University of Stuttgart
Germany

Ioan Scheffel

Asymptotics of Peaks-over-Threshold Estimators in Long Memory Linear Time Series

joint work with M. Oesting & G. Stupfler

GPSD 2025, March 12th

What are POT estimators? Why do we use them?

Our Setting:

What are POT estimators? Why do we use them?

Our Setting:

- (strictly) stationary time series X_t
- threshold sequence (u_n) with $\mathbb{P}[X_0 > u_n] \to 0$ for $n \to \infty$

What are POT estimators? Why do we use them?

Our Setting:

- (strictly) stationary time series X_t
- threshold sequence (u_n) with $\mathbb{P}[X_0 > u_n] \to 0$ for $n \to \infty$

Example:

With the **simple POT** we count exceedances of (X_t) over u_n to estimate $\mathbb{P}[X_0 > u_n]$, i.e.,

$$\frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{X_t > u_n\} \qquad \text{to estimate} \qquad \mathbb{P}[X_0 > u_n]\,.$$

What are POT estimators? Why do we use them?

Our Setting:

- (strictly) stationary time series X_t
- threshold sequence (u_n) with $\mathbb{P}[X_0 > u_n] \to 0$ for $n \to \infty$

Example:

With the **simple POT** we count exceedances of (X_t) over u_n to estimate $\mathbb{P}[X_0 > u_n]$, i.e.,

$$\frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{X_t > u_n\} \qquad \text{to estimate} \qquad \mathbb{P}[X_0 > u_n]\,.$$

Mostly used together with other estimators, e.g., moving window estimator

$$\frac{\frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{X_t > u_n, X_{t+h} > u_n\}}{\frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{X_t > u_n\}} \qquad \text{to estimate} \qquad \chi(h)\,,$$

where χ is the tail-dependence-coefficient

$$\chi(h) := \lim_{n\to\infty} \mathbb{P}[X_h > u_n \mid X_0 > u_n] = \lim_{n\to\infty} \frac{\mathbb{P}[X_h > u_n, X_0 > u_n]}{\mathbb{P}[X_0 > u_n]} \qquad \text{for } h > 0\,.$$

What are POT estimators? Why do we use them?

Other Examples:

- **Hill Estimator** to estimate how heavy the tails are, that is,

$$\frac{1}{n}\sum_{t=1}^{n}\left(\log(X_t) \;-\; \log(u_n)\right)\mathbb{1}\{X_t > u_n\} \qquad \text{to estimate} \qquad \gamma\,,$$

where $1/\gamma$ is the index of regular variation of the tail-function $x \mapsto \mathbb{P}[X_0 > x]$.

What are POT estimators? Why do we use them?

Other Examples:

- **Hill Estimator** to estimate how heavy the tails are, that is,

$$\frac{1}{n}\sum_{t=1}^{n}\left(\log(X_t) \;-\; \log(u_n)\right)\mathbb{1}\{X_t > u_n\} \qquad \text{to estimate} \qquad \gamma\,,$$

  where $1/\gamma$ is the index of regular variation of the tail-function $x \mapsto \mathbb{P}[X_0 > x]$.

- **Expected shortfall** to estimate how much it goes wrong when it goes wrong, that is,

$$\frac{\frac{1}{n}\sum_{t=1}^{n}(X_t \;-\; u_n)\,\mathbb{1}\{X_t > u_n\}}{\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}\{X_t > u_n\}} \qquad \text{to estimate} \qquad \mathbb{E}[X_0 - u_n \mid X_0 > u_n]\,.$$

Deterministic and Random Thresholds

Deterministic and Random Thresholds

Practical Challenge: choose threshold sequence (u_n)

Deterministic and Random Thresholds

Practical Challenge: choose threshold sequence (u_n)
Practical Solution: Order your sample

$$X_{1:n} \le \ldots \le X_{n:n}$$

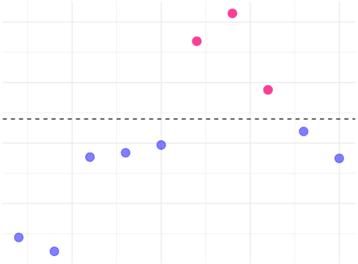
and choose $X_{n-k:n}$ as threshold with intermediate sequence $k = k(n)$.

Deterministic and Random Thresholds

Practical Challenge: choose threshold sequence (u_n)
Practical Solution: Order your sample

$$X_{1:n} \leq \ldots \leq X_{n:n}$$

and choose $X_{n-k:n}$ as threshold with intermediate sequence $k = k(n)$.

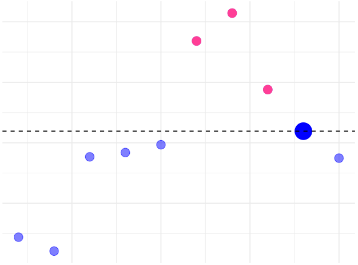


Deterministic and Random Thresholds

Practical Challenge: choose threshold sequence (u_n)
Practical Solution: Order your sample

$$X_{1:n} \leq \ldots \leq X_{n:n}$$

and choose $X_{n-k:n}$ as threshold with intermediate sequence $k = k(n)$.



Deterministic and Random Thresholds

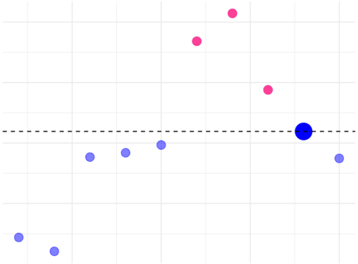
Practical Challenge: choose threshold sequence (u_n)
Practical Solution: Order your sample

$$X_{1:n} \leq \ldots \leq X_{n:n}$$

and choose $X_{n-k:n}$ as threshold with intermediate sequence $k = k(n)$.

Takeaways:

- deterministic thresholds more of a theoretical device
- random thresholds more common in practice



Outline

1. Asymptotics for i.i.d. Data
2. Asymptotics for Long Memory Linear Time Series
3. Random Thresholds

Asymptotics for i.i.d. Data

1

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

Takeaway:

What to expect from dependent data:

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

What to expect from dependent data:

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$

What to expect from dependent data:

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$
- usual **convergence rate** $\sqrt{n}$ **slows down** to $\sqrt{n\mathbb{P}[X_0 > u_n]}$

What to expect from dependent data:

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$
- usual **convergence rate** $\sqrt{n}$ **slows down** to $\sqrt{n\mathbb{P}[X_0 > u_n]}$
  $\rightsquigarrow$ We get **punished** for cutting out non-extreme data

What to expect from dependent data:

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$
- usual **convergence rate** $\sqrt{n}$ **slows down** to $\sqrt{n\mathbb{P}[X_0 > u_n]}$

  $\leadsto$ We get **punished** for cutting out non-extreme data

What to expect from dependent data:

- slow down of convergence rate due to decreased effective sample size
- weak dependence (short memory) behavior same as i.i.d.
- strong dependence (long memory) behavior different from i.i.d.

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$
- usual **convergence rate** $\sqrt{n}$ **slows down** to $\sqrt{n\mathbb{P}[X_0 > u_n]}$

  $\rightsquigarrow$ We get **punished** for cutting out non-extreme data

What to expect from dependent data:

- slow down of convergence rate due to decreased effective sample size ✓
- weak dependence (short memory) behavior same as i.i.d.
- strong dependence (long memory) behavior different from i.i.d.

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$
- usual **convergence rate** $\sqrt{n}$ **slows down** to $\sqrt{n\mathbb{P}[X_0 > u_n]}$

  $\rightsquigarrow$ We get **punished** for cutting out non-extreme data

What to expect from dependent data:

- slow down of convergence rate due to decreased effective sample size ✓
- weak dependence (short memory) behavior same as i.i.d. ×
- strong dependence (long memory) behavior different from i.i.d.

Asymptotics of POT Estimators for i.i.d. Data

Theorem (CLT for i.i.d. data)

Let (X_n) be i.i.d. and (u_n) a sequence with $u_n \to \infty$ such that $n\mathbb{P}[X_0 > u_n] \to \infty$.

$$\sqrt{n\mathbb{P}[X_0 > u_n]}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t > u_n\}}{\mathbb{P}[X_0 > u_n]} - 1\right) \to_d Z \sim \mathcal{N}(0,1) \qquad \text{for } n \to \infty\,.$$

Takeaway:

- when looking at extremes the **effective sample size reduces** to $n\mathbb{P}[X_0 > u_n]$
- usual **convergence rate** $\sqrt{n}$ **slows down** to $\sqrt{n\mathbb{P}[X_0 > u_n]}$
  $\rightsquigarrow$ We get **punished** for cutting out non-extreme data

What to expect from dependent data:

- slow down of convergence rate due to decreased effective sample size ✓
- weak dependence (short memory) behavior same as i.i.d. ×
- strong dependence (long memory) behavior different from i.i.d. ✓

2

Asymptotics for Long Memory Linear Time Series

Asymptotics for Long Memory Linear Time Series

Our Notion of Long Memory:

Asymptotics for Long Memory Linear Time Series

Our Notion of Long Memory: (Causal) linear time series

$$X_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}, \qquad t \in \mathbb{N}_0$$

with

Asymptotics for Long Memory Linear Time Series

Our Notion of Long Memory: (Causal) linear time series

$$X_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}\,, \qquad t \in \mathbb{N}_0$$

with

- symmetric i.i.d. innovations (ε_k) with notion of heavy tails

$$\alpha := \sup\{m \in (1,2] \mid \text{the } m\text{-th moment of } \varepsilon \text{ exists}\}$$

$\rightsquigarrow$ we allow for distribution with infinite variance

Asymptotics for Long Memory Linear Time Series

Our Notion of Long Memory: (Causal) linear time series

$$X_t \;=\; \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}\,, \qquad t \in \mathbb{N}_0$$

with

- symmetric i.i.d. innovations (ε_k) with notion of heavy tails

$$\alpha \;:=\; \sup\{m \in (1,2] \mid \text{the } m\text{-th moment of } \varepsilon \text{ exists}\}$$

  $\rightsquigarrow$ we allow for distribution with infinite variance

- slowly decaying coefficients (a_k) with $a_k \sim k^{-1+d}$ with long memory parameter

$$0 \;<\; d \;<\; 1 - 1/\alpha$$

  $\rightsquigarrow$ reminiscent of fractional differencing parameter in fractional ARIMA

Asymptotics for Long Memory Linear Time Series

Theorem (S., Oesting & Stupfler 2025+)

Let (X_t) be a long memory linear time series. Under technical conditions on the distribution of the innovations (ε_k) and the growth rate of (u_n) it holds

$$n^{1-d-1/\alpha}\frac{\mathbb{P}[X_0>u_n]}{f_{X_0}(u_n)}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t>u_n\}}{\mathbb{P}[X_0>u_n]}-1\right)\to_d C(\mathbb{P}_\varepsilon,d,\alpha)\cdot Z_\alpha \qquad \text{for } n\to\infty\,,$$

where

$$Z_\alpha \sim \begin{cases} S\alpha S\,, & \alpha\in(1,2)\,,\\ \mathcal{N}(0,1), & \alpha=2 \text{ and variance exists.}\end{cases}$$

Asymptotics for Long Memory Linear Time Series

Theorem (S., Oesting & Stupfler 2025+)

Let (X_t) be a long memory linear time series. Under technical conditions on the distribution of the innovations (ε_k) and the growth rate of (u_n) it holds

$$n^{1-d-1/\alpha}\frac{\mathbb{P}[X_0>u_n]}{f_{X_0}(u_n)}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\mathbb{1}\{X_t>u_n\}}{\mathbb{P}[X_0>u_n]}-1\right)\to_d C(\mathbb{P}_\varepsilon,d,\alpha)\cdot Z_\alpha \qquad \text{for } n\to\infty\,,$$

where

$$Z_\alpha \sim \begin{cases} S\alpha S\,, & \alpha\in(1,2)\,,\\ \mathcal{N}(0,1), & \alpha=2 \text{ and variance exists.}\end{cases}$$

Remarkable:
The density of the time series f_{X_0} pops up in the ratio $\mathbb{P}[X_0>u_n]/f_{X_0}(u_n)$.

Idea of Proof

We extend ideas from (Koul & Surgailis 2001):

- Control growth of

$$\mathbb{E}\left[\left|\sum_{t=1}^{n} \mathbb{1}\{X_t > u_n\} - \mathbb{P}[X_0 > u_n] + f_{X_0}(u_n) X_t\right|^r\right]$$

  for some carefully chosen $r \in (1, \alpha)$.

- Use CLT for partial sums

$$n^{-d-1/\alpha} \sum_{t=1}^{n} X_t \to_d C(\mathbb{P}_\varepsilon, d, \alpha) \cdot Z_\alpha \qquad \text{for } n \to \infty \,,$$

  together with Slutzky's theorem.

Idea of Proof

We extend ideas from (Koul & Surgailis 2001):

- Control growth of

$$\mathbb{E}\left[\left|\sum_{t=1}^{n} \mathbb{1}\{X_t > u_n\} - \mathbb{P}[X_0 > u_n] + f_{X_0}(u_n) X_t\right|^r\right]$$

  for some carefully chosen $r \in (1, \alpha)$.

- Use CLT for partial sums

$$n^{-d-1/\alpha} \sum_{t=1}^{n} X_t \to_d C(\mathbb{P}_\varepsilon, d, \alpha) \cdot Z_\alpha \qquad \text{for } n \to \infty \, ,$$

  together with Slutzky's theorem.

Also: We have proved more general results, e.g., for Hill estimator.

Ratio of Tail Function and Density

From i.i.d. section: expect to get punished for looking at extremes.
How severe punishment is we can read off:

$$\frac{\mathbb{P}[X_0 > u_n]}{f_{X_0}(u_n)} \sim \left\{ \right.$$

Observation:

Ratio of Tail Function and Density

From i.i.d. section: expect to get punished for looking at extremes.
How severe punishment is we can read off:

$$\frac{\mathbb{P}[X_0 > u_n]}{f_{X_0}(u_n)} \sim \begin{cases} 1/u_n & \text{if} \quad X_0 \sim \mathcal{N}(0,1)\,, \\ \end{cases}$$

Observation:

- for light tails the convergence rate slows down, e.g., with factor $1/u_n$

Ratio of Tail Function and Density

From i.i.d. section: expect to get punished for looking at extremes.
How severe punishment is we can read off:

$$\frac{\mathbb{P}[X_0 > u_n]}{f_{X_0}(u_n)} \sim \begin{cases} 1/u_n & \text{if} \quad X_0 \sim \mathcal{N}(0,1)\,, \\ u_n/\nu & \text{if} \quad f_{X_0} \in \mathrm{RV}_{-\nu-1}\,. \end{cases}$$

Observation:

- for light tails the convergence rate slows down, e.g., with factor $1/u_n$

Ratio of Tail Function and Density

From i.i.d. section: expect to get punished for looking at extremes.
How severe punishment is we can read off:

$$\frac{\mathbb{P}[X_0 > u_n]}{f_{X_0}(u_n)} \sim \begin{cases} 1/u_n & \text{if} \quad X_0 \sim \mathcal{N}(0,1)\,, \\ u_n/\nu & \text{if} \quad f_{X_0} \in \mathrm{RV}_{-\nu-1}\,. \end{cases}$$

Observation:

- for light tails the convergence rate slows down, e.g., with factor $1/u_n$
- for heavy tails the convergence rate **speeds up**, e.g., with factor u_n

Ratio of Tail Function and Density

From i.i.d. section: expect to get punished for looking at extremes.
How severe punishment is we can read off:

$$\frac{\mathbb{P}[X_0 > u_n]}{f_{X_0}(u_n)} \sim \begin{cases} 1/u_n & \text{if} \quad X_0 \sim \mathcal{N}(0,1)\,, \\ u_n/\nu & \text{if} \quad f_{X_0} \in \mathrm{RV}_{-\nu-1}\,. \end{cases}$$

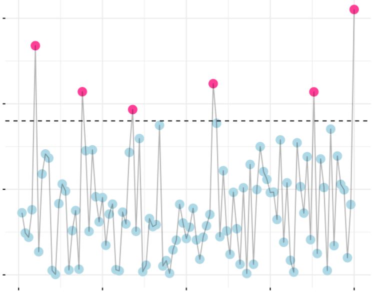
Observation:

- for light tails the convergence rate slows down, e.g., with factor $1/u_n$
- for heavy tails the convergence rate **speeds up**, e.g., with factor u_n
  $\rightsquigarrow$ **reward for looking at extremes!**

Ratio of Tail Function and Density

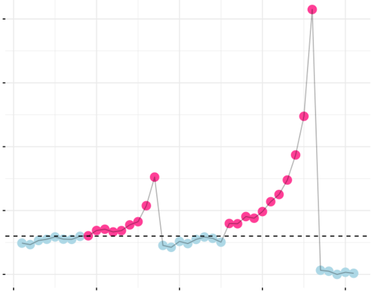
From i.i.d. section: expect to get punished for looking at extremes.
How severe punishment is we can read off:

$$\frac{\mathbb{P}[X_0 > u_n]}{f_{X_0}(u_n)} \sim \begin{cases} 1/u_n & \text{if} \quad X_0 \sim \mathcal{N}(0,1)\,, \\ u_n/\nu & \text{if} \quad f_{X_0} \in \mathrm{RV}_{-\nu-1}\,. \end{cases}$$



Observation:

- for light tails the convergence rate slows down, e.g., with factor $1/u_n$
- for heavy tails the convergence rate **speeds up**, e.g., with factor u_n
  $\rightsquigarrow$ **reward for looking at extremes!**



Random Thresholds

3

Tail Empirical Process

- Version with deterministic threshold

- Version with random threshold

Tail Empirical Process

- Version with deterministic threshold

$$\widetilde{T}_n(s) \;=\; \frac{1}{n\mathbb{P}[X_0 > u_n]} \sum\nolimits_{t=1}^{n} \mathbb{1}\{X_t > u_n \cdot s\} \qquad \text{for } s > 0$$

- Version with random threshold

Tail Empirical Process

- Version with deterministic threshold

$$\widetilde{T}_n(s) \;=\; \frac{1}{n\mathbb{P}[X_0 > u_n]} \sum\nolimits_{t=1}^{n} \mathbb{1}\{X_t > u_n \cdot s\} \qquad \text{for } s > 0$$

- Version with random threshold

$$\widehat{T}_n(s) \;=\; \frac{1}{k} \sum\nolimits_{t=1}^{n} \mathbb{1}\{X_t > X_{n-k:n} \cdot s\} \qquad \text{for } s > 0$$

Tail Empirical Process

- Version with deterministic threshold

$$\widetilde{T}_n(s) = \frac{1}{n\mathbb{P}[X_0 > u_n]} \sum_{t=1}^{n} \mathbb{1}\{X_t > u_n \cdot s\} \qquad \text{for } s > 0$$

- Version with random threshold

$$\widehat{T}_n(s) = \frac{1}{k} \sum_{t=1}^{n} \mathbb{1}\{X_t > X_{n-k:n} \cdot s\} \qquad \text{for } s > 0$$

Theorem (Kulik & Soulier 2011)

In long memory stochastic volatility model:

- $\widetilde{T}$ shows non-standard behavior due to long memory
- $\widehat{T}$ shows **i.i.d.** behavior

Tail Empirical Process

- Version with deterministic threshold

$$\widetilde{T}_n(s) = \frac{1}{n\mathbb{P}[X_0 > u_n]} \sum_{t=1}^{n} \mathbb{1}\{X_t > u_n \cdot s\} \qquad \text{for } s > 0$$

- Version with random threshold

$$\widehat{T}_n(s) = \frac{1}{k} \sum_{t=1}^{n} \mathbb{1}\{X_t > X_{n-k:n} \cdot s\} \qquad \text{for } s > 0$$

Theorem (Kulik & Soulier 2011)

In long memory stochastic volatility model:

- $\widetilde{T}$ shows non-standard behavior due to long memory
- $\widehat{T}$ shows **i.i.d.** behavior

$\rightsquigarrow$ random threshold eliminates long memory in stochastic volatility model

Tail Empirical Process

- Version with deterministic threshold

$$\widetilde{T}_n(s) = \frac{1}{n\mathbb{P}[X_0 > u_n]} \sum_{t=1}^{n} \mathbb{1}\{X_t > u_n \cdot s\} \qquad \text{for } s > 0$$

- Version with random threshold

$$\widehat{T}_n(s) = \frac{1}{k} \sum_{t=1}^{n} \mathbb{1}\{X_t > X_{n-k:n} \cdot s\} \qquad \text{for } s > 0$$

Theorem (Kulik & Soulier 2011)

In long memory stochastic volatility model:

- $\widetilde{T}$ shows non-standard behavior due to long memory
- $\widehat{T}$ shows **i.i.d.** behavior

$\rightsquigarrow$ random threshold eliminates long memory in stochastic volatility model

Question: Do we find this in **long memory linear time series**?

Hill Estimator with random thresholds

We consider the Hill estimator with deterministic ($\widetilde{\gamma}_n$) and random ($\widehat{\gamma}_n$) thresholds.

Theorem (S., Oesting & Stupfler 2025+)

With long memory linear time series and second order condition as $n \to \infty$ it holds

$$n^{1-(d+1/(\nu\wedge 2))} \quad u_n \quad \left(\widetilde{\gamma}_n - \frac{1}{\nu}\right) \quad \to_d \quad C \quad \frac{\nu}{1+\nu} \quad Z_{\nu\wedge 2}\,,$$

$$n^{1-(d+1/(\nu\wedge 2))} \quad F^{\leftarrow}_{X_0}\left(1 - \frac{k}{n}\right) \quad \left(\widehat{\gamma}_n - \frac{1}{\nu}\right) \quad \to_d \quad C \quad \frac{1}{1+\nu} \quad Z_{\nu\wedge 2}\,.$$

Stochastic Volatility Model

- no long memory for random thresholds

Linear Time Series

- long memory for random thresholds vanishes at the transition between heavy/light tails

Hill Estimator with random thresholds

We consider the Hill estimator with deterministic ($\widetilde{\gamma}_n$) and random ($\widehat{\gamma}_n$) thresholds.

Theorem (S., Oesting & Stupfler 2025+)

With long memory linear time series and second order condition as $n \to \infty$ it holds

$$n^{1-(d+1/(\nu\wedge 2))} \quad u_n \quad \left(\widetilde{\gamma}_n - \frac{1}{\nu}\right) \quad \to_d \quad C \quad \frac{\nu}{1+\nu} \quad Z_{\nu\wedge 2}\,,$$

$$n^{1-(d+1/(\nu\wedge 2))} \quad F^{\leftarrow}_{X_0}\left(1-\frac{k}{n}\right) \quad \left(\widehat{\gamma}_n - \frac{1}{\nu}\right) \quad \to_d \quad C \quad \frac{1}{1+\nu} \quad Z_{\nu\wedge 2}\,.$$

Stochastic Volatility Model

- no long memory for random thresholds
- asymptotic independence $\chi(h) = 0$

Linear Time Series

- long memory for random thresholds vanishes at the transition between heavy/light tails
- asymptotic dependence $\chi(h) > 0$

Summary & Work in Progress

In Long Memory Linear Time Series...

- ... convergence rate of POT estimators depends on heavy/light tails.
- ... with heavy tails we get rewarded for looking at extremes.
- ... transfer to random thresholds is beneficial both in practice and theory.

Work in Progress:

- Simulation study: Finite-sample behavior for Gaussian ARFIMA and beyond...

Thank You!

References

H. L. Koul and D. Surgailis.
Asymptotics of empirical processes of long memory moving averages with infinite variance.
Stochastic Processes and their Applications, 91(2):309–336, Feb. 2001.

R. Kulik and P. Soulier.
Tail empirical process for some long memory sequences.
Stochastic Processes and their Applications, 2011.

R. Kulik and P. Soulier.
Heavy-Tailed Time Series.
Springer, 2020.

G. Samorodnitsky.
Stochastic Processes and Long Range Dependence.
Springer, 2016.