Bootstrapping Block Maxima Estimators

Torben Staud - Ruhr University Bochum
Axel Bücher - Ruhr University Bochum

Computational and Methodological Statistics 2024
King's College London
14 December, 2024

Generated[1] by the prompt: *Create a happy machine kangaroo with a very heavy tail, wearing a shoe in London, trying to pull itself out by the bootstraps in comic style*

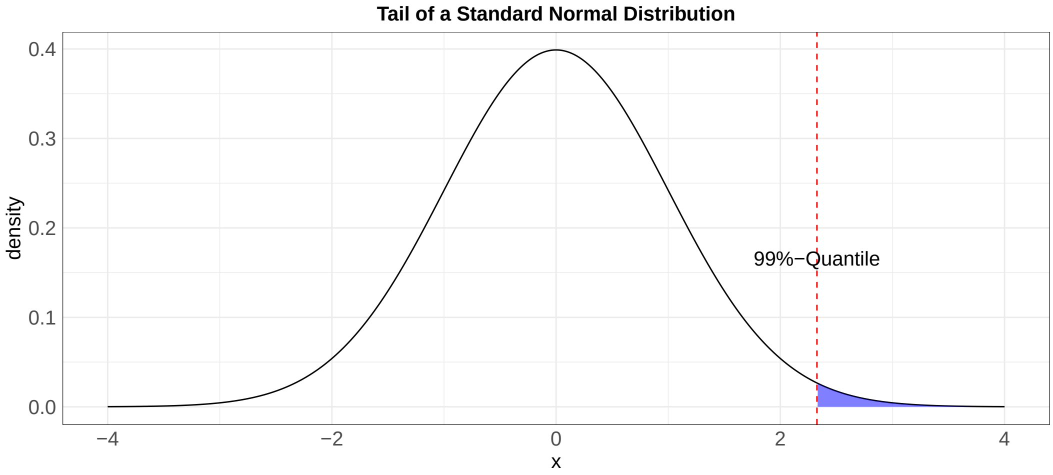
[1] DALL-E

Motivation: Extreme Value Statistics

- Extreme value theory is concerned with the tail of a distribution

Motivation: Extreme Value Statistics

- Extreme value theory is concerned with the tail of a distribution

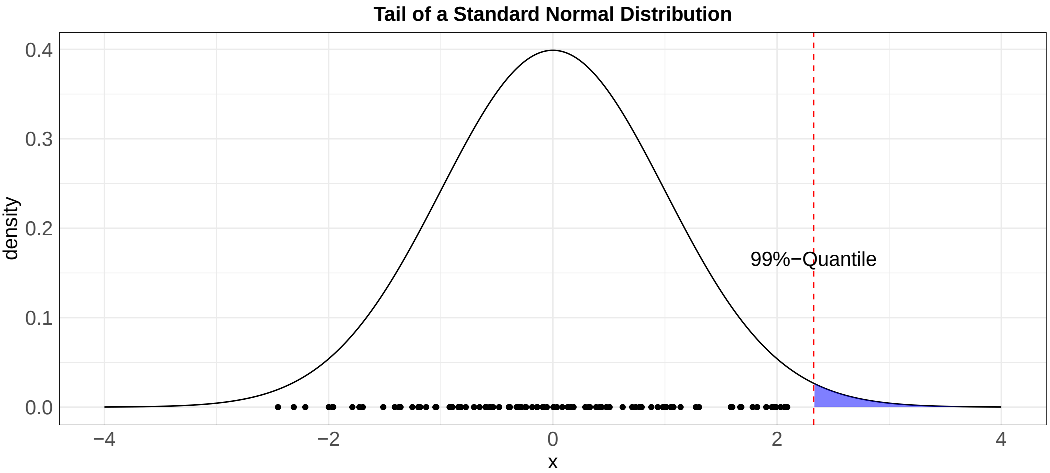


Motivation: Extreme Value Statistics

- Extreme value theory is concerned with the tail of a distribution
- Problem: Bulk of the data is not in the tail

Motivation: Extreme Value Statistics

- Extreme value theory is concerned with the tail of a distribution
- Problem: Bulk of the data is not in the tail



Motivation I: Extreme Value Statistics

- Extreme value theory is concerned with the tail of a distribution
- Problem: Bulk of the data is not in the tail
- Not to ignore: Events may be rare, but have a large impact: e.g. floods and earthquakes

Motivation I: Extreme Value Statistics

- Extreme value theory is concerned with the tail of a distribution
- Problem: Bulk of the data is not in the tail
- Not to ignore: Events may be rare, but have a large impact: e.g. floods and earthquakes

$\rightarrow$ Extreme Value Statistics is of high relevance

Motivation II: why bootstrapping?

Tukey once proposed to call the bootstrap in statistics shotgun as it could blow off the head of every statistical problem if we as statisticians could stand the resulting mess.[2]

[2] See the acknowledgements in Efron (1979)

Motivation II: why bootstrapping?

Tukey once proposed to call the bootstrap in statistics shotgun as it could blow off the head of every statistical problem if we as statisticians could stand the resulting mess.[2]

Indeed bootstrap consistency results often lead to

- confidence intervals
- estimators for asymptotic standard deviations of estimators
- statistical hypothesis tests

[2] See the acknowledgements in Efron (1979)

Motivation II: why bootstrapping?

Tukey once proposed to call the bootstrap in statistics shotgun as it could blow off the head of every statistical problem if we as statisticians could stand the resulting mess.[2]

Indeed bootstrap consistency results often lead to

- confidence intervals
- estimators for asymptotic standard deviations of estimators
- statistical hypothesis tests

$\rightarrow$ **Aim: use bootstraps in extremes when asymptotic variances are out of reach**

[2] See the acknowledgements in Efron (1979)

Motivation III: why block maxima?

- From now on: tail = maximum of a block of a certain size (block length) of the sample: $M_{r,t} = \max(X_t, \ldots, X_{t+r-1})$
- if block size is large: results from extreme value theory[3] state

$$\forall t \colon \mathcal{L}(M_{r,t}) \approx \mathrm{GEV}(\theta_r),$$

  where $\theta_r = (\mu_r, \sigma_r, \gamma) \in \mathbb{R} \times (0, \infty) \times \mathbb{R}$ are the parameters of the *Generalized Extreme Value* distribution
- μ_r, σ_r are location-scale parameters, while γ[4] determines the shape of the distribution

[3] Fisher-Tippett-Gnedenko Theorem
[4] *extreme value index*

Motivation III: why block maxima?

- From now on: tail = maximum of a block of a certain size (block length) of the sample: $M_{r,t} = \max(X_t, \dots, X_{t+r-1})$
- if block size is large: results from extreme value theory[3] state

$$\forall t \colon \mathcal{L}(M_{r,t}) \approx \mathrm{GEV}(\theta_r),$$

  where $\theta_r = (\mu_r, \sigma_r, \gamma) \in \mathbb{R} \times (0, \infty) \times \mathbb{R}$ are the parameters of the *Generalized Extreme Value* distribution
- μ_r, σ_r are location-scale parameters, while γ[4] determines the shape of the distribution

This gives rise to the block maxima methods: Disjoint and Sliding block maxima (more: later)

[3] Fisher-Tippett-Gnedenko Theorem
[4] *extreme value index*

Motivation III: Why bootstrap block maxima?

Variance of block maxima based estimators[5] might look trivial like

[5] Here: estimating the variance of a block maximum based on sliding block maxima

Variance of block maxima based estimators[5] might look trivial like

$$\sigma^2_{\mathrm{sb}} = \begin{cases} \frac{2}{3\gamma^3}\left(-3g_4 I_{2,2} + 8g_1 g_3 I_{2,1} - 6g_1^2 g_2 I_{1,1}\right), & \gamma > 0 \\ \frac{8}{\gamma^2}\left(\Gamma(-4\gamma) I_{2,2} - 2g_1\Gamma(-3\gamma) I_{2,1} + g_1^2\Gamma(-2\gamma) I_{1,1}\right), & \gamma < 0 \\ 2\zeta(3) - 48 - \frac{8}{3}\pi^2 + \frac{32}{3}\log^3(2) - 48\log^2(2) + 96\log(2) + \frac{16}{3}\pi^2\log(2), & \gamma = 0 \end{cases},$$

where $g_j := \Gamma(1 - j\gamma), j < 1/\gamma$;

$$I_{i,k} := \int_0^{1/2} \left(\alpha_{(j+k)\gamma}(w) - 1\right)\left\{w^{-j\gamma-1}(1-w)^{-k\gamma-1} + w^{-k\gamma-1}(1-w)^{-j\gamma-1}\right\} \mathrm{d}w$$

and

$$\alpha_\beta : (0,1) \to (0,\infty), \quad w \mapsto \alpha_\beta(w) = \begin{cases} \frac{1-(1-w)^{\beta+1}}{w(\beta+1)}, & \beta \neq -1 \\ -\frac{\log(1-w)}{w}, & \beta = -1 \end{cases}.$$

[5] Here: estimating the variance of a block maximum based on sliding block maxima

Grab the free lunch and bootstrap sliding block estimators!

Outline

Block maxima:

- Disjoint block maxima
- Sliding block maxima

Bootstrapping (sliding) block maxima:

- Naive approaches
- The circular block maxima approach
- Resampling algorithm
- Consistency results

Block maxima

Block maxima

Basic model assumptions:

- Strictly stationary time series excerpt $\mathcal{X}_n = (X_1, \ldots, X_n)$ with values in $\mathbb{R}$ and c.d.f. F.
- Short range dependency structures allowed
- F in the domain of attraction of G_γ, that is:

$$\mathcal{L}\Big(\max(X_1, \ldots, X_r)\Big) \approx \mathrm{GEV}(\mu_r, \sigma_r, \gamma)$$

Block maxima

Basic model assumptions:

- Strictly stationary time series excerpt $\mathcal{X}_n = (X_1, \ldots, X_n)$ with values in $\mathbb{R}$ and c.d.f. F.
- Short range dependency structures allowed
- F in the domain of attraction of G_γ, that is:

$$\mathcal{L}\Big(\max(X_1, \ldots, X_r)\Big) \approx \mathrm{GEV}(\mu_r, \sigma_r, \gamma)$$

Statistical challenges:

- Estimate γ (many estimators well known: Hill, PWM, (Pseudo-)MLE), extreme quantiles/return levels

Block maxima

Basic model assumptions:

- Strictly stationary time series excerpt $\mathcal{X}_n = (X_1, \ldots, X_n)$ with values in $\mathbb{R}$ and c.d.f. F.
- Short range dependency structures allowed
- F in the domain of attraction of G_γ, that is:

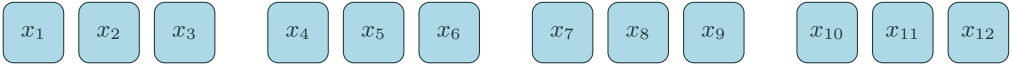
$$\mathcal{L}\Big(\max(X_1, \ldots, X_r)\Big) \approx \mathrm{GEV}(\mu_r, \sigma_r, \gamma)$$

Statistical challenges:

- Estimate γ (many estimators well known: Hill, PWM, (Pseudo-)MLE), extreme quantiles/return levels
- **Confidence intervals/Variance of Estimators** $\leftarrow$ our focus

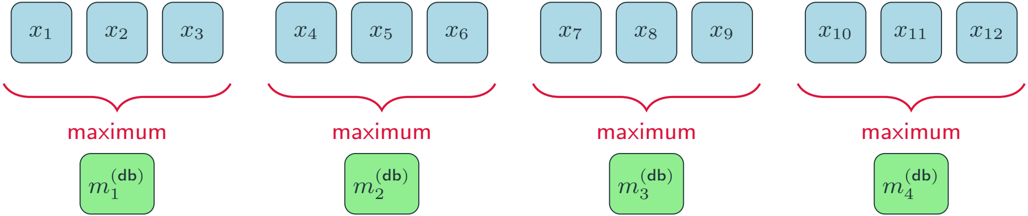
Disjoint block maxima

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



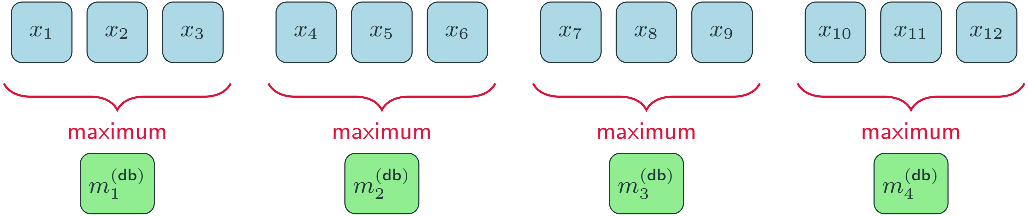
Disjoint block maxima

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



Disjoint block maxima

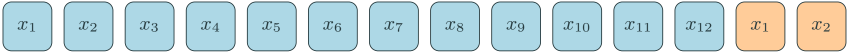
Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



Disjoint block maxima sample $\mathcal{M}^{(\mathbf{db})} = (m_1^{(\mathbf{db})}, \ldots, m_4^{(\mathbf{db})})$

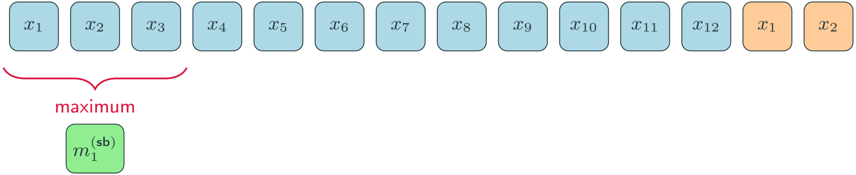
Sliding block maxima

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



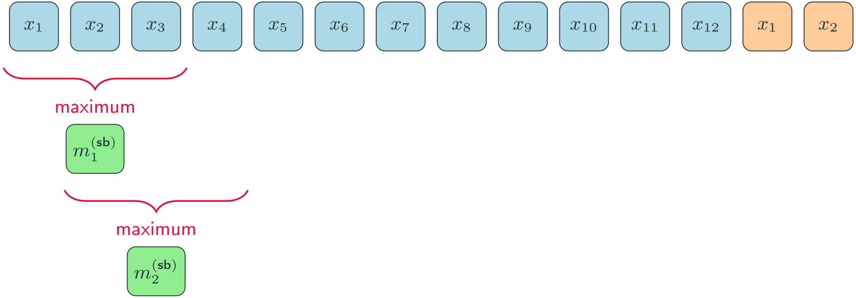
Sliding block maxima

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



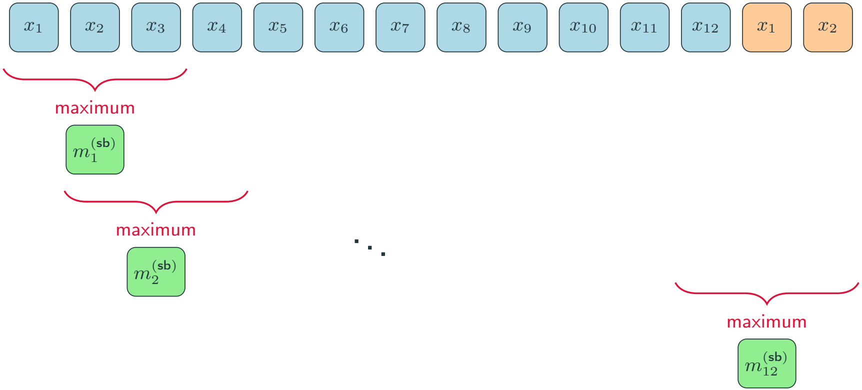
Sliding block maxima

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



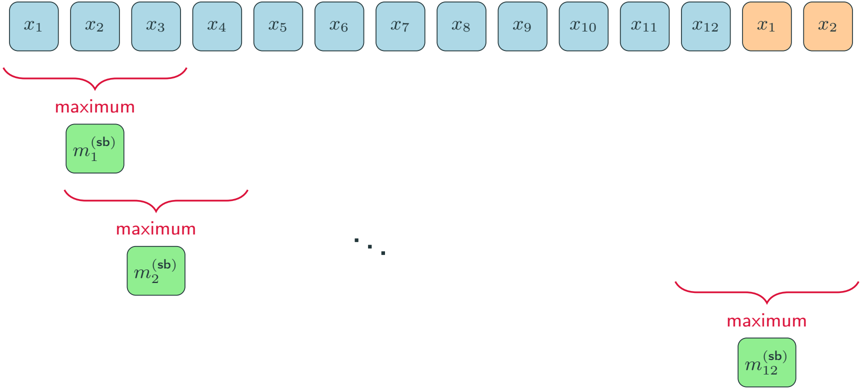
Sliding block maxima

Observations $x_1, \dots, x_{12}$; **block size** $r = 3$



Sliding block maxima

Observations $x_1, \ldots, x_{12}$; block size $r = 3$



Sliding block maxima sample $\mathcal{M}^{(\mathsf{sb})} = (m_1^{(\mathsf{sb})}, \ldots, m_{12}^{(\mathsf{sb})})$

Bootstrapping block maxima

Statistical setting

Estimation of $\boldsymbol{\theta}_r = \mathrm{E}[\boldsymbol{h}(\boldsymbol{M}_{r,1})]$where $\boldsymbol{h}\colon \mathbb{R} \to \mathbb{R}^q$ satisfies minimal regularity conditions

- Obtain the following sensible estimators for $\boldsymbol{\theta}_r$

Statistical setting

Estimation of $\boldsymbol{\theta}_r = \mathrm{E}[\boldsymbol{h}(\boldsymbol{M}_{r,1})]$where $\boldsymbol{h}\colon \mathbb{R} \to \mathbb{R}^q$ satisfies minimal regularity conditions

- Obtain the following sensible estimators for $\boldsymbol{\theta}_r$

$$\hat{\boldsymbol{\theta}}_n^{(\mathrm{mb})} = \frac{1}{n_{(\mathrm{mb})}} \sum_{i=1}^{n_{(\mathrm{mb})}} \boldsymbol{h}(\boldsymbol{M}_{r,i}^{(\mathrm{mb})}), \quad \mathrm{mb} \in \{\mathrm{db}, \mathrm{sb}\},$$

where $n_{(\mathrm{db})} = n/r, n_{(\mathrm{sb})} = n$.

Statistical setting

Estimation of $\boldsymbol{\theta}_r = \mathrm{E}[\boldsymbol{h}(\boldsymbol{M}_{r,1})]$where $\boldsymbol{h}\colon \mathbb{R} \to \mathbb{R}^q$ satisfies minimal regularity conditions

- Obtain the following sensible estimators for $\boldsymbol{\theta}_r$

$$\hat{\boldsymbol{\theta}}_n^{(\mathrm{mb})} = \frac{1}{n_{(\mathrm{mb})}} \sum_{i=1}^{n_{(\mathrm{mb})}} \boldsymbol{h}(\boldsymbol{M}_{r,i}^{(\mathrm{mb})}), \quad \mathrm{mb} \in \{\mathrm{db}, \mathrm{sb}\},$$

where $n_{(\mathrm{db})} = n/r, n_{(\mathrm{sb})} = n$.

Aim: Bootstrap $\hat{\boldsymbol{\theta}}_n^{(\mathrm{mb})} - \boldsymbol{\theta}_r$

Naive approach I

Consider first $\hat{\boldsymbol{\theta}}_n^{(\mathrm{db})}$

- Time series structure $\rightsquigarrow$ have to bootstrap blocks of observations
- Bootstrap block size of r is natural

$\rightsquigarrow$ Draw with replacement[6] from the sample of disjoint block maxima $\mathcal{M}^{(\mathrm{db})}$

[6] Multiplier bootstraps also possible

Naive approach I

Consider first $\hat{\boldsymbol{\theta}}_n^{(\mathrm{db})}$

- Time series structure $\rightsquigarrow$ have to bootstrap blocks of observations
- Bootstrap block size of r is natural

$\rightsquigarrow$ Draw with replacement[6] from the sample of disjoint block maxima $\mathcal{M}^{(\mathrm{db})}$

Algorithm 1 Disjoint block maxima bootstrap

Require: $n/r \in \mathbb{N}, \mathcal{M}^{(\mathrm{db})} = (m_1, \ldots, m_{n/r}), B \in \mathbb{N}$

1: **for** $b = 1$ to B **do**

2: Draw n/r times with replacement from $\mathcal{M}^{(\mathrm{db})}$ and concatenate to obtain $m^*_{b,1}, \ldots, m^*_{b,n/r}$

3: Compute $\hat{\boldsymbol{\theta}}^{*,(\mathrm{db})}_{n,b} = r/n \sum_{i=1}^{n/r} \boldsymbol{h}(\boldsymbol{m}^*_{b,i})$

4: **end for**

5: **return** $\hat{\boldsymbol{\theta}}^{*,(\mathrm{db})}_{n,1}, \ldots \hat{\boldsymbol{\theta}}^{*,(\mathrm{db})}_{n,B}$ ▷ Bootstrap replicates

---

[6] Multiplier bootstraps also possible

Naive approach II

Does it work?

Naive approach II

Does it work?

Yes; formally this means:

Theorem (Bücher and S., 2024) Under regularity conditions, as $n \to \infty$

$$d_K\Big(\mathcal{L}\big(\hat{\boldsymbol{\theta}}_n^{(\mathrm{db}),*} - \hat{\boldsymbol{\theta}}_n^{(\mathrm{db})} \mid \mathcal{X}_n\big), \mathcal{L}\big(\hat{\boldsymbol{\theta}}_n^{(\mathrm{db})} - \boldsymbol{\theta}_r\big)\Big) = o_{\mathbb{P}}(1).$$

Naive approach III

Now consider $\hat{\boldsymbol{\theta}}_n^{(\mathrm{sb})}$

- Problem: $\text{sliding-max}(x_1, \dots, x_r)$ does not make sense as opposed to its disjoint counterpart
- Instead one could draw r-blocks $M_{I_i}^{(\mathrm{sb})} = \{m_{(i-1)r+1}^{(\mathrm{sb})}, \dots, m_{ir}^{(\mathrm{sb})}\}$ of the sliding sample, $i = 1, \dots, n/r$

[7] c.f. slide 26

Naive approach III

Now consider $\hat{\boldsymbol{\theta}}_n^{(\mathrm{sb})}$

- Problem: $\text{sliding-max}(x_1, \ldots, x_r)$ does not make sense as opposed to its disjoint counterpart
- Instead one could draw r-blocks $M_{I_i}^{(\mathrm{sb})} = \{m_{(i-1)r+1}^{(\mathrm{sb})}, \ldots, m_{ir}^{(\mathrm{sb})}\}$ of the sliding sample, $i = 1, \ldots, n/r$

Does this work?[7]

[7] c.f. slide 26

Naive approach III

Now consider $\hat{\theta}_n^{(\mathrm{sb})}$

- Problem: $\text{sliding-max}(x_1, \ldots, x_r)$ does not make sense as opposed to its disjoint counterpart
- Instead one could draw r-blocks $M_{I_i}^{(\mathrm{sb})} = \{m_{(i-1)r+1}^{(\mathrm{sb})}, \ldots, m_{ir}^{(\mathrm{sb})}\}$ of the sliding sample, $i = 1, \ldots, n/r$

Does this work?[7]

No!

[7] c.f. slide 26

Circmax

Observations $x_1, \dots, x_{12}$; **block size** $r = 3$

x_1 x_2 x_3 | x_4 x_5 x_6 | x_7 x_8 x_9 | x_{10} x_{11} x_{12}

Circmax

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$

Divide observations into disjoint blocks of $2 * r = 6$

Circmax

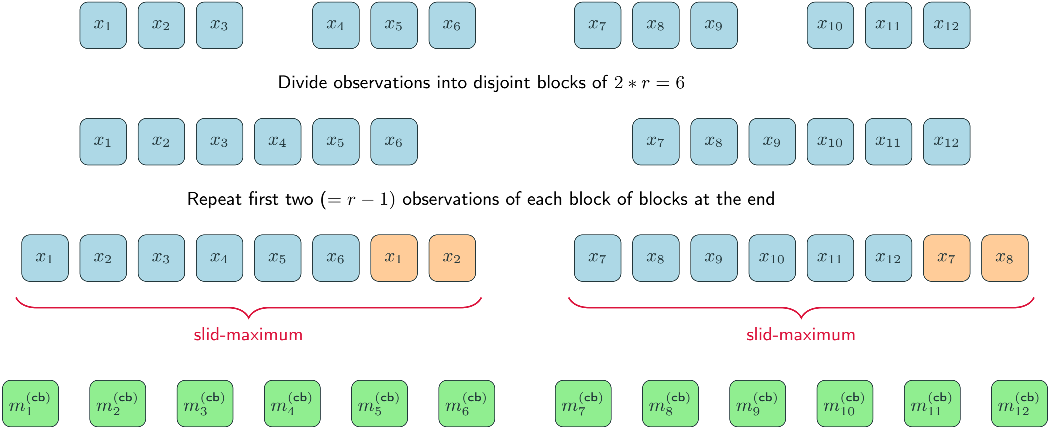
Observations $x_1, \ldots, x_{12}$; block size $r = 3$

Divide observations into disjoint blocks of $2 * r = 6$

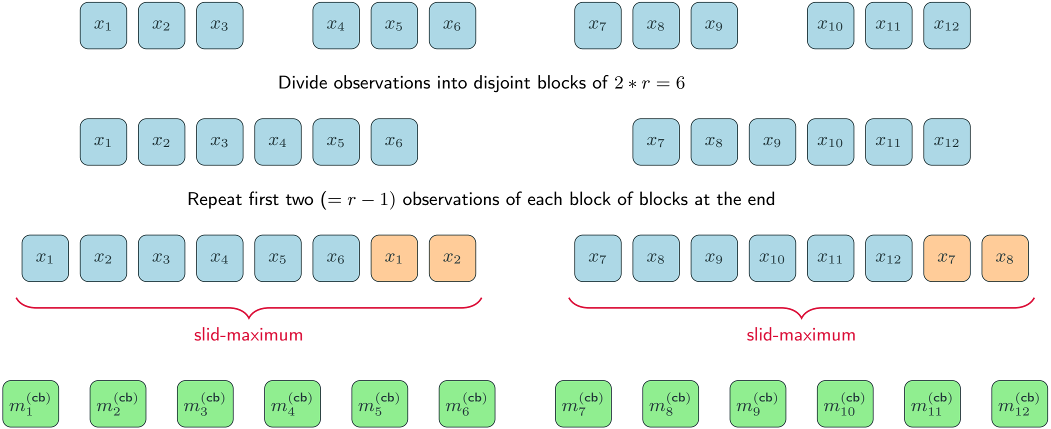
Repeat first two $(= r - 1)$ observations of each block of blocks at the end

Circmax

Observations $x_1, \ldots, x_{12}$; **block size** $r = 3$



Observations $x_1, \ldots, x_{12}$; block size $r = 3$



Circular block maxima sample $\mathcal{M}^{(\mathrm{cb})} = (m_1^{(\mathrm{cb})}, \ldots, m_{12}^{(\mathrm{cb})})$

Circmax II

The natural estimator for estimating $\boldsymbol{\theta}_r$ is $n^{-1}\sum_{i=1}^n \boldsymbol{h}(M_{r,i}^{(\mathrm{cb})})$

Lemma (Bücher, S., 2024) Under regularity conditions, as $n \to \infty$,

$$\frac{\mathrm{Var}(\hat{\boldsymbol{\theta}}_n^{(\mathrm{cb})})}{\mathrm{Var}(\hat{\boldsymbol{\theta}}_n^{(\mathrm{sb})})} \to 1.$$

Circmax bootstrap I

$\rightsquigarrow$ bootstrap $\hat{\boldsymbol{\theta}}_n^{(\mathrm{sb})}$ via $\hat{\boldsymbol{\theta}}^{(\mathrm{cb})}$

denote by $\mathcal{M}_i^{(\mathrm{cb})} = \{m_{(i-1)2r+1}, \ldots, m_{2ri}\}$ the ith $2r$ block of the circmax sample; $i = 1, \ldots, n/(2r)$

Algorithm 3 circmax block maxima bootstrap

Require: $n/r \in \mathbb{N}, \mathcal{M}^{(\mathrm{sb})} = (m_1, \ldots, m_n), B \in \mathbb{N}$

for $b = 1$ to B **do**

  Draw $n/(2r)$ times with replacement from $\{\mathcal{M}_i^{(\mathrm{cb})} : i = 1, \ldots, n/(2r)\}$ and concatenate to obtain

  $m_{b,1}^*, \ldots, m_{b,n}^*$

  Compute $\hat{\boldsymbol{\theta}}_{n,b}^{*,(\mathrm{cb})} = 1/n \sum_{i=1}^n \boldsymbol{h}(\boldsymbol{m}_{b,i}^*)$

end for

return $\hat{\boldsymbol{\theta}}_{n,1}^{*,(\mathrm{cb})}, \ldots \hat{\boldsymbol{\theta}}_{n,B}^{*,(\mathrm{cb})}$ ▷ Bootstrap replicates

Circmax bootstrap II

Does that one work?

Does that one work?

Theorem (Bücher, S., 2024) Under regularity conditions, as $n \to \infty$

$$d_K\Big(\mathcal{L}\big(\hat{\boldsymbol{\theta}}_n^{(\mathrm{cb}),*} - \hat{\boldsymbol{\theta}}_n^{(\mathrm{cb})} \mid \mathcal{X}_n\big), \mathcal{L}\big(\hat{\boldsymbol{\theta}}_n^{(\mathrm{sb})} - \theta_r\big)\Big) = o_{\mathbb{P}}(1).$$

Does that one work?

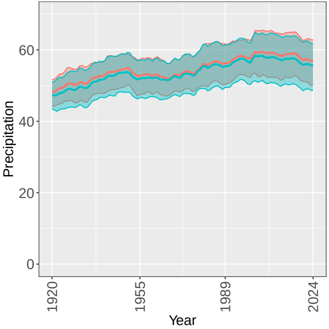
Theorem (Bücher, S., 2024) Under regularity conditions, as $n \to \infty$

$$d_K\Big(\mathcal{L}(\hat{\boldsymbol{\theta}}_n^{(\mathrm{cb}),*} - \hat{\boldsymbol{\theta}}_n^{(\mathrm{cb})} \mid \mathcal{X}_n), \mathcal{L}(\hat{\boldsymbol{\theta}}_n^{(\mathrm{sb})} - \theta_r)\Big) = o_{\mathbb{P}}(1).$$

- Allows for many applications: PWM-Estimator, Pseudo-MLE for Fréchet/GEV(γ), Moment estimators for the block maxima distribution (mean, variance ...)

Case Study

Confidence Intervals for the expected yearly maximum precipitation (in mm) at a fixed location[8]



- Left: Estimates (lines in the middle) and confidence intervals (ribbons) for the estimation of $\theta_r = \mathrm{E}[M_t]$
- Right: Width of the respective confidence intervals

[8] Hohenpeißenberg, Germany; data from 1879–2023

Conclusion

Conclusion

- Naive bootstrapping fails for sliding block maxima
- Introduced a new method **circular block maxima**
- Circmax enjoys advantages from disjoint and sliding world[9]
- Circmax based bootstraps are consistent for the sliding max estimation error

[9] small drawback: additional bias, but was found to be insignificant in simulation studies

References

- **Bücher, S. (2024). Bootstrapping Estimators based on the Block Maxima Method. arXiv:2409.08661.** Submitted for publication
- Bücher, Segers (2018). Inference for heavy tailed stationary time series based on sliding blocks. *Electron. J. Statist. 12(1): 1098-1125*
- de Haan, Zhou (2024). Bootstrapping Extreme Value Estimators. *Journal of the American Statistical Association, 119:545, 382-393*
- Efron (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist. 7(1): 1-26*
- Ferreira, de Haan (2015). On the block maxima method in extreme value theory: PWM estimators. *Ann. Statist. 43(1): 276-298*
- Zou, Volgushev, Bücher (2021). Multiple block sizes and overlapping blocks for multivariate time series extremes. *Ann. Statist. 49(1): 295-320*

Generated[10] by the prompt: *Create a comic style picture of a very happy block which is sliding down a slide shouting "thank you"*

[10] DALL-E

Disjoint block maxima II

Notes about $\mathcal{M}^{(\mathbf{db})}$:

- Only $n/r = o(n)$ disjoint block maxima
- No overlap between between blocks (disjoint)
- Asymptotic independence between $m_i^{(\mathbf{db})}, m_j^{(\mathbf{db})}$ for $i \neq j$ (**desirable property**)
- Asymptotic theory for many estimators established (Ferreira, de Haan, 2015)

Sliding block maxima II

Notes about $\mathcal{M}^{(\mathrm{sb})}$:

- Small modification: repeat the first $r-1$ observations at the end to ensure fair weighing
- After modifying we have n sliding block maxima
- Large overlap between blocks nearby
- Asymptotic **dependence** between $m_i^{(\mathrm{sb})}, m_j^{(\mathrm{sb})}$ for $i \neq j$, which can be described by a Marshall-Olkin type copula (in the one-dimensional case)
- **Linear estimators based on sliding blocks have smaller variance than their disjoint counterpart**: (Zou et. al. 2021)

Circmax III

Notes about $\mathcal{M}^{(\text{cb})}$:

- Essentially combines disjoint block with sliding block method
- Size of the circmax sample is n
- Large overlap between blocks nearby (sliding effect) but no overlap between $2 * r$ blocks (disjoint effect)
- Repeating observations induces non-stationarity but does not hurt[11]

[11] does not hurt to much: asymptotic variance stays the same but there is negligible (compared to classical) bias

Outlook

- Extendable to U-statistics of circular block maxima
- Extendable to non-stationary situations: piecewise stationary time series
- Purely of mathematical interest: one can define circular blocks with irrational outer block length $k \rightsquigarrow$ then circmax defines a spectrum of maxima methods with $k = 1$ corresponding to disjoint block maxima $k = n/r$ corresponding to sliding block maxima; non-trivial things happening for $k \in (1, 2)$

Example: MLE for Fréchet based on block maxima extracted from a time series

Statistical model:

- $\mathcal{X}_n = (X_1, \ldots, X_n)$ strictly stationary time series and $\mathbb{R}$-valued
- $\mathcal{X}_n$ belongs to the Fréchet DoA, that is: $\exists \alpha_0 > 0, (\sigma_r)_r \in (0, \infty)^{\mathbb{N}}$, s.t.

$$\max\left(\frac{X_1}{\sigma_r}, \ldots, \frac{X_r}{\sigma_r}\right) \rightsquigarrow P_{\alpha_0},$$

  where $P_{\alpha_0} \sim$ Fréchet(α_0)
- Basic model assumptions still hold (block size, dependency structure)

- Goal: estimation of α_0, σ_r

Example: MLE for Fréchet based on block maxima extracted from a time series

Statistical model:

- $\mathcal{X}_n = (X_1, \dots, X_n)$ strictly stationary time series and $\mathbb{R}$-valued
- $\mathcal{X}_n$ belongs to the Fréchet DoA, that is: $\exists \alpha_0 > 0, (\sigma_r)_r \in (0,\infty)^{\mathbb{N}}$, s.t.

$$\max\left(\frac{X_1}{\sigma_r}, \dots, \frac{X_r}{\sigma_r}\right) \rightsquigarrow P_{\alpha_0},$$

  where $P_{\alpha_0} \sim$ Fréchet(α_0)
- Basic model assumptions still hold (block size, dependency structure)

- Goal: estimation of α_0, σ_r

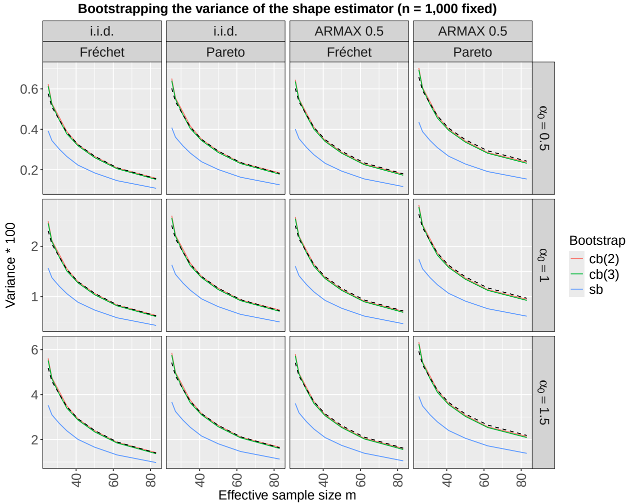
Good performing estimator (Bücher, Segers, 2018)

$$\hat{\theta}_n^{(\mathrm{sb})} := (\hat{\alpha}_n^{(\mathrm{sb})}, \hat{\sigma}_n^{(\mathrm{sb})})^\top := \operatorname*{argmax}_{\theta=(\alpha,\sigma)\in(0,\infty)^2} \sum_{M_i \in \mathcal{M}_{n,r}^{(\mathrm{sb})}} \ell_\theta(M_i \vee c),$$

where $c > 0$ arbitrary, ℓ_θ denotes log-likelihood of Fréchet(α_0, σ)

Finite sample results

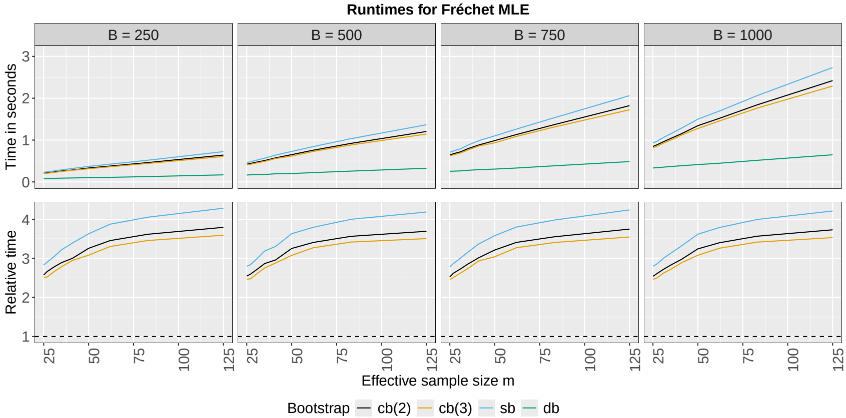
Bootstrap the variance of $\hat{\alpha}_n^{(\mathrm{sb})}$ (asymptotically: complicated function of α_0)



- variance of $\hat{\alpha}_n^{(\mathrm{sb})}$ obtained via presimulating 10^6 time series of sample size $n = 10^3$, calculating for each sample $\hat{\alpha}_n^{(\mathrm{sb})}$ and taking the empirical variance (black dashed line)
- different bootstrap procedures displayed; based on $B = 10^3$ bootstrap replicates, $N = 5 * 10^3$ repetitions (averaged)
- **inconsistency of sliding visible**
- **circmax bootstrapping works**

Runtime comparison

In applications important: large bootstrap replicate numbers not too expensive



Absolute and relative median runtimes of different bootstrap algorithms for bootstrapping $\hat{\theta}_n^{(\mathrm{mb})}$ (relative to the runtime of the disjoint blocks bootstrap) for fixed sample size $n = 1,000$ as a function of the effective sample size and for different numbers of bootstrap replicates B; based on 500 runs.