Heavy-tailed and long memory time series - simulation study and applications

Elżbieta Gajecka-Mirek\textsuperscript{a},
Oskar Knapik\textsuperscript{b}

\textsuperscript{a} PWSZ Nowy Sacz, Poland
\textsuperscript{b} CREATEES, Aarhus University, Denmark

February 5 – 9, 2017, Gródek nad Dunajcem, Poland
Research subject
Research subject

To find confidence interval for the mean with optimal chosen block length for specific time series:
Research subject

To find confidence interval for the mean with optimal chosen block length for specific time series:

- heavy tails
- long memory
Motivation

- simultaneous occurrence in practice heavy tails and long memory
- statistical inference based on asymptotic distributions do not always work (dependent data, resampling method)
Motivation

- simultaneous occurrence in practice heavy tails and long memory
- statistical inference based on asymptotic distributions do not always work (dependent data, resampling method)
- construction of confidence intervals without referring to the asymptotic distribution (resampling methods, block length)
Motivation

- simultaneous occurrence in practice heavy tails and long memory
- statistical inference based on asymptotic distributions do not always work (dependent data, resampling method)
- construction of confidence intervals without referring to the asymptotic distribution (resampling methods, block length)
Motivation

Subsampling - tool to construct CI

The idea of subsampling: Let \((X_1, \ldots, X_n)\) be observed sample.

- let \(T_n = \vartheta_n(\hat{\theta}_n - \theta)\) be tested statistics
- subsampling version of estimator \(\hat{\theta}_n\) is calculated on the subsample sample \((X_t, \ldots, X_{t+b-1})\), as \(\hat{\theta}_{n,b,t} = \hat{\theta}_b(X_t, \ldots, X_{t+b-1})\)
- empirical distribution

\[ L_{n,b}(x) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} 1\{\vartheta_b(\hat{\theta}_{n,b,t}-\hat{\theta}_n) \leq x\} \]

we use to approximate the asymptotic distribution of the estimator

\[ T_n = \vartheta_n(\hat{\theta}_n - \theta) \]
The advantage of subsampling is its insensitivity to the form of the asymptotic distribution.
Using resampling methods we need to answer the questions:

- If the empirical distribution (obtained from resampling methods) is close enough to real distribution? We need to know if our method is consistent and
Using resampling methods we need to answer the questions:

- If the empirical distribution (obtained from resampling methods) is close enough to real distribution? We need to know if our method is consistent and
- How long is the block?
Let us consider the model \( \{X_t\}_{t \in \mathbb{Z}} \) defined as:

\[
X_t = \sigma_t G_t + \eta,
\]

**A1** \( \sigma_t \) and \( G_t \) are independent,

**A2** \( \sigma_t \) is i.i.d. and have the marginal distribution of a heavy tail, ex.: an \( \alpha \)-stable or not as heavy as stable ex.: GED family, t-Student with an adequate number of degrees of freedom

**A3** \( G_t \) is long memory, purely non-deterministic, stationary mean zero Gaussian process \( G_t \) is a long memory with parameter \( \beta \in (0, 1) \),

**A4** \( \eta \) is the mean.
The properties of the model

1. \( \{X_t\}_{t \in \mathbb{Z}} \) has long memory.
2. \( \{X_t\}_{t \in \mathbb{Z}} \) is \( \lambda \)-weak dependence.
3. \( \{X_t\}_{t \in \mathbb{Z}} \) satisfies the condition:
   \[
   \lambda_r = O(r^{\beta-1}), \quad \beta \in [0, 1).
   \]
4. If the sequence \( \sigma_t \) is i.i.d., and its marginal distribution comes from for example the GED family, then the marginal distribution of \( \{X_t\}_{t \in \mathbb{Z}} \) has heavy tails.
Pr. 4’ If the sequence $\sigma_t$ is i.i.d. and its marginal distribution comes from stable family and is defined as follows:

$$\sigma_t = \sqrt{\epsilon_t},$$ (1)

where $\epsilon_t$ are i.i.d. $\alpha/2$-stable, $\alpha \in (1, 2)$, with skewness parameter equal to one, and scale parameter $(\cos(\pi \alpha/4))^{2/\alpha}$, and location parameter $\epsilon_t = 0$, for all $t$, then the marginal distribution of $X_t$ is $S\alpha S$ with scale parameter

$$\tau(X_t) = \sqrt{\gamma_G(0)/2}.$$ 

The variance of $X_t$ do not exist, but there exists periodic autocovariance function $\gamma(h) < \infty$ for $h \neq 0$.

Pr. 5 $\{X_t\}_{t \in \mathbb{Z}}$ is stationary.
The long range dependence - The long memory time series

**Definition**

The time series \( \{X_t\}_{t \in \mathbb{Z}} \) has **long memory time series** if its autocovariance function \( \gamma \) satisfies:

\[
\sum_{0 < |h| < n} \gamma(h) \sim C n^\beta
\]

as \( n \to \infty \), \( \beta \in [0, 1) \), and \( C > 0 \).

The long memory means that there is a correlation between observations far distant time.
**Long memory - Gegenbauer process**

**Definition**

(Gray, Zhang, Woodward, 1989) Let $\varepsilon_t$ i.i.d. $\{G_t\}_{t \in \mathbb{Z}}$ define as:

$$\prod_{1 \leq i \leq k} (1 - 2\nu_i B + B^2)^{d_i} G_t = \varepsilon_t,$$

(2)

is $k$-factor Gegenbauer process.

$0 < d_i < 1/2$ if $|\nu_i| < 1$ or $0 < d_i < 1/4$ if $|\nu_i| = 1$ for $i = 1, ..., k$. 

If \( \{\varepsilon_t\}_{t \in \mathbb{Z}} \) is the Gaussian white noise, then \( \{G_t\}_{t \in \mathbb{Z}} \) is Gaussian time series.
Weakly dependent sequences

Let $(E, \| \cdot \|)$ be a normed space.

$$\mathcal{L} = \{ h : E^u \to \mathbb{R}, \| h \|_\infty \leq 1, \text{Lip}(h) < \infty \},$$

where $\text{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|_1}$ and $\| x \|_1 = \sum_{i=1}^{u} \| x_i \|$. 
Weakly dependent sequences

Let \((E, \| \cdot \|)\) be a normed space.

\[
\mathcal{L} = \{ h : E^u \to \mathbb{R}, \| h \|_\infty \leq 1, \text{Lip}(h) < \infty \},
\]

where \(\text{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|_1}\) and \(\| x \|_1 = \sum_{i=1}^{u} \| x_i \|\).
**Definition**

(Doukhan, 1999). A sequence \( \{X_n\}_{n \in \mathbb{N}} \) of random variables taking values in \( E = \mathbb{R}^d \) is \((\theta, \mathcal{L}, \Psi)\)-weakly dependent if there exists \( \Psi : \mathcal{L} \times \mathcal{L} \times \mathbb{N}^* \times \mathbb{N}^* \rightarrow \mathbb{R} \) and a sequence \( \{\theta_r\}_{r \in \mathbb{N}} \) (\( \theta_r \to 0 \)) such that for any \((f, g) \in \mathcal{L} \times \mathcal{L}\), and \((u, v, r) \in \mathbb{N}^2 \times \mathbb{N}\)

\[
|\text{Cov}(f(X_{i_1}, ... , X_{i_u}), g(X_{j_1}, ... , X_{j_v}))| \leq \Psi(f, g, u, v)\theta_r
\]

whenever \( i_1 < i_2 < ... < i_u \leq r + i_u \leq j_1 < j_2 < ... < j_v \).
The model

GED distribution

Let us consider density function

\[ f(x; \mu, \alpha) = \frac{\alpha}{2A(\alpha)\Gamma(1/\alpha)} \exp\left\{-\left|\frac{x - \mu}{A(\alpha)}\right|^\alpha\right\}, \]

where \( A(\alpha) = \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} \), \( \alpha > 0 \), \( \mu \in (-\infty, \infty) \), and \( x \in \mathbb{R} \).

**Definition**

Random variable \( X \) has a GED distribution, ie. \( X \sim \mathcal{G}(\mu, 1, \alpha) \) if \( f(x) \) the density function of the variable \( X \) is defined by the equation (3).
α-stable variables

**Definition**

(Taqqu, 1994) Random variable $X$ has a **stable** distribution if there exist the parameters: $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\sigma > 0$ and $\mu \in \mathbb{R}$: characteristic function of the variable distribution $X$ has a form:

$$
\varphi(t) = \begin{cases} 
\exp\{-\sigma^\alpha |t|^\alpha (1 - i\beta(t) \tan \frac{\pi \alpha}{2}) + i\mu t\}, & \alpha \neq 1 \\
\exp\{-\sigma |t|(1 + i\beta \frac{2}{\pi}(t) \ln |t|) + i\mu t\}, & \alpha = 1 
\end{cases}
$$

$\alpha \in (0, 2]$ is the stability index, $\beta \in [-1, 1]$ is the skewness parameter, $\tau > 0$ is the scale parameter and $\mu \in \mathbb{R}$ is the location parameter.
The model

**Definition**

Estimator for the mean $\eta$:

$$\hat{\eta} = \frac{1}{n} \sum_{p=1}^{n} X_p,$$  \hspace{1cm} (4)
In the "stable" case let us define $\zeta = \max\{1/\alpha, (\beta + 1)/2\}$, where $\alpha$ is stable parameter, and $\beta$ is memory parameter. Let us consider the statistic:

$$A = n^{1-\zeta}(\hat{\eta} - \eta).$$

In not stable, but heavy tails case let us consider the statistic:

$$B = n^{-1/2}(\hat{\eta} - \eta).$$
In both cases the Central Limit Theorems can be shown for **A1** through **A4**.

- stable case: A. Jach, T. McElroy and D.N. Politis, *Subsampling inference for the mean of heavy-tailed long memory time series*. 
For the main model the subsampling is considered to approximate an asymptotic distribution of a sample mean.
THEOREM

(Consistency of the subsampling)
Let A1 - A4 hold and let CLT assumptions hold, subsampling is consistent, i.e.:

1. If \( x \) is point of continuity \( L \), then \( L_{m_n,b}(x) \overset{P}{\rightarrow} L(x) \).
2. If \( L \) is continuous, then \( \sup_x |L_{m_n,b}(x) - L(x)| \overset{P}{\rightarrow} 0 \).
3. If \( L \) is continuous in \( c(1 - q) \) (where \( c(1 - q) \) is \( q \)-th quantile) then, if \( m_n \to \infty \)

\[
P[m_n^{-1/2}(\hat{\eta}_{m_n} - \eta) \leq c_{m_n,b}(1 - q)] \to 1 - p
\]
Block length

To choose the subsampling block size $b$ - the method of P. Bertail (2011) was implemented.
(Block selection algorithm of Gotze and Rackauskas (2001), Jach, McElroy, Politis (2012).)
Recall that:

\[ L_{n,b}(x) = \frac{1}{n - b + 1} \sum_{i=1}^{n-b+1} I_{\{\tau_{b}(\hat{\eta}_n, b - \hat{\eta}) \leq x\}} \]

\[ \bar{L}_{n,b}(x) = \frac{1}{n - b + 1} \sum_{i=1}^{n-b+1} I_{\{\hat{\eta}_{n,b} - \hat{\eta} \leq x\}} \]
Theorem

Let $\mathbf{A1 - A4}$ hold and let CLT assumptions hold.

$$n^{-\gamma} \bar{L}_{n,b}^{-1}(x) = L^{-1}(x) + o(1),$$

if $L(x)$ is continuous and strictly increasing on $(l_0, l_1)$, where $l_0 = \sup\{x : L(x) = 0\}$ and $l_1 = \inf\{x : L(x) = 1\}$. 
The model

We have

\[ \log(|\overline{L}_b^{-1}(x)|) = \log(|L^{-1}(x)|) + \gamma \log(b) + o_P(1). \]

If we take any \( p_i \neq p_j \in (0, 1) \) and draw log of some quantile range of subsampling distribution

\[ -\log(|\overline{L}_b^{-1}(p_i) - \overline{L}_b^{-1}(p_j)|) = \log(|L^{-1}(p_i) - L^{-1}(p_j)|) + \gamma \log(b) + o_P(1). \]
Choosing the length of the block

It is possible to use the equation above to choose \( b \). The best choice of \( b \) is the largest one before the "unstable" behavior.
$b$ can’t be too small of course, but also it can’t be too big else the subsampling method do not work.

**Rysunek**: Stability of the subsampling distribution
For the simulation study we chose the Gaussian Gegenbauer process with $k=2$, $u_1 = 1$, zero-frequency and $u_2 = \text{acos}(2 \times \pi/8)$, Gegenbauer frequency $= 2 \times \pi/8$. Memory parameters: $d_1 = 0.4$, and $d_2 = 0.2$. Moreover $\sigma_t$ is from the GED family with $\nu = 1$. 
The model
Simulations
Choosing $b$:

**RYSUNEK:** Stability of the subsampling distribution
The model

Simulations

For the simulation study we chose the Gaussian Gegenbauer process with $k=1$, innovations with mean zero and variance 1, $\nu = 1$

For the $\epsilon_t$ we chose $\alpha/2$–stable i.i.d. random variables with the skewed parameter 1, the location parameter 0 and the scale $(\cos(\pi \alpha/4))^{2/\alpha}$. On the picture the $\eta$ is 0.

The number of observations is $NT = 10320$, period $T = 24$. In the first case we took $\beta = 0.3$ and $\alpha = 1.5$.

This is the ”tail” case.

For each $s = 1,\ldots,24$, we found subsample size by the P. Bertail method and then draw the equal-tailed and symmetric 95% confidence intervals.

In the second case we took $\beta = 0.4$ and $\alpha = 1.6$.

This is the ”memory” case.

And for each $s = 1,\ldots,24$, we have done the same as in previous case.
Symulacje modelu heavy tails

Rysunek: PLMHT
Rysunek: Equal tailed confidence interval for the mean parameter of the process $X_{s+p+T}$ with parameters $\alpha = 1.5$, $\beta = 0.3$ and $T = 24$. 
**Rysunek:** Symmetric confidence intervals for the mean parameter of the process $X_{s+p+T}$ with parameters $\alpha = 1.5$, $\beta = 0.3$ and $T = 24$. 
**RYSUNEK:** Equal-tail confidence interval for the mean parameter of the process $X_{s+p+T}$ with parameters $\alpha = 1.6$, $\beta = 0.4$ and $T = 24$. 
Rysunek: Symmetric confidence intervals for the mean parameter of the process $X_{s+p+T}$ with parameters $\alpha = 1.6$, $\beta = 0.4$ and $T = 24$. 
Let us consider the prices from Nord Pool electricity from January 1, 2000 to January 30, 2014. It gives 5144 observations. The Figure 8 shows the data. The daily prices are computed as the average of 24 hourly prices.

**Rysunek**: Daily price for Nord Pool: Prices in Norwegian kroner (NOK) per MWh. The period of observation: January 1, 2000 - January 30, 2014; number of observations is 5144
The period is equal to $T = 7$. For each $s = 1, \ldots, 7$, we calculated the estimator $P$.

**Rysunek:** Stability of the subsampling distribution for daily price for Nord Pool; The period of observation: January 1, 2000 - January 30, 2014

The size of $b$ was calculated by the method proposed by Bertail [2] and is 200.
The period is equal to $T = 7$. For each $s = 1,\ldots,7$, we calculated the estimator $P$.


The size of $b$ was calculated by the method proposed by Bertail [2] and is 200.
Construction of confidence intervals. 
Note that the construction of $P$ does not require knowledge of parameters of memory - $\beta$ and tails - $\alpha$. 
The confidence intervals for $\eta$. (The scale on vertical axle is 100.)

**Rysunek:** The confidence interval for daily price for Nord Pool; The period of observation: January 1, 2000 - January 30, 2014
References
 References


- Bertail P., *Comments on ”Subsampling weakly dependent time series and application to extremes”,* by P. Doukhan, R. Prohl


Thank you for your attention