Asymptotic expansion for block estimators in the PoT framework

Rafał Kulik

Department of Mathematics and Statistics (University of Ottawa)

Extreme Value Analysis conference, Milan, Italy

Joint work with Zaoli Chen

30 June 2023

Plan

1. Introduction
   - Cluster indices
   - Estimation of cluster indices
2. Expansion for blocks estimators
   - Internal clusters
   - Internal clusters
   - Expansion result
3. Summary

Motivation - estimation of extremal index

Let $\{X_j^\dagger, j \in \mathbb{Z}\}$ be a regularly varying sequence of i.i.d. nonnegative random variables with the tail distribution function $\overline{F}$.

Motivation - estimation of extremal index

Let $\{X_j^\dagger, j \in \mathbb{Z}\}$ be a regularly varying sequence of i.i.d. nonnegative random variables with the tail distribution function $\overline{F}$. In particular:

- $\lim_{x\to\infty} \overline{F}(tx)/\overline{F}(x) = t^{-\alpha}$ for some $\alpha > 0$ (e.g. Pareto, Student).

Motivation - estimation of extremal index

Let $\{X_j^\dagger, j \in \mathbb{Z}\}$ be a regularly varying sequence of i.i.d. nonnegative random variables with the tail distribution function $\overline{F}$. In particular:

- $\lim_{x\to\infty} \overline{F}(tx)/\overline{F}(x) = t^{-\alpha}$ for some $\alpha > 0$ (e.g. Pareto, Student).
- There exists a sequence $a_n \to \infty$ such that

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

Motivation - estimation of extremal index

Let $\{X_j^\dagger, j \in \mathbb{Z}\}$ be a regularly varying sequence of i.i.d. nonnegative random variables with the tail distribution function $\overline{F}$. In particular:

- $\lim_{x\to\infty} \overline{F}(tx)/\overline{F}(x) = t^{-\alpha}$ for some $\alpha > 0$ (e.g. Pareto, Student).
- There exists a sequence $a_n \to \infty$ such that

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

Let now $\{X_j, j \in \mathbb{Z}\}$ be a stationary regularly varying sequence with the same marginal tail df $\overline{F}$. Then

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

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

The 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

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

The 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

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

Notation: $\boldsymbol{x}_{i,j} = (x_i,\ldots,x_j)$, $\boldsymbol{x}^*_{i,j} = \max\{x_1,\ldots,x_j\}$.

The 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

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

Notation: $\boldsymbol{x}_{i,j} = (x_i,\ldots,x_j)$, $\boldsymbol{x}^*_{i,j} = \max\{x_1,\ldots,x_j\}$.
We will refer to $r_n w_n \to 0$ as "PoT condition".

Motivation - estimation of 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 a 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

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

- a 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)

Cluster indices - existence and representation

Define:

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

[4] Planinic and Soulier (2018); Chapter VI of Kulik and Soulier (2020); Planinic (2022)

Cluster indices - existence and representation

Define:

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

Theorem 1

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 $\mathbf{0}$.

[4] Planinic and Soulier (2018); Chapter VI of Kulik and Soulier (2020); Planinic (2022)

Cluster indices - existence and representation

Define:

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

Theorem 1

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 $\mathbf{0}$.

Representation:

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

where $\boldsymbol{Y}$ is the tail process and $\boldsymbol{Z}$ is a Palm version of $\boldsymbol{Y}$. [4]

[4] Planinic and Soulier (2018); Chapter VI of Kulik and Soulier (2020); Planinic (2022)

Blocks estimators

Define $m_n = [n/r_n]$ and consider the statistic

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

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

Blocks estimators

Define $m_n = [n/r_n]$ and consider the statistic

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

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

Note that

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

Central Limit Theorem(s) for blocks estimators

Theorem 2

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 for both disjoint blocks. [5] **This is in contrast to Block Maxima method** [6]

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

[6] Bücher and Segers (2018a, 2018b)

Expansion for blocks estimators

Goal: expand

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

Expansion for blocks estimators

Goal: expand

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

The expansion will have the form

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) \approx \frac{1}{nr_nw_n}\mathcal{IC}(H) + \frac{1}{nr_nw_n}\mathcal{BC}(H) + \text{smaller terms} ,$$

where $\mathcal{IC}$ and $\mathcal{BC}$ are internal and boundary clusters statistics.

Expansion for blocks estimators

Goal: expand

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

The expansion will have the form

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) \approx \frac{1}{nr_nw_n}\mathcal{IC}(H) + \frac{1}{nr_nw_n}\mathcal{BC}(H) + \text{smaller terms} \; ,$$

where $\mathcal{IC}$ and $\mathcal{BC}$ are internal and boundary clusters statistics. Once we obtain the (precise) rates for $\mathcal{IC}$ and $\mathcal{BC}$, we will get the precise rates for the expansion.

Internal clusters - set-up

- A large value in block j:

$$A_j^c = \{\boldsymbol{X}^*_{(j-1)r_n+1,jr_n} > u_n\} \ .$$

Internal clusters - set-up

- A large value in block j:

$$A_j^c = \{\boldsymbol{X}^*_{(j-1)r_n+1,jr_n} > u_n\} \ .$$

- For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $T_i(\boldsymbol{x})$ be locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$.

Internal clusters - set-up

- A large value in block j:

$$A_j^c = \{\boldsymbol{X}^*_{(j-1)r_n+1,jr_n} > u_n\} \ .$$

- For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $T_i(\boldsymbol{x})$ be locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$.
- **Cluster length functional**

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

Internal clusters - set-up

- A large value in block j:

$$A_j^c = \{\boldsymbol{X}^*_{(j-1)r_n+1,jr_n} > u_n\} \ .$$

- For $\boldsymbol{x} \in \mathbb{R}_+^{\mathbb{Z}}$, let $T_i(\boldsymbol{x})$ be locations of consecutive exceedences over 1 and $\Delta T_i(\boldsymbol{x}) = T_{i+1}(\boldsymbol{x}) - T_i(\boldsymbol{x})$.
- **Cluster length functional**

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

- The number of exceedences over 1: $\mathcal{E}(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\}$.

Internal clusters - definition

Define

$$\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\} \; .$$

Internal clusters - definition

Define

$$\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\} .$$

Example 3

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

Internal clusters - definition

Define

$$\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\} .$$

Example 3

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 $\widetilde{H}_{\mathcal{IC}}$.**

Internal clusters - definition

Define

$$\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\} .$$

Example 3

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 $\widetilde{H}_{\mathcal{IC}}$.**

We have

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

$$\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{A_{j-1} \cap A_j^c \cap A_{j+1}\} .$$

Internal clusters - definition

Define

$$\widetilde{H}_{\mathcal{IC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{E}(\boldsymbol{x})-1} \Delta T_i(\boldsymbol{x})\{H(\boldsymbol{x}_{-\infty,T_i(\boldsymbol{x})}) + H(\boldsymbol{x}_{T_{i+1}(\boldsymbol{x}),\infty}) - H(\boldsymbol{x})\} .$$

Example 3

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 $\widetilde{H}_{\mathcal{IC}}$.**

We have

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

$$\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{A_{j-1} \cap A_j^c \cap A_{j+1}\} .$$

We can view $\mathcal{IC}$ as disjoint blocks statistics acting on $\widetilde{H}_{\mathcal{IC}}$.

Why internal clusters?

Lemma 4

Assume that $\mathcal{AC}(r_n, c_n)$ *hold. Let* T_{first} *and* T_{last} *be the location of the first and the last block of size* r_n*. Let a random variable* U *be* $\mathrm{Uniform}(0,1)$*. Then, conditionally on* A_1^c*,*

$$\left(\frac{T_{\mathrm{first}}}{r_n}, \frac{T_{\mathrm{last}}}{r_n}, \mathbb{1}\{A_0\}, \mathbb{1}\{A_2\}\right) \Longrightarrow (U_1, U_1, 1, 1) \ .$$

Extension of vague convergence

Recall:

Theorem 5

Let

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}]}{r_n\mathbb{P}(X_0 > u_n)} = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}]}{r_n w_n} .$$

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 $\mathbf{0}$.

Extension of vague convergence

Recall:

Theorem 5

Let

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}]}{r_n\mathbb{P}(X_0 > u_n)} = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}]}{r_n w_n} .$$

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 $\widetilde{H}$.

Extension of vague convergence

Definition 6

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

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

Extension of vague convergence

Definition 6

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.

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

Lemma 7

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

$$\lim_{n\to\infty} \frac{1}{r_n w_n} \mathbb{E}\left[\widetilde{H}(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}\right] = \boldsymbol{\nu}^*(\widetilde{H}) \,.$$

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

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

Lemma 7

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

$$\lim_{n\to\infty} \frac{1}{r_n w_n} \mathbb{E}\left[\widetilde{H}(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}\right] = \boldsymbol{\nu}^*(\widetilde{H}) \,.$$

In particular, the cluster length (unlike the jump locations) is tight under the limiting conditional law. Moreover, $\mathbb{E}[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mid A_1^c]$ converges to a finite constant. [7]

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

If large blocks are considered, the situation changes:

Lemma 8

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*

$$\lim_{n\to\infty} \frac{1}{r_n^{\gamma+2} w_n^2} \mathbb{E}\left[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}\right] = \frac{1}{(\gamma+1)(\gamma+2)}\vartheta^2 \, .$$

If large blocks are considered, the situation changes:

Lemma 8

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*

$$\lim_{n\to\infty} \frac{1}{r_n^{\gamma+2} w_n^2} \mathbb{E}\left[\mathcal{L}^{\gamma}(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}\right] = \frac{1}{(\gamma+1)(\gamma+2)}\vartheta^2 .$$

Hence, $\mathbb{E}[\mathcal{L}^{\gamma}(\boldsymbol{X}_{1,r_n}/u_n) \mid A_1^c] \to \infty$.

If large blocks are considered, the situation changes:

Lemma 8

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*

$$\lim_{n\to\infty} \frac{1}{r_n^{\gamma+2} w_n^2} \mathbb{E}\left[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{A_1^c\}\right] = \frac{1}{(\gamma+1)(\gamma+2)} \vartheta^2 .$$

Hence, $\mathbb{E}[\mathcal{L}^\gamma(\boldsymbol{X}_{1,r_n}/u_n) \mid A_1^c] \to \infty$.

From statistical perspective, the disjoint blocks estimators will not be consistent in the large blocks scenario.

Convergence of internal clusters

Recall:

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

$$\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1}\boldsymbol{X}_{(j-1)r_n+1,jr_n})\mathbb{1}\{A_{j-1} \cap A_j^c \cap A_{j+1}\} \ .$$

Proposition 9

Assume that $\mathcal{S}_{2\gamma+3}(r_n, u_n)$ holds. Assume that $\boldsymbol{X}$ is mixing. Let $H \leq \mathcal{L}^\gamma$. Then

$$\frac{\mathcal{IC}(H)}{nw_n} \xrightarrow{\mathbb{P}} \boldsymbol{\nu}^*(\widetilde{H}_{\mathcal{IC}}) \ ,$$

Convergence of internal clusters

Recall:

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

$$\mathcal{IC}_j(H) = \widetilde{H}_{\mathcal{IC}}(u_n^{-1} \boldsymbol{X}_{(j-1)r_n+1,jr_n}) \mathbb{1}\{A_{j-1} \cap A_j^c \cap A_{j+1}\} .$$

Proposition 9

Assume that $\mathcal{S}_{2\gamma+3}(r_n, u_n)$ holds. Assume that $\boldsymbol{X}$ is mixing. Let $H \leq \mathcal{L}^\gamma$. Then

$$\frac{\mathcal{IC}(H)}{n w_n} \xrightarrow{\mathbb{P}} \boldsymbol{\nu}^*(\widetilde{H}_{\mathcal{IC}}) ,$$

- CLT uniform in H is also established.
- Corresponding result (with different rates) for large blocks scenario.

Boundary clusters - definition

We consider

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

with

$$\mathcal{BC}_j(H) = r_n \left\{ H(\boldsymbol{X}_{(j-1)r_n+1,jr_n}) + H(\boldsymbol{X}_{jr_n+1,(j+1)r_n}) - H(\boldsymbol{X}_{(j-1)r_n+1,(j+1)r_n}) \right\} \times \mathbb{1}\left\{A_{j-1} \cap A_j^c \cap A_{j+1}^c \cap A_{j+2}\right\} .$$

Boundary clusters - definition

We consider

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

with

$$\mathcal{BC}_j(H) = r_n \left\{ H(\boldsymbol{X}_{(j-1)r_n+1, jr_n}) + H(\boldsymbol{X}_{jr_n+1,(j+1)r_n}) - H(\boldsymbol{X}_{(j-1)r_n+1,(j+1)r_n}) \right\} \times \mathbb{1}\left\{ A_{j-1} \cap A_j^c \cap A_{j+1}^c \cap A_{j+2} \right\} .$$

Example 10

If $H(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$, then $\mathcal{BC}_j(H) = r_n \mathbb{1}\left\{ A_{j-1} \cap A_j^c \cap A_{j+1}^c \cap A_{j+2} \right\}$.

Why boundary clusters?

In the small blocks scenario, large values occur at the end of one block and beginning of the next block.

Lemma 11

Assume that $\mathcal{S}_1(r_n, u_n)$ (hence $r_n^2 w_n \to 0$) holds. Then

$$\lim_{n\to\infty} \frac{\mathbb{P}(A_1^c \cap A_2^c)}{w_n} = \vartheta \mathbb{E}\left[(\mathcal{L}(\boldsymbol{Z}) - 1)\right] .$$

Why boundary clusters?

In the small blocks scenario, large values occur at the end of one block and beginning of the next block.

Lemma 11

Assume that $\mathcal{S}_1(r_n, u_n)$ *(hence* $r_n^2 w_n \to 0$*) holds. Then*

$$\lim_{n\to\infty} \frac{\mathbb{P}(A_1^c \cap A_2^c)}{w_n} = \vartheta \mathbb{E}\left[(\mathcal{L}(\boldsymbol{Z}) - 1)\right] .$$

If the blocks are large, then the blocks behave like independent.

Lemma 12

Assume that $r_n^2 w_n \to \infty$*. Then*

$$\lim_{n\to\infty} \frac{\mathbb{P}(A_1^c \cap A_2^c)}{r_n^2 w_n^2} = \vartheta^2 .$$

Extension of vague convergence

From Lemmas 11 and 12 one builds convergence of functional, in parallel to internal clusters:

Extension of vague convergence

From Lemmas 11 and 12 one builds convergence of functional, in parallel to internal clusters:

Corollary 13

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

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[\widetilde{H}(u_n^{-1}\boldsymbol{X}_{1,2r_n})\mathbb{1}\{A_1^c \cap A_2^c\}\right]}{w_n} = \vartheta \mathbb{E}\left[(\mathcal{L}(\boldsymbol{Z}) - 1)\widetilde{H}(\boldsymbol{Z})\right] .$$

Convergence of boundary clusters

Let

$$\widetilde{H}_{\mathcal{BC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{L}(\boldsymbol{x})-1} \{H(\boldsymbol{x}) - H(\boldsymbol{x}_{-\infty,i-1}) - H(\boldsymbol{x}_{i,\infty})\} \ .$$

Convergence of boundary clusters

Let

$$\widetilde{H}_{\mathcal{BC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{L}(\boldsymbol{x})-1} \{H(\boldsymbol{x}) - H(\boldsymbol{x}_{-\infty,i-1}) - H(\boldsymbol{x}_{i,\infty})\} \ .$$

Proposition 14

Assume that $\mathcal{S}_{2\gamma+3}(r_n,u_n)$ holds. Assume that $\boldsymbol{X}$ is mixing. Let $H \leq \mathcal{L}^{\gamma}$. Then

$$\frac{\mathcal{BC}(H)}{nw_n} \xrightarrow{\mathbb{P}} \boldsymbol{\nu}^*(\widetilde{H}_{\mathcal{BC}}) \ ,$$

Convergence of boundary clusters

Let

$$\widetilde{H}_{\mathcal{BC}}(\boldsymbol{x}) := \sum_{i=1}^{\mathcal{L}(\boldsymbol{x})-1} \{H(\boldsymbol{x}) - H(\boldsymbol{x}_{-\infty,i-1}) - H(\boldsymbol{x}_{i,\infty})\} \; .$$

Proposition 14

Assume that $\mathcal{S}_{2\gamma+3}(r_n, u_n)$ *holds. Assume that* $\boldsymbol{X}$ *is mixing. Let* $H \leq \mathcal{L}^{\gamma}$*. Then*

$$\frac{\mathcal{BC}(H)}{nw_n} \xrightarrow{\mathbb{P}} \boldsymbol{\nu}^*(\widetilde{H}_{\mathcal{BC}}) \; ,$$

- CLT uniform in H is also established.
- Corresponding result (with different rates) for large blocks scenario.

Expansion result

Theorem 15

- *Under the small blocks scenario*

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) = O_P\left(\frac{1}{r_n}\right) .$$

- *Under the large blocks scenario*

$$\widetilde{\mathrm{DB}}_n(H) - \widetilde{\mathrm{SB}}_n(H) = O_P\left(r_n w_n\right) .$$

Summary

- We explained what happens in PoT framework.

Summary

- We explained what happens in PoT framework.
- We extended vague convergence of clusters to unbounded functionals: dichotomy between small and large blocks.

Summary

- We explained what happens in PoT framework.
- 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 explained what happens in PoT framework.
- 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.
- Conditioning on the event "at least two large values" changes the asymptotic behaviour of clusters.

Summary

- We explained what happens in PoT framework.
- 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.
- Conditioning on the event "at least two large values" changes the asymptotic behaviour of clusters.
- General problem: convergence of $\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n) \mid C_n]$, where $\mathbb{P}(C_n) \to 0$.

Thank you!!!!