

Subsampling for Non-stationary Time Series with Long Memory and Heavy Tails Using Weak Dependence Condition

Elżbieta Gajecka-Mirek and Oskar Knapik

Abstract Statistical inference for unknown distributions of statistics or estimators may be based on asymptotic distributions. Unfortunately, in the case of dependent data the structure of such statistical procedures is often ineffective. In the last three decades we can observe an intensive development the of so-called resampling methods. Using these methods, it is possible to directly approximate the unknown distributions of statistics and estimators. A problem that needs to be solved during the study of the resampling procedures is the consistency. Their consistency for independent or stationary observations has been extensively studied. Resampling for time series with a specific non-stationarity, i.e. the periodic and almost periodic strong mixing dependence structure also been the subject of research. Recent research on resampling methods focus mainly on the time series with the weak dependence structure, defined by Paul Doukhan, Louhichi et al. and simultaneously Bickel and Bühlmann. In the article a time series model with specific features i.e.: long memory, heavy tails (with at least a fourth moment, e.g.: GED, t-Student), weak dependence and periodic structure is presented. and the estimator of the mean function in the above-mentioned time series is investigated. In the article the necessary central limit theorems and consistency theorems for the mean function estimator (for one of the resampling techniques—the subsampling) are proven.

E. Gajecka-Mirek (✉)

Department of Economics, State Higher Vocational School in Nowy Sącz,
ul. Jagiellonska 61, Nowy Sącz, Poland
e-mail: egajecka-mirek@pwsz-ns.edu.pl; egajecka@gmail.com

O. Knapik

CREATES, Aarhus University, Fuglesangs Alle 4, 8210 Aarhus V, Denmark
e-mail: oknapik@econ.au.dk

1 Introduction

Possibility to construct sampling distributions of estimators for time series is very important in statistical studies. Traditional statistical inference based on asymptotic distributions does not always lead to effective statistical procedures. There are several reasons for this, e.g.:

- the convergence of the estimator to the asymptotic distribution is slow and often requires a large collection of observations. In practice, there is not always the possibility to receive enough data because of the costs or technical restrictions.
- The asymptotic distribution is often very complicated and depends on the unknown parameters, which in the case of dependent data is difficult to estimate.

In such situations, the so-called resampling methods (Efron [14]) are helpful. Moreover, in many cases these methods are the only effective technique. These methods allow us to approximate the unknown distributions (or characteristics) of the statistics and estimators without a reference to the form of the distribution. These approximations are used to construct the confidence intervals for the parameters and testing statistical hypothesis. (One can compute the confidence intervals and critical values from the resampling distributions instead of the asymptotic distributions.)

In the nineties of the last century the research was focused on stationary time series. At the present time, the efforts of researchers are concentrated on the non-stationary series, with discrete and continuous time. See, e.g. [13, 16, 26, 29] or [35] among many others.

One of the specific form of non-stationarity is periodicity. Gladyshev [18] initiated the development of research on periodicity in time series and stochastic process.

In 2006 Gardner et al. [17] have provided a general overview of research on periodicity and time series, considering over 1500 published papers on the topic. It was shown that the models with periodic structure are widely applicable e.g. in communication, signal processing, vibromechanics, econometrics, climatology and biology.

The resampling methods for periodic time series is an open research area, where many fundamental properties have yet to be proven.

In the article, we will deal only with one form of resampling—subsampling, since we will work with long memory time series.

The attention will be focused on the class of time series which simultaneously deals with three features: periodic structure, heavy tails (finite at least fourth moment) and long memory.

The heavy-tailed random variables are variables with distributions whose extreme values are “more probable than normal”. Examples of such distributions are the Generalized Error Distribution (GED) distributions or stable distributions. Both classes will be discussed in this thesis. Additionally, in real data sets one has to deal with long range dependence as well.

The presence of long range dependence in time series means that there exists dependence between observations which are distant in time from each other.

It is obvious that among the observations of the time series there is a relationship—the dependence. Over the years, the most popular way for studying this dependence

have been the mixing conditions. The fact is that under natural restrictions on the process parameters, many processes of interest fulfill mixing conditions [11]. On the other hand, there is a large class of processes for which mixing properties do not hold.

Simple example of such a time series is a stationary $AR(1)$ process:

$$X_t = aX_{t-1} + \varepsilon_t,$$

where the innovations are independent with $P(\varepsilon_t = 1) = P(\varepsilon_t = -1) = 1/2$ and $0 < |a| \leq 1/2$.

This process has a stationary distribution on $[-2, 2]$ and X_t has always the same sign as ε_t . It is possible to recover $X_{t-1}, X_{t-2}, \dots$ from X_t , it means that the process $\{X_t\}_{t \in \mathbb{Z}}$ is purely deterministic going backwards in time, so it cannot be strong mixing (proof is given in [1]).

More examples of weakly dependence sequences can be found in the research of Doukhan et al. [8, 21].

In 1999 Doukhan and Louhichi [12] and simultaneously Bickel and Bühlmann [6] proposed an alternative condition for the dependence in time series called weak dependence and ν -dependence, respectively.

The definition of weak dependence in comparison to, for example, mixing is very general. It includes very general data sets and models like causal, non causal linear, bilinear, strong mixing processes, dynamical systems or Markov processes driven by discrete innovations.

The main objective of this article is to introduce the theoretical and practical results describing the consistency of subsampling method and to show how to use them in statistical inference for time series with periodic behavior.

Three specific features of time series will be studied: heavy tails, long memory and periodic behavior. The construction of described in the thesis process entails the weak dependence property.

The central limit theorem for the mean estimator will be given. The subsampling method to estimate the mean vector will be presented and the applications of the central limit theorem to prove the consistency of subsampling method will be shown.

The structure of the article is as follows. In the second section the definitions, the examples and the main ideas will be introduced. The third section contains the construction of long memory, heavy tailed and periodically stationary model and its properties. The fourth section contains the central limit theorems. In the next section the consistency of the subsampling is shown.

2 Basic Concepts and Definitions

In this section the basic concepts and definitions will be presented. Some of them will be illustrated by examples.

Many real life phenomena are characterized by a seasonal behavior which, obviously, is not non-stationary. Seasonal data appear in such fields as: economics, biology, climatology, telecommunications and many others. If seasonality is not easily removable it means that we are dealing with a particular type of non-stationarity, for example the periodic structure. In such cases it is not just the mean that has a periodic rhythm. A periodic rhythm also describes the behavior of covariance.

Popular models used for describing such phenomena are periodically nonstationary processes.

Below are introduced the formal definitions of periodicity of the time series.

For the time series $\{X_t\}_{t \in \mathbb{Z}}$ we define the autocovariance of the pair (X_t, X_{t+h}) to be

$$\gamma_X(t, h) = \text{Cov}(X_t, X_{t+h}).$$

Definition 1 ([25], p. 5) Time series $\{X_t\}_{t \in \mathbb{Z}}$ is periodically correlated (PC) in Gladyshev sense, if the mean $\mu_X(t)$ is periodic ($\mu_X(t) = \mu_X(t + T)$) and the autocovariance function $\gamma_X(t, h)$ is periodic in t for all $h \in \mathbb{Z}$ ($\gamma_X(t, h) = \gamma_X(t + T, h)$).

If there is no ambiguity, we will write $\gamma(t, h)$ (or $\gamma(h)$ if we deal with classical weak stationarity) instead of $\gamma_X(t, h)$ for the autocovariance function of time series $\{X_t\}_{t \in \mathbb{Z}}$.

In many periodic phenomena the existence of the long range dependence is observed [5, 24, 31]. The presence of long range dependence in time series means that there exists a relationship between observations which are far away from each other in time. Long memory occurs in the sense that a hyperbolic behavior of the autocorrelations holds for almost all lags and frequencies respectively.

The long range dependence can be defined as long memory. Note that the definition of long memory introduced below is one of many possible definitions.

Definition 2 ([7], p. 520) A stationary, time series $\{X_t\}_{t \in \mathbb{Z}}$ has long memory if its autocovariance function γ satisfies the following formulas:

$$\sum_{0 < |h| < n} \gamma(h) \sim Cn^\beta$$

where $\beta \in [0, 1)$, and $C \neq 0$.

Definition 3 A PC time series $\{X_t\}_{t \in \mathbb{Z}}$ has a long memory if the autocovariance function $\gamma(s)(h) = \text{Cov}(X_{s+qT}, X_{s+(q+h)T})$ for each $q \in \mathbb{Z}$ satisfies the following formula

$$\sum_{0 < |h| < n} \gamma(s)(h) \sim C(s)n^\beta, \quad s \in \{1, \dots, T\}$$

where $\beta \in [0, 1)$. For each $s \in \{0, \dots, T-1\}$ $C(s)$ is the finite constant such that

$$C(s) = 2 \cdot \lim_{n \rightarrow \infty} \frac{\sum_{h=1}^{n-1} \gamma(s)(h)}{n^\beta} > 0.$$

Let us assume that the notation for the long memory with parameter $\beta \in [0, 1)$ will be $LM(\beta)$.

Suitable tool to describe the long memory behavior are so-called Gegenbauer processes [15, 23]. This kind of process will be used in modeling the long range dependence in this article.

Definition 4 ([19]) Let us assume that ε_t is i.i.d. innovation process. The process $\{G_t\}_{t \in \mathbb{Z}}$ defined by the equation:

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

is the k -factor Gegenbauer process.

$0 < d_i < 1/2$ if $|v_i| < 1$ or $0 < d_i < 1/4$ if $|v_i| = 1$ for $i = 1, \dots, k$ and I is identity operator, B is backshift operator.

Theorem 1 ([21]) *Process defined by the Definition 4 is long memory, stationary, causal and invertible and has a moving average representation:*

$$G_t = \sum_{j \geq 0} \psi_j(d, v) \varepsilon_{t-j},$$

with $\sum_{j=0}^{\infty} \psi_j^2(d, v) < \infty$, where $\psi_j(d, v), j \geq 0$, is defined by:

$$\psi_j(d, v) = \sum_{\substack{0 \leq l_1, \dots, l_n \leq j \\ l_1 + \dots + l_n = j}} C_{l_1}(d_1, v_1) \cdot \dots \cdot C_{l_k}(d_k, v_k),$$

where $C_{l_i}(d_i, v_i)$ are the Gegenbauer polynomials defined as follows:

$$(1 - 2vz + z^2)^{-d} = \sum_{j \geq 0} C_j(d, v) z^j, \quad |z| \leq 1, \quad |v| \leq 1.$$

Moreover, if $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ in the Definition 4 is the Gaussian white noise, then $\{G_t\}_{t \in \mathbb{Z}}$ is Gaussian time series.

The Gegenbauer processes are stationary, seasonal fractional models; see: [19, 20]. The form of the $\psi_j(d, v)$ also can be found in [19, 20].

The time series is a sequence of dependent observations. In 1999 Doukhan and Louhichi [12] and simultaneously Bickel and Bühlmann [6] introduced a new way of describing data—the weak dependence.

Below, the definition of weak dependence is introduced.

Let $(E, \|\cdot\|)$ be a normed space and $u \in \mathbb{N}^*$. We assume that a function $h : E^u \rightarrow \mathbb{R}$ belongs to the class $\mathcal{L} = \{h : E^u \rightarrow \mathbb{R}, \|h\|_\infty \leq 1, \text{Lip}(h) < \infty\}$, 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 5 ([4]) A sequence $\{X_t\}_{t \in \mathbb{Z}}$ of random variables taking values in $E = \mathbb{R}^d$ ($d \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$) is $(\varepsilon, \mathcal{L}, \Psi)$ -weakly dependent if there exists $\Psi : \mathcal{L} \times \mathcal{L} \times \mathbb{N}^* \times \mathbb{N}^* \rightarrow \mathbb{R}$ and a sequence $\{\varepsilon_r\}_{r \in \mathbb{N}}$ ($\varepsilon_r \rightarrow 0$) such that for any $(f, g) \in \mathcal{L} \times \mathcal{L}$, and $(u, v, r) \in \mathbb{N}^{*2} \times \mathbb{N}$

$$|\text{Cov}(f(X_{i_1}, \dots, X_{i_u}), g(X_{j_1}, \dots, X_{j_v}))| \leq \Psi(f, g, u, v) \varepsilon_r$$

whenever $i_1 < i_2 < \dots < i_u \leq r + i_u \leq j_1 < j_2 < \dots < j_v$.

The weak dependence notions are related to the initial time series and are measured in terms of covariance of the functions.

The asymptotic behavior of the covariance shows us the independence between “past” and “future”. Intuitively, the weak dependence is “forgetting” in time series.

There are several concepts of weakly dependent processes e.g.: λ -, θ -, η -weak dependence. The form of the Ψ function corresponds to particular cases of weak dependence.

The coefficient λ corresponds to:

$$\Psi(f, g, u, v) = uv \text{Lip}(f) \text{Lip}(g) + u \text{Lip}(f) + v \text{Lip}(g),$$

the coefficient η corresponds to:

$$\Psi(f, g, u, v) = u \text{Lip}(f) + v \text{Lip}(g),$$

the coefficient θ corresponds to:

$$\Psi(f, g, u, v) = v \text{Lip}(g).$$

Note there are other cases of weakly dependent, that are not quoted here.

In the definition of weak dependence we denote respectively λ_r , η_r or θ_r instead of ε_r .

The heavy-tailed random variables are variables with distributions whose extreme values are “more probable than normal”. The heavy tail phenomena occur frequently in real life. In contradiction to Gaussian phenomena which do not allow for large fluctuations, “heavy tails” can be used to describe high variability. The data with “heavy tails” appear in such different fields as economics, telecommunications, meteorology, physics and signal processing.

If we define kurtosis as μ_4/σ^4 and μ_4 is the fourth central moment (if it exists), while σ is the standard deviation then we can say that heavy-tailed variables are those

with kurtosis greater than three, and whose tails go to zero slower than in the normal distribution.

One of the example of heavy tailed distributed random variables are GED—random variables.

The Generalized Error Distribution (GED) is a parametric model of a heavy tailed distribution. All moments of the GED are finite and the GED has a relatively simple form of a probability density function.

The Generalized Error Distribution is a symmetric unimodal member of the exponential family. The domain of the probability distribution function is $(-\infty, \infty)$.

The original concept of the GED was introduced by Subbotin in 1923 [34], so it is known as “*Subbotin’s family of distributions*”. However, Subbotin proposed a two-parameters GED model:

$$f(x; h, m) = \frac{mh}{2\Gamma(1/m)} \exp\{-h^m |x|^m\}, \quad (2)$$

where $x \in \mathbb{R}$ and $h > 0$ and $m \geq 1$ are scale and shape parameters, respectively.

In 1963 Lunetta ([30]) has defined a three-parameters GED class, as follows:

$$f(x; \mu, \sigma, \alpha) = \frac{1}{2\sigma\alpha^{1/\alpha}\Gamma(1+1/\alpha)} \exp\left\{-\frac{1}{2}\left|\frac{x-\mu}{\sigma}\right|^\alpha\right\}, \quad (3)$$

where $\mu \in \mathbb{R}$ is the location parameter, $\tau > 0$ is the scale and $\alpha > 0$ is the shape (power).

Of course, the m in the Eq. (2) is equal to the α in the Eq. (3) while $h = (\alpha^{1/\alpha}\tau)^{-1}$. Taking into account the fact that the Euler gamma function Γ satisfies the formula $r\Gamma(r) = \Gamma(r+1)$, the Eqs. (2) and (3) are equivalent, whenever the location parameter μ in (3) is equal to zero.

The GED is also called the generalized normal class. The reason is that for the random variable X with density function as in the formula (3) we have the following equation:

$$\tau = \{E|X - \mu|^\alpha\}^{1/\alpha},$$

which for $\alpha = 2$, gives the standard deviation in the normal case. Note that if $\alpha \neq 2$, τ must not be confused with the standard deviation of X .

Below we give the definition in which τ stands for the standard deviation:

Definition 6 ([36]) The random variable X has GED distribution ($X \sim G(\mu, \tau, \alpha)$) if the density function, $f(x)$, of X is given by the equation:

$$f(x; \mu, \tau, \alpha) = (2\Gamma(1+1/\alpha)A(\tau, \alpha))^{-1} \exp\left\{-\left|\frac{x-\mu}{A(\tau, \alpha)}\right|^\alpha\right\} \quad (4)$$

with $A(\tau, \alpha) = \tau \sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)}$.

In the sequel of this dissertation it is enough to consider the case $\tau = 1$ in the Eq. (4). Therefore, we will be considering the density function

$$f(x; \mu, \alpha) = \frac{\alpha}{2A(\alpha)\Gamma(1/\alpha)} \exp\left\{-\left|\frac{x-\mu}{A(\alpha)}\right|^\alpha\right\}, \quad (5)$$

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

Our definition of the GED is as follows:

Definition 7 The random variable X has a GED distribution $X \sim G(\mu, 1, \alpha)$ if the density function, $f(x)$, of X follows the Eq. (5).

The r th central moment of a random variable $X \sim G(\mu, 1, \alpha)$ can be calculated as

$$E(X - EX)^r = \frac{1}{\sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)}\Gamma(1 + 1/\alpha)} \int_{-\infty}^{\infty} (x - EX)^r e^{-\frac{1}{2}|x-EX|^\alpha} dx,$$

where $r \in \mathbb{N}$. When r is odd then the r th moments are equal to zero, by symmetry. For r even the r th moments are as follow:

$$EX^r = (\sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)})^r \frac{\Gamma(1/\alpha(r+1))}{\Gamma(1/\alpha)}.$$

Notice that in the consequence of the Definition 7 the first four moments of GED distribution are: mean = μ , variance = 1, skewness = 0, kurtosis = $\frac{\Gamma(5/\alpha)\Gamma(1/\alpha)}{\Gamma^2(3/\alpha)}$.

Recall that the GED distribution has heavy tails, if $\alpha < 2$. When $\alpha > 2$ we get tails lighter than normal.

It is clear to see that the normal distribution is GED with $\alpha = 2$. Below are other examples of the GED distribution.

Example 1 If we choose $\alpha = 1$ in the Definition 6 then the GED distribution is so-called Double Exponential, or Laplace, distribution, i.e. $G(\mu, \tau^2, 1) = L(\mu, 4\tau^2)$.

Note that the Subbotin's model (2) does not allow for the tails heavier than those in the Laplace distribution. Unlike the formula (3), where the tails heavier than those in the Laplace distribution are allowed.

Example 2 If we consider the Definition 6 and $\alpha \rightarrow \infty$, then the GED distribution tends to the uniform distribution $U(\mu - \tau, \mu + \tau)$.

3 The Model and Its Properties

In many applications of time series analysis one is confronted separately with heavy tailed and long memory behavior. The non-stationarity of the time series, and its special case—the periodicity is also a feature that researchers are dealing with.

Below we present a model which will simultaneously be dealing with three features: periodicity, long memory and heavy tails. We build it by adjusting a long memory and a heavy tailed stationary model to the T-variate process.

Let the time series $\{X_t\}_{t \in \mathbb{Z}}$ be defined as:

$$X_t = \sigma_t GG_t + \eta_t, \quad (6)$$

where

- A1** The volatility time series σ_t and the Gaussian-Gegenbauer time series GG_t are independent
- A2** The sequence of random variables σ_t is i.i.d and its marginal distribution comes from a heavy tail family, with finite at least fourth moment, for example.
- A3** GG_t is periodic Gaussian-Gegenbauer time series. We put that $GG_t = f_t \cdot G_t$, where G_t is Gaussian-Gegenbauer mean zero time series with $k = 1$, $|v| \leq 1$, $LM(\beta)$ with $\beta \in [0, 1)$. The function f_t is a periodic, deterministic, bounded with a known period T . The autocovariance of G_t is γ_G .
- A4** The deterministic function η_t is periodic with the same period T as f_t .

$\beta = 2d$, where d is a memory parameter from the Definition 4.

Properties of the model

Lemma 1 *The process $\{X_t\}_{t \in \mathbb{Z}}$ defined by the Eq. (6) is a long memory process in the sense of Definition 3, with $\beta \in [0, 1)$.*

Proof

$$\sum_{0 < |h| < n} \gamma(s)(h) = \sum_{0 < |h| < n} (E\sigma)^2 f_h^2 \gamma_G(h) \sim C(s)n^\beta.$$

The last asymptotic equivalence follows from Theorem 2 [10]. □

Moreover:

Lemma 2 ([21], Theorem 2—[11]) *The long memory stationary Gaussian-Gegenbauer time series is not strong mixing.*

Lemma 3 *Assume A1 through A4. Then X_t defined by the Eq. (6) is λ -weakly dependent.*

Proof It follows from Lemma 2 that the Gaussian-Gegenbauer time series $G_{t \in \mathbb{Z}}$, is not strong mixing.

We already know that Gegenbauer (in the sense of the Definition 4) time series has a long memory. And from the [4], p. 8 follows that the stationary Gaussian long memory time series has the λ -weak dependence properties.

Finally the λ -weak dependence of X_t is implied from the Proposition 1 in [27].

Lemma 4 *Assume A1 through A4. Then the weak dependence coefficients of the model defined by the Eq. (6) satisfy the following relationship:*

$$\lambda_r = O(r^{\beta-1}), \quad \beta \in [0, 1).$$

Proof of above lemma follows from the Lemma 4 [9].

The Lemma 3 provide clear motivation to study weakly dependent structures.

GED volatility or other heavy tailed distribution with finite at least fourth moment

Assume that the volatility process σ_t in the model (6) comes from a GED distribution, as in Definition 7, i.e. $\sigma_t \sim G(\mu, 1, \alpha)$.

The following fact provides the information about the heaviness of X_t tails in the GED case:

Lemma 5 *Assume A1 through A4. Then X_t defined by the Eq. (6) has a heavy tailed marginal distribution for $\alpha > 0$.*

We do not have information about the marginal distribution of X_t , but since all the moments of both distribution GG_t and σ_t exist the model X_t is $(2 + \delta)$ -order, with $\delta > 0$.

In particular X_t is second-order, so it is periodically correlated in the sense of Gladyshev. The formalization of this statement is as follows:

Lemma 6 *X_t defined by the Eq. (6) with σ_t coming from the GED has a periodic mean, a periodic variance and a periodic autocovariance. Moreover, the autocovariance of X_t has a form:*

$$\gamma(t, h) = (Cov(\sigma_t, \sigma_{t+h}) + \varphi^2)|f_t||f_{t+h}|\gamma_G(h).$$

The estimator and its properties

For the model defined by the Eq. (6) one of the resampling method—subsampling—is considered to approximate an asymptotic distribution of the seasonal trend components, the overall mean and the vector of the seasonal trend components.

We start with the definition.

Definition 8 We define the estimator of the seasonal trend components $\eta(s)$ as follows:

$$\hat{\eta}_N(s) = \frac{1}{N} \sum_{p=0}^{N-1} X_{s+pT}, \quad s = 1, 2, \dots, T, \quad (7)$$

where T is the known period.

Definition 9 We define the estimator of the overall mean $\bar{\eta} = \frac{1}{T} \sum_{s=1}^T \eta(s)$ as follows:

$$\hat{\bar{\eta}} = \frac{1}{T} \sum_{s=1}^T \hat{\eta}_N(s) \quad (8)$$

where T is the known period.

4 Central Limit Theorems in the GED Case

In this subsection, the central limit theorems for the seasonal and overall mean for the variables from the GED are introduced.

Let us define $\Sigma_s = \text{Var}(X_s)$. In the GED case $\Sigma_s = |f_s|^2 \gamma_G(h)$, $s = 1, \dots, T$.

Let us define:

$$B_N(s) = N^{1/2}(\hat{\eta}_N(s) - \eta(s)).$$

For a sequence $\{m_N\}_{N \in \mathbb{N}}$ such that $m_N \rightarrow \infty$, if $N \rightarrow \infty$ let us define $\{k_N\}_{N \in \mathbb{N}}$ such that

$$k_N = \lfloor \frac{N}{m_N} \rfloor \rightarrow \infty, \text{ if } N \rightarrow \infty$$

and for λ -weakly dependent model $\{X_t\}_{t \in \mathbb{Z}}$ defined by the Eq. (6) following condition holds

$$\lambda_{m_N} k_N^{\frac{3}{2}} \rightarrow 0, \quad N \rightarrow \infty.$$

Let us consider a subsample $(X_{m_N}, \dots, X_{k_N m_N})$ of $(X_s, \dots, X_{s+(N-1)T})$.

Theorem 2 (Central Limit Theorem—for the seasonal means) *Assume A1 through A4 and the volatility process σ_t is as in the Definition 7. Then, for a sequence $\{m_N\}_{N \in \mathbb{N}}$ such that $m_N \rightarrow \infty$ and $k_N = \lfloor \frac{N}{m_N} \rfloor \rightarrow \infty$, if $N \rightarrow \infty$ and for each $s = 1, \dots, T$ following convergence holds:*

$$B_{k_N}(s) \rightarrow \mathcal{N}(0, \Sigma_s), \quad N \rightarrow \infty.$$

The proof follows from Proposition 4.1, [4].

Since the model defined by the Eq. (6) satisfies property $LM(\beta)$ (Definition 2), for $\beta \in [0, 1)$ the time series X_t defined by the Eq. (6) does not satisfy any known Central Limit Theorem. The reason for this fact is that it has the long memory property. But the Central Limit Theorem will be satisfied if we choose a subsample of the observation with the appropriate asymptotic step of sampling. In the model (6) the subsampled time series $\{Y_{sm_N} = X_{sm_N} - \eta_{sm_N}\}$, with a “subsampling” step m_N such that

$$o(m_N) = N^{3/(2\beta+1)}, \quad (9)$$

satisfies the Central Limit Theorem with a convergence rate $o(N^{(1-\beta)/(4\beta+2)})$.

The Eq. (9) gives us the restriction for β it means: $\beta \in (0, 1/2)$.

Following [4] there are objections to this “subsampling” method: only a part of the sample is used and the choice of the convergence rate of the “subsampling” implies the knowledge of the convergence rate of λ_r . But the convergence rate of λ_r , in long memory processes is connected with the the long memory parameter β . It could give us a step of “subsampling”.

Let us define:

$$B_{k_N, N} = k_N^{1/2}(\hat{\eta} - \eta).$$

Theorem 3 (Central Limit Theorem—for the vector of the means) *Assume A1 through A4. Moreover, assume that*

$$\lambda_{m_N} k_N^{\frac{3}{2}} \rightarrow 0, \quad N \rightarrow \infty.$$

Let $Z_N = (Y_{s+pT}, \dots, Y_{s+(p+1)T})$, $p = 0, \dots, N-1$ be a sequence of zero mean random variables with values in $\mathbb{R}^T$, where T is the period. Then, for a sequence $\{m_N\}_{N \in \mathbb{N}}$ such that $m_N \rightarrow \infty$ and $k_N = \lfloor \frac{N}{m_N} \rfloor \rightarrow \infty$, if $N \rightarrow \infty$

$$B_{k_N, N} = \frac{1}{\sqrt{k_N}} \sum_{i=1}^{k_N} Z_{im_N} \rightarrow \mathcal{N}_T(0, \text{Cov}(X_0)), \quad N \rightarrow \infty.$$

The proof of Theorem 3 implies from the Proposition 4.1, [4], but applied to the vectors.

Theorem 4 (Central Limit Theorem—for the overall mean) *Assume A1 through A4. Moreover assume that*

$$\lambda_{m_N} k_N^{\frac{3}{2}} \rightarrow 0, \quad N \rightarrow \infty.$$

Then, for a sequence $\{m_N\}_{N \in \mathbb{N}}$ such that $m_N \rightarrow \infty$ and $k_N = \lfloor \frac{N}{m_N} \rfloor \rightarrow \infty$, if $N \rightarrow \infty$

$$B = \frac{1}{\sqrt{k_N T}} \sum_{s=1}^T \left(\sum_{p=0}^{k_N-1} (X_{sm_N+pT} - \eta(s)) \right) \rightarrow \mathcal{N}(0, \Sigma), \quad N \rightarrow \infty.$$

The proof of the Theorem 4 implies from the Proposition 4.1, [4] for the vectors and the Cramer–Wald theorem.

5 Consistency of the Subsampling Method for the Mean

From research of Hall and Lahiri [22, 28] we know that, for long-range dependent process, the bootstrap (MBB, CBB, PBB, SBB, GSBB) do not work, whereas subsampling still works asymptotically.

All we need to know to use subsampling is if there exists a non-degenerated asymptotic distribution of the statistic (we do not have to know the form of the asymptotic distribution)—we need to have the Central Limit Theorem.

Recall that our sample size n is equal $n = NT$.

The idea of subsampling in our model for the seasonal components is as follows:

1. For each $s = 1, \dots, T$ the estimator $\hat{\eta}_N(s)$ is recomputed from the $(X_s, \dots, X_{s+(N-1)T})$ over “short” overlapping blocks of length b_s (b_s depends on N -the length of the sample)
2. From Step 1 $N - b_s + 1$ statistics are obtained for each s . In our context those will be $a_{b_s}(\hat{\eta}_{N,b_s,i}(s) - \hat{\eta}_N(s))$ where $\hat{\eta}_{N,b_s,i}(s)$ is subsampling version of the estimator $\hat{\eta}_N(s)$ and a_{b_s} is the normalize sequence.
3. then the empirical distributions:

$$L_{N,b_s}(x, s) = \frac{1}{N - b_s + 1} \sum_{i=1}^{N-b_s+1} 1_{\{a_{b_s}(\hat{\eta}_{N,b_s,i}(s) - \hat{\eta}_N(s)) \leq x\}}$$

are used to approximate the asymptotic distribution $L(s)(x)$ of the estimator $a_N(\hat{\eta}_N(s) - \eta(s))$.

The idea of subsampling in our model for the vector of seasonal components is as follows:

1. For each $s = 1, \dots, T$ the estimator $\hat{\eta}_N$ is recomputed from the $(X^1, \dots, X^N)$, where $X^i = (X_{1+(i-1)T}, \dots, X_{iT})$ over “short” overlapping blocks of length b (b depends on N -the length of the sample)
2. From Step 1 $N - b + 1$ statistics are obtained. In our context those will be $a_b(\hat{\eta}_{N,b} - \hat{\eta}_N)$ where $\hat{\eta}_{N,b}$ is subsampling version of the estimator $\hat{\eta}_N$ and a_b is the normalizing sequence.
3. then the empirical distributions: $L_{N,b}(x) = \frac{1}{N-b+1} \sum_{i=1}^{N-b+1} 1_{\{a_b(\hat{\eta}_{N,b,i} - \hat{\eta}_N) \leq x\}}$ are used to approximate the asymptotic distribution $L(x)$ of the estimator $a_N(\hat{\eta}_N - \eta_N)$.

The main problem with the subsampling procedure is its consistency. We need to prove that the finite sample quantiles generated by the subsampling procedure converge asymptotically ($N \rightarrow \infty$) to the quantiles of the asymptotic distribution.

Now we consider the problem of consistency of the Subsampling.

To prove the consistency of the subsampling procedure we need to know if there exists a non-degenerated asymptotic distribution of the statistic which means that we need to have the central limit theorem.

In both the stable and the GED case we have obtained a weakly convergence to the limit random variables.

Denote the cumulative distribution functions of this limit random variables by $L(s)$ and L for the seasonal means and vector of the seasonal means, respectively.

To fulfill the Step 3 of the subsampling procedure the empirical distribution functions $L_N(s)(x) = P(P_N(s) \leq x)$ are computed from the subsamples $(X_s, \dots, X_{s+(N-1)T})$, for each $s = 1, \dots, T$ where X_t is defined by the Eq. (6).

The conclusion from the central limit theorems is that the empirical distribution functions converge weakly to the cumulative distribution functions of the limit random variables

$$L_N(s)(x) \rightarrow L(s)(x) \quad \text{if} \quad N \rightarrow \infty, \quad s = 1, \dots, T.$$

Denote the density of the limit distribution by $L'(s)$. It is obvious that in the GED case $\|L'(s)\|_\infty < \infty$.

Theorem 5 (Consistency theorem for seasonal means in GED cases) *Assume that CLT holds and assume **A1** through **A4**, then consistency of the subsampling method holds for the “subsampling” step (9):*

1. If x is the point of the continuity of $L(s)$, then $L_{k_N, b_s}(s)(x) \xrightarrow{P} L(s)(x)$.
2. If L is continuous then $\sup_x |L_{k_N, b_s}(s)(x) - L(s)(x)| \xrightarrow{P} 0$.
3. If $L(s)$ is continuous in $c(1 - q)$ (where $c(1 - q)$ is a q -quantile) then if $k_N \rightarrow \infty$

$$P[k_N^{-1/2}(\hat{\eta}_{k_N}(s) - \eta(s)) \leq c_{k_N, b}(1 - q)] \rightarrow 1 - q$$

in GED case.

Where $\alpha \in (0, 1)$ and

$$c_{k_N, b}(1 - \alpha) = \inf \{x : L_{k_N, b}(s)(x) \geq 1 - \alpha\},$$

$$c(1 - \alpha) = \inf \{x : L(s)(x) \geq 1 - \alpha\}.$$

The $\xrightarrow{P}$ denotes convergence in probability.

Theorem 6 (Consistency theorem for vector of the seasonal means in GED cases) *Assume that CLT holds and assume **A1** through **A4**, then consistency of the subsampling method holds for the subsequence (9):*

1. If x is the point of the continuity of L , then $L_{k_N, b}(x) \xrightarrow{P} L(x)$.
2. If L is continuous then $\sup_x |L_{k_N, b}(x) - L(x)| \xrightarrow{P} 0$.

6 Conclusions

In the paper the model of the periodically correlated, long memory and heavy tailed time series was described. Estimator of the vector sample means of this particular model was introduced and Central Limit Theorem was proved. Furthermore, consistency of chosen method of estimation, the subsampling, was investigated. The question to be answered in further research is the extension of the model (6) to the class almost periodically correlated processes.

Appendix

In the appendix the proofs of some results are presented.

Proof Lemma 5

From direct calculations we can obtain the strict formula for the kurtosis of the model X_t , which is:

$$\frac{E(X_t - EX_t)^4}{(E(X_t - EX_t)^2)^2} = \frac{E\sigma_t^4 E G_t^4}{(E\sigma_t^2)^2 (E G_t^2)^2} = 3 \frac{\Gamma(5/\alpha) \Gamma(1/\alpha)}{\Gamma^2(3/\alpha)}.$$

If we use the Stirling's formula for Γ function we will obtain the approximation as follows:

$$kurtosis \approx 3 \cdot 1.4 \cdot 4.3^{1/\alpha}.$$

The kurtosis is more than 3 for all $\alpha > 0$. □

Proof Lemma 6

The mean of X_t is η_t , so it is periodic. The variance is periodic:

$$\begin{aligned} \gamma(t + T, 0) &= (f_{t+T})^2 (1 + \varphi^2) \gamma_G(t + T, 0) = \\ &= (f_{t+T})^2 (1 + \varphi^2) \gamma_G(0) = (f_t)^2 (1 + \varphi^2) \gamma_G(t, 0) = \gamma(t, 0). \end{aligned}$$

The autocovariance also:

$$\begin{aligned} \gamma(t + T, h) &= |f_{t+T} f_{t+T+h}| \varphi^2 \gamma_G(t + T, h) = \\ &= |f_{t+T} f_{t+T+h}| \varphi^2 \gamma_G(h) = |f_t f_{t+h}| \varphi^2 \gamma_G(t, h) = \gamma(t, h). \end{aligned}$$

The form of the variance and autocovariance follows from the form of the variance of variable with the GED distribution. □

Proof Theorem 5

Let us consider a sequence of statistics $B_{k_N}(s)$, for fixed $s = 1, 2, \dots, T$ and k_N as in the Theorem 2.

$L_{k_N}(s)(x) = P(B_{k_N}(s) \leq x)$ is cumulative distribution function of $B_{k_N}(s)$.

From the assumptions

$$\sup_{x \in \mathbb{R}} |L_{k_N}(s)(x) - L(s)(x)| \longrightarrow 0, \quad k_N \rightarrow \infty$$

For overlapping samples the number of subsamples:

$Y_{b,q}(s) = (X_{s+qT}, X_{s+(q+1)T}, \dots, X_{s+(q+b-1)T})$, $q = 0, 1, \dots, k_N - b$ and the number of subsampling statistics:

$B_{k_N,b,q}(s) = \sqrt{b}(\hat{\eta}_{k_N,b,q}(s) - \hat{\eta}_{k_N}(s))$ is $k_N - b + 1$.

Above statistics are used to approximate the distributions $L_{k_N}(s)(x)$ by empirical distribution functions: $L_{k_N,b,q}(s)(x) = \frac{1}{k_N - b + 1} \sum_{q=0}^{k_N-b} \mathbb{I}_{\{B_{k_N,b,q}(s) \leq x\}}$.

Let us define subsampled distribution:

$$U_{k_N,b,q}(s)(x) = \frac{1}{k_N - b + 1} \sum_{q=0}^{k_N-b} \varphi(\sqrt{b}(\hat{\eta}_{k_N,b,q}(s) - \eta_{k_N}(s))\varepsilon_n).$$

The sequence ε_n is decreasing to zero and φ is the non-increasing continuous function such that $\varphi = 1$ or 0 according to $x \leq 0$ or $x \geq 1$ and which is affine between 0 and 1 .

From [32] it is known that

$$\forall x \in \mathbb{R} \quad |L_{k_N,b,q}(s)(x) - U_{k_N,b,q}(s)(x)| \xrightarrow{p} 0.$$

It follows that it is enough to investigate only the variance of $U_{k_N,b,q}(s)$, $s = 1, \dots, T$

It is enough to investigate the variance of $U_{k_N,b_s,p}(s)$, $s = 1, \dots, T$ (Theorem 3.2.1 ([32])).

$$\text{Var} U_{k_N,b_s,p}(s)(x) = (k_N - b_s + 1)^{-2} \left(\sum_{|h| < k_N - b_s + 1} (k_N - b_s + 1 - |h|) \gamma(h) \right)$$

here $\gamma(h) = \text{Cov}(\varphi(B_{k_N,b,p}(s)/\varepsilon_n), \varphi(B_{k_N,b,p+h}(s)/\varepsilon_n))$.

From the assumption that we have λ -weak dependence and under condition **A1** through **A4** and the Lemma 3.1 in ([2])

$$\text{Cov}(\varphi(B_{k_N,b,p}(s)/\varepsilon_n), \varphi(B_{k_N,b,p+h}(s)/\varepsilon_n)) \leq \sqrt{b_s} \lambda_{h-b+1} / \varepsilon_n$$

It implies that $\text{Var}(U_{k_N,b_s,p}(s)(x))$ tends to zero, it proves point 1. of the Theorem 5.

The proof of point 2. and 3. if the 1. holds and under assumption of the model (6) is the same as the proof of 3. in the Theorem 3.2.1 ([32]). $\square$

Proof Theorem 6

For any vector of constants $c \in \mathbb{R}^T$ we have the equation for the subsampling version of the characteristic functions of the distributions:

$$\phi_{B_{k_N, b, q}}^*(c) = \phi_{c^T B_{k_N, b, q}}^*(1) \text{ in GED case}$$

Let $Z_{s+pT} = c_s X_{s+pT}$, where $p = 0, \dots, k_N - 1$ and $s = 1, \dots, T$. The series $\{Z_t\}$ fulfills the assumptions of Theorem 6, which means that subsampling is consistent for the mean $(\eta_N)_Z$. By Theorem A in Athreya [3] we have:
in GED case

$$\phi_{c^T B_{k_N, b, q}}^*(1) \xrightarrow{P} \phi_{k_N(\eta, c^T \Sigma c)}(1) = \phi_{k_N(\eta, \Sigma)}(c).$$

Moreover

$$P^*(B_{k_N, b, q} \leq x) \xrightarrow{P} F_{k_N(\eta, \Sigma)}(x),$$

for any $x \in \mathbb{R}^T$, where $F_{k_N(\eta, \Sigma)}(x)$ is the cumulative distribution function of $k_N(\eta, \Sigma)$. The second point of the thesis of the Theorem 6 follows then from Pólya's theorem ([33], p. 447). $\square$

References

1. Andrews, D. W. K. (1984). Non-strong mixing autoregressive processes. *Journal of Applied Probability*, 21, 930–934.
2. Ango-Nze, P., Dupoirson, S., & Rios, R. (2003). Subsampling under weak dependence condition, Technical Report DT2003-42. CREST
3. Athreya, K. B. (1987). Bootstrap of the mean in the infinite variance case. *The Annals of Statistics*, 15, 724–731.
4. Bardet, J.-M., Doukhan, P., Lang, G., & Ragache, N. (2008). Dependent Lindeberg central limit theorem and some applications. *ESAIM: Probability and Statistics*, 12, 154–172
5. Beran, J., Feng, T., Ghosh, S., & Kulik, R. (2013). *Long-memory processes. Probabilistic properties and statistical methods*. New York: Springer.
6. Bickel, P., & Bühlmann, P. (1999). A new mixing notion and functional central limit theorems for a sieve bootstrap in time series. *Bernoulli*, 5–3, 413–446.
7. Brockwell, P., & Davis, R. (1991). *Time series: Theory and methods*. New York: Springer.
8. Dedecker, J., Doukhan, P., Lang, G., León, J. R., Louhichi, S., & Prieur, C. (2008). *Weak dependence: With examples and applications*. Lecture notes 190 in statistics. Springer
9. Doukhan, P., & Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, 84, 313–342
10. Doukhan, P., Prohl, S., & Robert, C. Y. (2011). Subsampling weakly dependent times series and application to extremes. *TEST*, 20(3), 487–490.
11. Doukhan, P. (1994). *Mixing: Properties and examples*. Lecture notes in statistics 85. Springer
12. Doukhan, P., & Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, 84, 313–342.

13. Dudek, A. E., Leśkow, J., Politis, D., & Paparoditis, E. (2014). A generalized block bootstrap for seasonal time series. *Journal of Time Series Analysis*, 35, 89–114.
14. Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7, 1–26.
15. Ferrara, L., & Guégan, D. (2001). Forecasting with k-factor Gegenbauer processes: theory and applications. *Journal of Forecasting*, 20(8), 581–601.
16. Gajeczka-Mirek, E. (2014). Subsampling for weakly dependent and periodically correlated sequences. In *Cyclostationarity: Theory and methods*. Lecture notes in mechanical engineering. Springer.
17. Gardner, W., Napolitano, A., & Paura, L. (2006). Cyclostationarity: Half a century of research. *Signal Processing*, 86
18. Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet Mathematics*, 2
19. Gray, H. L., Zhang, N.-F., & Woodward, W. A. (1989). On generalized fractional processes. *Journal of Time Series Analysis*, 10, 233–257.
20. Gray, H. L., Zhang, N.-F., & Woodward, W. A. (1994). On generalized fractional processesa correction. *Journal of Time Series Analysis*, 15(5), 561–562.
21. Guégan, D., Ladoucette S.(2001). Non-mixing properties of long memory sequences. *Comptes Rendus de l' Academie des Sciences Paris*, 333, 373–376
22. Hall, P., Jing, B.-Y., & Lahiri, S. (1998). On the sampling window method for long-range dependent data. *Statistica Sinica*, 8, 1189–1204.
23. Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68, 165–176.
24. Hui, Y. V., & Li, W. K. (1995). On Fractionally differenced periodic processes. *The Indian Journal of Statistics*, 57, Series B, Pt. 1, 19–31.
25. Hurd, H. L., & Miamer, A. G. (2007). *Periodically correlated random sequences: Spectral. Theory and practice*. Wiley
26. Hurd, H., & Leśkow, J. (1992). Strongly consistent and asymptotically normal estimation of the covariance for almost periodically correlated processes. *Statistics and Decisions*, 10, 201–225.
27. Jach, A., McElroy, T., & Politis, D. N. (2012). Subsampling inference for the mean of heavy-tailed long memory time series. *Journal of Time Series Analysis*, 33(1), 96–111.
28. Lahiri, S. (1993). On the moving block bootstrap under long-range dependence. *Statistics and Probability Letters*, 18, 405–413.
29. Leśkow, J., & Synowiecki, R. (2010). On bootstrapping periodic random arrays with increasing period. *Metrika*, 71, 253–279.
30. Lunetta, G. (1963). Di una generalizzazione dello schema della curva normale. *Annali della Facoltà di Economia e Commercio di Palermo*, 17, 237–244.
31. Philippe, A., Surgailis, D., & Viano M. C. (2006). *Almost periodically correlated processes with long-memory*. Lecture notes in statistics 187. Springer
32. Politis, D. N., Romano, J. P., & Wolf, M. (1999). *Subsampling*. New York: Springer.
33. Shao, J., & Tu, D. (1995). *The jackknife and bootstrap*. New York: Springer.
34. Subbotin, M. T. (1923). On the law of frequency of error. *Matematicheskii Sbornik*, 31(2), 296–301.
35. Synowiecki, R. (2007). Consistency and application of moving block bootstrap for non-stationary time series with periodic and almost periodic structure. *Bernoulli*, 13, 1151–1178.
36. Varanasi, M. K., & Aazhang, B. (1989). Parametric generalized gaussian density estimations. *Journal of the Acoustical Society of America*, 86(4), 1404–1415.

Cyclostationarity: Theory and Methods III
Contributions to the 9th Workshop on Cyclostationary
Systems and Their Applications, Grodek, Poland, 2016
Chaari, F.; Leskow, J.; Napolitano, A.; Zimroz, R.;
Wylomańska, A. (Eds.)
2017, VIII, 257 p. 115 illus., 97 illus. in color., Hardcover
ISBN: 978-3-319-51444-4