Communications for Statistical Applications and Methods
2017, Vol. 24, No. 1, 1–13

http://dx.doi.org/10.5351/CSAM.2017.24.1.001
Print ISSN 2287-7843 / Online ISSN 2383-4757

Bootstrap methods for long-memory processes: a review

Young Min Kim[a], Yongku Kim[1,a]

[a]Department of Statistics, Kyungpook National University, Korea

Abstract

This manuscript summarized advances in bootstrap methods for long-range dependent time series data. The stationary linear long-memory process is briefly described, which is a target process for bootstrap methodologies on time-domain and frequency-domain in this review. We illustrate time-domain bootstrap under long-range dependence, moving or non-overlapping block bootstraps, and the autoregressive-sieve bootstrap. In particular, block bootstrap methodologies need an adjustment factor for the distribution estimation of the sample mean in contrast to applications to weak dependent time processes. However, the autoregressive-sieve bootstrap does not need any other modification for application to long-memory. The frequency domain bootstrap for Whittle estimation is provided using parametric spectral density estimates because there is no current nonparametric spectral density estimation method using a kernel function for the linear long-range dependent time process.

Keywords: autoregressive-sieve bootstrap, block bootstrap, frequency domain bootstrap, long-memory

1. Introduction

The bootstrap method, introduced by Efron (1979), is a computer-intensive method for a large class of statistical inference issues without any strict structural assumptions on the underlying data process. The bootstrap method has developed its application into broad fields of statistical inference ever since Efron suggested it. The bootstrap is often utilized because it has better performance than the conventional approaches as well as provides an empirical and efficient statistical inference for complicated problems; consequently, existing methodology is unable to produce analytical answers for statistical inference (e.g. Shin and Hwang, 2015; Yoo, 2015). Nonetheless, the bootstrapping idea is not a panacea for all problems of statistical inferences because it cannot be equally effectively applicable to all of random processes. In this paper, we consider stationary and strongly linear time dependent processes as well as provide different types of bootstrap methods for effective applications and their limitations on time-domains and frequency-domains.

Singh (1981) provided an example to show that the independently and identically distributed (iid) bootstrap, proposed by Efron (1979) does not work under dependent data structure, which is an important consideration to construct bootstrap methods for time series that should appropriately reflect the dependent structure for underlying data. Two approaches of "data block" mechanism and data-transformation have been proposed to capture the underlying dependent structure. First, the data-blocking mechanism is that each block consists of consecutive groups of data points in time and we resample data-block, not the original time points with replacement as if the data-blocks are iid data. The data-block mechanism can help capture or preserve the dependent structure of the original time

[1] Corresponding author: Department of Statistics, Kyungpook National University, 80, Daehak-ro, Buk-gu, Daegu 41566, Korea. E-mail: kim.1252@knu.ac.kr

Published 31 January 2017 / journal homepage: http://csam.or.kr

series in each block. Another approach is the data-transformation approach to treat data dependence using data-transformation, which can help weaken the covariance structure, without perfectly twisting dependencebecause transformed points canbe regarded as independent or asymptotically independent.

Recently, the issue of the time series analysis is to reckon with the dependent-type or dependent-strength in a stationary time process $\{X_t\}$. If $r(k) = \mathrm{Cov}(X_0; X_k)$, $k \geq 0$, demonstrates the process autocovariance function with lag k, then in general we may categorize the processes as weakly or short-range dependent (SRD) if the autocovariance decays fast enough (i.e., $r(k) \to 0$ as $k \to \infty$) so that $\sum_{k=1}^{\infty} |r(k)| < \infty$ holds, or as strongly or long-range dependent (LRD) processes by a slow covariance decay, $r(k) \approx Ck^{-\theta}$ as $k \to \infty$ for some $C > 0$ and $0 < \theta < 1$ satisfying $\sum_{k=1}^{\infty} |r(k)| = \infty$ (Beran *et al.*, 2013). The LRD processes can generally be applied in astronomy, hydrology and economics (Beran *et al.*, 2013; Henry and Zaffaroni, 2003; Montanari, 2003). Statistical issues between SRD and LRD time processes can change dramatically; in addition, it is more complicate to develop appropriate resampling methods. For example, Lahiri (1993) showed that the block bootstrap, which is generally valid under weak dependence (Künsch, 1989; Liu and Singh, 1992), is not directly applicable to estimate the distribution of a sample mean of a class of long-memory processes. However, Kim and Nordman (2011) proposed block bootstrap methods with the adjustment factor for strongly dependent time processes and provided behaviors of optimal block sizes based on a large-sample variance of a sample mean of stationary linear LRD processes.

The autoregressive (AR)-sieve bootstrap method takes the idea of sieve approximation (Grenander, 1981) to generate the sieve-estimated process. The AR-sieve bootstrap for weakly dependent time processes was introduced by Kreiss (1992) and has been developed by Bühlmann (1997) and Bickel and Bühlmann (1999). In addition, Choi and Hall (2000) showed that the AR-sieve bootstrap was more powerful than block bootstrap methods under causal linear SRD processes based on the error in the coverage probability of a one-sided confidence interval. Kapetanios and Psaradakis (2006) and Poskitt (2008) applied the AR-sieve bootstrap to causal linear LRD time processes.

Dahlhaus and Janas (1996) established a frequency-domain bootstrap (FDB) method for ratio statistics under SRD using a data-transformation (i.e., Fourier transform) to weaken the data dependence structure. The FDB methodology calculates the periodogram ordinates and estimates the spectral density function feasibly by the nonparametric method. FDB methodology independently resamples scaled periodogram ordinates, (i.e., periodogram ordinates divided by corresponding spectral density estimates) to create bootstrap versions of spectral estimators. Similar approaches to the FDB methodology had been developed for nonparametric spectral density estimation (Franke and Härdle, 1992; Nordgaard, 1992). The nonparametric spectral density estimator under SRD is a critical factor for the FDB method of Dahlhaus and Janas (1996) and other frequency domain resampling methods under SRD (Kreiss and Paparoditis, 2003). Kim and Nordman (2013) proposed the FDB on Whittle estimation under LRD because appropriate nonparametric estimators of the spectral density are unavailable under LRD which is uniformly consistent on the entire spectrum $(0, \pi]$.

Many researches review bootstrap methods for iid data and weakly dependent time processes. Even though the bootstrap under LRD is still developing, this review can incur the interest of resampling methods under LRD for researchers who handle long-memory processes. This paper reviews bootstrap methods for strongly dependent time processes with block bootstrap, AR-sieve bootstrap and FDB methodologies. The remaining of the paper is as follows. Section 2 provides stationary linear LRD time processes. In Section 3, the bootstrap methods under LRD have been provided and described. We will apply the bootstrap methods to a real data in Section 4 and give concluding remarks and future research topics for the proposed methodologies in Section 5.

2. Long-memory processes

The stationary linear time series with the process mean $EX_t = \mu \in \mathbb{R}$ is defined as

$$X_t = \mu + \sum_{j\in\mathbb{Z}} b_j\varepsilon_{t-j}, \tag{2.1}$$

where $\{\varepsilon_t\}$ are iid innovations with the mean $E\varepsilon_t = 0$ and its variance $E\varepsilon_t^2 < \infty$. The real-valued $\{b_j\}$ sequence of constants satisfies $\sum_{j\in\mathbb{Z}} b_j^2 < \infty$. We classify the time series $\{X_t\}$ as exhibiting the LRD time series if the autocovariances $r(k) = \mathrm{Cov}(X_0, X_k)$ satisfy a slow decay condition

$$r(k) \sim \sigma^2 k^{-\theta}, \quad k \to \infty, \tag{2.2}$$

for some $\theta \in (0, 1]$ and $\sigma^2 > 0$. If $\theta = 1$, the process is called the short-memory. Another condition for the LRD time process is that the partial autocovariance sum $\sum_{k=1}^{n} |r(k)| = O(n^{1-\theta})$ diverges as $n \to \infty$ (Beran *et al.*, 2013; Robinson, 1995). However, SRD time series has the autocovariances $r(k)$ decaying rapidly to 0 as the lag $k \to \infty$ satisfying $\sum_{k=1}^{\infty} |r(k)| < \infty$.

To characterize the LRD property, this paper considers statistical inference about the behavior of a large-sample variance estimator of a sample mean, $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$. Denote $\sigma^2_{n,\theta} \equiv n^{\theta}\mathrm{Var}(\bar{X}_n)$. The condition of the LRD property (2.2) implies that

$$\lim_{n\to\infty} \sigma^2_{n,\theta} = \sigma^2_{\infty,\theta} > 0, \tag{2.3}$$

holds for a constant $\sigma^2_{\infty,\theta}$ depending on $\theta \in (0, 1]$. It means that the sample variance estimator, $\mathrm{Var}(\bar{X}_n)$, of a sample mean under the linear LRD, decays at a slower rate $O(n^{-\theta})$ as $n \to \infty$ than the typical $O(n^{-1})$ rate under SRD. Note that $\theta = 1$ for $\sigma^2_{n,\theta}$ gives SRD and (2.3) is appropriate for specifying sample mean behavior under SRD. For linear LRD processes (2.1), if the condition (2.3) holds, we have a normal limit for the scaled sample mean as follows (Davydov, 1970; Ibragimov and Linnik, 1971, Theorem 18.6.5):

$$n^{\frac{\theta}{2}}\left(\bar{X}_n - \mu\right) \xrightarrow{d} N\left(0, \sigma^2_{\infty,\theta}\right) \quad \text{as } n \to \infty. \tag{2.4}$$

Note that for the bootstrap methodology for a sample mean, a sum of independently resampled block averages has a large-sample normal distribution with a variance matching that of $n^{\theta/2}(\bar{X}_n - \mu)$.

Another property of long-memory processes, $\{X_t\}$, can be shown as integral spectral density defined as

$$g(\lambda) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} r(k)e^{-\iota k\lambda}, \tag{2.5}$$

where $\lambda \in \Pi \equiv (-\pi, \pi]$ and $\iota = \sqrt{-1}$ behaving

$$\lim_{\lambda\to 0} |\lambda|^{1-\theta} g(\lambda) = C, \tag{2.6}$$

for some $\theta \in (0, 1]$ and a positive constant $C > 0$. The process $\{X_t\}$ is also referred as SRD time processes when $\theta = 1$, and LRD processes when $0 < \theta < 1$. This dependence-type classification is a common, in which LRD spectral density entails a pole of $g(\cdot)$ at the origin, frequency zero (Beran

Table 1: True autocorrelation, $\rho(1)$, and finite autocorrelation, $\rho_n(1)$ with lag of 1 for four different FARIMA processes for true parameters, AR = 0.3 and MA = −0.4 with true long-memory parameters, $\theta = 0.1, 0.5, 0.9$ using standard normal innovation

Process	θ	$\rho(1)$	$\rho_n(1)$ with n =						
			100	250	500	1,000	5,000	10,000	100,000
FARIMA$(0,d,0)$	0.1	0.82	0.49	0.61	0.65	0.69	0.70	0.73	0.74
	0.5	0.33	0.26	0.30	0.31	0.32	0.33	0.33	0.33
	0.9	0.05	0.04	0.04	0.05	0.05	0.05	0.05	0.05
FARIMA$(1,d,0)$	0.1	0.92	0.73	0.80	0.82	0.84	0.86	0.87	0.87
	0.5	0.61	0.53	0.58	0.59	0.60	0.60	0.61	0.61
	0.9	0.36	0.32	0.35	0.35	0.36	0.36	0.36	0.36
FARIMA$(0,d,1)$	0.1	0.48	0.07	0.18	0.21	0.27	0.28	0.32	0.33
	0.5	−0.12	−0.16	−0.13	−0.13	−0.13	−0.13	−0.12	−0.12
	0.9	−0.31	−0.33	−0.31	−0.31	−0.31	−0.31	−0.31	−0.31
FARIMA$(1,d,1)$	0.1	0.75	0.40	0.50	0.53	0.60	0.61	0.64	0.65
	0.5	0.21	0.15	0.19	0.19	0.21	0.21	0.21	0.21
	0.9	−0.05	−0.04	−0.04	−0.05	−0.05	−0.05	−0.05	−0.05

et al., 2013; Hosking, 1981). The periodogram I_n is an estimator of true spectral density g. The periodogram ordinates are defined as

$$I_{\mathrm{n}}(\lambda) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n}(X_t - \bar{X}_n)e^{-\iota t\lambda}\right|^2, \quad \lambda \in \Pi.$$

The scaled periodogram, $I_{\mathrm{n}}(\lambda)/g(\lambda)$ has asymptotically exponential distribution with mean 1 for stationary time processes with a bound spectral density (Beran *et al.*, 2013). For linear long-memory, more complicated assumptions for the underlying processes are needed to investigate the asymptotic results of scaled periodogram behavior.

We show some property of linear LRD processes using the autocorrelation with lag 1 in the simulation study. We consider several types of FARIMA(p,d,q) processes (Adenstedt, 1974; Granger and Joyeux, 1980; Hosking, 1981) with

$$\phi(B)(1-B)^d X_t = \mu + \psi(B)\varepsilon_t, \tag{2.7}$$

where $\{\varepsilon_t\}$ are iid mean-zero variables, B is a lag operator, and the long-memory parameter is given by $\theta = 1 - 2d$ for $d \in [0, 1/2)$. Note that $\phi(B) = 1 - \phi_1 B - \phi_2 B - \cdots - \phi_p B^p$ and $\psi(B) = 1 + \psi_1 B + \cdots + \psi_q B^q$. The FARIMA processes are well-known stationary linear long-memory processes. We consider FARIMA$(0,d,0)$, FARIMA$(1,d,0)$ with an AR parameter 0.3, FARIMA$(0,d,1)$ with an MA parameter −0.4 and FARIMA$(1,d,1)$ with an AR parameter 0.3 and an MA parameter −0.4 focusing on a variety of long-memory parameters $\theta \in \{0.1, 0.5, 0.9\}$ and sample sizes $n \in \{100, 250, 500, 1000, 5000, 10000, 100000\}$. We also consider the autocorrelation estimator of lag of 1 as a function of sample means, sometimes called ratio statistics. First, we compute the true autocorrelation with lag 1 and then estimate autocorrelation estimates with lag 1 from the generated data. Finally, we examine behavior of finite autocorrelation estimates for each FARIMA process. The Monte Carlo (MC) simulation runs were considered as $M = 5,000$.

Table 1 shows that the finite autocorrelation values with lag of 1 slowly converges to the true autocorrelation under LRD where $\theta \in (0, 0.5)$, called *stronger LRD*. For example, the finite autocorrelation with lag of 1 for the FARIMA$(0,d,0)$ process with $\theta = 1 - 2d = 0.1$ is $\rho_n(1) = 0.49$ for small samples, $n = 100$, which is far from the theoretical autocorrelation with lag of 1, $\rho(1) = 0.82$. This difference may lead to lower coverage probabilities for 95% confidence intervals under stronger LRD when the theoretical autocorrelation with lag of 1 is used. That is, as the strength of dependence under LRD increases, the difference between finite variance-covariance and true variance-covariance may be larger than LRD processes for $\theta \in [0.5, 1)$, called *weaker LRD*. Therefore, as mentioned in Section 1, this is another key consideration for application and construction of bootstrap methodologies for long-memory.

3. Bootstrap methods under long-range dependence

3.1. Block bootstrap

In this subsection, we describe the block bootstrap methods to apply the process mean of a long-memory process. We denote ℓ as an integer block size, which is less than sample size n. For the moving block bootstrap (MBB) proposed and developed by Künsch (1989) and Liu and Singh (1992), we make data-blocks as $B_i = (X_i, \ldots, X_{i+\ell-1})$ with starting point $1 \leq i \leq n - \ell + 1$. Thus, the total number of blocks is $n - \ell + 1$. We independently resample $m = \lfloor n/\ell \rfloor$ data-blocks from the original series $X_1, \ldots, X_n$ with replacement from $\{B_1, \ldots, B_{n-\ell+1}\}$ and then concatenate the sampled blocks as if the blocks are regarded as the iid observations. Carlstein (1986) proposed the non-overlapping block bootstrap (NBB) using a non-overlapping data-block set $\{B_{1+\ell(i-1)} : i = 1, 2, \ldots, m\}$. For the NBB, the number of dat-blocks is $\lfloor n/\ell \rfloor$. The remaining process of the NBB is the same as MBB. In general, using the larger number of blocks for the MBB leads that the MBB has better efficiency than the NBB for bias or variance estimation under SRD (Hall *et al.*, 1995; Künsch, 1989; Lahiri, 1999). We provide the block bootstrap procedures as:

Block bootstrap procedure:

1. Define the block size $\ell < n$ and compute the number of blocks for the resampling procedure, $m \equiv \lfloor n/\ell \rfloor$
2. Construct data-blocks for the MBB with the number of total blocks, $b =\equiv n - \ell + 1$ and for the NBB with $b^{'} \equiv \lfloor n/\ell \rfloor$
3. Generate bootstrap replicates from the data-block set as:
   (a) For a MBB series $X_1^*, \ldots, X_N^*$, $N \equiv m\ell$, we generate $I_1^*, \ldots, I_m^*$ from iid uniform random variables $\{I_1, \ldots, I_{n-\ell+1}\}$.
   (b) For a NBB series $X_1^*, \ldots, X_N^*$, $N \equiv m\ell$, we generate $I_1^*, \ldots, I_m^*$ from iid uniform random variables $\{I_1, \ldots, I_{1+\ell(m-1)}\}$.
4. Make a MBB/NBB series $X_1^*, \ldots, X_N^*$, where $N \equiv m\ell$.

For long-memory, Lahiri (1993) proved that the MBB fails whenever it produces non-normal limits for $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ (i.e., by non-linear transformations of Gaussian processes). Kim and Nordman (2011) established the validity of block bootstrap distribution estimation for the process mean of linear LRD processes, without assumption of a specific form for the long-memory covariances, that is, the block bootstrap remains consistent over a practical class of linear LRD processes, which need not be a causal assumption ($b_j = 0$ for $j < 0$ in (2.1)) as assumed for the AR-sieve bootstrap (cf. Poskitt, 2008). Kim and Nordman (2011) showed that the bootstrap sample mean should be “inflated” by an adjustment factor $a_n \equiv m^{(1-\theta)/2}$ for long-memory. Then, the adjustment factor, a_n for the sample mean under LRD constructs $m^{(1-\theta)/2}N^{\theta/2}\bar{X}_N^* = (m\ell^\theta)^{1/2}\bar{X}_N^*$ as the correct version of $n^{\theta/2}\bar{X}_n$. For weak dependent time processes, there is no inflation because $\theta = 1$ leads to $m^{(1-\theta)/2} = 1$. Hence, for either MBB or NBB, under some assumptions (cf. Kim and Nordman, 2011, Theorem 1), we have the following result:

$$\sup_{x\in\mathbb{R}}\left|P_*\left(m^{\frac{1}{2}}\ell^{\frac{\theta}{2}}\left(\bar{X}_N^* - \mathrm{E}_*\bar{X}_N^*\right) \le x\right) - P\left(n^{\frac{\theta}{2}}\left(\bar{X}_n - \mu\right) \le x\right)\right| \xrightarrow{p} 0,$$

where E_* and P_* are probability and expectation of the bootstrap distribution given the data.

In general, the variance of the MBB estimator is smaller (2/3) than that of the NBB estimator under SRD (Künsch, 1989; Lahiri, 2003). However, Kim and Nordman (2011) studied that the large-sample variance of MBB and NBB estimators can match under stronger LRD, $\theta \in (0, 0.5)$ even though the MBB method has more available blocks compared to the NBB method, which implies that the MBB method completely loses its advantage of variance estimation over the NBB method.

3.2. Autoregressive-sieve bootstrap

This subsection describes an AR-sieve bootstrap method for stationary causal linear time processes, proposed by Bühlmann (1997) and Kreiss (1988,1992) under weak dependence and extended to long-memory series by Kapetanios and Psaradakis (2006) and Poskitt (2008). Suppose that we have a data realization $X_1, \ldots, X_n$ having the process mean $\mathrm{E}X_t = \mu$ and its autocovariances $r(k)$, $k \ge 0$, can first be approximated by a stationary autoregressive process $\{Y_t\}$ of the AR order p. Then, we minimize the distance

$$\mathrm{E}\left[(X_t - \mu) - \sum_{j=1}^{p}\beta_j\left(X_{t-j} - \mu\right)\right]^2$$

to obtain reasonable $(\beta_1, \ldots, \beta_p)$. Thus, the pseudo time process $\{Y_t\}$ can be defined as

$$Y_t = \mu + \sum_{j=1}^{p}\beta_j\left(Y_{t-j} - \mu\right) + e_t,$$

where $\beta \equiv (\beta_1, \ldots, \beta_p)^T = \Gamma_p^{-1}r_p$ are the coefficients of the best-linear predictor of $X_t - \mu$ in terms of $(X_{t-1}-\mu, \ldots, X_{t-p}-\mu)$, and $\{e_t\}$ are approximately iid random variables with mean $\mathrm{E}e_t = 0$ and variance $\mathrm{E}e_t^2 = r(0) - \beta'\Gamma_p\beta$. In addition, r_p is a vector of autocovariances defined as $r_p = (r(1), \ldots, r(p))^T$ and Γ_p is the $p \times p$ matrix with $r(k-j)$ with the (k, j)th entry of r_p. The AR-sieve bootstrap method manufactures a sample size n bootstrap rendition of $\{Y_t\}$ for intimating the distribution of the original time process, $\{X_1, \ldots, X_n\}$. In general, the sieve approximation with $\{Y_t\}$ to $\{X_t\}$ should improve as the AR order p increases according to the sample size n. The procedure of the AR-sieve bootstrap is provided as follows.

AR-sieve bootstrap procedure:

1. Define the autoregressive order $p \equiv p_n$ depending on the sample size.

2. Estimate the coefficients $(\hat{\beta}_{1n}, \ldots, \hat{\beta}_{pn})$ from observed data $X_1, \ldots, X_n$ utilizing the solution to the sample version of Yule-Walker equations (Brockwell and Davis, 1991).

3. Define related residuals $\hat{e}_t = (X_t - \bar{X}_n) - \sum_{j=1}^{p} \hat{\beta}_{jn}(X_{t-j} - \bar{X}_n)$, $p + 1 \leq t \leq n$.
4. Generate each e_t^* independently from a set of the centered residuals $\{e_t - (n-p)^{-1} \sum_{j=p+1}^{n} e_j : p + 1 \leq t \leq n\}$.
5. Construct the AR-sieve bootstrap replicates from

$$Y_t^* = \bar{X}_n + \sum_{j=1}^{p} \hat{\beta}_{jn} \left(Y_{t-j}^* - \bar{X}_n \right) + e_t^*, \quad t \geq p + 1,$$

where we define $Y_1^* = \cdots = Y_p^* = \bar{X}_n$.

Note that we generate $Y_1^*, \ldots, Y_{n+q}^*$ with a burn-in of length $q \geq 1$. Usually, we assign q as 300. Based on the generated last n sample, we construct the AR-sieve bootstrap mean such as $n^{\theta/2}(\bar{X}_n^* - \bar{X}_n)$ for causal linear LRD processes. It is valid to approximate the distribution of $n^{\theta/2}(\bar{X}_n - \mu)$ for linear, LRD processes under certain conditions (Kapetanios and Psaradakis, 2006). For short-memory, Bühlmann (1997) and Choi and Hall (2000) illustrated that the AR-sieve bootstrap method can give more accurate distribution estimation in comparison to the block bootstrap methods. However, the condition of the process for the AR-sieve bootstrap is more stringent than the block bootstrap methods, for example the process satisfies the conditions of causality, linearity and often invertibility. The AR order is $p_n = \lfloor 2(\log n)^2 \rfloor$ as fixed or $p_n = \lfloor 10 + 2\hat{h} \rfloor$ as estimated where $\hat{h}$ is obtained to minimize an information criterion over $1 \leq h \leq 10 \log_{10} n$ (Bühlmann,1997; Poskitt, 2008). Kapetanios and Psaradakis (2006) and Poskitt (2008) under LRD applied the fixed order to long-memory.

3.3. Frequency domain bootstrap on Whittle estimation

The frequency domain bootstrap method under LRD does not require any assumptions about the full probability structure of time series; in addition, the FDB procedures are inspired because the scaled periodogram $I_{\mathrm{n}}(\lambda)/g(\lambda)$ asymptotically has the exponential limiting distribution with the parameter of 1 at a set of $\lambda \in (-\pi, \pi]$ (Beran *et al.*, 2013; Brockwell and Davis, 1991). In addition, $I_{\mathrm{n}}(\lambda_j)$ for $0 < \lambda_1 < \cdots < \lambda_L < \pi$ are asymptotically independent at $\lambda_j = 2\pi j/n$, $j = 1, 2, \ldots, L \equiv \lfloor (n-1)/2 \rfloor$. The advantage of the FDB is not necessary to mimic the covariance structure of underlying processes, but the FDB is available for ratio statistics such as autocorrelation (Dahlhaus and Janas, 1996). As mentioned in Section 1, appropriate nonparametric spectral density estimators, in particular using a kernel function, are currently open problems under LRD. Thus, for analogous purposes of periodogram scaling in a FDB method under LRD, the FDB method cannot be applied to the LRD processes directly. Therefore, Kim and Nordman (2013) proposed the semiparametric approach for application to the FDB method under LRD using parametric spectral density estimators on the Whittle likelihood because Whittle estimation is possibly applicable to a valid FDB method under LRD, through re-scaling periodogram ordinates with a parametric spectral density estimator. For this semiparametric approach, the parametric spectral density is defined as

$$g(\lambda) \equiv g\left(\lambda; \sigma_0^2, \theta_0\right) = \frac{\sigma_0^2}{2\pi} f(\lambda; \theta_0), \quad \lambda \in \Pi \equiv (-\pi, \pi] \tag{3.1}$$

holds at some true parameters (σ_0^2, θ_0). Whittle estimation (Whittle, 1953) looks for determining the parameter values at which the theoretical distance measure as

$$W\left(\sigma^2, \theta\right) = \frac{1}{4\pi} \int_0^{\pi} \left\{ \log g\left(\lambda; \sigma^2, \theta\right) + \frac{g(\lambda)}{g(\lambda; \sigma^2, \theta)} \right\} d\lambda = \log\left[\frac{\sigma^2}{2\pi}\right] + \frac{1}{4\pi} \int_0^{\pi} \frac{g(\lambda)}{g(\lambda; \sigma^2, \theta)} d\lambda \tag{3.2}$$

achieves its minimum (Dzhaparidze, 1986). Under mild conditions, the true parameter values (σ_0^2, θ_0) are then determined as the unique solutions to the score equations $\partial W(\sigma, \theta)/\partial(\sigma, \theta) = 0_{p+1}$, or equivalently

$$\int_0^{\pi} \left(\frac{\partial}{\partial \theta} \frac{1}{f(\lambda; \theta)} \right) g(\lambda)\, d\lambda = 0_p \quad \text{and} \quad 2 \int_0^{\pi} \frac{g(\lambda)}{f(\lambda; \theta)} d\lambda = \sigma^2, \tag{3.3}$$

where $f(\lambda; \theta)$ is defined in (3.1). The Whittle estimators $\hat{\theta}_n$ of θ are formally defined as a solution of the periodogram-based estimating functions

$$\ell_n(\theta) \equiv \frac{2\pi}{n} \sum_{j=1}^{L} \left(\frac{\partial}{\partial \theta} \frac{1}{f\left(\lambda_j; \theta\right)} \right) I_{\mathrm{n}}\left(\lambda_j\right) = 0_p, \tag{3.4}$$

defined by a Riemann integral approximation to (3.3) that substitutes the periodogram $I_{\mathrm{n}}(\cdot)$ as an estimator of the true spectral density $g(\cdot)$ at discrete Fourier frequencies $\lambda_j = 2\pi j/n$ for $j = 1, 2, \ldots, L$. Moreover, the corresponding Whittle estimator of σ^2 is computed by

$$\hat{\sigma}_n^2 = \frac{4\pi}{n} \sum_{j=1}^{L} \frac{1}{f\left(\lambda_j; \hat{\theta}_n\right)} I_{\mathrm{n}}\left(\lambda_j\right). \tag{3.5}$$

Now, the statistical inference with the FDB estimator concentrates on the distribution $P(\sqrt{n}(\hat{\theta}_n - \theta_0) \leq x)$, $x \in \mathbb{R}^p$ for the Whittle estimator $\hat{\theta}_n$. An bootstrap estimator $\hat{\theta}_n^*$ for the Whittle estimator is also obtained according to the following FDB procedure. Conditional on the original time series data $X_1, \ldots, X_n$ which has properties of a linear long-memory process,

FDB procedure:

1. Obtain the Whittle estimates $(\hat{\theta}_n, \hat{\sigma}_n^2)$ from the original data.
2. Define the parametric spectral density estimates $\hat{g}_n(\lambda_j) \equiv g(\lambda_j; \hat{\sigma}_n^2, \hat{\theta}_n)$ at discrete Fourier frequencies $j = 1, 2, \ldots, L$ using Whittle estimates.
3. Studentized periodogram ordinates $\hat{e}_j = I_{\mathrm{n}}(\lambda_j)/\hat{g}_n(\lambda_j)$, $j = 1, 2, \ldots, L$.
4. Obtain rescaled $\tilde{e}_j = \hat{e}_j/\hat{e}.$ where $\hat{e}. = \sum_{j=1}^{L} \hat{e}_j/L$.
5. Randomly sample $e_1^*, \ldots, e_L^*$ from the empirical distribution of $\{\tilde{e}_j; j = 1, 2, \ldots, L\}$.
6. Define a bootstrap version of periodogram ordinates as $\mathrm{I}_{\mathrm{n}}^*(\lambda_j) = e_j^* \hat{g}_n(\lambda_j)$ for $j = 1, 2, \ldots, L$.
7. Calculate bootstrap Whittle estimators $\hat{\theta}_n^*$ from the equation

$$\sum_{j=1}^{L} \left[\left(\frac{\partial}{\partial \theta} \frac{1}{f\left(\lambda_j; \theta\right)} \right) \mathrm{I}_{\mathrm{n}}^*\left(\lambda_j\right) - \left(\frac{\partial}{\partial \theta} \frac{1}{f\left(\lambda_j; \hat{\theta}_n\right)} \right) \hat{g}_n\left(\lambda_j\right) \right] = 0_p. \tag{3.6}$$

8. Construct the bootstrap versions $\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)$ of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ so that $P_*(\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) \leq x)$ is the bootstrap estimator of $P(\sqrt{n}(\hat{\theta}_n - \theta_0) \leq x)$, $x \in \mathbb{R}^p$.

Since the periodogram scaling scheme for the FDB under LRD using the parametric spectral density estimates is different from the general one in the original FDB method for short-memory (Dahlhaus and Janas, 1996) using the nonparametric spectral density estimates using the kernel method. Currently, the similar kernel density estimation under LRD estimators does not seemingly exist. However, the Whittle estimation method is valid for either short- and long-memory allowing a different re-scaling technique for the periodogram. The reason to apply the rescaling technique of studentized periodogram ordinates in Step 4 is to eliminate unnecessary bias in the FDB re-creation (Dahlhaus and Janas, 1996), as the studentized periodogram at a fixed frequency $0 < \lambda < \pi$ has an asymptotic exponential distribution with mean parameter 1 under SRD/LRD (Yajima, 1989). Steps 6–8 reproduce the structural relationship between Whittle estimators $\hat{\theta}_n$ and the true parameters θ_0 at the level of the bootstrap.

4. Data example

We illustrate the MBB, the AR-sieve and the FDB for Whittle estimation under LRD with data about the yearly minimal water levels of the Nile River from 622 to 1284 measured at the Roda gauge in Egypt (Tousson, 1925). The total observed sample size is $n = 663$. Figure 1(a) displays Nile River data assuming this data is a realization of a stationary linear LRD process $\{X_t\}$ satisfying (2.1), (2.2), and (2.5). In this data example, we are interested in obtaining confidence interval estimates for the process mean, $\mathrm{E}X_t = \mu$ and the long-memory parameter, θ using parametric approaches and bootstrap approaches. We compute 95% confidence interval estimates for the process mean using MBB and AR-sieve bootstrap methods and for the long-memory parameter using the FDB for Whittle estimation under LRD. For the parametric approach, we fit Nile River data to an ARFIMA model based on fractionally differenced white noise from the **R-package arfima**. Figures 1(b) and (c) show the partial autocorrelation and autocorrelation functions, which may lead us to consider Nile River data as the LRD time process. The AIC model selection method indicates that an FARIMA$(0, d, 0)$ model with $d = (1 - \theta)/2$ is an appropriate model for Nile River data. The estimate of the long-memory parameter is $\hat{\theta} = 0.245$, which implies Nile River data is a stronger LRD process. The confidence interval for $\hat{\theta}$ is $(0.124, 0.346)$ and the estimated process mean is $\hat{\mu} = 1148.13$ and its confidence interval is $(1063.77, 1232.49)$. Figure 1(d) shows that the theoretical ACF based on the estimated values, $\hat{\theta}$ and $\hat{\mu}$ looks similar to the ACF in Figure 1(c). From this result, Nile River data can be regarded as the FARIMA$(0, \hat{d}, 0)$ with $\hat{d} = 0.382$ and $\hat{\mu} = 1148.13$.

First, we compute the confidence interval estimates for the process mean, μ using the MBB and AR-sieve bootstrap methods. For the MBB, we use 4 fixed block sizes $\ell = \lfloor Cn^{1/3} \rfloor$ where $C = 0.2, 0.5, 1, 2$, displayed in Table 2. We also use 4 fixed AR orders, $p_n = \lfloor C(\log_{10}(n))^2 \rfloor$ with $C = 0.5, 1, 2, 3$. The calculated orders are illustrated in Table 2. Table 2 shows that the estimated confidence from the parametric approach is narrower than other bootstrap results. As Kim and Nordman (2011) mentioned that for the stronger LRD process, the smaller block size is close to the optimal block size for the statistical inference of the process mean, Table 2 shows that the confidence interval using the smaller block size is closer than those with larger block sizes. The AR-sieve bootstrap is relatively insensitive to various AR-orders compared to the MBB for a diversity of block sizes, which reflects the finding of Kim and Nordman (2011).

We estimated the confidence interval for the long-memory parameter, θ using the FDB for Whittle estimation under LRD. The confidence interval estimate for the long-memory parameter is (0.147, 0.32 3), which is slightly narrower than the confidence interval estimates from the parameter approach, which is one of results from Kim and Nordman (2013).

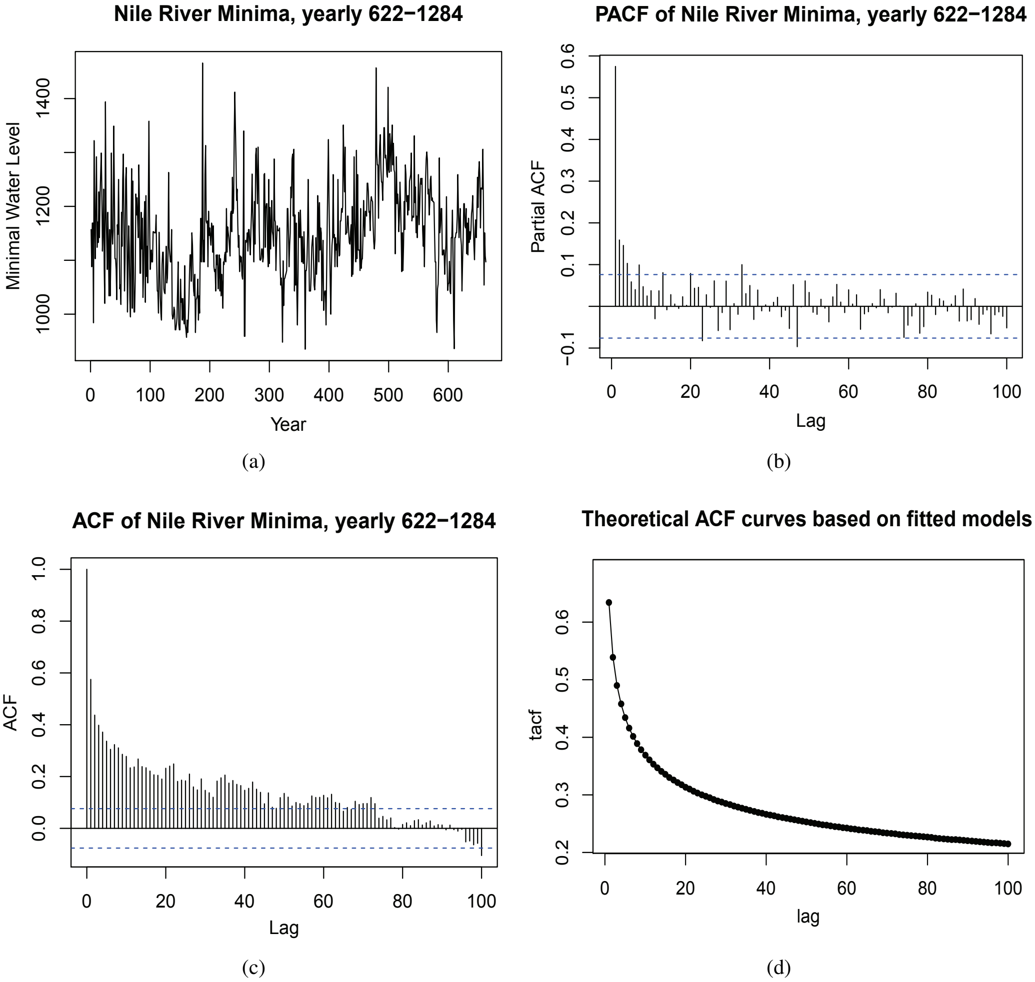

Figure 1: *Plots of the yearly minimal water levels of the Nile River from 622 to 1284 (a), its PACF (b), ACF (c) and theoretical ACF based (d) on the parameter estimates.*

Table 2: Confidence intervals for the process mean of the yearly minimal water levels of the Nile River from 622 to 1284 using the MBB and AR-sieve bootstrap methods considering various block sizes and AR orders, respectively

	Moving block bootstrap				AR-sieve bootstrap			
	$\ell = 2$	$\ell = 4$	$\ell = 8$	$\ell = 27$	$p_n = 3$	$p_n = 7$	$p_n = 15$	$p_n = 23$
Lower limit	1045.48	1020.58	989.29	936.28	1130.67	1124.33	1118.88	1115.97
Upper limit	1250.78	1275.69	1306.97	1359.98	1165.59	1171.93	1177.38	1180.29

5. Discussion and conclusion

A main challenge in extension of the block, AR-sieve and FDB bootstraps under weak dependence to LRD time processes is that the autocovariance is not summable and the spectral density has a pole at zero under LRD, which indicates the need to research if bootstrap methods constructed under weak dependence are workable for long-memory. Therefore, we cannot directly apply the bootstrap techniques under weak dependence to LRD time processes since the properties of the two weakly and strongly dependent processes have different dependent structures. For example, Subsection 3.1 describes that the adjustment factor for the block bootstrap under LRD is necessary to capture dependent structure of the underlying process in contrast to block bootstraps under weak dependence. In addition, Kim *et al.* (2013) showed that the block size for the blockwise empirical likelihood depends on the underling processes that implies that optimal block sizes depend on the magnitude of dependence of the original processes as well as the sample size. For example, data from the underlying process AR(1) with AR = 0.9 or AR(1) with AR = −0.2 need the different optimal block sizes with the same sample size. Hence, the optimal block sizes both depend on the sample size and the strength of dependence of underlying processes. Thus, for the block bootstrap under LRD, the optimal block selection method is an important and open question. The nonparametric spectral density estimation under LRD should be utilized in order to apply the FDB proposed by Dahlhaus and Janas (1996) to long-memory processes. However, the nonparametric method does not exist under LRD and remains a very important and open question for frequency-based bootstrap methods under LRD. If the method is developed, we can use the nonparametric spectral density estimator for the original FDB proposed by Dahlhaus and Janas (1996) and the AR-aided periodogram bootstrap (Kreiss and Paparoditis, 2003) under LRD.

Acknowledgement

This research was supported by Kyungpook National University Fund, 2016.

References

Adenstedt RK (1974). On large-sample estimation for the mean of a stationary sequence, *The Annals of Statistics*, **2**, 1095–1107.

Beran J, Feng Y, Ghosh S, and Kulik R (2013). *Long-Memory Processes : Probabilistic Properties and Statistical Methods*, Springer, Berlin.

Bickel PJ and Bühlmann P (1999). A new mixing notion and functional central limits theorems for a sieve bootstrap in time series, *Bernoulli*, **5**, 413–446.

Brockwell PJ and Davis RA (1991). *Time Series: Theory and Methods* (2nd ed), Springer, New York.

Bühlmann P (1997). Sieve bootstrap for time series, *Bernoulli*, **3**, 123–148.

Carlstein E (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence, *The Annals of Statistics*, **14**, 1171–1179.

Choi E and Hall P (2000). Bootstrap confidence regions constructed from autoregressions of arbitrary order, *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, **62**,461–477.

Dahlhaus R and Janas D (1996). A frequency domain bootstrap for ratio statistics in time series analysis, *The Annals of Statistics*, **24**, 1934–1963.

Davydov YA (1970). The invariance principle for stationary processes, *Theory of Probability and Its Application*, **15**, 487–498.

Dzhaparidze K (1986), *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer, New York.

Efron B (1979). Bootstrap methods: another look at the jackknife, *The Annals of Statistics*, **7**, 1–26.

Franke J and Härdle W (1992). On bootstrapping kernel spectral estimates, *The Annals of Statistics*, **20**, 121–145.

Granger CWJ and Joyeux R (1980). An introduction to long-memory time series models and frac-

tional differencing, *Journal of Time Series Analysis*, **1**, 15–29.
Grenander U (1981). *Abstract Inference*, Wiley, New York.
Hall P, Horowitz JL, and Jing BY (1995). On blocking rules for the bootstrap with dependent data, *Biometrika*, **82**, 561–574.
Henry M and Zaffaroni P (2003). The long range dependence paradigm for macroeconomics and finance. In P. Doukhan, G. Oppenheim and MS Taqqu (Eds.), *Theory and Applications of Long-Range Dependence* (pp. 417–438), Birkhäuser, Boston.
Hosking JRM (1981). Fractional differencing, *Biometrika*, **68**, 165–176.
Ibragimov IA and Linnik YV (1971). *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen.
Kapetanios G and Psaradakis Z (2006). *Sieve Bootstrap for Strongly Dependent Stationary Processes*, Queen Mary University of London, Department of Economics, London.
Kim YM and Nordman DJ (2011). Properties of a block bootstrap under long-range dependence, *Sankhya A*, **73**, 79–109.
Kim YM and Nordman DJ (2013). A frequency domain bootstrap for Whittle estimation under long-range dependence, *Journal of Multivariate Analysis*, **115**, 405–420.
Kim YM, Lahiri SN, and Nordman DJ (2013). A progressive block empirical likelihood method for time series, *Journal of the American Statistical Association*, **108**, 1506–1516.
Kreiss JP (1992). Bootstrap procedures for AR(∞)-processes. In KH Jockel, G Rothe, and W Sendler (Eds), *Bootstrapping and Related Techniques*(pp. 107-113), Springer, Heidelberg.
Kreiss JP (1988). *Asymptotic Statistical Inference for a Class of Stochastic Processes*, Habilitationsschrift, Universitat Hamburg.
Kreiss JP and Paparoditis E (2003). Autoregressive aided periodogram bootstrap for time series, *The Annals of Statistics*, **31**, 1923–1955.
Künsch HR (1989). The jackknife and the bootstrap for general stationary observations, *The Annals of Statistics*, **17**, 1217–1241.
Lahiri SN (1993). On the moving block bootstrap under long range dependence, *Statistics & Probability Letters*, **18**,405–413.
Lahiri SN (1999). Theoretical comparisons of block bootstrap methods, *The Annals of Statistics*, **27**, 386–404.
Lahiri SN (2003). *Resampling Methods for Dependent Data,* Springer, New York.
Liu RY and Singh K (1992). Moving blocks jackknife and bootstrap capture weak dependence. In R LePage and L Billard (Eds), *Exploring the Limits of Bootstrap* (pp. 225-248), John Wiley & Sons, New York.
Montanari A (2003). Long-range dependence in hydrology. In P Doukhan, G Oppenheim and M Taqqu (Eds), *Theory and Applications of Long-Range Dependence* (pp. 461–472), Birkhäuser, Boston.
Nordgaard A (1992). Resampling stochastic-processes using a bootstrap approach. In KH Jockel, G Rothe, and W Sendler (Eds), *Bootstrapping and Related Techniques* (pp. 181-185), Springer, Heidelberg.
Poskitt DS (2008). Properties of the sieve bootstrap for fractionally integrated and non-invertible processes, *Journal of Time Series Analysis*, **29**, 224–250.
Robinson PM (1995). Gaussian semiparametric estimation of long range dependence, *The Annals of Statistics*, **23**, 1630–1661.
Shin DW and Hwang E (2015). Stationary bootstrapping for the nonparametric AR-ARCH model, *Communications for Statistical Application and Methods*, **22**, 463–473.

Singh K (1981). On the asymptotic accuracy of the Efron's bootstrap, *The Annals of Statistics*, **9**, 1187–1195.

Tousson O (1925). *Mémoire sur l'Histoire du Nil*, Institut français d'archéologie orientale, Cairo.

Whittle P (1953). Estimation and information in stationary time series, *Arkiv för Matematik*, **2**, 423–434.

Yajima Y (1989). A central limit theorem of Fourier transforms of strongly dependent stationary processes, *Journal of Time Series Analysis*, **10**, 375–383.

Yoo JK (2015). Generalized partially double-index model: bootstrapping and distinguishing values, *Communications for Statistical Application and Methods*, **22**, 305–312.

Received December 30, 2016; Revised January 6, 2017; Accepted January 6, 2017