf{E}(1/\Gamma_{n+1})^{p\cdot(2/(2-p))}\right]^{(2-p)/2} \\ &:= I_1 I_2. \end{aligned}$$

Since $\mathsf{E}(\Gamma_{n+1}) = n+1$ and $\mathrm{Var}(\Gamma_{n+1}) = n+1$, $I_1 = (n+2)^{p/2}$. For I_2, since Γ_{n+1} has the density $x^n e^{-x} 1_{x \geq 0}/n!$, we have

$$I_2 = \left[\mathsf{E}\left(\Gamma_{n+1}^{-(2p)/(2-p)}\right)\right]^{(2-p)/2} = \left[\int_0^\infty x^{n-(2p)/(2-p)} e^{-x}/n!\, dx\right]^{(2-p)/2} < \infty.$$

Using the asympotic approximations of the gamma function and $n!$ by Stirling's formulas, it follows that $I_2 = O(n^{-p})$ as $n \to \infty$. Thus

$$\mathsf{E}(v_p(x) 1_{G(n)}) \leq \mathsf{E} v_p(x) = \mathsf{E}|1 - n/\Gamma_{n+1}|^p = O(n^{-p/2}). \tag{8}$$

Next we will obtain an upper bound for $\mathsf{E}(v_p(y) 1_{G(n)})$. To do so, let $z_i := \Gamma_{n+1} y_i$ for $i = 1, 2, \ldots, 2n+1$.

CLAIM 3.1. *Let $w_i := z_{2i-1} + z_{2i} = 1 - \gamma_i$ for $i = 1, 2, \ldots, n$. Then $\mathsf{E}(v_p(\{w_i\}_{i=1}^n)) \leq \gamma_p \cdot n$, where γ_p is some positive constant depending only on p.*

PROOF. $\{w_i\}_{i=1}^n = \{1 - \gamma_i\}_{i=1}^n$ are i.i.d. with $\mathsf{E} w_i = 0$ and $\mathsf{E} w_i^2 < \infty$. Applying Theorem C to $\{w_i\}_{i=1}^n$, Claim 3.1 follows. □

The next claim is obvious, since each $z_{2i} = 1$ and $z_{2i-1} = -\gamma_i < 0$ a.s.

CLAIM 3.2. *Let $p > 1$, let $T_j := \sum_{i=1}^{j} z_i$ for $j \geq 1$, $T_0 := 0$, and let $0 = j_0 < j_1 < \cdots < j_k = 2n+1$ be such that $\sum_{i=1}^{k} |T_{j_i} - T_{j_{i-1}}|^p$ reaches the maximum $v_p(\{z_i\}_{i=1}^{2n+1})$. Let $\Delta_i := T_{j_i} - T_{j_{i-1}}$ for $i = 1, 2, \ldots, k$. Then a.s. Δ_i alternate in sign. Since $z_1, z_3, \ldots, z_{2n+1} < 0$ and $z_2, z_4, \ldots, z_{2n} > 0$, we have that k is an odd integer, $\Delta_1, \Delta_3, \ldots, \Delta_k < 0$ and $\Delta_2, \Delta_4, \ldots, \Delta_{k-1} > 0$. Furthermore, each Δ_i $(i = 1, 2, \ldots, k)$ is a sum of an odd number of consecutive z_i's beginning and ending with terms having the same sign as Δ_i itself.*

CLAIM 3.3. *Let $1 \leq p < 2$. Then $\mathrm{E}v_p(\{z_i\}_{i=1}^{2n+1}) = \mathrm{E}(\sum_{i=1}^{k} |T_{j_i} - T_{j_{i-1}}|^p) \leq \beta_p n$ for some constant $\beta_p > 0$.*

PROOF OF CLAIM 3.3. For $p = 1$, we have a.s. $j_i \equiv i$ and the claim holds. For $1 < p < 2$, by Claim 3.2, $v_p(\{z_i\}_{i=1}^{2n+1}) = \sum_{i=1}^{k} |\Delta_i|^p$, where Δ_i has the form

$$\Delta_i := \begin{cases} z_{2l_1(i)-1} + z_{2l_1(i)} + \cdots + z_{2l_2(i)-1}, & \text{for } i = 1, 3, \ldots, k, \\ z_{2m_1(i)} + z_{2m_1(i)+1} + \cdots + z_{2m_2(i)}, & \text{for } i = 2, 4, \ldots, k-1, \end{cases}$$

with $1 \leq l_1(i) \leq l_2(i) \leq n+1$ and $1 \leq m_1(i) \leq m_2(i) \leq n$ for each i. Thus

$$\Delta_i = \begin{cases} \displaystyle\sum_{j=l_1(i)}^{l_2(i)-1} w_j + z_{2l_2(i)-1}, & \text{for } i = 1, 3, \ldots, k, \\[2ex] \displaystyle z_{2m_1(i)} + \sum_{j=m_1(i)+1}^{m_2(i)} w_j, & \text{for } i = 2, 4, \ldots, k-1. \end{cases}$$

Using the fact that $|a+b|^p \leq 2^{p-1}(|a|^p + |b|^p)$, it follows that

$$|\Delta_i|^p \leq \begin{cases} \displaystyle 2^{p-1}\left(\left|\sum_{j=l_1(i)}^{l_2(i)-1} w_j\right|^p + |z_{2l_2(i)-1}|^p\right), & \text{for } i = 1, 3, \ldots, k, \\[2ex] \displaystyle 2^{p-1}\left(|z_{2m_1(i)}|^p + \left|\sum_{j=m_1(i)+1}^{m_2(i)} w_j\right|^p\right), & \text{for } i = 2, 4, \ldots, k-1. \end{cases}$$

Noting that each Δ_i $(i = 1, 2, \ldots, k)$ is a sum of disjoint consecutive z_j's, we have

$$v_p(\{z_i\}_{i=1}^{2n+1}) \leq 2^{p-1}(\Sigma_1 + \Sigma_2),$$

where Σ_1 is part of a p-variation sum of $\{w_1, w_2, \ldots, w_n\}$, and Σ_2 is a sum of terms $|z_j|^p$ for k distinct values of j. Thus, since $\mathrm{E}|z_1|^p \geq 1 = \mathrm{E}|z_2|^p$ and likewise for other z_i, i odd or even,

$$\begin{aligned} \mathrm{E}v_p(\{z_i\}_{i=1}^{2n+1}) &\leq 2^{p-1}(\mathrm{E}\Sigma_1 + \mathrm{E}\Sigma_2) \leq 2^{p-1}[\mathrm{E}v_p(\{w_i\}_{i=1}^{n}) + (2n+1)\mathrm{E}|z_1|^p] \\ &\leq 2^{p-1}[\gamma_p n + (2n+1)\mathrm{E}|z_1|^p] \leq \beta_p n \end{aligned}$$

for some constant $\beta_p < \infty$, thereby establishing Claim 3.3. □

Recall that $\mathrm{E}v_p(\{x_i\}_{i=1}^{2n+1}) = O(n^{-p/2}) = o(n^{1-p})$ and that $z_i = y_i\Gamma_{n+1}$. From the definitions,

$$\mathrm{E}(v_p(y)1_{G(n)}) = \mathrm{E}[v_p(z/\Gamma_{n+1})1_{G(n)}] \leq (2/n)^p \mathrm{E}v_p(z) = O(n^{1-p}).$$

Now Theorem 3.2 follows from (6), (7) and (8). □

4. The almost sure asymptotic behavior of $\|F_n - F\|_{[p]}$.

In the previous sections, we investigated the p-variation for partial sum processes and empirical processes and obtained bounds in expectation ($p < 2$) or in probability ($p \leq 2$) for them. In this section, we will look at the almost sure asymptotic behavior of the p-variation norm $\|F_n - F\|_{[p]}$. In other words, we will find numbers a_n such that $0 < \limsup_{n\to\infty} \|F_n - F\|_{[p]}/a_n < \infty$ almost surely.

For $p = 1$, clearly $\|F_n - F\|_{(1)} \equiv 2$ for any continuous F. For $p = \infty$, the p-variation seminorm of $F_n - F$, defined as $(\sup - \inf)(F_n - F)$, is within a factor of 2 of the sup norm, for which there is the Smirnov–Chung LIL (law of the iterated logarithm) as follows:

THEOREM D (Smirnov [17] and Chung [3]). *For $p = \infty$,*

$$\limsup_{n\to\infty} \sqrt{2n}\,\|F_n - F\|_\infty/\sqrt{\mathrm{LL}\,n} = 1$$

almost surely.

THEOREM 4.1. *For $2 < p < \infty$, $1 \leq \limsup_{n\to\infty} \sqrt{2n}\,\|F_n - F\|_{[p]}/\sqrt{\mathrm{LL}\,n} < \infty$ almost surely.*

PROOF. From [4], $\sqrt{n}\|F_n - F\|_{(p)}$ is bounded in probability uniformly in n for $p > 2$. It follows that the lim sup in the statement is finite a.s. (Theorem 4.1 in [8]). The lim sup is at least 1 by considering $F_n(t) - F(t)$ at $t = 1/2$ and the one-dimensional log log law. □

We see that for $p = 1$ and $p > 2$, actually the almost sure asymptotic order of magnitude of the p-variation norm $\|F_n - F\|_{[p]}$ is either known or follows directly from known results. Next, we will look at the case where p is between 1 and 2.

THEOREM 4.2. *For $1 < p < 2$ and some constant $\lambda_p < \infty$, if F is any continuous d.f. on* R *and Y_n are any random variables having the distribution of $\|F_n - F\|_{[p]}$ for each n but an arbitrary joint distribution,*

$$1 \leq \inf_{n\geq 1} \frac{Y_n}{n^{(1/p)-1}} \leq \limsup_{n\to\infty} \frac{Y_n}{n^{(1/p)-1}} \leq \lambda_p \quad \text{a.s.}$$

PROOF. As discussed in Section 3, we can assume that F is the uniform $[0,1]$ d.f. U. Let $X_1, X_2, \dots, X_n$ be the n observations on which F_n is based.

Let $G_i(t) := 1_{\{X_i \le t\}} - t$ for $0 \le t \le 1$ and $i = 1, 2, \dots, n$. Then $(F_n - F)(t) = n^{-1}\sum_{i=1}^n G_i(t)$. It is obvious that $\|G_i\|_{[p]} \le 3$ for $i = 1, 2, \dots, n$. So $\mathsf{E}\|G_i\|_{[p]}^2 \le 9$ for $i = 1, 2, \dots, n$. By Theorem 3.2, for $1 < p < 2$, $\mathsf{E}v_p(\alpha_n) \le C_p n^{1-p/2}$ for n large. Then

$$\begin{aligned}\mathsf{E}\left\|\sum_{i=1}^n G_i\right\|_{[p]} &= \sqrt{n}\,\mathsf{E}\|\alpha_n\|_{[p]} = \sqrt{n}\,\mathsf{E}\big((v_p(\alpha_n))^{1/p} + \|\alpha_n\|_\infty\big)\\ &\le \sqrt{n}\,\mathsf{E}(2(v_p(\alpha_n))^{1/p}) \le 2\sqrt{n}\,(\mathsf{E}v_p(\alpha_n))^{1/p}\\ &\le 2\sqrt{n}\,C_p^{1/p} n^{(1/p)-1/2} = \lambda_1 n^{1/p} \qquad (\lambda_1 := 2C_p^{1/p}),\end{aligned}$$

where the second inequality follows from Jensen's inequality and the third from Theorem 3.2.

Now, p-variation norms can be evaluated over rational arguments and so are measurable. So, we can apply Lemma 2.1 of Kuelbs [10] (cf. also Lemma 2.6 in [8]). Let $\tau_n := 9n \ge \sum_{i=1}^n \mathsf{E}\|G_i\|_{[p]}^2$ and $M := 3 \ge \|G_i\|_{[p]}$, $i = 1, 2, \dots, n$. Then for any $0 \le \gamma \le 1/(2M) = 1/6$ and for any $\Lambda > 0$, we have

$$\begin{aligned}\mathsf{P}\big(\|F_n - F\|_{[p]}/n^{(1/p)-1} > \Lambda\big) &= \mathsf{P}\left(\left\|\sum_{i=1}^n G_i\right\|_{[p]} > \Lambda n^{1/p}\right)\\ &\le \exp\left(3\gamma^2\tau_n - \gamma\left(\Lambda n^{1/p} - \mathsf{E}\left\|\sum_{i=1}^n G_i\right\|_{[p]}\right)\right)\\ &\le \exp\big(27\gamma^2 n - \gamma(\Lambda n^{1/p} - \lambda_1 n^{1/p})\big).\end{aligned}$$

Take $\gamma := n^{(1/p)-1} \le 1/6$ for n large. Then

$$\begin{aligned}&\mathsf{P}\big(\|F_n - F\|_{[p]}/n^{(1/p)-1} > \Lambda\big)\\ &\quad\le \exp\big(27n^{(2/p)-1} - (\Lambda - \lambda_1)n^{(2/p)-1}\big)\\ &\quad= \exp\big(-(\Lambda - \lambda_p)n^{(2/p)-1}\big) \quad \text{where } \lambda_p := \lambda_1 + 27.\end{aligned}$$

By the Borel–Cantelli lemma, since for $\Lambda > \lambda_p$, $\sum_{n=1}^\infty \exp(-(\Lambda - \lambda_p)n^{(2/p)-1})$ converges, we have $\limsup_{n\to\infty} \|F_n - F\|_{[p]}/n^{(1/p)-1} \le \Lambda$ almost surely for any $\Lambda > \lambda_p$, and hence for $\Lambda = \lambda_p$.

On the other hand, by only considering the jumps of $F_n - F$ at each X_i, we have $\|F_n - F\|_{[p]} \ge n^{(1/p)-1}$ almost surely since F is continuous. The proof holds for any Y_n with the same distribution as $\|F_n - F\|_{[p]}$ for each n, so Theorem 4.2 follows. $\square$

REMARK. For $p = 2$, the almost sure asymptotic order of $\|F_n - F\|_{[p]}$ remains an open problem.

Acknowledgments. This paper is based on my Ph.D. dissertation written under the supervision of Professor Richard M. Dudley. I thank him for his invaluable guidance during the research process and for his many helpful suggestions during the writing of this paper. I also thank the referee for very helpful suggestions on shortening proofs, and Professor Jon A. Wellner and Professor Ronald Pyke for a discussion about spacings via Professor Dudley.

REFERENCES

[1] BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, MA.
[2] BRETAGNOLLE, J. (1972). p-variation de fonctions aléatoires. II. Processus à accroissements indépendants. *Lecture Notes in Math.* 258 64–71. Springer, Berlin.
[3] CHUNG, K. L. (1949). An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.* 67 36–50.
[4] DUDLEY, R. M. (1992). Fréchet differentiability, p-variation and uniform Donsker classes. *Ann. Probab.* 20 1968–1982.
[5] DUDLEY, R. M. (1993). *Real Analysis and Probability*, 2nd ed. Chapman and Hall, New York.
[6] DUDLEY, R. M. (1994). The order of the remainder in derivatives of composition and inverse operators for p-variation norms. *Ann. Statist.* 22 1–20.
[7] DUDLEY, R. M. (1997). Empirical processes and p-variation. In *Festschrift for Lucien Le Cam* (D. Pollard, E. Torgersen and G. L. Yang, eds.) 219–233. Springer, New York.
[8] DUDLEY, R. M. and PHILIPP, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. Verw. Gebiete* 62 509–552.
[9] HUANG, Y.-C. (1994). Speed of convergence of classical empirical processes in p-variation-norm. Ph.D. dissertation, MIT.
[10] KUELBS, J. (1977). Kolmogorov's law of the iterated logarithm for Banach space valued random variables. *Illinois J. Math.* 21 784–800.
[11] LÉPINGLE, D. (1976). La variation d'ordre p des semi-martingales. *Z. Wahrsch. Verw. Gebiete* 36 295–316.
[12] LOVE, E. R. and YOUNG, L. C. (1937). Sur une classe de fonctionelles linéaires. *Fund. Math* 28 243–257.
[13] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York.
[14] PISIER, G. and XU, Q. (1988). The strong p-variation of martingales and orthogonal series. *Probab. Theory Related Fields* 77 497–514.
[15] QIAN, J. (1997). The p-variation of partial sum processes and empirical processes. Ph.D. dissertation, Tufts Univ.
[16] SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.
[17] SMIRNOV, N. V. (1944). Approximate laws of distribution of random variables from empirical data. *Uspekhi Mat. Nauk* 10 179–206 (in Russian).
[18] TAYLOR, S. J. (1972). Exact asymptotic estimates of Brownian path variation. *Duke Math. J.* 39 219–241.
[19] WIENER, N. (1924). The quadratic variation of a function and its Fourier coefficients. *J. Math. Phys.* 3 72–94.
[20] YOUNG, L. C. (1936). An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* 67 251–282.

DEPARTMENT OF MATHEMATICS
BUCKNELL UNIVERSITY
LEWISBURG, PENNSYLVANIA 17837
E-MAIL: jqian@bucknell.edu