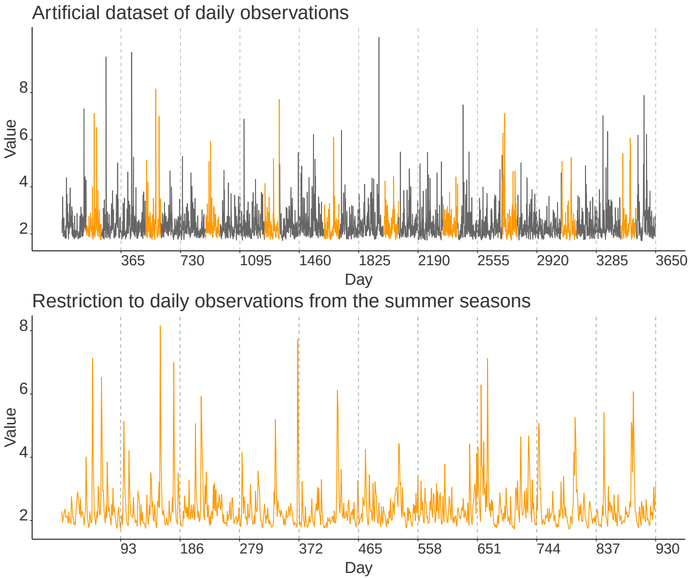
U-statistics of extremes based on disjoint and sliding block maxima

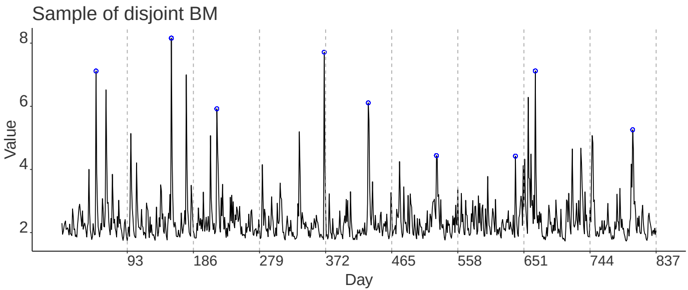
Torben Staud – (Heinrich-Heine-University Düsseldorf)
Axel Bücher – (Heinrich-Heine-University Düsseldorf)

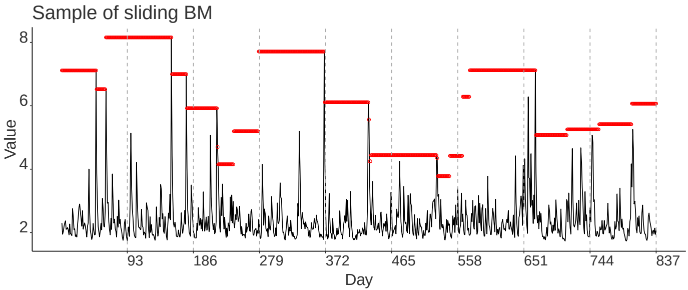
16th German Probability and Statistics Days 2023, Essen
08.03.2023

Extreme observations I



Extreme observations II





Outline

- Statistics of time series extremes
  disjoint and sliding block maxima
- U-statistics
  classic setting and weakly dependent data
- U-statistics of block maxima
  asymptotic normality and comparison of block methods

Statistics of time series extremes

Foundation for block maxima

Theorem (Fisher-Tippett-Gnedenko, 1928-1943): Suppose the X_i are i.i.d. $\sim F$, there are normalizing sequences $a_r > 0, b_r \in \mathbb{R}$ and a non-degenerate limiting distribution G satisifying

$$\frac{M_{1:r} - b_r}{a_r} \xrightarrow[r \to \infty]{d} G,$$

where $M_{1:r} := \max(X_1, \ldots, X_r)$. Then $G \sim \text{GEV}(\mu, \sigma, \gamma)$ for a shape parameter $\gamma \in \mathbb{R}$ depending on F and location-scale parameter $\mu, \sigma \in \mathbb{R} \times \mathbb{R}^+$.

Foundation for block maxima

Theorem (Fisher-Tippett-Gnedenko, 1928-1943): Suppose the X_i are i.i.d. $\sim F$, there are normalizing sequences $a_r > 0, b_r \in \mathbb{R}$ and a non-degenerate limiting distribution G satisifying

$$\frac{M_{1:r} - b_r}{a_r} \xrightarrow[r \to \infty]{d} G,$$

where $M_{1:r} := \max(X_1, \ldots, X_r)$. Then $G \sim \text{GEV}(\mu, \sigma, \gamma)$ for a shape parameter $\gamma \in \mathbb{R}$ depending on F and location-scale parameter $\mu, \sigma \in \mathbb{R} \times \mathbb{R}^+$.

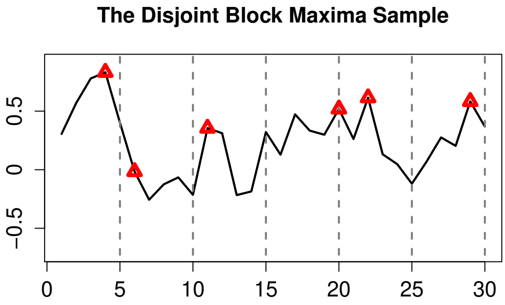
Maximum Domain of Attraction condition: There exist sequences $a_r > 0, b_r \in \mathbb{R}$ and a $\gamma \in \mathbb{R}$ such that

$$\frac{\max\{X_1, \ldots, X_r\} - b_r}{a_r} \xrightarrow[r \to \infty]{d} \text{GEV}(\gamma). \qquad \text{(DoA)}$$

Types of block maxima I

- $X_1, X_2, \ldots, X_n$ excerpt from a stationary time series satisfying (DoA)

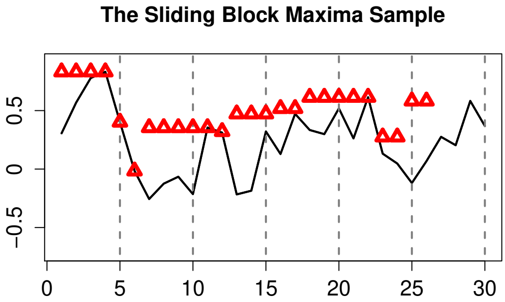
Definition: Define $M_{r,i}^{\text{db}} := \max\left(X_{(i-1)\cdot r+1}, \ldots, X_{r\cdot i}\right)$ as the (i−th) **disjoint block maximum**, for $i = 1, \ldots, n/r$.



Types of block maxima II

- $X_1, X_2, \ldots, X_n$ excerpt from a stationary time series satisfying (DoA)

Definition: Define $M_{r,i}^{\mathrm{sb}} := \max(X_i, \ldots, X_{i+r-1})$ as the (i–th) **sliding block maximum**, for $i = 1, \ldots, n - r + 1$.



U-statistics

U-statistics I

Estimation Problem:

- F unknown c.d.f. from a c.d.f.-class $\mathcal{F}$ and for known $\rho \in \mathbb{N}$, $h\colon \mathbb{R}^\rho \to \mathbb{R}$ one can write:

  $$\theta = \theta(F) = \int \dots \int h(x_1, \dots, x_\rho)\, \mathrm{d}F(x_1), \dots \mathrm{d}F(x_\rho) = \mathrm{E}[h(Z_1, \dots, Z_\rho)],$$

  if $Z_1, \dots Z_\rho \sim F$ i.i.d.

U-statistics I

Estimation Problem:

- F unknown c.d.f. from a c.d.f.-class $\mathcal{F}$ and for known $\rho \in \mathbb{N}$, $h\colon \mathbb{R}^\rho \to \mathbb{R}$ one can write:

$$\theta = \theta(F) = \int \dots \int h(x_1, \dots, x_\rho)\, \mathrm{d}F(x_1), \dots \mathrm{d}F(x_\rho) = \mathrm{E}[h(Z_1, \dots, Z_\rho)],$$

  if $Z_1, \dots Z_\rho \sim F$ i.i.d.

Examples:

- $\theta = \mathrm{E}[Z_1]$ for $h(x) = x, \rho = 1$.
- $\theta = \mathrm{Var}(Z_1)$ for $h(x,y) = (x-y)^2/2, \rho = 2$.
- Probability weighted moments.
- If Z_i are $\mathbb{R}^2-$valued: Covariance, Kendall's τ.

U-statistics I

Estimation Problem:

- F unknown c.d.f. from a c.d.f.-class $\mathcal{F}$ and for known $\rho \in \mathbb{N}$, $h\colon \mathbb{R}^\rho \to \mathbb{R}$ one can write:

$$\theta = \theta(F) = \int \dots \int h(x_1, \dots, x_\rho)\,\mathrm{d}F(x_1), \dots \mathrm{d}F(x_\rho) = \mathrm{E}[h(Z_1, \dots, Z_\rho)],$$

  if $Z_1, \dots Z_\rho \sim F$ i.i.d.

Examples:

- $\theta = \mathrm{E}[Z_1]$ for $h(x) = x, \rho = 1$.
- $\theta = \mathrm{Var}(Z_1)$ for $h(x, y) = (x - y)^2/2, \rho = 2$.
- Probability weighted moments.
- If Z_i are $\mathbb{R}^2-$valued: Covariance, Kendall's τ.

How to estimate θ?

U-statistics II

Definition and Theorem (Hoeffding 1948): For a kernel function $h\colon \mathbb{R}^2 \to \mathbb{R}$

$$U_n := \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} h(Z_i, Z_j)$$

is called **U-statistic** (of order 2 with kernel h). Under a non *degeneracy condition* and if the Z_i are i.i.d. it holds that

$$\sqrt{n}\,\{U_n - \theta\} \rightsquigarrow \mathcal{N}(0, \sigma^2),$$

where σ^2 depends on h and F.

U-statistics II

Definition and Theorem (Hoeffding 1948): For a kernel function $h\colon \mathbb{R}^2 \to \mathbb{R}$

$$U_n := \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} h(Z_i, Z_j)$$

is called **U-statistic** (of order 2 with kernel h). Under a non *degeneracy condition* and if the Z_i are i.i.d. it holds that

$$\sqrt{n}\,\{U_n - \theta\} \rightsquigarrow \mathcal{N}(0, \sigma^2),$$

where σ^2 depends on h and F.

Extensions to dependent data:

- $\star/\psi$-mixing (1972 Sen).
- $\beta-$mixing (1976 Yoshihara).
- $\alpha-$mixing (2010 Dehling, Wendler).

U-statistics of extremes

Object of interest

Plug block maxima into a u-statistic:

$$U_{n,r}^{\mathrm{mb}} := \binom{\tilde{n}}{2}^{-1} \sum_{1 \leq i < j \leq \tilde{n}} h(M_{r,i}^{\mathrm{mb}}, M_{r,j}^{\mathrm{mb}}),$$

where $\mathrm{mb} \in \{\mathrm{db}, \mathrm{sb}\}$, $\tilde{n} =$ number of blocks.

Object of interest

Plug block maxima into a u-statistic:

$$U_{n,r}^{\mathrm{mb}} := \binom{\tilde{n}}{2}^{-1} \sum_{1 \le i < j \le \tilde{n}} h(M_{r,i}^{\mathrm{mb}}, M_{r,j}^{\mathrm{mb}}),$$

where $\mathrm{mb} \in \{\mathrm{db}, \mathrm{sb}\}$, $\tilde{n} =$ number of blocks.

Objectives:

- Asymptotic distribution/variance
- Comparison between disjoint and sliding

Kernel transformation condition

Problem:

Only the rescaled block maxima $Z_{r,i}^{\mathrm{mb}} := (M_{r,i}^{\mathrm{mb}} - b_r)/a_r$ have distributional limits and in general $h(M_{r,i}^{\mathrm{mb}}, M_{r,j}^{\mathrm{mb}}) \neq h(Z_{r,i}^{\mathrm{mb}}, Z_{r,j}^{\mathrm{mb}})$.

Kernel transformation condition

Problem:
Only the rescaled block maxima $Z_{r,i}^{\text{mb}} := (M_{r,i}^{\text{mb}} - b_r)/a_r$ have distributional limits and in general $h(M_{r,i}^{\text{mb}}, M_{r,j}^{\text{mb}}) \neq h(Z_{r,i}^{\text{mb}}, Z_{r,j}^{\text{mb}})$.

Solution:
Suppose h satisfies the following kernel transformation condition:

(Simplified) Kernel condition: There exists a function $f : \mathbb{R}^+ \to \mathbb{R}^*$ such that for $b \in \mathbb{R}, a > 0, m_1, m_2 \in \mathbb{R}$

$$h\left(\frac{m_1 - b}{a}, \frac{m_2 - b}{a}\right) = \frac{h(m_1, m_2)}{f(a)}. \tag{KT}$$

Kernel transformation condition

Problem:
Only the rescaled block maxima $Z_{r,i}^{\mathrm{mb}} := (M_{r,i}^{\mathrm{mb}} - b_r)/a_r$ have distributional limits and in general $h(M_{r,i}^{\mathrm{mb}}, M_{r,j}^{\mathrm{mb}}) \neq h(Z_{r,i}^{\mathrm{mb}}, Z_{r,j}^{\mathrm{mb}})$.

Solution:
Suppose h satisfies the following kernel transformation condition:

(Simplified) Kernel condition: There exists a function $f : \mathbb{R}^+ \to \mathbb{R}^*$ such that for $b \in \mathbb{R}, a > 0, m_1, m_2 \in \mathbb{R}$

$$h\left(\frac{m_1 - b}{a}, \frac{m_2 - b}{a}\right) = \frac{h(m_1, m_2)}{f(a)}. \tag{KT}$$

Examples:

- variance kernel, PWM, Kendall's τ, covariance

Main result

Theorem (Bücher, S., work in progress) Under regularity conditions including (KT) we have for a strictly stationary time series $(X_n)_n$ satisfying (DoA)

$$\sqrt{n/r} \cdot \frac{U_{n,r}^{\mathrm{mb}} - \mathrm{E}[U_{n,r}^{\mathrm{mb}}]}{f(a_r)} \rightsquigarrow \mathcal{N}(0, \sigma_{\mathrm{mb}}^2),$$

and assuming a bias condition:

$$\sqrt{n/r} \cdot \left\{ \frac{U_{n,r}^{\mathrm{mb}}}{f(a_r)} - \theta \right\} \rightsquigarrow \mathcal{N}(B, \sigma_{\mathrm{mb}}^2),$$

where f is from the (KT) condition, B the asymptotic bias and $\theta := \mathrm{E}[h(Z_1, Z_2)]$ with $Z_1, Z_2 \sim \mathrm{GEV}(\gamma)$ i.i.d. σ_{mb}^2 depends on h, mb, and γ from (DoA). Furthermore it holds that $\sigma_{db}^2 \geq \sigma_{sb}^2$.

Main result

Theorem (Bücher, S., work in progress) Under regularity conditions including (KT) we have for a strictly stationary time series $(X_n)_n$ satisfying (DoA)

$$\sqrt{n/r} \cdot \frac{U_{n,r}^{\mathrm{mb}} - \mathrm{E}[U_{n,r}^{\mathrm{mb}}]}{f(a_r)} \rightsquigarrow \mathcal{N}(0, \sigma_{\mathrm{mb}}^2),$$

and assuming a bias condition:

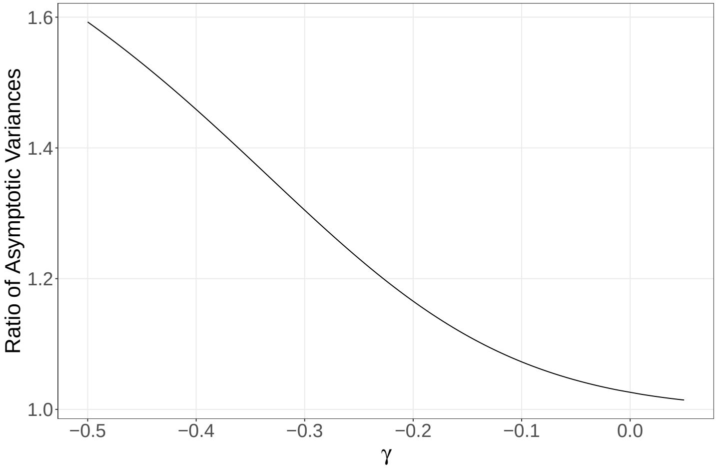
$$\sqrt{n/r} \cdot \left\{ \frac{U_{n,r}^{\mathrm{mb}}}{f(a_r)} - \theta \right\} \rightsquigarrow \mathcal{N}(B, \sigma_{\mathrm{mb}}^2),$$

where f is from the (KT) condition, B the asymptotic bias and $\theta := \mathrm{E}[h(Z_1, Z_2)]$ with $Z_1, Z_2 \sim \mathrm{GEV}(\gamma)$ i.i.d. σ_{mb}^2 depends on h, mb, and γ from (DoA). Furthermore it holds that $\sigma_{db}^2 \geq \sigma_{sb}^2$.

- Most of the time $\sigma_{\mathrm{db}}^2 > \sigma_{\mathrm{sb}}^2$.

Theoretical comparison for the variance-kernel

Variance kernel $h(x, y) = (x - y)^2/2$, plot of ratio $\sigma^2_{\text{db}}/\sigma^2_{\text{sb}}$



Generalizations

- The $X_1, \ldots, X_n$ may be an excerpt of a piecewise stationary time series:

$$(X_1, \ldots, X_n) = (Y_1^{(1)}, \ldots, Y_r^{(1)}, Y_1^{(2)}, \ldots, Y_r^{(2)}, \\ \ldots, Y_1^{(m)}, \ldots, Y_r^{(m)}).$$

- X_i may be multivariate and the kernel $\mathbb{R}^d-$valued.
- U-statistic may be of higher order.

Summary

- Analyzed u-statistics where we plugged in block maxima.
- Established asymptotic normality.
- Sliding is *preferable* over disjoint.

Thank you!

Asymptotic variances

$$\sigma^2_{\text{db}} := 4\,\text{Var}\big(h_1(Z_1)\big),$$

where $Z_1, Z_2 \sim \text{GEV}(\gamma)$ i.i.d. and $h_1(z) := \text{E}[h(z, Z_2)]$.

$$\sigma^2_{\text{sb}} := 8 \int_0^1 \text{Cov}\left(h_1(Z_{1,\xi}, h_1(Z_{2,\xi})\right) \text{d}\xi,$$

where $(Z_{1,\xi}, Z_{2,\xi}) \sim G_{\gamma,\xi}$ and $G_{\gamma,\xi}$ is a bivariate extreme value distribution with $\text{GEV}(\gamma)$ marginals and a certain Pickands-dependence function A_ξ. ξ governs the overlap between $Z_{1,\xi}$ and $Z_{2,\xi}$ meaning that there is an $\big((1-\xi) \wedge 0\big)$%-overlap.

Proof ideas of main result

- Decompose $U_{n,r}$ into projected term

$$A_n := \frac{2}{\widetilde{n}} \sum_{i=1}^{\widetilde{n}} h_{1,r}(Z_{r,i})$$

and degenerate term

$$B_n := \frac{2}{\widetilde{n}(\widetilde{n}-1)} \sum_{1 \le i < j \le \widetilde{n}} h_{2,r}(Z_{r,i}, Z_{r,j}),$$

where

$$h_{1,r}(z) := \mathrm{E}[h(z, Z_{r,1})] - \mathrm{E}[h(Z_{r,1}, Z_{r,2})]$$

and

$$h_{2,r}(x,y) := h(x,y) - h_{1,r}(x) - h_{1,r}(y) - \mathrm{E}[h(Z_{r,1}, Z_{r,2})].$$

- $B_n \xrightarrow{\mathbb{P}} 0$: long range independency, Bradley-coupling, stochastic continuity type arguments .
- $A_n \xrightarrow{d} \mathcal{N}$: blocking of blocks (long range independency), wichura strategy.

Outlook

- Estimating the asymptotic variance.
- Choosing the block size r in finite sample situations.

Literature: u-statistics of extremes

Existing literature

- similar object **extremal u-statistic**: consider $h_r(x_1, \ldots, x_r)$ for $r \to \infty$. (2001 Segers)
- recently (2022 Oorschot, Segers, Zhou) asymptotic theory for extreme U-Statistics.
- no (direct) literature on u-Statistics of block maxima.
- recently (2023 Dehling, Giraudo, Schmidt) investigated u-statistics of sample moments of blocks.

Selected references

Bücher, A. and Zanger, L. (2021+): On the Disjoint and Sliding Block Maxima method for piecewise stationary time series. To appear in: Ann. Statist.

Dehling, H. and Wendler, M. (2010): Central limit theorem and the bootstrap for U-statistics of strongly mixing data. Journal of Mult. Analysis 10 126-137.

Hoeffding, W. (1948): A class of statistics with asymptotically normal distributions. Ann. of Statist, 19: 293–32.

Oorschot, J. et al (2022): Tail inference using extreme U-statistics.

Segers, J. (2001): Extremes of a Random Sample: Limit Theorems and Applications, Katholieke Universiteit Leuven.

Sen, P.K. (1972): Limiting behavior of regular functionals of empirical distributions for stationary *-mixing processes. Z. Wahrsch. verw. Gebiete 25, 71-82

Yoshihara, K. (1976): Limiting Behavior of U-Statistics for Stationary, Absolutely Regular Processes. Z. Wahrsch. verw. Gebiete 35, 237-252