Limit theorems for unbounded cluster functionals

Rafał Kulik

Department of Mathematics and Statistics (University of Ottawa)

New Perspectives in the Theory of Extreme Values, Dubrovnik, Croatia

Joint work with Zaoli Chen

24 October 2023

Plan

General motivation

Let X_j be a sequence of random vectors with values in $\mathbb{R}^d$. Let $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ be a *cluster functional*. Let $r_n \to \infty$ be a sequence of integers (*block size*) and let $u_n \to \infty$ (*threshold*).

General motivation

Let X_j be a sequence of random vectors with values in $\mathbb{R}^d$. Let $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ be a *cluster functional*. Let $r_n \to \infty$ be a sequence of integers (*block size*) and let $u_n \to \infty$ (*threshold*). We are interested in the limiting behaviour of

$$\mathbb{E}\left[H\left(\frac{X_1,\dots,X_{r_n}}{u_n}\right)\right] = \mathbb{E}\left[H\left(\frac{X_1,\dots,X_{r_n}}{u_n}\right)\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] ,$$

where $\boldsymbol{x}_{i,j} = (x_i,\dots,x_j)$, $\boldsymbol{x}^*_{i,j} = \max_{k=i,\dots,j}|x_k|$, with $|\cdot|$ being a norm on $\mathbb{R}^d$.

General motivation

Let X_j be a sequence of random vectors with values in $\mathbb{R}^d$. Let $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ be a *cluster functional*. Let $r_n \to \infty$ be a sequence of integers (*block size*) and let $u_n \to \infty$ (*threshold*). We are interested in the limiting behaviour of

$$\mathbb{E}\left[H\left(\frac{X_1,\dots,X_{r_n}}{u_n}\right)\right] = \mathbb{E}\left[H\left(\frac{X_1,\dots,X_{r_n}}{u_n}\right)\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] ,$$

where $\boldsymbol{x}_{i,j} = (x_i,\dots,x_j)$, $\boldsymbol{x}^*_{i,j} = \max_{k=i,\dots,j} |x_k|$, with $|\cdot|$ being a norm on $\mathbb{R}^d$. Why unbounded functionals?

- Inhomogeneous random graphs;
- Blocks estimators for *cluster indices* stationary regularly varying time series;
- Statistical inference for stationary, continuous-time, regularly varying processes.

Cluster indices - extremal index

Let $\boldsymbol{X} = \{X_j, j \in \mathbb{Z}\}$ be a nonnegative, stationary regularly varying sequence. Then

$$\lim_{n\to\infty} \mathbb{P}\left(\max\{X_1, \ldots, X_n\} \le a_n x\right) = \exp(-\theta x^{-\alpha}) \ ,$$

where $\theta \in (0, 1]$ is called the *extremal index* (whenever exists).

Cluster indices - extremal index

Let $\boldsymbol{X} = \{X_j, j \in \mathbb{Z}\}$ be a nonnegative, stationary regularly varying sequence. Then

$$\lim_{n\to\infty} \mathbb{P}\left(\max\{X_1,\ldots,X_n\} \le a_n x\right) = \exp(-\theta x^{-\alpha}) ,$$

where $\theta \in (0,1]$ is called the *extremal index* (whenever exists). The (candidate) extremal index can be represented as

$$\theta = \lim_{n\to\infty} \frac{\mathbb{P}(\max\{X_1,\ldots,X_{r_n}\} > u_n)}{r_n \mathbb{P}(X_0 > u_n)} = \lim_{n\to\infty} \frac{\mathbb{P}(\boldsymbol{X}^*_{1,r_n} > u_n)}{r_n w_n} ,$$

where

$$r_n \to \infty , \quad r_n/n \to 0 , \quad u_n \to \infty , \quad r_n w_n \to 0 , \quad n w_n \to \infty .$$

Cluster indices - extremal index

Let $\boldsymbol{X} = \{X_j, j \in \mathbb{Z}\}$ be a nonnegative, stationary regularly varying sequence. Then

$$\lim_{n\to\infty} \mathbb{P}\left(\max\{X_1, \ldots, X_n\} \le a_n x\right) = \exp(-\theta x^{-\alpha}) \,,$$

where $\theta \in (0, 1]$ is called the *extremal index* (whenever exists). The (candidate) extremal index can be represented as

$$\theta = \lim_{n\to\infty} \frac{\mathbb{P}(\max\{X_1, \ldots, X_{r_n}\} > u_n)}{r_n \mathbb{P}(X_0 > u_n)} = \lim_{n\to\infty} \frac{\mathbb{P}(\boldsymbol{X}^*_{1,r_n} > u_n)}{r_n w_n} \,,$$

where

$$r_n \to \infty \,, \quad r_n/n \to 0 \,, \quad u_n \to \infty \,, \quad r_n w_n \to 0 \,, \quad n w_n \to \infty \,.$$

We will refer to $r_n w_n \to 0$ as "Peak-over-Threshold (PoT) condition".

Cluster indices

Note that we can-rewrite

$$\theta = \lim_{n\to\infty} \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n w_n}$$

with the cluster functional $H : \mathbb{R}_+^{\mathbb{Z}} \to \mathbb{R}$:

$$H(\boldsymbol{x}) = \mathbb{1}\left\{\max_{j\in\mathbb{Z}} x_j > 1\right\} = \mathbb{1}\{\boldsymbol{x}^* > 1\} .$$

- the cluster size distribution obtained with

$$H(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\} = m\right\}, \quad m \in \mathbb{N} ;$$

- (***) the large deviation index of a univariate time series obtained with [1]

$$H(\boldsymbol{x}) = \mathbb{1}\{K(\boldsymbol{x}) > 1\} , \quad K(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} x_j ;$$

[1] Mikosch and Wintenberger (2013, 2014)

Cluster indices - existence and representation

When does the cluster index exist? We need assumptions on r_n, u_n; time series; and functionals H.

[2] Davis and Hsing (1995)

[3] Kulik, Soulier, Wintenberger (2019)

Cluster indices - existence and representation

When does the cluster index exist? We need assumptions on r_n, u_n; time series; and functionals H.

- Let $r_n w_n = r_n \mathbb{P}(X_0 > u_n) \to 0$;

[2] Davis and Hsing (1995)

[3] Kulik, Soulier, Wintenberger (2019)

Cluster indices - existence and representation

When does the cluster index exist? We need assumptions on r_n, u_n; time series; and functionals H.

- Let $r_n w_n = r_n \mathbb{P}(X_0 > u_n) \to 0$;
- Anticlustering condition (extremes cannot persists for infinite horizon time): We say that Condition $\mathcal{AC}(r_n, c_n)$ holds if for every $x, y \in (0, \infty)$, [2]

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{\ell \le |j| \le r_n} X_j > u_n x \mid X_0 > u_n y \right) = 0 \,.$$

[2] Davis and Hsing (1995)

[3] Kulik, Soulier, Wintenberger (2019)

Cluster indices - existence and representation

When does the cluster index exist? We need assumptions on r_n, u_n; time series; and functionals H.

- Let $r_n w_n = r_n \mathbb{P}(X_0 > u_n) \to 0$;
- Anticlustering condition (extremes cannot persists for infinite horizon time): We say that Condition $\mathcal{AC}(r_n, c_n)$ holds if for every $x, y \in (0, \infty)$, [2]

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{\ell \le |j| \le r_n} X_j > u_n x \mid X_0 > u_n y \right) = 0 .$$

  The condition is valid for e.g. geometrically ergodic Markov chains. [3]

[2] Davis and Hsing (1995)

[3] Kulik, Soulier, Wintenberger (2019)

Define:

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n\mathbb{P}(X_0 > u_n)} = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n w_n} .$$

[4] Basrak and Segers (2009), Basrak, Planinić and Soulier (2018); Chapter VI of Kulik and Soulier (2020)

Define:

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n\mathbb{P}(X_0 > u_n)} = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n w_n} .$$

Let $\boldsymbol{Y}$ be the tail process:

$$\frac{\boldsymbol{X}}{x} \mid X_0 > x \xrightarrow{w} \boldsymbol{Y} \text{ as } x \to \infty .$$

Note that $Y_0 > 1$.

[4] Basrak and Segers (2009), Basrak, Planinić and Soulier (2018); Chapter VI of Kulik and Soulier (2020)

Define:

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n\mathbb{P}(X_0 > u_n)} = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n w_n} \ .$$

Let $\boldsymbol{Y}$ be the tail process:

$$\frac{\boldsymbol{X}}{x} \mid X_0 > x \stackrel{w}{\Longrightarrow} \boldsymbol{Y} \text{ as } x \to \infty \ .$$

Note that $Y_0 > 1$.

Theorem 1

Assume that $\mathcal{AC}(r_n, c_n)$ holds. Then $\boldsymbol{\nu}_n^*(H) \to \boldsymbol{\nu}^*(H)$ for all bounded, continuous (w.r.t $\boldsymbol{Y}$), shift invariant functions H with support separated from $\mathbf{0}$.

[4]

[4] Basrak and Segers (2009), Basrak, Planinić and Soulier (2018); Chapter VI of Kulik and Soulier (2020)

Cluster indices - "proof"

- First jump decomposition (we can include $\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$ for free)

$$\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)] = \sum_{j=1}^{r_n} \mathbb{E}[H(\boldsymbol{X}_{j,r_n}/u_n)\mathbb{1}\{\boldsymbol{X}^*_{1,j-1} \le u_n\}\mathbb{1}\{X_j > u_n\}] \,.$$

- Stationarity and conditioning:

$$w_n \sum_{j=1}^{r_n} \mathbb{E}[H(\boldsymbol{X}_{-j,r_n-j}/u_n)\mathbb{1}\{\boldsymbol{X}^*_{1-j,-1} \le u_n\} \mid X_0 > u_n]$$

- Divide by w_n and r_n to get

$$\int_0^1 g_n(s)\mathrm{d}s$$

with
$$g_n(s) = \mathbb{E}\left[H(\boldsymbol{X}_{-[r_n s],r_n-[r_n s]}/u_n)\mathbb{1}\left\{\boldsymbol{X}^*_{1-[r_n s],-1} \le u_n\right\} \mid X_0 > u_n\right].$$

Cluster indices - proof and representation

- Use the weak convergence to the tail process along with the anticlustering condition to get (uniformly in $s \in [0,1]$)

$$g_n(s) \to \mathbb{E}[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}] .$$

[5] Planinić (2022), Last (2023)

Cluster indices - proof and representation

- Use the weak convergence to the tail process along with the anticlustering condition to get (uniformly in $s \in [0,1]$)

$$g_n(s) \to \mathbb{E}[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \leq 1\}] .$$

We have

$$\boldsymbol{\nu}^*(H) = \mathbb{E}\left[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \leq 1\}\right] = \theta\mathbb{E}\left[H(\boldsymbol{Z})\right] ,$$

where $\boldsymbol{Z}$ is a Palm version of $\boldsymbol{Y}$. [5]

[5] Planinić (2022), Last (2023)

Cluster indices - proof and representation

- Use the weak convergence to the tail process along with the anticlustering condition to get (uniformly in $s \in [0,1]$)

$$g_n(s) \to \mathbb{E}[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}] .$$

We have

$$\boldsymbol{\nu}^*(H) = \mathbb{E}\left[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}\right] = \theta\mathbb{E}\left[H(\boldsymbol{Z})\right] ,$$

where $\boldsymbol{Z}$ is a Palm version of $\boldsymbol{Y}$. [5]

Example 2

- With $H(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$ we get

$$\theta = \boldsymbol{\nu}^*(H) = \mathbb{E}\left[\mathbb{1}\{\boldsymbol{Y}^* > 1\}\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}\right] = \mathbb{P}(\boldsymbol{Y}^*_{-\infty,-1} \le 1) ,$$

[5] Planinić (2022), Last (2023)

Unbounded cluster functionals - continuous-time processes

For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $\mathcal{E}(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\}$.

[6] Soulier (2021)

Unbounded cluster functionals - continuous-time processes

For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $\mathcal{E}(\boldsymbol{x}) = \sum_{j \in \mathbb{Z}} \mathbb{1}\{x_j > 1\}$. We can write

$$\theta = \mathbb{E}\left[\frac{1}{\mathcal{E}(\boldsymbol{Y})}\right] = \mathbb{E}\left[\frac{1}{\sum_{j \in \mathbb{Z}} \mathbb{1}\{Y_j > 1\}}\right] \in (0, 1] \,.$$

[6] Soulier (2021)

Unbounded cluster functionals - continuous-time processes

For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $\mathcal{E}(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\}$. We can write

$$\theta = \mathbb{E}\left[\frac{1}{\mathcal{E}(\boldsymbol{Y})}\right] = \mathbb{E}\left[\frac{1}{\sum_{j\in\mathbb{Z}} \mathbb{1}\{Y_j > 1\}}\right] \in (0,1] \,.$$

Let now $\widetilde{\boldsymbol{X}} = \{\widetilde{X}(t), t \in \mathbb{R}\}$ be a continuous time, stationary regularly varying process and let $\widetilde{\boldsymbol{Y}}$ be the corresponding tail process. For $\boldsymbol{x} = \{x(t), t \in \mathbb{R}\}$, let $\mathcal{E}_c(\boldsymbol{x}) = \int_{\mathbb{R}} \mathbb{1}\{x(t) > 1\}\mathrm{d}t$. Then [6]

$$\theta_c = \mathbb{E}\left[\frac{1}{\mathcal{E}_c(\widetilde{\boldsymbol{Y}})}\right] \in (0,\infty] \,.$$

[6] Soulier (2021)

Unbounded cluster functionals - continuous-time processes

For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $\mathcal{E}(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\}$. We can write

$$\theta = \mathbb{E}\left[\frac{1}{\mathcal{E}(\boldsymbol{Y})}\right] = \mathbb{E}\left[\frac{1}{\sum_{j\in\mathbb{Z}} \mathbb{1}\{Y_j > 1\}}\right] \in (0,1] \,.$$

Let now $\widetilde{\boldsymbol{X}} = \{\widetilde{X}(t), t \in \mathbb{R}\}$ be a continuous time, stationary regularly varying process and let $\widetilde{\boldsymbol{Y}}$ be the corresponding tail process. For $\boldsymbol{x} = \{x(t), t \in \mathbb{R}\}$, let $\mathcal{E}_c(\boldsymbol{x}) = \int_{\mathbb{R}} \mathbb{1}\{x(t) > 1\}\mathrm{d}t$. Then [6]

$$\theta_c = \mathbb{E}\left[\frac{1}{\mathcal{E}_c(\widetilde{\boldsymbol{Y}})}\right] \in (0,\infty] \,.$$

Implication: let $X_j := \widetilde{X}(t_j)$ be a sample from $\widetilde{\boldsymbol{X}}$. Estimation of the extremal index θ_c involves the unbounded functional.

[6] Soulier (2021)

Unbounded cluster functional - blocks estimators

Let $m_n = [n/r_n]$ and consider the disjoint and sliding estimators of $\boldsymbol{\nu}^*(H)$:

$$\widetilde{\mathrm{DB}}_n(H) = \frac{1}{n\mathbb{P}(X_0 > u_n)} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/u_n) \ ,$$

$$\widetilde{\mathrm{SB}}_n(H) = \frac{1}{nr_n\mathbb{P}(X_0 > u_n)} \sum_{i=1}^{n-r_n} H(\boldsymbol{X}_{i,r_n+i-1}/u_n) \ .$$

Unbounded cluster functional - blocks estimators

Let $m_n = [n/r_n]$ and consider the disjoint and sliding estimators of $\boldsymbol{\nu}^*(H)$:

$$\widetilde{\mathrm{DB}}_n(H) = \frac{1}{n\mathbb{P}(X_0 > u_n)} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/u_n) ,$$

$$\widetilde{\mathrm{SB}}_n(H) = \frac{1}{nr_n\mathbb{P}(X_0 > u_n)} \sum_{i=1}^{n-r_n} H(\boldsymbol{X}_{i,r_n+i-1}/u_n) .$$

Note that

$$\boldsymbol{\nu}^*(H) = \lim_{n\to\infty} \mathbb{E}[\widetilde{\mathrm{DB}}_n(H)] = \lim_{n\to\infty} \mathbb{E}[\widetilde{\mathrm{SB}}_n(H)] .$$

Unbounded cluster functional - blocks estimators

Theorem 3

Let $\{X_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying univariate time series. Under the "appropriate" conditions

$$\sqrt{nw_n}\left\{\widetilde{\mathrm{DB}}_n(H) - \boldsymbol{\nu}^*(H)\right\} \xrightarrow{\mathrm{d}} \mathbb{G}(H)\;,$$

where $\mathbb{G}$ is a centered Gaussian process with the variance $\boldsymbol{\nu}^*(H^2)$.

The same asymptotics holds also for sliding blocks. [7] **This is in contrast to Block Maxima method.** [8]

[7] Drees and Rootzen (2010); Chapter X of Kulik and Soulier (2020); Cissokho and Kulik (2021,2022); Drees and Neblung (2021)

[8] Bücher and Segers (2018a, 2018b)

Unbounded cluster functional - blocks estimators

Goal: expand

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) .$$

Unbounded cluster functional - blocks estimators

Goal: expand

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) \ .$$

The expansion has the form

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) \approx \frac{1}{n r_n w_n} \mathcal{IC}(H) + \text{the remainder} \ ,$$

where $\mathcal{IC}$ is given by (approximately)

Unbounded cluster functional - blocks estimators

Goal: expand

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) .$$

The expansion has the form

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) \approx \frac{1}{n r_n w_n} \mathcal{IC}(H) + \text{the remainder} ,$$

where $\mathcal{IC}$ is given by (approximately)

$$\mathcal{IC}(H) = \sum_{j=1}^{m_n} \mathcal{IC}_j(H) ,$$

with

Unbounded cluster functional - blocks estimators

- $\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$;

Unbounded cluster functional - blocks estimators

- $\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$;
- $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\}$;

Unbounded cluster functional - blocks estimators

- $\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$;
- $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\}$;
- $T_i(\boldsymbol{x})$ are the locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$;

Unbounded cluster functional - blocks estimators

- $\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$;
- $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\}$;
- $T_i(\boldsymbol{x})$ are the locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$;
- **Cluster length functional**

$$\mathcal{L}(\boldsymbol{x}) = T_{\max}(\boldsymbol{x}) - T_{\min}(\boldsymbol{x}) + 1 .$$

Unbounded cluster functional - blocks estimators

- $\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$;
- $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\}$;
- $T_i(\boldsymbol{x})$ are the locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$;
- **Cluster length functional**

$$\mathcal{L}(\boldsymbol{x}) = T_{\max}(\boldsymbol{x}) - T_{\min}(\boldsymbol{x}) + 1 \ .$$

Example 4

Take $H(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$. Then $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{x}) - 1$.

Unbounded cluster functional - blocks estimators

- $\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}$;
- $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\}$;
- $T_i(\boldsymbol{x})$ are the locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$;
- **Cluster length functional**

$$\mathcal{L}(\boldsymbol{x}) = T_{\max}(\boldsymbol{x}) - T_{\min}(\boldsymbol{x}) + 1 \ .$$

Example 4

Take $H(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$. Then $\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{x}) - 1$. **Important: Starting with bounded H, we get unbounded $\widetilde{H}_{\mathcal{IC}}$.**

Jump locations

Note: T_i are neither bounded nor shift-invariant.

Jump locations

Note: T_i are neither bounded nor shift-invariant.

Theorem 5

Assume that $\mathcal{AC}(r_n, c_n)$ hold. Let a random variable U be $\mathrm{Uniform}(0,1)$. Then, conditionally on $\boldsymbol{X}^*_{1,r_n} > u_n$,

$$r_n^{-1}\left(T_{\min}(\boldsymbol{X}_{1,r_n}/u_n), T_{\max}(\boldsymbol{X}_{1,r_n}/u_n)\right) \stackrel{w}{\Longrightarrow} (U, U) \ .$$

Jump locations

Note: T_i are neither bounded nor shift-invariant.

Theorem 5

Assume that $\mathcal{AC}(r_n, c_n)$ hold. Let a random variable U be $\mathrm{Uniform}(0,1)$. Then, conditionally on $\boldsymbol{X}^*_{1,r_n} > u_n$,

$$r_n^{-1}\left(T_{\min}(\boldsymbol{X}_{1,r_n}/u_n), T_{\max}(\boldsymbol{X}_{1,r_n}/u_n)\right) \stackrel{w}{\Longrightarrow} (U, U)\ .$$

Corollary 6

Let $\gamma > 0$. Under the same conditions,

$$\lim_{n\to\infty} \frac{\mathbb{E}[T^{\gamma}_{\min}(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n^{\gamma+1} w_n} = \lim_{n\to\infty} \frac{\mathbb{E}[T^{\gamma}_{\max}(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n^{\gamma+1} w_n} = \frac{1}{\gamma+1}\theta\ .$$

Jump locations

Note: T_i are neither bounded nor shift-invariant.

Theorem 5

Assume that $\mathcal{AC}(r_n, c_n)$ *hold. Let a random variable* U *be* $\mathrm{Uniform}(0,1)$*. Then, conditionally on* $\boldsymbol{X}^*_{1,r_n} > u_n$*,*

$$r_n^{-1}\left(T_{\min}(\boldsymbol{X}_{1,r_n}/u_n),\, T_{\max}(\boldsymbol{X}_{1,r_n}/u_n)\right) \stackrel{w}{\Longrightarrow} (U, U)\ .$$

Corollary 6

Let $\gamma > 0$*. Under the same conditions,*

$$\lim_{n\to\infty} \frac{\mathbb{E}[T^{\gamma}_{\min}(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n^{\gamma+1} w_n} = \lim_{n\to\infty} \frac{\mathbb{E}[T^{\gamma}_{\max}(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n^{\gamma+1} w_n} = \frac{1}{\gamma+1}\theta\ .$$

Application: an alternative estimator of the extremal index (*with a smaller bias*).

Unbounded, shift-invariant functionals

Note: $\mathcal{L}$ is not bounded but shift-invariant. Recall:

Theorem 7

Let

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{\boldsymbol{X}_{1,r_n}^* > u_n\}]}{r_n w_n} .$$

Assume that $\mathcal{AC}(r_n, c_n)$ holds. Then $\boldsymbol{\nu}_n^*(H) \to \boldsymbol{\nu}^*(H)$ for all bounded continuous shift invariant functions H with support separated from $\mathbf{0}$.

Unbounded, shift-invariant functionals

Note: $\mathcal{L}$ is not bounded but shift-invariant. Recall:

Theorem 7

Let

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{\boldsymbol{X}_{1,r_n}^* > u_n\}]}{r_n w_n} .$$

Assume that $\mathcal{AC}(r_n, c_n)$ holds. Then $\boldsymbol{\nu}_n^*(H) \to \boldsymbol{\nu}^*(H)$ for all bounded continuous shift invariant functions H with support separated from $\mathbf{0}$.

We need to extend it to unbounded functionals H.

Unbounded, shift-invariant functionals

Definition 8

Condition $\mathcal{S}_\gamma(r_n, u_n)$ holds if for all $s, t > 0$

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \frac{1}{\mathbb{P}(X_0 > u_n)} \sum_{i=\ell}^{r_n} i^\gamma \mathbb{P}(X_0 > u_n s, X_i > u_n t) = 0 \,. \quad (\mathcal{S}_\gamma(r_n, u_n))$$

Unbounded, shift-invariant functionals

Definition 8

Condition $\mathcal{S}_\gamma(r_n, u_n)$ holds if for all $s, t > 0$

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \frac{1}{\mathbb{P}(X_0 > u_n)} \sum_{i=\ell}^{r_n} i^\gamma \mathbb{P}(X_0 > u_n s, X_i > u_n t) = 0 \, . \quad (\mathcal{S}_\gamma(r_n, u_n))$$

- The stronger anticustering condition can be viewed as uniform integrability condition.
- The condition is roughly equivalent to the small blocks assumption: $r_n^{\gamma+1} w_n \to 0$.
- It can be verified for many time series models.

Unbounded, shift-invariant functionals

Under the small blocks condition, the business is as usual:

Theorem 9

Assume that $\mathcal{S}_\gamma(r_n, u_n)$ holds (hence $r_n^{\gamma+1} w_n \to 0$). Then for any $H \leq \mathcal{L}^\gamma$

$$\lim_{n\to\infty} \frac{1}{r_n w_n} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n) \mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] = \boldsymbol{\nu}^*(H) \ .$$

[9] See Drees and Rootzen (2010) along with the correction note

Unbounded, shift-invariant functionals

Under the small blocks condition, the business is as usual:

Theorem 9

Assume that $\mathcal{S}_\gamma(r_n, u_n)$ holds (hence $r_n^{\gamma+1} w_n \to 0$). Then for any $H \le \mathcal{L}^\gamma$

$$\lim_{n\to\infty} \frac{1}{r_n w_n} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] = \boldsymbol{\nu}^*(H) \,.$$

This is basically the uniform integrability.

[9] See Drees and Rootzen (2010) along with the correction note

Unbounded, shift-invariant functionals

Under the small blocks condition, the business is as usual:

Theorem 9

Assume that $\mathcal{S}_\gamma(r_n, u_n)$ holds (hence $r_n^{\gamma+1} w_n \to 0$). Then for any $H \leq \mathcal{L}^\gamma$

$$\lim_{n\to\infty} \frac{1}{r_n w_n} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n) \mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] = \boldsymbol{\nu}^*(H) \; .$$

This is basically the uniform integrability. In particular, the cluster length (unlike the jump locations) is tight under the limiting conditional law. Moreover, $\mathbb{E}[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mid \boldsymbol{X}^*_{1,r_n} > u_n]$ converges to a finite constant. [9]

[9] See Drees and Rootzen (2010) along with the correction note

Unbounded, shift-invariant functionals

If large blocks are considered, the situation changes:

Theorem 10

Assume that $\mathcal{AC}(r_n, c_n)$ holds and $r_n^{\gamma+1} w_n \to \infty$. Assume that $\boldsymbol{X}$ is mixing. Then

$$\lim_{n\to\infty} \frac{1}{r_n^{\gamma+2} w_n^2} \mathbb{E}\left[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] = \frac{1}{(\gamma+1)(\gamma+2)} \theta^2 \ .$$

Unbounded, shift-invariant functionals

If large blocks are considered, the situation changes:

Theorem 10

Assume that $\mathcal{AC}(r_n, c_n)$ holds and $r_n^{\gamma+1} w_n \to \infty$. Assume that $\boldsymbol{X}$ is mixing. Then

$$\lim_{n\to\infty} \frac{1}{r_n^{\gamma+2} w_n^2} \mathbb{E}\left[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] = \frac{1}{(\gamma+1)(\gamma+2)} \theta^2 \,.$$

Hence, in the large blocks scenario, $\mathbb{E}[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mid \boldsymbol{X}^*_{1,r_n} > u_n] \to \infty$.

Unbounded, shift-invariant functionals

If large blocks are considered, the situation changes:

Theorem 10

Assume that $\mathcal{AC}(r_n, c_n)$ *holds and* $r_n^{\gamma+1} w_n \to \infty$. *Assume that* $\boldsymbol{X}$ *is mixing. Then*

$$\lim_{n\to\infty} \frac{1}{r_n^{\gamma+2} w_n^2} \mathbb{E}\left[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mathbb{1}\{\boldsymbol{X}^*_{1,r_n} > u_n\}\right] = \frac{1}{(\gamma+1)(\gamma+2)} \theta^2 .$$

Hence, in the large blocks scenario, $\mathbb{E}[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mid \boldsymbol{X}^*_{1,r_n} > u_n] \to \infty$. **From the statistical perspective, the disjoint blocks estimators will not be consistent in the large blocks scenario.**

Summary

- We extended vague convergence of clusters to unbounded functionals, non-shift invariant functionals (jump locations are uniform over each block);

Summary

- We extended vague convergence of clusters to unbounded functionals, non-shift invariant functionals (jump locations are uniform over each block);
- We extended vague convergence of clusters to unbounded functionals: dichotomy between small and large blocks;

Summary

- We extended vague convergence of clusters to unbounded functionals, non-shift invariant functionals (jump locations are uniform over each block);
- We extended vague convergence of clusters to unbounded functionals: dichotomy between small and large blocks; "Statistical" implication: inconsistency of the block estimators in the large blocks scenario.

Summary

- We extended vague convergence of clusters to unbounded functionals, non-shift invariant functionals (jump locations are uniform over each block);
- We extended vague convergence of clusters to unbounded functionals: dichotomy between small and large blocks; "Statistical" implication: inconsistency of the block estimators in the large blocks scenario.
- Extension (not discussed here): Conditioning on the event "at least one large value in two consecutive blocks" changes the asymptotic behaviour of clusters.

References

- Z. Chen, R. Kulik. Limit theorems for unbounded cluster functionals of regularly varying time series. *Submitted, arxiv*.
- Z. Chen, R. Kulik. Asymptotic expansions for blocks estimators: PoT framework. *Submitted, arxiv*.

Thank you!!!!