Extremes of Time Series

Johan Segers

joint work with B. Basrak, A. Bücher, R. Davis, H. Drees, C.A.T. Ferro, A. Janßen, D. Krizmanić, T. Meinguet, C.Y. Robert, M. Warchoł, Y. Zhao

Standing on the shoulders of many others colleagues

Université catholique de Louvain
Institut de Statistique, Biostatistique et Sciences Actuarielles

Workshop on "Extremes, copulas, and actuarial science"
CIRM – Centre International de Rencontres Mathématiques
Luminy, Marseille, February 23, 2016

Extremes of Time Series

I Block maxima

II Clusters

III Tail processes

Part I

Block maxima

Extremes of Time Series

The annual maximum method for time series

Univariate maxima: maximum likelihood

Multivariate maxima: extreme-value copula

Max-stability

A distribution is *max-stable* iff maxima of iid samples from it have the same distribution up to location and scale.

Max-stability: cdf G solves functional equation

$$G^n(a_n x + b_n) = G(x)$$

Solve for G?

Solution: generalized extreme-value distribution.

- univariate: three-parameter family
- multivariate: additionally, extreme-value copula

[Fisher and Tippett 1928; de Haan and Resnick 1977]

Extremal types theorem(s)

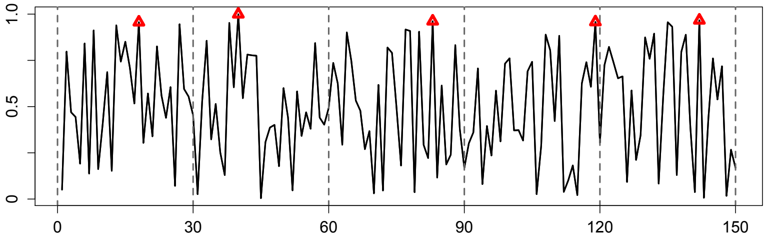
Under weak dependence, limit laws (if any) of affinely normalized block maxima are max-stable, hence GEV.

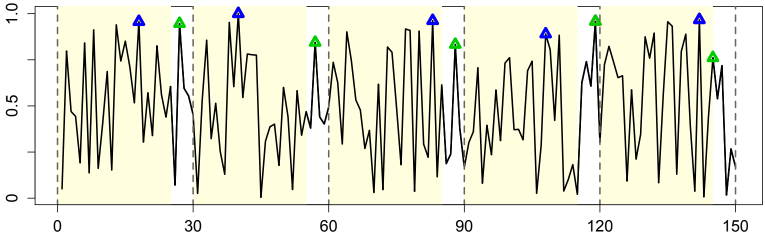
Idea:

- Write maximum as a maximum over maxima over (approximately) independent smaller blocks.
- Maxima over the smaller blocks converge to the same distribution as the global maximum (convergence of types theorem).
- Two different ways of obtaining the same limit.

[Gnedenko 1943; Leadbetter 1974; Hsing 1989; Hüsler 1990]

Weak dependence heuristic: big blocks, small blocks





Annual maximum method

Method: fit a GEV to a sample of block maxima.

Univariate:

- matching probabilities or quantiles
- matching (probability weighted) moments
- maximum likelihood

Multivariate:

- nonparametric techniques
- assume parametric model and do maximum likelihood

Asymptotic frameworks

Data generating process?

iid random variables or vectors sampled from:

- the limiting GEV
- a distribution with an extreme-value copula
  - margins partially (un)known

Triangular array of block maxima extracted from

- an iid time series
- a stationary time series
- a general, non-specified time series

Extremes of Time Series

The annual maximum method for time series

Univariate maxima: maximum likelihood

Multivariate maxima: extreme-value copula

Fréchet MLE

Given: $\boldsymbol{M}_n = (M_{n,1}, \ldots, M_{n,k_n})$, sample of 'block maxima', not all tied.

Think of $M_{n,i}$ as approximately (i)id Fréchet(α_0, σ_n)

$$G_{\alpha,\sigma}(x) = \exp\{-(x/\sigma)^{-\alpha}\}, \qquad x > 0$$

Estimate Fréchet parameters by maximum likelihood:

$$(\hat{\alpha}_n, \hat{\sigma}_n)$$

Fréchet MLE: consistency

$\boldsymbol{M}_n = (M_{n,1}, \ldots, M_{n,k_n})$, sample of 'block maxima', not all tied.

Assumption: there exists $\sigma_n > 0$ such that

$$\frac{1}{k_n} \sum_{i=1}^{k_n} f(M_{n,i}/\sigma_n) \xrightarrow{p} \int_0^\infty f(x)\, \mathrm{d}G_{\alpha_0,1}(x)$$

for all f of the form $f(x) = x^{-\alpha}$ or $f(x) = x^{-\alpha} \log x$, all $\alpha > 0$

Then the Fréchet MLE $(\hat{\alpha}_n, \hat{\sigma}_n)$ exists, is unique, and

$$\hat{\alpha}_n \xrightarrow{p} \alpha_0,$$
$$\hat{\sigma}_n/\sigma_n \xrightarrow{p} 1$$

Proof: investigate asymptotic properties of score equations.

[Bücher and S. 2015]

Fréchet MLE: Asymptotic normality

$\boldsymbol{M}_n = (M_{n,1}, \ldots, M_{n,k_n})$, sample of ‘block maxima’, not all tied.

Assumption: convergence in probability or asymptotic normality of statistics of the form

$$\frac{1}{k_n} \sum_{i=1}^{k_n} f(M_{n,i}/\sigma_n)$$

for functions $f : (0, \infty) \to \mathbb{R}$ arising in the score equations

Then the Fréchet MLE $(\hat{\alpha}_n, \hat{\sigma}_n)$ exists, is unique, and

$$v_n(\hat{\alpha}_n - \alpha_0, \hat{\sigma}_n/\sigma_n - 1) \xrightarrow{d} \text{weak limit}$$

Rate of convergence $v_n \to \infty$ determined by assumption; typically $v_n = \sqrt{k_n}$

[Bücher and S. 2015]

Special case: stationary time series

Specialize previous theorem to

$$M_{n,i} = \max\{X_t : t = (i-1)r_n + 1, \ldots, ir_n\}$$

(ith block of size r_n, for block $i = 1, \ldots, k_n$, with $k_n = \lfloor n/r_n \rfloor$ blocks)

Conditions on $(X_t)_{t\in\mathbb{Z}}$ to ensure that the general theorem applies:

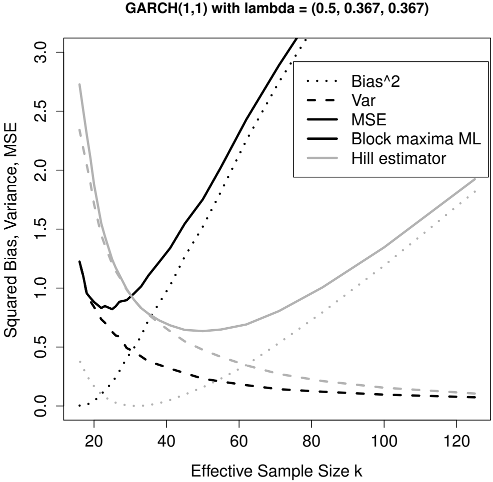
- rescaled maxima M_n/σ_n attracted by Fréchet distribution
- control on the rate of convergence
- moment conditions
- mixing

$\Longrightarrow$ Asymptotic normality of Fréchet MLE $(\hat{\alpha}_n, \hat{\sigma}_n)$

- non-zero mean possible due to rate of convergence
- covariance matrix: inverse of Fisher information matrix as if iid Fréchet random sample

Finite-sample performance

Comparison of Fréchet MLE $\hat{\alpha}_n$ based on k block maxima with Hill estimator based on k largest upper order statistics



$$X_t = |Z_t|$$

$$Z_t = \varepsilon_t \sigma_t, \qquad \varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$$

$$\sigma_t^2 = \lambda_0 + \lambda_1 Z_{t-1}^2 + \lambda_2 \sigma_{t-1}^2$$

$$\alpha_0 = 5$$

Sample size $n = 1000$
3000 repetitions

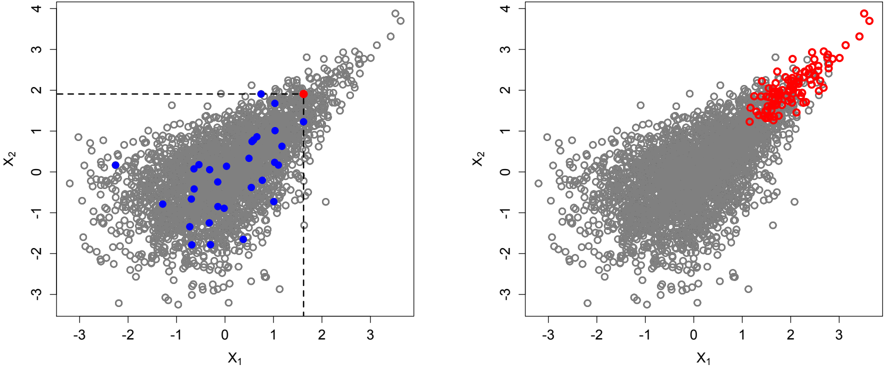
Extremes of Time Series

The annual maximum method for time series

Univariate maxima: maximum likelihood

Multivariate maxima: extreme-value copula

Block maxima



one year

all years

Copula extremal types theorem

Multivariate stationary time series, weakly dependent.
Vector of componentwise sample maxima.

Copula extremal types theorem

Limit copula (if any) of a vector of componentwise sampl maxima is an extreme-value copula, C_E.

- This C_E could be different from the extreme-value copula attractor of the stationary distribution (multivariate extremal index).

[Hsing 1989; Hüsler 1990]

Empirical copula

Sample of k_n vectors of componentwise block maxima: estimate C_E?

Empirical copula $\hat{C}_n$: empirical cdf of vectors of normalized ranks.

[Deheuvels 1979]

Empirical copula process:

$$\sqrt{k_n}(\hat{C}_n - C_E)$$

Converges weakly to the same limit as if $\hat{C}_n$ were the empirical copula from an iid sample from C_E, plus possible bias term.

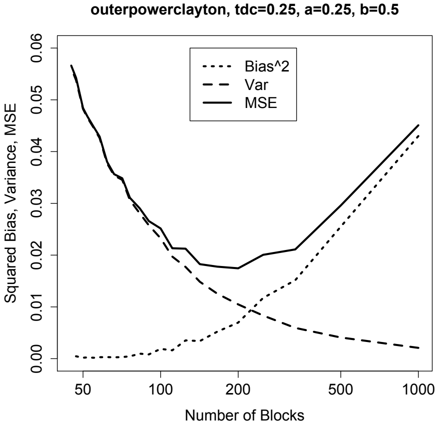
Conditions:

- mixing conditions
- growth of block sizes
- rate of convergence in copula extremal type theorem
- smoothness of C_E

[Bücher and S. 2014]

Estimating the Pickands dependence function

Weak convergence of empirical copula process yields weak convergence of estimators of other dependence objects, e.g., Pickands dependence function



Bivariate moving maximum process, sample size $n = 1000$

Estimator from Bücher, Dette, Volgushev (2011)

Part II

Clusters of extremes

Extremes of Time Series

Extremal index

Cluster functionals and the cluster map

Approximate cluster distributions

Maxima of stationary time series

Stationary, real-valued time series $(X_t)_{t \in \mathbb{Z}}$. Weakly dependent.

Sample of size n, divided in k_n blocks of size r_n.

Levels u_n such that $n\mathrm{P}[X_0 > u_n] \to \tau \in (0, \infty)$.

$$\mathrm{P}[M_n \le u_n] \approx (\mathrm{P}[X_0 \le u_n])^{n\theta_n}$$

where

$$\begin{aligned} \theta_n &= \frac{\mathrm{P}[M_{r_n} > u_n]}{r_n \mathrm{P}[X_0 > u_n]} \\ &\approx \mathrm{P}[M_{1,r_n} \le u_n \mid X_0 > u_n] \end{aligned}$$

Limit, θ, of θ_n, if any, is the extremal index. [Leadbetter 1983; O'Brien 1987]
$\implies$ 'Blocks' and 'runs' estimators.

[Hsing 1991, 1993; Smith & Weissman 1994; Weissman & Novak 1998; Robert, S., Ferro 2009]

Other characterizations of the extremal index

Stationary cdf $F(x) = \mathrm{P}[X_0 \leq x]$.

Inter-exceedance times: $T(u_n) = \inf\{k \geq 1 : X_k > u_n\}$ given $X_0 > u_n$.

$$\mathrm{P}[T(u_n) > t/\overline{F}(u_n) \mid X_0 > u_n] \to \theta\, e^{-t\theta}, \qquad t > 0$$

[Ferro and S. 2003]

Distorted probability integral transform:

$$\mathrm{P}[(F(M_n))^n \leq x] \to x^{\theta}, \qquad 0 < x < 1$$

[Northrop 2015]

An informal view on clusters

For weakly dependent stationary sequences,
extremes arrive in clusters.

We are concerned with the asymptotic distribution of the 'block'

$$(X_1, \ldots, X_{r_n})$$

given that at least one 'extreme value' occurs

$$\sum_{i=1}^{r_n} \mathbf{1}(X_i \text{ hits an exceptional set}) \geq 1 \tag{C}$$

when the expected number of extremes is asymptotically negligible

$$r_n \, \mathrm{P}(X_1 \text{ hits an exceptional set}) = o(1)$$

Extremes of Time Series

Extremal index

Cluster functionals and the cluster map

Approximate cluster distributions

Cluster statistics

Ingredients

- Stationary process $(X_n)_n$ on $\mathbb{R}$
- High threshold u_n
- Block size r_n

Interest is in cluster statistics of the form

$$c(X_1 - u_n, \ldots, X_{r_n} - u_n) \quad \text{conditionally on} \quad M_{r_n} > u_n$$

that only depend on the 'cluster':
the stretch between the first and the last exceedance over u_n.

We require that

$$r_n \to \infty, \qquad r_n \, \mathrm{P}(X_1 > u_n) \to 0$$

Examples of cluster statistics

- Block maximum: maximal excess

$$c(y_1, \ldots, y_{r_n}) = \max(y_1, \ldots, y_{r_n})$$

- Aggregate excess: sum of excesses

$$c(y_1, \ldots, y_{r_n}) = \max(y_1, 0) + \cdots + \max(y_{r_n}, 0)$$

- Cluster size: number of excesses

$$c(y_1, \ldots, y_{r_n}) = \mathbf{1}(y_1 > 0) + \cdots + \mathbf{1}(y_{r_n} > 0)$$

- Cluster duration: time span between first and last excess

$$c(y_1, \ldots, y_{r_n}) = \max\{i : y_i > 0\} - \min\{i : y_i > 0\} + 1$$

- Number of threshold upcrossings

$$c(y_1, \ldots, y_{r_n}) = \mathbf{1}(y_1 > 0) + \mathbf{1}(y_1 \leq 0 < y_2) + \cdots + \mathbf{1}(y_{r_n - 1} \leq 0 < y_{r_n})$$

Cluster functionals

Desirable properties of $c(\,\cdot\,)$:

- Its domain is a vector of arbitrary length
  with at least one non-zero component.
- It depends only on the 'extreme' part of the vector

Definition

A cluster functional is a map $c : \boldsymbol{A} \to \mathbb{R}$ with

$$\boldsymbol{A} = \boldsymbol{A}_1 \cup \boldsymbol{A}_2 \cup \dots$$
$$\boldsymbol{A}_r = \mathbb{R}^r \setminus (-\infty, 0]^r = \{(y_1, \dots, y_r) \in \mathbb{R}^r : \max(y_1, \dots, y_r) > 0\}$$

and neglecting everything that happened
before or after the first or last positive value:

$$c(y_1, \dots, y_r) = c(y_\alpha, \dots, y_\omega)$$
$$\alpha = \min\{i : y_i > 0\}$$
$$\omega = \max\{i : y_i > 0\}$$

[Yun 2000; S. 2003; Drees & Rootzén 2010]

Cluster map

Definition

Recall $\boldsymbol{A} = \bigcup_{r \geq 1} \boldsymbol{A}_r$ and $\boldsymbol{A}_r = \mathbb{R}^r \setminus (-\infty, 0]^r$. Define the cluster map

$$C : \boldsymbol{A} \to \boldsymbol{A} : (y_1, \ldots, y_r) \mapsto (y_\alpha, \ldots, y_\omega)$$
$$\alpha = \min\{i : y_i > 0\}$$
$$\omega = \max\{i : y_i > 0\}$$

[Segers 2005]

Then $c : \boldsymbol{A} \to \mathbb{R}$ is a cluster functional if and only if

$$c = f \circ C \text{ for some } f : \boldsymbol{A} \to \mathbb{R}$$

Hence, to know the asymptotic distribution of cluster statistics,
it is sufficient to know the asymptotic distribution of the 'cluster' itself

$$C(X_1 - u_n, \ldots, X_{r_n} - u_n) \quad \text{conditionally on} \quad M_{r_n} > u_n$$

Extremes of Time Series

Extremal index

Cluster functionals and the cluster map

Approximate cluster distributions

Aim: switch to a simpler conditioning event

We are interested in the cluster distribution

$$\mathrm{P}[C(X_1 - u_n, \ldots, X_{r_n} - u_n) \in \cdot \mid M_{r_n} > u_n]$$

Recall $r_n \to \infty$ and $r_n \mathrm{P}(X_1 > u_n) \to 0$.

The conditioning event $\{M_{r_n} > u_n\}$ is awkward to work with: when exactly did the exceedances occur?

We'd rather prefer expressions in terms of the law of

$$(X_1, \ldots, X_k) \mid X_1 > u_n$$

This would be particularly convenient in the case of Markov chains.

Expected cluster size

Expected number of exceedances given that there is at least one:

$$\mathrm{E}\left[\sum_{i=1}^{r_n} \mathbf{1}(X_i > u_n) \,\middle|\, M_{r_n} > u_n\right] = \frac{r_n \,\mathrm{P}(X_1 > u_n)}{\mathrm{P}(M_{r_n} > u_n)} =: \frac{1}{\theta_n}$$

so

$$\theta_n = \frac{\mathrm{P}(M_{r_n} > u_n)}{r_n \,\mathrm{P}(X_1 > u_n)} \in (0, 1]$$

Example

In the iid case, since $r_n \,\overline{F}(u_n) \to 0$, we have

$$\theta_n = \frac{1 - (1 - \overline{F}(u_n))^{r_n}}{r_n \overline{F}(u_n)} \to 1$$

Finite-cluster condition

Suppose that the impact of a shock is somehow limited in time:

$$\underbrace{X_1}_{>u_n}, X_2, \ldots, X_m, \underbrace{X_{m+1}, \ldots, X_{r_n}}_{>u_n?}$$

$$\underbrace{X_1, \ldots, X_{r_n-m}}_{>u_n?}, X_{r_n-m+1}, \ldots, X_{r_n-1}, \underbrace{X_{r_n}}_{>u_n}$$

Formally, put $M_{i,j} = \max(X_i, \ldots, X_j)$ and suppose

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathrm{P}(M_{m+1,r_n} > u_n \mid X_1 > u_n) = 0 \qquad \text{(FiCl1)}$$

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathrm{P}(M_{1,r_n-m} > u_n \mid X_{r_n} > u_n) = 0 \qquad \text{(FiCl2)}$$

Sufficient condition:

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sum_{i=m+1}^{r_n} \mathrm{P}(X_i > u_n \mid X_1 > u_n) = 0 \qquad \text{(FiCl)}$$

Bounded expected cluster sizes

If (FiCl), the expected cluster size remains bounded:

$$\limsup_{n\to\infty} \frac{r_n \, \mathrm{P}(X_1 > u_n)}{\mathrm{P}(M_{r_n} > u_n)} < \infty$$

$$\text{i.e. } \liminf_{n\to\infty} \theta_n > 0.$$

Proof: observe that $M_{r_n} \geq \max(X_1, X_{m+1}, X_{2m+1}, \ldots, X_{km+1})$ with $k \sim r_n/m$.

The approximant

Consider a bounded, measurable cluster functional $c : \boldsymbol{A} \to \mathbb{R}$.
Apply c to different stretches of the process:

$$c_n(i,j) = c(X_i - u_n, \ldots, X_j - u_n) \text{ on the event } M_{i,j} > u_n$$

Consider the approximation error

$$\left| \underbrace{\mathrm{E}[c_n(1, r_n) \mid M_{r_n} > u_n]}_{\text{quantity of interest}} - \underbrace{\frac{\alpha_{n,m}(c)}{\theta_{n,m}}}_{\text{approximant}} \right|$$

where

$$\begin{aligned} \alpha_{n,m}(c) &= \mathrm{E}[c_n(1,m) \mid X_1 > u_n] \\ &\qquad - \mathrm{E}[c_n(2,m),\, M_{2,m} > u_n \mid X_1 > u_n] \\ \theta_{n,m} &= \mathrm{P}[M_{2,m} \le u_n \mid X_1 > u_n] \qquad \text{'runs'} \end{aligned}$$

The cluster approximation

Theorem

If (FiCl), then

$$\lim_{m\to\infty} \limsup_{n\to\infty} |\underbrace{\theta_{n,m}}_{\text{'runs'}} - \underbrace{\theta_n}_{\text{'blocks'}}| = 0$$

as well as

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sup_{c:|c|\le 1} \left| \mathrm{E}[c_n(1, r_n) \mid M_{r_n} > u_n] - \frac{\alpha_{n,m}(c)}{\theta_{m,n}} \right| = 0$$

[S. 2005]

Proof: elementary calculations, based on careful use of

- partitionings of the event $\{M_{r_n} > u_n\}$ and similar ones
- stationarity
- the cluster property
- (FiCl)

Part III

Tail processes

Extremes of Time Series

Probabilistic background

Estimation

Regular variation: a convenient hypothesis

Regularly varying tails:

$$P(X > u) = u^{-\alpha} L(u)$$

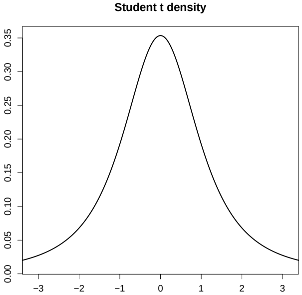
with L 'slowly varying' (think of it as constant)

Both positive and negative spikes:

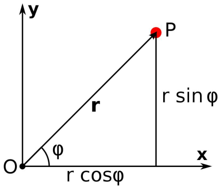
$$P(|X| > u) = u^{-\alpha}\, L(u)$$

$$P(X > u) \sim p_+\, P(|X| > u)$$
$$P(X < -u) \sim p_-\, P(|X| > u)$$

Example: Student t ($p_+ = p_- = 1/2$)



Far far away



$$(X, Y) \to \begin{cases} r = \sqrt{X^2 + Y^2} & \text{how big} \\ \varphi = \arctan(Y, X) & \text{in which direction} \end{cases}$$

Multivariate regular variation:

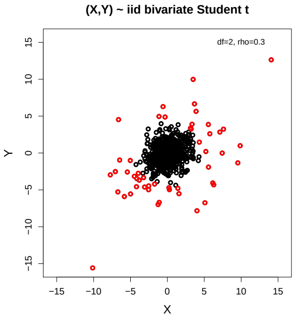
$$P[r > u] = u^{-\alpha} L(u)$$

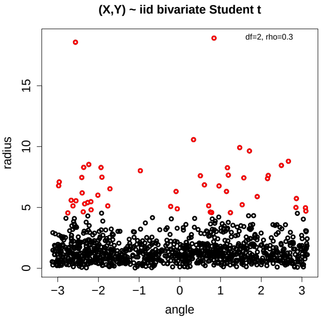
$$P[\varphi \in \cdot \mid r > u] \xrightarrow{d} H(\cdot)$$

Stability at infinity

As the radius grows larger, it becomes independent of the angle:

$$P[r > uy\ ,\ \varphi \in \cdot\ \mid r > u] \xrightarrow{d} y^{-\alpha} H(\,\cdot\,)$$



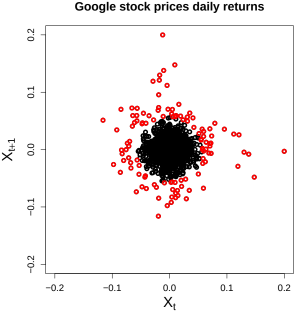


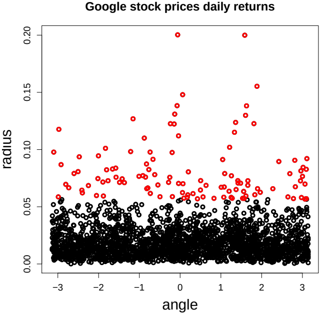
Time is just another form of space

- Stationary time series $X_0, X_1, X_2, \ldots$.
- Impact of a shock on the future?

Assumption

$(X_t, X_{t+1}, \ldots, X_{t+h})$ is regularly varying $\forall t, h$





Troubles ahead

Regular variation is equivalent to the existence of an 'extreme regime':

$$\mathcal{L}\big((X_h/u)_h \mid |X_0| > u\big) \xrightarrow{d} (Y_h)_h \qquad \text{tail process}$$

$$\mathcal{L}\big((X_h/|X_0|)_h \mid |X_0| > u\big) \xrightarrow{d} (\Theta_h)_h \qquad \text{spectral tail process}$$

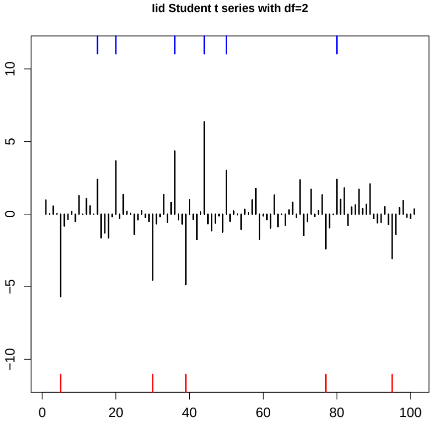
[Basrak & S. 2009, 2011]

- What are Y_h and Θ_h for usual time series models?
- How can we estimate them?
- What can we learn from them?

Independence: isolated spikes

$X_0, X_1, X_2, \dots$ iid F regularly varying:

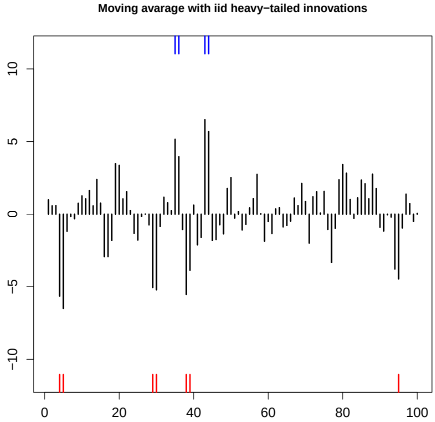
$$X_0/|X_0| \mid |X_0| > u \xrightarrow{d} \begin{cases} +1 & \text{wp } p_+ \\ -1 & \text{wp } p_- \end{cases} \quad \text{and} \quad X_h/|X_0| \mid |X_0| > u \xrightarrow{d} 0$$



Spill-overs: a bunch of spikes

Moving average X_t with iid heavy-tailed innovations Z_t:

$$X_t = Z_{t-1} + Z_t$$



Tail processes of more involved processes

Linear process:

$$X_t = \sum_{j \in \mathbb{Z}} a_j Z_{t-j}$$

$\implies$ Tail process dominated by biggest-shock heuristic

[Davis & Resnick 1985; Meinguet & S. 2010]

Markov process, e.g. stochastic recurrence equation:

$$X_t = A_t X_{t-1} + B_t$$

$\implies$ Tail process is multiplicative random walk

[Smith 1992; Perfekt 1994; Yun 1998; S. 2007; Janßen & S. 2014]

Change of measure due to a time-shift

Recall $\mathcal{L}\big((X_h / |X_0|)_h \mid |X_0| > u\big) \xrightarrow{d} (\Theta_h)_h$

For all integer $s \le 0 \le t$, all $k \in \mathbb{Z}$, all measurable f such that $f(y_s, \ldots, y_t) = 0$ whenever $y_0 = 0$:

$$\mathrm{E}[f(\Theta_{s-k}, \ldots, \Theta_{t-k})] = \mathrm{E}\left[f\left(\frac{\Theta_s}{|\Theta_k|}, \ldots, \frac{\Theta_t}{|\Theta_k|}\right) |\Theta_k|^\alpha\right]$$

Consequence of stationarity of $(X_t)_{t \in \mathbb{Z}}$ and polar decomposition.

Special case:

$$\mathrm{P}[\Theta_h > x] = \mathrm{E}[|\Theta_{-h}|^\alpha \, \mathbf{1}(\Theta_0 / |\Theta_{-h}| > x)], \qquad x > 0.$$

'forward' versus 'backward' estimators of law of Θ_h.

Extends to abstract metric spaces endowed with scalar multiplication and a norm-like functional. Perhaps even for more general index sets endowed with a group action.

[Hult & Lindskog 2006; Meinguet, S. & Zhao 2016]

Extremes of Time Series

Probabilistic background

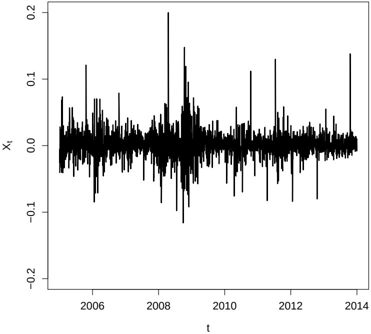
Estimation

Estimating Θ_0

$$X_0/|X_0| \mid |X_0| > x \xrightarrow{d} \Theta_0 = \begin{cases} +1 & \text{wp } p_+ \\ -1 & \text{wp } p_- \end{cases}$$

$$\hat{p}_{+n} = \frac{\sum_{i=1}^{n} \mathbf{1}(X_i > u_n)}{\sum_{i=1}^{n} \mathbf{1}(|X_i| > u_n)}$$

- For suitable sequence $u_n \to \infty$ such that $n\mathrm{P}[|X_0| > u_n] \to \infty$.

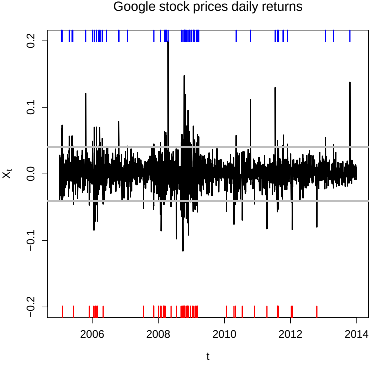


Estimating Θ_0

$$X_0/|X_0| \mid |X_0| > x \xrightarrow{d} \Theta_0 = \begin{cases} +1 & \text{wp } p_+ \\ -1 & \text{wp } p_- \end{cases}$$

$$\hat{p}_{+n} = \frac{\sum_{i=1}^{n} \mathbf{1}(X_i > u_n)}{\sum_{i=1}^{n} \mathbf{1}(|X_i| > u_n)}$$

- For suitable sequence $u_n \to \infty$ such that $n\mathrm{P}[|X_0| > u_n] \to \infty$.
- Example: u_n is a 95% quantile of $|X_t|$.

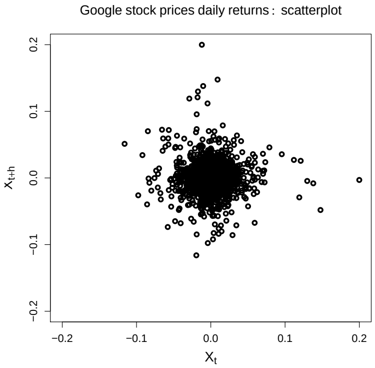


Estimating Θ_h

$$X_h / |X_0| \mid |X_0| > u \xrightarrow{d} \Theta_h$$

$$\hat{F}_n^{\Theta_h}(x) = \frac{\sum_{t=1}^n \mathbf{1}\left(X_{t+h}/|X_t| \le x,\ |X_t| > u_n\right)}{\sum_{t=1}^n \mathbf{1}\left(|X_t| > u_n\right)}$$

- For suitable sequence $u_n \to \infty$ such that $n\mathrm{P}[|X_0| > u_n] \to \infty$.
- Example: $h = 1$, $x = 1$...

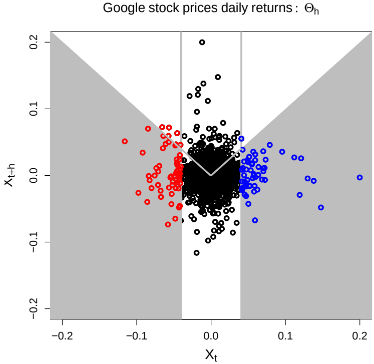


Estimating Θ_h

$$X_h / |X_0| \mid |X_0| > u \xrightarrow{d} \Theta_h$$

$$\hat{F}_n^{\Theta_h}(x) = \frac{\sum_{t=1}^{n} \mathbf{1}\left(X_{t+h}/|X_t| \leq x,\ |X_t| > u_n\right)}{\sum_{t=1}^{n} \mathbf{1}\left(|X_t| > u_n\right)}$$

- For suitable sequence $u_n \to \infty$ such that $n\mathrm{P}[|X_0| > u_n] \to \infty$.
- Example: $h = 1$, $x = 1$, and u_n is a 95% quantile of $|X_t|$.

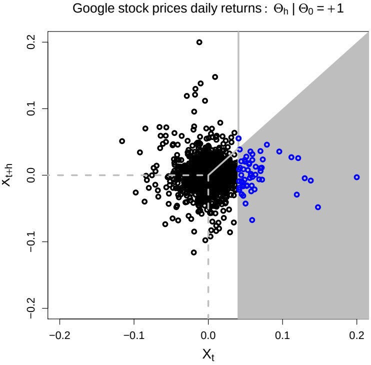


Estimating Θ_h

$$X_h / |X_0| \mid X_0 > u \xrightarrow{d} \Theta_h \mid \Theta_0 = +1$$

$$\hat{F}_n^{\Theta_h|\Theta_0=1}(x) = \frac{\sum_{t=1}^{n} \mathbf{1}\left(X_{t+h}/|X_t| \le x,\ X_t > u_n\right)}{\sum_{t=1}^{n} \mathbf{1}\left(X_t > u_n\right)}$$

- For suitable sequence $u_n \to \infty$ such that $n\mathrm{P}[|X_0| > u_n] \to \infty$.
- Example: $h = 1$, $x = 1$, and u_n is a 95% quantile of $|X_t|$.

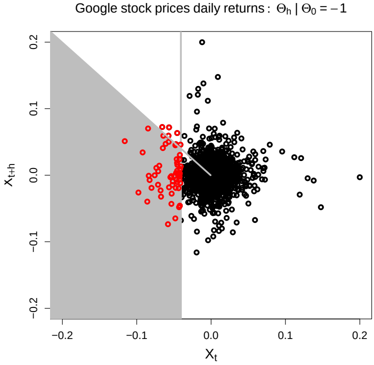


Estimating Θ_h

$$X_h / |X_0| \mid X_0 < -u \xrightarrow{d} \Theta_h \mid \Theta_0 = -1$$

$$\hat{F}_n^{\Theta_h | \Theta_0 = -1}(x) = \frac{\sum_{t=1}^{n} \mathbf{1}\left(X_{t+h}/|X_t| \leq x,\ X_t < -u_n\right)}{\sum_{t=1}^{n} \mathbf{1}\left(X_t < -u_n\right)}$$

- For suitable sequence $u_n \to \infty$ such that $n\mathrm{P}[|X_0| > u_n] \to \infty$.
- Example: $h = 1$, $x = 1$, and u_n is a 95% quantile of $|X_t|$.



Enhancements

Possibly better estimation by exploiting time-change formula: backward estimator [Drees, S. and Warchoł 2015]

Confidence intervals:

- by stationary bootstrap [Politis & Romano 1994]
- or by multiplier block bootstrap.

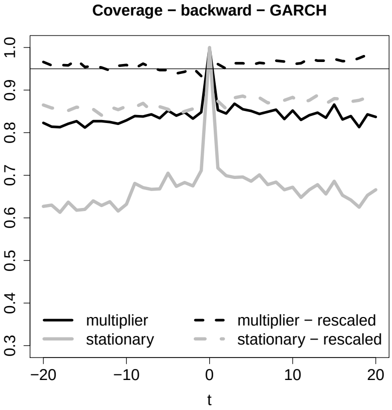
Better coverage via lower thresholds and 'upscaling' [Drees 2015].

Asymptotic justification: empirical processes of cluster functionals [Drees & Rootzén 2010; Drees 2015; Davis, Drees, S. & Warchoł 2016].

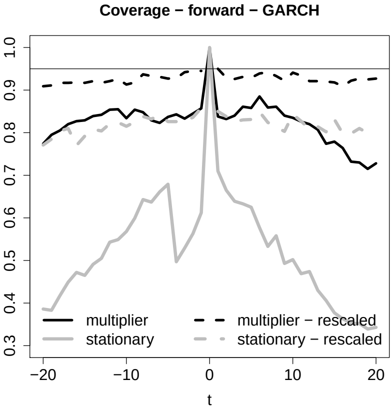
Estimators and resampling

pseudo-random samples from GARCH(1,1) with t_4 innovations

Coverage probabilities of confidence intervals for $\mathrm{P}[|\Theta_t| > 1]$
threshold at 98% empirical quantile



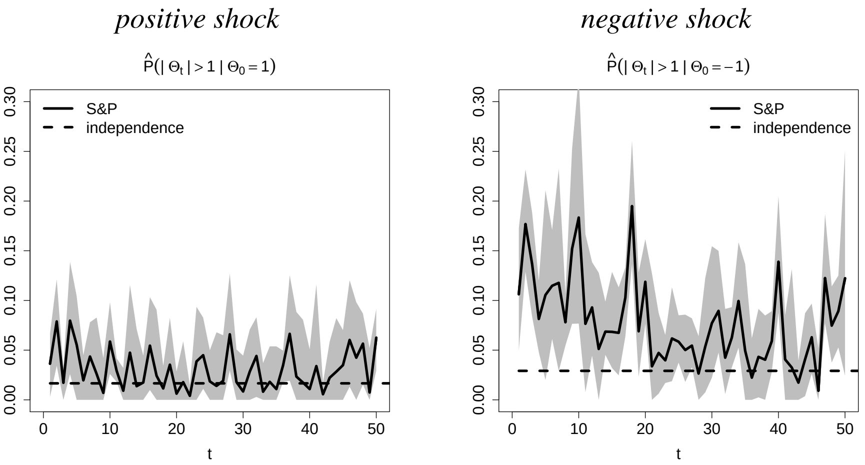
forward estimator



backward estimator

Leverage effect: empirical evidence

S&P500 daily returns (1995–2004)



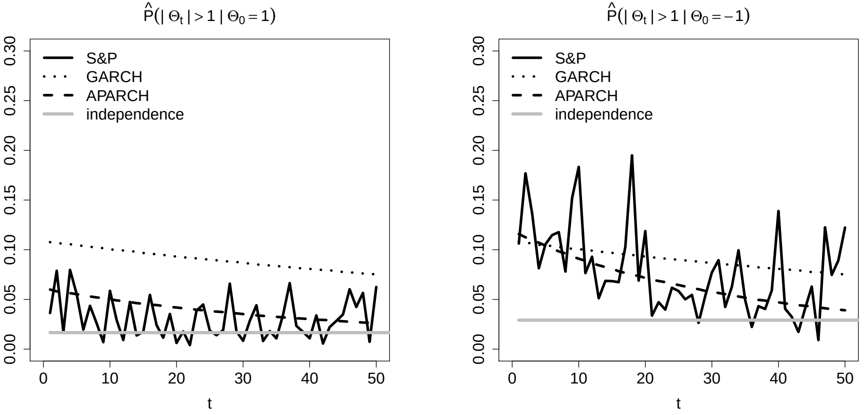
Leverage effect: model fit

S&P500 daily returns: GARCH(1,1) versus APARCH(1,1) $X = \sigma_t Z_t$ with

$$\sigma_t^2 = \omega + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \qquad \text{GARCH}$$

$$\sigma_t^\delta = \omega + \alpha_1 (|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \beta_1 \sigma_{t-1}^\delta \qquad \text{APARCH}$$ [Ding, Granger, Engle 1993]

positive shock *negative shock*



Conclusion: Time series extremes

- Interesting probabilistic structures
  - generalized extreme-value distributions
  - point processes
  - clusters of extremes
  - tail processes
- Challenging statistical questions:
  - Methods: (non/semi-)parametric estimators, resampling
  - Asymptotic justification: empirical processes

Thank you!