Partial sum processes in p-variation norm

Rimas Norvaiša (Vilnius)

Sandbjerg, 2010 January 27

For a function $f\colon [0,1] \to R$ and a number $0 < p < \infty$, the p-*variation* of f is

$$v_p(f) := \sup \left\{ \sum_{i=1}^{m} |f(t_j) - f(t_{j-1})|^p \right\} \le +\infty,$$

where the supremum is taken over all partitions $0 = t_0 < t_1 < \cdots < t_m = 1$, $m = 1, 2, \ldots$, of the interval $[0,1]$.

If $v_p(f) < +\infty$ then we say that f has finite p-variation and $\mathcal{W}_p[0,1]$ is the set of all such functions. This set is a *Banach* space with the norm

$$\|f\|_{[p]} := \|f\|_{\sup} + v_p(f)^{1/p}.$$

For a comparison with the α-*Hölder*, $\alpha \in (0,1]$, property of f, if $p := 1/\alpha$, then

$$\sum_{i=1}^{m} |f(t_j) - f(t_{j-1})|^p \le C^p \sum_{j=1}^{m} (t_j - t_{j-1}) = C^p$$

and so $v_p(f) \le C^p < +\infty$. But note that a finite p-variation property can have discontinuous functions, such as sample functions of stable processes.

Let $X_1, X_2, \ldots$ be real random variables. For each $n = 1, 2, \ldots$, let S_n be the n-th partial sum process

$$S_n(t) := X_1 + \cdots + X_{\lfloor tn \rfloor}, \quad t \in [0, 1],$$

Thus for each $n = 1, 2, \ldots$ and $t \in [0, 1]$,

$$S_n(t) = \begin{cases} 0, & \text{if } t \in [0, 1/n), \\ X_1 + \cdots + X_k, & \text{if } t \in [\frac{k}{n}, \frac{k+1}{n}), \\ & k \in \{1, \ldots, n-1\}, \\ X_1 + \cdots + X_n, & \text{if } t = 1. \end{cases}$$

Then for any $p \in (0, \infty)$,

$$v_p(S_n) = \max \left\{ \sum_{j=1}^{m} |X_{k_{j-1}+1} + \cdots + X_{k_j}|^p \right\},$$

where the maximum is taken over $0 = k_0 < \cdots < k_m = n$ and $1 \leq m \leq n$.

J. Bretagnolle (1972): given $p \in (0, 2)$ there exists a finite constant C_p such that

$$\left(\sum_{i=1}^{n} E|X_i|^p \leq \right) Ev_p(S_n) \leq C_p \sum_{i=1}^{n} E|X_i|^p,$$

provided $X_1, X_2, \ldots$ are independent, $E|X_i|^p < \infty$ and $EX_i = 0$ if $p > 1$.

Suppose that $X_1, X_2, \ldots$ are independent identically distributed real random variables, $EX_1 = 0$ and $EX_1^2 = 1$. Let $Lx := \max\{1, \log x\}$, $x > 0$.

J. Qian (1998): $v_2(S_n) = O_P(nLLn)$ as $n \to \infty$. Also, $O_P(nLLn)$ cannot be replaced by $o_p(nLLn)$, if in addition $E|X_1|^{2+\epsilon} < \infty$ for some $\epsilon > 0$.

Let W be a standard *Wiener* process on the interval $[0, 1]$. Due to results of *N. Wiener* (1923) and *P. Lévy* (1940):

$$v_p(W) < +\infty \quad \text{almost surely iff } p > 2,$$

and $v_2(W) = +\infty$ almost surely.

More precise information can be obtained in terms of ϕ-variation, defined as p-variation except that the power function $x \mapsto x^p$, $x \geq 0$, is replaced by a function ϕ.

S. J. Taylor (1972): $v_{\psi_1}(W) < +\infty$ a. s., where

$$\psi_1(x) := x^2/LL(1/x), \quad 0 < x \leq e^{-e}.$$

Also, $v_\psi(W) = +\infty$ a. s., for any ψ such, that $\psi_1(x) = o(\psi(x))$ as $x \downarrow 0$.

Given a sequence $X_1, X_2, \ldots$ of i.i.d. real random variables having a d.f. F, let F_n be the empirical d.f. based on $X_1, \ldots, X_n$.

R.M. Dudley (1992): Let $2 < p < \infty$ and let F be a uniform d.f. The convergence in law

$$\sqrt{n}(F_n - F) \Rightarrow B \quad \text{in } \mathcal{W}_p[0,1],$$

as $n \to \infty$ holds, where B is a Brownian bridge.

Y.-Ch. Huang and R.M. Dudley (2001): For $2 < p < \infty$ there is a finite constant C_p such that if F is any d.f. on R, then on some probability space there exist $X_1, X_2, \ldots$ i.i.d. r.v.'s with d.f. F and Brownian bridges B_n such that for all n,

$$E\|\sqrt{n}(F_n - F) - B_n \circ F\|_{[p]} \le C_p n^{(2-p)/(2p)},$$

and the order of bound is best possible in general.

J. Qian (1998): Let $1 < p < 2$. There exists a finite constant c such, that a. s.

$$1 \le \liminf_{n\to\infty} \frac{\|F_n - F\|_{[p]}}{n^{1/p-1}} \le \limsup_{n\to\infty} \frac{\|F_n - F\|_{[p]}}{n^{1/p-1}} \le c.$$

These are the main facts in the present context known up till recently.

R. Norvaiša and *A. Račkauskas*: Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables and let S_n be the n-th partial sum process. The convergence in law

$$n^{-1/2} S_n \Rightarrow \sigma W \quad \text{in } \mathcal{W}_p[0,1],$$

as $n \to \infty$ holds if and only if $EX_1 = 0$ and $\sigma^2 := EX_1^2 < \infty$.

It is interesting to compare this fact with the related convergence of smoothed partial sum processes with respect to the α-Hölder norm. Let $\tilde{S}_n$ be a (random) function obtained from S_n by linear interpolation between points

$$\left(\frac{k}{n}, S_n\left(\frac{k}{n}\right)\right) \quad \text{ir} \quad \left(\frac{k+1}{n}, S_n\left(\frac{k+1}{n}\right)\right)$$

$k = 0, 1, \ldots, n-1$.

A. Račkauskas and *C. Suquet* (2004): Let $p > 2$. Convergence in law

$$n^{-1/2} \tilde{S}_n \Rightarrow \sigma W \quad \text{in } \mathcal{H}^0_{1/p}[0,1],$$

as $n \to \infty$ holds if and only if $EX_1 = 0$ and

$$\lim_{t \to \infty} t \Pr(\{|X_1| > t^{1/2-1/p}\}) = 0.$$

Sketch of proof. The proof rests on some results concerning a problem of representing linear bounded functionals on the Banach space $\mathcal{W}_q[0,1]$. The following fact (with a different constant) is due to *Love, E. R.* and *Young L. C.* (1937):

Theorem: Let $1 < q < \infty$, $1/p + 1/q = 1$, let $L\colon \mathcal{W}_q[0,1] \to R$ be a linear bounded functional and let $F(t) := L(1_{[0,t]})$, $t \in [0,1]$. Then

$$\|F\|_{[p]} \leq 4 \sup\{|L(f)|\colon f \in \mathcal{F}_q\} = 4\|L\|_{\mathcal{F}_q},$$

where $\mathcal{F}_q := \{f \in \mathcal{W}_q[0,1]\colon \|f\|_{[q]} \leq 1\}$.

To use this fact we represent the n-th partial sum process as follows

$$n^{-1/2} S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i 1_{[0,t]}(i/n) = \nu_n(1_{[0,t]}),$$

where for any function $f\colon [0,1] \to R$

$$\nu_n(f) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i f(i/n).$$

Then we use a theory of stochastic processes indexed by functions.

For each $f \in \mathcal{L}^2([0,1], \lambda)$, let

$$\nu(f) := \int_0^1 f\, dW,$$

where the integral is defined in the Itô sense. Then ν is the isonormal *Gauss*ian process in the *Hilbert* space $L^2([0,1], \lambda)$. When $f \in \mathcal{W}_q[0,1]$ and $q < 2$, then $\nu(f)$ exists as the *Riemann-Stieltjes* integral and

$$Var(\nu_n(f)) = \frac{\sigma^2}{n} \sum_{i=1}^{n} f^2(\frac{i}{n}) \to \sigma^2 \int_0^1 f^2\, d\lambda =: \sigma_f^2$$

By *Lindeberg*'s CLT, $\mathcal{L}(\nu_n(f)) \to N(0, \sigma_f)$ as $n \to \infty$ for each (fixed) $f \in \mathcal{W}_q[0,1]$.

In fact we need a convergence which is uniform over a class of functions

$$\mathcal{F}_q = \{f \in \mathcal{W}_q[0,1] \colon \|f\|_{[q]} \le 1\},$$

which is the unit ball in the Banach space $\mathcal{W}_q[0,1]$.

Convergence in law (basic facts).

Let M be a metric space and let $\mathcal{B}$ be a σ-algebra of its Borel sets. Let $(\Omega, \mathcal{A}, P)$ be a probability space. If a function $X\colon \Omega \to M$ is $\mathcal{A} - \mathcal{B}$ measurable, then it is called a random variable (r.v.). A law of a r.v. X is a measure $\mathcal{L}(X)$ on $\mathcal{B}$ with values

$$\mathcal{L}(X)(B) = P(\{\omega \in \Omega\colon X(\omega) \in B\}), \quad B \in \mathcal{B}.$$

A sequence $\mathcal{L}(Z_n)$ of laws on a metric space converges (weakly) to a law $\mathcal{L}(Z)$, if for each $h \in C_b(M)$,

$$Eh(Z_n) = \int_M h\, d\mathcal{L}(Z_n) \to \int_M h\, d\mathcal{L}(Z) = Eh(Z).$$

The problem comes from the fact that many interesting functions are not r.v.'s. This is the case when M is not separable metric space since the Borel σ-algebra $\mathcal{B}$ in such a space is too big to carry a σ-additive measure.

The *Banach* space $(\mathcal{W}_p[0,1], \|\cdot\|_{[p]})$, as well as the *Banach* space $(\ell_\infty(\mathcal{F}), \|\cdot\|_{\mathcal{F}})$ are not seprable. The set of indicator functions $\{1_{[0,t]}\colon t \in [0,1]\}$ is not countable and not dense.

The non-separability problem is solved by using the following extension of the classical weak convergence notion.

Hoffmann-Jørgensen (1984): Let M be a metric space, $(\Omega, \mathcal{A}, P)$ be a probability space, let Z_n: $\Omega \to M$, $n = 1, 2, \ldots$, be functions and let Z be a function from Ω to a separable subspace of M which is Borel measurable. It is said that Z_n converge in law to Z, written as $Z_n \Rightarrow Z$ in M, if for each $h \in C_b(M)$,

$$E^*h(Z_n) = \int_\Omega^* h \circ Z_n \, dP \to \int_\Omega h \circ Z \, dP = Eh(Z),$$

where $E^*T := \inf\{EU\}$ is the upper integral.

(So the laws need not exist to converge in law, except for the limit function).

Theorem: If a metric space M is separable and Z_n: $\Omega \to M$ are r.v.'s then convergence $Z_n \Rightarrow Z$ in M is equivalent to the usual weak convergence of laws $\mathcal{L}(Z_n) \to \mathcal{L}(Z)$.

We are interested in $\nu_n \Rightarrow \nu$ in $\ell_\infty(\mathcal{F}_q)$ with $q < 2$, here

$$\nu_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i f(\frac{i}{n}), \quad \nu(f) = \int_0^1 f\, dW$$

and $f \in \mathcal{F}_q = \{f \in \mathcal{W}_q[0,1] \colon \|f\|_{[q]} \leq 1\}$.

The first question is when does the limit law exists? Let Q be a probability on $[0,1]$. For each $f, g \in \mathcal{F} \subset \mathcal{L}_2([0,1], Q)$, let

$$\rho_{2,Q}(f,g) := (\int_{[0,1]} [f-g]^2\, dQ)^{1/2}.$$

Then $\rho_{2,Q}$ is the pseudometric on $\mathcal{F}$. If λ is the Lebesgue measure on $[0,1]$, then let $\rho_2 := \rho_{2,\lambda}$. Let $UC(\mathcal{F})$ be the set of functions $h \colon \mathcal{F} \to R$, which are uniformly continuous w.r.t. ρ_2. Then $UC(\mathcal{F})$ is the separable subspace of $\ell_\infty(\mathcal{F})$ with $\|\cdot\|_{\mathcal{F}}$.

Dudley (1973): Let $\mathcal{F} \subset \mathcal{L}_2([0,1], \lambda)$. There exists a version of $\nu = \{\nu(f) \colon f \in \mathcal{F}\}$ such, that $\nu \colon \Omega \to UC(\mathcal{F})$, provided

$$\int_0^1 \sqrt{\log N(\epsilon, \mathcal{F}, \rho_2)}\, d\epsilon < \infty,$$

where $N(\epsilon, \mathcal{F}, \rho_2)$ is the minimal number of balls of radius ϵ needed to cover $\mathcal{F}$.

For each n, let $Z_{n1}, \ldots, Z_{nn}$ be independent stochastic processes indexed by a class $\mathcal{F}$ and defined on a probability space $(\Omega_n, \mathcal{A}_n, P_n)$. Next are conditions for

$$Z_n := \sum_{i=1}^{n} (Z_{ni} - EZ_{ni}) \Rightarrow Z \quad \text{in } \ell_\infty(\mathcal{F}). \qquad (1)$$

The following is from the book of *Van der Vaart* and *Wellner* (1996).

Theorem: For each n, let $\{Z_{ni}\colon 1 \le i \le n\}$ be independent stochastic processes indexed by a totally bounded semimetric space $(\mathcal{F}, \rho)$. Assume that the sums Z_n are „properly measurable" and that

$$\lim_{n\to\infty} \sum_{i=1}^{n} E^* \|Z_{ni}\|_{\mathcal{F}}^2 1_{\{\|Z_{ni}\|_{\mathcal{F}} > \eta\}} = 0, \quad \forall \eta > 0,$$

$$\lim_{n\to\infty} \sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{n} E[Z_{ni}(f) - Z_{ni}(g)]^2 = 0, \quad \forall \delta_n \downarrow 0,$$

$$P_n^* - \lim_{n\to\infty} \int_0^{\delta_n} \sqrt{\log N(\epsilon, \mathcal{F}, d_n)}\, d\epsilon = 0, \quad \forall \delta_n \downarrow 0,$$

and the sequence of covariance functions of $Z_n - EZ_n$ converge pointwise on $\mathcal{F} \times \mathcal{F}$ to the covariance of Z. Then (1) holds true.

Let $F_{\mathcal{F}}$ be a function with values

$$F_{\mathcal{F}}(x) := \sup\{|f(x)|: f \in \mathcal{F}\}, \quad x \in [0,1].$$

If $F_{\mathcal{F}}$ is measurable, then it is called the envelope function of $\mathcal{F}$.

Theorem: Let $X_1, X_2, \ldots$ be i.i.d. real r.v.'s with $EX_1 = 0$ and $EX_1^2 = \sigma^2 < \infty$. Let $1 \leq q < 2$ and $\mathcal{F} \subset \mathcal{W}_q[0,1]$ be "image admissible Suslin", $\|F_{\mathcal{F}}\|_{\sup} < \infty$ and

$$\int_0^1 \sup_{Q \in \mathcal{Q}} \sqrt{\log N(\epsilon, \mathcal{F}, \rho_{2,Q})}\, d\epsilon < \infty, \tag{2}$$

where $\mathcal{Q}$ is the set of all probability measures on $[0,1]$. Then

$$\nu_n \Rightarrow \nu \quad \text{in} \quad \ell_\infty(\mathcal{F}).$$

Dudley: Let $\mathcal{F}_q = \{f \in \mathcal{W}_q[0,1]: \|f\|_{[q]} \leq 1\}$ with $1 \leq q < 2$. Then $\mathcal{F}_q$ is "image admissible Suslin", the envelope function $F_{\mathcal{F}_q} \equiv 1$ and (2) holds for $\mathcal{F} = \mathcal{F}_q$.

Application of the result.

Consider the model of nonlinear regression:

$$y_i = \beta f(i/n) + \epsilon_i, \quad i = 1, \ldots, n,$$

where ϵ_i are i.i.d. r.v.'s $E\epsilon_1 = 0$ and $E\epsilon_1^2 = 1$. The function f: $[0,1] \to R$ is known, while the coefficient β is estimated by $\widehat{\beta}_n$ obtained by least square method. Then r.v.'s

$$\widehat{\epsilon}_i := y_i - \widehat{\beta}_n f(i/n), \quad i = 1, \ldots, n.$$

are called residuals. Let $\widehat{S}_n(t) := \widehat{\epsilon}_1 + \ldots + \widehat{\epsilon}_{\lfloor tn \rfloor}$, $n = 1, 2, \ldots$, and $t \in [0,1]$.

Teorema: Let $p > 2$ and $q \geq 1$ be such that $1/p + 1/q > 1$ and let $f \in \mathcal{W}_q[0,1]$ be continuous. Then

$$n^{-1/2}\widehat{S}_n \Rightarrow W - g\int_0^1 \widetilde{f}\, dW \quad \text{in } \mathcal{W}_p[0,1],$$

where $g(t) := \int_0^t \widetilde{f}(s)\, ds$ and $\widetilde{f} := f/\|f\|_{L_2}$.

For another application consider a problem of estimating a change in the mean

$$X_{ni} := a_{nj} + \epsilon_i, \quad \begin{cases} i \in (\tau^*_{j-1} n, \tau^*_j n], \\ j = 1, \ldots, m, \end{cases}$$

where $0 = \tau^*_0 < \tau^*_1 < \cdots < \tau^*_m = 1$, ϵ_i are i.i.d. r.v.'s $E\epsilon_1 = 0$, $E\epsilon_1^2 = 1$ and $a_{n1}, \ldots, a_{nm}$ are real numbers. Assume that m and $\tau^*_1, \ldots, \tau^*_{m-1}$ are not known. We would like to separate the null hypothesis

$$H_0 : \quad m = 1$$

from its alternative

$$H_A : \quad 1 < m \leq n.$$

For this aim we consider the functional

$$T_{p,n} := \max \left\{ \sum_{j=1}^{m} |Y_{n,k_j} - Y_{n,k_{j-1}}|^p \right\},$$

here the maximum is taken over $0 = k_0 < \cdots < k_m = n$, $1 \leq m \leq n$, $p > 0$, and

$$Y_{n,k} := \sum_{i=1}^{k} X_{ni} - \frac{k}{n} \sum_{i=1}^{n} X_{ni}.$$

To verify the null hypothesis we can use the fact

Theorem: Let $X_{ni} = a_n + \epsilon_i$ for each $i = 1, \ldots, n$ and $n \in N$ (i.e. $m = 1$, no change). If $p > 2$, then

$$\mathcal{L}(n^{-p/2} T_{n,p}) \to \mathcal{L}(v_p(B)),$$

as $n \to \infty$; here $B(t) = W(t) - tW(1)$, $t \in [0, 1]$.

To verify the alternative let $0 = k_0 < k_1 < \cdots < k_m = n$, $1 \leq m \leq n$, $\Delta\tau^*_{nj} := (k_j - k_{j-1})/n$, $j = 0, 1, \ldots, m$, and

$$\Delta_n := n \left(\sum_{j=1}^{m} (\Delta\tau^*_{nj})^p |a_{nj} - \sum_{l=1}^{m} \Delta\tau^*_{nl} a_{nl}|^p \right)^{1/p}.$$

Theorem: Let $n^{-1/2}\Delta_n \to \infty$ and $p > 2$. Then for each $0 < M < \infty$

$$\lim_{n\to\infty} P(\{n^{-p/2} T_{p,n} < M\}) = 0.$$