REGRESSION MODELS WITH LOCALLY STATIONARY LONG-MEMORY ERRORS

By

Guillermo Patricio Ferreira Cabezas

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
SANTIAGO , CHILE
AUGUST 2010

© Copyright by Guillermo Patricio Ferreira Cabezas, 2010

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
DEPARTMENT OF
STATISTICS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "**Regression Models with Locally Stationary Long-Memory errors** " by **Guillermo Patricio Ferreira Cabezas** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: August 2010

Research Supervisor: ______________________
Wilfredo Palma

Examing Committee: ______________________
Reinaldo Arellano

____
Natalia Bahamonde

____
Manuel Galea

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Date: **August 2010**

Author: **Guillermo Patricio Ferreira Cabezas**

Title: **Regression Models with Locally Stationary Long-Memory errors**

Department: **Statistics**

Degree: **Ph.D.** Convocation: **August** Year: **2010**

Permission is herewith granted to Pontificia Universidad Católica de Chile to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

THE AUTHOR RESERVES OTHER PUBLICATION RIGHTS, AND NEITHER THE THESIS NOR EXTENSIVE EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT THE AUTHOR'S WRITTEN PERMISSION.

THE AUTHOR ATTESTS THAT PERMISSION HAS BEEN OBTAINED FOR THE USE OF ANY COPYRIGHTED MATERIAL APPEARING IN THIS THESIS (OTHER THAN BRIEF EXCERPTS REQUIRING ONLY PROPER ACKNOWLEDGEMENT IN SCHOLARLY WRITING) AND THAT ALL SUCH USE IS CLEARLY ACKNOWLEDGED.

To my parents:

Guillermo Ferreira Valenzuela
Maria Rosa Cabezas Fernandez

and to my brothers:

Luis Alberto Ferreira Cabezas
Jose Ricardo Ferreira Cabezas

Table of Contents

Abstract

This thesis addresses the statistical analysis of regression models with locally stationary disturbances. This methodology allows for the fitting of non-stationary time series data displaying both trends and time-varying long-range dependent errors. Data with such features arise in many fields, including for example economy, climatology and hydrology, among others.

In order to deal with the non-stationary behavior of the regression errors, a locally stationary approach is proposed. This statistical framework allows for the modeling of a time-varying autocovariance structure. In this context, the parameters of the non-stationary model are allowed to vary smoothly over time so that it can be locally approximated by stationary processes.

The study conducted in this thesis focuses on the analysis of some statistical properties of the least squares estimates (LSE) of the regression models described above. These estimators are widely used in practice because they can be readily calculated. Observe that other techniques such as, for example the best linear unbiased estimators (BLUE), make the unrealistic assumption that the dependence structure of the errors in known a priori. This critical assumption is even harder to justify in practice since the dependence structure of the errors is not necessarily stationary.

The behavior of the LSE is studied in this work from three complementary points of view. First, the large sample behavior of the LSE is analyzed. In particular, conditions for the consistency of these estimators are provided. Besides, precise convergence rates for the asymptotic variance of the LSE of this regression model are established. It is shown that these estimators satisfy a central limit theorem. In addition, the asymptotic normality of the estimates of the error model parameters is established. Second, the finite sample performance of the LSE is studied by means of several Monte Carlo simulations. Finally, the application of the proposed regression methodologies is illustrated with real-life data examples.

Acknowledgements

I would like to thank Wilfredo Palma, my supervisor, for his many suggestions and constant support during this research. Professor Palma has been a cornerstone in the development of my thesis and during my time as a researcher, with his great ability to teach, taught me the mathematical foundations of time series. That is why I will thank him very much for being a great teacher and a good friend.

I should also mention that my graduate studies at Pontificia Universidad Católica de Chile were supported by the VRAID (*Vicerectoría Adjunta de Investigación y Doctorado*) and CONICYT (*Comisión Nacional de Investigación Cientifica y Tecnológica*) with their program of Support Grant for doctoral thesis 2010 academic year, for making possible the assistance to some important meetings and afford useful bibliography for my research.

Of course, I am very grateful to my girlfriend Claudia for her patience and *love*. Besides for her continued unconditional support and for all the good and bad times we had together.

I wish to thank the following friends: Ricardo Olea, Luis Guitierrez, Rodrigo Cofre, Mauricio Castro, Jorge Figueroa, with whom we share some discussions on statistics. I also would like to thank the professors who gave me a solid theoretical basis of statistics: Maria Valenzuela, Bernardo Lagos, Ernesto San Martin, Fernando Quintana. Thank my friends, who made my life very fun: Joel Muñoz and his wife Carolina, Jose Alejandro Torres, Alejandro Sagardia, Raul Ortega, Andre Toledo and his wife Andrea; *and* Ernestina Maripangui, and Soledad Alcaino who helped me a lot in the difficult times in the period of Ph.D.

Finally, I would like to thank my family: Maria, Guillermo, Luis, and José, who have always supported and encouraged me in all my decisions. Thank you for your love and for never losing faith in me.

Guillermo Ferreira

Introduction

Let the observed process $\{Y_{t,T}\}$ follow the regression model

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \varepsilon_{t,T},$$

where $X(\frac{t}{T}) = (x_{t1}, \ldots, x_{tp})'$ is a p-vector of non-stochastic regressors $\beta = (\beta_1, \ldots, \beta_p)'$ is a vector of unknown regression parameters, and $\varepsilon_{t,T}$ the sequence of a class of locally stationary long-memory (LSLM) processes. We discuss the asymptotic properties of the LSE for the unknown parameter, more specifically the consistency, asymptotic variance and normality of $\widehat{\beta}$ under the family of LSLM processes. However this model does not have a stationary property, which is crucial in the standard estimation theory and asymptotic theory of time series models. Spectral analysis of time series is a large field, presenting a great interest from both theoretical and practical viewpoints. The fundamental starting point of this analysis is the Cramer representation, therefore a stationary time series can also be viewed as a sum of an infinite number of randomly weighed complex exponentials, Fourier basis functions, through the use of the Cramer representation see Brillinger (1981)

$$X_t = \int_{-\pi}^{\pi} A(\lambda)e^{i\lambda t}\mathrm{d}\xi(\lambda), \quad t \in \mathbb{Z},$$

where $A(\lambda)$ is the transfer function and $\xi(\lambda)$ is zero-mean random process with variance one and orthogonal increments i.e. $E(\mathrm{d}\xi(\lambda), \overline{\mathrm{d}\xi(\lambda)}) = \delta_0(\lambda - \mu)$, see Priestley (1965).

The Cramer representation of a stationary time series is in terms of the Fourier functions which are perfectly localized in the frequency domain but not localized in time. Moreover, the transfer function is independent of time. The above statement about the distribution of power over frequencies not changing in time is not true for a non-stationary time series data.

Locally stationary processes are becoming an important tools to analyse non-stationary time series data. Many authors have suggested definitions for this type of processes, including Silverman (1957), Priestley (1965) and Dahlhaus (1996), among others. Furthermore, the theory of locally stationary processes has been recently extended to encompass non-stationary long-range dependent time series data, see for example Beran (2008), Genton and Perrin (2004) (2004) and Jensen and Witcher (2000). Long-memory time series has attracted a great deal attention in the last decades, see for example the monographs Beran (1994) and Palma (2007). In particular, characterization of long-memory has been studied by Parzen (1992) and Hall (1997). The estimation parameter of LSLM processes has been studied by Beran (2008) and Jensen and Witcher (2000), among others. However, it seems that the estimation of the regression parameters of such processes has received far less attention.

In this work , we establish conditions for the asymptotic variance of vector parameter estimates establishes precise convergence rates for a family of LSLM processes with general time-varying long memory parameter. Apart from establishing these asymptotical results, this work explores the finite sample calculation of the variance of LSE of a LSLM process.

This thesis is organized as follows. Chapter 1 is devoted to provide definitions of long-memory and locally stationary processes discussed in this work. Chapter 2 studies the statistical properties of the sample mean as an estimate of a locally stationary process with constant mean. Chapter 3 extends these results to the case where the locally stationary process has a time-varying scalar trend. Further extensions to the multivariate case are considered in Chapter 4. An application of these techniques to real-life data is discussed in Chapter 5. Conclusions and some guidelines for future work are addressed in Chapter 6. This work ends with a technical appendix containing several auxiliary lemmas.

Chapter 1

Locally stationary long-memory processes

Most time series analyses are based on the assumption that the probabilistic properties of the underlying process are time-invariant. Even if this assumption is very useful in order to construct simple predictors and asymptotic properties over the parameters are satisfied, it seems not to be the best strategy in practice, actually, many time series are not covariance stationary and exhibit a time-varying or evolutionary second order structure [cf. Preistley, 1965]. In the following we will give notions of how this non-stationary behavior can be modeled and derive some basic properties of these processes.

1.1 Introduction

The stationarity property of a time series is important in the theory of estimation and asymptotic of time series models. A discrete time series $\{X_t\ ,\ t = 0, \pm 1, \pm 2, \ldots\}$ is said to be *strictly stationary* if for any $t_1, t_2, \ldots, t_n$ and for any k, the joint probability distribution of $\{X_{t1}, X_{t2}, \ldots, X_{tn}\}$ is identical with the joint probability distribution of

$\{X_{t1+k}, X_{t2+k}, \ldots, X_{tn+k}\}$ Brockwell and Davis (1996). This is a difficult requirement to satisfy or to verify for any time series. It can be relaxed by requiring stationarity only for moments up to some order. A zero-mean discrete time random process X_t is said to be *weakly stationary or, simply, stationary* if the auto-covariance function of lag k,

$$\gamma_X(k) = E(X_t X_{t+k}),$$

between X_t and X_{t+k} depends only on k, but not on t. The *spectral density function* $f_X(\cdot)$ of a stationary process is defined as the discrete Fourier transform of the auto-covariance function,

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-ih\lambda), \quad -\pi < \lambda < \pi. \tag{1.1.1}$$

The summability of $|\gamma(\cdot)|$ implies that the series in (1.1.1) converges absolutely. A common nonparametric estimator of the spectral density function of X_t of length n is the periodogram,

$$I_X(\lambda) = \frac{1}{2\pi n} |J_X(\lambda)|^2 = \frac{1}{2\pi n} J_X(\lambda) J_X(-\lambda),$$

where $J_X(\cdot)$ is the discrete Fourier transform of X_t and it is defined as

$$J_X(\lambda) = \sum_{t=0}^{n} X_t \exp(-it\lambda).$$

In general, $I_X(\lambda)$ is asymptotically unbiased but inconsistent estimator of $f_X(\lambda)$ [cf. Brockwell and Davies, 1996]. Consistency of this estimate can be improved by applying a proper tapering to the data Dahlhaus and Giraitis (1998). Besides X_t can be represented as

$$X_t = \int_{-\pi}^{\pi} A(\lambda) \exp(it\lambda) \mathrm{d}\xi(\lambda),$$

discussed in the Introduction (previous section). The spectral density function of the stationary time series X_t can then be expressed as

$$f_X(\lambda) = |A(\lambda)|^2 = A(\lambda)A(-\lambda). \tag{1.1.2}$$

The above statement about the distribution of power over frequencies not changing in time is not true for a non-stationary time series . If X_t is non-stationary then the auto-covariance function of X_t and X_{t+k} is a function of t and k. In practice, many time series, especially hydrology, climatology, and financial time series, exhibit non-stationary behavior. In such cases, various techniques, such as specialized transformations (*differencing*) of data or considering the data over small piecewise stationary time intervals, can be employed to make the analysis of stationary techniques applicable for non-stationary time series. As referenced in his paper Dahlhaus (1997) *" If one abandons the assumption of stationary, the number of possible models for time series data explodes. For example, one may consider ARMA models with time varying coefficients. In this case the time behavior of the coefficients may again be modeled in different ways"*. To improve the understanding concept Dahlhaus proposes the following model

$$X_t = g(t)X_{t-1} + Z_t \quad \text{with} \quad Z_t \overset{iid}{\sim} N(0, \sigma^2),$$

for $t = 1, \ldots, T$.

Inference in this case means inference for the unknown function g on the grid $\{1, \ldots, T\}$. It is obvious that an asymptotic approach where $T \to \infty$ in not suitable for describing a statistical method since future "observations " of $g(t)$ do not necessarily contain any information on $g(t)$ on $\{1, \ldots, T\}$. This mean that as X_t is non-stationary and may seem contradictory to construct a asymptotic or forecasting theory, since a best linear predictor exploits generally when exist a varying time structure in the unconditional moments of the process.

This problem is overcome if we add regularity assumptions on the deterministic function $g(t)$. For instance, we can impose that $g(t)$ is a piecewise constant function. More generally, we can assume that $g(t)$ is nearly constant along intervals of a certain length τ. However, this approach is not satisfactory since it implicitly imposes that the function $g(t)$ is estimable only using τ observations. In this framework, when the length of the data set increases, no improvement is possible in the estimation of $g(t)$ over this interval of length τ. This implies that asymptotic considerations can not be used in the statistical inference of such process. This is a substantial drawback, because the usual statistical properties of estimators such as consistency, efficiency or central limit theorems cannot be used to measure and to compare the quality of different estimators.

To overcome this problem, Dahlhaus introduced a concept of *"local stationarity"*, he suppose observe the series from time 0 up to $T-1$ (T observations). The local stationary assumption postulates the existence of a deterministic function $g(u)$ defined for $u \in [0,1)$ such that the approximation $g(t) \approx g(u)$ holds in an appropriate way, we will define below. In this approach, two scales of time are defined: The *observed time* , which is the usual scale of time $0,\ldots,T-1$, and the *rescaled time* defined on the interval $[0,1)$. The resulting non-stationary process is doubly indexed

$$X_{t,T} = g\left(\tfrac{t}{T}\right) X_{t-1,T} + Z_t \quad \text{with} \quad Z_t \overset{iid}{\sim} N(0,\sigma^2).$$

The regularity assumptions are now made on the function $g(u)$ defined on $[0,1)$. Due to the mapping between $0,\ldots,T-1$ and $[0,1)$, the estimation of $g(u)$ becomes a standard statistical problem: For instance if $g(u)$ is constant on an interval of length $\tau < 1$ in the rescaled time, then it may be estimated using $\tau \cdot T$ observed data in the real time.

An important consequence of the rescaled time is the interpretation of *asymptotics*. When T tends to infinity, we get more information on the local structure of $X_{t,T}$ process in the rescaled time, because the mapping defines a finer grid in the rescaled time. However, it does not mean that we look into the future, because the rescaled time has a fixed bounded support $[0, 1)$.

1.2 Long-memory processes

In literature, autocovariances and autocorrelations are often referred to as memory indicators. A simple way to classify the memory type of a stationary time series is by quantifying the rate of decay of autocovariances or autocorrelations. In mathematical terms long-memory process autocorrelations have a power type decay to zero as the lag increases. The autocorrelations decay to zero in a short memory process, such as $ARMA$ processes [cf. Beran (1994), Brockwell and Davis (1996)], occurs at a much more rapid, exponential, rate.

An ARMA process X_t is a short memory process since the autocovariance between X_t and X_{t+k} decreases exponentially as $k \to \infty$. In fact the autocorrelation function, $\rho(k)$, is exponentially bounded, i.e. $|\rho(k)| \leq Cr^k$, for $k = 1, 2, \ldots$, where $C > 0$ and $0 < r < 1$.

1.2.1 Definition

There exist many definitions for the long-memory process see Palma (2007) for more details, a particular definitions of this process is given below.

Definition 1.2.1. A long-memory process X_t can be defined by specifying a hyperbolic decay of the auto-covariances

$$E(X_t X_{t+k}) = \gamma_X(k) \sim k^{2d-1} l_1(k),$$

as $k \to \infty$, where d is the so-called *long-memory parameter* and $l_1(\cdot)$ is a slowly varying function.

For any real number $d > -1$, be let define the difference operator $\nabla^d = (1-B)^d$ where B is the backward shift operator, using the binomial expansion,

$$\nabla^d = (1-B)^d = \sum_{j=0}^{\infty} \pi_j B^j,$$

where

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{0<k\leq j} \frac{k-1-d}{k} \quad j = 0, 1, 2, \ldots, \tag{1.2.1}$$

and $\Gamma(\cdot)$ is the gamma function,

$$\Gamma(x) = \begin{cases} \int_{-\infty}^{\infty} t^{x-1} e^{-t}\, dt, & x > 0 \\ \infty, & x = 0 \\ x^{-1}\Gamma(1+x), & x < 0 \end{cases}$$

Similarly, we can also define the operator ∇^{-d}, the counterpart of ∇^d as

$$\nabla^{-d} = (1-B)^{-d} = \sum_{j=0}^{\infty} \psi_j B^j, \tag{1.2.2}$$

where

$$\psi_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} = \prod_{0<k\leq j} \frac{k-1+d}{k}, \quad j = 0, 1, 2, \ldots$$

By applying the Stirling's formula, $\Gamma(x) \sim \sqrt{(2\pi)}e^{-x+1}(x-1)^{x-1/2}$ as $x \to \infty$, we obtain

$$\begin{aligned} \pi_j &\sim j^{-d-1}/\Gamma(-d) \\ \psi_j &\sim j^{d-1}/\Gamma(d), \end{aligned} \tag{1.2.3}$$

as $j \to \infty$, where $a_T \sim b_T$ means that $a_T/b_T \to 1$, as $T \to \infty$.

1.2.2 ARFIMA Processes

A well know class of long-memory models is the autoregressive fractionally integrated moving-average (ARFIMA) processes introduced by Granger and Joyeux (1980) and Hosking (1981). An ARFIMA process X_t may be defined by

$$\Phi(B)X_t = \Theta(B)(1-B)^{-d}Z_t, \tag{1.2.4}$$

where $\Phi(B) = (1+\phi_1 B+\phi_2 B^2+\ldots+\phi_p B^p)$ and $\Theta(B) = (1+\theta_1 B+\theta_2 B^2+\ldots+\theta_q B^q)$ are the autoregressive and moving-average operators, respectively, $(1-B)^{-d}$ is a fractional differencing operator defined in (1.2.2).

The next Theorem examines the existence of a stationary solution of the ARFIMA process defined by equation (1.2.4), including its uniqueness, causality, and invertibility.

Theorem 1.2.1. *Considerer the ARFIMA process defined by* (1.2.4). *Assume that the polynomials* $\Phi(\cdot)$ *and* $\Theta(\cdot)$ *have no common zeros and that* $d \in (-1, 1/2)$. *Then,* *(a) If the zeros of* $\Phi(\cdot)$ *lie outside the unit circle* $\{z : |z| = 1\}$, *then there is a unique stationary solution of* (1.2.4) *given by*

$$X_t = \sum_{j=-\infty}^{\infty} \phi_j Z_{t-j},$$

where $\phi(z) = (1-z)^{-d}\Theta(z)/\Phi(z)$.

(b) If the zeros of $\Phi(\cdot)$ *lie outside the closed unit disk* $\{z : |z| \leq 1\}$, *then the solution* X_t *is causal.*

(c) If the zeros of $\Theta(\cdot)$ *lie outside the closed unit disk* $\{z : |z| \leq 1\}$, *then the solution* X_t *is invertible.*

(d) If the solution X_t *is causal and invertible, then its autocorrelation function* $\rho(\cdot)$ *and spectral density* $f(\cdot)$ *satisfy, for* $d \neq 0$,

$$\rho(k) \sim Ck^{2d-1}, \quad as \quad k \to \infty,$$

where $C \neq 0$, *and*

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(e^{-i\lambda})\right|^2}{\left|\Phi(e^{-i\lambda})\right|^2} \left|1 - e^{-i\lambda}\right|^{-2d} \sim \frac{\sigma^2}{2\pi} \left[\frac{\Theta(1)}{\Phi(1)}\right]^2 \lambda^{-2d},$$

as $\lambda \to 0$.

For the Proof see Palma (2007). A class particular of ARFIMA process are the call fractionally integrated noise, a definition of this process is give at the next subsection.

1.2.3 Fractionally integrated noise

Definition 1.2.2. (*The ARIMA(0,d,0) Process*) The process X_t is said to be an ARIMA(0,d,0) process with $d \in (-.05, 0.5)$ if X_t is a stationary solution with zero mean of the difference equations,

$$\nabla^d X_t = Z_t, \quad \text{where} \quad Z_t \sim WN(0, \sigma^2). \tag{1.2.5}$$

The process X_t is often called *fractionally integrated noise* . Implicit in Definition (1.2.2) is the requirement that the series $\nabla^d X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$ with π_j as in (1.2.1), should be mean square convergent. In the case if Z_t is Gaussian then we call X_t fractionally integrated Gaussian noise. In Cramer representation,

$$\nabla^d X_t = \int_{-\pi}^{\pi} e^{it\lambda}(1 - e^{-i\lambda})^d d\, Z_X(\lambda).$$

In view of the representation (1.2.5) of Z_t we say that X_t is invertible, even though the coefficients π_j may not be absolutely summable as in the corresponding representation of Z_t for an invertible $ARMA$ process. We shall say that is causal if X_t can be expressed as

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

The existence of a stationary causal solution are established for [cf. Brockwell and Davis, 1996, Theorem 13.2.1]. From (1.2.3) it follows that $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ so that

$$\sum_{j=0}^{n} \psi_j e^{-ij\lambda} \longrightarrow (1 - e^{-i\lambda})^{-d},$$

as $n \to \infty$. Since X_t is obtained from ψ_j by the application of linear filter [cf. Brillinger(1981), and Brockwell and Davis(1996)], the transfer function and the spectral density function of fractionally integrated noise can be derived as

$$\begin{aligned} A(\lambda) &= (1 - e^{-i\lambda})^{-d} \\ f(\lambda) &= |A(\lambda)|^2 f_Z(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - e^{-i\lambda}\right|^{-2d}, \quad -\pi \le \lambda \le \pi. \end{aligned}$$

The representations spectral of this process is given by

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t}(1 - e^{-i\lambda})^{-d} d\,\xi(\lambda).$$

The autocovariance function of fractionally integrated noise process is given by

$$\gamma(h) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)} \frac{\Gamma(h+d)}{\Gamma(1+h-d)},$$

and the autocorrelation function is

$$\rho(h) = \frac{\Gamma(1-d)}{\Gamma(d)} \frac{\Gamma(h+d)}{\Gamma(1+h-d)}.$$

1.3 The model of locally stationary

1.3.1 Definition

Definition 1.3.1. A sequence of stochastic processes $X_{t,T}(t = 1, 2, \ldots, T)$ is called locally stationary with transfer function A^0 if there exists a representation

$$X_{t,T} = \int_{-\pi}^{\pi} A^0_{t,T}(\lambda) \exp{(i\lambda t)} d\,B(\lambda), \tag{1.3.1}$$

where $B(\lambda)$ is a Brownian motion on $[-\pi, \pi]$ and there exists a constant K and a 2π periodic function $A : (0, 1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$ such that

$$\sup_{t,\lambda} \left| A^0_{t,T}(\lambda) - A\left(\tfrac{t}{T}, \lambda\right) \right| \leq \frac{K}{T}, \tag{1.3.2}$$

for all T, $A(u, \lambda)$ is assumed to be continuous in u, where t and $u = \frac{t}{T}$ denote time points in the interval $[1, T]$ and the rescaled interval $[0, 1]$ respectively. The smoothness of A in u guarantees that the process has locally a stationary behavior. The idea behind this representation is, essentially, that, for each fixed T, one implicitly assumes some local interval of stationarity about each time point and a smooth change from one interval to the next. We also require additional smoothness conditions on A, namely differentiability, to develop asymptotic theory. Dahlhaus (1997) defined the time-varying (evolutionary) spectral density function of a locally stationary process at time $u \in [0,1]$ and frequency $\lambda \in [-\pi, \pi]$ by the formula

$$f(u,\lambda) = |A(u,\lambda)|^2 .$$

A good example of locally stationary processes is time-varying long-memory process, is the case of fractionally integrate noise [cf. Section 1.2.3].

$$\begin{aligned} X_{t,T} &= \sigma\left(\tfrac{t}{T}\right) \nabla^{-d\left(t/T\right)} Z_t \\ &= \sigma\left(\tfrac{t}{T}\right) \sum_{j=0}^{T} \psi_{t,T,j} Z_{t-j}, \end{aligned}$$

where $\nabla^{-d(\cdot)}$ is the fractional integration operator, with

$$\psi_{t,T,j} = \frac{\Gamma\left(j + d\left(t/T\right)\right)}{\Gamma\left(j+1\right)\Gamma\left(d(t/T)\right)}.$$

By applying the Stirling's formula, we obtain

$$\psi_{t,T,j} \sim \frac{j^{d(t/T)} - 1}{\Gamma\left(d(t/T)\right)}, \quad \text{as} \quad j \to \infty.$$

Therefore, the sequence $\psi_{t,T,j}$ is square summable, $\sum_{j=0}^{\infty} \psi_{t,T,j}^2 < \infty$, for $d \in (-0.5, 0.5)$. Hence, the time-varying transfer function is defined as the discrete Fourier transform of $\psi_{t,T,j}$ as $T \to \infty$, just as in the case of fractionally integrated noise ,

$$\sum_{j=0}^{T} \psi_{t,T,j} e^{-ij} \longrightarrow (1 - e^{-i\lambda})^{-d(\,u)}, \quad \text{for} \quad \text{all} \quad u \in [0,1],$$

as $T \to \infty$. Since $X_{t,T}$ is obtained from $\psi_{t,T,j}$ by the application of linear filter [cf. Brillinger(1981), and Brockwell and Davis(1996)], the transfer function and the spectral density function of fractionally integrated noise can be derived as

$$A(u,\lambda) = \frac{\sigma(u)}{\sqrt{2\pi}}\left(1-e^{-i\lambda}\right)^{-d(u)}$$
$$f(u,\lambda) = |A(u,\lambda)|^2 = \frac{\sigma^2(u)}{2\pi}\left|1-e^{-i\lambda}\right|^{-2d(u)},$$

where $\frac{1}{\sqrt{2\pi}}$and $\frac{1}{2\pi}$ are the transfer function and the spectral density function of Z_t respectively. Since Z_t is a i.i.d sequence , its Cramer representation is as follows

$$Z_t = \int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}}e^{i\lambda t}d\,\xi(\lambda).$$

Then,

$$\nabla^{-d(t/T)}Z_t = \int_{-\pi}^{\pi}e^{i\lambda t}\left|1-e^{-i\lambda}\right|^{-d(t/T)}(2\pi)^{-1/2}d\,\xi(\lambda).$$

Hence $X_{t,T}$ has the transfer function

$$A^0_{t,T}(\lambda) = \frac{\sigma(t/T)}{\sqrt{2\pi}}(1-e^{-i\lambda})^{-d(t/T)},$$

and

$$\sup_{t,\lambda}\left|A^0_{t,T}(\lambda) - A\left(\tfrac{t}{T},\lambda\right)\right| = 0 \leq \tfrac{K}{T}.$$

Therefore, the time-varying fractionally integrated noise is locally stationary. This result can be extended to a general locally stationary ARFIMA processes for more details see Palma and Olea (2010).

Chapter 2

Analysis of the sample mean of LSLM processes

Some asymptotic statistical properties of the sample mean of a class of LSLM process are studied in this chapter. Conditions for consistency are investigated and precise convergence rates of the variance of the sample mean are established for a class of time-varying long-memory parameter functions. A central limit theorem for the sample mean is also established. Furthermore, the calculation of the variance of the sample mean is illustrated by several numerical and simulated experiments.

2.1 Introduction

This chapter discusses the statistical properties of the sample mean of a class of LSLM processes. The analysis of the sample mean is an essential part of the theory and application of stochastic processes. As stated by Parzen (1986), *"The behavior of sample means, which needs to be understood by all applied statisticians and users of simulation methods, can be considered to be the most basic question of both classical and modern probability and statistics"*. The asymptotic behavior of the sample mean has been well established in the context of linear stationary processes see for example Section 5.6.1 of Pourahmadi (2001). In particular, several authors have studied the problem for stationary long-memory models, see for example Adenstedt (1974), Samarov and Taqqu (1988). In addition, the behavior of the sample mean has been studied in the context of short-memory locally stationary processes, see for example Dahlhaus (1996, 1997). However, to the best of our knowledge, no general asymptotic results are available yet in the context of long-memory locally stationary processes.

Locally stationary processes are becoming an important tool for analyzing non-stationary time series data. Many authors have suggested definitions for this type of processes, including Silverman (1957), Priestley (1965) and Dahlhaus (1996), among others. Furthermore, the theory of locally stationary processes has been recently extended to encompass non-stationary long-range dependent time series data, see for example Beran (2008), Genton and Perrin (2004) and Jensen and Witcher (2000). Long-memory time series has attracted a great deal attention in the last decades, see for example the monographs Beran (1994) and Palma (2007). In particular, characterization of long-memory has been studied by Parzen (1992) and Hall (1997).

The parameter estimation of LSLM processes has been studied by Beran (2008) and Jensen and Witcher (2000), among others. However, it seems that the estimation of the mean of such processes has received far less attention. In this work, we establish conditions to ensure the consistency of the sample mean and establish precise convergence rates for a family of LSLM processes with linear, quadratic or general time-varying long memory parameter. Apart from establishing these asymptotical results, this work explores the finite sample calculation of the theoretical variance of the sample mean of a LSLM. These empirical studies show that in order to be precise, the use of the asymptotic formula for the variance of the sample mean requires very large sample sizes. Thus, we offer alternative approximation formulas which work well for moderate sample sizes.

The remaining of this chapter is structured as follows. Section 3.2 discusses a class of LSLM processes. Section 3.3 establishes the consistency of the sample mean of this family of LSLM models. Section 3.4 we provides convergence rates for the variance of this estimator and shows its asymptotic normality in the Section 3.5. Section 2.6 illustrates the use of the asymptotic formulas for the variance of the sample mean as well as finite sample approximations.

2.2 Locally stationary long-memory processes

Definition 2.2.1. A sequence of stochastic processes $Y_{t,T}(t = 1, ..., T)$ is called locally stationary with transfer function A^0 and constant mean μ if there exists a spectral representation

$$Y_{t,T} = \mu + \int_{-\pi}^{\pi} e^{i\lambda t} A^0_{t,T}(\lambda) \mathrm{d}\xi(\lambda), \tag{2.2.1}$$

where the following holds.

(a) $\xi(\lambda)$ is a Brownian motion on $[-\pi, \pi]$.

(b) There exists a constant K and a 2π period function $A : (0, 1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$,

and

$$\sup_{t,\lambda} \left| A^0_{t,T}(\lambda) - A\left(\tfrac{t}{T}, \lambda\right) \right| \leq \frac{K}{T}, \tag{2.2.2}$$

for all T.

The transfer function $A^0_{t,T}(\lambda)$ of this class of non-stationary processes changes smoothly over time so that they can be locally approximated by stationary processes. An example of this class of locally stationary processes is given by the infinite moving average expansion

$$Y_{t,T} = \mu + \sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\psi_j\left(\tfrac{t}{T}\right)Z_{t-j}, \tag{2.2.3}$$

where $\{Z_t\}$ is a zero-mean and unit variance white noise and $\{\psi_j(u)\}$ are coefficients satisfying $\sum_{j=0}^{\infty}\psi_j\,(u)^2 < \infty$ for all $u \in [0,1]$. The model defined by (2.2.3) generalizes the usual Wold expansion for a linear stationary process allowing the coefficients of the infinite moving average expansion vary smoothly over time. A particular case is the generalized version of the fractional noise process described by the discrete-time equation

$$Y_{t,T} = \mu + \sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\eta_j\left(\tfrac{t}{T}\right)Z_{t-j}, \tag{2.2.4}$$

for $t = 1, 2.\ldots, T$ where $\{Z_t\}$ is a white noise sequence with zero mean and unit variance the infinite moving average coefficients $\{\eta_j(u)\}$ are given by

$$\eta_j(u) = \frac{\Gamma\left[j + d(u)\right]}{\Gamma\left(j+1\right)\Gamma\left[d(u)\right]}, \tag{2.2.5}$$

where $\Gamma(\cdot)$ is the Gamma function and $d(\cdot)$ is a smoothly time-varying long-memory parameter. For simplicity, the locally stationary fractional noise process 2.2.4 will be denoted as LSFN . Lemma A.1.1 provides a closed-form formula for calculating the covariance function $\kappa_T(s,t) = \operatorname{cov}\left(Y_{s,T}, Y_{t,T}\right)$ for a LSFN, which is useful for simulating this class of processes, see Section 2.6 for details. The class of LSFN models can be extended to the locally stationary ARFIMA processes, see Jensen and Witcher Jensen and Witcher (2000) for details. As an example, consider the locally stationary $ARFIMA(0,d,1)$ model defined by

$$Y_{t,T} = \sigma\left(\tfrac{t}{T}\right)\left[1-\theta\left(\tfrac{t}{T}\right)B\right](1-B)^{-d(t/T)}Z_t, \tag{2.2.6}$$

where $\theta(\cdot)$ is a smoothly varying moving average coefficient satisfying $|\theta(u)|<1$ for $u\in[0,1]$. Similarly to Lemma A.1.1, it can be readily proved that the covariance $\kappa_T(s,t)$ of the process (2.2.6) is given by

$$\kappa_T(s,t) = \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\frac{\Gamma\left[1-d\left(\frac{s}{T}\right)-d\left(\frac{t}{T}\right)\right]\Gamma\left[s-t+d\left(\frac{s}{T}\right)\right]}{\Gamma\left[1-d\left(\frac{s}{T}\right)\right]\Gamma\left[d\left(\frac{s}{T}\right)\right]\Gamma\left[s-t+1-d\left(\frac{t}{T}\right)\right]} \tag{2.2.7}$$
$$\times\left[1+\theta\left(\tfrac{s}{T}\right)\theta\left(\tfrac{t}{T}\right)-\theta\left(\tfrac{s}{T}\right)\frac{s-t-d\left(\frac{t}{T}\right)}{s-t-1+d\left(\frac{s}{T}\right)}-\theta\left(\tfrac{t}{T}\right)\frac{s-t-d\left(\frac{s}{T}\right)}{s-t-1+d\left(\frac{t}{T}\right)}\right],$$

for $s,t=1,\ldots,T, s\geq t$.

2.3 Consistency

In what follows we study some of the asymptotic properties of the sample mean as an estimate of μ under the following regularity conditions.

A1. *The time-varying covariance function of the process* (2.2.1) *satisfies*

$$\kappa_T(s,t)\sim g(\tfrac{s}{T},\tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1},$$

for large $s-t>0$, where $d:[0,1]\to(0,\frac{1}{2})$ and g is a $\mathcal{C}^1(\mathbb{R}\times\mathbb{R})$ function which is uniformly bounded over $[0,1]\times[0,1]$.

A2. *(*Linear Case *) The function $d(\cdot)$ is linear with positive slope, then reaches its maximum value, d_1, at $u=1$ and if $d(\cdot)$ has negative slope, then reaches its maximum value, d_0, at $u=0$.*

A3. (General Case) *The function $d(\cdot)$ reaches its maximum value, d_0, at u_0 with $d''(u_0) < 0$ and continuous third derivative.*

A4. *There exist a positive constant K such that $|\sigma(u)\psi_j(u)| \leq K j^{d_1 - 1}$.*

Note that according to Lemma A.1.1, the elements $\kappa_T(s,t)$ of the variance-covariance matrix of a locally stationary fractional noise process described by (2.2.4)-(2.2.5) are given by

$$\kappa_T(s,t) = \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\frac{\Gamma\left[1-d\left(\frac{s}{T}\right)-d\left(\frac{t}{T}\right)\right]\Gamma\left[s-t+d\left(\frac{s}{T}\right)\right]}{\Gamma\left[1-d\left(\frac{s}{T}\right)\right]\Gamma\left[d\left(\frac{s}{T}\right)\right]\Gamma\left[s-t+1-d\left(\frac{t}{T}\right)\right]},$$

for $s, t = 1, \ldots, T$, $s \geq t$. Thus, an application of the Stirling's approximation yields,

$$\kappa_T(s,t) \sim \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\frac{\Gamma\left[1-d\left(\frac{s}{T}\right)-d\left(\frac{t}{T}\right)\right]}{\Gamma\left[1-d\left(\frac{s}{T}\right)\right]\Gamma\left[d\left(\frac{s}{T}\right)\right]}(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1},$$

for large $s - t > 0$. Hence, this locally stationary fractional noise process satisfies Assumption A1. The next theorem establishes the consistency of the estimate $\widehat{\mu}_T$.

Theorem 2.3.1. (Consistency) *Assume that the process $\{Y_{t,T}\}$ satisfies* (2.2.1). *Then, under Assumptions A1–A2 the estimator $\widehat{\mu}_T$ is consistent, that is,*

$$\widehat{\mu}_T \to \mu,$$

in probability, as $T \to \infty$.

Proof. By definition, the variance of the estimator $\widehat{\mu}_T$ can be written as

$$\mathrm{Var}(\widehat{\mu}_T) = \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\kappa_T(s,t) = \frac{1}{T^2}\left[2\sum_{s>t}^{T}\kappa_T(s,t) + \sum_{s=1}^{T}\kappa_T(s,s)\right].$$

Therefore,

$$\mathrm{Var}(\widehat{\mu}_T) \sim \frac{2}{T^2}\sum_{s>t}^{T}\kappa_T(s,t),$$

as $T \to \infty$. Furthermore, given that by Assumption A1

$$\kappa_T(s,t) \sim g(\tfrac{s}{T}, \tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1},$$

for large $s - t > 0$, we conclude that

$$\mathrm{Var}(\widehat{\mu}_T) \sim \frac{2}{T^2}\sum_{s>t}^{T} g(\tfrac{s}{T}, \tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}. \tag{2.3.1}$$

Since by Assumption A1, $|g(x,y)|$ is uniformly bounded for all $(x,y) \in [0,1] \times [0,1]$ we have that

$$\begin{aligned}\mathrm{Var}(\widehat{\mu}_T) &\leq \frac{K}{T^2}\sum_{s>t}^{T}(s-t)^{2d_1-1} \leq \frac{K}{T^{1-2d_1}}\sum_{s>t}^{T}(\tfrac{s}{T}-\tfrac{t}{T})^{2d_1-1}\tfrac{1}{T^2} \\ &\leq \frac{K}{T^{1-2d_1}}\int_0^1\int_0^x (x-y)^{2d_1-1}\,dy\,dx \leq \frac{K'}{T^{1-2d_1}},\end{aligned}$$

where K' is a positive constant. Now, by Chebyshev's inequality, for any $\varepsilon > 0$ we have

$$\mathbb{P}(|\widehat{\mu}_T - \mu| > \varepsilon) \leq \frac{\mathrm{Var}(\widehat{\mu}_T)}{\varepsilon^2} \leq \frac{K'}{\varepsilon^2 T^{1-2d_1}}.$$

Since $0 < d_1 < \frac{1}{2}$, $\mathbb{P}(|\widehat{\mu}_T - \mu| > \varepsilon) \to 0$ as $T \to \infty$, proving the result. □

Observe that Theorem 2.3.1 involving the consistency of $\widehat{\mu}_T$ is valid for a general time-varying long-memory function $d(\cdot)$ satisfying Assumption 3. Thus, it is not restricted only to the linear cases.

2.4 Asymptotic Variance

We study the behavior of the variance of the sample mean for a LSLM process satisfying some regularity assumptions given by A1–A3. Before exploring that situation, recall that for a stationary long-memory process $\{y_1, \ldots, y_T\}$ with long-memory parameter d, the variance of the sample mean $\bar{y}_T = \frac{1}{T}\sum_{t=1}^{T} y_t$ behave like

$$\mathrm{Var}(\bar{y}_T) \sim c\, T^{2d-1},$$

as $T \to \infty$. Given a sample $\{Y_{1,T}, \ldots, Y_{t,T}\}$ of the process (2.2.1) we can estimate the mean of the process μ by using its sample mean $\widehat{\mu}_T = \frac{1}{T}\sum_{t=1}^{T} y_{t,T}$. The objective is to know the asymptotic variance of $\widehat{\mu}_T$.

In this section we state the convergence rate of the sample mean variance. Theorem 2.4.1 deals with a linear case and Theorem 2.4.2 deals the quadratic case while Theorem 2.4.3 focuses on a general case for time-varying long-memory parameter.

Theorem 2.4.1. *(*Linear Case *) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (2.2.1) *and* $d(u) = \alpha_0 + \alpha_1 u$ *with* $\alpha_1 > 0$*. Then, under Assumptions A1–A2 the estimator* $\widehat{\mu}_T$ *satisfies*

$$T^{1-2d_1}(\alpha_1 \log T)^{2d_1+1}\,\mathrm{Var}(\widehat{\mu}_T) \to g(1,1)\Gamma(2d_1),$$

as $T \to \infty$*. If* $\alpha_1 < 0$*, then*

$$T^{1-2d_0}(\alpha_1 \log T)^{2d_0+1}\,\mathrm{Var}(\widehat{\mu}_T) \to g(0,0)\Gamma(2d_0),$$

as $T \to \infty$*.*

Proof. By definition, the variance of the sample mean can be written as

$$\begin{aligned}
\mathrm{Var}(\widehat{\mu}_T) &= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\kappa_T(s,t) = \frac{1}{T^2}\left[2\sum_{s>t}^{T}\kappa_T(s,t) + \sum_{s=1}^{T}\kappa_T(s,s)\right] \\
&\sim \frac{2}{T^2}\sum_{s>t}^{T}\kappa_T(s,t) = \frac{2}{T^2}\sum_{s>t}^{T}\frac{\Gamma(1-d_s-d_t)}{\Gamma(1-d_s)\Gamma(d_s)}(s-t)^{(d_s+d_t-1)} \\
&= \frac{2}{T^2}\sum_{s>t}^{T} g(\tfrac{s}{T},\tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}, \qquad (2.4.1)
\end{aligned}$$

as $T \to \infty$. Therefore the sum approximation for integrates we have

$$\begin{aligned}
\operatorname{Var}(\widehat{\mu}_T) &\sim \frac{2}{T^2}\sum_{s>t}^{T} g(\tfrac{s}{T},\tfrac{t}{T})(\tfrac{s}{T}-\tfrac{t}{T})^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1} T^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1} \\
&\sim 2\int_0^1\int_0^x g(x,y)(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\, dy\, dx. \qquad (2.4.2)
\end{aligned}$$

Similarly to the proof of Lemma A.1.2, the asymptotic value of $\operatorname{Var}(\widehat{\mu}_T)$ depends only on the evaluation of the double integral (2.4.2) in a neighborhood of $(x,y)=(1,1)$. Consequently, let us define any $\varepsilon>0$ the set

$$\begin{aligned}
A_T = \{(x,y)|1-\varepsilon \le x, y \le 1+\varepsilon, 1/T < x-y, |d(x)-d_1| < \delta,\\
|d(y)-d_1| < \delta, |g(x,y)-g(1,1)| < \delta\},
\end{aligned}$$

for some $\delta > 0$. This is a nonempty set since $d(\cdot)$ and $g(\cdot)$ are continuous functions in a neighborhood of 1. Let C_T be defined as

$$C_T = T^{1-2d_1}\left(\alpha_1 \log T\right)^{2d_1+1}. \qquad (2.4.3)$$

Then,

$$\begin{aligned}
\lim_{T\to\infty} C_T \operatorname{Var}(\widehat{\mu}_T) &= \lim_{T\to\infty} 2C_T\int_0^1\int_0^x \tilde{g}(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\
&= \lim_{T\to\infty} 2C_T\int_{A_T}\int \tilde{g}(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx.
\end{aligned}$$

Therefore $1 < (x-y)T$ we have that

$$\begin{aligned}
&\lim_{T\to\infty} 2C_T\int_{A_T}\int \tilde{g}(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\
&\le [\tilde{g}(1,1)+\delta]\lim_{T\to\infty} 2C_T\int_{A_T}\int \left[(x-y)T\right]^{2\delta+2d_1-1}\, dy\, dx.
\end{aligned}$$

Therefore, by virtue of Lemma A.1.2, we conclude that

$$\lim_{T\to\infty} 2C_T\int_0^1\int_0^x g(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \le [g(1,1)+\delta]\,\Gamma(2d_1).$$

By an analogous argument, we can also conclude that

$$\lim_{T\to\infty} 2C_T \int_0^1 \int_0^x g(x,y)\,[(x-y)T]^{d(x)+d(y)-1}\,dy\,dx \geq [g(1,1)-\delta]\,\Gamma(2d_1). \tag{2.4.4}$$

Now, since ε and δ can be chosen arbitrarily small, we have that

$$\lim_{T\to\infty} 2C_T \int_0^1 \int_0^x g(x,y)\,[(x-y)T]^{d(x)+d(y)-1}\,dy\,dx = g(1,1)\Gamma(2d_1).$$

□

The negative case is similar to the previous proof, thus the evaluation of the double integral (2.4.2) is a neighborhood of point $(x,y)=(0,0)$. Therefore utilizing the Lemma A.1.3 and defining to any $\varepsilon>0$ the set

$$A_T = \{(x,y)|1-\varepsilon \leq x, y \leq 1+\varepsilon, 1/T < x-y, |d(x)-d_0| < \delta,$$
$$|d(y)-d_0| < \delta, |g(x,y)-g(0,0)| < \delta\},$$

for some $\delta>0$, and $C_T = T^{1-2d_0}(\alpha_1 \log T)^{2d_0+1}$, the result is obtained.

Theorem 2.4.2. *(*Quadratic Case*) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (2.2.1) *and* $d(u) = a + b\,u - c\,u^2$ *with* $c>0$. *If assumptions A1 and A3 are fulfilled, then the variance of* $\widehat{\mu}_T$ *satisfies*

$$T^{1-2d_0}(c/2\log T)^{d_0+\frac{1}{2}}\operatorname{Var}(\widehat{\mu}_T) \to \begin{cases} \frac{1}{2}\sqrt{\pi} g(u_0,u_0)\Gamma(d_0) & \text{if} \quad u_0 \in (0,1) \\ \frac{1}{4}\sqrt{\pi} g(u_0,u_0)\Gamma(d_0) & \text{if} \quad u_0 = 0,1. \end{cases}$$

as $T\to\infty$.

Proof. From expression (2.4.1) we have

$$\begin{aligned} \operatorname{Var}(\widehat{\mu}_T) &\sim 2\sum_{s>t}^{T} g(\tfrac{s}{T},\tfrac{t}{T})(\tfrac{s}{T}-\tfrac{t}{T})^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1} T^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}\tfrac{1}{T^2} \\ &\sim 2\int_0^1\int_0^x g(x,y)(x-y)^{d(x)+d(y)-1}T^{d(x)+d(y)-1}\,dy\,dx. \end{aligned} \tag{2.4.5}$$

Now, by means of the variable transformation $u = x + y$ and $v = x - y$, we can write

$$d(x) + d(y) - 1 = 2d_0 - 1 - 2\beta[(x - u_0)^2 + (y - u_0)^2] = \alpha(u) - \beta v^2,$$

where $\alpha(u) = 2d_0 - 1 - \beta(u - 2u_0)^2$ and $\beta = c/2$. Thus,

$$\begin{aligned}\mathrm{Var}(\widehat{\mu}_T) \ &\sim \ \int_0^1 \int_0^u \tilde{g}(u,v) v^{\alpha(u)-\beta v^2} T^{\alpha(u)-\beta v^2}\, dv\, du \\ &+ \ \int_1^2 \int_0^{2-u} \tilde{g}(u,v) v^{\alpha(u)-\beta v^2} T^{\alpha(u)-\beta v^2}\, dv\, du,\end{aligned}$$

where $\tilde{g}(u,v) = g(\frac{u+v}{2}, \frac{u-v}{2})$. Therefore,

$$\begin{aligned}&T^{1-2d_0}(\beta \log T)^{d_0+\frac{1}{2}}\, \mathrm{Var}(\widehat{\mu}_T) \\ &\quad \sim \int_0^1 T^{-\beta(u-2u_0)^2} \left(\sqrt{\beta \log T}\right)^{1+\beta(u-2u_0)^2} h_T(u)\, du \\ &\quad + \ \int_1^2 T^{-\beta(u-2u_0)^2} \left(\sqrt{\beta \log T}\right)^{1+\beta(u-2u_0)^2} h_T(2-u)\, du,\end{aligned} \tag{2.4.6}$$

where

$$h_T(u) = \left(\sqrt{\beta \log T}\right)^{\alpha(u)+1} \int_0^u \tilde{g}(u,v) v^{\alpha(u)-\beta v^2} T^{-\beta v^2}\, dv. \tag{2.4.7}$$

Now, an application of Lemma A.1.8 yields,

$$\begin{aligned}&T^{1-2d_0}(\beta \log T)^{d_0+\frac{1}{2}}\, \mathrm{Var}(\widehat{\mu}_T) \\ &\quad \sim \int_0^1 T^{-\beta(u-2u_0)^2} \left(\sqrt{\beta \log T}\right)^{1+\beta(u-2u_0)^2} h(u)\, du \\ &\quad + \ \int_1^2 T^{-\beta(u-2u_0)^2} \left(\sqrt{\beta \log T}\right)^{1+\beta(u-2u_0)^2} h(u)\, du,\end{aligned}$$

where

$$h(u) = \frac{1}{2}\tilde{g}(u,0)\Gamma\left[\frac{\alpha(u)+1}{2}\right].$$

On the other hand,

$$\int_0^1 T^{-\beta(u-2u_0)^2}\left(\sqrt{\beta\log T}\right)^{1+\beta(u-2u_0)^2} h(u)\,du$$
$$= \sqrt{n}\int_0^1 \exp[c_n(u-2u_0)^2]h(u)\,du,$$

where $n=\beta\log T$ and $c_n = n-\beta\log\sqrt{n}$. Since $c_n/n\to 1$ as $n\to\infty$ and $c_n>0$, by Lemma A.1.9 we conclude that

$$\int_0^1 T^{-\beta(u-2u_0)^2}\left(\sqrt{\beta\log T}\right)^{1+\beta(u-2u_0)^2} h(u)\,du \to$$
$$\sqrt{\pi}h(2u_0)I_{(0,1)}(2u_0)+\frac{\sqrt{\pi}}{2}h(2u_0)I_{\{0,1\}}(2u_0),$$

as $T\to\infty$, where I_A is the indicator function of A. An analogous argument leads to

$$\int_1^2 T^{-\beta(u-2u_0)^2}\left(\sqrt{\beta\log T}\right)^{1+\beta(u-2u_0)^2} h(u)\,du \to$$
$$\sqrt{\pi}h(2u_0)I_{(0,1)}(2u_0-1)+\frac{\sqrt{\pi}}{2}h(2u_0)I_{\{0,1\}}(2u_0-1).$$

as $T\to\infty$. Now, by observing that

$$h(2u_0)=\frac{1}{2}\tilde{g}(2u_0,0)\Gamma\left[\frac{\alpha(2u_0)+1}{2}\right]=\frac{1}{2}g(u_0,u_0)\Gamma(d_0),$$

the result is proved. □

Theorem 2.4.3. (General Case) *Assume that the process $\{Y_{t,T}\}$ satisfies* (2.2.1). *If assumptions A1 and A3 are fulfilled, then the variance of $\widehat{\mu}_T$ satisfies*

$$T^{1-2d_0}(\log T)^{d_0+\frac{1}{2}}\operatorname{Var}(\widehat{\mu}_T)\to V(u_0),$$

as $T\to\infty$ where

$$V(u_0)=\begin{cases}\frac{4^{d_0}\sqrt{\pi}g(u_0,u_0)\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}} & \textit{if} \quad u_0\in(0,1)\\ \frac{1\sqrt{\pi}g(u_0,u_0)\Gamma(d_0)}{4^{d_0-1}[-d''(u_0)]^{d_0+1/2}} & \textit{if} \quad u_0=0,1.\end{cases} \tag{2.4.8}$$

Proof. Similarly to the proof of Theorem 2.4.2, the asymptotic value of $\mathrm{Var}(\widehat{\mu}_T)$ depends only on the evaluation of the integral (2.4.5) in a neighborhood of $(x, y) = (u_0, u_0)$. Consequently, let us define for any $\varepsilon > 0$ the set

$$\begin{aligned} A_T = \{(x,y) | u_0 - \varepsilon \leq x, y \leq u_0 + \varepsilon, 1/T < x - y, |d''(x) - d''(u_0)| < \delta, \\ |d''(y) - d''(u_0)| < \delta, |g(x,y) - g(u_0,u_0)| < \delta\}, \end{aligned}$$

for some $\delta > 0$. This is a nonempty set since $d''(\cdot)$ and $g(\cdot,\cdot)$ are continuous functions in a neighborhood of u_0. Define $C_T = T^{1-2d_0} (\log T)^{d_0 + \frac{1}{2}}$. Then,

$$\begin{aligned} \lim_{T\to\infty} C_T \,\mathrm{Var}\left(\widehat{\mu}_T\right) &= \lim_{T\to\infty} 2C_T \int_0^1 \int_0^x g(x,y) \left[(x-y)T\right]^{d(x)+d(y)-1} \, dy \, dx \\ &= \lim_{T\to\infty} 2C_T \int_{A_T}\!\!\int g(x,y) \left[(x-y)T\right]^{d(x)+d(y)-1} \, dy \, dx. \end{aligned}$$

Since $1 < (x - y)T$ we have that

$$\begin{aligned} &\lim_{T\to\infty} C_T \int_{A_T}\!\!\int g(x,y) \left[(x-y)T\right]^{d(x)+d(y)-1} \, dy \, dx \\ &\quad \leq [g(u_0,u_0) + \delta] \lim_{T\to\infty} C_T \int_{A_T}\!\!\int \left[(x-y)T\right]^{2d_0 + [d''(u_0)-\delta][(x-u_0)^2 + (y-u_0)^2]/2 - 1} \, dy \, dx. \end{aligned}$$

Then it follows by Theorem 2.4.2 we have that

$$\begin{aligned} &\lim_{T\to\infty} C_T \int_{A_T}\!\!\int g(x,y) \left[(x-y)T\right]^{d(x)+d(y)-1} \, dy \, dx \\ &\qquad \leq [g(u_0,u_0) + \delta] \frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{\left[\delta - d''(u_0)\right]^{d_0+1/2}}. \end{aligned}$$

By an analogous argument, we can also conclude that

$$\begin{aligned} &\lim_{T\to\infty} C_T \int_{A_T}\!\!\int g(x,y) \left[(x-y)T\right]^{d(x)+d(y)-1} \, dy \, dx \\ &\qquad \geq [g(u_0,u_0) - \delta] \frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{\left[-\delta - d''(u_0)\right]^{d_0+1/2}}. \qquad (2.4.9) \end{aligned}$$

Now, since ε and δ can be chosen arbitrarily small, we have that

$$\lim_{T\to\infty} C_T \int_{A_T}\int g(x,y)\,[(x-y)T]^{d(x)+d(y)-1}\,dy\,dx$$
$$= g(u_0,u_0)\frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}}.$$

A similar argument yields the result for $u_0 = 0, 1$. □

2.5 Normality

The next theorem establishes the asymptotic normality of $\widehat{\mu}_T$. Observe that we have added the assumption that the input noise $\{Z_t\}$ in the generalized Wold expansion (2.2.3) is a sequence of independent identically distributed random variables. As noted by Hosking (1996)(p.264), this assumption seems to be essential for the existence of a central limit theorem for the sample mean.

Theorem 2.5.1. *(*Normality*) Assume that the process $\{Y_{t,T}\}$ satisfies* (2.2.3) *where $\{Z_t\}$ is a sequence of independent identically distributed random variables. Then under Assumptions A1, A3 and A4*

$$T^{1-2d_0}(\log T)^{2d_0+1}(\widehat{\mu}_T-\mu)\to N[0,V(u_0)],$$

as $T\to\infty$, where $V(u_0)$ is given by (2.4.8)*.*

Proof. We adapt the Theorem 18.6.5 by Ibragimov and Linnik (1971), as corrected by Hosking (1996). Without loss of generality, assume that $\mu = 0$ and define $S_T = \sum_{t=1}^{T} Y_{t,T}$. Then, we can write

$$S_T = \sum_{k=-\infty}^{T} c_{k,T} Z_k,$$

where the coefficients $\{c_{k,T}\}$ are given by

$$c_{k,T} = \sum_{j=\max\{1,k\}}^{T} \sigma\left(\tfrac{j}{T}\right)\psi_{j-k}\left(\tfrac{j}{T}\right).$$

Let $\sigma_T^2 = \mathrm{Var}(S_T)$ just as Ibragimov and Linnik's show that the ratio $\frac{c_{k,T}}{\sigma_T}$ converges to zero uniformly as $T \to \infty$. In what follows, we prove that this is indeed the case for the class of locally stationary processes under study. First, observe that from Assumption A3 we may conclude that

$$|c_{k,T}| \leq KT^{d_0}, \tag{2.5.1}$$

for $k \leq T$. On the other hand, note that $\sigma_T^2 = T^2\,\mathrm{Var}(\widehat{\mu}_T)$. Hence, by (2.4.9) we have that

$$\frac{C_T}{T^2}\sigma_T^2 \geq \frac{4^{d_0}\sqrt{\pi}\Gamma\left(d_0\right)\left[g(u_0,u_0)-\delta\right]}{\left[-\delta - d''(u_0)\right]^{d_0+1/2}},$$

for large T, where C_T is defined in (2.4.3). Since $d''(u_0) < 0, g(u_0,u_0) > 0$, $\Gamma(u_0) > 0$ for any $u_0 \in [0,1]$ and δ can be chosen arbitrarily small, there exists a constant $K > 0$ such that

$$\frac{C_T}{T^2}\sigma_T^2 \geq K,$$

for large T. Hence

$$\frac{1}{\sigma_T} \leq K\frac{\sqrt{C_T}}{T}. \tag{2.5.2}$$

Now, by (2.5.1) and (2.5.2) we conclude that

$$a_T = \frac{|c_{k,T}|}{\sigma_T} \leq K\frac{(\log T)^{d_0/2+1/4}}{\sqrt{T}},$$

in which $a_T \to 0$ uniformly as $T \to \infty$. $\square$

2.6 Numerical and Simulation Studies

This section discusses the calculation of the variance of the sample mean of LSLM processes, assessing the accuracy of the asymptotic formula provided by Theorems 2.4.1–2.4.3 comparing the sample variance obtained from several simulations to their theoretical counterparts. These calculations are illustrated with a locally stationary fractional noise process with linear, quadratic and general long-memory function. Given that the calculation of the exact variance of the sample mean is a highly demanding task for large sample sizes, we examine other approximate methods.

2.6.1 Linear long-memory function

Example 2.6.1. *For the LSFN process with time-varying long-memory parameter given by*

$$d(u) = 0.2 + 0.15\,u, \tag{2.6.1}$$

for $u \in [0, 1]$.

This function is depicted in Figure2.1.

A realization of this process with $4,000$ observation is shown in Figure 2.2. The samples of this LSFN process used in these simulations are generated by means of the innovation algorithm, see for example (Brockwell and Davis, 1991, p.172). In this implementation, the variance-covariance matrix of the process, $\kappa_T(s, t)$, is given by Lemma A.1.1.

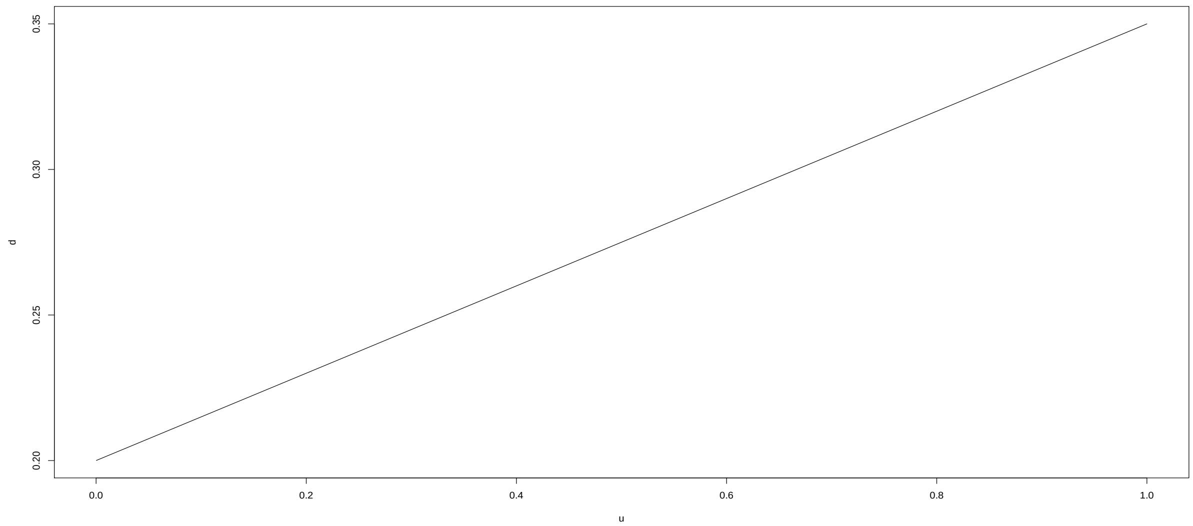

Figure 2.1: *Time varying long-memory function* $d(u) = a + bu$, $u \in [0, 1]$ *with* $a = 0.2$ *and* $b = 0.15$ *.*

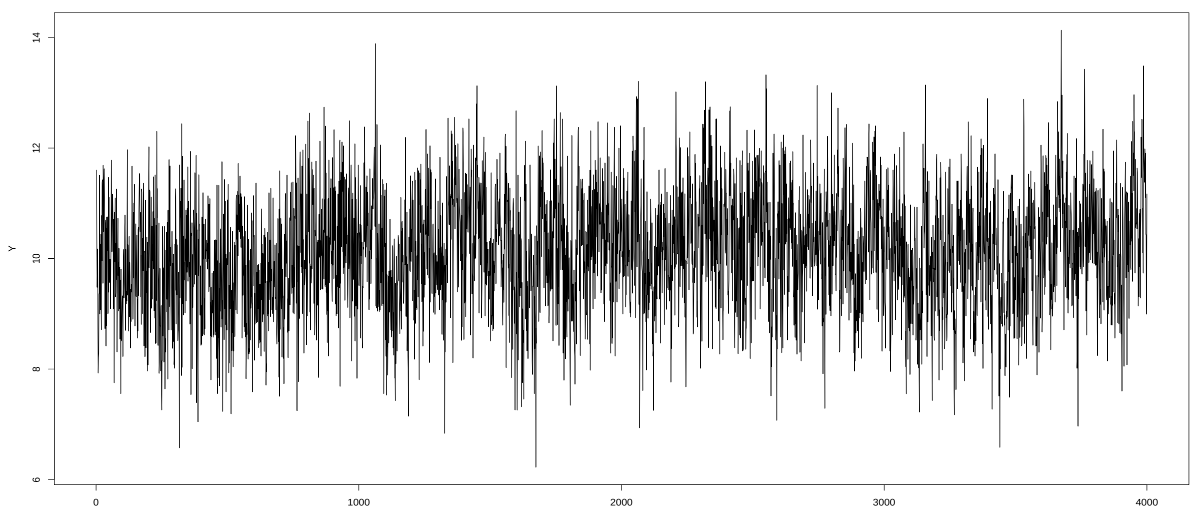

Figure 2.2: *Simulated locally stationary process with* $4,000$ *observations, with linear long-memory function.*

The following tables report a set of simulation and numerical experiments to illustrate the calculation of the variance of the sample mean. We consider locally stationary fractional noise models with time-varying parameter specified by (2.6.1) and different sample sizes. On the other hand, calculating the exact value of the variance of the sample mean is a demanding computational task, especially for large sample sizes. The exact value of the variance of the sample mean is given by

$$\operatorname{Var}(\widehat{\mu}_T) = \frac{1}{T^2} \sum_{s=1}^{T} \sum_{t=1}^{T} \kappa_T(s,t), \tag{2.6.2}$$

and using formula (2.4.2) we may obtain the approximation,

$$\operatorname{Var}(\widehat{\mu}_T) \quad \sim \quad 2 \int_0^1 \int_0^x g(x,y)(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\, dy\, dx, \tag{2.6.3}$$

where

$$g(x,y) = \frac{\Gamma\left[1-d(x)-d(y)\right]}{\Gamma\left[1-d(x)\right]\Gamma\left[d(x)\right]}.$$

For simplicity, this formula will be denoted as *Approximation*. On the other hand we have the asymptotic variance given by the Theorem 2.4.1

$$T^{1-2\alpha_0-2\alpha_1}(\alpha_1 \log T)^{2\alpha_0+2\alpha_1+1} \operatorname{Var}(\widehat{\mu}_T) \to g(1,1)\Gamma(2\alpha_0+2\alpha_1), \tag{2.6.4}$$

as $T \to \infty$. we will denote formula *Asymptotic.*

Table 2.1 reports the values of the variance of the sample mean for three sample sizes $T = 1,000$, $T = 2,000$ and $T = 4,000$ obtained from the following four approaches: Exact, Sample, Approximation and the Asymptotic formula. The first row of the table provides the exact values of the variance of $\widehat{\mu}_T$ given by formula (2.6.2). The second row corresponds to the average of this value over 1,000 repetitions. The third row corresponds to Approximation given by (2.6.3) and asymptotic method give

Table 2.1: *Estimation of the mean: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.0588310	0.04444876	0.03358069
Sample	0.0583592	0.04644588	0.03292527
Approximation	0.0554740	0.04168669	0.03139149
Asymptotic	0.1305255	0.09011318	0.06310254

Table 2.2: *Estimation of the mean: Ratio of Approximation and Asymptotic Variance*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation/Asymptotic	0.5723458	0.9803262	1.001104

by (2.6.4) for the variance of $\widehat{\mu}_T$. From this table, note that the sample mean variance values from the simulations (second row) and Approximation Variance(third row) are relatively close to their theoretical counterparts displayed in the first row. On the other hand, Asymptotic Variance formula seems to be far off from the exact value for these three sample sizes. Thus, for these sample sizes, the asymptotic formula is not very useful for calculating the variance of $\widehat{\mu}_T$. In order to evaluate the accuracy of the asymptotic formula for larger sample sizes, the Table 2.2 reports the ratio between approximation given for formula (2.6.3) and Asymptotic Variance formula (2.6.4). Due to the large sample sizes involved in this table, in these experiments we have not calculated the exact variance de $\widehat{\mu}_T$ nor the sample values. From this table, we observe that the values Approximation and Asymptotic Variance is quite close when the number of sample size increases.

2.6.2 Quadratic long-memory function

Example 2.6.2. *Consider the following illustrative example consisting of a LSFN process defined by* (2.2.4)–(2.2.5) *with time-varying long-memory parameter given by*

$$d(u) = \tfrac{4}{17} + \tfrac{2}{17}\,u - \tfrac{4}{17}\,u^2, \tag{2.6.5}$$

for $u \in [0, 1]$.

This function, depicted in Figure 2.3, has a maximum value $d_0 = 0.25$ reached at $u_0 = 0.25$.

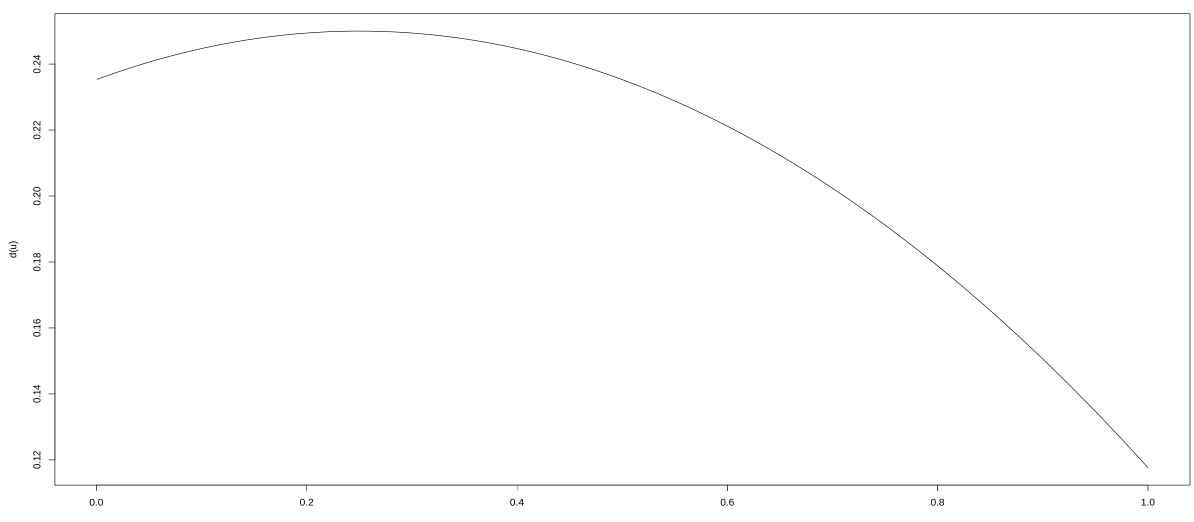

Figure 2.3: *Time varying long-memory function* $d(u) = a + bu - cu^2$, $u \in [0, 1]$ *with* $a = 4/17$, $b = 2/17$ *and* $c = 4/17$.

The following two tables report a set of simulation and numerical experiments to illustrate the calculation of the variance of the sample mean. We consider locally stationary fractional noise models with time-varying parameter specified by (2.6.5) and different sample sizes. In this section we also discuss other approximated method for calculating the variance of the mean.

As previous case the exact value of the variance of the sample mean is given by the formula (2.6.2). Another approximation of the variance of the sample mean can be obtained as follows (For more details see Appendix Lemma A.1.10). From expression (2.4.6) with $u_0 = 0.25$ we have that

$$
\begin{aligned}
\mathrm{Var}(\widehat{\mu}_T) &\sim \frac{T^{2d_0-1}}{(\beta\log T)^{d_0+\frac{1}{2}}}\int_0^1 T^{-\beta(u-2u_0)^2}\left(\sqrt{\beta\log T}\right)^{1+\beta(u-2u_0)^2} h_T(u)\,du \\
&\sim \int_0^1 T^{\alpha(u)}\left(\sqrt{\beta\log T}\right)^{-\alpha(u)-1} h_T(u)\,du.
\end{aligned}
$$

But, from Lemma A.1.8 equations (A.1.8)-(A.1.9) we have that

$$
h_T(u) \sim \tfrac{1}{2} g(\tfrac{u}{2},\tfrac{u}{2})\,\gamma\left[\tfrac{\alpha(u)+1}{2}, \beta(\log T)u^2\right],
$$

where $\gamma(x,a)$ corresponds to the incomplete Gamma function

$$
\gamma(x,a) = \int_0^x t^{a-1}\exp(-t)\;dt.
$$

Hence,

$$
\mathrm{Var}(\widehat{\mu}_T) \sim \frac{1}{2}\int_0^1 \frac{T^{\alpha(u)}}{\left(\sqrt{\beta\log T}\;\right)^{\alpha(u)+1}}\, g(\tfrac{u}{2},\tfrac{u}{2})\,\gamma\left[\tfrac{\alpha(u)+1}{2}, \beta(\log T)u^2\right]\;du. \tag{2.6.6}
$$

This formula will be denoted as *Approximation 1*. Finally, we can approximate the value of the variance of the sample mean by the asymptotic expression provided by Theorem 2.4.2,

$$
\mathrm{Var}(\widehat{\mu}_T) \sim \frac{1}{2}\sqrt{\pi} g(u_0,u_0)\Gamma(d_0)T^{2d_0-1}(\beta\log T)^{-d_0-\frac{1}{2}}. \tag{2.6.7}
$$

Table 2.3: *Estimation of the mean: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.02034246	0.01419231	0.009870528
Sample	0.01879493	0.01329861	0.009766211
Approximation	0.02283783	0.01566236	0.01075406
Approximation 1	0.01477839	0.01033301	0.006440145
Asymptotic	0.04735846	0.03116997	0.02064407

For simplicity, this expression will be denoted as *Asymptotic* formula. Table 2.3 reports the values of the variance of the sample mean for three sample sizes $T = 1,000$, $T = 2,000$ and $T = 4,000$ obtained from the following five approaches: exact, sample, Approximation, Approximation 1 and the Asymptotic formula. The first row of the table provides the exact values of the variance of $\widehat{\mu}_T$ given by (2.6.2). The second row corresponds to the average of this value over 1,000 repetitions. The third and fourth rows correspond to the sample mean variances obtained from Approximation and Approximation 1, given by formulas (2.6.3) and (2.6.6), respectively. The fifth column shows the approximated values of the variance of $\widehat{\mu}_T$ provided by the asymptotic formula (2.6.7). From this table, note that the sample mean variance values from the simulations (second row) and Approximation 1 (third row) are relatively close to their theoretical counterparts displayed in the first row. On the other hand, Approximation 1 and the Asymptotic formula seems to be far off from the exact value for these three sample sizes. Thus, for these sample sizes, the asymptotic formula is not very useful for calculating the variance of $\widehat{\mu}_T$.

In order to evaluate the accuracy of the asymptotic formula for larger sample sizes, Table 2.4 reports the ratios between Approximation and Approximation 1 to the asymptotic formula. From this table, the asymptotic formula seems to produce accurate values, but for quite large sample sizes.

Table 2.4: *Estimation of the mean: Ratio of variances*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation /Asymptotic	0.5595917	0.9391358	0.9822582
Approximation 1/Asymptotic	0.3921184	0.9463146	1.000580

2.6.3 Cubic long-memory function

Example 2.6.3. *Extending the previous example let $\{y_{t,T} : t \in \mathbb{Z}\}$ be a LSFN process with time-varying long-memory parameter given by*

$$d(u) = \tfrac{3}{17} + \tfrac{2}{17}\,u + \tfrac{1}{17}\,u^2 - \tfrac{3}{17}\;u^3, \tag{2.6.8}$$

for $u \in [0,1]$.

This function, depicted in Figure 2.4, has a maximum value $d_0 = 0.230$ reached at $u_0 = 0.595$.

The following two tables, report a set of simulation and numerical experiments to illustrate the calculation of the variance of the sample mean. We can also approximate

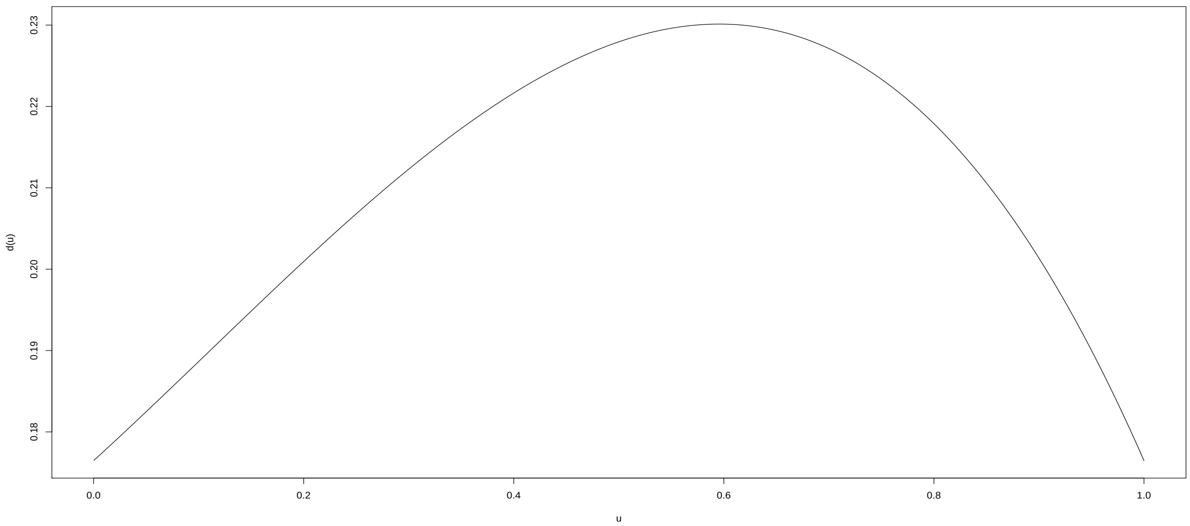

Figure 2.4: *Time varying long-memory function* $d(u) = a + bu + cu^2 - eu^3$, $u \in [0, 1]$ *with* $a = 3/17$, $b = 2/17$,$c = 1/17$ *and* $c = 3/17$.

the value of the variance of the sample mean by the asymptotic expression given by

$$\mathrm{Var}(\widehat{\mu}_T) \sim \frac{1}{2}\sqrt{\pi} g(u_0, u_0)\Gamma(d_0) T^{2d_0-1} \left[-\frac{d''(u_0)}{2^2} \log T \right]^{-d_0-\frac{1}{2}}. \tag{2.6.9}$$

For simplicity, this expression will be denoted as *Asymptotic* formula. Table 2.5 reports the values of the variance of the sample mean for three sample sizes $T = 1,000$, $T = 2,000$ and $T = 4,000$. In this table the fourth row show the approximated values of the variance of $\widehat{\mu}_T$ provided by the asymptotic formula (2.6.9). One can see that the variance of the sample mean given in the first three rows are very close. To evaluate the accuracy of the asymptotic formula for larger sample sizes, Table 2.6 reports the ratio between approximation given for formula (2.6.3) and Asymptotic Variance formula (2.6.9). From this table, we observe that the values Approximation and Asymptotic Variance is quite close when the number of sample size increases.

Table 2.5: *Estimation of the mean: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.01811858	0.01238319	0.008430995
Sample	0.01869189	0.01245654	0.008560451
Approximation	0.01902065	0.01280057	0.008617324
Asymptotic	0.03191383	0.02047275	0.013213600

Table 2.6: *Estimation of the mean: Ratio of Approximation and Asymptotic Variance*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation/Asymptotic	0.7098559	1.014965	1.004061

2.6.4 General long-memory function

The rate of convergence of the asymptotic variance for the sample mean established for a general class of time-varying long memory parameter functions, turns out to be accurate for large sample sizes, as shown in the following examples.

Example 2.6.4. *Consider the locally stationary fractional noise process with time-varying long-memory parameter given by*

$$d(u) = 0.1 + u\exp(-2\,u), \tag{2.6.10}$$

for $u \in [0, 1]$.

This function, depicted in Figure 2.5, has a maximum value $d_0 = 0.284$ reached at $u_0 = 0.500$.

In this example we evaluate the asymptotic formula for the variance of the sample mean given by Theorem 2.4.3. The Tables 2.7 and 2.8 show us that the asymptotic variance formula is accurate but this is accomplished with a very large sample sizes.

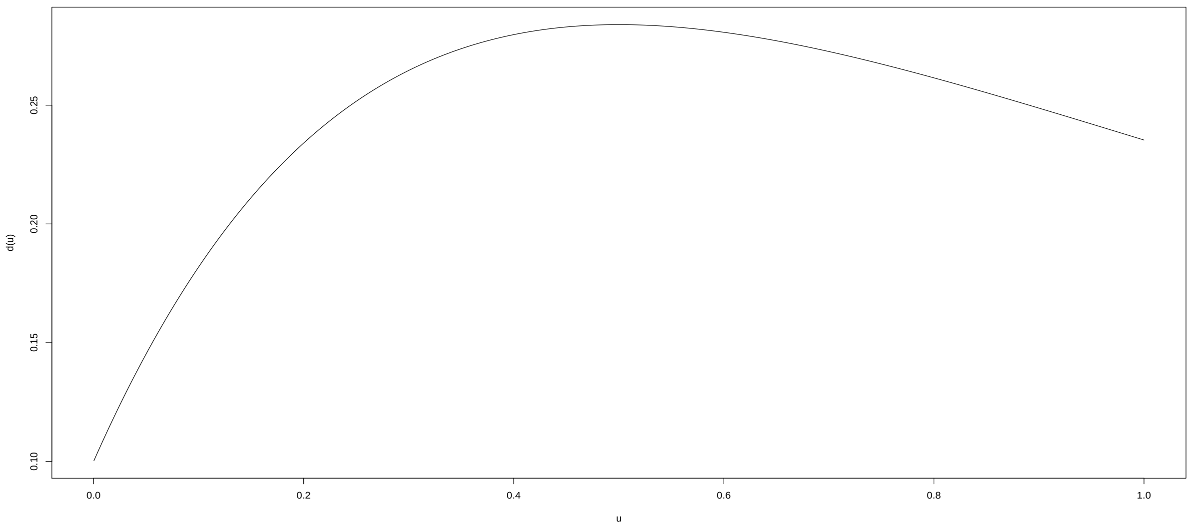

Figure 2.5: *Time varying long-memory function* $d(u) = a + bu\exp(-cu),\ u \in [0,1]$ *with* $a = 0.1,\ b = 1$ *and* $c = 2$.

Table 2.7: *Estimation of the mean: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.03664112	0.02657624	0.01923992
Sample	0.03794652	0.02698353	0.01941801
Approximation	0.03683037	0.02651523	0.01910474
Asymptotic	0.05977541	0.04110415	0.02845064

Table 2.8: *Estimation of the mean: Ratio of Approximation and Asymptotic Variance*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation/Asymptotic	0.7279555	1.025402	1.006205

Example 2.6.5. *Figure 2.6 show us an example of time-varying long-memory parameter, given by*

$$d(u) = 0.5 - 0.2\exp(-\,u) - 0.15\,u^2, \tag{2.6.11}$$

for $u \in [0,1]$.

From the figure we can observe that the function $d(\,u)$ has a maximum value $d_0 = 0.342$ reached at $u_0 = 0.433$.

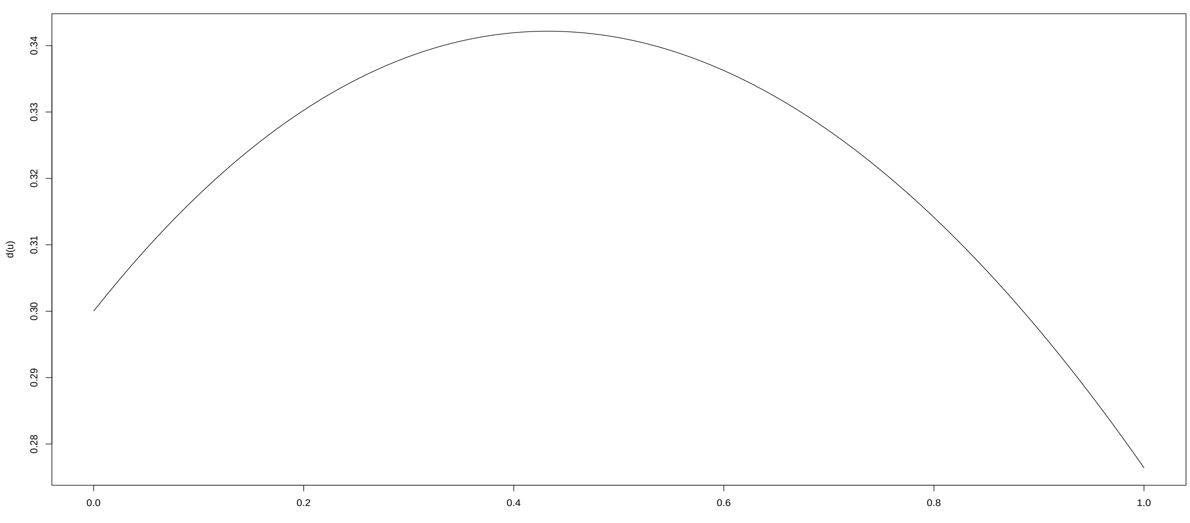

Figure 2.6: *Time varying long-memory function* $d(u) = a - b\exp(-u) - cu$, $u \in [0,1]$ *with* $a = 0.5$, $b = 0.2$ *and* $c = 0.15$.

Table 2.9 reports the values of the variance of the sample mean for three sample sizes $T = 1,000$, $T = 2,000$ and $T = 4,000$. In the same way as in the examples above, the sample mean variance and Approximation are relatively close to their theoretical counterparts.

Table 2.9: *Estimation of the mean: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.11413985	0.09024096	0.07124653
Sample	0.1117166	0.09347473	0.07115020
Approximation	0.11670980	0.09183322	0.07227960
Asymptotic	0.26638680	0.19747530	0.14742360

The Table 2.10 reports the variance ratio between Approximation given by formula (2.6.3) and Asymptotic Variance formula given by Theorem 2.4.3. From this table, we observe that the values Approximation and Asymptotic Variance is quite close when the number of sample size increases.

Table 2.10: *Estimation of the mean: Ratio of Approximation and Asymptotic Variance*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation/Asymptotic	0.5461717	0.981355	0.9978066

Chapter 3

Estimation of a time-varying trend for a LSLM process

It is apparent from the time series graphs of many economic and climatology series, that they share certain characteristics. In particular there is a tendency, especially noticeable for price index for example, that increase (or decrease) over time.

In this chapter we analyze the problem of relating a time series with a time-varying trend, in many practical applications, the behavior of a time series may be related to the behavior of other components or regressors. A widely used approach to model these relationships is the linear regression analysis.

We explore some asymptotic statistical properties for the LSE of a linear regression model with LSLM disturbances. In this chapter we analyze the simple case of a vector non-stochastic regressor, if this vector is equal to unit vector, then we have the case of a time series with a constant mean, where the LSE is the sample mean. In Chapter 4 we extend the asymptotic statistical properties introduced in this chapter to a p-vector of non-stochastic regressors.

3.1 Introduction

Figure 3.1 displayed a simulated LSFN process with 2,000 observations and $d(u) = a + b\,u$, it is evident from a visual inspection of this serie that there is a trend upward over time. However, in this chapter we consider trends in a more general class of processes that vary over time.

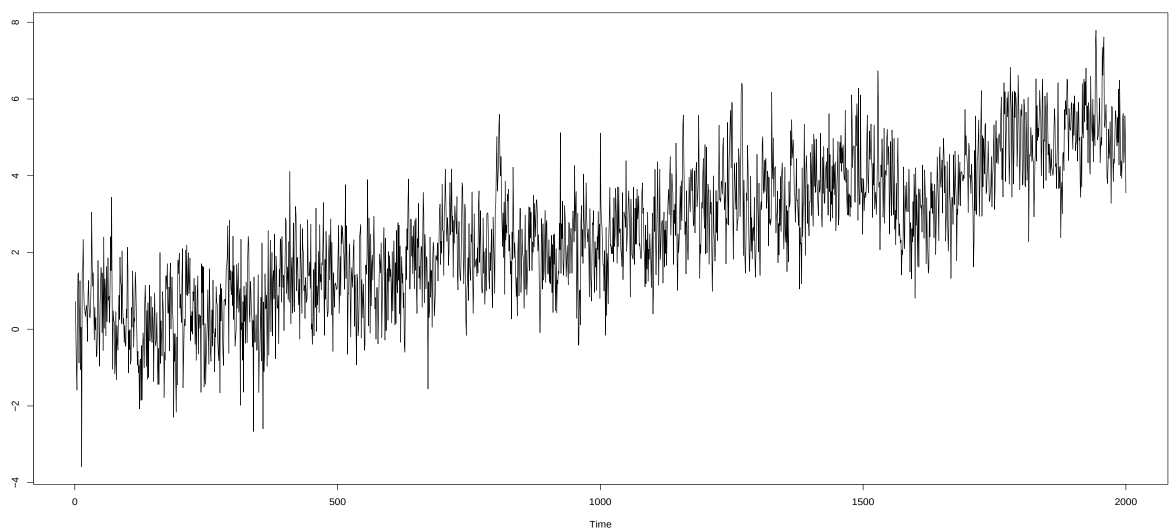

Figure 3.1: *Simulated LSFN process with* $d(u) = a + bu$, $u \in [0, 1]$ *with* $a = 0.3$ *and* $b = 0.1$ *.*

In the same way as in the previous chapter, we provide asymptotic results for calculation of the theoretical variance of the LSE of a LSLM process. These empirical studies show that in order to be precise, the use of the asymptotic formula for the variance of the LSE requires very large sample sizes. Thus, we offer alternative approximation formulas which work well for moderate sample sizes.

3.2 Locally stationary long-memory processes

In this chapter, we go one step further with the analysis in Chapter 2 and study how this non-stationary behavior can be modeled. We focus on a very simple model of non-stationarity with a scalar vector no-stochastic regressors and the sequence of errors from a LSLM evolving with time. In this section, we also derive some basic properties of these processes. We adapt the definition of locally stationary processes given in Section 1.3 to a more general case, where $Y_{t,T}$ corresponds to a model with a time-varying trend.

Definition 3.2.1. A sequence of stochastic processes $Y_{t,T}(t = 1, ..., T)$ is called locally stationary with transfer function A^0 and time-varying trend if there exists a spectral representation

$$Y_{t,T} = x(\tfrac{t}{T})\beta + \int_{-\pi}^{\pi} e^{i\lambda t} A^0_{t,T}(\lambda) \mathrm{d}\xi(\lambda), \tag{3.2.1}$$

where the following holds.

(a) $\xi(\lambda)$ is a Brownian motion on $[-\pi, \pi]$

(b) There exists a constant K and a 2π period function $A : (0, 1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$,

and

$$\sup_{t,\lambda} \left| A^0_{t,T}(\lambda) - A\left(\tfrac{t}{T}, \lambda\right)\right| \leq \frac{K}{T}, \tag{3.2.2}$$

for all T.

An example more general of this class of locally stationary processes is given by the infinite moving average expansion

$$Y_{t,T} = x\left(\tfrac{t}{T}\right)\beta + \sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\psi_j\left(\tfrac{t}{T}\right)Z_{t-j}, \tag{3.2.3}$$

where $\{Z_t\}$ is a zero-mean and unit variance white noise and $\{\psi_j(u)\}$ are coefficients satisfying $\sum_{j=0}^{\infty}\psi_j(u)^2 < \infty$ for all $u \in [0,1]$. The model defined by (3.2.3) generalizes the usual Wold expansion for a linear stationary process allowing the coefficients of the infinite moving average expansion vary smoothly over time. A particular case is the generalized version of the fractional noise process described by the discrete-time equation

$$Y_{t,T} = x\left(\tfrac{t}{T}\right)\beta + \sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\eta_j\left(\tfrac{t}{T}\right)Z_{t-j}, \tag{3.2.4}$$

for $t = 1, 2. \ldots, T$ where $\{Z_t\}$ is a white noise sequence with zero mean and unit variance the infinite moving average coefficients $\{\eta_j(u)\}$ are given by (2.2.5)

3.3 Consistency

Consider the following linear regression model

$$Y_{t,T} = x(\tfrac{t}{T})\beta + \varepsilon_{t,T}, \tag{3.3.1}$$

for $t = 1, 2, \ldots, T$ where $\{Y_{t,T}\}$ is an observed sequence, $x(\frac{t}{T})$ a scalar of no-stochastic regressors, $\{\varepsilon_{t,T}\}$ sequence of errors is a LSLM processes and β is unknown regression parameter. We then proceed to the analysis of some large sample properties of the LSE under the following regularity conditions.

A5. *Parameter regression is a function of* $(\frac{t}{T})$ *which satisfies*

$$x_{t,T} = x\left(\tfrac{t}{T}\right) \to x(u),$$

as $\frac{t}{T} \to u$ *for all* $u \in [0,1]$ *and function which is uniformly bounded over* $[0,1]$.

A6. *There exist a positive constant* K *such that* $|x(u)\sigma(u)\psi_j(u)| \leq K j^{d_1-1}$.

The LSE the β is given by

$$\widehat{\beta}_T = \frac{1}{\sum_{t=1}^{T} x_{t,T}^2} \sum_{t=1}^{T} y_{t,T} x_{t,T}.$$

The definition of LSE is equivalent to

$$\widehat{\beta}_T = v_T^{-1} \sum_{t=1}^{T} y_{t,T} x_{t,T},$$

where v_T is given by

$$v_T = \sum_{t=1}^{T} x_{t,T}^2 \sim T \int_0^1 x^2(u)\mathrm{d}u \sim Tk. \tag{3.3.2}$$

The consistency of the least square estimator $\widehat{\beta}_T$ is established in the next theorem.

Theorem 3.3.1. *(*Consistency*) Consider the linear model* (3.3.1) *where the process* $\{Y_{t,T}\}$ *satisfies* (3.2.1)*. Then, under Assumptions A1, A3 and A5 the estimator* $\widehat{\beta}_T$ *is consistent, that is,*

$$\widehat{\beta}_T \to \beta,$$

in probability ,as $T \to \infty$.

Proof. By definition, the variance of the estimator $\widehat{\beta}_T$ can be written as

$$\begin{aligned}
\mathrm{Var}\left(\widehat{\beta}_T\right) &= \mathrm{Var}\left(v_T^{-1}\sum_{t=1}^{T} y_t x\left(\tfrac{t}{T}\right)\right) \\
&= [v_T^{-1}]^2 \sum_{t=1}^{T}\sum_{s=1}^{T} x\left(\tfrac{t}{T}\right)\kappa_T(s,t)x\left(\tfrac{s}{T}\right) \\
&= [v_T^{-1}]^2\left[2\sum_{s>t}^{T} x\left(\tfrac{t}{T}\right)x\left(\tfrac{s}{T}\right)\kappa_T(s,t) + \sum_{s=1}^{T}\kappa_T(s,s)x\left(\tfrac{s}{T}\right)^2\right].
\end{aligned}$$

Therefore,

$$\mathrm{Var}\left(\widehat{\beta}_T\right) \sim 2[v_T^{-1}]^2\sum_{s>t}^{T} x\left(\tfrac{t}{T}\right)x\left(\tfrac{s}{T}\right)\kappa_T(s,t),$$

as $T\to\infty$. Furthermore, given that by Assumption A1

$$\kappa_T(s,t) \sim g(\tfrac{s}{T},\tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1},$$

for large $s-t>0$, we conclude that

$$\mathrm{Var}\left(\widehat{\beta}\right) \sim 2[v_T^{-1}]^2\sum_{s>t}^{T} x\left(\tfrac{t}{T}\right)x\left(\tfrac{s}{T}\right)g(\tfrac{s}{T},\tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}.$$

Since by Assumption A1, $|g(x,y)|$ is uniformly bounded for all $(x,y)\in[0,1]\times[0,1]$ and by Assumption A5 the scalar $x_{t,T}(u)$ for $u\in[0,1]$ is bounded, next we have that

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}_T) &\leq K[v_T^{-1}]^2\sum_{s>t}^{T}(s-t)^{2d_0-1} \leq K[v_T^{-1}]^2\sum_{s>t}^{T}(\tfrac{s}{T}-\tfrac{t}{T})^{2d_0-1}T^{2d_0-1} \\
&\leq \frac{K}{T^{1-2d_0}}\int_0^1\int_0^x (x-y)^{2d_0-1}\,dy\,dx \leq \frac{K'}{T^{1-2d_0}}.
\end{aligned}$$

Therefore we conclude that

$$\mathrm{Var}(\hat{\beta}_T) \quad\leq\quad K'T^{2d_0-1}\to 0,$$

as $T\to\infty$. Now, by Chebyshev's inequality, for any $\varepsilon>0$ we have

$$\mathbb{P}(|\widehat{\beta}_T-\beta|>\varepsilon) \leq \frac{\mathrm{Var}(\widehat{\beta}_T)}{\varepsilon^2} \leq \frac{KT^{2d_0-1}}{\varepsilon^2}.$$

Since , $\mathbb{P}(|\widehat{\beta}_T-\beta|>\varepsilon)\to 0$ as $T\to\infty$, proving the result. □

3.4 Asymptotic Variance of the LSE

In this section we analyze the asymptotic variance of the LSE for the regression parameter of the process defined in (3.3.1). Given a sample $\{y_{1,T},\ldots,y_{t,T}\}$ we know the LSE for the regression parameter is given by

$$\begin{aligned}\hat{\beta} &= \frac{1}{\sum_{t=1}^{T} x(\frac{t}{T})^2}\sum_{t=1}^{T} y_{t,T}\, x(\frac{t}{T}) \\ &= \sum_{t=1}^{T} y_{t,T}\,\rho_{t,T},\end{aligned}$$

where $\rho_{t,T} = \frac{x(\frac{t}{T})}{\sum_{t=1}^{T} x(\frac{t}{T})^2}$. An application of Assumption A5 yields,

$$T\rho\left(\tfrac{t}{T}\right) \to \frac{x(u)}{\int_0^1 x^2(y)\mathrm{d}y}.$$

The next results specify the convergence rate of the asymptotic variance of the estimator.

Theorem 3.4.1. *(*Linear Case *) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (3.2.1) *and* $d(u)=\alpha_0+\alpha_1\,u$ *with* $\alpha_1>0$. *Then, under Assumptions A1, A2 and A5 the estimator* $\widehat{\beta}_T$ *satisfies*

$$T^{1-2d_1}(\alpha_1\log T)^{2d_1+1}\,\mathrm{Var}(\widehat{\beta}_T) \to \frac{x(1)^2 g(1,1)\Gamma(2d_1)}{\left[\int_0^1 x^2(y)\,dy\right]^2},$$

If $\alpha_1<0$, *then*

$$T^{1-2d_0}(\alpha_1\log T)^{2d_0+1}\,\mathrm{Var}(\widehat{\beta}_T) \to \frac{x^2(0) g(0,0)\Gamma(2d_0)}{\left[\int_0^1 x^2(y)\,dy\right]^2},$$

as $T\to\infty$.

Proof. The variance of the estimator $\widehat{\beta}_T$ can be written as

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}_T) &= \sum_{s=1}^{T}\sum_{t=1}^{T}\rho_T(s)\rho_T(t)\kappa_T(s,t) \\
&= \left[2\sum_{s>t}^{T}\rho_T(s)\rho_T(t)\kappa_T(s,t)+\sum_{s=1}^{T}\rho_T(s)^2\kappa_T(s,s)\right] \\
&\sim 2\sum_{s>t}^{T}\rho_T(s)\rho_T(t)\kappa_T(s,t) \\
&= 2\sum_{s>t}^{T}\rho_T(s)\rho_T(t)\frac{\Gamma(1-d_s-d_t)}{\Gamma(1-d_s)\Gamma(d_s)}(s-t)^{(d_s+d_t-1)} \\
&= 2\sum_{s>t}^{T}\rho(\tfrac{s}{T})\rho(\tfrac{t}{T})g(\tfrac{s}{T},\tfrac{t}{T})(s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1},
\end{aligned}$$

as $T\to\infty$. Therefore the sum approximation for integrates we have

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}_T) &\sim 2\sum_{s>t}^{T}\rho(\frac{s}{T})\rho(\frac{t}{T})g(\tfrac{s}{T},\tfrac{t}{T})(\tfrac{s}{T}-\tfrac{t}{T})^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}T^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1} \\
&\sim 2\int_0^1\int_0^x T^2\rho(x)\rho(y)g(x,y)(x-y)^{d(x)+d(y)-1}T^{d(x)+d(y)-1}\,dy\,dx \\
&\sim 2\int_0^1\int_0^x \tilde{g}(x,y)(x-y)^{d(x)+d(y)-1}T^{d(x)+d(y)-1}\,dy\,dx. \qquad (3.4.1)
\end{aligned}$$

where

$$\tilde{g}(u,v)=\frac{x\left(\frac{u+v}{2}\right)x\left(\frac{u-v}{2}\right)}{\left[\int_0^1 x^2(y)\,dy\right]^2}g\left(\tfrac{u+v}{2},\tfrac{u-v}{2}\right). \qquad (3.4.2)$$

Similarly to the proof of Theorem 2.4.1, the asymptotic value of $\mathrm{Var}(\widehat{\beta}_T)$ depends only on the evaluation of the double integral (3.4.1) in a neighborhood of $(x,y)=(1,1)$. Consequently, let us define any $\varepsilon>0$ the set

$$\begin{aligned}
A_T = \{(x,y)|1-\varepsilon\le x,y\le 1+\varepsilon, 1/T<x-y, |d(x)-d_1|<\delta, \\
|d(y)-d_1|<\delta, |\tilde{g}(x,y)-\tilde{g}(1,1)|<\delta\},
\end{aligned}$$

for some $\delta > 0$. This is a nonempty set since $d(\cdot)$ and $\tilde{g}(\cdot)$ are continuous functions in a neighborhood of 1 the proof follows of the Theorem 2.4.1. $\square$

Theorem 3.4.2. *(*General Case*) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (3.2.1)*. If assumptions A1, A3 and A5 are fulfilled, then the variance of* $\widehat{\beta}_T$ *satisfies*

$$T^{1-2d_0}(\log T)^{d_0+\frac{1}{2}}\operatorname{Var}(\widehat{\beta}_T) \to V(u_0),$$

as $T \to \infty$ *and*

$$V(u_0) = \begin{cases} \frac{4^{d_0}\sqrt{\pi}h(u_0,u_0)\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}} & \text{if} \quad u_0 \in (0,1) \\ \frac{\sqrt{\pi}h(u_0,u_0)\Gamma(d_0)}{4^{d_0-1}[-d''(u_0)]^{d_0+1/2}} & \text{if} \quad u_0 = 0,1, \end{cases} \tag{3.4.3}$$

where $h(\cdot,\cdot)$ *is given by*

$$h(u_0,u_0) = \frac{g(u_0,u_0)x(u_0)^2}{\left[\int_0^1 x^2(y)\,dy\right]^2}$$

Proof. Similar to the proof of Theorem 2.4.3 $\square$

3.5 Normality

Theorem 3.5.1. *(*Normality*) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (3.2.3) *where* $\{Z_t\}$ *is a sequence of independent identically distributed random variables. Then under Assumptions A1, A3 and A6*

$$T^{1-2d_0}\,(\log T)^{2d_0+1}\left(\widehat{\beta}_T - \beta\right) \to N[0, V(u_0)],$$

as $T \to \infty$*, where* $V(u_0)$ *is given by* (3.4.3)*.*

Proof. Define $S_T = \sum_{t=1}^{T} x\left(\frac{t}{T}\right) Y_{t,T}$. Then, we can write

$$S_T = \sum_{k=-\infty}^{T} c_{k,T} Z_k,$$

where the coefficients $\{c_{k,T}\}$ are given by

$$c_{k,T} = \sum_{j=\max\{1,k\}}^{T} x\left(\tfrac{j}{T}\right)\sigma\left(\tfrac{j}{T}\right)\psi_{j-k}\left(\tfrac{j}{T}\right).$$

Let$\sigma_T^2 = \operatorname{Var}(S_T)$. As pointed out by Hosking (1996), the key aspect of Ibragimov and Linnik's proof is showing that $\frac{c_{k,T}}{\sigma_T}$ converges to zero uniformly as $T \to \infty$. In what follows, we prove that this is indeed the case for the class of locally stationary processes under study. First, observe that from Assumption A6 we may conclude that

$$|c_{k,T}| \leq KT^{d_0}, \tag{3.5.1}$$

for $k \leq T$. On the other hand, note that $\sigma_T^2 = v_T^2 \operatorname{Var}(\widehat{\beta}_T)$ where v_T is as (3.3.2). Hence, by (2.4.9) we have

$$\frac{C_T}{T^2}\sigma_T^2 \geq \frac{4^{d_0}\sqrt{\pi}\Gamma\left(d_0\right)\left[h(u_0,u_0)-\delta\right]}{\left[-\delta - d''(u_0)\right]^{d_0+1/2}},$$

for large T, where C_T is defined in (2.4.3). Since $h(u_0,u_0) > 0$, $\Gamma(d_0) > 0$ and δ can be chosen arbitrarily small, there exists a constant $K > 0$ such that

$$\frac{C_T}{v_T^2}\sigma_T^2 \geq K,$$

for large T. Hence

$$\frac{1}{\sigma_T} \leq K\frac{\sqrt{C_T}}{v_T}. \tag{3.5.2}$$

Now, by (3.5.1) and (3.5.2) we conclude that

$$a_T = \frac{|c_{k,T}|}{\sigma_T} \leq K\frac{\left(\alpha_1 \log T\right)^{d_0/2+1/4}\sqrt{T}}{v_T},$$

in which $a_T \to 0$ uniformly as $T \to \infty$. $\square$

3.6 Numerical and Simulation Studies

Example 3.6.1. *Let $\{y_{t,T}\}$ be harmonic Model given by*

$$Y_{t,T} = \sin\left(\omega \tfrac{t}{T}\right)\beta + \varepsilon_{t,T},$$

where $\omega = 1.6$ and $\varepsilon_{t,T}$ is a LSFN process descriptive by (3.2.1). Consider the time-varying long-memory parameter given by

$$d(u) = 0.2 + 0.15\,u, \tag{3.6.1}$$

for $u \in [0, 1]$.

To illustrate the calculation of the variance of the LSE, consider a trajectory $\{y_{1,T}, \ldots, y_{t,T}\}$ of a simulated LSFN process with time-varying long-memory parameter given in (3.6.1), time-varying trend $\sin\left(1.6\frac{t}{T}\right)$ and sample size $T = 4,000$ displayed in Figure 3.2. As in the previous case the sample of this LSFN processes are generated by means of the innovation algorithm. In this implementation, the variance-covariance matrix of the process, $\kappa_T(s,t)$, is given by Lemma A.1.1.

Note what in this case the Asymptotic Variance is given by

$$(\alpha_1 \log T)^{2\alpha_0+2\alpha_1+1} T^{1-2\alpha_0-2\alpha_1}\,\mathrm{Var}(\widehat{\beta}_T) \to \frac{x(1)^2 g(1,1)\Gamma(2\alpha_0+2\alpha_1)}{\left[\int_0^1 x^2(y)\,dy\right]^2}, \tag{3.6.2}$$

as $T \to \infty$ and Approximation is give by

$$\mathrm{Var}(\widehat{\beta}_T) \sim 2\int_0^1 \int_0^x \tilde{g}(x,y)(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\,dy\,dx. \tag{3.6.3}$$

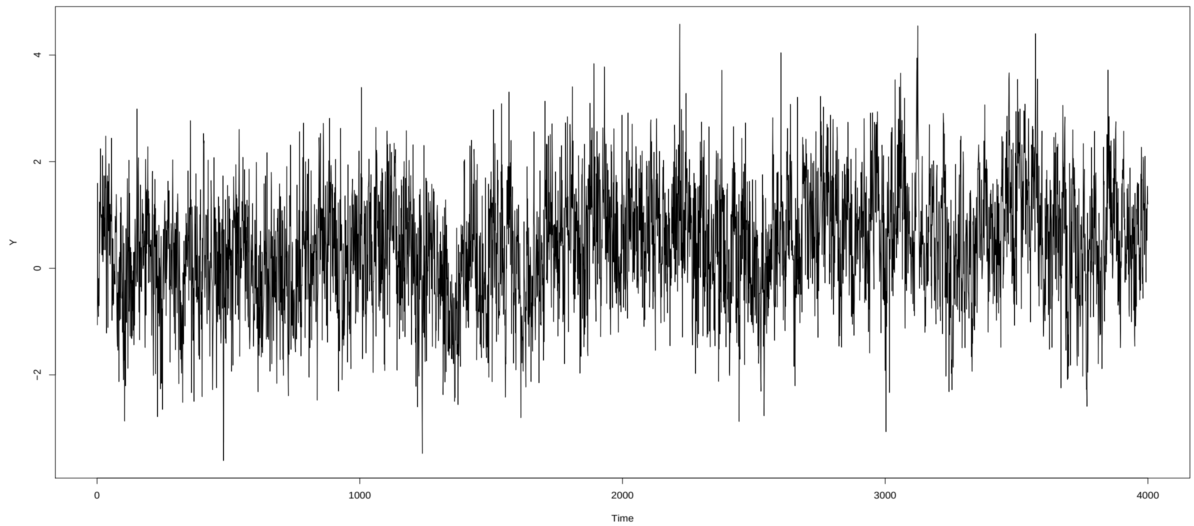

Figure 3.2: *Simulated locally stationary process with* $4,000$ *observations.*

where

$$\tilde{g}(x,y) = \frac{x(x)x(y)}{\left[\int_0^1 x^2(y)\,dy\right]^2} g(x,y).$$

The Table 3.1 reports the values of the variance for the LSE of β. From the Table 3.1 we can see that the values of the first row are similar to the values of the Approximation (third row). To reproduce the Theoretical Variance (first row) for T very large computational cost is very high we will use the ratio between the Approximation formula and the Asymptotic variance to establish the accuracy of the asymptotic variance formula. Table 3.2 shows the ratio between the Asymptotic Variance (3.6.2) and the Approximation (3.6.3) for large enough values of T. It is observed that this ratio is closer to one. In following example discusses the calculation of the variance of the LSE, assessing the accuracy of the asymptotic formula provided by Theorem 3.4.2.

Example 3.6.2. *Extending the previous example, let* $\{y_{t,T}\}$ *be an harmonic model*

Table 3.1: *Estimation Harmonic Model: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.13783469	0.10643133	0.08208923
Sample	0.12648179	0.1043008	0.0775631
Approximation	0.1325978	0.1018098	0.07825691
Asymptotic	0.5031331	0.3473569	0.2432397

Table 3.2: *Estimation Harmonic Model: Ratio of Approximation and Asymptotic Variance*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation/Asymptotic	0.3878928	0.9702865	1.001508

given by

$$Y_{t,T} = \cos\left(\omega \tfrac{t}{T}\right)\beta + \varepsilon_{t,T},$$

where $\omega = \frac{\pi}{2}$ *and* $\varepsilon_{t,T}$ *is a LSLM process described by* (3.2.1) *and the time-varying long-memory parameter is given by*

$$d(u) = 0.15 + 1.5\, u \exp(-2\, u), \tag{3.6.4}$$

for $u \in [0, 1]$.

This function, depicted in Figure 2.5, has a maximum value $d_0 = 0.284$ reached at $u_0 = 0.500$. The Table 3.3 reports the values of the variance for the LSE for three sample sizes $T = 1,000$, $T = 2,000$ and $T = 4,000$. In this case the asymptotic formula is give by Theorem 3.4.2 for the variance of $\widehat{\beta}_T$. As in the previous case we

Table 3.3: *Estimation Harmonic Model: Variance of the estimate*

Method	Sample Size		
	$T = 1,000$	$T = 2,000$	$T = 4,000$
Exact	0.0607423	0.04384365	0.0316032
Sample	0.0606481	0.04143008	0.0315631
Approximation	0.0589109	0.04220783	0.0302805
Asymptotic	0.2296406	0.15791080	0.1092995

Table 3.4: *Estimation Harmonic Model: Ratio of Approximation and Asymptotic Variance*

Method	Sample Size		
	$\log T = 10$	$\log T = 100$	$\log T = 500$
Approximation/Asymptotic	0.5608171	0.9462125	0.9983411

will use the ratio between the approximation and the asymptotic variance to establish the accuracy of the asymptotic variance formula. Table 3.4 show us the ratio between the Asymptotic Variance and the Approximation (3.6.3) for large enough values of T. From this table, the Asymptotic formula seems to produce accurate values for very large sample sizes.

Chapter 4

Regression estimation with LSLM disturbances

The LSE of the chapter 3 has been deliberately simple, restricted to the bivariate case involving just the variables $Y_{t,T}$ and $x_{t,T}$, the observed process and regressor respectively. While there are very few applications in which only two variables are involved most of the important principles in estimation can be illustrated with this simple case. The extension to the multivariate case is straightforward given the framework outlined in the bivariate case. For example, the method of least squares still proceeds by defining the residual sum of squares and seeking the estimators that result in a minimum. Similarly the principles of asymptotic theory are the same. In particular in this chapter we are interested in extend the asymptotic properties for LSE to linear regression model to more than two regressors. This extension is made very much easier if matrix-vector notations are used.

4.1 Introduction

Let the observed process $\{Y_{t,T}\}$ follow the regression model

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \varepsilon_{t,T},$$

where $X(\frac{t}{T}) = (x_{t1}, \ldots, x_{tp})'$ is a p-vector of non-stochastic regressors $\beta = (\beta_1, \ldots, \beta_p)'$ is a vector of unknown regression parameters, and $\varepsilon_{t,T}$ the sequence of errors of a LSLM processes. We discuss the asymptotic properties of the LSE for the unknown parameter more specifically the consistency, asymptotic variance and normality of the LSE under the family of LSLM processes. However this model does not have a stationary property, which is crucial in the theory of estimation and asymptotic of time series models. Asymptotic properties of the LSE in a regression model with long memory stationary errors ε_t has been studied by Yajima (1991). However, the case that $\varepsilon_{t,T}$ is a long-memory locally stationary process has not been clarified fully yet since this process causes considerable mathematical difficulties. In this chapter, we establish conditions for consistency and establishes precise convergence rates of the variance of the LSE for a family of LSLM processes with general time-varying long memory parameter. Apart from establishing these asymptotical results, this chapter explores the finite sample calculation of the theoretical variance of the LSE of a LSLM. The remaining of this chapter is structured as follows: Section (4.2) discusses a class of LSLM processes; Section (4.3) establishes the consistency of the LSE of this family of LSLM models; In section (4.4) will discuss the asymptotic variance of this estimator differentiating the linear case and the general case for the time-varying long-memory parameter; Section (4.5) we discuss an asymptotic distribution for the LSE. Here we impose the condition that the white noise process are a sequence of independent identically distributed random; Section (4.6) we discuss the parameter estimate by minimization of a generalization of the Whittle function where the usual periodogram is replaced by local periodograms over data segments, here the unknown parameter β is estimated by LSE $\widehat{\beta}$; Section (4.7) illustrates the use of the asymptotic formulas for the variance of the LSE as well as finite sample approximations.

4.2 Locally stationary long-memory processes

Definition A sequence of stochastic processes $\{Y_{t,T}\}$ is called locally stationary with transfer function A^0 if there exists a spectral representation

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \int_{-\pi}^{\pi} e^{i\lambda t} A^0_{t,T}(\lambda) \mathrm{d}\xi(\lambda), \tag{4.2.1}$$

for $t = 1, \ldots, T$, where the following holds.

(a) $\xi(\lambda)$ is a Brownian motion on $[-\pi, \pi]$

(b) There exists a constant K and a 2π period function $A : [0,1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, -\lambda) = \overline{A(\mu, \lambda)}$, and

$$\sup_{t,\lambda} \left| A^0_{t,T}(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq \frac{K}{T},$$

for all T. In this definition two different functions $A^0_{t,T}(\lambda)$ and $A\left(\frac{t}{T}, \lambda\right)$ are defined. This complicated construction is necessary if we want to model a class of processes which is rich enough to cover interesting applications. In particular, if we do not define these two functions in the above definition, then the class does no longer include time-varying $AR(p)$ processes (as showed in Dahlhaus (1996)).

Observe that we have used the same convention for the asymptotic concept than in Chapter 1. This implies that the non-stationary process is doubly-indexed. The smoothness of A in u defines the departure from stationarity and ensures the locally stationary behavior of the process. An example of this class of locally stationary processes is given by the infinite moving average expansion

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\psi_j\left(\tfrac{t}{T}\right)Z_{t-j}, \tag{4.2.2}$$

where $\{Z_t\}$ is a zero-mean and unit variance white noise and $\{\psi_j(u)\}$ are coefficients satisfying $\sum_{j=0}^{\infty}\psi_j\left(u\right)^2 < \infty$ for all $u \in [0,1]$. The model defined by (4.2.2) generalizes the usual Wold expansion for a linear stationary process allowing the coefficients of the infinite moving average expansion vary smoothly over time. A particular case is the generalized version of the fractional noise process described by the discrete-time equation

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \sigma\left(\frac{t}{T}\right)\sum_{j=0}^{\infty}\eta_j\left(\frac{t}{T}\right)Z_{t-j}, \tag{4.2.3}$$

for $t = 1, 2. \ldots, T$ where $\{Z_t\}$ is a white noise sequence with zero mean and unit variance the infinite moving average coefficients $\{\eta_j(u)\}$ are given by

$$\eta_j(u) = \frac{\Gamma\left[j + d(u)\right]}{\Gamma\left(j+1\right)\Gamma\left[d(u)\right]}, \tag{4.2.4}$$

where $\Gamma\left(\cdot\right)$ is the Gamma function and $d(\cdot)$ is a smoothly time-varying long-memory coefficient. A natural extension of the LSFN model is the locally stationary autoregressive fractionally integrated moving average LS-ARFIMA process is defined by

$$\Phi\left(\frac{t}{T}\right)Y_{t,T} = \Theta\left(\frac{t}{T}\right)(1-B)^{-d\left(\frac{t}{T}\right)}\sigma\left(\frac{t}{T}\right)Z_t, \tag{4.2.5}$$

for $t = 1, 2, \ldots n$, where for $u \in [0,1]$, $\Phi\left(u, B\right) = 1 + \phi_1(u)B + \ldots. + \phi_p(u)B^p$ is an autoregressive polynomial, $\Theta\left(u, B\right) = 1 + \theta_1(u)B + \ldots. + \theta_q B^q$ is a moving average polynomial, $d(u)$ is a long-memory parameter, $\sigma(u)$ is a scale factor and Z_t is a zero-mean and unit variance white noise.

4.3 Consistency

We shall consider the regression model of the form

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \varepsilon_{t,T}, \tag{4.3.1}$$

for $t = 1, 2, \ldots, T$ where $Y_{t,T}$ is an observed sequence, $X(\frac{t}{T})$ is a p-vector no-stochastic regressors, $\{\varepsilon_{t,T}\}$ sequence of errors under the class of LSLM processes and β is unknown regression parameter. In what follows we study some of the asymptotic properties of the LSE under the following regularity conditions.

A7. *The time-varying spectral density of the process* (4.2.1) *is strictly positive and satisfies*

$$f_\theta(u,\lambda) = \frac{C_f(\theta,u)}{|1-e^{-i\lambda}|^{2d(u)}},$$

where $C_f(\theta,u)$ *is a nonnegative bounded function,* $\lambda \in [-\pi,\pi]$ *and* $0 < d(u) < \frac{1}{2}$ *for all* $u \in [0,1]$ *and* $\theta \in \Theta$. *As a particular case of the assumption, consider the extension of the usual fractional noise process with time-varying long parameter, described by* (4.2.3)-(4.2.4). *The spectral density of this LS-FN process given by*

$$f_\theta(u,\lambda) = \frac{\sigma^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-d_\theta(u)}.$$

By $f_\theta(u,\lambda) = |A(u,\lambda)|^2$ *we denote the time varying spectral density of our process, from* (4.2.4) *the transfer function of this process satisfies*

$$\int_{-\pi}^{\pi} A(u,\lambda)A(u,-\lambda)\exp(ik\lambda) = \frac{\Gamma[1-d(u)-d(v)]\Gamma[k+d(u)]}{\Gamma[1-d(u)]\Gamma[d(u)]\Gamma[k+1-d(v)]},$$

for $k \geq 0$. Thus, by Stirling's approximation we get

$$\int_{-\pi}^{\pi} A(u,\lambda)A(u,-\lambda)\exp(ik\lambda) \sim \frac{\Gamma[1-d(u)-d(v)]}{\Gamma[1-d(u)]\Gamma[d(u)]} k^{d(u)+d(v)-1},$$

for $k \to \infty$.

A8. *The time-varying covariance function of the process* (4.2.1) *satisfies*

$$\kappa_T(s,t) \sim g(\tfrac{s}{T},\tfrac{t}{T})(s-t)^{d(\frac{s}{T})+d\left(\frac{t}{T}\right)-1},$$

for large $s-t>0$, where $d:[0,1] \to (0,\frac{1}{2})$ and g is a $\mathcal{C}^1(\mathbb{R}\times\mathbb{R})$ function which is uniformly bounded over $[0,1]\times[0,1]$.

A9. *The p-vector no-stochastic regressors are a function continuous what satisfied*

$$X(\tfrac{t}{T}) \to X(u),$$

as $\frac{t}{T} \to u$ for all $u \in [0,1]$ and function which is uniformly bounded over $[0,1]$.

A10. *There exist a positive constant K such that $|\alpha_i x_i\,(u)\,\sigma\,(u)\,\psi_{j-k}\,(u)| \leq Kj^{d_0-1}$ for $i=1,\ldots,p$, for all $u \in [0,1]$ and $j \geq 1$.*

Given a sample $\{y_{1,T},\ldots,y_{t,T}\}$ of the process (4.2.1) we know the least squares estimators for the regression parameters β is given for

$$\widehat{\beta} = (X_T X_T')^{-1} X_T Y_T,$$

where $X_T X_T'$ is given by

$$X_T X_T' = \sum_{t=1}^{T} x_{t,T} x_{t,T}' = V^T,$$

where $x_{t,T} = (x_{t,1}, \ldots, x_{t,p})' = \left(x_1\left(\frac{t}{T}\right), \ldots, x_p\left(\frac{t}{T}\right)\right)' = x\left(\frac{t}{T}\right)$ is a sequence of regressors. Hence V^T can be write

$$V_{i,j}^T = \sum_{t=1}^{T} x_i\left(\tfrac{t}{T}\right) x_j'\left(\tfrac{t}{T}\right), \tag{4.3.2}$$

to $i, j = 1, \ldots p$. Therefore V^T is a matrix the $p \times p$ that satisfy be positive definite and under assumption A8 we have that V^T satisfies the following condition

$$T\left[V_{i,j}^T\right]^{-1} \to \left[\int_0^1 x(y)x'(y)\,dy\right]_{i,j}^{-1},$$

as $T \to \infty$. Analogy to $X_T Y_T$ be can written

$$X_T Y_T = \sum_{t=1}^{T} x_{t,T} y_{t,T}.$$

Therefore the least squares estimator can be write

$$\hat{\beta}_T = \left[V^T\right]^{-1} \sum_{t=1}^{T} x_{t,T} y_{t,T}.$$

The consistency of the LSE is established in the next theorem.

Theorem 4.3.1. *Consider the linear model* (4.3.1) *where the sequence of observations* $\{Y_{t,T}\}$ *satisfies* (4.2.1) *with the spectral density* $f_\theta(u, \lambda)$ *of A7. Then, under Assumptions A3, A8 and A9 the estimator* $\widehat{\beta}_T$ *is consistent, that is,*

$$\widehat{\beta}_T \to \beta,$$

in probability , as $T \to \infty$

Proof. Let α be a fixed vector, the variance of the estimator $\widehat{\beta}_T$ can be written as

$$\begin{aligned}
\mathrm{Var}(\alpha'\widehat{\beta}_T) &= \mathrm{Var}(\alpha' V_T^{-1}\sum_{t=1}^{T} x_t y_t) \\
&= \alpha' V_T^{-1}\sum_{t=1}^{T} x_t \,\mathrm{Var}(y_t) x_t' V_T^{-1}\alpha \\
&= \alpha' V_T^{-1}\sum_{s,t=1}^{T} x_t \gamma(s-t) x_t' V_T^{-1}\alpha \\
&= \int_{-\pi}^{\pi} \left| \sum_{t=1}^{T} \alpha' V_T^{-1} x_t \exp{(i\lambda t)} \right|^2 f_\theta(u,\lambda)\, d\lambda \\
&\leq KT^{2d(u)} \sum_{t=1}^{T} (\alpha' V_T^{-1} x_t)^2 \leq KT^{2d_0} \sum_{t=1}^{T} (\alpha' V_T^{-1} x_t)^2 \\
&= KT^{2d_0} \alpha' V_T^{-1}\alpha.
\end{aligned}$$

The last result is given by Lemma A.2.1. Therefore we conclude that

$$\mathrm{Var}(\widehat{\beta}_T) \leq KT^{2d_0} V_T^{-1} \leq KT^{2d_0 - 1} \to 0,$$

as $T \to \infty$. Now, by Chebyshev's inequality, for any $\varepsilon > 0$ we have

$$\mathbb{P}(|\widehat{\beta}_T - \beta| > \varepsilon) \leq \frac{\mathrm{Var}(\widehat{\beta}_T)}{\varepsilon^2} \leq \frac{KT^{2d_0-1}}{\varepsilon^2}.$$

Since , $\mathbb{P}(|\widehat{\beta}_T - \beta| > \varepsilon) \to 0$ as $T \to \infty$, proving the result.

$\square$

4.4 Asymptotic variance

The asymptotic variance of the LSE is analyzed in this section. The next results specify rate of convergence for asymptotic variance of $\widehat{\beta_T}$.

Theorem 4.4.1. *(*Linear Case *) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (4.2.1) *and* $d(u) = \alpha_0 + \alpha_1 u$ *with* $\alpha_1 > 0$*. Then, under Assumptions A2, A8 and A9, the estimator* $\widehat{\beta}_T$ *satisfies*

$$T^{1-2d_1}(\alpha_1 \log T)^{2d_1+1}\,\mathrm{Var}(\widehat{\beta}_T) \quad \to \quad G(1,1)\Gamma(2d_1),$$

as $T \to \infty$*. If* $\alpha_1 < 0$*, then*

$$T^{1-2d_0}(\alpha_1 \log T)^{2d_0+1}\,\mathrm{Var}(\widehat{\beta}_T) \quad \to \quad G(0,0)\Gamma(2d_0),$$

as $T \to \infty$*. Where* G *is an* $p \times p$ *matrix with elements given by*

$$G(u,u) = \left[\int_0^1 x(v)x'(v)\,dv\right]^{-1} x(u)x'(u)\left[\int_0^1 x(v)x'(v)\,dv\right]^{-1} g(u,u). \qquad (4.4.1)$$

Proof. By definition , the variance of the estimator $\widehat{\beta}_T$ can be written as

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}_T) &= \mathrm{Var}\left([V_T]^{-1}\sum_{t=1}^{T} x_t y_t\right) \\
&= [V_T]^{-1}\,\mathrm{Var}\left(\sum_{t=1}^{T} x_t y_t\right)[V_T]^{-1} \\
&= V_T^{-1}\left[\sum_{t=1}^{T} \mathrm{cov}(x_t y_t, x_s y_s)\right] V_T^{-1} \\
&= V_T^{-1}\left[\sum_{s=1}^{T}\sum_{t=1}^{T} x_t\,\mathrm{cov}(y_t, y_s) x_t'\right] V_T^{-1} \\
&= [V_T]^{-1}\left[\sum_{s=1}^{T}\sum_{t=1}^{T} x\left(\tfrac{t}{T}\right)\kappa_T(t,s) x'\left(\tfrac{t}{T}\right)\right][V_T]^{-1} \\
&\sim [V_T]^{-1}\left[\sum_{s>t}^{T} x\left(\tfrac{t}{T}\right)\kappa_T(t,s) x'\left(\tfrac{s}{T}\right) + \sum_{s<t}^{T} x\left(\tfrac{t}{T}\right)\kappa_T(s,t) x'\left(\tfrac{s}{T}\right)\right][V_T]^{-1} \\
&\sim [V_T]^{-1}\sum_{s>t}^{T} x\left(\tfrac{t}{T}\right)\frac{\Gamma(1-d_s-d_t)}{\Gamma(1-d_s)\Gamma(d_s)}(s-t)^{(d_s+d_t-1)} x'\left(\tfrac{s}{T}\right)[V_T]^{-1} \\
&\quad + [V_T]^{-1}\sum_{s>t}^{T} x\left(\tfrac{s}{T}\right)\frac{\Gamma(1-d_s-d_t)}{\Gamma(1-d_s)\Gamma(d_s)}(s-t)^{(d_s+d_t-1)} x'\left(\tfrac{t}{T}\right)[V_T]^{-1} \\
&\sim [V_T]^{-1}\sum_{s>t}^{T}\left[x\left(\tfrac{t}{T}\right)x'\left(\tfrac{s}{T}\right) + x\left(\tfrac{s}{T}\right)x'\left(\tfrac{t}{T}\right)\right] g(\tfrac{s}{T},\tfrac{t}{T}) \\
&\quad \times (s-t)^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}[V_T]^{-1},
\end{aligned}$$

as $T \to \infty$. Therefore the sum approximation for integrate we have

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}_T) \;\sim\; & [V_T]^{-1}\sum_{s>t}^{T} T^2\left[x\left(\tfrac{t}{T}\right)x'\left(\tfrac{s}{T}\right) + x\left(\tfrac{s}{T}\right)x'\left(\tfrac{t}{T}\right)\right] g(\tfrac{s}{T},\tfrac{t}{T}) \\
& \times(\tfrac{s}{T}-\tfrac{t}{T})^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1} T^{d\left(\frac{s}{T}\right)+d\left(\frac{t}{T}\right)-1}[V_T]^{-1}\tfrac{1}{T^2} \\
\sim\; & [V_T]^{-1}\int_0^1\int_0^x T^2\left[x(x)\,x'(y) + x(y)\,x'(x)\right] g(x,y) \\
& \times(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\,dy\,dx\,[V_T]^{-1} \\
\sim\; & \int_0^1\int_0^x G(x,y)(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\,dy\,dx, \qquad (4.4.2)
\end{aligned}$$

where

$$G(x,y) = \left[\int_0^1 x(z)x'(z)\,dz\right]^{-1} \left[x\,(x)\,x'\,(y) + x\,(y)\,x'\,(x)\right] \\ \times \left[\int_0^1 x(z)x'(z)\,dz\right]^{-1} g\,(x,y)\,. \tag{4.4.3}$$

Hence, similarly to the proof of Lemma A.1.2, the asymptotic variance value of $\mathrm{Var}(\widehat{\beta}_T)$ depends only on the evaluation of the double integral (4.4.2) in a neighborhood of $(x,y) = (1,1)$. Consequently, let us define any $\varepsilon > 0$ the set

$$A_T = \{(x,y)|1-\varepsilon \leq x, y \leq 1+\varepsilon, 1/T < x-y, |d(x)-d_1| < \delta, \\ |d(y)-d_1| < \delta, |G(x,y)-G(1,1)| < \delta\},$$

for some $\delta > 0$. This is a nonempty set since $d(\cdot)$ and $G(\cdot)$ are continuous functions in a neighborhood of 1. Let C_T be defined as

$$C_T = T^{1-2d_1}\left(\alpha_1 \log T\right)^{2d_1+1}. \tag{4.4.4}$$

Then,

$$\begin{aligned} \lim_{T\to\infty} C_T \,\mathrm{Var}\left(\widehat{\beta}_T\right) &= \lim_{T\to\infty} C_T \int_0^1 \int_0^x G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\,dy\,dx \\ &= \lim_{T\to\infty} C_T \int_{A_T}\!\!\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\,dy\,dx. \end{aligned}$$

Therefore $1 < (x-y)T$ we have that

$$\lim_{T\to\infty} C_T \int_{A_T}\!\!\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\,dy\,dx \\ \leq \left[G(1,1)+\delta\right] \lim_{T\to\infty} C_T \int_{A^T_{i,j}}\!\!\int \left[(x-y)T\right]^{2\delta+2d_1-1}\,dy\,dx.$$

Therefore, by virtue of Lemma A.1.2, we conclude that

$$\lim_{T\to\infty} C_T \int_0^1 \int_0^x G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\,dy\,dx \leq \left[G(1,1)+\delta\right]\Gamma(2d_1).$$

By an analogous argument, we can also conclude that

$$\lim_{T\to\infty} C_T \int_0^1 \int_0^x G(x,y)\,[(x-y)T]^{d(x)+d(y)-1}\; dy\, dx$$
$$\geq [G_{i,j}(1,1)-\delta]\,\Gamma(2d_1). \tag{4.4.5}$$

Now, since ε and δ can be chosen arbitrarily small, we have that

$$\lim_{T\to\infty} C_T \int_0^1 \int_0^x G(x,y)\,[(x-y)T]^{d(x)+d(y)-1}\; dy\, dx = G(1,1)\Gamma(2d_1).$$

□

Theorem 4.4.2. (General Case *) Assume that the process* $\{Y_{t,T}\}$ *satisfies* (4.2.1). *Then, under Assumptions A3, A8 and A9 the estimator* $\widehat{\beta}_T$ *satisfies*

$$T^{1-2d_0}\,(\log T)^{d_0+\frac{1}{2}}\,\mathrm{Var}\left(\widehat{\beta}_T\right) \to V(u_0),$$

as $T\to\infty$. *Where* $V(u_0)$ *is given by*

$$\begin{cases} G(u_0,u_0)\frac{2^{2d_0}\sqrt{\pi}\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}} & \textit{if} \quad u_0\in(0,1) \\ G(u_0,u_0)\frac{2^{2d_0-1}\sqrt{\pi}\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}} & \textit{if} \quad u_0 = 0,1. \end{cases}$$

Where G *is an* $p\times p$ *matrix with elements given by*

$$G_{i,j}(u_0,u_0) = \left[\int_0^1 x(y)x'(y)\,dy\right]^{-1} x(u_0)x'(u_0)\left[\int_0^1 x(y)x'(y)\,dy\right]^{-1} g(u_0,u_0).$$

Proof. Similarly to the proof of Theorem 2.4.2, the asymptotic value of $\mathrm{Var}(\widehat{\beta}_T)$ depends only on the evaluation of the integral (4.4.2) in a neighborhood of u_0. Consequently, let us define for each element of the matrix $G_{i,j}(x,y)$ and for any $\varepsilon>0$ the set

$$A_T = \{(x,y)|u_0-\varepsilon\leq x,y\leq u_0+\varepsilon, 1/T < x-y, |d''(x)-d''(u_0)|<\delta,$$
$$|d''(y)-d''(u_0)|<\delta, |G(x,y)-G(u_0,u_0)|<\delta\},$$

for some $\delta > 0$. Define $C_T = T^{1-2d_0}\left(\log T\right)^{d_0+\frac{1}{2}}$. Then,

$$\begin{aligned}\lim_{T\to\infty} C_T \operatorname{Var}\left(\widehat{\beta}_T\right) &= \lim_{T\to\infty} C_T \int_0^1 \int_0^x G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\ &= \lim_{T\to\infty} C_T \int_{A_T}\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx.\end{aligned}$$

Therefore $1 < (x-y)T$ we have that

$$\begin{aligned}&\lim_{T\to\infty} C_T \int_{A_T}\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\ &\quad\le \left[G(u_0,u_0)+\delta\right]\lim_{T\to\infty} C_T \int_{A_T}\int \left[(x-y)T\right]^{2d_0+[d''(u_0)-\delta][(x-u_0)^2+(y-u_0)^2]/2-1}\, dy\, dx,\end{aligned}$$

for each $i, j = 1, \ldots p$. Then it follows by Theorem 2.4.2 we have that

$$\begin{aligned}&\lim_{T\to\infty} C_T \int_{A_T}\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\ &\qquad\le \left[G(u_0,u_0)+\delta\right]\frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{\left[\delta - d''(u_0)\right]^{d_0+1/2}}.\end{aligned}$$

By an analogous argument, we can also conclude that

$$\begin{aligned}&\lim_{T\to\infty} C_T \int_{A_T}\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\ &\qquad\ge \left[G(u_0,u_0)-\delta\right]\frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{\left[-\delta - d''(u_0)\right]^{d_0+1/2}}.\end{aligned} \tag{4.4.6}$$

Now, since ε and δ can be chosen arbitrarily small, we have that

$$\begin{aligned}&\lim_{T\to\infty} C_T \int_{A_T}\int G(x,y)\left[(x-y)T\right]^{d(x)+d(y)-1}\, dy\, dx \\ &\qquad= G(u_0,u_0)\frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{\left[-d''(u_0)\right]^{d_0+1/2}}.\end{aligned}$$

A similar argument yields the result for $u_0 = 0, 1$. $\square$

4.5 Normality

In this section we discuss the asymptotic normality of $\widehat{\beta}$ for this we assume that the input noise $\{Z_t\}$ in the generalized Wold expansion is a sequence of independent identically distributed random variables.

Theorem 4.5.1. *Assume that the process $\{Y_{t,T}\}$ satisfies (4.2.2) where $\{Z_t\}$ is a sequence of independent identically distributed random variables. Under the Assumptions A3, A7–A10 the estimator satisfies*

$$T^{1-2d_0}\left(\log T\right)^{2d_0+1}\left(\widehat{\beta}_T-\beta\right)\to N[0,V(u_0)],$$

where $V(u_0)$ is given by

$$V(u_0)=\begin{cases} G(u_0,u_0)\frac{2^{2d_0}\sqrt{\pi}\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}} & \textit{if} \quad u_0\in(0,1) \\ G(u_0,u_0)\frac{2^{2d_0-1}\sqrt{\pi}\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}} & \textit{if} \quad u_0=0,1. \end{cases}$$

Proof. We adapt the proof of Theorem 18.6.5 of Ibragimov and Linnik (1971), as corrected by Hosking (1996). Let α be a fixed vector and define

$$S_T=\sum_{t=1}^{T}\sum_{i=1}^{p}\alpha_i x_i\left(\tfrac{t}{T}\right)Y\left(\tfrac{t}{T}\right)$$

for $i=1,\ldots,p$. Then, we can write

$$S_T=\sum_{k=-\infty}^{T}C_{k,T}Z_k,$$

where the coefficients $C_{k,T}$ are given by

$$C_{k,T}=\sum_{j=\max\{1,k\}}^{T}\sum_{i=1}^{p}\alpha_i x_i\left(\tfrac{j}{T}\right)\sigma\left(\tfrac{j}{T}\right)\psi_{j-k}\left(\tfrac{j}{T}\right).$$

Let $\sigma_T^2=\operatorname{Var}\left(S_T\right)$. As pointed out by Hosking 1996, the key aspect of Ibragimov and Linnik's proof is showing that $C_{k,T}/\sigma_T$ converges to zero uniformly as $T\to\infty$.

In what follows, we prove that this is indeed the case for the class of locally stationary processes. Under the Assumption A10 we have conclude that

$$|C_{k,T}| \leq KT^{d_0}, \tag{4.5.1}$$

for all $k \leq T$. On other hand, we have that $\mathrm{Var}(\alpha'\widehat{\beta}_T) = \frac{K}{T^2}\sigma_T^2$. Hence by (4.4.6)

$$C_T\frac{\sigma_T^2}{T^2} \geq [\alpha' G(u_0,u_0)\alpha - \delta]\,\frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{[-\delta - d''(u_0)]^{d_0+1/2}},$$

for large T, where C_T is defined in the proof of Theorem 4 for any $u_0 \in [0,1]$ and δ arbitrarily small, there exists a constant $K > 0$ such that

$$\frac{1}{\sigma_T} \leq K\sqrt{C_T}T^{-1}. \tag{4.5.2}$$

Combining the equation (4.5.1) and (4.5.2) we conclude that

$$\frac{|C_{k,T}|}{\sigma_T} \leq K\frac{[\log T]^{d_0/2+1/4}}{\sqrt{T}},$$

for $i, j = 1, \ldots, p$ which tends to zero uniformly as $T \to \infty$. □

When the time-varying long-memory parameter $d(\cdot)$ is a linear function the asymptotic normality of $\widehat{\beta}_T$ is similar the cases previous.

4.6 Estimation of the error parameters

In this section we discuss the fitting of a locally model with time-varying spectral density $f_\theta(u,\lambda), \theta \in \Theta \subset \mathbb{R}^p$ to observations $\{Y_{t,T}\}$ in the family of models given by (4.2.1). We shall construct as estimator for the errors by substituting $\widehat{\beta}_T$ for the unknown parameters β. We obtain the parameter estimate by minimization of a generalization of the Whittle function where the usual periodogram is replaced by local periodograms. Let

$$I_N^{x(t/T)\beta}(u,\lambda) := \frac{1}{2\pi H_{2,N}(0)}\left|d_N^{Y-x(t/T)\beta}(u,\lambda)\right|^2,$$

$$\mathcal{L}_T(\theta,\beta) = \frac{1}{4\pi}\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\left\{\log f_\theta(u_j,\lambda) + \frac{I_N^{x(t/T)\beta}(u_j,\lambda)}{f_\theta(u_j,\lambda)}\right\}d\lambda,$$

$$\widehat{\theta}_T := arg\min_{\theta\in\Theta}\mathcal{L}_T(\theta,\beta) \quad and \quad \widetilde{\theta}_T := arg\min_{\theta\in\Theta}\mathcal{L}_T(\theta,\widehat{\beta}_T),$$

where

$$d_N^Y(u,\lambda) = \sum_{s=0}^{N-1} h\left(\tfrac{s}{N}\right) Y_{[uT]-N/2+s+1}e^{-i\lambda s}, \quad H_{k,N}(\lambda) = \sum_{s=0}^{N-1} h\left(\tfrac{s}{N}\right)^k e^{-i\lambda s}.$$

Thus, $I_N(u,\lambda)$ is the periodogram over a segment of length N with midpoint $[uT]$. For study of the asymptotic properties of $\widetilde{\theta}$ we need following regularity conditions.

A11. *The time-varying spectral density is strictly positive and satisfies*

$$f_\theta(u,\lambda) \sim C_f(\theta,u)|\lambda|^{-2\alpha(\theta,u)},$$

as $|\lambda|\to 0$, where $C_f(\theta,u)$ is strictly positive function and $\alpha(\theta,u)\in(0,1)$.

A12. *The data taper $h(u)$ is a positive, bounded function for $u\in[0,1]$ and symmetric around $\frac{1}{2}$ with a bounded derivative.*

A13. *The sample size T and the subdivisions integers N, S and M tend to infinity satisfying $S/N\to 0$, $\sqrt{T}\log^2 N/N\to 0$, $\sqrt{T}/M\to 0$, and $N^3\log^2 N/T^2\to 0$.*

A14. *There exist a constant positive K such that*

$$\left|\sum_{j=0}^{t} h\left(\tfrac{j}{N}\right)\gamma(j-s)\right| < K$$

where $t=1,\ldots,T$ and $s=1,\ldots,N$

Theorem 4.6.1. *Suppose that assumptions A11–A14 holds and in addition that*

$$\left\| X'\left(\tfrac{t}{T}\right)\widehat{\beta}_T - X'\left(\tfrac{t}{T}\right)\beta \right\| = o_p\left(\frac{\sqrt{N}T^{d_0-1/2}}{\log T^{d_0+1/2}}\right)$$

and

$$\left\| \left\{X'\left(\tfrac{t}{T}\right)\widehat{\beta}_T - X'\left(\tfrac{t}{T}\right)\beta\right\} - \left\{X'\left(\tfrac{t-1}{T}\right)\widehat{\beta}_T - X'\left(\tfrac{t-1}{T}\right)\beta\right\} \right\| = o_p\left(\frac{T^{d_0-1/2}}{\sqrt{N}\log T^{d_0+1/2}}\right)$$

uniformly in t. Then

$$\sqrt{T}\left(\widetilde{\theta}_T - \widehat{\theta}_T\right) \to_p 0$$

that is $\widetilde{\theta}_T$ is consistent and has the same asymptotic distribution as $\widehat{\theta}_T$.

The result is proved in the Appendix.

4.7 Numerical and Simulation Studies for a Regression Model

In this section we discuss the application of the previous theoretical results to the analysis of the large sample properties of LSE for a regression model in LSLM process, the model is given by ,

$$Y_{t,T} = \beta_1 x_{t1} + \beta_2 x_{t2} + \varepsilon_{t,T},$$

where $[x_{t1}, x_{t2}]' = \left[\left(\frac{t}{T}\right), \sin\left(\omega\frac{t}{T}\right)\right]'$ is a vector of non-stochastic regressors with $\omega = 4$ and $\beta = (\beta_1, \beta_2)'$ is a vector of unknown regression parameters, and $\varepsilon_{t,T}$ the sequence of errors is a LSLM processes. In this section we will be verifying the asymptotic formula provided by Theorem 4.4.1 and Theorem 4.4.2, comparing the sample variance obtains from several simulations to their theoretical counterparts. The calculations are illustrated with locally stationary fractional noise process with lineal and general case long-memory function. The samples of this LSFM process used in these simulations are generated by mean of the innovation algorithm, see (Brockwell and Davis, 1991, p.172). In this implementation, the variance-covariance matrix of the process, $\kappa_T(s,t)$, is given by Lemma A.1.2 . Given that the calculation of the exact variance of the LSE is a highly demanding task for large sample sizes, we examine other approximate methods. The exact value of the variance of vector parameters is given by

$$\operatorname{Var}(\widehat{\beta}_T) = V_T^{-1}\left[\sum_{s=1}^{T}\sum_{t=1}^{T} x\left(\tfrac{t}{T}\right)\kappa_T(t,s)x'\left(\tfrac{t}{T}\right)\right]V_T^{-1}, \tag{4.7.1}$$

where V_T is give by the equation (4.3.2), and using formula (4.4.2) we may obtain the approximation,

$$\operatorname{Var}(\widehat{\beta}_T) \sim \int_0^1\int_0^x G(x,y)(x-y)^{d(x)+d(y)-1}T^{d(x)+d(y)-1}\,dy\,dx, \tag{4.7.2}$$

with $G(x,y)$ given by

$$\begin{aligned} G(x,y) = {} & \left[\int_0^1 x_t(z)x_t'(z)\,dz\right]^{-1}\left[x\left(x\right)x'\left(y\right)+x\left(y\right)x'\left(x\right)\right] \\ & \times\left[\int_0^1 x_t(z)x_t'(z)\,dz\right]^{-1} g\left(x,y\right). \end{aligned} \tag{4.7.3}$$

For simplicity, this formula will be denoted as *Approximation.* On the other hand we have the asymptotic variance for linear case given by

$$T^{1-2d_1}(\log T)^{2d_1+1}\operatorname{Var}(\widehat{\beta}_T) \to G(1,1)\frac{\Gamma(2d_1)}{[\alpha_1]^{2d_1+1}}, \tag{4.7.4}$$

as $T\to\infty$ and for general case given by

$$T^{1-2d_0}\left(\log T\right)^{d_0+\frac{1}{2}}\operatorname{Var}(\widehat{\beta}_T) \to G(u_0,u_0)\frac{4^{d_0}\sqrt{\pi}\Gamma(d_0)}{[-d''(u_0)]^{d_0+1/2}}, \tag{4.7.5}$$

as $T\to\infty$ for simplicity, this expression will be denoted as *Asymptotic* formula.

4.7.1 Numerical and Simulation Studies for the Linear Case

Consider the following illustrative example consisting of a LSFN process defined by (4.2.3)–(4.2.4) with time-varying long-memory parameter given by

$$d(u) = 0.3 + 0.15\,u, \qquad (4.7.6)$$

for $u \in [0,1]$. The Figure (4.1) shows the evolution of time-varying long-memory parameter.

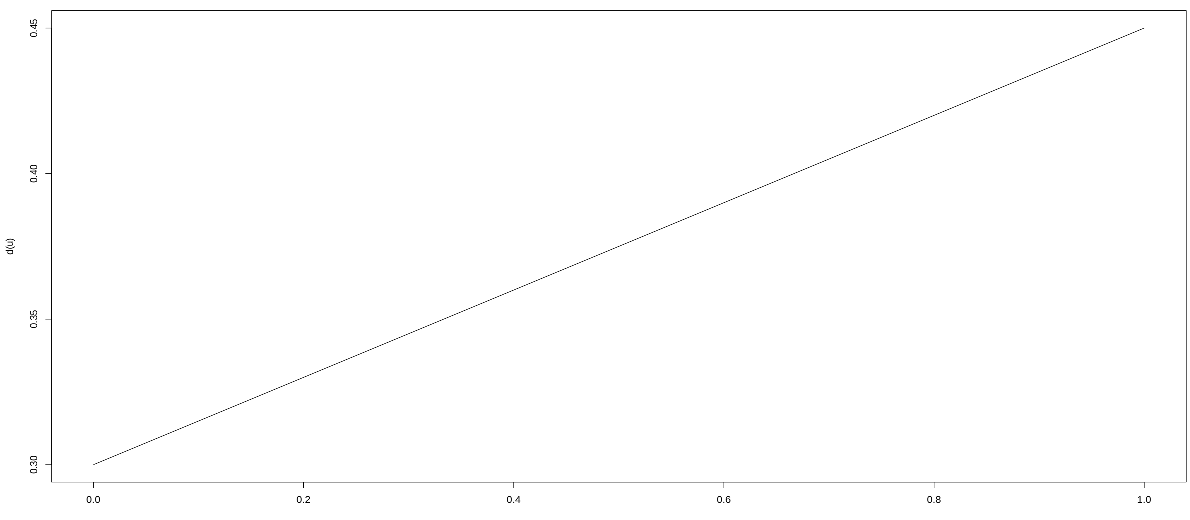

Figure 4.1: Time-Varying long memory parameter

The Table 4.1 reports the values of the variance for three sample sizes $T = 1000$, $T = 2000$ and $T = 4000$ obtained from the following four approaches: exact, sample, Approximation and the Asymptotic formula. The Exact value is given for (4.7.1) and represent the value of the Variance of $\widehat{\beta}_T$. The second row corresponds to be average of the variance $\widehat{\beta}_T$ estimate for least-square estimator over 1000 repetitions. The third row correspond to the Variance the $\widehat{\beta}_T$ obtained from Approximation,

Table 4.1: *Estimation Model LS : Variance of the estimate*

Method	Variance		Covariance
T=1000	Var ($\hat{\beta}_1$)	Var ($\hat{\beta}_2$)	cov ($\hat{\beta}_1, \hat{\beta}_2$)
Exact	1.4518342	0.1641284	-0.1107668
Sample	1.4104739	0.1716930	-0.1166340
Approximation	1.4351997	0.1842634	-0.1147786
Asymptotic	11.787343	5.780556	-8.254538
T=2000	Var ($\hat{\beta}_1$)	Var ($\hat{\beta}_2$)	cov ($\hat{\beta}_1, \hat{\beta}_2$)
Exact	1.3102114	0.1441131	-0.1245641
Sample	1.2197547	0.1530794	-0.1341820
Approximation	1.2920312	0.1598947	-0.1236997
Asymptotic	9.170844	4.497415	-6.422234
T=4000	Var ($\hat{\beta}_1$)	Var ($\hat{\beta}_2$)	cov ($\hat{\beta}_1, \hat{\beta}_2$)
Exact	1.1815534	0.1269124	-0.1332342
Sample	1.1887740	0.1348541	-0.1398417
Approximation	1.1736965	0.1398363	-0.1388113
Asymptotic	7.249257	3.555062	-5.076570

given by formula is (4.7.2). The four row report the approximated values of the variance $\widehat{\beta}_T$ provided by the asymptotic formula (4.7.4). The table 4.1 shows that the parameters variances from the simulations (second row) and Approximation (third row) are relatively close to their theoretical counterparts displayed in the first row. On the other hand, the Asymptotic formula seems to be far off from the exact value for these three sample sizes. Thus, for these sample sizes, the asymptotic formula is not very useful for calculating the variance of $\widehat{\beta}_T$. In order to evaluate the accuracy of

Table 4.2: *Ratio between Approximation formula and Asymptotic Variance*

Sample Size	Variance		Covariance
$\log T = 100$	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Approximation/Asymptotic	0.73141	0.5856191	0.6539831
$\log T = 1200$	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Approximation/Asymptotic	0.9772955	0.9644594	0.9709982
$\log T = 1500$	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Approximation/Asymptotic	1.000292	0.9851984	0.9905511

the asymptotic formula for larger sample sizes, Table 4.2 reports the variance ratios between the Approximation and Asymptotic formula. From this table, the asymptotic formula seems to produce accurate values, but for quite large sample sizes.

4.7.2 Numerical and Simulation Studies for the General Case

To illustrate the asymptotic variance formula given in Theorem (4.4.2), consider the following locally stationary fractional noise process with time-varying long-memory parameter given by

$$d(u) = \; 0.1 + \; u \exp{(-2\,u)}, \tag{4.7.7}$$

for $u \in [0, 1]$. This function, depicted in Figure 2.5, has a maximum value $d_0 = 0.284$ reached at $u_0 = 0.500$.

Table 4.3: *Estimation Model LS : Variance of the estimate*

Method	Variance		Covariance
T=1000	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Exact	0.08455518	0.05036215	0.01626905
Sample	0.09009177	0.05792289	0.01714857
Approximation	0.09137119	0.05013938	0.01569106
Asymptotic	0.04381562	0.20421934	0.09459385
T=2000	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Exact	0.06156288	0.03713725	0.01206201
Sample	0.06529732	0.03982817	0.01236225
Approximation	0.06557527	0.03633117	0.01118837
Asymptotic	0.03012951	0.14043003	0.06504682
T=4000	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Exact	0.044632273	0.027242047	0.008995782
Sample	0.045688292	0.029589303	0.008883724
Approximation	0.047069891	0.026341244	0.008502723
Asymptotic	0.02085443	0.09720000	0.04502278

In this case the long- memory parameter does not belong to the class of polynomials generated by the basis $\{g_j(u) = u^j\}$, so the fourth row corresponds to the approximated values of the variance of $\widehat{\beta}_T$ provided by the asymptotic formula (4.7.5). The Table 4.3 reports Exact Value, Sample, Approximation and Asymptotic Variance formula the regression parameters estimation for LSE methods. Similar to previous case the Table 4.3 shows that the theoretical variance, sample and approximation are very close, however the asymptotic formula seems far from the exact values for these sample sizes. The Table 4.4 reports the ratios between Approximation and the asymptotic formula. Due to the large sample sizes involved in this table, in these experiments we have not calculated the exact variance of $\widehat{\beta}_T$ nor the sample values. From this table, the asymptotic formula seems to produce accurate values, but for quite large sample sizes.

Table 4.4: *Ratio between Approximation formula and Asymptotic Variance*

Sample Size	Variance		Covariance
$\log T = 100$	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Approximation/Asymptotic	1.614918	0.7483252	0.9605854
$\log T = 500$	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Approximation/Asymptotic	1.108186	0.9456312	1.002092
$\log T = 1500$	$\mathrm{Var}(\hat{\beta}_1)$	$\mathrm{Var}(\hat{\beta}_2)$	$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$
Approximation/Asymptotic	1.035321	0.9819726	1.001368

Chapter 5

Application

5.1 The Tree-Ring data

In this chapter we focus our attention to analyze the signification of the vector of parameters using of the asymptotic formula for the variance of $\widehat{\beta}_T$. The Tree-Ring data of BRISTLECONE PINE measurements at NEVADA, from 0AD to 1967. The model for this data is given by

$$Y_{t,T} = X'\left(\tfrac{t}{T}\right)\beta + \varepsilon_{t,T}, \tag{5.1.1}$$

where $X\left(\frac{t}{T}\right) = \left(1, \frac{t}{T}, \sin\left(\omega \frac{t}{T}\right), \cos\left(\omega \frac{t}{T}\right)\right)'$ is a vector regressors, β_j with $j = 1, \ldots, 4$ is the vector of parameters, $\omega = 364$ and $\varepsilon_{t,T}$ the sequence of errors of a LSLM processes.

The data, available at the National Climatic Data Center, are report by V.C. Lamarche and C. Ferguson and displayed in the Figure 5.1, the measurement can be used to indicate the chances of temperature, precipitation, climate and environmental change derived from tree ring measurements.

The least squares fitting assuming uncorrelated errors is shown in Table 5.1. Observe that according to this table all the regression coefficients in model (5.1.1) are

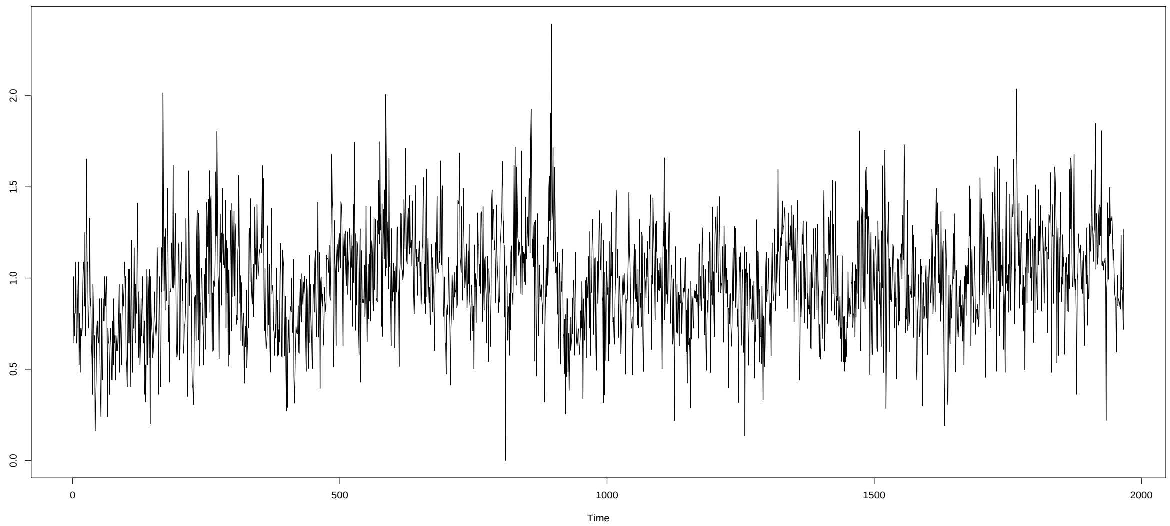

Figure 5.1: *Tree Ring Data*

significant at the 5% level. Hence we estimated the vector of parameters by LSE and

Table 5.1: *Tree-Ring Data: Least Square Fit*

Parameters	Estimates	SD	t value	P-value
β_1	0.895042	0.012681	70.581	0.0000
β_2	0.164763	0.021956	7.504	0.0000
β_3	0.005594	0.008956	0.625	0.532
β_4	-0.013969	0.008970	-1.557	0.120

obtain the residuals $Y_{t,T} - X'\left(\frac{t}{T}\right)\widehat{\beta}$, the sample ACF of the residuals for this model is displayed in Panel (a) of Figure 5.2, and it shows significant autocorrelations at large lags. In addition, the corresponding variance plot is shown in Panel (b) of Figure 5.2. In a variance plot, the broken line represents the expected behavior of the variance of the sample mean of a block of k observation for the short-memory case. On the other hand, the heavy line represents the expected behavior of the variance for a long-memory process. From both panels, this series seems to exhibit long-range dependence behavior.

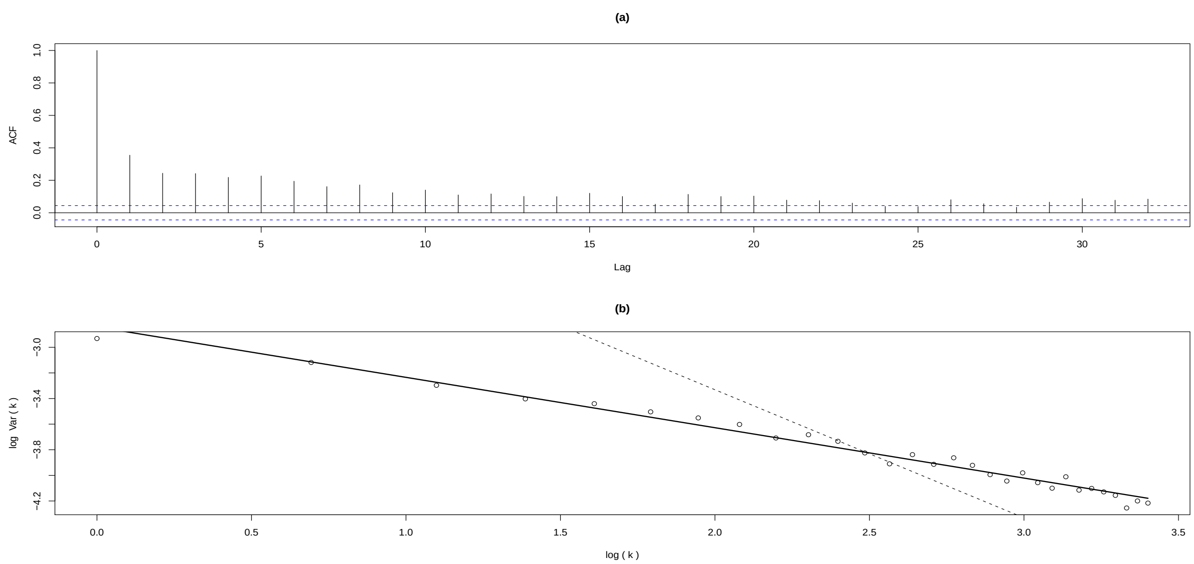

Figure 5.2: *Tree Ring Data. (a) Sample ACF, (b) Variance plot*

Nevertheless, a closer look to the empirical ACF of the data reveals that the degree of long-memory does not seem to be constant over time see Figure 5.3. In fact, the values of the sample ACF of the first 500 observations, see Panel (a), are higher than the corresponding sample ACF values for the other two 500-year periods considered, see panels (b)-(c).

As a result from these two plots, it seems that the disturbances $\varepsilon_{t,T}$ in the linear regression model (5.1.1) may be locally stationary long memory correlation structure and the LSE fitting may not be adequate.

To account for the possible locally stationary long-memory behavior of the errors,

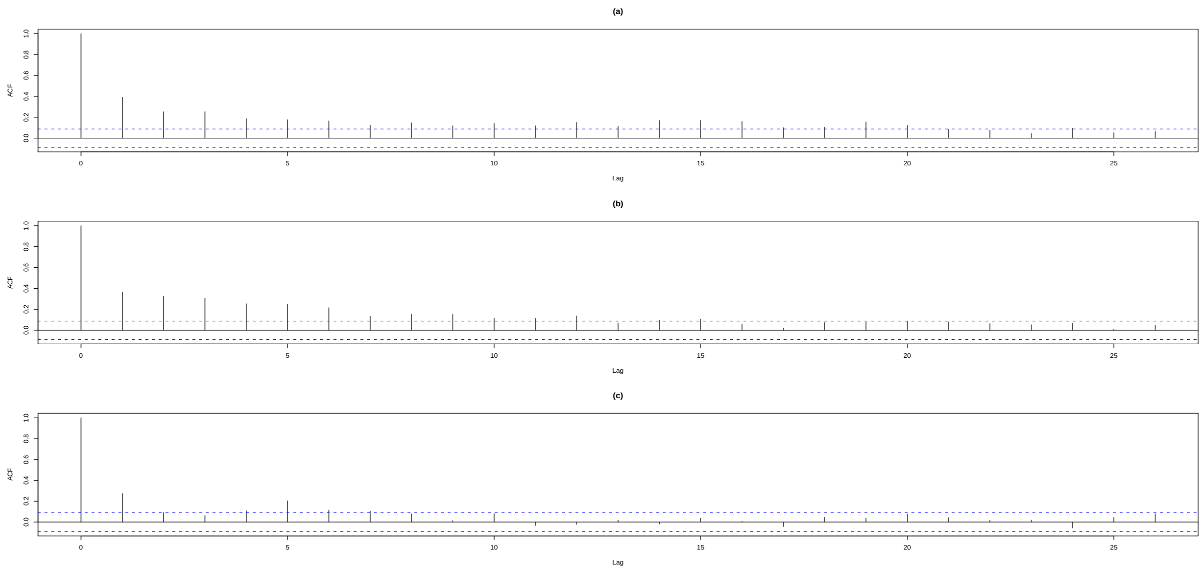

Figure 5.3: *Tree Ring Data. Sample ACF: (a) Observations 1 to 500,(b) Observations 750 to 1250, (c) Observations 1490 to 1967 .*

the following LS-ARFIMA model is proposed for the regression disturbances $\varepsilon_{t,T}$.

$$\Phi\left(\frac{t}{T}\right)\varepsilon_{t,T} = \Theta\left(\frac{t}{T}\right)(1-B)^{-d\left(\frac{t}{T}\right)}\,\sigma\left(\frac{t}{T}\right)Z_t. \tag{5.1.2}$$

The model selected according to the Akaike's information criterion (AIC) is the LS-ARFIMA(0,d,0)-(2,0), the parameter estimates and significance are reported in Table 5.1. Note that according to the fourth column of this table, all the parameters of this model are statistically significant at the 5% level. In this case the time varying long-memory coefficient is a quadratic function i.e. $d(u) = \alpha_0 + \alpha_1 u + \alpha_2 u^2$. Panels (a) and (b) of Figure 5.4 show the evolution of the long-memory parameter, $d(u)$, and the variance scale, $\sigma(\mu)^2$ which, in our case is constant. In both panels the heavy line represents the locally stationary ARFIMA model, the horizontal broken line indicates the stationary ARFIMA model. Figure 5.5 exhibits three panels exploring the structure of the residuals. Panel (a) of this figure displays the residuals from the fitted

Table 5.2: *Model Estimation : Tree-Ring Data*

Parameters	Estimates	SD	t value
α_0	0.1987363	0.0527406	3.7682
α_1	0.4987270	0.2435983	2.0473
α_2	-0.6299850	0.2358630	-2.6710
β_0	0.2555662	0.0040746	62.7216

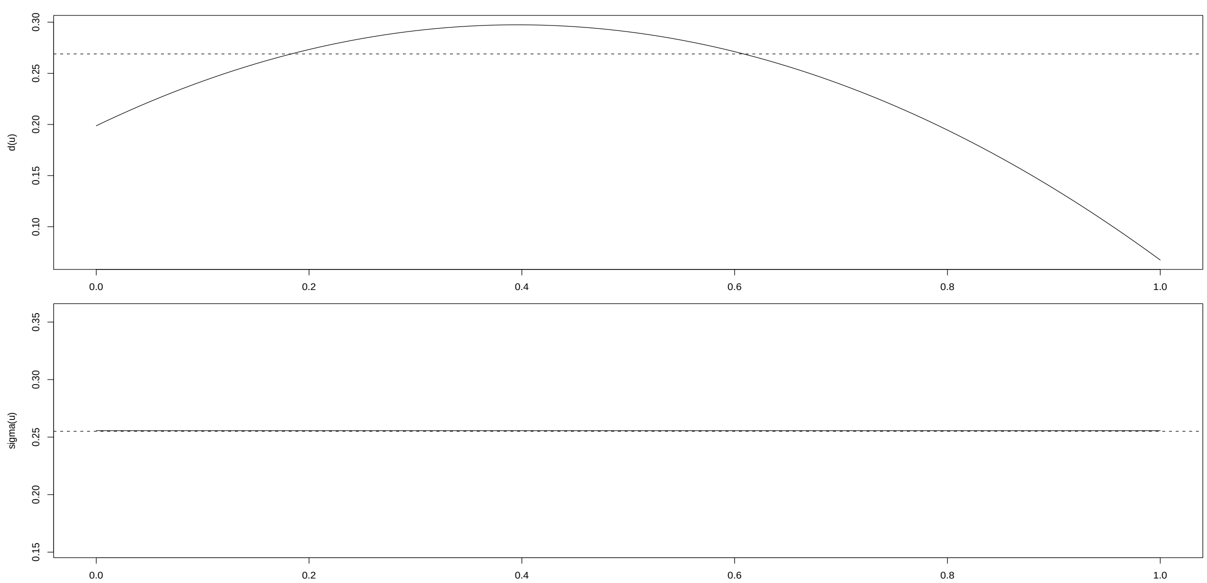

Figure 5.4: *Tree Ring Data. (a) Estimates of the long-memory parameter. (b) Estimates the noise variance.*

LS-ARFIMA model. Panel (b) shows the sample ACF, and Panel (c) exhibits the Ljung-Box whiteness tests. From the figure we can see no significant autocorrelations, this conclusion is formally supported by the Ljung-Box tests where we consideration $K = 16$ windows, see Panel (c) which indicates that the white noise null hypothesis is not rejected for all the lags considered in this case $Lag = 30$, at the 5% level of significance. Now, to analyze the significance the vector parameter we'll use the formula

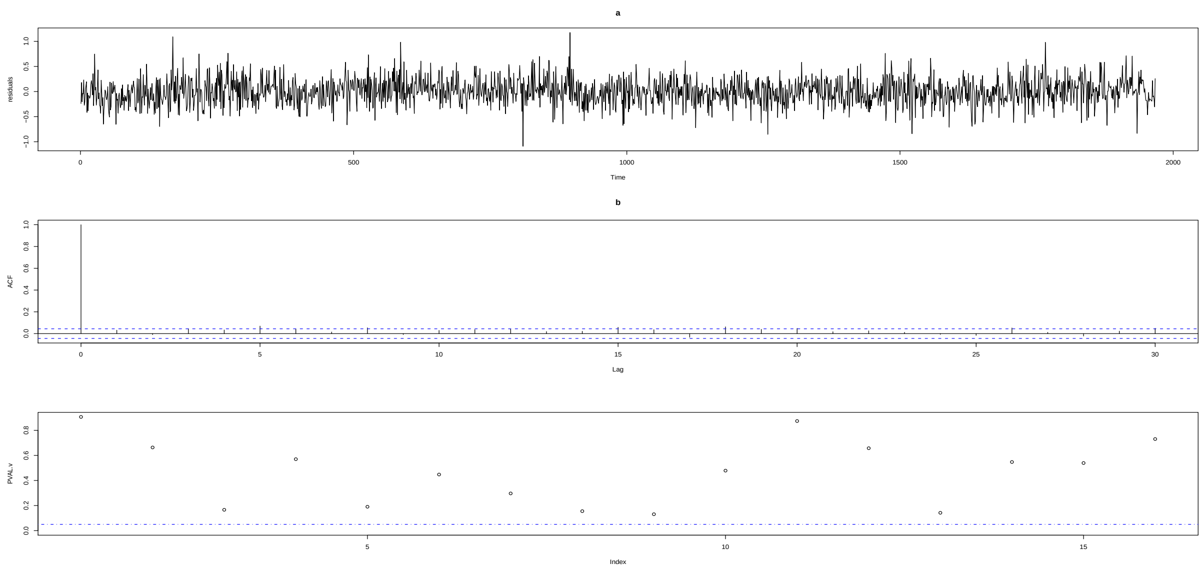

Figure 5.5: *Tree Ring Data: Residual analysis (a) Residuals from the fitted model, (b) Sample ACF, (c) Ljung-Box tests.*

of the asymptotic variance the $\widehat{\beta}_T$ for the general case the time varying long-memory parameter given by Theorem 4.4.2. Using the estimate of Whittle we have that

$$\widehat{d}(u) = 0.1987363 + 0.4987270\, u - 0.6299850\, u^2,$$

and test of significance approximate for $\widehat{\beta}_T$ is given the $t-$ test

$$t_{c,j} = \frac{\widehat{\beta}_{T,j}}{\sqrt{\operatorname{Var}\left(\widehat{\beta}_{T,j}\right)}}$$

with $j = 1, 2, 3$. The significance of the parameters and standard deviation are reported in Table 5.2. Note that according to the fourth column of this table, the

Table 5.3: *Estimation Model: Mean Sample*

Parameters	Estimates	SD	t value
β_0	0.895042441	0.27411285	3.2652334
β_1	0.164763042	0.30873077	0.5336787
β_2	0.005593744	0.03345162	0.1672189
β_3	-0.013968661	0.00684155	-2.0417380

first and fourth parameter that correspond to the media and to the component $cos(\cdot)$ of a harmonic function respectively, are significant, however the other parameters are not statistically significant at the 5% level.

Chapter 6

Conclusions and further work.

6.1 Concluding Remarks

In this work we have established some asymptotic statistical properties to the vector of parameter of a regression model with errors belonging to the class of LSLM process. Is important to remember that these models do not have a fixed structure in time, unless these processes show a time-varying second-order structure. As in Silverman (1957) definition, each model on non-stationary covariance has to define explicitly its departure from stationarity. However, from a statistical viewpoint, many questions remain. For example, with this lack of an invariant second-order structure, how can we estimate the time-varying parameters with a high accuracy? so a serious problem here is that we cannot build an asymptotic theory for the estimation of time varying parameters. Consequently, the standard statistical properties like consistency, efficiency or central limit theorems cannot be use to measure and compare the quality of different estimators. As we mentioned before in the Chapter 1 to overcome this problem, Dahlhaus introduced a concept of locally stationary.

In this context we have investigated the asymptotic of the sample mean of a class of LSLM process with a general specification for the time-varying long-memory parameter. As evidenced by the Theorems 2.4.1–2.4.3, the asymptotic behavior of the variance of the sample mean of a LSLM process is more complex than its stationary long memory counterpart.

The statistical properties for the least square estimator LSE for a vector of parameter was analyzed in Chapter 4. The Theorem (4.3.1) study the consistency of LSE, while the Theorem (4.4.1) and Theorem (4.4.2) give us an explicit formula for the asymptotic variance of the LSE.

Finally, a central limit theorem is established in Theorem (4.5), where we assume that input noise in the generalized wold expansion is a sequence of independent identically distributed random variables.

6.1.1 Further Research

Further work in this framework is the study of asymptotic efficiency for the LSE. Several authors have studied the problem of the asymptotic efficiency of the LSE $\overline{Y}_t$ in ARIMA process, relative to the BLUE see for example Grenander and Rosenblatt (1954) and Grenander and Rosenblatt (1957) they considered a spectral density $f(\lambda)$ piecewise continuous, with no discontinuities at $\lambda = 0$ and $0 < f(\lambda) < \infty$, then $\overline{Y_t}$ is asymptotic efficient. Adenstedt (1974) established certain criteria for asymptotic efficiency for the sample mean for a spectral density of the form $\lambda^{-2d}L(\lambda)$ as $\lambda \to 0$, where $L(\lambda)$ is a slowly varying function at the origin with $0 < L(0) < \infty$. One possible approach to this issue is given by Samarov and Taqqu (1988), they have obtained results for the efficiency of LSE when the time series is fractional ARIMA(0,d,0) for all $d < \frac{1}{2}$. If the time series has a spectral density f, the efficiency of the LSE estimator $\overline{Y}_t$ is defined as

$$e(n, f) = \frac{\mathrm{Var}(\widehat{m}_f)}{\mathrm{Var}(\overline{Y}_t)}, \tag{6.1.1}$$

where $\widehat{m}_f$ is the BLUE for μ. The asymptotic efficiency is

$$e(\infty, f) = \lim_{n\to\infty} e(n, f).$$

Therefore, a pending job is to analyze the asymptotic behavior of the minimum variance for the sample mean of a class locally stationary fractional noise process, where the time-varying long-memory parameter through of a lineal process or a more general behavior. We should find the asymptotic efficiency for spectral density $f(u, \lambda)$ given by

$$f_\theta(u, \lambda) = \frac{\sigma^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-d_\theta(u)}.$$

Appendix A

Technical Appendix

A.1 Supplementary Material for Chapter 2

Lemma A.1.1. *The variance-covariance matrix* $[\kappa_T(s,t)]_{s,t=1,\ldots,T}$ *of the process* (2.2.4) *is given by*

$$\kappa_T(s,t) = \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\frac{\Gamma\left[1-d\left(\frac{s}{T}\right)-d\left(\frac{t}{T}\right)\right]\Gamma\left[s-t+d\left(\frac{s}{T}\right)\right]}{\Gamma\left[1-d\left(\frac{s}{T}\right)\right]\Gamma\left[d\left(\frac{s}{T}\right)\right]\Gamma\left[s-t+1-d\left(\frac{t}{T}\right)\right]},$$

for $s,t = 1,\ldots,T,\ s \geq t$.

Proof. By definition, the elements $\kappa_T(s,t)$ of the variance-covariance matrix of the process (2.2.4) are given by

$$\begin{aligned}
\kappa_T(s,t) &= \mathrm{E}[Y_{s,T}Y_{t,T}] = \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\eta_{s-t+j}\left(\tfrac{s}{T}\right)\eta_j\left(\tfrac{t}{T}\right) \\
&= \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\sum_{j=0}^{\infty}\frac{\Gamma\left[s-t+j+d\left(\frac{s}{T}\right)\right]\Gamma\left[j+d\left(\frac{t}{T}\right)\right]}{\Gamma\left[s-t+j+1\right]\Gamma\left[j+1\right]} \\
&= \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\frac{\Gamma\left[s-t+d\left(\frac{s}{T}\right)\right]}{\Gamma\left[d\left(\frac{s}{T}\right)\right]\Gamma\left[s-t+1\right]} \\
&\quad\times\sum_{j=0}^{\infty}\frac{\Gamma\left[s-t+j+d\left(\frac{s}{T}\right)\right]\Gamma\left[j+d\left(\frac{t}{T}\right)\right]\Gamma\left[s-t+1\right]}{\Gamma\left[s-t+d\left(\frac{s}{T}\right)\right]\Gamma\left[d\left(\frac{t}{T}\right)\right]\Gamma\left[s-t+j+1\right]\Gamma\left[j+1\right]}.
\end{aligned}$$

Therefore, by an application of the hypergeometric function $F(a, b; c, z)$ with $z = 1$ we get

$$\begin{aligned}\kappa_T(s,t) &= \sigma\left(\tfrac{s}{T}\right)\sigma\left(\tfrac{t}{T}\right)\frac{\Gamma\left[s-t+d\left(\frac{s}{T}\right)\right]}{\Gamma\left[d\left(\frac{s}{T}\right)\right]\Gamma\left[s-t+1\right]} \\ &\quad \times F(s-t+d\left(\tfrac{s}{T}\right), d\left(\tfrac{t}{T}\right); s-t+1, 1).\end{aligned}$$

Now, by Gradshteyn and Ryzhik (2000) [Eq. 9.122] the result is obtained. □

Lemma A.1.2. *Let $d(u)$ be a function lineal and define the double integral*

$$I_T = \int_0^1 \int_0^x \tilde{g}_{i,j}(x,y)(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\, dy\, dx,$$

where $\tilde{g}(x,y)$ is a function defined in (3.4.2). Then,

$$T^{1-2d_1}\left(\alpha_1 \log T\right)^{2d_1+1} I_T \to \frac{x(1)^2}{\left[\int_0^1 x^2(y)\, dy\right]} g(1,1)\Gamma(2d_1),$$

as $T \to \infty$

Proof. By means of the variable transformation $u = x + y$ and $v = x - y$, we can write

$$d(x) + d(y) = 2\alpha_0 + \alpha_1 u,$$

where $\alpha_1 > 0$. Thus,

$$\begin{aligned}I_T &\sim \int_0^1 \int_0^u \tilde{g}(u,v) v^{2\alpha_0+\alpha_1 u-1} T^{2\alpha_0+\alpha_1 u-1}\, dv\, du \\ &+ \int_1^2 \int_0^{2-u} \tilde{g}(u,v) v^{2\alpha_0+\alpha_1 u-1} T^{2\alpha_0+\alpha_1 u-1}\, dv\, du,\end{aligned}$$

Note that $\tilde{g}(u,v)$ reaches its maximum value when v is close to zero. Therefore, analyzing I_1 and an application Lemma A.1.4 we have

$$\begin{aligned}
T^{1-2\alpha_0-\alpha_1}&(\alpha_1 \log T)\, I_1 \\
&\sim \int_0^1 \int_0^u T^{\alpha_1(u-1)} (\alpha_1 \log T)\, \tilde{g}(u,v) v^{2\alpha_0+\alpha_1 u-1}\, dv\, du \\
&\sim \int_0^1 T^{\alpha_1(u-1)} (\alpha_1 \log T) \int_0^u \tilde{g}(u,v) v^{2\alpha_0+\alpha_1 u-1}\, dv\, du \\
&\sim \alpha_1 \log T \int_0^1 T^{\alpha_1(u-1)} \tilde{g}(u,0) \frac{u^{2\alpha_0+\alpha_1 u}}{2\alpha_0+\alpha_1 u}\, du.
\end{aligned}$$

Now, an application of Lemma A.1.5 yields,

$$\begin{aligned}
T^{1-2\alpha_0-\alpha_1}&(\alpha_1 \log T)\, I_1 \\
&\sim \frac{\alpha_1 \log T}{2\alpha_0+\alpha_1} \int_0^1 \tilde{g}(u,0) u^{2\alpha_0+\alpha_1 u} T^{\alpha_1(u-1)}\, du \\
&\sim \frac{\tilde{g}(1,0)}{2\alpha_0+\alpha_1} \\
&\sim \frac{x\left(\frac{1}{2}\right)^2}{\left[\int_0^1 x^2(y)\, dy\right]^2} \frac{g\left(\frac{1}{2},\frac{1}{2}\right)}{(2\alpha_0+\alpha_1)},
\end{aligned}$$

In summary we have,

$$(\alpha_1 \log T) T^{1-2\alpha_0-\alpha_1} I_1 \to g\left(\tfrac{1}{2},\tfrac{1}{2}\right) \frac{x\left(\frac{1}{2}\right)^2}{\left[\int_0^1 x^2(y)\, dy\right]^2 (2\alpha_0+\alpha_1)}, \tag{A.1.1}$$

as $T \to \infty$. Considering the integrate I_2, and an application of Lemma A.1.4 yields, we have

$$
\begin{aligned}
&[\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1} T^{1-2\alpha_0-2\alpha_1} I_2 \\
&\quad\sim \int_1^2 \int_0^{2-u} T^{\alpha_1 u-2\alpha_1} [\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1} \tilde{g}(u,v) v^{2\alpha_0+\alpha_1 u-1}\, dv\, du \\
&\quad\sim \int_1^2 T^{\alpha_1 u-2\alpha_1} [\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1} \int_0^{2-u} \tilde{g}(u,v) v^{2\alpha_0+\alpha_1 u-1}\, dv\, du \\
&\quad\sim \int_1^2 T^{\alpha_1 u-2\alpha_1} [\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1} \tilde{g}(u,0) \frac{(2-u)^{2\alpha_0+\alpha_1 u}}{2\alpha_0+\alpha_1 u}\, du \\
&\quad\sim \frac{[\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1}}{2\,(\alpha_0+\alpha_1)} \int_1^2 \tilde{g}(u,0) T^{-\alpha_1(2-u)} (2-u)^{2\alpha_0+\alpha_1 u}\, du \\
&\quad\sim \frac{[\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1}}{2\,(\alpha_0+\alpha_1)} \int_0^1 \tilde{g}(x+1,0) T^{-\alpha_1(1-x)} (1-x)^{2\alpha_0+\alpha_1(x+1)}\, dx \\
&\quad\sim \frac{[\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1}}{2\,(\alpha_0+\alpha_1)} \int_0^1 \tilde{g}(x+1,0) T^{-\alpha_1(1-x)} (1-x)^{2\alpha_0+2\alpha_1}\, dx.
\end{aligned}
$$

An application of Lemma A.1.6 yields

$$
\begin{aligned}
&[\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1} T^{1-2\alpha_0-2\alpha_1} I_2 \\
&\qquad\sim \frac{[\alpha_1 \log T]^{1+2\alpha_0+2\alpha_1}}{2(\alpha_0+\alpha_1)} \\
&\qquad\qquad \times \int_0^1 \tilde{g}(x+1,0) T^{-\alpha_1(1-x)} (1-x)^{2\alpha_0+2\alpha_1}\, dx \\
&\qquad\sim \frac{\tilde{g}(2,0)\Gamma(1+2\alpha_0+2\alpha_1)}{2(\alpha_0+\alpha_1)}.
\end{aligned}
$$

On the other hand,

$$
[\alpha_1 \log T]^{2\alpha_0+2\alpha_1+1} T^{1-2\alpha_0-2\alpha_1} I_2 \to \frac{x(1)^2 g(1,1)\Gamma(2\alpha_0+2\alpha_1)}{\left[\int_0^1 x^2(y)\, dy\right]^2}, \tag{A.1.2}
$$

as $T \to \infty$.Therefore of (A.1.1) and (A.1.2) be conclude

$$\begin{aligned}
&\lim_{T\to\infty} T^{1-2d_1}(\alpha_1 \log T)^{2d_1+1} I_T \\
&= \lim_{T\to\infty}\left[\frac{(\alpha_1 \log T)^{2d_1}}{T^{\alpha_1}} T^{1-2\alpha_0-\alpha_1}(\alpha_1 \log T) I_1 + T^{1-2d_1}(\alpha_1 \log T)^{2d_1+1} I_2\right] \\
&= \lim_{T\to\infty}\left[\frac{x\left(\frac{1}{2}\right)^2 g\left(\frac{1}{2},\frac{1}{2}\right)}{2\alpha_0+\alpha_1}\frac{(\alpha_1 \log T)^{2d_1+1}}{T^{\alpha_1}} + \frac{x(1)^2 g(1,1)\Gamma(2d_1)}{\left[\int_0^1 x^2(y)\,dy\right]^2}\right].
\end{aligned}$$

The first term on the equation converges to zero when $T \to \infty$ the result is proven. $\square$

Lemma A.1.3. *Let $d = \alpha_0 - \alpha_1 x$ with $\alpha_1 > 0$ and define the double integral*

$$I_T = \int_0^1 \int_0^x \tilde{g}(x,y)(x-y)^{d(x)+d(y)-1} T^{d(x)+d(y)-1}\,dy\,dx,$$

where $\tilde{g}(x,y)$ is a function defined in (3.4.2). *Then,*

$$T^{1-2\alpha_0}\left(\alpha_1 \log T\right)^{2\alpha_0+1} I_T \to \frac{x(0)^2}{\left[\int_0^1 x^2(y)\,dy\right]} g(0,0)\Gamma(2\alpha_0),$$

as $T \to \infty$

Proof. Now d is a lineal function and has negative slope, by means of the variable transformation $u = x + y$ and $v = x - y$ we can write

$$\begin{aligned}
I_T \quad &\sim \int_0^1 \int_0^u \tilde{g}(u,v) v^{2\alpha_0-\alpha_1 u-1} T^{2\alpha_0-\alpha_1 u-1}\,dv\,du \\
&+ \int_1^2 \int_0^{2-u} \tilde{g}(u,v) v^{2\alpha_0-\alpha_1 u-1} T^{2\alpha_0-\alpha_1 u-1}\,dv\,du.
\end{aligned}$$

A utilization the Lemma A.1.4 we can approximation the integrate I_1 by

$$\begin{aligned}
&[\alpha_1 \log T]^{2\alpha_0+1} T^{1-2\alpha_0} I_1 \\
&\quad \sim \int_0^1 \int_0^u [\alpha_1 \log T]^{2\alpha_0+1} T^{-\alpha_1 u} \tilde{g}(u,v) v^{2\alpha_0-\alpha_1 u-1}\, dv\, du \\
&\quad \sim \int_0^1 T^{-\alpha_1 u} [\alpha_1 \log T]^{2\alpha_0+1} \int_0^u \tilde{g}(u,v) v^{2\alpha_0-\alpha_1 u-1}\, dv\, du \\
&\quad \sim \int_0^1 [\alpha_1 \log T]^{2\alpha_0+1} T^{-\alpha_1 u} \tilde{g}(u,0) \frac{u^{2\alpha_0-\alpha_1 u}}{2\alpha_0-\alpha_1 u}\, du \\
&\quad \sim \frac{[\alpha_1 \log T]^{2\alpha_0+1}}{2\alpha_0} \int_0^1 \tilde{g}(u,0) u^{2\alpha_0} T^{-\alpha_1 u}\, du.
\end{aligned}$$

Now an application Lemma A.1.7 yields, we have

$$\begin{aligned}
&[\alpha_1 \log T]^{2\alpha_0+1} T^{1-2\alpha_0} I_1 \\
&\qquad \sim \frac{1}{2\alpha_0} (\alpha_1 \log T)^{2\alpha_0+1} \int_0^1 \tilde{g}(u,0) T^{-\alpha_1 u} u^{2\alpha_0}\, du \\
&\qquad \sim \frac{\tilde{g}(0,0)\Gamma(1+2\alpha_0)}{2\alpha_0}.
\end{aligned}$$

Therefore the approximation for I_1 is give by

$$(\alpha_1 \log T)^{2\alpha_0+1} T^{1-2\alpha_0} I_1 \quad \to \quad \frac{x^2(0) g(0,0) \Gamma(2\alpha_0)}{\left[\int_0^1 x^2(y)\, dy\right]^2},$$

as $T \to \infty$. Now we considers the approximation for I_2

$$
\begin{aligned}
&(\alpha_1 \log T)\, T^{1-2\alpha_0+\alpha_1} I_2 \\
&\qquad \sim \int_1^2 \int_0^{2-u} (\alpha_1 \log T)\, T^{\alpha_1(1-u)} \tilde{g}(u,v) v^{2\alpha_0-\alpha_1 u-1}\, dv\, du \\
&\qquad \sim \int_1^2 (\alpha_1 \log T)\, T^{\alpha_1(1-u)} \int_0^{2-u} \tilde{g}(u,v) v^{2\alpha_0-\alpha_1 u-1}\, dv\, du \\
&\qquad \sim \int_1^2 (\alpha_1 \log T)\, T^{\alpha_1(1-u)} \tilde{g}(u,0) \frac{(2-u)^{2\alpha_0-\alpha_1 u}}{2\alpha_0-\alpha_1 u}\, du \\
&\qquad \sim (\alpha_1 \log T) \int_0^1 T^{-\alpha_1 x} \tilde{g}(x+1,0) \frac{(1-x)^{2\alpha_0-\alpha_1(x+1)}}{2\alpha_0-\alpha_1(x+1)}\, dx \\
&\qquad \sim (\alpha_1 \log T) \int_0^1 T^{-\alpha_1(1-y)} \tilde{g}(2-y,0) \frac{y^{2\alpha_0-\alpha_1(2-y)}}{2\alpha_0-\alpha_1(2-y)}\, dy \\
&\qquad \sim \frac{(\alpha_1 \log T)}{2(\alpha_0-\alpha_1)} \int_0^1 T^{-\alpha_1(1-y)} \tilde{g}(2-y,0) y^{2\alpha_0-\alpha_1(2-y)}\, dy \\
&\qquad \sim \frac{(\alpha_1 \log T)}{2(\alpha_0-\alpha_1)} \int_0^1 \tilde{g}(2-y,0) T^{-\alpha_1(1-y)} y^{2\alpha_0-\alpha_1}\, dy.
\end{aligned}
$$

An application of Lemma A.1.5 replacing we have

$$
\begin{aligned}
&(\alpha_1 \log T) T^{1-2\alpha_0+\alpha_1} I_2 \\
&\qquad \sim \frac{\alpha_1 \log T}{2(\alpha_0-\alpha_1)} \int_0^1 \tilde{g}(2-y,0) T^{\alpha_1(y-1)} y^{2\alpha_0-\alpha_1}\, dy \\
&\qquad \sim \frac{\tilde{g}(1,0)}{2(\alpha_0-\alpha_1)}.
\end{aligned}
$$

Thus the approximation for I_2 is given by

$$
(\alpha_1 \log T) T^{1-2\alpha_0+\alpha_1} I_2 \to g\left(\tfrac{1}{2}, \tfrac{1}{2}\right) \frac{x\left(\frac{1}{2}\right)^2}{\left[\int_0^1 x^2(y)\, dy\right]^2 2(\alpha_0-\alpha_1)},
$$

Hence the approximation of the two integrates are given by

$$
\begin{aligned}
&\lim_{T\to\infty} T^{1-2d_0}(\alpha_1 \log T)^{2d_0+1} I_T \\
= \quad &\lim_{T\to\infty} \left[T^{1-2d_0}(\alpha_1 \log T)^{2d_0+1} I_1 + \tfrac{(\alpha_1 \log T)^{2d_0}}{T^{\alpha_1}} T^{1-2d_0+\alpha_1}(\alpha_1 \log T) I_2\right] \\
= \quad &\lim_{T\to\infty} \left[\frac{x(0)^2 g(0,0)\Gamma(2d_0)}{\left[\int_0^1 x^2(y)\,dy\right]^2} + \frac{x\,(1/2)^2\, g\,(1/2,1/2)}{2(\alpha_0-\alpha_1)} \frac{(\alpha_1 \log T)^{2d_0}}{T^{\alpha_1}}\right].
\end{aligned}
$$

The second term of the equation converges to zero when $T \to \infty$ the result is proven.

□

Lemma A.1.4. *Let $g(x)$ be a $\mathcal{C}^1(\mathbb{R})$ function such that $g(0) \neq 0$ and let $h(u)$ be a continuous function. Then*

$$
\lim_{h(u)\to 0} \frac{1}{[h(u)]^{2\alpha+\beta u}} \int_0^{h(u)} g(x) x^{2\alpha+\beta u-1}\, dx = \frac{g(0)}{2\alpha+\beta u}.
$$

Proof. Since the function $g \in \mathcal{C}^1(\mathbb{R})$, then by Taylor's Theorem we can write

$$
g(x) = g(0) + g'(x)x.
$$

Hence, we can write

$$
\begin{aligned}
&\frac{1}{[h(u)]^{2\alpha+\beta u}} \int_0^{h(u)} g(x) x^{2\alpha+\beta u-1} \mathrm{d}x \\
&\qquad = \frac{g(0)}{[h(u)]^{2\alpha+\beta u}} \int_0^{h(u)} x^{2\alpha+\beta u-1} \mathrm{d}x \\
&\qquad\qquad + \frac{1}{[h(u)]^{2\alpha+\beta u}} \int_0^{h(u)} g'(x) x^{2\alpha+\beta u} \mathrm{d}x \\
&\qquad = \frac{g(0)}{2\alpha+\beta u} + \frac{1}{[h(u)]^{2\alpha+\beta u}} \int_0^{h(u)} g'(x) x^{2\alpha+\beta u} \mathrm{d}x. \qquad \text{(A.1.3)}
\end{aligned}
$$

The second term of (A.1.3) is zero, in effect

$$\begin{aligned}\frac{1}{[h(u)]^{2\alpha+\beta u}} &\int_0^{h(u)} g'(x)x^{2\alpha+\beta u}\mathrm{d}x \\ &\leq \frac{K}{[h(u)]^{2\alpha+\beta u}}\int_0^{h(u)} x^{2\alpha+\beta u}\mathrm{d}x \\ &\leq \frac{K}{[h(u)]^{2\alpha+\beta u}}\frac{[h(u)]^{2\alpha+\beta u+1}}{2\alpha+\beta u+1} \\ &\leq \frac{Kh(u)}{2\alpha+\beta u+1} \to 0,\end{aligned}$$

as $h(u) \to 0$. □

Lemma A.1.5. *Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1(\mathbb{R})$ function such that $g(1) \neq 0$ and define the integral I_n as*

$$I_n = [b \log n] \int_0^1 g(x)x^{2a+b}n^{b(x-1)}\, dx.$$

Then $I_n \to g(1)$, as $n \to \infty$.

Proof. Since the function $\mathcal{C}^1(\mathbb{R})$, then by Taylor's theorem we can write

$$g(x) = g(1) + g'(\xi_x)(x-1), \tag{A.1.4}$$

for some $0 \leq \xi_x \leq x$, for positive x. Hence integrating both sides of the equation (A.1.4) can be written as

$$\begin{aligned}I_n = g(1)b\log n &\int_0^1 x^{2a+b}n^{b(x-1)}\, dx \\ &+ b\log n \int_0^1 g'(\xi_x)(x-1)x^{2a+b}n^{b(x-1)}\, dx.\end{aligned} \tag{A.1.5}$$

The first integrate on the right can be calculated

$$\begin{aligned}g(1)b\log n &\int_0^1 x^{2a+b}n^{-b(1-x)}\, dx \\ &= g(1)\int_0^{b\log n}\left[1-\frac{t}{b\log n}\right]^{2a+b}\exp(-t)\, dt \\ &\sim g(1)\int_0^{\infty}\exp(-t)\, dt = g(1),\end{aligned}$$

as $n \to \infty$. In other words the second integrate of (A.1.5)is zero as $n \to \infty$ in effect

$$\begin{aligned}
&\left| b \log n \int_0^1 g'(\xi_x)(x-1)x^{2a+b} n^{b(x-1)}\, dx \right| \\
&\qquad \le Kb \log n \int_0^1 (x-1)x^{2a+b} n^{b(x-1)}\, dx \\
&\qquad \sim Kb \log n \int_0^1 (x-1) n^{-b(1-x)}\, dx \\
&\qquad \sim Kb \log n \int_0^1 y \exp(-yb \log n)\, dy \\
&\qquad \sim \frac{K}{b \log n} \int_0^{b \log n} t \exp(-t)\, dt \to 0,
\end{aligned}$$

as $n \to \infty$. $\square$

Lemma A.1.6. *Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1(\mathbb{R})$ function such that $g(1) \neq 0$ and define the integrate I_n as*

$$I_n = [b \log n]^{1+2a+2b} \int_0^1 g(x)(1-x)^{2a+2b} n^{b(x-1)}\, dx$$

Then $I_n \to g(1)\Gamma[1+2a+2b]$, as $n \to \infty$.

Proof. Since the function $\mathcal{C}^1(\mathbb{R})$, then by Taylor's theorem we can write

$$g(x) = g(1) + g'(\xi_x)(x-1), \tag{A.1.6}$$

for some $0 \le \xi_x \le x$, for positive x. Hence integrating both sides of the equation (A.1.6) can be written as

$$\begin{aligned}
I_n = {} & g(1)\, [b \log n]^{1+2a+2b} \int_0^1 (1-x)^{2a+2b} n^{b(x-1)}\, dx \\
& + [b \log n]^{1+2a+2b} \int_0^1 g'(\xi_x)(x-1)(1-x)^{2a+2b} n^{b(x-1)}\, dx. \qquad \text{(A.1.7)}
\end{aligned}$$

The first integrate in the expression on the right above is calculated as follows,

$$
\begin{aligned}
&g(1)\,[b\log n]^{1+2a+2b}\int_0^1 (1-x)^{2a+2b} n^{b(x-1)}\,dx \\
&\quad= g(1)\,[b\log n]^{1+2a+2b}\int_0^1 y^{2a+2b}\exp\left(-yb\log n\right)dy \\
&\quad= g(1)\,[b\log n]^{2a+2b}\int_0^{b\log n}\left[\frac{t}{b\log n}\right]^{2a+2b}\exp\left(-t\right)dt \\
&\quad= g(1)\int_0^{b\log n} t^{2a+2b}\exp\left(-t\right)dt \to g(1)\Gamma(2a+2b+1),
\end{aligned}
$$

as $n \to \infty$. On other hand, the second integrate in (A.1.7) converge to zero since

$$
\begin{aligned}
&\left|[b\log n]^{1+2a+2b}\int_0^1 g'(\xi_x)(x-1)(1-x)^{2a+2b} n^{b(x-1)}\,dx\right| \\
&\quad\leq K\,[b\log n]^{1+2a+2b}\int_0^1 (1-x)^{2a+2b+1} n^{-b(1-x)}\,dx \\
&\quad\leq K\,[b\log n]^{1+2a+2b}\int_0^1 y^{2a+2b+1} n^{-by}\,dy \\
&\quad\leq K\,[b\log n]^{1+2a+2b}\int_0^{b\log n}\left[\frac{t}{b\log n}\right]^{2a+2b+1}\exp\left(-t\right)dt \\
&\quad\leq \frac{K}{[b\log n]}\int_0^{b\log n} t^{2a+2b+1}\exp\left(-t\right)dt \to 0,
\end{aligned}
$$

as $n \to \infty$. $\square$

Lemma A.1.7. *Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1(\mathbb{R})$ function such that $g(1) \neq 0$ and define the integrate I_n as*

$$
I_n = [b\log n]^{1+2a}\int_0^1 g(x)x^{2a} n^{-bx}\,dx
$$

Then $I_n \to g(0)\Gamma[1+2a]$, as $n \to \infty$.

Proof. The proof uses similar arguments as previous case. $\square$

Lemma A.1.8. *Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1(\mathbb{R})$ function such that $g(0) \neq 0$ and let $\alpha : [0,1] \to \mathbb{R}$ be a continuous function such that $\alpha(u) > -1$ for all $u \in [0,1]$. Then,*

$$I_n = (\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)} g(x) \exp(-nx^2)\, dx \to \frac{1}{2} g(0) \Gamma\left[\frac{\alpha(u)+1}{2}\right],$$

as $n \to \infty$, for any $u \in (0,1)$.

Proof. Since the function $g \in \mathcal{C}^1(\mathbb{R})$, then by Taylor's theorem we can write

$$g(x) = g(0) + g'(\xi_x)x,$$

for some $0 \leq \xi_x \leq x$, for positive x. Hence, we can write

$$\begin{aligned} I_n &= g(0)(\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)} \exp(-nx^2)\, dx \\ &\quad + (\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)+1} g(\xi_x) \exp(-nx^2)\, dx. \end{aligned} \tag{A.1.8}$$

The first integral in the expression above can be written as

$$g(0)(\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)} \exp(-nx^2)\, dx = g(0) \int_0^{u\sqrt{n}} y^{\alpha(u)} \exp(-y^2)\, dy. \tag{A.1.9}$$

Consequently,

$$\begin{aligned} g(0)(\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)} \exp(-nx^2)\, dx &\to \\ \frac{1}{2} g(0) \int_0^\infty y^{\alpha(u)} \exp(-y^2)\, dy &= g(0) \frac{1}{2} \Gamma\left[\tfrac{\alpha(u)+1}{2}\right], \end{aligned}$$

as $n \to \infty$. On the other hand, the second integral in (A.1.8) converges to zero since

$$\begin{aligned} &\left| (\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)+1} g(\xi_x) \exp(-nx^2)\, dx \right| \\ &\leq K(\sqrt{n})^{\alpha(u)+1} \int_0^u x^{\alpha(u)} \exp(-nx^2)\, dx \leq \frac{K}{\sqrt{n}} \int_0^{u\sqrt{n}} y^{\alpha(u)} \exp(-y^2)\, dy \to 0, \end{aligned}$$

as $n \to \infty$. $\square$

Lemma A.1.9. *Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1(\mathbb{R})$ function and define the integral I_n as*

$$I_n = \sqrt{n} \int_0^1 \exp[-c_n(x - x_0)^2] g(x)\, dx,$$

where $\{c_n\}$ is a sequence of positive real numbers such that $c_n/n \to 1$, as $n \to \infty$. Then,

$$I_n \to \sqrt{\pi} g(x_0) I_{(0,1)}(x_0) + \frac{\sqrt{\pi}}{2} g(x_0) I_{\{0,1\}}(x_0),$$

as $n \to \infty$.

Proof. Consider first the case $x_0 \in (0, 1)$. Since $g \in \mathcal{C}^1(\mathbb{R})$, we can write $g(x) = g(x_0) + g'(\xi_x)(x - x_0)$, for some ξ_x between x_0 and x. Consequently,

$$\begin{aligned} I_n = \quad & g(x_0)\sqrt{n} \int_0^1 \exp[-c_n(x - x_0)^2]\, dx \\ + \quad & \sqrt{n} \int_0^1 g'(\xi_x)(x - x_0) \exp[-c_n(x - x_0)^2]\, dx. \end{aligned}$$

Note that the first integral in the expression above can be written as

$$g(x_0)\sqrt{\pi} \left\{ \Phi\left[(1 - x0)\sqrt{2n}\right] - \Phi\left[-x0\sqrt{2n}\right] \right\} \left(\frac{n}{c_n}\right)^{1/2}$$

where $\Phi(\cdot)$ is the Gaussian distribution function. Thus, for any $x_0 \in (0, 1)$ we have that

$$g(x_0)\sqrt{\pi} \left\{ \Phi[(1 - x0)\sqrt{2n}] - \Phi[-x0\sqrt{2n}] \right\} \left(\frac{n}{c_n}\right)^{1/2} \to \sqrt{\pi} g(x_0),$$

as $n \to \infty$. On the other hand,

$$\begin{aligned} & \left| \sqrt{n} \int_0^1 g'(\xi_x)(x - x_0) \exp[-c_n(x - x_0)^2]\, dx \right| \\ & \leq K \sqrt{\frac{n}{c_n}} \int_0^1 \sqrt{c_n} |x - x_0| \exp(-c_n(x - x_0)^2)\, dx \\ & = K \sqrt{\frac{n}{c_n}} \int_{-x_0\sqrt{2c_n}}^{(1-x_0)\sqrt{2c_n}} |y| \exp(-y^2/2)\, \frac{dy}{\sqrt{c_n}} \\ & \leq \frac{K}{\sqrt{c_n}} \int_{-\infty}^{\infty} |y| \exp(-y^2/2)\, dy \to 0, \end{aligned}$$

as $n \to \infty$. Now, if $x_0 = 0$ or $x_0 = 1$, then

$$g(x_0)\sqrt{\pi}\left\{\Phi[(1-x_0)\sqrt{2n}] - \Phi[-x_0\sqrt{2n}]\right\}\left(\frac{n}{c_n}\right)^{1/2} \to \frac{1}{2}\sqrt{\pi}g(x_0),$$

as $n \to \infty$. Therefore, in these two cases

$$I_n \to \frac{1}{2}\sqrt{\pi}g(x_0),$$

as $n \to \infty$. Finally, if x_0 is outside the interval $[0,1]$, then $(x-x_0)^2 > \varepsilon > 0$ for all $x \in [0,1]$. Hence, $\exp[-c_n(x-x_0)^2] < \exp(-c_n\varepsilon)$ and then

$$|I_n| \leq \sqrt{n}\exp(-c_n\varepsilon)\int_0^1 |g(x)|dx \to 0,$$

as $n \to \infty$. □

Lemma A.1.10. *Assume that the process $\{Y_{t,T}\}$ satisfies* (2.2.1) *and* (2.6.5), *and that Assumption A1 holds. Then the variance of $\widehat{\mu}_T$ satisfies*

$$\mathrm{Var}(\widehat{\mu}_T) \sim \frac{1}{2}\int_0^1 \frac{T^{\alpha(u)}}{\left(\sqrt{\beta\log T}\right)^{\alpha(u)+1}} g\left(\tfrac{u}{2},\tfrac{u}{2}\right)\gamma\left[\frac{\alpha(u)+1}{2},\beta(\log T)u^2\right]d\,u,$$

where $\beta = c/2$ and $\gamma(x,a)$ corresponds to the incomplete Gamma function

$$\gamma(x,a) = \int_0^x t^{a-1}\exp(-t)d\,t.$$

Proof. From expression (2.4.6) we have that

$$\begin{aligned}\mathrm{Var}(\widehat{\mu}_T) &\sim \frac{T^{2d_0-1}}{(\beta\log T)^{d_0+1/2}}\int_0^1 T^{-\beta(u-2u_0)^2}\left(\sqrt{\beta\log T}\right)^{1+\beta(u-2u_0)^2} h_T(u)d\,u \\ &\sim \int_0^1 T^{\alpha(u)}\left(\sqrt{\beta\log T}\right)^{-\alpha(u)-1} h_T(u)d\,u. \qquad \text{(A.1.10)}\end{aligned}$$

But, from (4.5.1) with $n = \beta\log T$ we can write

$$h_T(u) = \left(\sqrt{n}\right)^{\alpha(u)+1}\int_0^u \tilde{g}(u,v)v^{\alpha(u)-\beta v^2}\exp(-nv^2)d\,v.$$

Now, by a similar arguments leading to (A.1.8) we have that

$$h_T(u) \sim \left(\sqrt{n}\right)^{\alpha(u)+1} \tilde{g}(u,0) \int_0^u v^{\alpha(u)-\beta v^2} \exp\left(-nv^2\right) d\, v$$
$$+ \left(\sqrt{n}\right)^{\alpha(u)+1} \int_0^u \tilde{g}(u,\xi_v) \xi_v^{-\beta \xi_v^2} v^{\alpha(u)} \exp\left(-nv^2\right) d\, v,$$

for some $\xi_v \in [0,u]$. But, analogously to the proof of Lemma (A.1.8), the second integral in the expression above is negligible for large n. Thus,

$$h_T(u) \sim \left(\sqrt{n}\right)^{\alpha(u)+1} \tilde{g}(u,0) \int_0^u v^{\alpha(u)-\beta v^2} \exp\left(-nv^2\right) d\, v.$$

Now, by replacing this expression in (A.1.10), the result follows. □

A.2 Supplementary Material for Chapter 4

Lemma A.2.1. *Define R and $\tilde{R}$ be the $T \times T$ matrices with (i,j)th entry $\gamma(h,u)$ and $\tilde{\gamma}(h,u)$, respectively, where*

$$\tilde{\gamma}(h,u) = \int_{-\pi}^{\pi} \frac{e^{i\lambda h}}{\left|1-e^{i\lambda}\right|^{2d(u)}}\, d\lambda,$$

and I_T be the $T \times T$ identity matrix. Then

$$R_T \leq K\tilde{R}_T \leq KT^{2d_s} I_T \qquad \forall u \in [0,1].$$

Proof. Let x be a fixed vector, since

$$\begin{aligned} x'R_T x &= \sum_{t,s=1}^{T} \gamma(t-s) x_t x_s \\ &= \sum_{t,s=1}^{T} \int_{-\pi}^{\pi} e^{i\lambda(t-s)} x_t x_s f(u,\lambda) \mathrm{d}\lambda \\ &= \int_{-\pi}^{\pi} \sum_{t,s=1}^{T} e^{(i\lambda(t-s))} x_t x_s \frac{f_0(\lambda)}{|1-e^{i\lambda}|^{2d(u)}} \mathrm{d}\lambda \\ &= \int_{-\pi}^{\pi} \left| \sum_{t=1}^{T} e^{i\lambda t} x_t \right|^2 \frac{f_0(\lambda)}{|1-e^{i\lambda}|^{2d(u)}} \mathrm{d}\lambda \\ &\leq K \int_{-\pi}^{\pi} \left| \sum_{t=1}^{T} e^{i\lambda t} x_t \right|^2 \frac{1}{|1-e^{i\lambda}|^{2d(u)}} \mathrm{d}\lambda \\ &= K x' \tilde{R}_T x. \end{aligned}$$

Let $y = \tilde{R}_T^{1/2} x$, then

$$y'y = x' \tilde{R}_T^{1/2} \tilde{R}_T^{1/2} x = x' \tilde{R}_T x = \int_{-\pi}^{\pi} \frac{\left| \sum_{t=1}^{T} e^{i\lambda t} x_t \right|^2}{|1-e^{i\lambda}|^{2d(u)}} \, d\lambda.$$

Thus

$$\frac{y'y}{x'x} = \frac{\displaystyle\int_{-\pi}^{\pi} \frac{\left| \sum_{t=1}^{T} e^{i\lambda t} x_t \right|^2}{|1-e^{i\lambda}|^{2d(u)}} \, d\lambda}{\displaystyle 1/2\pi \int_{-\pi}^{\pi} \left| \sum_{t=1}^{T} e^{i\lambda t} x_t \right|^2 d\lambda}. \tag{A.2.1}$$

Now, for $0 < d(u) < \frac{1}{2}$ we have

$$\frac{|\lambda|^{2d(u)}}{|1-e^{i\lambda}|^{2d(u)}} = \left| \frac{\lambda/2}{\sin(\lambda/2)} \right|^{2d(u)}, \quad \lambda \in [-\pi, \pi].$$

Therefore for $d(u) \in (0, 1/2)$,

$$\left| \frac{\lambda/2}{\sin(\lambda/2)} \right|^{2d(u)} \leq \left(\frac{\pi}{2} \right)^{2d(u)} = C(u).$$

Since $0 < d(u) < \frac{1}{2} \Longrightarrow C(u) \leq \frac{\pi}{2} \equiv C$. Hence

$$\frac{|\lambda|^{2d(u)}}{|1-e^{i\lambda}|^{2d(u)}} = \left|\frac{\lambda/2}{\sin(\lambda/2)}\right|^{2d(u)} \leq C.$$

Thus in (A.2.1)

$$\frac{y'y}{x'x} = \frac{\int_{-\pi}^{\pi} \frac{\left|\sum_{t=1}^{T} e^{i\lambda t} x_t\right|^2}{|1-e^{i\lambda}|^{2d(u)}} \mathrm{d}\lambda}{1/2\pi \int_{-\pi}^{\pi} \left|\sum_{t=1}^{T} e^{i\lambda t} x_t\right|^2 \mathrm{d}\lambda} \leq 2\pi C \frac{\int_{-\pi}^{\pi} |\lambda|^{-2d(u)} \left|\sum_{t=1}^{T} e^{i\lambda t} x_t\right|^2 \mathrm{d}\lambda}{\int_{-\pi}^{\pi} \left|\sum_{t=1}^{T} e^{i\lambda t} x_t\right|^2 d\lambda}.$$

Define $h^*(\lambda) = \frac{|\sum_{j=1}^{T} e^{i\lambda j} x_j|^2}{\int_{-\pi}^{\pi} |\sum_{j=1}^{T} e^{i\lambda j} x_j|^2 \, d\lambda}$. Then, this is a probability function over $[-\pi, \pi]$ satisfying $\int_{-\pi}^{\pi} h^*(\lambda) d(\lambda) = 1$ and $h^*(\lambda) \leq \frac{T}{2\pi}$. Consequently,

$$\begin{aligned} \frac{y'y}{x'x} &\leq \int_{-\pi}^{\pi} |\ \lambda\ |^{-2d(u)}\ h^*(\lambda) d(\lambda) \leq CT \int_{0}^{\pi/T} \lambda^{-2d(u)} d(\lambda) \\ &= \frac{C\pi^{1-d(u)}}{1-2d(u)} T^{2d(u)} \leq KT^{2d_0}. \end{aligned}$$

Hence $\tilde{R}_T \leq KT^{2d_0} I \quad \forall u \in [0,1]$. □

Proof. 4.6.1. The proof uses argument analogues as in Theorem 3.2 and Theorem 3.6 of Dahlhaus (1997). The consistency of $\widetilde{\theta}_T$ follows with the proof of Theorem 3.2, hence if we show that

$$\sup_{\theta} \left| \mathcal{L}_T(\theta, x_j(t/T)\widehat{\beta}) - \mathcal{L}_T(\theta, x_j(t/T)\beta) \right| \to_p 0$$

for $j = 1, \ldots, p$ that is, if we show

$$\sup_{\theta} \left| \frac{1}{M} \sum_{j=1}^{M} \int_{-\pi}^{\pi} \left\{ I_N^{x_j(t/T)\widehat{\beta}}(u_j, \lambda) - I_N^{x_j(t/T)\beta}(u_j, \lambda) \right\} \phi_\theta(u_j, \lambda)\, d\lambda \right| \to_p 0,$$

where $\phi_\theta(u_j, \lambda) = f_\theta(u_j, \lambda)^{-1}$. By the mean value theorem there exist a vector $\overline{\theta}_T$ satisfying $|\overline{\theta}_T - \theta_0| \leq |\widehat{\theta}_T - \theta_0|$, such that

$$\sqrt{T} \left\{ \nabla \mathcal{L}_T(\widetilde{\theta}_T, x_j(t/T)\widehat{\beta}) - \nabla \mathcal{L}_T(\theta_0, x_j(t/T)\widehat{\beta}) \right\} = \nabla^2 \mathcal{L}_T(\overline{\theta}_T, x_j(t/T)\widehat{\beta}) \sqrt{T}(\widetilde{\theta}_T - \theta_0)$$

it suffices show that

(A.1) $\sqrt{T}\nabla\mathcal{L}_T(\theta_0, x_j(t/T)\widehat{\beta}) - \sqrt{T}\nabla\mathcal{L}_T(\theta_0, x_j(t/T)\beta) \to_p 0$

(A.2) $\nabla^2\mathcal{L}_T(\overline{\theta}, x_j(t/T)\widehat{\beta}) \to_p \Gamma.$

Therefore the result follows if we show that

$$\sqrt{T}\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\left\{I_N^{x_j(t/T)\widehat{\beta}_T}(u_j,\lambda) - I_N^{x_j(t/T)\beta}(u_j,\lambda)\right\}\phi_{\theta_0}(u_j,\lambda)\,d\lambda \to_p 0 \quad \text{(A.2.2)}$$

for $j = 1,\ldots,p$ and $\phi_\theta(u,\lambda) = \nabla f_\theta(u,\lambda)^{-1}$ and

$$\sup_\theta\left|\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\left\{I_N^{x_j(t/T)\widehat{\beta}}(u_j,\lambda) - I_N^{x_j(t/T)\beta}(u_j,\lambda)\right\}\phi_{\theta_0}(u_j,\lambda)\,d\lambda\right| \to_p 0 \quad \text{(A.2.3)}$$

for $\phi_\theta(u,\lambda) = f_\theta(u,\lambda)^{-1}$ and $\phi_\theta(u,\lambda) = \nabla^2 f_\theta(u,\lambda)^{-1}$. The last expression is equal to

$$\begin{aligned}\sup_\theta\Bigg|\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}&\phi_\theta(u_j,\lambda)\left\{2\pi H_{2,N}(0)\right\}^{-1}\\ &\times\Big\{d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)d_N^{x_j(t/T)\left(\beta-\widehat{\beta}\right)}(u_j,-\lambda)\\ &\quad+d_N^{x_j(t/T)\left(\beta-\widehat{\beta}\right)}(u_j,\lambda)d_N^{Y-x_j(t/T)\beta}(u_j,-\lambda)\\ &\qquad+d_N^{x_j(t/T)\left(\beta-\widehat{\beta}\right)}(u_j,\lambda)d_N^{x_j(t/T)\left(\beta-\widehat{\beta}\right)}(u_j,-\lambda)\Big\}\,d\lambda\Bigg|\end{aligned} \quad \text{(A.2.4)}$$

which by means of the Cauchy-Schwarz inequality is with

$$\delta_T := \frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\left\{2\pi H_{2,N}(0)\right\}^{-1}\left|d_N^{x_j(t/T)\left(\beta-\widehat{\beta}\right)}(u_j,\lambda)\right|^2\,d\lambda$$

bounded by

$$\sup_{\theta,u,\lambda}|\phi_\theta(u,\lambda)|\left\{2\left(\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}I_N^{x_j(t/T)\beta}(u_j,\lambda)\,d\lambda\right)^{1/2}\delta_T^{1/2} + \delta_T\right\}.$$

The expression $(1/M)\sum_{j=1}^{M}\int_{-\pi}^{\pi} I_N^{x_j(t/T)\beta}(u_j,\lambda)\,d\lambda$ is bounded in probability (Theorem 2) see Palma and Olea (2010) and

$$\begin{aligned}
\delta_T &= \frac{1}{M}\sum_{j=1}^{M} H_{2,N}(0)^{-1}\sum_{s=1}^{N} h\left(\tfrac{s-1}{N}\right)^2\left\{x_j\left(\frac{t_j-N/2+s}{T}\right)\beta\right.\\
&\qquad \left.-x_j\left(\frac{t_j-N/2+s}{T}\right)\widehat{\beta}_T\right\}^2\\
&= o_p\left(\tfrac{NT^{2d_0-1}}{\log T^{2d_0+1}}\right)
\end{aligned}$$

therefore (A.2.3) is proved. To prove (A.2.2) we note that $\sqrt{T}\delta_T \to 0$. Since $\sqrt{T}\delta_T^{1/2} \nrightarrow 0$ we need a better estimate for the first and second term of (A.2.4). Let $c_T := \sqrt{T}\left\{2\pi H_{2,N}(0)\right\}^{-1}$, $\overline{H}_{t,N}(\lambda) := \sum_{s=0}^{t-1} h(s/N)e^{-i\lambda s}$ and $\bar{t}_j = t_j - N/2$,

$$\begin{aligned}
&\frac{c_T}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)d_N^{x_j(t/T)\left(\beta-\widehat{\beta}_T\right)}(u_j,-\lambda)\,d\lambda\\
&= c_T\sum_{j=1}^{M}\sum_{t=0}^{N-1}\left\{x_j\left(\frac{\bar{t}_j+t+1}{T}\right)\beta - x_j\left(\frac{\bar{t}_j+t+1}{T}\right)\widehat{\beta}_T\right\}\\
&\quad\times\int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)\left\{\overline{H}_{t+1,N}(-\lambda)-\overline{H}_{t,N}(-\lambda)\right\}\,d\lambda\\
&= -c_T\sum_{j=1}^{M}\sum_{t=0}^{N-1}\left[\left\{x_j\left(\frac{\bar{t}_j+t+1}{T}\right)\beta - x_j\left(\frac{\bar{t}_j+t+1}{T}\right)\widehat{\beta}_T\right\}\right.\\
&\quad\left.-\left\{x_j\left(\frac{\bar{t}_j+t}{T}\right)\beta - x_j\left(\frac{\bar{t}_j+t}{T}\right)\widehat{\beta}_T\right\}\right]\\
&\quad\times\int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)\overline{H}_{t,N}(-\lambda)\,d\lambda\\
&+c_T\sum_{j=1}^{M}\left\{x_j\left(\frac{\bar{t}_j+t}{T}\right)\beta - x_j\left(\frac{\bar{t}_j+t}{T}\right)\widehat{\beta}_T\right\}\\
&\quad\times\int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)\overline{H}_{t,N}(-\lambda)\,d\lambda.
\end{aligned}$$

On other hand we have $\overline{H}_{t,N}(-\lambda) \le KL_N(\lambda)$ uniformly in t. Utilizing the Lemma (A.2.2) we have that

$$\operatorname{Var}\int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)\overline{H}_{t,N}(-\lambda)\,d\lambda = O(N),$$

uniformly in u_j and t. Since $E(d_N^{Y-x_j(t/T)\beta}(u_j,\lambda)) = 0$ the whole expressions tends to zero in probability. The second term of (A.2.4) is treated in the same way, which proves the result. □

Lemma A.2.2. *Let $I(u,\lambda)$ be defined by*

$$I(u,\lambda) = \int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-\mu}(u_j,\lambda)\overline{H}_{t,N}(-\lambda)\,d\lambda$$

where $\phi_{\theta_0}(u_j,\lambda) = \nabla f_\theta(u,\lambda)^{-1}$. Then, there exist a constant $K>0$ such that

$$|\operatorname{Var} I(u,\lambda)| \le KN.$$

Proof.

$$\begin{aligned}
I(u,\lambda) &= \int_{-\pi}^{\pi}\phi_{\theta_0}(u_j,\lambda)d_N^{Y-\mu}(u_j,\lambda)\overline{H}_{t,N}(-\lambda)\,d\lambda \\
&= \int_{-\pi}^{\pi}\nabla f_\theta(u_j,\lambda)^{-1}\sum_{s=0}^{N-1} h\left(\tfrac{s}{N}\right)h\left(\tfrac{t}{N}\right)\left(Y_{[uT]-N/2+s+1}-\mu\right) \\
&\qquad \times e^{-i\lambda s}\sum_{j=0}^{t-1} h\left(\tfrac{j}{N}\right)e^{i\lambda j}\,d\lambda \\
&= \sum_{s=0}^{N-1}\sum_{j=0}^{t-1} h\left(\tfrac{s}{N}\right)h\left(\tfrac{j}{N}\right)\left(Y_{[uT]-N/2+s+1}-\mu\right)\int_{-\pi}^{\pi}\nabla f_\theta(u_j,\lambda)^{-1}e^{i\lambda(j-s)}\,d\lambda \\
&= \sum_{s=0}^{N-1} h\left(\tfrac{j}{N}\right)\left(Y_{[uT]-N/2+s+1}-\mu\right)c(s),
\end{aligned}$$

where $c(s) = \sum_{j=0}^{t-1} h\left(\frac{s}{N}\right)\gamma_\mu(j-s)$. Then, under assumption A7 and A9 we have that

$$
\begin{aligned}
|\mathrm{Var}(I(u,\lambda))| &= \left|\mathrm{Var}\left(\sum_{s=0}^{N-1} h\left(\tfrac{s}{N}\right)\left(Y_{[uT]-N/2+s+1}-\mu\right)c(s)\right)\right| \\
&= \left|\sum_{s,k=0}^{N-1} h\left(\tfrac{s}{N}\right) h\left(\tfrac{k}{N}\right) c(s)c(k)\gamma_{\mu}(s-k)\right| \\
&\leq K\sum_{s=0}^{N}\sum_{k=0}^{N}|s-k|^{-2d_\theta(u)-1} \leq K\sum_{j=0}^{N}(N-j)j^{-2d_\theta(u)-1} \leq KN
\end{aligned}
$$

$\square$

Bibliography

Adenstedt, R. K. (1974). On large-sample estimation for the mean of a stationary random sequence. *Ann. Statist. 2*, 1095–1107.

Beran, J. (1994). Statistics for long-memory processes.

Beran, J. (2008). On parameter estimation for locally stationary long-memory processes. *Journal of Statistical Planning and Inference,*. in press.

Brillinger, D. R. (1981). *Time series* (Second ed.). Oakland, Calif.: Holden-Day Inc. Data analysis and theory, Holden-Day Series in Time Series Analysis.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (second ed.). New York: Springer.

Brockwell, P. J. and R. A. Davis (1996). *Introduction to Time Series and Forecasting.* New York: Springer.

Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl. 62*(1), 139–168.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist. 25*(1), 1–37.

Dahlhaus, R. and L. Giraitis (1998). On the optimal segment length for parameter estimates for locally stationary time series. *J. Time Ser. Anal. 19*(6), 629–655.

Genton, M. and O. Perrin (2004). On a time deformation reducing nonstationary stochastic processes to local stationarity. *J. Appl. Probab. 41*(1), 236–249.

Gradshteyn, I. S. and I. M. Ryzhik (2000). *Table of Integrals, Series, and Products* (Sixth ed.). San Diego, CA: Academic.

Granger, C. W. J. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal. 1*(1), 15–29.

Grenander, U. and M. Rosenblatt (1954). An extension of a theorem of G. Szegö and its application to the study of stochastic processes. *Trans. Amer. Math. Soc. 76*, 112–126.

Grenander, U. and M. Rosenblatt (1957). *Statistical Analysis of Stationary Time Series*. New York: Wiley.

Hall, P. (1997). Defining and measuring long-range dependence. In *Nonlinear Dynamics and Time Series (Montreal, PQ, 1995)*, Volume 11 of *Fields Inst. Commun.*, pp. 153–160. Providence, RI: Amer. Math. Soc.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika 68*(1), 165–176.

Hosking, J. R. M. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *Journal of Econometrics 73*.

Ibragimov, I. A. and Y. V. Linnik (1971). Independent and stationary sequences of random variables. *Wolters-Noordhoff, Groningen ..*

Jensen, M. and B. Witcher (2000). Time-varying long memory in volatility: detection and estimation with wavelets. Technical report, EURANDOM.

Palma, W. (2007). *Long-Memory Time Series: Theory and Methods*. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley.

Palma, W. and R. Olea (2010). An efficient estimator for locally stationary Gaussian long-memory preocesses. *Ann. Statist.*, In press.

Parzen, E. (1986). Behavior of sample means and nonparametric time-series estimation. In *WSC '86: Proceedings of the 18th conference on Winter simulation*, New York, NY, USA, pp. 337–342.

Parzen, E. (1992). Time series, statistics, and information. In *New directions in time series analysis, Part I*, Volume 45 of *IMA Vol. Math. Appl.*, pp. 265–286. New York: Springer.

Pourahmadi, M. (2001). *Foundations of Time Series Analysis and Prediction Theory.* New York: Wiley.

Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *J. R. Stat. Soc. Ser. B Stat. Methodol. 27*, 204–237.

Samarov, A. and M. S. Taqqu (1988). On the efficiency of the sample mean in long-memory noise. *J. Time Ser. Anal. 9*(2), 191–200.

Silverman, R. (1957). Locally stationary random processes. *Information Theory, IRE Transactions on 3*(3), 182–187.

Yajima, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Ann. Statist. 19*(1), 158–177.