Applications of weak dependence to extreme value theory and resampling

Paul Doukhan
doukhan@u-cergy.fr
http:www.doukhan.u-cergy.fr

UMR 8088 AGM, University Cergy-Pontoise

Toulouse, TSE

September 22, 2016

Presentation

Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, a natural question is to quantify the asymptotic independence of this process at increasing times:

This problem is considered through elementary ideas and applications adapted to large sample data

Outline:

- From independence to dependence
- Models
- Technique
- Applications
- Application to extremes

Independence

We wish to answer the question
How to weaken the independence relation

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \quad ?$$

relating the events $A \in \sigma(P)$ of the past history with those $B \in \sigma(F)$ in a (not so close) future.
This relation is also restated as:

$$\mathrm{Cov}(f(P), g(F)) = 0, \qquad \forall f, g, \quad \|f\|_\infty, \|g\|_\infty \le 1$$

(Variables P, F denote here Past and Future)

Mixing (Rosenblatt, 1956)

$$\alpha(\sigma(P),\sigma(F)) = \sup_{A,B} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| = \frac{1}{2} \sup_{\|f\|_\infty,\|g\|_\infty\le 1} |\mathrm{Cov}(f(P),g(F))|$$

$X = (X_t)_{t\in\mathbb{Z}}$, $P = (X_{i_1},\dots,X_{i_u})$, $F = (X_{j_1},\dots,X_{j_v})$,
$i_1 \le \cdots \le i_u, j_1 \le \cdots \le j_v$ and $r = j_1 - i_u$ is large:

$$\alpha(r) = \sup_{P,F} \alpha(\sigma(P),\sigma(F)) \to_{r\to\infty} 0$$

See Rio 2000 for sharp technical results, see also Doukhan 1994 and Bradley 2007

Some nonmixing models

$X_t = \frac{1}{2}(X_{t-1}+\xi_t)$, $\xi_t \sim b\left(\frac{1}{2}\right)$ iid, Andrews-Rosenblatt (1984) ($X_{t-1} = \mathrm{frac}(2X_t)$)
$X_t = \xi_t(1+aX_{t-1})$, $\mathbb{P}(\xi_0 = \pm 1) = 1/2$, $a \in \left(\frac{3-\sqrt{5}}{2}, \frac{1}{2}\right]$, ($X_t = \sum_{j\ge 0} a^j \xi_t \cdots \xi_{t-j}$)

Covariances versus independence

Independence sometimes coincides with orthogonality
$\mathrm{Cov}(P,F)=0 \Longrightarrow$ independence of a random vector (P,F) if

$P,F\in\{0,1\}$	admit Bernoulli distributions
(P,F)	is a Gaussian vector
Z=(P,F)	is an associated vector (see below)

$Z\in\mathbb{R}^p$ associated $\Leftrightarrow$ $\mathrm{Cov}(f(Z),g(Z))\geq 0$ for $f,g:\mathbb{R}^p\to\mathbb{R}$ (coordinatewise $\uparrow$)

Then $|\mathrm{Cov}(f(P),g(F))|\leq \sum_{i,j} a_i b_j |\mathrm{Cov}(P_i,F_j)|$,

for $(P,F)\in\mathbb{R}^{p+q}$ associated or Gaussian

$$\begin{aligned}|f(x_1,\ldots,x_p)-f(y_1,\ldots,y_p)| &\leq a_1|x_1-y_1|+\cdots+a_p|x_p-y_p|\\ |g(x_1,\ldots,x_q)-g(y_1,\ldots,y_q)| &\leq b_1|x_1-y_1|+\cdots+b_q|x_q-y_q|\end{aligned}$$

Counterexamples: independent vectors, stability through $\uparrow$ images

A linear process

$$X_t = \sum_{j=-\infty}^{\infty} a_j \xi_{t-j}, \quad \sum_{j=-\infty}^{\infty} |a_j| < \infty, \|\xi_0\|_m < \infty, \quad (\xi_t)_{t\in\mathbb{Z}} \text{ iid}$$

$$X_t^p = \sum_{|j|<p} a_j \xi_{t-j} \Rightarrow \|X_t - X_t^p\|_m \le \|\xi_0\|_m \sum_{|j|\ge p} |a_j|,$$

$t - s > 2p \Rightarrow (X_s^p, X_t^p)$ independent.

$$\begin{aligned} |\mathrm{Cov}(f(X_s), g(X_t))| &\le |\mathrm{Cov}(f(X_s) - f(X_s^p), g(X_t))| \\ &+ |\mathrm{Cov}(f(X_s^p), g(X_t^p))| + |\mathrm{Cov}(f(X_s^p), g(X_t) - g(X_t^p))| \\ &\le 2\mathrm{Lip}\, g \|f\|_\infty \|X_s - X_s^p\|_1 + 2\mathrm{Lip}\, f \|g\|_\infty \|X_t - X_t^p\|_1 \end{aligned}$$

A definition of weak dependence should be flexible enough to include both this example (which includes ARMA models) and that of associated processes.
It should also yield *reasonable* limit theory in order to work out the consistency of statistical procedures.
Bickel & Bühlmann (1999) also define weak dependence to bootstrap such models: in this case innovations do not admit a density.

General formulation (Doukhan & Louhichi, 1999)

$(X_t)_{t\in\mathbb{Z}}$ $(\in E)$, $f : E^u \to \mathbb{R}$ from $\mathcal{F}$, $g : E^v \to \mathbb{R}$ from $\mathcal{G}$:

$$|\mathrm{Cov}\,(f(X_{i_1},\ldots,X_{i_u}),g(X_{j_1},\ldots,X_{j_v}))| \le \Psi(f,g)\epsilon(r), \quad \epsilon(r)\downarrow 0$$

$$\begin{aligned}\Psi(f,g) &= v\mathrm{Lip}\,g, & \epsilon(r)=\theta(r),\\ &= u\mathrm{Lip}\,f + v\mathrm{Lip}\,g + uv\mathrm{Lip}\,f\cdot\mathrm{Lip}\,g, & \epsilon(r)=\lambda(r)\end{aligned}$$

$$\mathrm{Lip}\,f = \sup_{(y_1,\ldots,y_u)\neq(x_1,\ldots,x_u)} \frac{|f(y_1,\ldots,y_u)-f(x_1,\ldots,x_u)|}{\|y_1-x_1\|+\cdots+\|y_u-x_u\|}.$$

Noncausal coefficients correspond to symmetric Ψ's.

Random fields or metric index sets are also considered (think of point processes).

vector LARCH(∞) models

$$X_t = \xi_t\Big(a + \sum_{j=1}^{\infty} a_j X_{t-j}\Big), \quad X_t(n\times 1), \xi_t(n\times p), a(p\times 1), a_j(p\times n)$$

$\phi = \|\xi_0\|_m \sum_j \|a_j\| < 1$, a $\mathbb{L}^m$-solution for (8) writes

$$X_t = \xi_t\Big(a + \sum_{k=1}^{\infty} \sum_{j_1,\ldots,j_k\geq 1} a_{j_1}\xi_{t-j_1}\cdots a_{j_k}\xi_{t-j_1-\cdots-j_k}a\Big)$$

Then	$\theta(t)$	$\leq$	$C\,t^{-b}$,	$C\,(q\vee\phi)^{\sqrt{t}}$,		$C\,e^{-bt}$	
if respectively	$A(s)$	$\leq$	$C's^{-b}$,	$C'q^s$,	or	$a_j = 0,\ j > C'$	
	$A(s) = \|\xi_0\|_m \sum_{j\geq s}\|a_j\|$						

- GARCH(p,q) (Engle, Granger) $r_t = \sigma_t\epsilon_t,\ \sigma_t^2 = \sum_{j=1}^p \beta_j\sigma_{t-j}^2 + \gamma_0 + \sum_{j=1}^q \gamma_j r_{t-j}^2$
- ARCH(∞) (Surgailis *et al.* 2001) $r_t = \sigma_t\epsilon_t, \quad \sigma_t^2 = \beta_0 + \sum_{j=1}^\infty \beta_j r_{t-j}^2$
- Bilinear (Giraitis, Surgailis, 2003) $X_t = \zeta_t\Big(a + \sum_{j=1}^\infty a_j X_{t-j}\Big) + b + \sum_{j=1}^\infty b_j X_{t-j}$

General memory models

$X_t = F(X_{t-1}, X_{t-2}, X_{t-3}, \ldots; \xi_t)$, $\qquad (\xi_t)_{t\in\mathbb{Z}}$ iid, $F : (\mathbb{R}^d)^{\mathbb{N}} \times \mathbb{R}^D \to \mathbb{R}^d$

with $\|F(x_1, x_2, x_3, \ldots; \xi_t) - F(y_1, y_2, y_3, \ldots; \xi_t)\|_m \le \sum_{j=1}^{\infty} a_j \|x_j - y_j\|$, then:

$\|F(0, 0, 0, \ldots; \xi_t)\|_m < \infty$, $a = \sum_{j=1}^{\infty} a_j < 1$ $(m \ge 1)$ imply existence in $\mathbb{L}^m$, stationarity and weak dependence:

$$\theta(r) \le C \inf_{N>0} \left(\sum_{j \ge N} a_j + a^{\frac{r}{N}} \right)$$

- Regression models $\quad X_t = f(X_{t-1}, \ldots, X_{t-k}) + \zeta_t g(X_{t-1}, \ldots, X_{t-k}) + \xi_t$
- variations on LARCH $\quad X_t = \xi_t \left(a + \sum_{j=1}^{\infty} a_j(X_{t-j}) \right)$, a_j Lipschitz
- Mean fields type models $\quad X_t = f\left(\xi_t, \sum_{s\ge 1} a_s X_{t-s}\right)$, f Lipschitz

Integer valued models

Thining, Steutel & van Harn operator is defined as

$$a \circ X = sign(X) \sum_{i=1}^{|X|} Y_i \qquad \text{for } a > 0, \quad X \in \mathbb{Z},$$

$(Y_i)_i$ is iid, context-independent, $\mathbb{E} Y_0 = a$ (e.g. Poisson or Bernoulli).

- Galton-Watson process with immigration, INAR $X_t = a \circ X_{t-1} + \xi_t$
- Integral bilinear models $X_t = a \circ X_{t-1} + b \circ (\varepsilon_{t-1} X_{t-1}) + \varepsilon_t$ Estimation from moments (Doukhan, Latour, Oraichi, 2006).
- INLARCH(∞) $X_t = \xi_t \left(a_0 + \sum_{j=1}^{\infty} a_j \circ X_{t-j} \right)$ QMLE (Latour, Truquet 2008).
- Random INAR models $X_t = a_t \circ X_{t-1} + \xi_t$, stationary (a_t) such $\mathbb{E}(a_t | \mathcal{F}_{t-1}) < 1$ (working paper with Lang).
- GLM integer models $X_t | \mathcal{F}_{t-1} \sim \mathcal{P}(\lambda_t)$ with $\lambda_t = g(\lambda_{t-1}, X_{t-1}, \ldots)$ with Fokianos and Tjostheim, 2011.

Existence of strictly stationary solutions, weak dependence properties $\Longrightarrow$ limit theory in estimation procedures.

Allowing $X_t \leq 0$ also gives non-associated and perhaps non-mixing processes

Limit theorems are fundamental to prove consistencies

- Moment inequalities
  - for integer moments, Doukhan & Louhichi use combinatorial methods
  - for causal coefficients Louhichi, Prieur use Lindeberg method
  - for $(2+\delta)-$order Doukhan & Wintenberger extend Ibragimov (1975) argument
- Exponential inequalities
  - For iid rvs, Bernstein inequality writes $\mathbb{P}(S_n \geq t\sqrt{n}) \leq C \exp\left\{-\frac{t^2}{2\sigma^2 + K\frac{t}{\sqrt{n}}}\right\}$
  - Doukhan, Louhichi use moment combinatorics to get $\leq Ce^{-c\sqrt{t}}$,
  - Doukhan, Neumann use cumulant techniques $\leq C \exp\left\{-\dfrac{t^2}{2\sigma^2 + K(t/\sqrt{n})^\alpha}\right\}$,
  - Rio (2000) and Dedecker (1999) extend Nagaev-Fuk maximal inequalities
  - Dedecker & Prieur use coupling arguments under causality. See also Rio, Merlevède and Peligrad (2010).

Limits in distribution enable goodness of fit tests I

A) Donsker invariance principles,

X_n stationary, with $\mathbb{E}X_0 = 0$, with $\sigma^2 = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_0, X_k) \geq 0$ (well defined), then

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} X_k \overset{D[0,1]}{\longrightarrow}_{n\to\infty} \sigma W_t$$

if one of those conditions holds

- $\mathbb{E}|X_0|^{2+\delta} < \infty$ and $\lambda(i) = O(i^{-a})$ for $a > 2 + 2/\delta$
- $\mathbb{E}|X_0|^{2+\delta} < \infty$ and $\kappa(i) = O(i^{-a})$ for $a > 2$
- $\mathbb{E}|X_0|^{2+\delta} < \infty$ and $\sum_{i>0} i^{1/\delta}\theta(i) < \infty$,
- $\mathbb{E}|X_0|^2 \log_+ |X_0| < \infty$ and $\theta(i) = O(a^i)$ for some $0 < a < 1$.

Dedecker, Doukhan, Louhichi, Prieur, Wintenberger

Limits in distribution enable goodness of fit tests II

B) Empirical Central Limit Theorem

X_n stationary, then $\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(\mathbf{1}(X_k \le x) - F(x)) \xrightarrow{D[\overline{\mathbb{R}}]}_{n\to\infty} Z(x)$ where $(Z(x))_{x\in\mathbb{R}}$ is the centered Gaussian process with covariance

$$\Gamma(x,y) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(\mathbf{1}(X_0 \le x), \mathbf{1}(X_k \le y))$$

if $F(x) \equiv x$, and a weak dependence condition is assumed

- $\theta(i) = O(i^{-a})$ for $a > 1$ (Dedecker and Prieur)
- $\lambda(i) = O(i^{-a})$ for $a > 15/2$ (under association: $a > 4$ is enough: Louhichi)
- $\eta(i) = O(i^{-a})$ for $a > 2 + 2\sqrt{2} \approx 4.8\cdots$ (Prieur)

Applications

- Estimation
  - Moment method for integer valued bilinear models (with Latour, Oraichi),
  - QMLE for ARCH(∞), INLARCH(∞)(Bardet, Latour, Truquet, Wintenberger)
  - Whittle estimator, empirical periodogram contrast (with Bardet, & León)
  - Kernel estimation $X_n = f(X_{n-1}, \ldots, X_{n-p}) + \xi_n g(X_{n-1}, \ldots, X_{n-q})$ (with Ango Nze, Dedecker, Louhichi, Prieur, Ragache, & Wintenberger), and prediction...
- Random fields, reliability of multicomponent systems (with Lang, Louhichi, Truquet, Ycart)
- Hard resampling is possible under nonparametric autoregression, since innovations dont need to have a density (with Neumann 2008, Neumann, Paparoditis, 2006, and with Mtibaa on progress)
- Stochastic algorithms, Sparsity, regression and density estimation (with Brandière 2004, and with Alquier 2011)
- Ripley statistics for point processes, uses spatial definitions for the dependence of such models (with Lang 2016)

Extreme Value Theory (EVT) I

Under simple dependence assumptions we prove that a distribution function G named a distribution function (O'Brien, 1987) satisfies

$$\sup_x |\mathbb{P}(M_n \leq x) - G^n(x)| \to_{n\to\infty} 0$$

which means a behavior 'as' in the iid case.
G is unique up to extremal equivalence $G \sim H$ in case

$$\sup_x |G^n(x) - H^n(x)| \to 0, \quad \text{as } n \to \infty.$$

O'Brien 1974 tail regularity is also assumed, for each $\gamma \in (0,1)$ there is a sequence with $\lim_n \mathbb{P}(M_n \leq v_n(\gamma)) = \gamma \iff \lim_{x\to G_-^*} \frac{1-G(x)}{1-G(x_-)} = 1$ $(G(G_-^*) = 1)$.

Extremal index

Suppose that $\{X_j\}$ admits a phantom distribution function G of the form $G(x) = F^{\theta_X}(x)$, for some $\theta_X \in (0,1]$, i.e.

$$\sup_u |\mathbb{P}(M_n \leq u) - (F^{\theta_X}(u))^n| \to 0, \quad \text{as} \quad n \to \infty$$

Then we say that $\{X_j\}$ has the extremal index θ_X (in the sense of Leadbetter (1983)).
In many cases the extremal index is the reciprocal of the mean size of clusters of big values occurring in the sequence $\{X_j\}$.

Zero extremal index I

Following Leadbetter (1983) we say that $\{X_j\}$ has the extremal index $\theta_X = 0$ if

$$\mathbb{P}(M_n \le u_n(\tau)) \to 1, \quad \text{if } n(1 - F(u_n(\tau)) \to \tau \in (0, \infty).$$

Intuitively this means that the partial maxima M_n increase much slower comparing with the independent case and that information on F alone cannot determine the limit behavior of laws of maxima M_n

Zero extremal index II

There exists non recurrent Markov chains with $0-$extremal index and which admits a phantom distribution
Asmussen (1998) considers the stationary Markov chain

$$X_{j+1} = (X_j + \xi_j)^+, \quad j = 1, 2, \dots,$$

with $\{\xi_j\}$ i.i.d. independent of X_0, with distribution function H and mean $-m < 0$.
If H is subexponential, and

$$\lim_{x\to\infty} \frac{1 - \mathbb{P}^2(\xi_0 \leq x)}{1 - \mathbb{P}(\xi_0 \leq x)} = 2,$$

then $\{X_j\}$ has the extremal index zero.

EVT Main result (Doukhan, Jakubowski, Lang, 2015)

In order to achieve such asymptotic behaviors:

Theorem

Let $\{X_j\}$ be stationary. Then the sequence $\{X_j\}$ admits a continuous *phantom distribution function if and only if $\exists\gamma \in (0,1)$ such that $\forall c > 0$:*

- $\mathbb{P}(M_n \leq v_n) \to_{n\to\infty} \gamma,$
- $\sup_{p+q\leq cn} \left|\mathbb{P}\left(M_{p+q} \leq v_n\right) - \mathbb{P}\left(M_p \leq v_n\right)\mathbb{P}\left(M_q \leq v_n\right)\right| \to_{n\to\infty} 0.$

.

This follows under strong mixing or under weak dependence with convenient rates if the marginal distribution of X is regular enough.
In order to fit such phantoms, a main project relies on the estimation of quantiles for the extremes of block.

Extreme Values Theory II

In some situations there exist sequences u_b, v_b with:

$$\mathbb{P}\left(u_b \max_{1\le i\le b} X_i - v_b \le x\right) \to \mathbb{K}(x)$$

A well known case is more demanding and corresponds to the existence of extremal indices θ_X.
Here $\mathbb{K}(x) \equiv \mathbb{G}_\alpha^{\theta_X}(x)$ for $\mathbb{G}_\alpha$ the law of attraction for extremes parametrized by its Hill index $\alpha \in \mathbb{R}$.
Extremal and Pareto indices are estimated for stationary times series (X_t), they entail the knowledge of the asymptotic behavior of the extremes in most of the cases,
but Gumbel limiting distribution necessitates an additional nonparametric estimation!
For the case of Markov chains such asymptotic behaviors need recurrence.

Subsampling (Doukhan, Prohl, Robert, 2011)

Besides direct estimates of the asymptotic behaviors for extremes, alternative techniques may be used.

Subsampling will ensure some estimated confidence intervals for extremes in order to estimate Values at Risks and others financial quantities.

A first step:

Convergent statistics sequences $T_b = t_b(x_1, \ldots, x_b)$ may be subsampled with

$$\frac{1}{n - b_n + 1} \sum_{i=0}^{n-b_n} \mathbb{1}_{\{T_{b_n,i} \leq x\}}, \qquad T_{b_n,i} = t_{b_n}(X_{i+1}, \ldots, X_{i+b_n})$$

$$\frac{b_n}{n - b_n + 1} \sum_{i=0}^{n/b_n - 1} \mathbb{1}_{\{T_{b_n,i} \leq x\}}, \qquad T_{b_n,i} = t_{b_n}(X_{ib_n+1}, \ldots, X_{(i+1)b_n})$$

$t \mapsto \mathbb{1}_{t \leq x}$ is replaced by a continuous $1/\epsilon_n-$Lipschitz approximation.

Centered higher moments yield uniform almost sure convergence

Renormalization (Doukhan, Prohl, Robert, 2011)

The divergent case $T_b = t_b(x_1, \ldots, x_b) = \max\{x_1, \ldots, x_b\}$ needs more attention

If $\mathbb{P}\left(u_b \max_{1\le i\le b} X_i - v_b \le x\right) \to \mathbb{K}(x) \equiv \mathbb{G}^{\theta x}(x)$,

Set $\widetilde{\mathbb{H}}_{b,n}(x) = \dfrac{1}{N}\displaystyle\sum_{i=0}^{N-1} \varphi\left(\frac{m_b(Y_{b,i}) - x}{\epsilon_n}\right)$,

$\widehat{v}_{b,n} = \widetilde{\mathbb{H}}_{b,n}^{-1}\left(\frac{1}{2}\right)$, $\widehat{u}_{b,n} = \left(\widetilde{\mathbb{H}}_{b,n}^{-1}(t_2) - \widetilde{\mathbb{H}}_{b,n}^{-1}(t_1)\right)^{-1}$ estimate v_b, u_b

up to the normalization $\mathbb{K}^{-1}\left(\frac{1}{2}\right) = 0$ median equals 0 and

up to a multiplicative constant $C = \mathbb{K}^{-1}(t_2) - \mathbb{K}^{-1}(t_1)$ fix interquartiles.

Now

$$\widetilde{\mathbb{H}}_{b,n}\left(\widehat{v}_{b,n} + \frac{x}{\widehat{u}_{b,n}}\right) \to_{n\to\infty} \mathbb{K}\left(\frac{x}{C}\right), \qquad (b = b(n) \to \infty \quad \text{conveniently})$$

convergence is uniform, either in probability or a.s.

Further projects with extremes

In many cases EVT for nonmixing sequences are an important question.
For this we plan further works on:

- From the construction of phantoms in 2015's paper with Jakubowski and Lang, a sequence of $2^{-k}-$quantiles is enough to fit one of them. This is easy to provide consistent blockwise estimations of those quantiles. A smoothing procedure should allow this estimation..
- Multidimensional extremes may not be directly obtained but a copula vision should allow to fit EVT (with Chautru and Segers)
- Self driven subsampling procedures (with Bertail)
- Clusters determined through Drees and Rootzen 2011's blockwise general procedure (FCLT for clusters) are extended with Gomez under weak dependence for specific applications including the estimation of Hill index

Wild Bootstrap (D., Lang, Neumann, Leucht, 2014) I

Suppose that we observe a stretch $X_1, \ldots, X_n$ from a (strictly) stationary and real-valued process $(X_t)_{t\in\mathbb{Z}}$. We denote by F the common cumulative distribution function of the X_t and by F_n the empirical distribution function, i.e.

$$F_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \mathbb{1}(X_t \le x).$$

The empirical process $G_n(x) = \sqrt{n}\,(F_n(x) - F(x)) \xrightarrow{d} G(x)$ under some assumptions for some Gaussian process with $EG(x) = 0$ and $\mathrm{Cov}(G(x), G(y)) = \sum_{t=-\infty}^{\infty} \mathrm{Cov}(\mathbb{1}(X_0 \le x), \mathbb{1}(X_t \le y))$. For tails analysis, consider a weighted version of the empirical process, e.g.

$$H_n(x) = \frac{\sqrt{n}}{\sqrt{F(x)(1-F(x))}}\,(F_n(x) - F(x))\,.$$

(strong mixing, absolute regularity, $\phi-$mixing, weak dependence, and others may be considered)

Wild Bootstrap (D., Lang, Neumann, Leucht, 2014) II

The so-called dependent wild bootstrap was introduced by Shao (2010) for smooth functions of the mean. Originally, the idea of the dependent wild bootstrap is to construct the pseudo-observations as follows:

$$X_t^* = \bar{X}_n + \left(X_t - \bar{X}_n\right) \varepsilon_{t,n}^*, \qquad t = 1, \ldots, n.$$

Here $\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t$ and $(\varepsilon_{t,n}^*)_{1 \le t \le n}$ is a triangular scheme of weakly dependent random variables:

- the $\varepsilon_{t,n}^*$ are independent of $X_1, \ldots, X_n$,
- $E^* \varepsilon_{n,t}^* = 0$,
- $\mathrm{Cov}^*(\varepsilon_{s,n}^*, \varepsilon_{t,n}^*) = \rho(|s-t|/l_n)$, where $\rho(u) \to_{u \to 0} 1$
- $\sum_{r=1}^n \rho(r/l_n) = O(l_n)$, $l_n \to_{n \to \infty} \infty$ and $l_n = o(n)$.

Shao (2010) verified that under certain regularity conditions
$\sup_x \left| P\left(\sqrt{n}\left[H(\bar{X}_n) - H(EX_1)\right] \le x\right) - P^*\left(\sqrt{n}\left[H(\bar{X}_n^*) - H(\bar{X}_n)\right] \le x\right)\right| \stackrel{P}{\to} 0$
where H is a smooth function and $\bar{X}_n^* = \frac{1}{n} \sum_{t=1}^n X_t^*$.

Wild Bootstrap (D., Lang, Neumann, Leucht, 2014) III

In our case of the empirical process, the role of the X_t above is taken by $Z_t = \mathbb{1}(X_t \leq x)$.
Following 'Shao's idea' we define bootstrap counterparts of the Z_t as

$$Z_t^* = \bar{Z}_n + (Z_t - \bar{Z}_n)\, \varepsilon_{t,n}^* = F_n(x) + (\mathbb{1}(X_t \leq x) - F_n(x))\; \varepsilon_{t,n}^*.$$

This leads to the following bootstrap version of the empirical process:

$$G_n^*(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (Z_t^* - \bar{Z}_n) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (Z_t - F_n(x))\; \varepsilon_{t,n}^* = G_n^{*,0}(x) - R_n^*(x),$$

$$G_n^{*,0}(x) = \textstyle\sum_{t=1}^n (Z_t - F(x))\, \varepsilon_{t,n}^* / \sqrt{n}, \quad R_n^*(x) = (F_n(x) - F(x)) \sum_{t=1}^n \varepsilon_{t,n}^* / \sqrt{n}.$$

Since $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = O_P(1)$ *we obtain* $\sup_{x\in\mathbb{R}} |R_n^*(x)| = O_{P^*}(\sqrt{l_n/n})$.
$\{Y_n^* = O_{P^*}(r_n)\}$ if $\{\forall \epsilon > 0, \exists K(\epsilon)$ with $P(P^*(|Y_n^*/r_n| > K(\epsilon)) > \epsilon) \underset{n\to\infty}{\longrightarrow} 0\}$.

Hence, we can analyze $G_n^{*,0}$ instead of G_n^* in the sequel.

Wild Bootstrap (D., Lang, Neumann, Leucht, 2014) IV

Applications are important for statistics

- Confidence sets for quantiles; besides the inversion of the cumulative distribution function, a statistical validation is needed and obtained through bootstraping techniques.
- Kolmogorov Smirnov type tests of hypothesis for time series: in this case indeed the supremum of the limit process G do depend on the marginal distribution of the sample, contrary to the independent case.
- Order statistics: asymptotics are still reachable here through the weighted empirical process.

Estimating a variance: D., Jakubowicz, León (2009) I

$$\text{If}\quad \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k \to_{n\to\infty} \mathcal{N}_d(0,\Sigma),\quad \text{with}\quad \Sigma = \sum_{k=-\infty}^{\infty} \mathbb{E}X_0X_k'$$

Self-normalized results yield asymptotic confidence sets, Σ is estimated by:

- Spectrum: $\widehat{\Sigma} = \widehat{f}(0)$ if the matrix-spectral density is estimated
- Donsker: $\frac{1}{\sqrt{n}}\sum_{ns<i<nt} X_i \to Z(t) - Z(s)$ Brownian, $Z(1) \sim \mathcal{N}_d(0,\Sigma)$

$$\Delta_{j,n} = \frac{1}{\sqrt{n}}\sum_{i\in B_j} X_i \to Z(t_j) - Z(s_j) \qquad (B_j = [ns_j, nt_j] \cap \mathbb{N})$$

Then for suitable choices of F, and $0 = s_1 < t_1 \le s_2 < \cdots \le s_m < t_m = 1$

$$\widetilde{F}_n = \frac{1}{m}\sum_{j=1}^{m} F(\Delta_{j,n}) \to \mathbb{E}F\big(\mathcal{N}_d(0,\Sigma)\big)$$

Carlstein (1986) mixing, Peligrad-Shao (1995) ρ-mixing use both $t_i = s_{i+1}$

Estimating a variance: D., Jakubowicz, León (2009) II

In order to derive a self-normalized CLT, D., Jakubowicz, León (2009) set $t_i < s_{i+1}$ and, under weak dependence:

$$\frac{\sqrt{N_n}}{\sqrt{(\widehat{G}_n - \widehat{F}_n^2)^+}}\left(\widetilde{F}_n - \mathbb{E}F(\mathcal{N}_d(0,\Sigma)\right) \to \mathcal{N}(0,1), \qquad (G \equiv F^2)$$

Applications to
- Linear models with dependent inputs
- Sea waves modeling, $X_t = F(Y_t)$ for F approximately linear
- Crossing numbers of oscillatory systems

For such explicit examples for which such procedures is proved to be useful through simulation studies.

References

Andrews, D. (1984) *Non strong mixing autoregressive processes.* J. Appl. Probab. 21, 930-934.

Dedecker, J., Doukhan, P., León, J. R., Lang, G., Louhichi, S., Prieur, C. (2007) *Weak dependence: models, theory and applications.* Lecture Notes in Statistics 190, Springer-Verlag.

Doukhan, P. (1994) *Mixing: Properties and Examples.* LNS 85. Springer Verlag.

Doukhan, P. (2016) *Stochastic Models for Time Series.* To appear 343 pages. Series Mathématiques et Applications. Springer Verlag.

Doukhan, P., Jakubowski, A., Lang, G. (2015) Phantom distribution functions for some stationary sequences. Extremes 18:697?725.

Doukhan, P., Jakubowicz, J., León, J.R.(2009) *Estimation of the limit variance for sums of* $\lambda-$*dependent vector valued sequences.* Volume to W. Philipp memory.

Doukhan, P., Lang, G., Leucht, A., Neumann, M. (2014) Dependent Wild Bootstrap for the empirical process. J.T.S.A..

Doukhan, P., Prohl, S., Robert, C. (2011) *Subsampling extremes of weakly dependent sequences.* TEST.

Leadbetter, M. R., Lindgren, G., Rootzen, H. (1983) *Extremes and related properties of random sequences and processes.* Springer Series in Statistics.

Martins, A.P., Ferreira, H. (2005) *The multivariate extremal index and the dependence structure of a multivariate extreme value distribution.* Test 14, 433–448.

Rio, E. (2000) *Théorie asymptotique pour des processus aléatoires faiblement dépendants,* SMAI, Mathématiques et Applications 31, Springer. English version 2012, on ArXiv.