République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université A.MIRA-BEJAIA

Faculté des Sciences Exactes
Département de Mathématiques
Laboratoire de Mathématiques Appliquées

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT

Domaine : Mathématiques et informatique Filière : Mathématiques
Spécialité : Probabilités et Statistiques

Présentée par
Seba Djillali

Thème

Étude Des Processus Fractionnaires à Erreurs Mélangeantes

Soutenue le : 10 Novembre 2022 **Devant le Jury composé de :**

Nom et Prénom	Grade		
Mme Zerouati Halima	Professeur	Univ. De Bejaia	Présidente
Mme Karima Belaide	Professeur	Univ. De Bejaia	Rapporteur
Mr Kandouci Abdeldjebbar	Professeur	Univ. De Saida	Examinatrice
Mme Lagha Karima	Professeur	Univ. De Bejaia	Examinatrice
Mme Saadi Nora	MCA	Univ. De Bejaia	Examinatrice
Mme Sadki Ourida	Professeur	Univ. USTHB	Examinatrice

Année Universitaire : 2021-2022

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Reaserch
University of A.MIRA-BEJAIA

Faculty of Exact Sciences
Departement of Mathematics
Laboratory of Applied Mathematics

THESIS

IN VIEW OF OBTAINING A DOCTORATE DIPLOMA

Domain : Mathematics and Computer Science Field : Mathematics
Speciality : Probability et Statistics

Presented by
Seba Djillali

Theme

A Study On Fractional Autoregressive Process With Mixing Errors

Sustained : 10 November 2022 **Board of the Assembled Jury :**

Full Name	Grade		
Ms. Zerouati Halima	**Professor**	Univ. De Bejaia	Chairwoman
Ms. Karima Belaide	**Professor**	Univ. De Bejaia	Supervisor
Mr. Kandouci Abdeldjebbar	**Professor**	Univ. De Saida	Examiner
Ms. Lagha Karima	**Professor**	Univ. De Bejaia	Examiner
Ms. Saadi Nora	**MCA**	Univ. De Bejaia	Examiner
Ms. Sadki Ourida	**Professor**	Univ. USTHB	Examiner

Academic Year : 2021-2022

Acknowledgements

First of all, I thank ALLAH for giving me the courage and the determination, as well as guidance in conducting this research study despite all difficulties. I want to express my great appreciation and deepest gratitude to my supervisor, Ms. Belaide Karima, without whom this dissertation would not have been emerged. It was a great honor and an immense pleasure to be able to accomplish this work under her direction. I could benefit from her sound advice and careful guidance, her availability. Again thank you !

I want to thank very warmly Ms. Zerouati Halima for chairing my committee members. I wish to extend my sincere thanks to Mr Kandouci Abdeldjebbar, Sadki Ourida, Lagha Karima, Saadi Nora. for agreeing to judge and criticize this work assuming roles of examiners.

I am grateful to my family for their continuous support, their encouragement, and especially for their putting up with me during all my study period. I would like to thank my parents for their wise counsel and sympathetic ear. You are always there for me.

I could not have completed this thesis without the support of my colleagues and my friends, who provided stimulating discussions as well as happy distractions to rest my mind outside of my research.

Last but certainly not least, I would like to thank everybody who was important to the successful realization of this Thesis.

Dedication

To my beloved mother

For the special little things you do, for all the words that sometimes go unspoken, for your kind support throughout the years, for teaching me the meaning of compassion, and sharing my triumphs and my tears, I truly hope you see that nothing you have done has been forgotten, and day by day you just mean more to me thank you for all that you have done and you are still doing for me.

To my dear father

Thank you for your help, encouragement and support.

To my brother and my sisters

To all my friends

Contents

List of Figures

List of Tables

List of Works

Publications

- On several local asymptotic properties for Fractional Autoregressive process with strong mixing noises. (Published in Communication in Statistics- Simulation and Computation.) doi: 10.1080/03610918.2022.2055069
- On probabilistic properties of fractional autoregressive process of order 1 with strong mixing errors (Submitted)

Communications

- On Fractional Autoregressive Process Of Order 1 With Strong Mixing Errors. 1st International Conference on Pure and Applied Mathematics IC-PAM'21, May 26-27, 2021, Ouargla, Algeria (Virtual conference).
- Local Asymptotic Normality Property For ARFIMA Model. Mini Congress of the Algeria Mathematicians, MCMA'2021, October 27-28, 2021, Msila, Algeria (Virtual conference).
- On Local Asymptotic Normality. Seminar of Laboratory (LMA) Bejaia. 1st June, 2021, Bejaia.

Notations

$\mathbb{E}$	Mathematical expectation
P	Probability
AR	AutoRegressive
MA	Moving Average
$ARIMA$	AutoRegressive Integrated Moving Average
$ARFIMA$	AutoRegressive Fractional Integrated Moving Average
$\Gamma(.)$	Gamma function
$F(.,.,.,.)$	Hypergeometric function
I_f	Fisher information
$i.i.d$	Independent and Identically Distributed
LAN	Local Asymptotic Normality
LAQ	Local Asymptotic Quadratic
LAM	Local Asymptotic Minimaxity
$\longrightarrow^d$	Convergence in distribution.
$\longrightarrow^P$	Convergence in probability.

General Introduction

Time series analysis has developed tremendously and is now widely used in both theoretical and applied fields. Throughout the previous four decades a significant number of studies have been devoted to understanding and applying these techniques.

Time series data is an ordered sequence of observations of well defined data items at regular time intervals. The number of research in this area continues to increase because time is component of everything that can be observed. Time series analysis has two basic goals: determining the nature of the phenomenon represented by a set of data and forecasting (predicting future values of the time series variable). Such data has numerous applications in finance, hydrology, biology, physics...etc. For example: In the study of weather records; time series analysis is used frequently by weatherman to explain and predict what the temperatures will be during different months and seasons throughout the year. In economic indicators, we have the example of stock prices we can gain a better understanding of the patterns in various stock prices. In the medical industry, time series analysis is used to monitor the heart rate of patients who are taking particular medications to ensure that their heart rate does not fluctuate too much at any given time of day.

The theoretical developments started early with stochastic processes, the first actual application of autoregressive models to data can be brought back to the work of Yule [89] and Slutsky [75] in the (1926) and (1927). Moving average is a method of smoothing data points that has been in use for decades before this, it was known as "instantaneous averages" Hooker [46] in (1901), and Yule defined them as "moving-averages" in (1909). Although they did not use the phrase in his textbook, it was popularized by King's Elements of Statistical Method [51](1912). Wold [85] "process of moving average" (A Study in the Analysis of Stationary Time Series (1938)) is an abbreviation for a sort of stochastic process, Wold described how special cases of the process had been studied in the (1927) by Yule [89], the moving average was created to remove periodic fluctuations in time series, such as those caused by seasonality. For stationary series, Wold [85] introduced ARMA (AutoRegressive Moving Average) models.

The expression ARMA(p,q) is written as follows

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{1}$$

Where φ_i and θ_i are the parameters of the model, ε_i are i.i.d. errors.

ARMA(p,0) is AR(p), ARMA(0,q) is MA(q).

The Box-Jenkins [16] method considered a fundamental contribution to time series analysis appeared only in (1970). Three principles guide this model's handling of data: autoregression, differencing, and moving average. These three principles are known as p,d and q, respectively. the autoregression (p) process tests the data for its level of stationarity, if the data being used is stationary, it can simplify the forecasting process, if the data are non-stationary it will need to be differenced (d). The data is also tested for its moving average fit (which is done in part q of the analysis process).

Overall, initial analysis of the data prepares it for forecasting by determining the parameters (p, d, and q), which are then applied to develop a forecast. It popularized the autoregressive integrated moving average ARIMA(p,d,q) model by using an iterative modeling procedure consisting of identification, estimation, and model checking.

Box-Jenkins multivariate models, are used to analyze more than one time dependent variable, such as cases of COVID-19, lockdown over time and prices of food.

The first generalization was to accept multivariate ARMA models, among which especially VAR models (Vector AutoRegressive) have become popular, Litterman and Doan [26](1986) discussed VAR model , Litterman and Sims [52] (1984) shed light on BVAR model. These techniques, however, are only applicable for stationary time serie. Peiris studied an autoregressive univariate model of order 1 with time dependent coefficient, they established an extension of the model in multivariate case. Osborn [62](1988), Birchenhall et al. [12] were among the first to introduce periodic models into economics (1989), other developments were extensions to multivariate periodic models in Boswijk et al. [14](1997). A thorough account of the developments in application are reported in Xirasagar [87](2005), Anderoni et al. [4](2006), Tsitsika et al. [81](2009).

Time series analysis was originally divided into frequency domain and time domain approach, The data autocorrelation function and parametric models, such as ARIMA models, are used in the time domain approach to describe the dynamic dependence of the series Box, Jenkins, and Reinsel [18](1994), The frequency domain approach focuses on spectral analysis or power distribution over frequency to study theory and applications of time series analysis.

The approach of Box-Jenkins can only capture short range dependence which is characterized by an exponential decay of the autocorrelation function Brockwell and Davis [20](1991). On the other hand long range dependence (LRD) means that the effect of a shock could not disappear quickly but it takes a long time to vanish in other words the sum of the auto-correlations is infnite Baillie [5](1996), the classical models cited above can not be suitable to describe (LRD).

Fractional AutoRegressive Integrated Moving Average (ARFIMA) is a well known model which can be accurate in modeling (LRD), it presents an extension of usual ARIMA by allowing for fractional degrees of integration.

"Fractional differencing is a crucial step in the building of ARFIMA model, However, due to difficulties in fractional differencing, most researchers use first-order differencing as an alternative, such technique will certainly cause over-differencing, which will lead to the loss of information Huang" [48](2010).

Empirical evidence of long memory are numerous in different fields such as astronomy, chemistry, agriculture, and geophysics originates from considerably earlier eras, for example Newcomb [61] (1886), Student [78](1927), Fairfield Smith [76](1938),

Jeffreys [50](1939). This model has been initially treated in the field of hydrology with Hurst's works [49] (1951) on the floods Nile, Hosking [47](1981), Granger and Joyeux [35](1980) introduced ARFIMA model, they based on the works of Mandelbrot and Van Ness [60](1968) who dealt with Fractional Brownian motion.
The model ARFIMA(0,d,0) is defined as follow

$$(1-L)^d X_t = \varepsilon_t \tag{2}$$

L is lag operator, $d \in \mathbb{R}$, ε_t i.i.d. errors.
Gonçalves [33](1987) presented an extension of the model ARMA(p,q) where p and q can take real values, this model includes ARIMA(p,d,q) short and long range dependence. Beran [9] (1994) dealt with long memory processes and its probabilistic and statistical properties, Further information on ARFIMA(p,d,q) models was given by Sowell [74] (1992), Chung [23] (1994), Brockwell and Davis [21](2002).
Identifying the existence of long memory feature via techniques such as Autocorrelation Function (ACF) test is possible, this method is one of the most popular tests identifying the memory of the time series first introduced by Ding and Granger [36](1996). In this test, autocorrelation graph decreases from a certain value very slowly or hyperbolically. Therefore, such time series have long memory feature.
There are other techniques used in the identifying of the long memory feature such as Gewek and Porter-Hudak (GPH), this method is based on frequency domain analysis.
The concept Fractional differentiation of ARIMA process was further expanded by Andel [3] (1986) and Gray, Zhang, and Woodward [86] (1989, 1994), who introduced the Gegenbauer ARMA abbreviated as GARMA class of time series models based on the theory of Gegenbauer polynomials (see also Giraitis and Leipus, [32] (1995), and Woodward, Cheng, and Gray, [37](1998)).
A GARMA process is a long memory process generated by

$$(1+2uL+L^2)^d X_t = \varepsilon_t \tag{3}$$

Where $|u| < 1$, $0 < d < \dfrac{1}{2}$, ε_t i.i.d. errors.
When $u = 1$, we obtain ARFIMA(0,2d,0) process, it means that GARMA is Generalized model of ARFIMA.

Sabzikar et al has recently presented a new time series model called AutoRegressive Tempered Fractionally Integrated Moving Average (ARTFIMA).

$$(1 - e^{-\lambda} L)^d X_t = \varepsilon_t \tag{4}$$

λ, d are real parameters, ε_t are i.i.d. errors.
This model exhibit semi-long range dependence, their autocovariance function resembles long range dependence for a number of lags, depending on the tempering parameter, but eventually it decreases quickly. This model has a summable covariance function, since the tempering parameter can be made as small as we like, the mathematically more tractable ARTFIMA model can fit data that is usually modeled using the ARFIMA model.
An important step in the Box-Jenkins approach is the estimation of parameters, one of the easiest methods is the method of moments (MOM), model parameters are estimated by equating population and sample moments, it was introduced by Pafnuty Chebyshev in (1887). Least squares estimation (LSE) was viewed as generalization of method of moments, it consists to minimize the sum of the squared vertical distances, the maximum likelihood estimator (MLE) estimate model parameters by maximizing the likelihood function, this method was developed by Fisher in (1922). In the case of ARFIMA models, Geweke and Porter-Hudak [31] proposed a semi-parametric estimator for ARFIMA(0,d,0), Yajima [88] (1985) discussed the estimation of long memory parameter using maximum likelihood estimator, Bloomfield and Sastry [13](1992) worked on the estimation of the parameters of ARFIMA(p,d,q), Sowell [74] (1992) derived a procedure to compute the unconditional exact likelihood function, Beran [10] (1995) developed an approximation based on conditional least squares.
The adaptive estimation was also widely treated by the researchers due to its performance, an adaptive estimator is an estimator in a parametric or semiparametric model with nuisance parameters such that the presence of these nuisance parameters does not affect efficiency of estimation. This technique start with stein [77] (1956) who dealt with traditional models in the independent case, Fabian and Hannan (1982) [28] established results for a family of model that fulfill the property of

Local Asymptotic Normality (LAN) criterion which is introduced by Le Cam [58] (1960), this notion consists in approaching a series of statistical experiments by a gaussian family, it is fundamental in parametric theory and it is used to describe the asymptotic optimality of estimators. Le Cam [59](1986) developed this technique in the case of limit experiences, Bickel [11](1982) established some results on adaptive estimation in regression models.

Swensen [79](1985) prove the asymptotic normality of the likelihood ratio of autoregressive time series with a regression, Kreiss [55](1987) established LAN of stationary ARMA(p,q) process with independent and identically distributed but not necessary gaussian, then they treated Local Asymptotic Minimaxity (LAM), it leads us to show that the normal distribution is the best limit and describes the optimality of the estimator, this concept was introduced by Haejk [42](1972). Kreiss [56](1987) generalized adaptive estimation for non gaussian autoregressive models with infinite order. Furthermore Hallin and Puri [45](1994) dealt with ARMA processes with a linear regression trend under unspecified innovations then they showed LAN property, Garel and Hallin [30](1995) established LAN result for vector ARMA models with linear trend, a regression model ARFIMA(p,d,q) with long memory disturbances are studied by Hallin et al [44](1999), we have also the contributions of Hallin and Lotfi (2005), Bentarzi et al. [6](2009), Guta and Kiukas [38](2015), Kara-terki and Mourid [53](2016) treated the local asymptotic normality of function autoregressive processes, Haddad and Belaide [40](2019) extended the results of Serroukh [72](1996).

Brief Outline of the Thesis

This dissertation is devoted to study a fractional autoregressive model of order 1, FAR(1).

$$(1 - aL)^d X_t = \varepsilon_t \tag{5}$$

L is lag operator, a and d are real parameters, ε_t are strong mixing noises, with $\sigma^2 < \infty$.
In our case, we insure the invertibility and causality using the following conditions

- (1) $|a| < 1$ and $d \in \mathbb{R}$
- (2) $a = 1$ and $|d| < \dfrac{1}{2}$.

The errors in this model are assumed to be mixing, in reality is very fruitful to assume the dependence of innovations, because it is applicable in economics, finance and other fields, that is why there is a vast literature on the notion of mixing coefficients of the random variables, especially strong mixing process ($\alpha-$mixing), introduced by Rosenblatt [68](1956)
The main objective of this work is to establish some results about special ARFIMA model and the derivation of several local asymptotic properties (Local Asymptotic Normality (LAN) and Local Asymptotic Minimaxity (LAM)).
This thesis is set out as follows
In the first chapter, we cope with preliminaries, basic notions, definitions and key concepts that will be used in the next chapters to arrive at the primary conclusions, then we discuss some properties of ARIMA and ARFIMA models, especially Fractional Autoregressive model of order 1. (Invertibility and causality conditions, autocovariance function, autocorrelation function and their asymptotic behavior)
The second chapter is dedicated to generalize some probabilistic properties in the case of Fractional Autoregressive process (FAR(1)) with strong mixing noises. We treat the geometrical strong mixing case and the arithmetical strong mixing case, then we check the validity of the theoretical results by a simulation study, we show

the effect of strong mixing coefficients on the behavior of the autocorrelation function, this chapter was the subject of an article, which is submitted.
The third chapter is devoted to study several local asymptotic properties, we explore the Local Asymptotic Normality (LAN) property, give the local Asymptotic Quadratic expression, and prove the primary result using Swensen's conditions. Finally, we demonstrate our method with a simulation study. After that, we'll look at the properties of Local Asymptotic Minimaxity (LAM) and Local Asymptotic Linearity, This chapter was the subject of an article, which is published in Communication in statistics simulation and computation.

Chapter 1

Preliminaries on Time Series Processes

Introduction

This chapter is devoted to study some characteristics of time series model, we answer basic and important questions. Is the series stationary, invertible, causal? Fluctuations in the series take long time or short time to vanish?
We use the Box-Jenkins Methodology to investigate various statistical and probabilistic aspects of time series processes.
We start with the study of classical time series models AR, MA, ARMA and ARIMA, then we deal with ARFIMA process and their properties basing on some definitions and tools.
Furthermore, we discuss a key idea in asymptotic theory known as Local Asymptotic Normality, which was introduced by Lucien Le Cam, this idea dates back to Wald [83](1943) they used a Gaussian family to approximate a series of statistical experiments, the results of projecting the (LAN) property on a Fractional Autoregressive Model of order 1 (Long Memory Model) in the case of independent noises and mixing noises are then presented.

1.1 Definition and Tools in Time Series

In what follows, we regroup some definitions and tools that are going to be necessary established by this thesis.

1.1.1 Generalities on Time Series Process

Definition 1.1. (Stochastic Process) Let $(\Omega, \mathcal{F}, P)$ be a probability space and let T be an index set, any collection of random variables $X = X_t : t \in T$ defined on $(\Omega, \mathcal{F}, P)$ is called a stochastic process with index set T.

Definition 1.2. (Time Series Process) Time series is a sequence taken at successive equally spaced points in time, thus it is a sequence of discrete-time data, it can be decomposed into unobservable components. In the most complete case, these components are the trend (T_t), the seasonal (S_t) and the irregular components (ε_t). We use a dataset of weekly mortality in Algeria caused by the Coronavirus as an example we refer the reader to the website World Health Organization

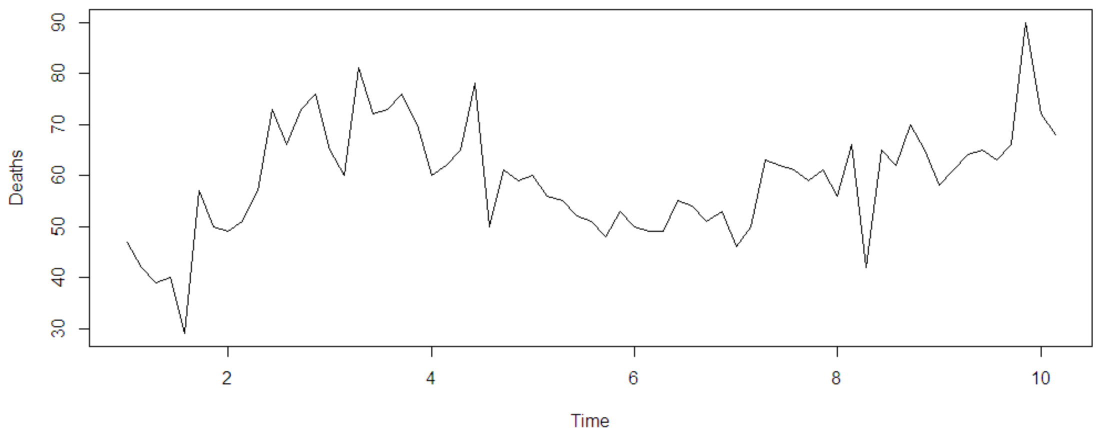

Figure 1.1: Weekly mortality in Algeria caused by the Coronavirus. from 2^{nd} March 2020 to 30^{th} May 2021

Now, we present the trend, seasonal, and random components of weekly mortality in Algeria caused by the Coronavirus dataset.

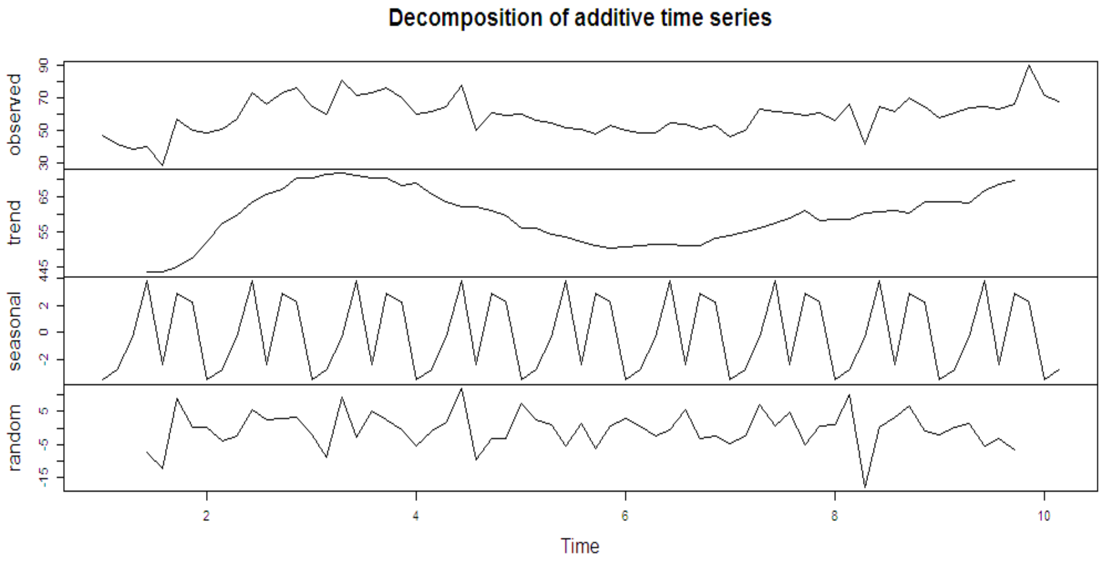

Figure 1.2: Decomposition of weekly mortality in Algeria caused by the Coronavirus

Additive and multiplicative models

The amplitude of both the seasonal and irregular variations do not change as the level of the trend rises or falls.

We have used dataset in R:"International airline passengers: monthly totals in thousands" to illustrate Additive and multiplicative models.

In such cases, an additive model is appropriate.

$$X_t = T_t + S_t + \varepsilon_t$$

The amplitude of both the seasonal and irregular variations increase as the level of the trend rises. In this situation, a multiplicative model is usually appropriate.

$$X_t = T_t \times S_t \times \varepsilon_t$$

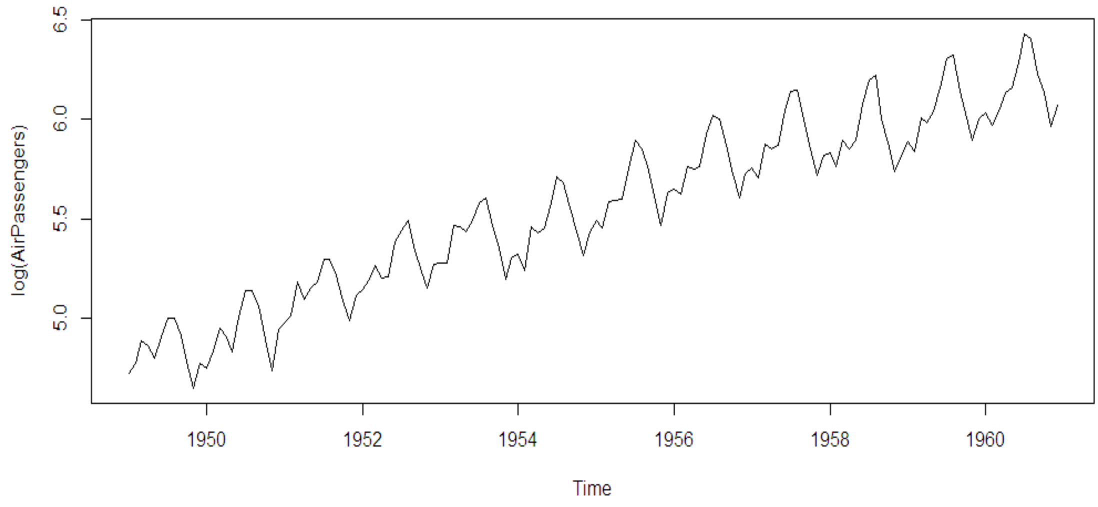

Figure 1.3: Log of Monthly totals of international airline passengers, 1949 to 1960.

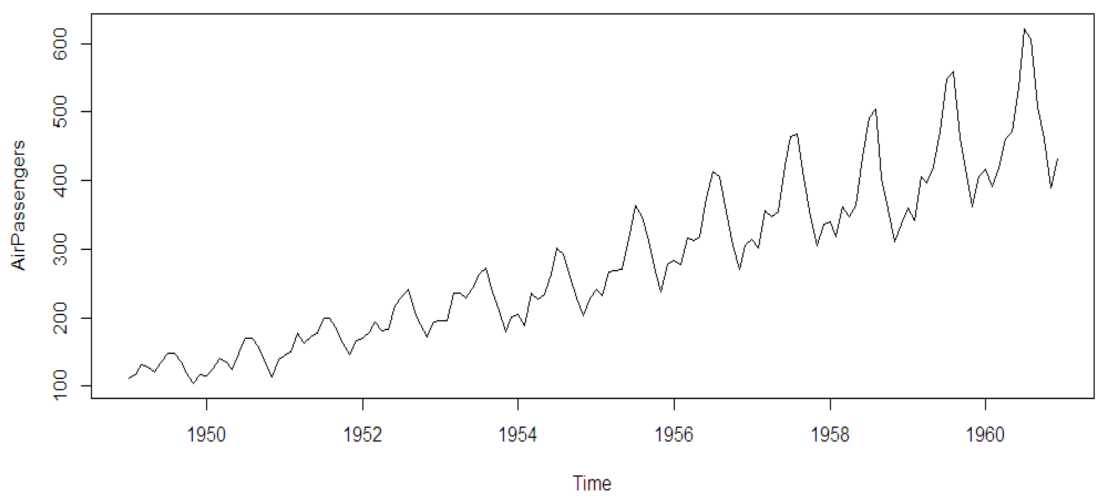

Figure 1.4: Monthly totals of international airline passengers, 1949 to 1960.

Definition 1.3. (Lag Operator) lag operator (L) or backshift operator operates on an element of a time series to produce the previous element.

$$LX_t = X_{t-1} \qquad L^j X_t = X_{t-j} \tag{1.1}$$

Stationarity, in the most basic sense, refers to the statistical features of a process that generates a time series that do not change over time.

Definition 1.4. (Strong Stationary Process)[63] Formally, the discrete stochastic process $\{X_t, t_1, ..., t_n \in \mathbb{R}\}$, $h \in \mathbb{R}$ is strongly stationary if

$$F_X(X_{t_1+h}, ..., X_{t_n+h}) = F_X(X_{t_1}, ..., X_{t_n}) \tag{1.2}$$

Where $n \in \mathbb{N}$ and $F_X(X_{t_1+h}, ..., X_{t_n+h})$ represent the cumulative distribution function of the unconditional joint distribution of $\{X_t\}$ at times $t_1 + h, \ldots, t_n + h$. Since h does not affect $F_X(\cdot)$, F_X is not a function of time.

Definition 1.5. (Weak stationary Process)[19] The process $\{X = x_t, t \in \mathbb{Z}\}$ is weakly stationary if the expected value and the variance of xt are constant over time and if the covariance between X_t and X_{t+h} (called autocovariance) depends on the time lag h only, where we assume these moments to exist. Then we write

- The first moment of x_t is constant, $\forall t, \mathbb{E}[X_t] = \mu$
- The second moment of x_t is finite, $\forall t, \mathbb{E}[X_t^2] < \infty$
- The autocovariance depends only on h, $\forall t, Cov(X_t, X_{t+h}) = \gamma(h)$

Example 1.1. We present two examples to clarify the stationary time series and non-stationary time series.
The first case is an Autoregressive process of order 1, AR(1) $(1 - 0.9L)Y_t = \varepsilon_t$ is weakly stationary time series.
The second case is an is an Autoregressive process of order 2, AR(2)
$(1 - 1.5L - 0.5L^2)Y_t = \varepsilon_t$ is a non stationary time series.

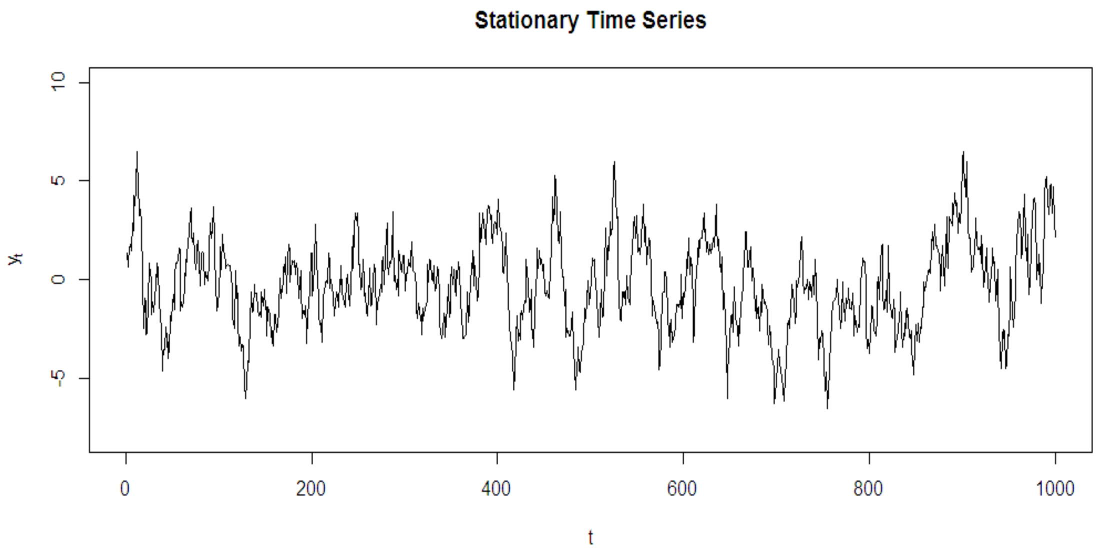

Figure 1.5: AR(1), $Y_t - 0.9Y_{t-1} = \varepsilon_t$

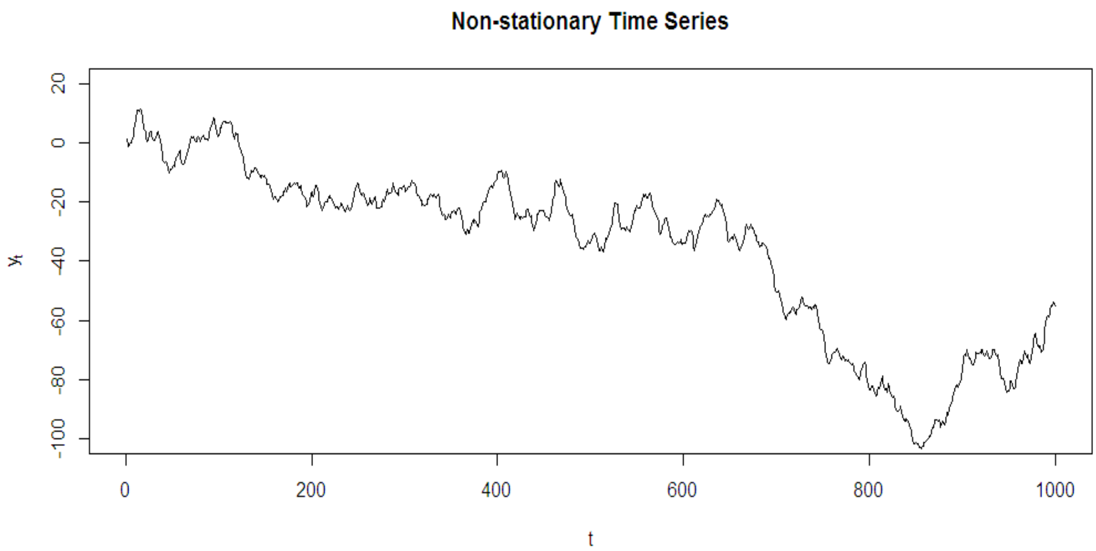

Figure 1.6: AR(2), $Y_t - 1.5Y_{t-1} - 0.5Y_{t-2} = \varepsilon_t$

1.1.2 Autocorrelation Properties

Definition 1.6. (The Autocovariance Function of a Stochastic Process) Let $\{X_t, t \in \mathbb{Z}\}$ be a stochastic process such that $Var(X_t) < \infty \ \forall t \in \mathbb{Z}$. The function

$$\gamma_X : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$$

defined by

$$\gamma_X(t_1, t_2) = Cov(X_{t_1}, X_{t_2}) \tag{1.3}$$

is called Autocovariance Function of the stochastic process X_t.
The autocovariance function describes the strength of the linear relationship between the random variables X_{t_1} and X_{t_2}.

$$Cov(X_{t_1}, X_{t_2}) = \mathbb{E}[(X_{t_1} - \mathbb{E}(X_{t_1}))(X_{t_2} - \mathbb{E}(X_{t_2}))]$$

Autocovariance function evaluated in (t, t) gives the variance, because

$$\gamma_x(t, t) = \gamma_X(0) = \mathbb{E}[(X_t - \mu_t)^2] = Var(X_t)$$

Where $\mathbb{E}(X_t) = \mu_t$

Definition 1.7. (The Autocorrelation Function of a stationary process) [39] Consider a weakly stationary stochastic process $X_t, t \in \mathbb{Z}$, we have that

$$\gamma_X(t + k, t) = Cov(X_{t+k}, X_t) = \gamma_X(k) \quad \forall t, k \in \mathbb{Z}$$

We observe that does not depend on t $\gamma_x(t + k, t)$. It depends only on the time difference k.
The function $\gamma_X(k)$ is called autocovariance function of the weakly stationary stochastic process $X_t, t \in \mathbb{Z}$
Autocorrelation function is defined by

$$\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)} \tag{1.4}$$

1.1.3 Mathematical Tools

Here we gather some mathematical tools (Special functions, inequalities and convergence lemmas and theorems.).

Definition 1.8. (Gamma Function)[1] The Gamma function is defined as follows

$$\Gamma(z+1) = \int_0^\infty z^a e^{-z} dt$$

The Gamma function is an analogue of factorial for non-integers. The Gamma function satisfies the functional equations

$$\Gamma(z) = (z-1)!$$

$$\Gamma(1+z) = z\Gamma(z)$$

$$\Gamma(1-z) = -z\Gamma(-z)$$

$$\frac{\Gamma(-z+j)}{\Gamma(-z)} = (-1)^j \frac{\Gamma(z+j)}{\Gamma(z-j+1)} \text{for} j = 1, 2, 3...$$

Definition 1.9. (Some important formulas)[1] Hypergeometric series were studied by Leonhard Euler, but the first full systematic treatment was given by Carl Friedrich Gauss (1813).

The hypergeometric function is defined for $|z| < 1$, we can write it as follows

$$F(a, b, c, z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)j!} z^j \tag{1.5}$$

$$F(a, b, c, 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \tag{1.6}$$

$$F(-\delta, b, c, \alpha z) = (1-\alpha z)^\delta \tag{1.7}$$

Sheppard formula

$$\frac{\Gamma(j+a)}{\Gamma(j+b)} \approx j^{a-b}$$

Definition 1.10. (Fisher Information)[82] Fisher information informs us how much information we can gather from a sample about an unknown parameter.
To put it another way, it informs us how well we can measure a parameter given a set of data.
the partial derivative with respect to θ of the natural logarithm of the likelihood function is called the score. Under certain regularity conditions ($f(x;\theta)$ are continuously differentiable). The Fisher information is defined to be the variance of the score

$$I(\theta) = \mathrm{E}\left[\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2 \middle| \theta\right] = \int_{\mathbb{R}} \left(\frac{\partial}{\partial\theta}\log f(x;\theta)\right)^2 f(x;\theta)\,dx, \tag{1.8}$$

Proposition 1.1. *(Markov inequality)([82]) If X is any nonnegative random variable, then*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}, \tag{1.9}$$

for any $a > 0$

Proposition 1.2. *If X is any random variable, then for any $a > 0$ we have*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{Var(X)}{a^2}$$

We generalize the previous inequality

Proposition 1.3. *(Tchebychev inequality)([54]) If X is any random variable, then for any $a > 0$ we have*

$$\mathbb{P}\left(|X| \geq a\right) \leq \frac{1}{a^p}\mathbb{E}\left(|X|^p\right) \tag{1.10}$$

1.1.4 Convergence Theorems

There are different definitions of random variable convergence, a sequence of random variables follows a fixed behavior when repeated for a large number of times. This concept is a crucial in statistics and stochastic process.

Proposition 1.4. *(Convergence in Distribution)([71]) A sequence $X_1, X_2, \ldots$ of real-valued random variables is said to converge in distribution, or converge weakly, or converge in law to a random variable X if*

$$\lim_{n\to\infty} F_n(X) = F(X), \lim_{n\to\infty} F_n(X) = F(X) \tag{1.11}$$

Convergence in distribution is denoted by $X_n \to^d X$
F_n and F are the cumulative distribution functions of random variables X_n and X, respectively.

Proposition 1.5. *(Convergence in Probability)([71]) A sequence X_n of random variables converges in probability towards the random variable X if for all $\varepsilon > 0$*

$$\lim_{n\to\infty} P\Big(|X_n - X| > \varepsilon\Big) = 0. \lim_{n\to\infty} P\Big(|X_n - X| > \varepsilon\Big) = 0. \tag{1.12}$$

Convergence in probability is denoted by $X_n \to^P X$
Convergence in probability implies convergence in distribution.

Proposition 1.6. *(L^p Spaces) ([54] P 119) Given a measure space $(\Omega, \mathcal{F}, \mu)$. For $1 \le p < \infty$, we consider the set of all measurable functions from Ω to $\mathbb{C}$ or $\mathbb{R}$ whose absolute value raised to the p^{th} power has a finite integral, or equivalently, that*

$$\|f\|_p \equiv \left(\int_\Omega |f|^p \, \mathrm{d}\mu\right)^{1/p} < \infty \tag{1.13}$$

For random variable X, we can define it in the following way

$$\mathbb{E}(|X|^p)^{\frac{1}{p}} < \infty \tag{1.14}$$

Theorem 1.7. *(The Monotone Convergence Theorem for Random Variables)([54]) Let $(X_n)_n$ be random variables such that $X_n \geq 0$ for all n and X_n coverges to X as $n \to \infty$ a.s. Then*

$$\mathbb{E}[X_n] = \mathbb{E}[X] \quad \text{as} \quad n \to \infty. \tag{1.15}$$

Lemma 1.1. *(Fatou's Lemma for Random Variables)([54]) Let Y be a random variable that satisfies $E[|Y|] < \infty$. Then the following holds,*

- *If* $Y \leq X_n$ *for all* n*, then* $\mathbb{E}(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} \mathbb{E}(X_n)$

- *If* $Y \geq X_n$ *for all* n*, then* $\mathbb{E}(\limsup_{n\to\infty} X_n) \geq \limsup_{n\to\infty} \mathbb{E}(X_n)$

Definition 1.11. (Stochastic o and O symbols)[82] For a given sequence of random variables R_n,

$$X_n = o(R_n) \quad \text{means} \quad X_n = Y_n R_n \quad \text{and} \quad Y_n \xrightarrow{P} 0,$$

$$X_n = O(R_n) \quad \text{means} \quad X_n = Y_n R_n \quad \textit{and} \quad Y_n = O(1).$$

There are many rules of calculus with o and O symbols, which we apply without comment. For instance

$$o(1) + o(1) = o(1),$$

$$o(1) + O(1) = O(1),$$

$$o(1)O(1) = o(1),$$

$$(1 + o(1))^{-1} = O(1),$$

$$o(R_n) = R_n o(1),$$

$$O(R_n) = R_n O(1),$$

$$o(O(1)) = o(1).$$

1.1.5 Mixing Variables

Mixing conditions are usual structures for modeling dependence for a sequence of random variables, this notion is defined in the following way

Definition 1.12. (Mixing Processes)([68]) Mixing conditions are usual structures for modeling dependence for a sequence of random variables. Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{B}$ and $\mathcal{C}$ be two sub $\sigma-$field of $\mathcal{F}$.
In order to estimate the correlation between $\mathcal{B}$ and $\mathcal{C}$ various a coefficient are used

(1) $\alpha = \alpha(\mathcal{B}, \mathcal{C}) = \sup_{B\in\mathcal{B}, C\in\mathcal{C}} \mid P(B \cap C) - P(B) \cap (C) \mid,$

(2) $\beta = \beta(\mathcal{B},\mathcal{C}) = \sup_{C\in\mathcal{C}} \mid P(C) - P(C|B) \mid,$

(3) $\varphi = \varphi(\mathcal{B},\mathcal{C}) = \sup_{B\in\mathcal{B},P(B),C\in\mathcal{C}} \mid P(B\cap C) - P(B)\cap(B) \mid,$

(4) $\rho = \rho(\mathcal{B},\mathcal{C}) = \sup_{X\in L^2(\mathcal{B}),X\in L^2(\mathcal{C})} \mid corr(X,Y) \mid.$ where $corr$ is correlation.

Proposition 1.8. *These coefficient satisfy the following inequalities*

$$2\alpha \leq \beta \leq \varphi,$$

$$4\alpha \leq \rho \leq 2\varphi^{\frac{1}{2}}.$$

Then

$$\varphi\text{-mixing} \Rightarrow \beta\text{-mixing} \Rightarrow \alpha\text{-mixing},$$

$$\varphi\text{-mixing} \Rightarrow \rho\text{-mixing} \Rightarrow \alpha\text{-mixing}.$$

Proof

For more details see [27].

Remark 1.1. **(1)** $0 \leq \alpha(\mathcal{B},\mathcal{C}) \leq \dfrac{1}{4}$

(2) $0 \leq \beta(\mathcal{B},\mathcal{C}) \leq 1$

(3) $0 \leq \varphi(\mathcal{B},\mathcal{C}) \leq \infty$

(4) $0 \leq \rho(\mathcal{B},\mathcal{C}) \leq 1$

In this work we use the α-mixing (or strong mixing) notion, which is one of the most treated among the different mixing structures introduced in the literature.

Definition 1.13. A process $(X_t, t \in \mathbb{Z})$ is said to be $\alpha-$mixing if

$$\alpha(k) = \sup_{B\in\mathcal{B},C\in\mathcal{C}} \mid P(B\cap C) - P(B)\cap(C) \mid \quad k \geq 1$$

Where $\mathcal{B} = \sigma(X_s, s \leq t)$, $\mathcal{B} = \sigma(X_s, s \geq t+k)$ and

$$\lim_{k\to\infty} \alpha(k) = 0.$$

Definition 1.14. (Geometrical and arithmetical α-mixing)([29] P 154) We will mainly consider both of the following subclasses of mixing processes.
The sequence $X_n, n \in \mathbb{Z}$, is said to be arithmetically α-mixing with rate $a > 0$ if

$$\exists C > 0, \quad \alpha(k) \leq Ck^{-a}$$

It is called geometrically α-mixing if

$$\exists C > 0, \quad \exists t \in (0,1) \quad \alpha(k) \leq Ct^k$$

Theorem 1.9. *(Davydov's inequality)([15]) Let X and Y be two real valued random variables such that $X \in L^q(P), Y \in L^r(P)$ where $q > 1, r > 1$ and $\frac{1}{q}+\frac{1}{r} = 1-\frac{1}{p}$ then*

$$|Cov(X,Y)| \leq 2p(2\alpha)^{\frac{1}{p}} \parallel X \parallel_q \parallel Y \parallel_r \tag{1.16}$$

1.2 ARIMA Processes

To model a phenomenon, Box and Jenkins provide processes. We have gone over the most important time series models.

Definition 1.15. [16] A stochastic process $(X_t)_{t\geq 0}$ is said to be an ARIMA(p,d,q) an integrated mixture autoregressive moving average model if it satisfies the following equation

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \quad \forall t \geq 0 \tag{1.17}$$

Where $d \in \mathbb{N}$, L is lag operator, $\varepsilon_t \sim \mathcal{N}(0,\sigma^2)$ i.i.d. errors, with $\sigma^2 < \infty$.
$\phi(L) = (1 - \phi_1 L - ... - \phi_p L^p)$ with $\phi_p \neq 0$
$\theta(L) = (1 - \theta_1 L - ... - \theta_q L^q)$ with $\theta_q \neq 0$

In the case of $d = 0$, we obtain ARMA(p,q) process

$$\phi(L)X_t = \theta(L)\varepsilon_t \quad \forall t \geq 0 \tag{1.18}$$

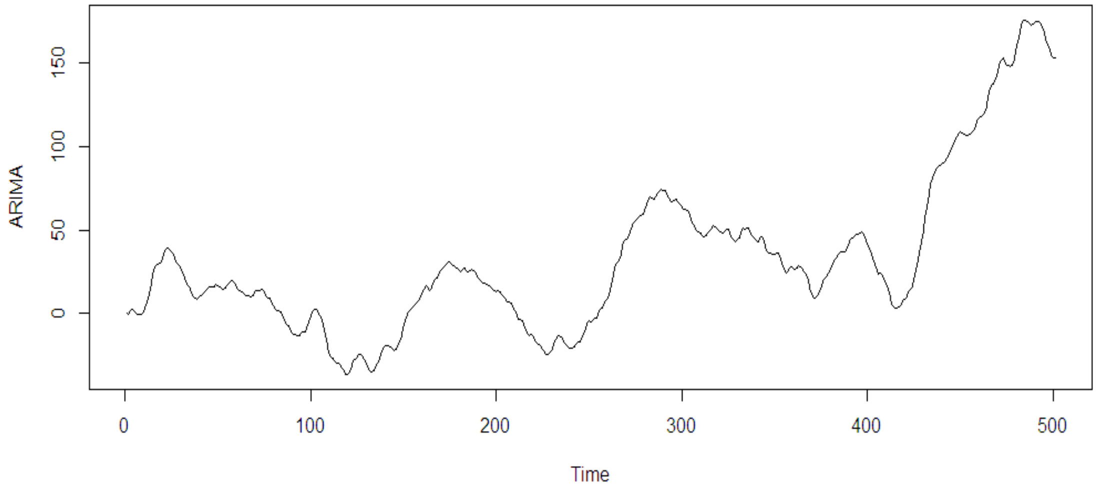

Figure 1.7: ARIMA(1,1,1), $\phi_1 = 0.7$, $\theta_1 = 0.8$

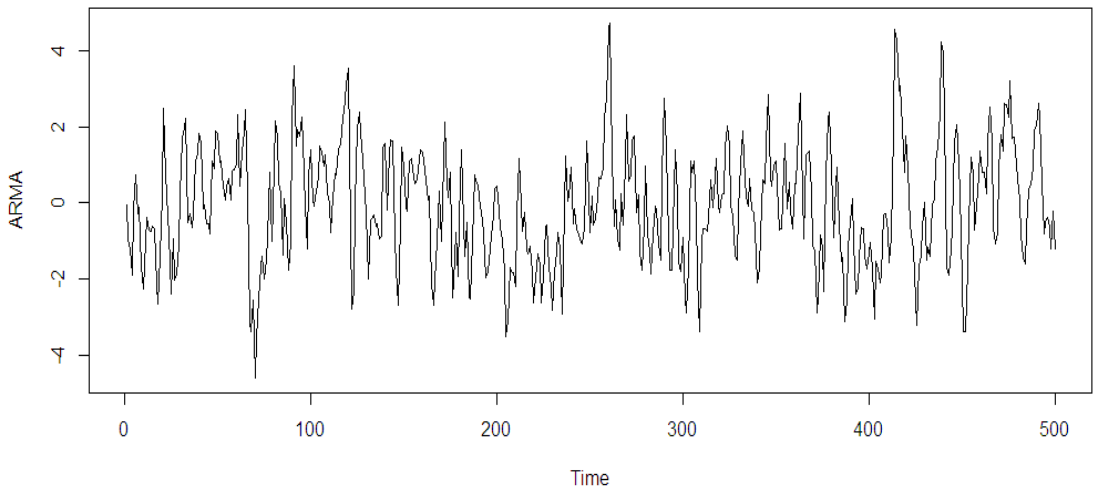

Figure 1.8: ARMA(1,1), $\phi_1 = 0.5$, $\theta_1 = 0.2$

Autoregressive order p is written as follows

$$\phi(L)X_t = \varepsilon_t \tag{1.19}$$

Moving average of order q

$$X_t = \theta(L)\varepsilon_t \tag{1.20}$$

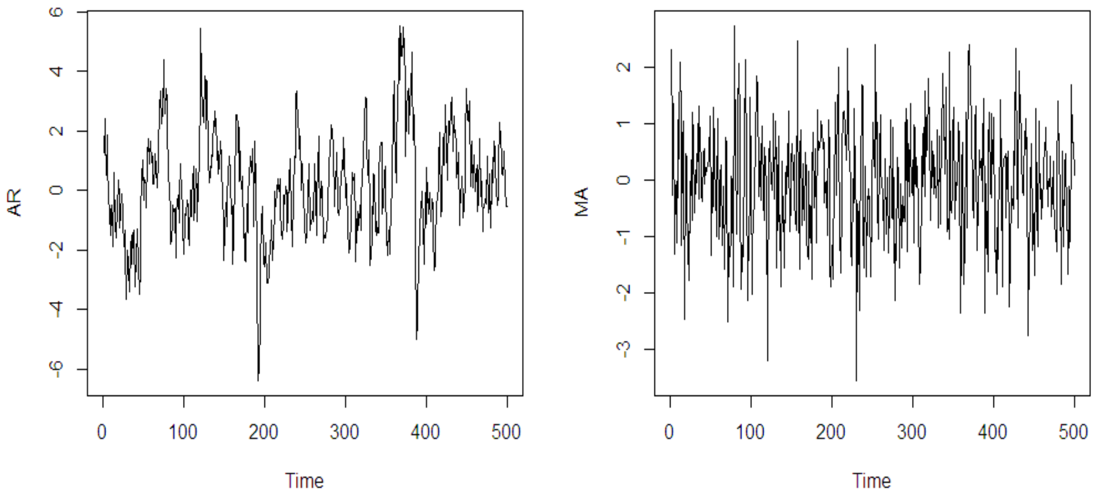

Figure 1.9: AR(1),$\phi_1 = 0.8$, MA(1), $\theta_1 = 0.3$

1.2.1 Causality and Invertibility Conditions

Causality of a stationary time series indicates that the time series is dependent on past values. Essentially, ARMA(p, q) model can be written in the following form

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \tag{1.21}$$

Where $\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \dfrac{\theta(z)}{\phi(z)}, \quad |z| < 1$

This would give rise to the property of causality, where sum of the coefficients of

the infinite MA expressions is finite $\sum_{j=0}^{\infty} |\psi_j| < \infty$.
For MA models, there is the element of non-uniqueness in their representation regarding the autocorrelation function, can be the same for equations of various MA coefficients and variance, and thus it is hard to distinguish between them.
Invertibility comes into play when one should pick the best representation by making ε_t the subject and expressing the time series in an infinite AR representation.

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \tag{1.22}$$

where $\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \dfrac{\phi(z)}{\theta(z)}, \quad |z| < 1$
It is only invertible where the infinite sum of the coefficients of the infinite AR expression is finite. $\sum_{j=0}^{\infty} |\pi_j| < \infty$
MA(q) process is causal by definition because $\sum_{i=0}^{q} \theta_i^2 < \infty$
MA(q) process is invertible if and only if $\theta_z = \sum_{j=0}^{q} \theta_j z^j = 0$, for all $z \in \mathbb{C}$ and $z > 1$.
AR(p) is causal if and only if $\phi_z = \sum_{j=0}^{q} \phi_j z^j = 0$, for all $z \in \mathbb{C}$ and $z > 1$
AR(p) is invertible by definition because $\sum_{i=0}^{q} \phi_i^2 < \infty$
These models are called short memory processes, which means that the shock disappears quickly, in other words the sum of autocorrelation is finite.
We present the example of an autocorrelation function of ARMA process. It is clear that the process has a fast decay.

1.3 Properties of ARFIMA Processes

ARFIMA(p, d, q) [35] is defined in the following equation

$$\phi(L)(1 - L)^d X_t = \theta(L)\varepsilon_t \quad \forall t \geq 0 \tag{1.23}$$

L is lag operator, $d \in \mathbb{R}$, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ i.i.d. errors, with $\sigma^2 < \infty$

$$(1 - L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j - d)}{\Gamma(-d)\Gamma(j + 1)} L^j \tag{1.24}$$

Where $\Gamma(.)$ is gamma function.

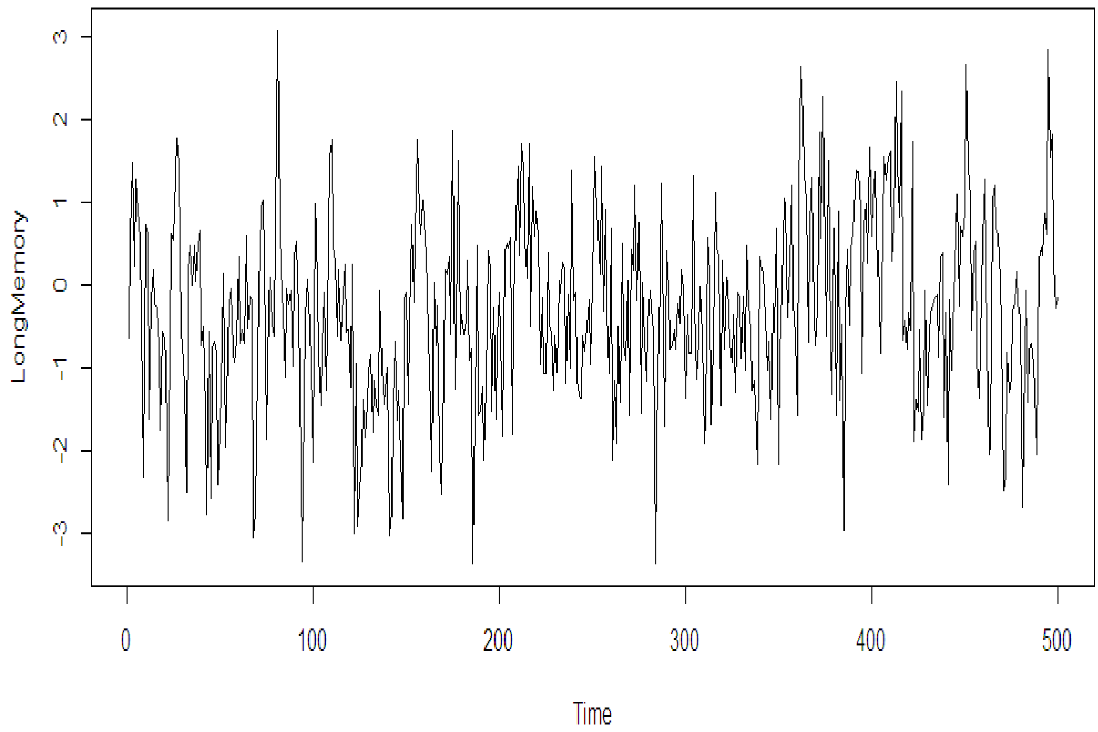

Figure 1.10: FAR(1), $d = 0.3$

The stochastic process X_t is both stationary and invertible if all roots of $\theta(L)$ and $\phi(L)$ lie outside the unit circle and $|d| < 0.5$, The process is non-stationary for $d \geq 0.5$, The process exhibits short memory for $d = 0$, corresponding to stationary and invertible ARMA modeling.
In this case, ARFIMA(p, d, q) we have

$$X_t = (1-L)^{-d}\frac{\theta(L)}{\phi(L)}\varepsilon_t \tag{1.25}$$

and

$$\varepsilon_t = (1-L)^{d}\frac{\phi(L)}{\theta(L)}X_t \tag{1.26}$$

1.3.1 Autocorrelation Function of long memory model

We deal with Fractional Autoregressive process of order 1 (long memory process)

$$(1-L)^d X_t = \varepsilon_t, \tag{1.27}$$

where $|d| < \dfrac{1}{2}$ and ε_t i.i.d errors and normally distributed $\mathcal{N}(0,\sigma^2)$, $\sigma^2 < \infty$

Proposition 1.10. *The model is causal, if $d < \dfrac{1}{2}$ It can be stated as follows*

$$X_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \tag{1.28}$$

Where $\psi_i = \dfrac{\Gamma(d+i)}{\Gamma(d)\Gamma(i+1)}$

Conditions of Causality [33]

1. $d > 0$, $\quad \sum_{i=0} |\psi_i| < \infty$
2. $-\dfrac{1}{2} < d < 0 \quad \sum_{i=0} |\psi_i| = \infty$ and $\sum_{i=0} \psi_i^2 < \infty$
3. $d \leq -\dfrac{1}{2} \quad \sum_{i=0} |\psi_i| = \infty = \sum_{i=0} \psi_i^2 = \infty$

Proposition 1.11. *The model is invertible, if $-\dfrac{1}{2} < d$ it can be written as follows*

$$\varepsilon_t = \sum_{i=0}^{\infty} \pi_i X_{t-i} \tag{1.29}$$

Where $\pi_i = \dfrac{\Gamma(-d+i)}{\Gamma(-d)\Gamma(i+1)}$

Conditions of Invertibility [33]

1. $d < 0$, $\quad \sum_{i=0} |\pi_i| < \infty$
2. $0 < d < \frac{1}{2}$ $\quad \sum_{i=0} |\pi_i| = \infty$ and $\sum_{i=0} \pi_i^2 < \infty$
3. $d \geq \frac{1}{2}$ $\quad \sum_{i=0} |\pi_i| = \infty = \sum_{i=0} \pi_i^2 = \infty$

Proposition 1.12.

$$\gamma_X(h) = \sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} F(d, d+h, h+1, 1), \qquad h \geq 0 \tag{1.30}$$

Proof

$$\begin{aligned}
\gamma_X(h) &= \mathbb{E}([\sum_{i\geq 0} \psi_i \varepsilon_{t-i}][\sum_{i\geq 0} \psi_i \varepsilon_{t+h-i}]) \\
&= \sum_{i\geq 0} \psi_i^2 \mathbb{E}(\varepsilon_{t-i}, \varepsilon_{t+h-i}) + \sum_{j\geq 0} \sum_{i\geq 0} \psi_i \psi_j \mathbb{E}(\varepsilon_{t-j}, \varepsilon_{t+h-i})
\end{aligned}$$

$\mathbb{E}(\varepsilon_{t-i}, \varepsilon_{t+h-i}) = 0$

$\mathbb{E}(\varepsilon_{t-j}, \varepsilon_{t+h-i}) \begin{cases} \sigma^2, & j = i - h; \\ 0, & j \neq i - h. \end{cases}$

We replace ψ_i by its value.

$$\begin{aligned}
\gamma_X(h) &= \sigma^2 \sum_{i\geq 0} \psi_i \psi_{i+h} \\
&= \sigma^2 \sum_{i\geq 0} \frac{\Gamma(d+i+h)}{\Gamma(d)(h+i)!} \frac{\Gamma(d+i)}{\Gamma(d)i!} \\
&= \left[\sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)}\right] \left[\frac{\Gamma(h+1)}{\Gamma(h+d)\Gamma(d)} \sum_{i\geq 0} \frac{\Gamma(d+h)\Gamma(d+h+i)}{\Gamma(h+1+i)i!}\right] \\
&= \sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} F(d, h+d, h+1, 1)
\end{aligned}$$

Using 1.6, the autocovariance function $\gamma_X(h)$ converges

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

when $t \to \infty$

$$\gamma_X(0) = \sigma^2 F(d, d, 1, 1) \tag{1.31}$$

$$\gamma_X(h) \simeq \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)} h^{2d-1} \tag{1.32}$$

Autocorrelation function

$$\rho_X(h) \simeq \frac{\frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}}{F(d, d, 1, 1)} h^{2d-1} \tag{1.33}$$

$$\rho_X(h) \simeq C h^{2d-1} \tag{1.34}$$

where C is constant.

The autocovariance function describes the behavior of FAR(1) long memory model, this function decreases hyperbolically.

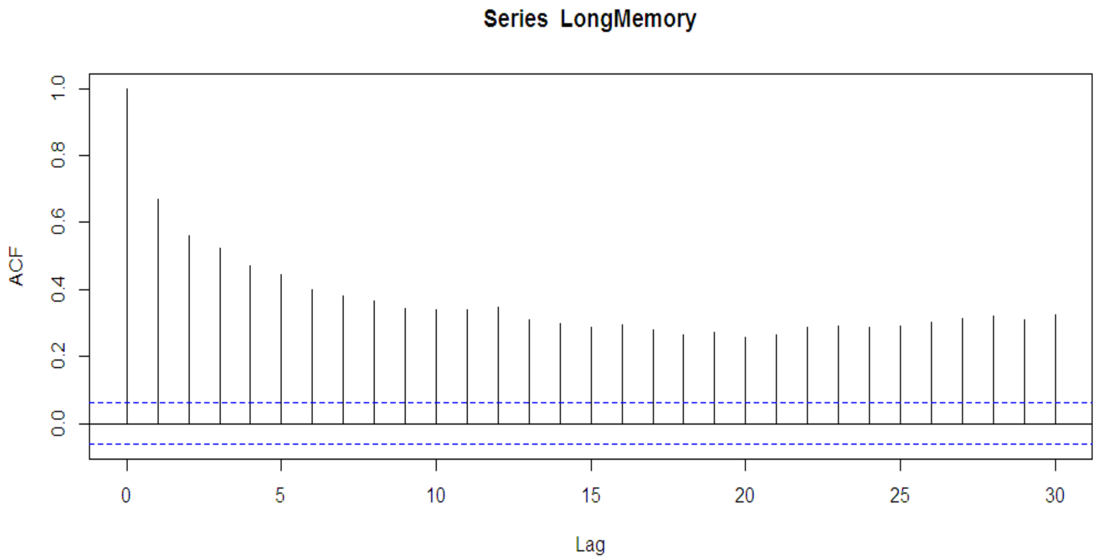

Figure 1.11: FAR(1), $d = 0.4$

1.4 Preliminaries On Local Asymptotic Normality (LAN) Property

This section introduces some fundamental concepts and definitions that will help you deal with the Local Asymptotic Normality (LAN) property for time series, we focus on ARFIMA model especially fractional Autoregressive process of order 1, FAR(1).

1.4.1 Converegence of Statistical Experiment

The statistical experiment is defined $\{P_\theta, \theta \in \Theta \quad \subset \mathbb{R}^k\}$ on sample space $(\Omega, \mathcal{F})$, the full observation is a single observation from the product $\{P_\theta^n\}$ of n copies. it can be approximated by Gaussian experiments after a suitable reparametrization (localization of the parameters).
We define a local parameter $h = \sqrt{n}(\theta - \theta_0)$ thus we obtain $P^n_{\theta_0 + h/\sqrt{n}}$ an experiment with parameter h have the same statistical properties for normal distribution, it consists of observing a single observation from a normal distribution with mean h and known covariance matrix equal to the inverse of the Fisher information matrix $\mathcal{N}(h, I_\theta^{-1})$.
We use Taylor expansion of the logarithm of the likelihood.
p_θ is density of P_θ, $\log p_\theta$ is twice differentiable

$$\log \prod_{i=1} \frac{p^n_{\theta_0+h/\sqrt{n}}}{p_{\theta_0}}(X_i) \quad = \quad \frac{h}{\sqrt{n}} \sum_{i=1}^{n} \frac{d \log p_\theta(X_i)}{d\theta} + \frac{h^2}{2n} \sum_{i=1}^{n} \frac{\partial^2 \log p_\theta(X_i)}{\partial \theta^2} + o_{X_i}(h^2)$$

the first term can be rewritten as $h\Delta_{n,\theta}$ is asymptotically normal with mean zero and variance I_θ , by the central limit theorem. Furthermore, the second term in the expansion is asymptotically equivalent to $-\frac{1}{2}h^2 I_\theta$, by the law of large numbers.

$$\log \prod_{i=1} \frac{p^n_{\theta_0+h/\sqrt{n}}}{p_{\theta_0}}(X_i) \quad = \quad h\Delta_{n,\theta} - \frac{1}{2}h^2 I_\theta + o_{P_\theta}(1)$$

this expansion concerns the likelihood process in a neighborhood of θ , we speak of "local asymptotic normality" of the sequence of models $\{P_\theta^n : \theta \in \Theta\}$.

On other side we establish a direct relationship between the local experiments and a normal limit experiment.

$$\log \frac{d\mathcal{N}(h, I_\theta^{-1})}{d\mathcal{N}(0, I_\theta^{-1})}(X_i) = h^t I_\theta X_i - \frac{1}{2} h^t I_\theta h \tag{1.35}$$

Fixing θ, we get $P_{\theta+h/\sqrt{n}}, h \in \mathbb{R}^k$ as a statistical model with parameter h, for "known" (θ). We show that this can be approximated by the statistical model $\mathcal{N}(h, I_\theta^{-1}, h \in \mathbb{R}^k)$.
Due to this expansion concerns the likelihood process in a neighborhood of θ and the approximation to normal distribution, we speak about (Local Asymptotic Normality).

Definition 1.16. [82](Local Asymptotic Normality) A sequence of parametric statistical models $P_{n,\theta}, \theta \in \Theta$ is said to be locally asymptotically normal (LAN) at θ if there exist matrices r_n and I_θ and a random vector $\Delta_{n,\theta} \sim \mathcal{N}(h, I_\theta)$

$$\log \frac{dP_{n,\theta+r_n^{-1}h}}{dP_{n,\theta}} = h^{'} \Delta_{n,\theta} - \frac{1}{2} h^{'} I_\theta \, h + o_{P_{n,\theta}}(1) \tag{1.36}$$

1.4.2 Le Cam and Swensen results on LAN property

Let $P_{0,n}$ and $P_{1,n}$ be two sequences of probability measures on the measurable spaces $(\Omega_n, \mathcal{F}_n)$ Suppose that for each n there is a filtration $\mathcal{F}_{n,k} \subset \mathcal{F}_{n,k+1}$ of σ-algebras with $\mathcal{F}_{n,k_n} = \mathcal{F}_n$, let $P_{0,n,k}$ and $P_{1,n,k}$ be the restriction of $P_{0,n}$ and $P_{1,n}$ respectively.
We consider

$$y_{n,k} = \left(\frac{\gamma_{n,k}}{\gamma_{n,k+1}}\right)^{\frac{1}{2}} - 1$$

Where $\gamma_{n,k}$ the Radon-Nikodym derivative on $\mathcal{F}_{n,k}$ of the part $P_{1,n,k}$.
Logarithm of likelihood function $\Lambda_n = \log \dfrac{dP_{1,n}}{dP_{0,n}}$ under $\mathcal{F}_n$ is defined by

$$\Lambda_n = 2 \sum_k \log(y_{n,k} + 1)$$

Theorem 1.13. *(Le Cam)[58] Assume the following conditions are satisfied, all convergence being in probability under $P_{0,n}$*

- *(i1)* $\max\limits_k y_{n,k} \to^P 0$
- *(i2)* $\sum\limits_k |y_{n,k}^2| \to^P \dfrac{(\lambda^{(n)})^2}{4}$
- *(i3)* $\sum\limits_k \mathbb{E}(y_{n,k}^2 + 2y_{n,k}/\mathcal{F}_{n,k-1}) \to^P 0$
- *(i4)* $\sum\limits_k \mathbb{E}\{(y_{n,k}^2 \mathbb{1}_{(|y_{n,k}|>\delta)}/\mathcal{F}_{n,k-1}) \to^P 0\}$ *for* $\delta > 0$

Thus

$$\Lambda_n = \mathcal{N}\left(-\frac{(\lambda^{(n)})^2}{2}, (\lambda^{(n)})^2\right) \tag{1.37}$$

We replace $y_{n,k}$ by martingale differences to prepare the way for the version of Swensen.

Definition 1.17. (Martingale Differences)[43] A stochastic series X is an Martingale Differences Series if its expectation with respect to the past is zero. Formally, consider an adapted sequence $\{X_t, \mathcal{F}_t\}_{-\infty}^{\infty}$ on a probability space $(\Omega, \mathcal{F}, P) X_t$ is an Martingale Differences Series if it satisfies the following two conditions

- $\mathbb{E}\,|X_t| < \infty$,
- $\mathbb{E}\,[X_t|\mathcal{F}_{t-1}] = 0, a.s.$

For all t.
By construction, this implies that if Y_t is a martingale, then $X_t = Y_t - Y_{t-1}$ will be a martingale differences series.

Now, we present Swensen's conditions in the following Lemma

Lemma 1.2. *(Swensen)[79] We consider* $\{Z_{n,t}, \mathcal{F}_{n,t}\}$ *which satisfy*

- *(C1)* $\sum\limits_{t=1}^{n} \mathbb{E}(Z_{n,t} - y_{n,t})^2 \longrightarrow 0$ *when* $n \to \infty$
- *(C2)* $\sup\limits_n \sum\limits_{t=1}^{n} \mathbb{E}(Z_{n,t})^2 < \infty$
- *(C3)* $\max\limits_{t \le n} |Z_{n,t}| \longrightarrow 0$ *under* P_θ *when* $n \to \infty$

- *(C4)* $\sum_{t=1}^{n} Z_{n,t}^2 - \frac{(\lambda)^2}{4} \longrightarrow 0$ *under* P_θ *when* $n \to \infty$
- *(C5)* $\sum_{t=1}^{n} \mathbb{E}(Z_{n,t}^2 \mathbb{1}_{Z_{n,t} \geq \frac{1}{2}} / \mathcal{F}_{n,t}) \longrightarrow 0$ *under* P_θ *when* $n \to \infty$
- *(C6)* $\sum_{t=1}^{n} \mathbb{E}(Z_{n,t}/\mathcal{F}_{n,t}) = 0$ *a.s*

P_θ is under $P_{0,n}$, thus $(i1)$, $(i2)$ and $(i4)$ of theorem above are satisfied as well as $(i3)$

$$\Lambda_n = \mathcal{N}\left(-\frac{(\lambda)^2}{2}, (\lambda)^2\right) \tag{1.38}$$

We can replace $(C4)$ and $(C5)$ by the following conditions

- $\sum_{t=1}^{n} \mathbb{E}(Z_{n,t}^2/\mathcal{F}_{n,t-1}) - \frac{(\lambda)^2}{4} \longrightarrow 0$ under P_θ when $n \to \infty$
- $\sum_{t=1}^{n} Z_{n,t}^2 \longrightarrow \frac{(\lambda)^2}{4}$ under P_θ when $n \to \infty$

1.5 Local Asymptotic Normality Property for Time series

(LAN) property was widely studied by many researchers. First with Swenesen [79], Kreiss ([55], [56]), this criterion was projected on ARFIMA processes (see Serroukh [72]) and recently Haddad et Belaide [40] generalize some results on long memory process in the case of dependent noises, Amimour et Belaide [2] assumed the periodicity of the memory parameter.

In this section we treat (LAN) for Fractional autoregressive model of order 1, we consider the model

$$(1 - L)^d X_t = \varepsilon_t$$

Where ε_t are i.i.d, $-\frac{1}{2} < d < \frac{1}{2}$ to insure the causality and invertibility conditions. After the following deviation $d^{(n)} = d + n^{-\frac{1}{2}}\delta^{(n)}$, we obtain

$$(1 - L)^{d+n^{-\frac{1}{2}}\delta^{(n)}} X_t = \varepsilon_t(d^{(n)}) \tag{1.39}$$

Proposition 1.14. *Under the hypothesis of regularity, for all* $0 < d < \frac{1}{2}$ *and for all* $\delta^{(n)}$ *with* $\sup_n \| \delta^{(n)} \| < \infty$*, we have*
Local asymptotic quadratic decomposition

$$\Lambda_{f,d+n^{-\frac{1}{2}}\delta^{(n)}/d} = \delta^{(n)}\Delta_f^n - \frac{1}{2}\sigma^2 I_f(\delta^n)^2\frac{\pi^2}{6} + o_{p_d}(1) \tag{1.40}$$

Local Asymptotic Normality of the central sequence $\Delta_f^n(d)$

$$\Delta_f^n(d) \sim \mathcal{N}(0, \sigma^2 I_f \frac{\pi^2}{6}) \tag{1.41}$$

Proof

For a detailed proof see (Serroukh [72])
On the other hand, Haddad and Belaide treated the fractional autoregressive model of order 1 but they assumed the dependency of the errors, case of mixing errors, we consider the model 1.5 ε_t are assumed to be strong mixing.
After suitable deviation of the parameter, we get

$$(1-L)^{d+n^{-\frac{1}{2}}\delta^{(n)}} X_t = \varepsilon_t(d^{(n)}) \tag{1.42}$$

Proposition 1.15. *Under the hypothesis of regularity, for all* $0 < d < \frac{1}{2}$ *and for all* $\delta^{(n)}$ *with* $\sup_n \| \delta^{(n)} \| < \infty$*, we have*
Local asymptotic quadratic decomposition

$$\Lambda_{f,d+n^{-\frac{1}{2}}\delta^{(n)}/d} = \delta^{(n)}\Delta_f^n - \frac{1}{2}\sigma^2(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})(\delta^n)^2\frac{\pi^2}{6} + o_{p_d}(1) \tag{1.43}$$

Local Asymptotic Normality of the central sequence $\Delta_f^n(d)$

$$\Delta_f^n(d) \sim \mathcal{N}\left([(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{4}} \frac{\pi}{\sqrt{6}}, \sigma^2[(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}} \frac{\pi^2}{6} \right) \tag{1.44}$$

Proof

For a detailed proof see (Haddad and Belaide [40])

1.6 Conclusion

In this chapter, we have presented some important definitions, tools and other important theories which helps us to deal with different probabilistic and statistical properties. Firstly we have focused on the classical processes ARIMA, ARMA, AR, MA then we move to the long memory process Fractional ARIMA and Fractional AR, then we present some results on Local Asymptotic Normality, we have started with historical review on the concept and some useful definitions, then we have presented examples about the application of LAN technique on ARFIMA process, these illustrations gives us a clear road map to show the results obtained in the last section.

Chapter 2

Probabilistic Properties for Fractional Autoregressive process with mixing errors

Introduction

There is a vast literature that dealt with Fractional Autoregressive Integrated Moving Average (ARFIMA) models, generally the researchers assume the independence of the errors in order to facilitate the calculation and simplify the model, but in reality the dependency of the errors is a crucial assumption because it is more accurate to describe the behavior of real phenomena. The researchers start to be interested in the notions of the dependency, especially the notion of mixing.
In this chapter, we treat the case of Fractional Autoregressive model of order 1 with dependent errors, we deal with strong mixing errors (Geometrical and arithmetical mixing errors cases).

2.1 Properties of Fractional Autoregressive Process with Independent Errors

Definition 2.1. [33] We consider the following model

$$(1 - aL)^d X_t = \varepsilon_t \tag{2.1}$$

L is lag opertor, a, d are real numbers, in this case the errors $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ are assumed to be i.i.d., with $\sigma^2 < \infty$

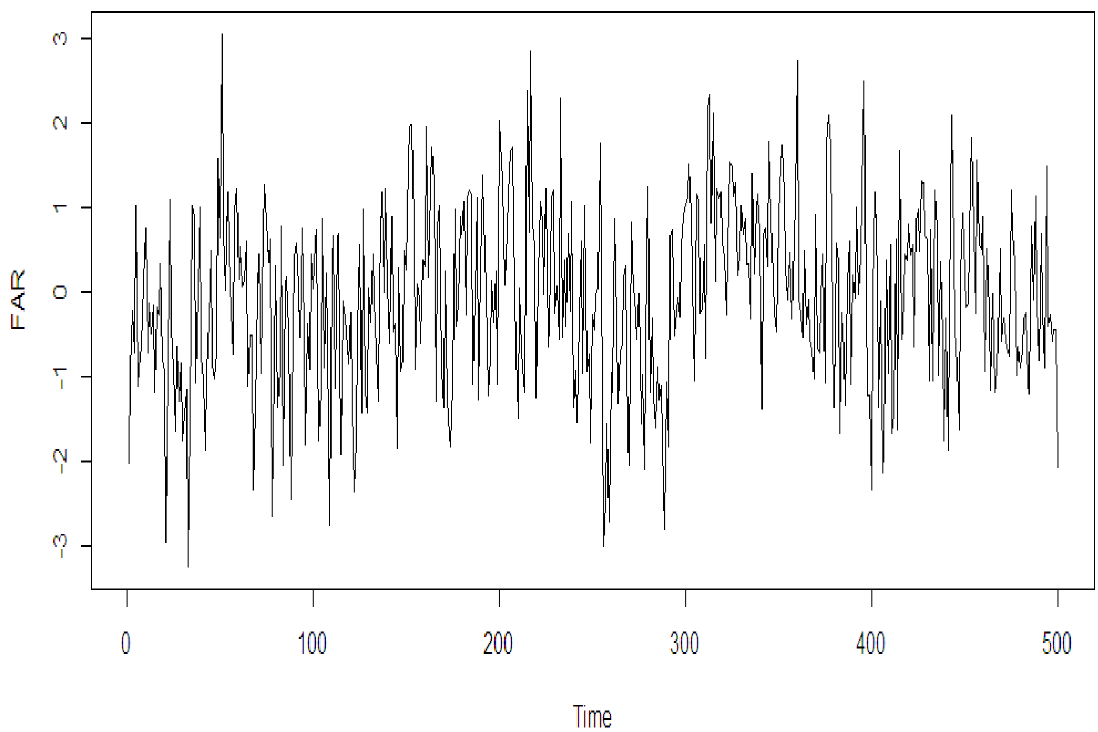

Figure 2.1: FAR(1), $a = 0.5$ and $d = 0.4$

Proposition 2.1. *[33] Let $a, d \in \mathbb{R}$, the process is invertible and causal if*

- $|a| < 1$ *and* $d \in \mathbb{R}$
- $a = 1 \; -\dfrac{1}{2} < d < \dfrac{1}{2}$

The model is stationary, invertible and causal, thus we can write

$$X_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \tag{2.2}$$

where $\psi_i = \dfrac{\Gamma(i+d)a^i}{\Gamma(d)\Gamma(i+1)} L^i$
and

$$\varepsilon_t = \sum_{i=0}^{\infty} \pi_i X_{t-i} \tag{2.3}$$

where $\pi_i = \dfrac{\Gamma(i-d)a^i}{\Gamma(-d)\Gamma(i+1)} L^i$

Proposition 2.2. *The autocovariance function $\gamma_X(.)$ of the process X_t is defined as follow*

$$\gamma_X(h) = \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h}, \qquad h \in \mathbb{Z} \tag{2.4}$$

Proof

Let $h \in \mathbb{Z}$, we have

$$\begin{aligned}
\gamma_X(h) &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-h} - \mathbb{E}(X_{t-h}))] \\
&= \mathbb{E}\left[\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} - \mathbb{E}(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i})\right)\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-h-i} - \mathbb{E}(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-h-i})\right)\right] \\
&= \mathbb{E}\left[\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \sum_{i=0}^{\infty} \psi_i \mathbb{E}(\varepsilon_{t-i})\right)\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-h-i} \sum_{i=0}^{\infty} \psi_i \mathbb{E}(\varepsilon_{t-h-i})\right)\right] \\
&= \mathbb{E}\left[\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}\right)\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-h-i}\right)\right] \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \mathbb{E}(\varepsilon_{t-i} \varepsilon_{t-h-j})
\end{aligned}$$

If $i = j + h$, we get $\mathbb{E}(\varepsilon_{t-i}\varepsilon_{t-h-j}) = \sigma^2$ and if $i \neq j + h$ we get $\mathbb{E}(\varepsilon_{t-i}\varepsilon_{t-h-j}) = 0$
Thus, we obtain

$$\gamma_X(h) = \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h}$$

Proposition 2.3. *The autocovariance function $\gamma_X(.)$ of the process X_t can be written in function of hypergeometric series*

$$\gamma_X(h) = \sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} a^h F(d, d+h, h+1, a^2), \qquad h \geq 0 \tag{2.5}$$

Where $F(.,.,.,.)$ is hypergeometric function.

Proof

Using the equation above

$$\begin{aligned}
\gamma_X(h) &= \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h} \\
&= \sigma^2 \sum_{i=0}^{\infty} \frac{\Gamma(i+d)}{\Gamma(d)} \frac{a^i}{\Gamma(i+1)} \frac{\Gamma(i+h+d)}{\Gamma(d)} \frac{a^{i+h}}{\Gamma(i+h+1)} \\
&= \sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)} \sum_{i=0}^{\infty} \frac{\Gamma(i+d)\Gamma(i+d+h)}{\Gamma(h+i+1)\Gamma(d)\Gamma(h+i)\Gamma(d)} \frac{\Gamma(h+1)}{\Gamma(h+d)} a^{2i+h} \\
&= \sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} a^h \sum_{i=0}^{\infty} \frac{\Gamma(i+d)\Gamma(i+d+h)}{\Gamma(h+i+1)\Gamma(h+i)\Gamma(d)} \frac{\Gamma(h+1)}{\Gamma(h+d)} a^{2i} \\
&= \sigma^2 \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} a^h F(d, d+h, h+1, a^2)
\end{aligned}$$

2.1.1 Asymptotic Behavior of Autocorrelation Function

Using 1.7 Sheppard formula [1], we obtain

$$\gamma_X(h) \simeq \frac{a^h}{\Gamma(d)} h^{d-1} (1-a^2)^{-d} \tag{2.6}$$

We know that the autocorrelation function is defined in the following form

$$\rho_X(h) = \frac{Cov(X_t, X_{t+h})}{Var(X_t)} \tag{2.7}$$

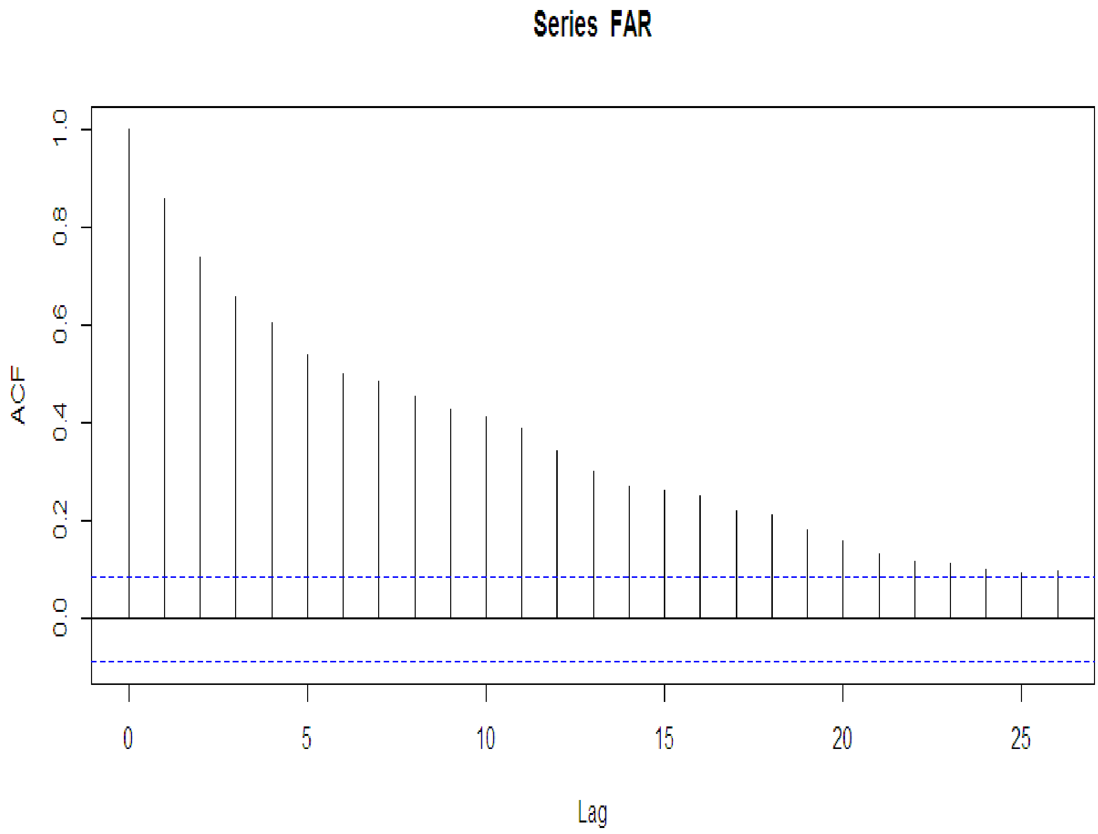

Figure 2.2: Auto-correlation function of FAR(1), $a = 0.9$ and $d = 0.8$

Using the proposition above

$$Var(X_t) = F(d, d, 1, a^2), \tag{2.8}$$

We get

$$\rho_X(h) \simeq \frac{\frac{a^h}{\Gamma(d)} h^{d-1}(1-a^2)^{-d}}{F(d, d, 1, a^2)} \tag{2.9}$$

$$\rho_X(h) \simeq C a^h h^{d-1} \tag{2.10}$$

where C is constant.

It is clear that the autocorrelation function has the part of hyperbolic decay h^{d-1} and the part of exponential decay a^h, it means that this model describes the hyperbolic decay when a is close to 1 and the exponential decay when a is further than 1 and (close to 0).

2.2 Main Results

2.2.1 Presentation of the Model

A stochastic process is called a fractional autoregressive process of order **1** with strong mixing errors if it is a solution of

$$(1 - aL)^d X_t = \varepsilon_t \tag{2.11}$$

Where $|a| < 1$,$d \in \mathbb{R}$ so the model is causal and invertible, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ are assumed to be strongly mixing, with $\sigma^2 < \infty$.

Assumptions

Our assumptions are gathered here for easy reference

(A1) The sequence of random variables $(\varepsilon_t)_{t\in\mathbb{Z}}$ satisfy the Cramer conditions given by

$$|\mathbb{E}\varepsilon_i^k| \leq \ell^{k-2}\frac{k!}{2}\mathbb{E}(\varepsilon_i) < \infty \tag{2.12}$$

$i = 1...n$, $k = 3, 4, ...$ and $\ell > 0$

(A2) We consider the following subclass of strong mixing process (Geometrically $\alpha-$mixing)

$$\exists r > 0 \quad and \quad 0 < b < 1 \quad \alpha(n) = rb^n$$

(A3) We consider the following subclass of strong mixing process (Arithmetic $\alpha-$mixing)

$$\exists s > 0 \quad and \quad l < 3 \quad \alpha(n) = sn^{-l}$$

Comments on the assumptions

The Assumptions $(A1)$ guarantees that $(\mathbb{E}(\varepsilon_i^p))^{\frac{1}{p}}$ with $p > 0$ exists. $(A2)$ is used to characterize the dependency structure of noises.
The Assumptions $(A1)$ combined with $(A2)$ and $(A3)$ ensures that the results converge.

Remark 2.1. [15] Davydov's inequality is modified in the case of $\alpha-$mixing

$$|Cov(\varepsilon_i, \varepsilon_j)| \leq C\alpha|i-j| \tag{2.13}$$

2.2.2 Geometrical Strong Mixing Case

Theorem 2.4. *Under assumptions $(A1)$ and $(A2)$, we have for any integer $k > 2$ and $0 < b < 1$*

$$|\gamma_X(h)| = \sigma^2 \psi_h a^h F(d, d+h, h+1, a^2) + C_p'' b^{\frac{k-2}{k}h} \tag{2.14}$$

Proof

For $h \in \mathbb{Z}$, we have

$$\begin{aligned} |Cov(X_t, X_{t+h})| &\leq \sum_{i=0}^{\infty} \psi_i^2 |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-i})| + \sum_{i \neq j} \psi_i \psi_j |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-j})| \\ &= P_1 + P_2 \end{aligned}$$

Firstly, we deal with P_1

$$P_1 = \sum_{i=0}^{\infty} \psi_i^2 |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-i})|$$

Under the assumption $(A1)$ and using Davydov's inequality modified for $\alpha-$mixing case, we get

$$|Cov(\varepsilon_{t-i}, \varepsilon_{t+h-i})| \leq 8(\mathbb{E}(\varepsilon_{t-i}\varepsilon_{t+h-i}))^k \alpha^{\frac{k-2}{k}}(h) \quad k = 3, 4, ... \tag{2.15}$$

Using Cramer condition A(1), we obtain

$$P_1 \leq 8(\sigma^2 \frac{k!}{2} \ell^{k-2})^{\frac{2}{k}} \alpha^{\frac{k-2}{k}}(h) \sum_{i=0}^{\infty} \psi_i^2$$

Thus

$$P_1 \leq C\alpha^{\frac{k-2}{k}}(h) \sum_{i=0}^{\infty} \psi_i^2 \tag{2.16}$$

where the constant C

$$C = 8(\sigma^2 \frac{k!}{2} \ell^{k-2})^{\frac{2}{k}}$$

according to $(A2)$, we get

$$P_1 \leq b^{\frac{k-2}{k}h} r^{\frac{k-2}{k}} \sum_{i=0}^{\infty} \psi_i^2.$$

$$P_1 \leq C_k b^{\frac{k-2}{k}h} \quad k = 3, 4, ... \tag{2.17}$$

where C_k is constant

$$C_k = Cr^{\frac{k-2}{k}} \sum_{i=0}^{\infty} \psi_i^2.$$

It remains to deal with P_2

$$P_2 \leq \sum_{i \neq j} \psi_i \psi_j |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-j})|$$

now, we treat two cases; when $i = j - h$ and $i \neq j - h$, following similar steps we get

$$P_2 \leq \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} |\psi_i \psi_j| \alpha^{\frac{k-2}{k}}(h + i - j) + \sigma^2 \sum_{i=0}^{\infty} |\psi_i \psi_{i+h}| \quad k = 3, 4, ...$$

using the properties of hypergeometric function and $(A2)$ and 2.5, we obtain

$$P_2 \leq C'_k b^{\frac{k-2}{k}h} + \sigma^2 \psi_h a^h F(d, d + h, h + 1, a^2) \quad k = 3, 4, ... \tag{2.18}$$

where the constant C'_k

$$C'_k = 2Cr^{\frac{k-2}{k}} C_2$$

and

$$C_2 = \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} |\psi_i \psi_j| b^{\frac{k-2}{k}(i-j)}$$

we combine 2.17 and 2.18, we take $C''_k = C_k + C'_k$, therefore the result will be

$$|Cov(X_t, X_{t+h})| \leq \sigma^2 \psi_h a^h F(d, d + h, h + 1, a^2) + C''_k b^{\frac{k-2}{k}h}$$

which completes the proof.

Corollary 1 *Let $|a| < 1$, using Theorem 2.4 we get*

$$Cov(X_t, X_{t+h}) \sim \frac{a^h}{\Gamma(d)} h^{d-1}(1-a^2)^{-d} \tag{2.19}$$

Proof

We know that $0 < b < 1$, thus

$$C_k'' b^{\frac{k-2}{k}h} \to 0, \tag{2.20}$$

when $h \to \infty$

Using Schepard formula [1] $\dfrac{\Gamma(h+a)}{\Gamma(h+b)} = h^{a-b} \quad h \to \infty$

Therefore

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

Moreover

We use (1.7) if $|a| < 1$ when h tends to infinity.

Thus

$$F(d, d+h, h+1, a^2) = (1-a^2)^{-d} \tag{2.21}$$

We combine (2.20) and (2.21) to obtain the result in Corollary 1.

Corollary 2 *Let $|a| < 1$ using Theorem 2.4 and the precedent corollary*

$$\rho_X(h) \sim \frac{\frac{a^h}{\Gamma(d)} h^{d-1}(1-a^2)^{-d}}{F(d, d, 1, a^2)} \tag{2.22}$$

Proof

We know that

$$\rho_X(h) = \frac{\gamma_X(h)}{Var(X_t)}$$

The autocovariance $Cov(X_t, X_{t+h})$ is calculated in 2, it suffices to compute $Var(X_t)$ which is expressed in equation (2.8), Finally, we obtain

$$\rho_X(h) \sim \frac{\frac{a^h}{\Gamma(d)} h^{d-1}(1-a^2)^{-d}}{F(d, d, 1, a^2)} \tag{2.23}$$

which completes the proof.

2.2.3 Arithmetical Strong Mixing Case

On the other hand, if the innovations (ε_i) are arithmetically α-mixing, we obtain this result

Theorem 2.5. *Under the assumption (A1) and A(3), we have for any integer* $p > 2$ *and reel* $l > 3$,

$$|\gamma_X(h)| = \sigma^2 \Psi_h a^h F(d, d+h, h+1, a^2) + V_p h^{l^{\frac{2-k}{k}}} + V_p' h^{-1} \tag{2.24}$$

Proof

We follow the same steps as in the case of geometrical strong mixing errors.
For $h \in \mathbb{Z}$, we have

$$\begin{aligned} |Cov(X_t, X_{t+h})| &\leq \sum_{i=0}^{\infty} \psi_i^2 |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-i})| + \sum_{i \neq j} \psi_i \psi_j |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-j})| \\ &= R_1 + R_2 \end{aligned}$$

On the first hand, we treat $R1$

$$R_1 = \sum_{i=0}^{\infty} \psi_i^2 |Cov(\varepsilon_{t-i}, \varepsilon_{t+h-i})|$$

Under the assumption $(A1)$ and using Davydov's inequality modified for $\alpha-$mixing case, we get

$$|Cov(\varepsilon_{t-i}, \varepsilon_{t+h-i})| \leq 8(\mathbb{E}(\varepsilon_{t-i}\varepsilon_{t+h-i}))^k \alpha^{\frac{k-2}{k}}(h) \quad k = 3, 4, ... \tag{2.25}$$

Using Cramer condition A(1), we obtain

$$R_1 \leq 8(\sigma^2 \frac{k!}{2} \ell^{k-2})^{\frac{2}{k}} \alpha^{\frac{k-2}{k}}(h) \sum_{i=0}^{\infty} \psi_i^2$$

Thus

$$R_1 \leq C \alpha^{\frac{k-2}{k}}(h) \sum_{i=0}^{\infty} \psi_i^2 \tag{2.26}$$

According to $(A2)$, we get

$$R_1 \leq s^{\frac{k-2}{k}h} n^{\frac{k-2}{k}} \sum_{i=0}^{\infty} \psi_i^2.$$

$$V_p = Cs^{\frac{k-2}{k}} \sum_{i=0}^{\infty} \psi_i^2 \tag{2.27}$$

Where C is constant

On the second hand we deal with R_2, $R_2 = R_{2,1} + R_{2,2}$

$$R_{2,1} = \sum_{i \neq j} |\psi_i \psi_j||Cov(\varepsilon_{t-i}, \varepsilon_{t+h-j})| \quad with \quad j = i + h$$

$$R_{2,1} = \sigma^2 \psi_h F(d, d+h, h+1, a^2) \tag{2.28}$$

$$R_{2,2} = \sum_{i \neq j} |\psi_i \psi_j||Cov(\varepsilon_{t-i}, \varepsilon_{t+h-j})| \quad with \quad j \neq i + h$$

$$R_{2,2} = V_p' h^{-1} \tag{2.29}$$

With $V_p' = 2Cs^{\frac{p-2}{p}} C_1$,

Where $C_1 = \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} |\psi_i \psi_j| h((i-j)+h)^{l(\frac{2-p}{p})}$

We combine (2.28), (2.29) and (2.26) to complete the proof.

The asymptotic behavior in this case can be deduced analogously to geometric strong mixing case.

Arithmetical Strong Mixing Case for long memory model

- In the case of $a = 1$, we refer the reader to the work of Haddad and Belaide [41].

$$(1 - L)^d X_t = \varepsilon_t \tag{2.30}$$

$|d| < \dfrac{1}{2}$ to insure the invertibility, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ are i.i.d., with $\sigma^2 < \infty$

Theorem 2.6. *Under the assumption (A1) and A(3), we have for any integer $p > 2$ and reel $l > 3$,*

$$|\gamma_x(h)| = \sigma^2 \Psi_h a^h F(d, d+h, h+1, 1) + V_p h^{l\frac{2-k}{k}} + V_p' h^{-1} \quad h > 0 \tag{2.31}$$

Where

$$V_p = C \sum_{i=0}^{\infty} \psi_i^2 \quad p \geq 3$$
$$V_p' = C \sum_{j=1}^{\infty} \psi_j^2 \quad p \geq 3$$

C positive constant

Proof

We treat analogously to the Arithmetic and geometrical mixing noises cases.

Lemma 2.1. $0 < d < \frac{1}{2}$, *using the previous result 2.6*

$$\rho(h) \simeq Ch^{2d-1} when h \to \infty \tag{2.32}$$

Proof

The proof is deduced immediately using Scheppard formula.

2.3 Simulation study

A simulation study with R is illustrated to check the results elaborated in the previous section.
We generate the process which satisfies the model (5) in dependent case ($\alpha-$mixing) when ($\alpha = 0.1$) and ($\alpha = 0.25$) then we compare the behavior of the model in this case to the independent case ($\alpha = 0$).

Remark 2.2. We generate strong mixing sequence of variables then we use it to simulate the model given by the equation (5)

- When $\alpha = 0$, the graph in black
- When $\alpha = 0.1$, the graph in red

- When $\alpha = 0$, the graph in blue

In order to measure the performance of our approach we use sample lags ($h = 500$) then we treat different cases, we take $a = 0.99, 0.2$ and $d = 0.49, 0.1$.
We summarize the results obtained in tables then we present the simulated data in graphs to show the effect of ($\alpha-$mixing) noise.

Table: Auto-correlation for $a = 0.99$ and $d = 0.49$

h	1	50	100	200	300	400	500
$\alpha = 0$	0.8757	0.2096	0.0934	0.0250	0.0024	0.0007	0.0004
$\alpha = 0.1$	0.9324	0.3346	0.2342	0.1732	0.1576	0.1530	0.1515
$\alpha = 0.25$	0.9499	0.3741	0.2775	0.2187	0.2037	0.1992	0.1978

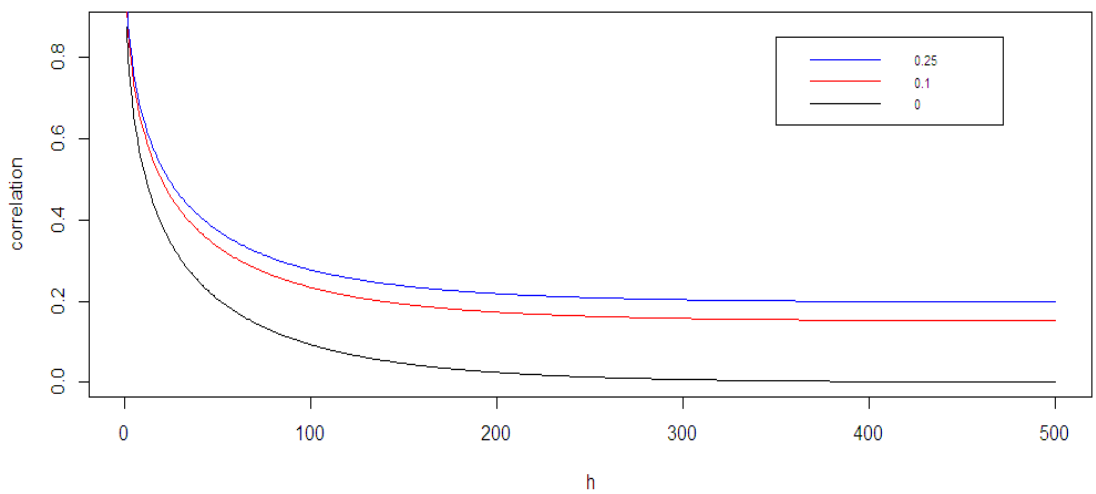

Figure 2.3: Auto-correlations for FAR(1) with $a = 0.99$ and $d = 0.49$.

Table: Auto-correlation for $a = 0.99$ and $d = 0.1$

h	1	50	100	200	300	400	500
$\alpha = 0$	0.6544	0.4250	0.1475	0.0030	0.0015	0.0004	0.0001
$\alpha = 0.1$	0.6805	0.1139	0.0882	0.0773	0.0753	0.0747	0.0745
$\alpha = 0.25$	0.6889	0.1369	0.1119	0.1013	0.0993	0.0988	0.0986

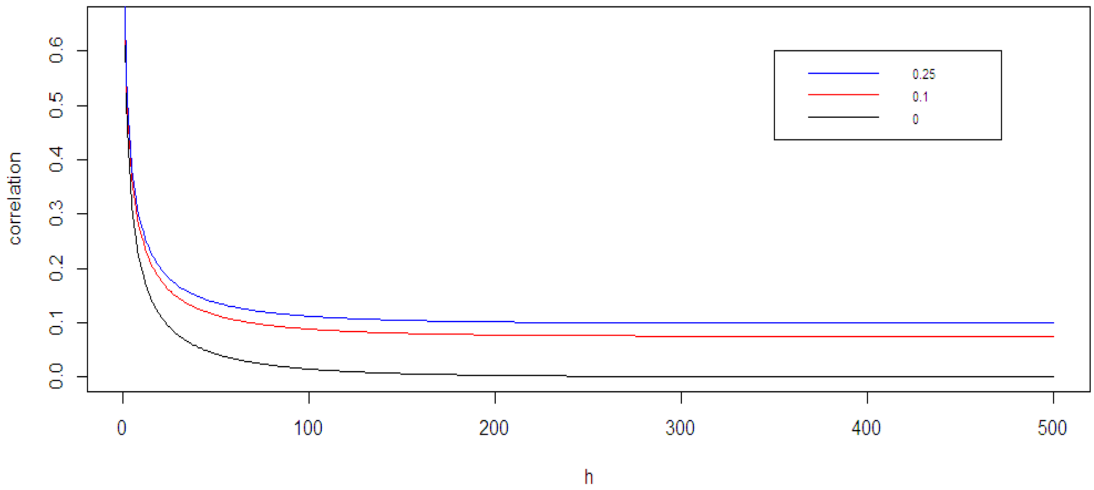

Figure 2.4: Auto-correlations for FAR(1) with $a = 0.99$ and $d = 0.1$.

Interpretation

In these figures, we present FAR(1) and we show the influence of the $\alpha-$mixing noise, we fix $a = 0.99$ then we variate d, it takes the values 0.49, 0.1.

- We remark clearly that the simulated model is affected by the variation of d.(when the parameter d is bigger the decay is slower).

- The results shown above illustrate how strong mixing can affect the behavior of the model, as the coefficient mixing is higher the decay will be slower.

- When $\alpha = 0.25, 0.1$ (dependent case) the auto-correlation of the model decreases more slowly than in the independent case $\alpha = 0$.

Table: Auto-correlation for $a = 0.2$ and $d = 0.49$

h	1	50	100	200	300	400	500
$\alpha = 0$	0.2248	1e-27	2e-53	2e-106	2e-151	3e-211	2e-263
$\alpha = 0.1$	0.2391	0.0183	0.0183	0.0182	0.0182	0.0182	0.0181
$\alpha = 0.25$	0.2441	0.0247	0.0247	0.0247	0.0247	0.0246	0.0246

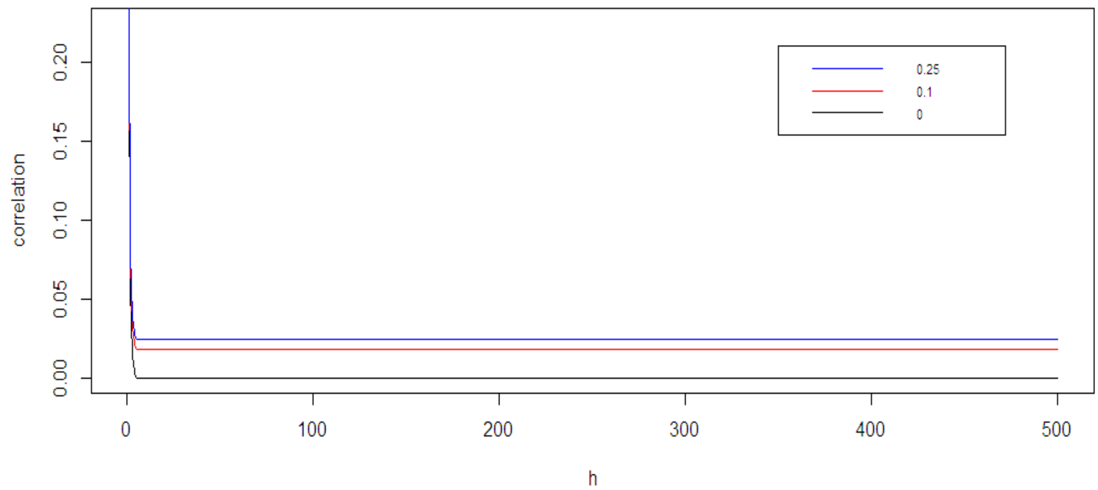

Figure 2.5: Auto-correlations for FAR(1) with $a = 0.2$ and $d = 0.49$.

Table: Auto-correlation for $a = 0.2$ and $d = 0.1$

h	1	50	100	200	300	400	500
$\alpha = 0$	0.1662	2e-28	8e-55	2e-107	8e-160	3e-212	2e-264
$\alpha = 0.1$	0.1721	0.0007	0.0007	0.0007	0.0007	0.0007	0.0006
$\alpha = 0.25$	0.1742	0.0009	0.0009	0.0009	0.0009	0.0008	0.0008

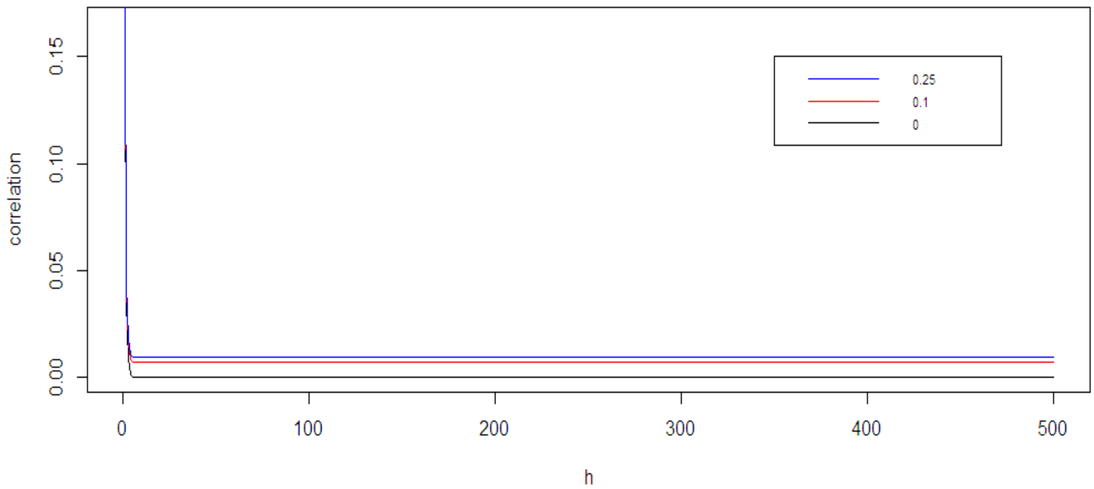

Figure 2.6: Auto-correlations for FAR(1) with $a = 0.2$ and $d = 0.1$.

Interpretation

Finally, we treat our model FAR($\mathbf{1}$), we take $a = 0.2$ and $d = 0.49$, 0.1.

- When the mixing coefficient is higher the auto-correlation decreases quickly but slower than in the case of $\alpha = 0$ (independent) $\rho_X(h)$.

- The effect of d is clear on the behavior of FAR(1), moreover we have a relation between the strong correlation and the parameter d.

- We remark that $\rho_X(h)$ decreases to zero as $h \to \infty$ (in all figures)

We remark clearly the influence of the parameter a in all figures because the auto-correlation is equal to $Va^h h^{d-1}$ where V is constant , when the value of a is higher (close to 1) the auto-correlation decay hyperbolically and when the value of a is lower(close to 0) the decay is fast, now we can say that the value of the parameter a play a major role in the behavior of the process.
In other words, when (a is close to 1) our model behaves as long-memory process but when (a is close to 0) the behavior of the model approach to the behavior of short-memory process.

2.4 Conclusion

In this chapter, we have treated several probabilistic properties of fractional autoregressive models with strong mixing errors, we shed light on autocovariance, autocorrelation functions and their asymptotic behavior, we remark the effect of strong mixing coefficients on the behavior of the model in other words memory is longer when the strong mixing coefficient is higher, furthermore we remark the effect of the parameters of models (d the parameter of the memory) especially the parameter a, our model behaves as long-memory when (a is close to 1) but it behaves as short-memory process when (a is close to 0).

Chapter 3

Local Asymptotic Properties for Fractional Autoregressive process with mixing noises

Introduction

Asymptotic approximation can be used theoretically to study the quality (efficiency) of statistical procedures, other asymptotic approach exists in addition to the usual one Local Asymptotic Normality (LAN) concept it is assumed that the value of the "true parameter" in the model varies slightly with n. This approach lets us study the regularity of estimators.

This approach start with Wald [83](1943) and developed by Le Cam [58](1960), we study also Local Asymptotic Minimaxity (LAM) and Local Asymptotic Linearity. In our work we use Swensen's Lemma [79] (1985), we follow the same methodology in Kreiss [55](1987).

Serroukh [72](1996) treated LAN and LAM properties for Fractional Autoregressive model FAR(1), Haddad and Belaide [40](2020) dealt with LAN property for long memory process.

We generalize the results of Haddad and Belaide on (LAN) for FAR(1) with long memory, we deal with

$$(1 - aL)^d X_t = \varepsilon_t \tag{3.1}$$

$|a| < 1$, $d \in \mathbb{R}$, ε_t are assumed to be strongly mixing.

3.1 Construction of the Variables

3.1.1 Notations and Hypothesis

We consider the following deviation of the parameter θ

$$\iota^{(n)} = n^{\frac{1}{2}}(\theta^{(n)} - \theta) \quad \text{where} \quad \theta^{(n)} = (a + n^{-\frac{1}{2}}\eta^{(n)}, d + n^{-\frac{1}{2}}\delta^{(n)})$$

After local deviation of the parameter of the model 3.1, we obtain

$$(1 - (a + n^{\frac{-1}{2}}\eta^{(n)})L)^{d+n^{\frac{1}{2}}\delta^{(n)}} X_t = \varepsilon_t \qquad n = 1...t \in \mathbb{Z} \tag{3.2}$$

Where $||\iota^{(n)}|| < \infty$ and $||.||$ is any norm on $\mathbb{R}^2$
$A_n(X)$ denote the $\sigma-$algebra generated by $(X_t, t \leq n)$.
$A_n(\varepsilon)$ denote the $\sigma-$algebra generated by $(\varepsilon_t, t \leq n)$.
$A_{n,t}$ denote the sub $\sigma-$algebra of $A_n(X)$ generated by the past of process until the moment t
P_θ the probability distribution of the random vector $X = (X_1, ..., X_n)$ under θ and $f(.)$ unknown probability density of the white noise ε_t

1. The moving average representation: Gonçalves [33]

$$X_t^{(n)} = \sum_{i=0}^{\infty} \psi_i(\theta^{(n)})\varepsilon_{t-i} \qquad t \in \mathbb{Z}$$

Where $\quad \psi_i(\theta^{(n)}) = \dfrac{\Gamma(i + d + n^{-\frac{1}{2}}\delta^{(n)})(a + n^{-\frac{1}{2}}\eta^{(n)})^i}{\Gamma(d + n^{-\frac{1}{2}}\delta^{(n)})\Gamma(i+1)}$

2. The autoregressive representation: Gonçalves [33]

$$\varepsilon_t^n(\theta^{(n)}) = \sum_{j=0}^{\infty} \pi_j(\theta^{(n)})X_{t-j}$$

Where $\quad \pi_j(\theta^{(n)}) = \dfrac{\Gamma(j - d - n^{-\frac{1}{2}}\delta^{(n)})(a + n^{-\frac{1}{2}}\eta^{(n)})^j}{\Gamma(-d - n^{-\frac{1}{2}}\delta^{(n)})\Gamma(j+1)}$

($\Gamma(.)$ is gamma function)

$$\begin{aligned}\varepsilon_t(\theta^n) &= \sum_{j=0}^{t-1}\pi_j(\theta^n)X_{t-j} + \sum_{k=0}^{\infty}\sum_{j=0}^{k}\pi_j(\theta^n)X_{t-j}^n \\ &= \sum_{j=0}^{t-1}\pi_j(\theta^n)X_{t-j} + \sum_{k=0}^{\infty}\sum_{j=0}^{k}\pi_j(\theta^n)\psi_{k-j}\varepsilon_{t-k} \\ &= \sum_{j=0}^{t-1}\pi_j(\theta^n)X_{t-j} + \sum_{k=0}^{\infty}\sum_{j=0}^{k}\pi_{j+t}(\theta^n)\psi_{k-j}\varepsilon_{-k}\end{aligned}$$

Indeed $\varepsilon_t(\theta^n) = \sum_{j=0}^{t-1}\pi_j(\theta^n)X_{t-j} + \sum_{k=0}^{\infty}\sum_{j=0}^{k}\pi_{j+t}(\theta^n)\psi_{k-j}\varepsilon_{-k}$

We assume that we have a finite length realization $X^{(n)} = (X_1^{(n)}...X_n^{(n)})$ of stationary solution of 3.1 and note by

$H_f^{(n)}(\theta)$ the hypothesis on which $X_t^{(n)}$ is generated by the model given by 3.1.

$H_f^{(n)}(\theta^{(n)})$ the counter-hypothesis sequence under which $X_t^{(n)}$ is generated by 3.2.

The Likelihood Function

To prove LAN property, we deal with the asymptotic distribution of the logarithm of conditional likelihood ratio.

Under $H_f^{(n)}(\theta)$ The likelihood function is given by

$$I_{\theta,f}(X^{(n)}) = f\left(\sum_{j=0}^{t-1}\pi_j(\theta)X_{t-j} + \sum_{k=0}^{\infty}\sum_{j=0}^{k}\pi_{j+t}(\theta)\psi_{k-j}\varepsilon_{-k}\right)$$

Now, we use **Swensen's** lemma [79]

$$\Lambda_{f,\theta+n^{-\frac{1}{2}}\iota^{(n)}/\theta} = \log \Upsilon_{f,\theta+n^{-\frac{1}{2}}\iota^{(n)}/\theta}$$

where $\Upsilon_{f,\theta+n^{-\frac{1}{2}}\iota^{(n)}/\theta}(X^{(n)}) = \dfrac{I_{\theta^{(n)},f}(X^{(n)})}{I_{\theta,f}(X^{(n)})}$

$Z_t(\theta)$ is the residual under the hypothesis $H_f^{(n)}(\theta)$, it coincides with white noise ε_t

$$Z_t(\theta) - \gamma_{n,t} = Z_t(\theta^{(n)})$$

Where

$$\gamma_{n,t} = \sum_{k=1}^{\infty}\sum_{j=1}^{k}(\pi_j(\theta^{(n)}) - \pi_j(\theta))\psi_{k-j}(\theta)Z_{t-k}(\theta^{(n)})$$

and

$$Z_t(\theta) = \sum_{j=0}^{t-1} \pi_j(\theta) X_{t-j} + \sum_{k=0}^{\infty} \sum_{j=0}^{k} \pi_{j+t}(\theta) \psi_{k-j} Z_{-k}$$

We pose

$$y_{n,t} = \frac{f^{\frac{1}{2}}(Z_t - \gamma_{n,t})}{f^{\frac{1}{2}}(Z_t)} - 1$$

Assumptions and comments

The assumptions are gathered here for easy references.

- (H1) $f(x) > 0, x \in \mathbb{R}$:
  - $\int_{+\infty}^{-\infty} x f(x) dx = 0$
  - $\int_{+\infty}^{-\infty} x^2 f(x) dx = \sigma^2$
  - $\int_{+\infty}^{-\infty} x^4 f(x) dx < \infty.$
- (H2) (Hajek and Sidak [42] 1967) there exists function f such that $-\infty < a < b < +\infty$, $f(b) - f(a) = \int_{+\infty}^{-\infty} f'(x) dx$ where f is absolutely continuous over finite intervals.
- (H3) Let $\phi_f = -\frac{f'}{f}$, f has a positif and finite Fischer information $I(f) = \int_{+\infty}^{-\infty} \phi_f^2 < \infty$ and $\int_a^b (-\frac{f'}{f})^4 f(x) dx < \infty$
- (H4)f is unimodal, ie $-\log(x)$ function is convex on the open intervals $]a; b[$, ϕ_f is increasing on $\mathbb{R}$.
- (H5)The sequence of random variables $(\varepsilon_t)_{t \in \mathbb{Z}}$ satisfy the Cramer conditions given by

  $$|\mathbb{E}\varepsilon_i^m| \leq \ell^{m-2} \frac{m!}{2} \mathbb{E}(\varepsilon_i) < \infty$$

  $i = 1...n$ and $m = 3, 4, ...$
- (H6) $\exists u_n \in \mathbb{N}^*, o(n[\alpha(u_n)]^{\frac{p-2}{p}}) \longrightarrow_{n \to \infty} 0$

Construction of the central sequence

We apply Taylor expansion on $\gamma_{n,t}$, we get

We write $\pi_j(\theta^{(n)})-\pi_j(\theta) = \dfrac{n^{-\frac{1}{2}}}{2}\displaystyle\sum_{k=1}^{\infty}\sum_{j=1}^{k}\left(\dfrac{\partial\pi_j(\theta)}{\partial a}\eta^{(n)} + \dfrac{\partial\pi_j(\theta)}{\partial d}\delta^{(n)}\right)+\dfrac{n^{-1}}{2}\partial^2\pi_j(\theta)(n^{-\frac{1}{2}}c\iota^{(n)}),\quad 0 < c < 1$

Where

$$\partial^2\pi_j(\theta)(n^{-\frac{1}{2}}c\iota^{(n)}) = \frac{\partial^2\pi_j(\theta)(n^{-\frac{1}{2}}c\iota^{(n)})}{\partial a^2}(\eta^n)^2+\frac{\partial^2\pi_j(\theta)(n^{-\frac{1}{2}}c\iota^{(n)})}{\partial a\partial d}(\eta^n\delta^n)+\frac{\partial^2\pi_j(\theta)(n^{-\frac{1}{2}}c\iota^{(n)})}{\partial d^2}(\delta^n)^2$$

$$\partial^2\pi_j(\theta)(n^{-\frac{1}{2}}c\iota^{(n)}) \to 0$$

Thus

$$Z_{n,t} \;=\; \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\left(\frac{\partial\pi_j(\theta)}{\partial a}\eta^{(n)} + \frac{\partial\pi_j(\theta)}{\partial d}\delta^{(n)}\right)\psi_{k-j}(\theta)\phi(Z_t)Z_{t-k}$$

We replace $\left(\dfrac{\partial\pi_j(\theta)}{\partial a}\eta^{(n)} + \dfrac{\partial\pi_j(\theta)}{\partial d}\delta^{(n)}\right)$ and ψ_{k-j} by their values, we obtain

$$\begin{aligned} Z_{n,t} &= \frac{n^{-\frac{1}{2}}}{2}(\sum_{k=1}^{\infty}\sum_{j=1}^{k}\eta^{(n)}\frac{-\Gamma(j-d)\Gamma(k-j+d)}{\Gamma(-d)\Gamma(d)\Gamma(k-j+1)\Gamma(j)}a^{k-1}\phi(Z_t)Z_{t-k} \\ &\quad +\delta^{(n)}\frac{\Gamma(j-d)\Gamma(k-j-d)}{\Gamma(-d)\Gamma(d)\Gamma(k-j+1)\Gamma(j+1)}\left(\sum_{h=1}^{j}\frac{1}{j-h-d}\right)a^k\phi(Z_t)Z_{t-k}) \end{aligned}$$

Because

- $\displaystyle\sum_{j=1}^{k}\frac{-\Gamma(j-d)\Gamma(k-j+d)}{\Gamma(-d)\Gamma(d)\Gamma(k-j+1)\Gamma(j)} = -d$
- $\displaystyle\sum_{j=1}^{k}\frac{\Gamma(j-d)\Gamma(k-j-d)}{\Gamma(-d)\Gamma(d)\Gamma(k-j+1)\Gamma(j+1)}\left(\sum_{h=1}^{j}\frac{1}{j-h-d}\right) = \frac{1}{k}$

For more details see Gonçalves [33]

We remark that

$$\begin{aligned} Z_{n,t} &= \frac{n^{-\frac{1}{2}}}{2}\eta^{(n)}d\sum_{k=1}^{q}a^{k-1}\phi_f(Z_t)Z_{t-k-1} + \frac{n^{-\frac{1}{2}}}{2}\delta^{(n)}d\sum_{k=1}^{q}\frac{a^k}{k}\phi_f(Z_t)Z_{t-k} + R_{n,t} \\ &= \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{q}a^{k-1}(d\eta^{(n)} + \frac{a}{k}\delta^{(n)})\sum_{t=k+1}^{n}\phi_f(Z_t)Z_{t-k} + \underbrace{R_{n,t}}_{o(1)} \end{aligned}$$

with

$$R_{n,t} = \frac{n^{-\frac{1}{2}}}{2} \sum_{k=q+1}^{\infty} a^{k-1}(d\eta^{(n)} + \frac{a}{k}\delta^{(n)}) \sum_{t=k+1}^{\infty} \phi_f(Z_t)Z_{t-k}$$

$R_{n,t} = o_P(1)$

Thus

$$2Z_{n,t} = (\iota^n)'\Delta_f^{(n)} + o_p(1)$$

$$\Delta_f^{(n)} = n^{-\frac{1}{2}} \begin{pmatrix} \sum_{k=1}^{q} da^{k-1} \sum_{t=q+1}^{n} \phi_f(Z_t)Z_{t-k} \\ \sum_{k=1}^{q} \frac{a^k}{k} \sum_{t=q+1}^{n} \phi_f(Z_t)Z_{t-k} \end{pmatrix} \tag{3.3}$$

$\Delta_f^{(n)}$ is called central sequence.

$q = q(n) < n$, q is truncation parameter.

Proposition 3.1. *We have*

$$(\log(1-L))^2(1-L)^{n^{-\frac{1}{2}}c\delta^{(n)}} = \sum_{\infty}^{l=2} h_{l,n}L^l \tag{3.4}$$

with $\quad h_{l,n} = \sum_{l-2}^{j=0} C_{l-j} \frac{\Gamma(j - n^{-\frac{1}{2}}c\delta^{(n)})}{\Gamma(-n^{-\frac{1}{2}}c\delta^{(n)})\Gamma(j+1)} \qquad l = 2, 3, ...$

Where $C_l = \frac{2}{l}\sum_{i=1}^{l-1}\frac{1}{i} \qquad l \geq 2 \quad$ *and* $\quad \sum_{l=2}^{\infty} h_{l,n}^2 < \infty \quad$ *if* $\quad n^{-\frac{1}{2}}c\delta^{(n)} > 0$

Proof

The proof is detailed in Page 58-60 in Serroukh [44].

3.2 Main results

Theorem 3.2. *Under the assumptions* $(H1)$ *to* $(H6)$ *and for any integer* $p > 2$ *and* $k \geq 1$

$$\mathbb{E}(\Delta_f^{(n)}) \leq \begin{pmatrix} \left(\frac{d^2}{(1-a^2)}\right)^{\frac{1}{2}} \left[(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})\right]^{\frac{1}{4}} \\ \left(\sum_{k=1}^{q} \frac{a^{2k}}{k^2}\right)^{\frac{1}{2}} \left[(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})\right]^{\frac{1}{4}} \end{pmatrix} \tag{3.5}$$

with $C_0 = 8C^{\frac{2}{p}}, C = \frac{p!}{2}k^{p-2}\alpha^{\frac{2-p}{p}}, p > 2, k > 0$

The covariance matrix is expressed as follows

$$\begin{pmatrix} V_{1,1} & V_{1,2} \\ V_{2,1} & V_{2,2} \end{pmatrix} \tag{3.6}$$

$V_{1,1} \leq \frac{d^2}{(1-a^2)}([(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}})$

$V_{1,2} \leq \frac{-d}{a}\log(1-a^2)\left([(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right)$

$V_{2,1} \leq \frac{-d}{a}\log(1-a^2)\left([(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right)$

$V_{2,2} \leq \sum_{k=1}^{q}\frac{a^{2k}}{k^2}\left([(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right)$

Proof

We know that

$$\mathbb{E}(\Delta_f^n) \leq -\frac{1}{2}\left(\begin{array}{c} \left[\sum_{k=1}^{q} d^2 a^{2k-2}\sum_{t=k+1}^{n} Var(\phi_f(Z_t)Z_{t-k}) + I_f\sigma^2\right]^{\frac{1}{2}}. \\ \left[\sum_{k=1}^{q}\frac{a^{2k}}{k^2}\sum_{t=k+1}^{n} Var(\phi_f(Z_t)Z_{t-k}) + I_f\sigma^2\right]^{\frac{1}{2}}. \end{array}\right)$$

Firstly, we deal with $Var(\phi_f(Z_t)Z_{t-k})$ using Davydov's inequality [27] modified for $\alpha-$mixing case, we take $p = q \neq 2$ it implies that $r = \frac{p}{p-2}$ where $p > 2$ and $q < \infty$, $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ and $p \in \mathbb{N}$ and using the $(H5)$ we obtain

$$Var(\phi_f(Z_t)Z_{t-k}) \leq (I_f^2 + C_0 I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4 + C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}} - I_f\sigma^2 \tag{3.7}$$

with $C_0 = 8C^{\frac{2}{p}}, C = \frac{p!}{2}k^{p-2}\alpha^{\frac{2-p}{p}}, p > 2, k > 0$

We replace 3.7 in the previous inequality of $\mathbb{E}(\Delta_f^n)$, we get the result.

Now, we prove the second part of the Theorem 3.2

On the first hand we calculate $V_{1,1}$ and $V_{2,2}$ using the following inequality

$Var(\phi_f(Z_t)Z_{t-k}) \leq [(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}} - I_f\sigma^2$

So

$$\begin{aligned} V_{1,1} &\leq \frac{d^2}{(1-a^2)}\left([(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}} - I_f\sigma^2 + I_f\sigma^2\right) \\ &\leq \frac{d^2}{(1-a^2)}(\ [(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}) \end{aligned}$$

Similarly, we calculate $V_{2,2}$
$V_{2,2} \leq \sum_{k=1}^{q} \frac{a^{2k}}{k^2}\left([(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right)$
On the other hand, we calculate $V_{1,2}$ and $V_{2,1}$, we get

$$\begin{aligned} V_{1,2} &\leq -d\sum_{k=1}^{q}\frac{a^{2k}-1}{k}\left([(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right) \\ &\leq \frac{-d}{a}\log(1-a^2)\left([(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right) \end{aligned}$$

Similarly we calculate $V_{2,1}$.

3.2.1 Local Asymptotic Normality

Proposition 3.3. *Under the assumptions $(H1)-(H6)$, for $|a| \leq 1$ and $d \in \mathbb{R}$, for all bounded sequences $\iota^{(n)}$, we have LAN property.*
Local Asymptotic Quadratic (LAQ) decomposition has the following form

$$\Lambda_{f,\theta+n^{-\frac{1}{2}}\iota^{(n)}/\theta} = (\iota^{(n)})'\Delta_f^{(n)}(\theta) - \frac{1}{2}(\iota^{(n)})'(I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\Xi(\theta)\iota^{(n)}$$

Where

$$\Xi(\theta) = \begin{pmatrix} \frac{d^2}{1-a^2} & \frac{-d}{a}\log(1-a^2) \\ \frac{-d}{a}\log(1-a^2) & \sum_{k=1}^{\infty}\frac{a^{2k}}{k^2} \end{pmatrix}$$

with $\theta = (a,d)$

The central sequence $\Delta_f^{(n)}(\theta)$ 3.3 is asymptotically normal under $H_n^f(\theta)$ with:

Mean $$\begin{pmatrix} \left(\frac{d^2}{1-a^2}\right)^{\frac{1}{2}} [(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{4}} \\ \left(\sum_{k=1}^{q}\frac{a^{2k}}{k^2}\right)^{\frac{1}{2}} [(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{4}} \end{pmatrix}$$

$$\textit{Covariance} \quad \left((I_f^2 + C_0 I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4 + C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\Xi(\theta)\right)$$

Where $C_0 = 8C^{\frac{2}{p}}$, $C = \dfrac{p!}{2}k^{p-2}\alpha^{\frac{2-p}{p}}, \quad p \in \mathbb{N}, p > 0, k > 0$

Proof

To prove the previous proposition we have to check the conditions of Swensen's lemma 1.2

Proof (C1)

$$\begin{aligned}
\mathbb{E}\sum_{t=1}^{n}(Z_{n,t} - y_{n,t})^2 &\le 2\sum_{t=1}^{n}\mathbb{E}\left(\frac{f^{\frac{1}{2}}(Z_t - \gamma_{n,t})}{f^{\frac{1}{1}}(Z_t)} - 1 + \frac{1}{2}\gamma_{n,t}\phi_f(Z_t)\right)^2 + 2\sum_{t=1}^{n}\mathbb{E}(Z_{n,t} - \frac{1}{2}\gamma_{n,t}\phi_f(Z_t))^2 \\
&\le A + B
\end{aligned}$$

We start with the second term B

Firstly

$$\begin{aligned}
Z_{n,t} - \frac{1}{2}\gamma_{n,t}\phi_f(Z_t) &= \phi_f(Z_t)\psi_{k-j}Z_{t-k} \times (\frac{n^{-1}}{4}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta + n^{-\frac{1}{2}}c\iota^{(n)})}{\partial a^2}(\eta^{(n)})^2 \\
&\quad + \frac{n^{-1}}{2}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta + n^{-\frac{1}{2}}c\iota^{(n)})}{\partial a\partial d}(\eta^{(n)}\delta^{(n)}) \\
&\quad + \frac{n^{-1}}{4}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta + n^{-\frac{1}{2}}c\iota^{(n)})}{\partial d^2}(\delta^{(n)})^2) \\
&= B_1 + B_2 + B_3
\end{aligned}$$

Where $B_1 = \phi_f(Z_t)\psi_{k-j}Z_{t-k} \times \left(\dfrac{n^{-1}}{4}\displaystyle\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j}{\partial a^2}(\theta + n^{-\frac{1}{2}}c\iota^{(n)})(\eta^{(n)})^2\right)$

$B_2 = \phi_f(Z_t)\psi_{k-j}Z_{t-k}\left(\dfrac{n^{-1}}{2}\displaystyle\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta + n^{-\frac{1}{2}}c\iota^{(n)})}{\partial a\partial d}(\eta^{(n)}\delta^{(n)})\right)$

$B_3 = \phi_f(Z_t)\psi_{k-j}Z_{t-k}\left(\dfrac{n^{-1}}{4}\displaystyle\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta + n^{-\frac{1}{2}}c\iota^{(n)})}{\partial d^2}(\delta^{(n)})^2\right)$

We know that $(Z_{n,t} - \dfrac{1}{2}\gamma_{n,t}\phi_f(Z_t))^2 \le 5(B_1^2 + B_2^2 + B_3^2)$

We deal with B_1

$$\begin{aligned}
B_1 &= \frac{n^{-1}}{4}(\eta^{(n)})^2[(1 - (a + n^{-\frac{1}{2}}c\eta^{(n)})L)^{d + n^{-\frac{1}{2}}c\delta^{(n)} - 2}(1 - aL)Z_{t-2}]\phi_f(Z_t) \\
&= \frac{n^{-1}}{4}(\eta^{(n)})^2[I_1 \times I_2 Z_{t-2}]\phi_f(Z_t)
\end{aligned}$$

Where
$I_1 = (1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-2}$ and $I_2 = (1-aL)$
Using Davydov's inequality, we get

$$\mathbb{E}(B_1^2) \leq \frac{n^{-1}}{4}(\eta^{(n)})^2[(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-2}(1-aL)]^2(I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}$$

The terms I_1 and I_2 are integrable and square integrable.
We deduce $\sum_{t=1}^{\infty}\mathbb{E}(B_1)^2 \longrightarrow 0$
Following the same steps, we calculate $\sum_{t=1}^{\infty}\mathbb{E}(B_2)^2$

$$\begin{aligned}
\mathbb{E}(B_2^2) &= (\frac{n^{-1}}{2}\eta^{(n)}\delta^{(n)})\mathbb{E}[(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-1}\log(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L) \\
&\quad (1-aL)^{-d}Z_t]^2(\phi_f(Z_t))^2 \\
&= (\frac{n^{-1}}{2}\eta^{(n)}\delta^{(n)})\mathbb{E}[(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-1}\log(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L) \\
&\quad (1-aL)^{-d}Z_{t-k}]^2[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}
\end{aligned}$$

The three terms are squared integrable, thus $\sum_{t=1}^{\infty}\mathbb{E}(B_2)^2 \longrightarrow 0$
Analogously we treat B_3.

$$\begin{aligned}
B_3 &= -\frac{n^{-1}}{2}(\delta^n)^2\sum_{k=1}^{\infty}\sum_{j=1}^{k}\left(\frac{\partial^2}{\partial d^2}\frac{\Gamma(j-d-n^{-\frac{1}{2}}c\delta^{(n)})(a+n^{-\frac{1}{2}}c\eta^{(n)})^j}{\Gamma(j-d-n^{-\frac{1}{2}}c\delta^{(n)})j!}\right) \\
&\quad \times\frac{\Gamma(k-j+d)a^{k-j}}{\Gamma(d)(k-j)!}\phi_f(Z_t)Z_{t-k} \\
&= -\frac{n^{-1}}{2}(\delta^n)^2\left[\left(\frac{\partial^2}{\partial d^2}(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}}(1-aL)^{-d}Z_t\right)\phi_f(Z_t)\right]
\end{aligned}$$

We calculate

$$\frac{\partial^2}{\partial d^2}(1-(a+n^{-\frac{1}{2}}c\delta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}}$$

$$\begin{aligned}
\frac{\partial^2}{\partial d^2}(1-(a+n^{-\frac{1}{2}}c\delta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}} &= \frac{\partial}{\partial d}\left[\frac{\partial}{\partial d}\exp(d+n^{-\frac{1}{2}}c\delta^{(n)})\log(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)\right] \\
&= \left[\log(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)\right]^2(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}}
\end{aligned}$$

We get

$$B_3 = -\frac{n^{-1}}{2}(\delta^{(n)})^2\left[\log(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^2(1-(a+n^{-\frac{1}{2}}c\eta^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}}(1-aL)^{-d}Z_t\right] \tag{3.8}$$

The components of 3.8 are square integrable, consequently $\sum_{t=1}^{\infty}\mathbb{E}(B_3)^2 \longrightarrow_{n\to\infty} 0$.

Now, using then results of B_1, B_2, B_3, we conclude that B converges to zero.

For the first term A, we can write $\gamma_{n,t}$ as follow

$$\gamma_{n,t} = n^{-\frac{1}{2}}U_t + V_{n,t}$$

We note

$$\varphi_{1,k}^{(n)} = \eta^{(n)}da^{k-1} + \delta^{(n)}\frac{a^k}{k}$$

$$\varphi_{2,k}^{(n)}(\theta+n^{-\frac{1}{2}}c\eta^{(n)}) = \sum_{j=1}^{k}\psi_{t-j}(\theta)\partial^2\pi_j(\theta+n^{-\frac{1}{2}}c\eta^{(n)})$$

and

$$U_t = \sum_{k=1}^{q}\varphi_{1,k}^{(n)}Z_{t-k} \qquad V_{n,t} = \sum_{k=1}^{q}\frac{n^{-1}}{2}\varphi_{2,k}^{(n)}(\theta+n^{-\frac{1}{2}}c\eta^{(n)})Z_{t-k}$$

We consider $A_{1,n}$ and $A_{2,n}$

and $A_1(m) = \{|Z_{t-i}| < m, i = 1...q\}$ $\qquad A_2(m) = \{|Z_{t-i}| \geq m, i = 1...q\}$

$$\begin{aligned} A_{1,n} &= \sum_{t=1}^{n}\mathbb{E}\{\mathbb{1}_{A_1}(\frac{f^{\frac{1}{2}}(Z_t-\gamma_{n,t})}{f^{\frac{1}{2}}(Z_t)} - 1 - Z_{n,t})^2\} \\ &= \sum_{t=1}^{n}\mathbb{E}\{\mathbb{1}_{A_1}\left(\frac{f^{\frac{1}{2}}(Z_t-n^{-\frac{1}{2}}U_t-V_{n,t})-f^{\frac{1}{2}}(z_t)}{(n^{-\frac{1}{2}}U_t+V_{n,t})f^{\frac{1}{2}}(Z_t)} - \frac{1}{2}\phi_f(Z_t)\right)^2\gamma_{n,t}^2\} \end{aligned}$$

We integrate with respect to z, u, v_n, than we obtain

$$A_{1,n} \leq \sup_{|u|<k_0, v_n\leq n^{-r}k_1}\int_{\mathbb{R}}\left(\frac{f^{\frac{1}{2}}(Z-n^{-\frac{1}{2}}u-v_n)-f^{\frac{1}{2}}(Z)}{(n^{-\frac{1}{2}}u+v_n)} - \frac{1}{2}\frac{f'(Z)}{f(Z)}\right)^2 dz\sum_{t=1}^{n}\mathbb{E}(\gamma_{n,t}^2)$$

According to Swensen it is suffices to show that $\mathbb{E}(\gamma_{n,t}^2) < \infty$

$$\begin{aligned} \mathbb{E}(\gamma_{n,t}^2) &< \mathbb{E}(n^{-\frac{1}{2}}\sum_{k=0}^{\infty}\varphi_{1,k}^{(n)}Z_t + \varphi_{2,k}^{(n)}(\theta+n^{-\frac{1}{2}}c\iota^{(n)})Z_{t-k})^2 \\ &< 2\sigma^2\sum_{k=0}^{\infty}(\varphi_{1,k}^{(n)})^2 + \frac{n^{-1}}{2}\sigma^2\sum_{k=0}^{\infty}(\varphi_{2,k}^{(n)})^2 \\ &< \infty \end{aligned}$$

We conclude that $A_{1,n} \longrightarrow 0$

Analogously, we treat $A_{2,n}$

$$\begin{aligned}
A_{2,n} &= \sum_{i=1}^{n} \mathbb{E}\left(\mathbb{1}_{A_2}\left(\frac{f^{\frac{1}{2}}(Z_t - n^{-\frac{1}{2}})U_t - V_{n,t} - f^{\frac{1}{2}}(Z_t)}{f^{\frac{1}{2}(Z_t)}}\right) - \frac{1}{2}\phi_f(Z_t)\gamma_{n,t} \right) \\
&= \int_{\mathbb{R}} \left(f^{\frac{1}{2}}(z - n^{-\frac{1}{2}}u - v_n) - f^{\frac{1}{2}}(z) - \frac{1}{2}(n^{-\frac{1}{2}}u - v_n)\phi_f(z) f^{\frac{1}{2}}(z) \right)^2 dz \\
&\leq \int_{\mathbb{R}} (f(z - n^{-\frac{1}{2}}u - v_n)dz + \int_{\mathbb{R}} f(z)dz + \int_{\mathbb{R}} (n^{-\frac{1}{2}}u - v_n)\phi_f^2(z) f(z)dz \\
&\leq (n^{-\frac{1}{2}}u - v_n)[(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}
\end{aligned}$$

using Holder's inequality

$$A_{2,n} \leq [(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}} P(|Z_{t-i}| \geq m, i = 1, ...,)^{\frac{1}{3}} \sum_{i=1}^{n} \mathbb{E}(|\gamma_{n,t}|^3)^{\frac{2}{3}}$$

Where $\mathbb{E}(|\gamma_{n,t}|^3)^{\frac{2}{3}}$ is bounded because
$\mathbb{E}(\gamma_{n,t}^4) \leq 4n^{-2}\mathbb{E}\left(\sum_{k=1}^{\infty} \varphi_{1,k} Z_{t-k}\right)^4 + \frac{n^{-4}}{4}\mathbb{E}\left(\sum_{k=1}^{\infty} \varphi_{2,k}(\theta + n^{-\frac{1}{2}}c\iota^{(n)})Z_{t-k}\right)$ The coefficients $\varphi_{1,k}$ and $\varphi_{2,k}(\theta + n^{-\frac{1}{2}}c\iota^{(n)})$ are square integrable, Thus

$$\mathbb{E}(\gamma_{n,t}^3)^{\frac{2}{3}}\mathbb{E}(\gamma_{n,t}^4)^{\frac{1}{2}}$$

and $P(|Z_{t-i}| \geq m, i = 1, ...,)^{\frac{1}{3}} \leq \sigma^{\frac{2}{3}} k^{-\frac{2}{3}} (\log q)^{-\frac{1}{3}}$ where q is truncation parameter. Finally we conclude that $A_{2,n} \to_{\infty} 0$ $A_{2,n} \longrightarrow 0$

Proof C2

$$Z_{n,t} = \frac{n^{-\frac{1}{2}}}{2} \sum_{k=1}^{\infty} \varphi_{1,k}^{(n)} \phi_f(Z_t) Z_{t-k}$$

$$\begin{aligned}
\mathbb{E}(Z_{n,t})^2 &= \frac{n^{-1}}{4} \lim_{m\to\infty} \inf \mathbb{E}(\sum_{k=1}^{m} \varphi_{1,k}^{(n)} \phi_f(Z_t) Z_{t-k})^2 \\
&= \frac{n^{-1}}{4}(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}}) \sum_{k=1}^{\infty} (\varphi_{1,k}^{(n)})^2
\end{aligned}$$

We have $\sum_{k=1}^{\infty} \varphi_{1,k}^{(n)} < \infty \Longrightarrow \mathbb{E}(Z_{n,t})^2 < \infty$

We conclude that $\sup_n \mathbb{E}(Z_{n,t}^2) < \infty$

Proof C3

We consider $\max_{t\leq n}|Z_{t,n}| = Z_{t_0,n}$

We have

$$\begin{aligned}\max_{t\leq n}|Z_{t,n}| > \varepsilon &= Z_{t_0,n} + \sum_{t=1,t\neq t_0}^{n} Z_{n,t}\mathbb{1}_{(|Z_{t,n}|>\varepsilon)} > \varepsilon \\ &= \sum_{t=1}^{n} Z_{n,t}\mathbb{1}_{(|Z_{t,n}|>\varepsilon)} > \varepsilon^2\end{aligned}$$

Using Markov inequality, we obtain

$$P\left(\max_{t\leq n}|Z_{t,n}| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\sum_{t=1}^{n}\mathbb{E}(Z_{n,t}^2\mathbb{1}_{(|Z_{t,n}|>\varepsilon)})$$

and

$$\begin{aligned}\mathbb{E}(Z_{n,t}^2\mathbb{1}_{(|Z_{t,n}|>\varepsilon)}) &= \frac{n^{-1}}{4}\mathbb{E}\left(\liminf_{m\to\infty}\sum_{t=1}^{n}\left(\sum_{k=1}^{m}\varphi_{1,k}(\theta)\phi_f(Z_t)Z_{t-k}\right)^2\mathbb{1}_{(|Z_{t,n}|>\varepsilon)}\right) \\ &= \frac{n^{-1}}{4}\sum_{k=1}^{n}\liminf_{m\to\infty}\mathbb{E}\left((\sum_{k=1}^{m}\varphi_{1,k}(\theta)\phi_f(Z_t)Z_{t-k})^2\mathbb{1}_{(|Z_{t,n}|>\varepsilon)}\right) \\ &= \frac{n^{-1}}{4}\liminf_{m\to\infty}\sum_{t=1}^{n}\int_{(|Z_{t,n}|>\varepsilon)}(\varphi_{1,k}(\theta)\phi_f(Z_t)Z_{t-k})^2 dP_\theta\end{aligned}$$

and

$\int_{(|Z_{t,n}|>\varepsilon)}(\sum_{k=1}^{n}\varphi_{1,k}(\theta)\phi_f(Z_t)Z_{t-k})^2 dP_\theta \leq \sum_{k=1}^{n}\varphi_{1,k}^2\int_{(|Z_{t,n}|>\varepsilon)}(\phi_f(Z_tZ_{t-k}))^2 dP_\theta$

We have also

$|Z_{t,n}| > \varepsilon \Leftrightarrow |\frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{\infty}\varphi_{1,k}\phi_f(Z_tZ_{t-k})| > \varepsilon$

Combining the results we get

$$\sum_{t=1}^{n}\int_{(|Z_{t,n}|>\varepsilon)}(\varphi_{1,k}^{n}\phi_f(Z_tZ_{t-k}))^2 dP_\theta \longrightarrow 0$$

When $n \longrightarrow 0$

Consequently $C3$ is verified.

Proof C4

$\sum Z_{n,t}^2 = \frac{n^{-1}}{4}\sum_{t=1}^{n}(\sum_{k=1}^{\infty}\varphi_{1,k}^{(n)}\phi_f(Z_t)Z_{t-k})^2$

$U_{m,n} = \frac{n^{-1}}{4}\sum_{t=1}^{n}\phi_f^2(Z_t)(\sum_{k=1}^{m}\varphi_{1,k}^{(n)}Z_{t-k})^2$

The process $\frac{n^{-1}}{4}\sum_{t=1}^{n}\phi_f^2(Z_t)(\sum_{k=1}^{m}\varphi_{1,k}^{(n)}Z_{t-k})^2$ is strictly $(m+1)$ dependent with the mean $(I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\sum_{k=1}^{m}(\varphi_{1,k}^{(n)})^2$ according to the Ergodic theorem
$U_{m,n}-\frac{1}{4}(I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\sum_{k=1}^{m}(\varphi_{1,k}^{(n)})^2=o_{p_\theta}(1)$
$\frac{1}{4}(I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\sigma^2\sum_{k=1}^{m}(\varphi_{1,k}^{(n)})^2\longrightarrow_{n\to\infty}\frac{(\lambda)^2}{4}<\infty$
On the other hand

$$\begin{aligned}
\mathbb{P}(|Z_{n,t}^2-U_{m,n}|>\varepsilon) &= \mathbb{P}(\frac{n^{-1}}{4}|\sum_{t=1}^{n}\phi_f^2Z_t[(\sum_{k=1}^{\infty}\phi_f^2Z_{t-k})^2-(\sum_{k=1}^{\infty}\phi_f^2Z_{t-k})^2]|>\varepsilon)\\
&\leq (I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\frac{n^{-1}}{4}\times\\
&\quad \sum_{t=1}^{n}\mathbb{E}|(\sum_{t=m+1}^{\infty}\varphi_{1,k}(\theta)Z_{t-k})^2+2(\sum_{k=1}^{m}\varphi_{1,k}(\theta)Z_{t-k})(\sum_{k=m+1}^{\infty}\varphi_{1,k}(\theta)Z_{t-k})|
\end{aligned}$$

We apply Cauchy-Schwartz inequality then we obtain

$$\begin{aligned}
\mathbb{P}(|Z_{n,t}^2-U_{m,n}|>\varepsilon) &\leq \frac{1}{4}(I_f^2+C_0I_f^{\frac{2}{p}})^{\frac{1}{2}}(\sigma^4+C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}}\\
&\quad \left(\sum_{k=m+1}^{\infty}\varphi_{1,k}^{n}(\theta)\sigma^2[\sum_{k=1}^{m}\varphi_{1,k}^2(\theta)\sigma^2]^{\frac{1}{2}}[\sum_{k=m+1}^{\infty}\varphi_{1,k}^2(\theta)\sigma^2]^{\frac{1}{2}}\right)
\end{aligned}$$

$\mathbb{P}(|Z_{n,t}^2-U_{m,n}|>\varepsilon)\longrightarrow_{m\to}0$ Thus $U_m-\frac{(\lambda^n)^2}{4}=o_{p_\theta}(1)$

Proof C5

$\mathbb{E}(\mathbb{E}(Z_{n,t}^2\mathbb{1}_{|Z_{n,t}|>\frac{1}{2}}/A_{n,t-1}))\leq\mathbb{E}(Z_{n,t}\mathbb{1}_{\max_{1\leq t\leq n}|Z_{n,t}|>\frac{1}{2}})$ based on (Swensen [79]), we have to show $\sum Z_{n,t}^2$ is uniformly integrable, for this purpose we use Serfling [71] lemma ([71] P 15) and $C4$ then we conclude under P_θ that $\sum Z_{n,t}^2<\infty$

Proof C6

$$\begin{aligned}
\mathbb{E}(Z_{n,t}/A_{n,t-1}) &= \frac{n^{\frac{-1}{2}}}{2}\mathbb{E}(\sum_{k=1}^{\infty}\varphi_{1,k}(\theta)\phi_f(Z_t)Z_{t-k}/A_{n,t-1})\\
&= \frac{n^{\frac{-1}{2}}}{2}\sum_{k=1}^{\infty}\varphi_{1,k}(\theta)Z_t\mathbb{E}(\phi_f(Z_t)/A_{n,t-1})\\
&= 0
\end{aligned}$$

Because $\mathbb{E}(\phi_f(Z_t)/A_{n,t-1}) = \int_0^1 \varphi(u,f)du = 0$ We use absolutely continuous densities compared to the Lebesgue measure, therefore the conditions of Swensen's lemma are satisfied, condition (iii) of Le Cam theorem stated in Swensen [79] is automatically satisfied.

Using Swensen's lemma and assumption (S6) in (Haddad and Belaide [40])

$$2\sum_{t=1}^{n} Z_{n,t} \quad is\, asymptotically \qquad \mathcal{N}(\mu, CV)$$

$$\mu = \begin{pmatrix} \left(\frac{d^2}{1-a^2}\right)^{\frac{1}{2}} [(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{4}} \\ \sum_{k=1}^{\infty} \left(\frac{a^{2k}}{k^2}\right)^{\frac{1}{2}} (I_f^2 + C_0 I_f^{\frac{2}{p}})^{\frac{1}{2}} (\sigma^4 + C_0\sigma^{\frac{4}{p}})^{\frac{1}{4}} \end{pmatrix}$$

$$CV = \left((I_f^2 + C_0 I_f^{\frac{2}{p}})^{\frac{1}{2}} (\sigma^4 + C_0\sigma^{\frac{4}{p}})^{\frac{1}{2}} \Xi(\theta)\right)$$

Where μ is the mean and CV is the covariance matrix.

3.3 Simulation study

In order to check the validity of the theoretical results presented above we conduct numerical simulations.

Firstly, we deal with central sequence then we show the local asymptotic normality using equation 3.3. For each Δ_f^n and $k \leq n$ we generate n sequence of geometrical strong mixing errors, f assumed to be normal distribution, we vary the values of $\alpha = 0, 0.15, 0.25$ and we vary the values of n.

Figures 1,2 and 3: We present the QQ-plot sample of Δ_f^n which shows that the central sequence has normal distribution limit.

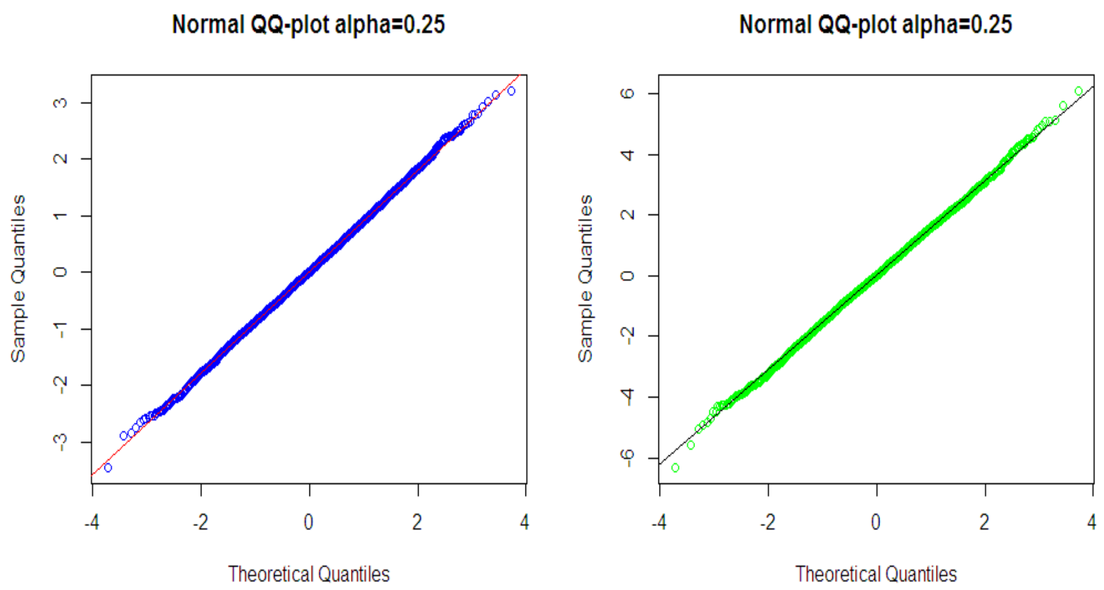

Figure 3.1: QQ-plot of Δ_f^n sample for $\alpha = 0.25$

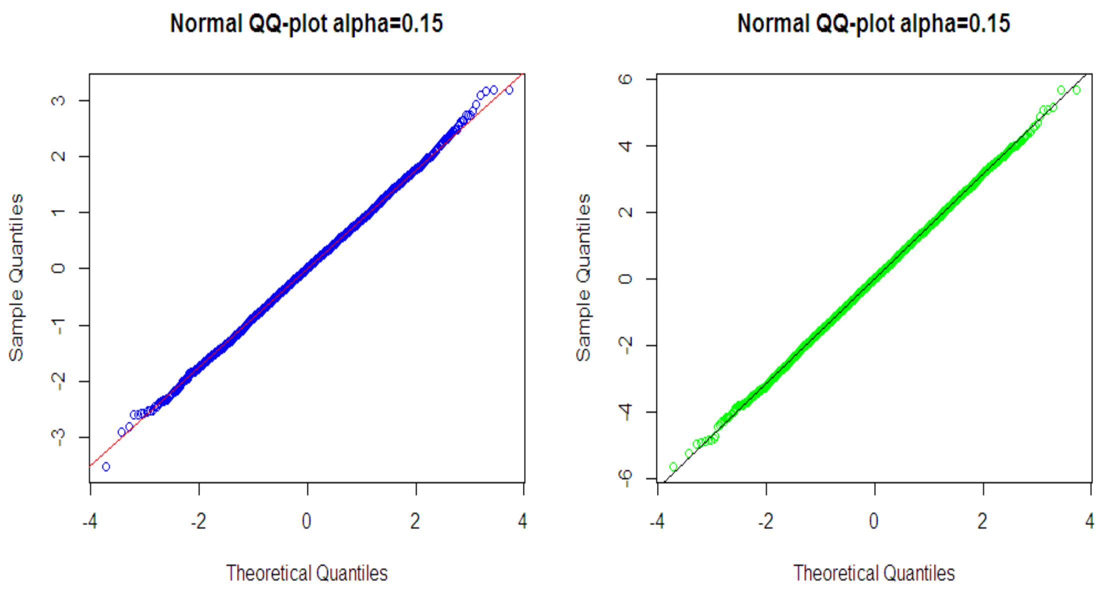

Figure 3.2: QQ-plot of Δ_f^n sample for $\alpha = 0.1$

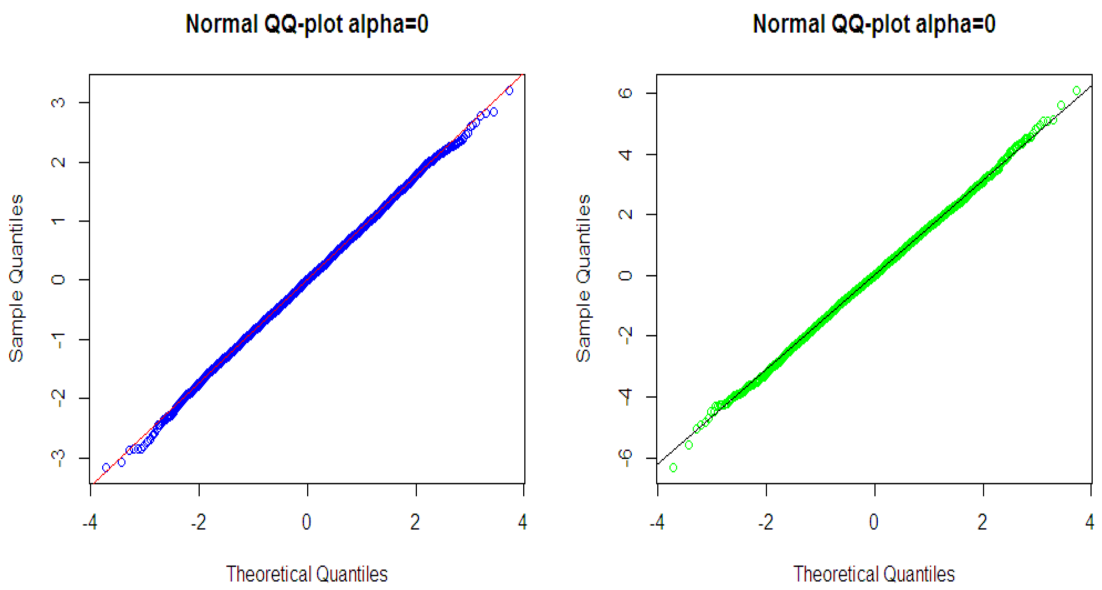

Figure 3.3: QQ-plot of Δ_f^n sample for $\alpha = 0$

Figures 4,5 and 6 : The density curve of the central sequence and the density curve of normal distribution with sample mean and sample variance are presented in figures below, we remark clearly that the two curves are very close to each other in the three graphs when α=0.25, 0.15 and also for the case $\alpha = 0$
We take the values $d = 0.4$, $a = 0.8$, $\sigma = 2$.
Therefore, the figures show that the central sequence has normal distribution.
We note that *alpha* means the value of $\alpha-$mixing parameter.

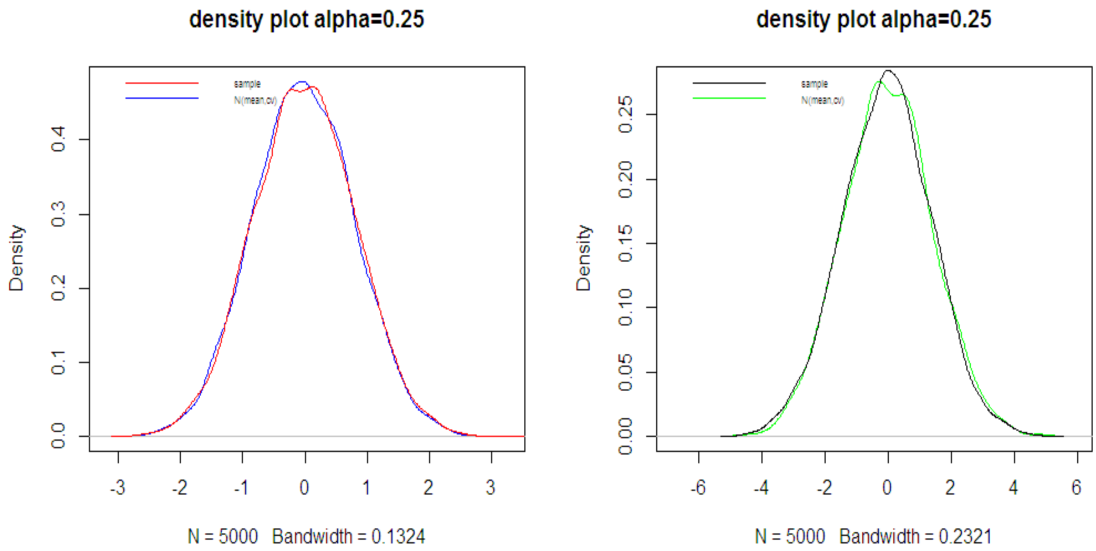

Figure 3.4: Density plot of Δ_f^n sample for $\alpha = 0.25$

To compare between the generated central sequence Δ_f^n and the gaussian process with same parameters we use Root Mean Squared Errors (RMSE) which is an important tool, it shows that our central sequence is very close to the Gaussian process $\mathcal{N}(\mu, CV)$.
nine different sample sizes are used namely $n=$ 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000.
Table: RMSE of generated central sequence and the gaussian process.

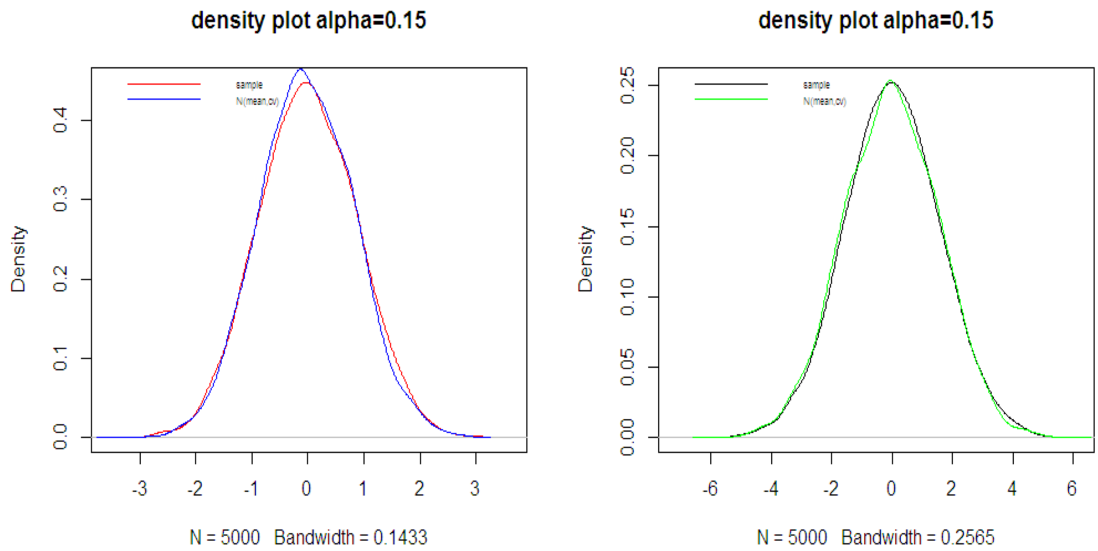

Figure 3.5: Density plot of Δ_f^n sample for $\alpha = 0.1$

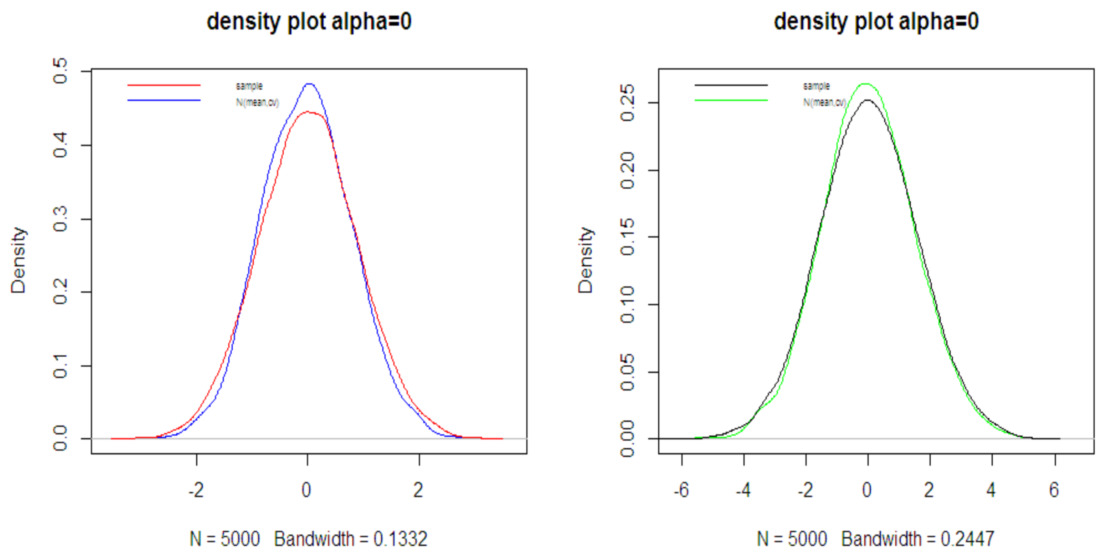

Figure 3.6: Density plot of Δ_f^n sample for $\alpha = 0$

n	500	1000	1500	2000	2500	3000	3500	4000	5000
$\alpha = 0$	3.104	2.579	2.372	2.092	1.844	1.755	1.683	1.541	1.423
$\alpha = 0$	5.389	3.962	3.644	3.055	2.891	2.717	2.582	2.453	2.201
$\alpha = 0.15$	2.914	2.142	1.808	1.614	1.516	1.411	1.366	1.293	1.259
$\alpha = 0.15$	5.126	3.742	3.279	2.872	2.698	2.598	2.444	2.385	2.225
$\alpha = 0.25$	2.914	2.142	1.808	1.599	1.449	1.375	1.299	1.246	1.178
$\alpha = 0.25$	5.086	3.686	3.218	2.816	2.618	2.474	2.418	2.318	2.160

Interpretation

- RMSE decreases when the size of n increases.

- We remark clearly that the values of RMSE are small, it means that Δ_f^n is close to the gaussian process with same parameters.

- We remark that the strong mixing coefficients has a significant influence on the behavior of LAN property, when $\alpha-$mixing coefficient is higher the RMSE is small.

- In the case of independent errors ($\alpha = 0$) the RMSE is higher than the case of dependent errors, when (α=0.25, 0.15).

3.4 Local Asymptotic Minimaxity For FAR(1) Process

Introduction

Under Local Asymptotic Normality property, an asymptotic efficiency criterion is often considered in the literature "Local Asymptotic Minimax" (LAM) property which is an important concept of optimality used in the estimation.
In this section, we treat several local asymptotic properties for Fractional Autoregressive process, we deal with Local Asymptotic Minimax and Local Asymptotic Linearity.

3.4.1 Preliminaries for Local Asymptotic Minimax

In this section, we study LAM property under Local Asymptotic Normality conditions. Let $\ell(.)$ (lost function) defined in $\mathbb{R}^2$ satisfied the following conditions:
(i) $\ell(x) > 0$ (ii) $\ell(x) = \ell(-x)$ (iii) $\{x \backslash \ell(x) \leq u\}$ is convex $\forall u \in \mathbb{R}_+^*$
We use LAN conditions to construct LAM estimator

Theorem 3.4. *Local Asymptotic Minimax Let Z_n be a sequence of an arbitrary estimators of θ and ℓ be a loss function on $\mathbb{R}$, we consider LAN condition $LAN(\theta, \Xi(\theta), \Delta_n^f(\theta))$ is satisfied*

$$\lim_{M\to\infty} \liminf_{n\to\infty} \inf_{Z_n} \sup_{\sqrt{n}|\theta-\theta_0|\leq M} \mathbb{E}(\ell\{\sqrt{n}(Z_n-\theta)\}) \geq \mathbb{E}(\ell(Z)) \tag{3.9}$$

A sequence of estimators Z_n is said to be Local Asymptotic Minimax (LAM) is the previous inequality is satisfied.

Lemma 3.1. *Under LAN conditions, for any sequences of estimators Z_n, if Z_n is $\theta-$regular, then Z_n is Local Asymptotic Minimax.*

Remark 3.1. The value of θ under the null hypothesis is unspecified, thus we replace θ by $\widehat{\theta}_n$ which is $\sqrt{n}-$consistent.

$\widehat{\theta}_n$ is $\sqrt{n}-$consistent it means that $\sqrt{n}(\widehat{\theta}_n - \theta) = O_p(1)$ if $\exists c > 0$
such that

$$\sqrt{n}|\widehat{\theta}_n - \theta| \leq c, \quad \forall n \in \mathbb{N}, \theta \in \Theta$$

We use also discrete sequence od estimators $\overline{\theta}_n$

Definition 3.1. A sequence $\overline{\theta}_n$ of estimators is said to be discrete is there is $K \in \mathbb{N}$ such that $\overline{\theta}_n$ takes at most K different values in the following set

$$Q_n = \{\theta \in \mathbb{R}^k : \sqrt{n|\theta - \theta_0| < c}\}, c > 0.$$

This property allows us to construct regular estimator using local asymptotic discrete estimator.

3.4.2 Local Asymptotic Properties for FAR(1) with Independent Noises

Local Asymptotic Minimaxity for FAR(1) is expressed in the following lemma, the errors are i.i.d noises.

Lemma 3.2. *Under the conditions of invertibility and LAN hypothesis, for all $\theta^{(n)}$*
If

$$\sqrt{n}(\theta^{(n)} - \theta) - \left(\frac{\Xi(\theta)^{-1}}{I_f}\right)\Delta_f^{(n)}(\theta) = o_{P_\theta}(1) \tag{3.10}$$

Then
$\theta^{(n)}$ is LAM under $H_f^{(n)}(\theta)$

On the other hand, result on Local Asymptotic Linearity is established in the following lemma

Lemma 3.3. *If $\theta^{(n)}$ is sequence of estimators of θ, $\sqrt{n}(\theta^{(n)} - \theta) = O_{p_\theta}(1)$*

$$\Delta_f^{(n)}(\theta^{(n)}) - \Delta_f^{(n)}(\theta) + I_f\Xi(\theta)\sqrt{n}(\theta^{(n)} - \theta) = o_{p_\theta}(1) \tag{3.11}$$

3.5 Local Asymptotic Properties for Fractional Autoregressive Model with Strong Mixing Noises

We generalize the results of Serroukh (1996) in the case of mixing noises.

3.5.1 Local Asymptotic Minimaxity

Lemma 3.4. *Under the conditions of invertibility and LAN hypothesis, for all $\theta^{(n)}$ If*

$$\sqrt{n}(\theta^{(n)} - \theta) - \left(\Xi(\theta)[(I_f^2 + C_0 I_f^{\frac{2}{p}})(\sigma^4 + C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right)^{-1} \Delta_f^{(n)}(\theta) = o_{P_\theta}(1) \tag{3.12}$$

Then
$\theta^{(n)}$ is LAM under $H_f^{(n)}(\theta)$

$\theta^{(n)}$ is said $\theta-$regular, if 3.12 is verified.

Proof

We based on Lemma (4.1) (Kreiss)[55] they dealt with the case of ARMA processes it suffices to use Davydov's inequality to obtain the result. ■
In order to construct regular estimates, we have to consider $\overline{\theta}^{(n)}$ which is $\sqrt{n}$-convergent estimators, it is essential to assume that

$$\sqrt{n}(\overline{\theta}^{(n)} - \theta) = O_{p_\theta}(1) \tag{3.13}$$

Remark 3.2. $\overline{\theta}^{(n)}$ is consistent estimator of θ and it is asymptotically equivalent to $\theta^{(n)}$.

On the other hand we establish local asymptotic linearity.

3.5.2 Local Asymptotic Linearity

We use Local Aymptotic Quadratic decomposition to deduce this property.

Lemma 3.5. *If $\theta^{(n)}$ is sequence of estimators of θ, $\sqrt{n}(\theta^{(n)}-\theta)=O_{p_\theta}(1)$*

$$\Delta_f^{(n)}(\theta^{(n)})-\Delta_f^{(n)}(\theta)+[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\Xi(\theta)\sqrt{n}(\theta^{(n)}-\theta)=o_{p_\theta}(1) \quad (3.14)$$

Proof

The proof can be deduced directly from LAN property

$$\Delta_f^{(n)}(\theta)-\Delta_f^{(n)}(\theta+n^{-1/2}\iota^{(n)})=\Xi_f(\theta)\iota^{(n)}+o_{P_\theta}(1) \quad (3.15)$$

We consider $\theta^{(n)}=\theta+n^{-\frac{1}{2}}(n^{\frac{1}{2}}(\theta^{(n)}-\theta))$ and $\iota^{(n)}=n^{\frac{1}{2}}(\theta^{(n)}-\theta)$

We replace $\theta^{(n)}$ and $\iota^{(n)}$ by the expressions above, we get

$$\Delta_f^{(n)}(\theta)-\Delta_f^{(n)}(\theta^{(n)})=\Xi_f(\theta)n^{\frac{1}{2}}(\theta^{(n)}-\theta)+o_{P_\theta}(1) \quad (3.16)$$

Proposition 3.5. *We consider $\overline{\theta}^{(n)}$ a sequences of estimators $\sqrt{n}-$convergent and Locally asymptotically discrete of θ, then*

$$\widehat{\theta}^{(n)}=\overline{\theta}^{(n)}+\frac{1}{\sqrt{n}}\left(\widehat{\Xi}^{(n)}(\overline{\theta}^{(n)})[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\right)^{-1}\Delta_f^{(n)}(\overline{\theta}^{(n)}) \quad (3.17)$$

is $\theta^{(n)}-$regular.

Where $\widehat{\Xi}(\theta)=\frac{1}{n}\sum_{t=1}^{n}Y_{t-1}(\theta)(Y_{t-1}(\theta))'$ *With* $Y_{t-1}(\theta)=\begin{pmatrix} d\sum_{j=1}^{q}a^{j-1}Z_{t-j-1} \\ \sum_{j=1}^{q}\frac{a^j}{j}Z_{t-j}\end{pmatrix}$

Proof

If 3.12 is verified, then $\widehat{\theta}_f^{(n)}$ is $\theta-$regular

$$\sqrt{n}(\widehat{\theta}^{(n)}-\theta)-\frac{(\Xi(\theta))^{-1}}{[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}}\Delta_f^{(n)}(\theta)=$$

$$\sqrt{n}(\overline{\theta}^{(n)}-\theta)-\frac{\left(\widehat{\Xi}^{(n)}(\overline{\theta}^{(n)})\right)^{-1}}{[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}}\Delta_f^{(n)}(\overline{\theta})-\frac{(\Xi(\theta))^{-1}}{[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}}\Delta_f^{(n)}(\theta)=$$

$$\frac{\left(\widehat{\Xi}^{n}(\overline{\theta}^{(n)})\right)^{-1}}{[(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}}([(I_f^2+C_0I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\widehat{\Xi}^{(n)}(\overline{\theta}^{(n)})\sqrt{n}(\overline{\theta}^{(n)}-\theta)$$

$$+\Delta_f^{(n)}(\overline{\theta})-\Xi^{-1}(\theta)\widehat{\Xi}^{(n)}(\overline{\theta}^{(n)})\Delta_f^{n}(\theta))=$$

$$\frac{\left(\widehat{\Xi}^{(n)}(\overline{\theta}^{(n)})\right)^{-1}}{[(I_f^2+C_0 I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}}([(I_f^2+C_0 I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}(\widehat{\Xi}^n(\overline{\theta}^{(n)})-\Xi(\theta))\sqrt{n}(\overline{\theta}^{(n)}-\theta)$$
$$+\Delta_f^{(n)}(\overline{\theta}^{(n)})-\Delta_f^{(n)}(\theta)+[(I_f^2+C_0 I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\Xi(\theta)\sqrt{n}(\overline{\theta}^{(n)}-\theta)+(1-\Xi^{-1}(\theta)\widehat{\Xi}^{(n)}(\overline{\theta}^n))\Delta_f^{(n)}(\theta))$$

after a suitable decomposition we use local asymptotic linearity of $\Delta_f^{(n)}(\theta)$ and lemma 3.5 which implies

$$\Delta_f^{(n)}(\theta^{(n)})-\Delta_f^{(n)}(\theta)+[(I_f^2+C_0 I_f^{\frac{2}{p}})(\sigma^4+C_0\sigma^{\frac{4}{p}})]^{\frac{1}{2}}\Xi(\theta)\sqrt{n}(\theta^{(n)}-\theta)=o_{p_\theta}(1)$$

We know that $(\Xi(\theta)-\widehat{\Xi}(\overline{\theta}))\Delta_f^{(n)}(\theta)=o_{P_\theta}(1)$.

We combine the results above to prove proposition 3.5.

3.6 Conclusion

In this chapter we have achieved the main goal which is the study of local asymptotic normality for fractional autoregressive model with strong mixing noises, we have proved asymptotic normality of the central sequence of our model using Swensen's lemma, then we have showed that FAR(1) satisfies local asymptotic minimaxity property and local asymptotic linearity. Finally we have used numerical simulation to check that the central sequence obeys a standard law and this result is elaborated theoretically.

General Conclusion

To sum up, this thesis a study on Fractional AutoRegressive model of order 1 with dependent errors, we have focused o the case of mixing errors because it is fruitful and more realistic to assume the dependency of the errors as we have done in our work. We have discussed several probabilistic and statistical properties such as autocovariance function, autocorrelation function and their asymptotic properties which allows us to detect the effect of mixing errors on the behavior of the model studied, we have implemented a simulation study in order to check the validity of the theoretical results and show us clearly the difference between the behavior of independent case and dependent case (strong mixing errors with different values)
Furthermore, we have studied an important asymptotic theory which is known as Local Asymptotic Normality (LAN) property basing on Swensen's lemma, then we have made a simulation study to prove the established results.
We have dealt with some local asymptotic properties such as (Local Asymptotic Minimax (LAM), Local Asymptotic Linearity), these properties are crucial in the construction of the adaptive estimator.
This work answered concisely and clearly the main research question and show us the impact of the dependency of the errors, this property was especially studied in finance and economic.

Perspectives

It will be interested to deal with

- Construction of the adaptive estimator of the model studied in this thesis.
- Extension of the results to the case of associated errors.
- Adaptive estimation for GARMA process with mixing errors.
- Local Asymptotic Normality for Tempered ARFIMA with mixing errors.
- Local Asymptotic Normality for superior order of ARFIMA process.

Bibliography

[1] Abramowitz, M., Stegun, I. A. (Eds.). (1964). Handbook of mathematical functions with formulas, graphs, and mathematical tables (Vol. 55). US Government printing office.

[2] Amimour, A., Belaide, K. (2020). Local asymptotic normality for a periodically time varying long memory parameter. Communications in Statistics-Theory and Methods, 1-17.

[3] Andel, J. (1986). Long memory time series models. Kybernetika (Prague) 22 105-123.

[4] Andreoni, A., Postorino, M. N. (2006). A multivariate ARIMA model to forecast air transport demand. Proceedings of the Association for European Transport and Contributors, 1-14.

[5] Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics.

[6] Bentarzi, M., Guerbyenne, H., Merzougui, M. (2009). Adaptive estimation of causal periodic autoregressive model. Communications in Statistics-Simulation and Computation, 38(8), 1592-1609.

[7] Bentarzi, M., Merzougui, M. (2009). Adaptive test for periodicity in self-exciting threshold autoregressive models. Communications in Statistics-Simulation and Computation, 38(8), 1723-1741.

[8] Bentarzi,M. Merzougui,M.(2010).Adaptive Test for Periodicity in Autoregressive Conditional Heteroskedastic Processes. Comm.Simulation Comput. 39,1735-1753.

[9] Beran, J. (1994). Statistics for Long-Memory Processes: Volume 61 of Monographs on statistics and applied probability. Chapman Hall, New York.

[10] Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. Journal of the Royal Statistical Society: Series B (Methodological), 57:659-672.

[11] Bickel, P. J. (1982). On adaptive estimation. The Annals of Statistics, 647-671.

[12] Birchenhall, C. R., Bladen-Hovell, R. C., Chui, A. P., Osborn, D. R., Smith, J. P. (1989). A seasonal model of consumption. The Economic Journal, 99(397), 837-843.

[13] Bloomfield, An, S., Pantula, S. P. (1992) . Asymptotic properties of the MLE in fractional ARIMA models. Institute of Statistics Mimeograph Series No. 2228, North Carolina State University

[14] Boswijk, H. P., Franses, P. H., Haldrup, N. (1997). Multiple unit roots in periodic autoregression. Journal of Econometrics, 80(1), 167-193.

[15] Bosq, D. (2000). Linear processes in function spaces: theory and applications (Vol. 149). Springer Science and Business Media.

[16] Box, G.E.P. and Jenkins, G.M. (1970) Time series analysis: forecasting and control. Holden-Day, San Francisco.

[17] Box, G.E.P. and Jenkins, G.M. (1976) Time series analysis, forecasting and control. The Clarendon Press, San Francisco : Holden-Day, 2nd edition.

[18] Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (2008). Time series analysis : forecasting and control. Wiley Series in Probability and Statistics, Oxford, 4th edition edition.

[19] Cox, D. R. (1965). On the estimation of the intensity function of a stationary point process. Journal of the Royal Statistical Society: Series B (Methodological), 27(2), 332-337.

[20] Brockwell, P. J., Davis, R. A. (1991). Stationary time series. In Time Series: Theory and Methods (pp. 1-41). Springer, New York, NY.

[21] Brockwell, P. J., Davis, R. A. (2002) Introduction to Time Series and Forecasting.Springer,NewYork.

[22] Chen, G., Abraham, B., Peiris, S. (1994). Lag window estimation of the degree of differencing in fractionally integrated time series models. Journal of Time Series Analysis, 15(5), 473-487.

[23] Chung, C.-F. (1996). A generalized fractionally integrated autoregressive moving-average process. J. Time Series Anal. 17 111-140.

[24] Davis, P. J. (1959). "Leonhard Euler's Integral: A Historical Profile of the Gamma Function". American Mathematical Monthly. 66 (10): 849-869.

[25] Diebold, F. X., Rudebusch, G. D. (1989). Long memory and persistence in aggregate output. Journal of monetary economics, 24(2), 189-209.

[26] Doan, T. A., Litterman, R. B,. (1986). Forecasting with Bayesian vector autoregressions-five years of experience. Journal of Business and Economic Statistics, vol. 4, no. 1 (January 1986), 25-38.

[27] Doukhan, P. (1994). Mixing: Properties and examples, lecture notes in statistics. Vol. 85. Berlin,Germany: Springer.

[28] Fabian, V., Hannan, J. (1982). On estimation and adaptive estimation for locally asymptotically normal families. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 59(4), 459-478.

[29] Ferraty, F. and Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statist. 16, 111-27.

[30] Garel, B., Hallin, M. (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend. Annals of the Institute of Statistical Mathematics, 47(3), 551-579.

[31] Geweke, J., Porter-Hudak, S. (1983). The estimation and application of long memory time series models. Journal of time series analysis, 4(4), 221-238.

[32] Giraitis, L. and Leipus, R. (1995). A generalized fractionally differencing approach in long memory modeling. Liet. Mat. Rink. 35 65-81.

[33] Gonçalves, E. (1987). Une généralisation des processus ARMA. Annales d'Economie et de Statistique, 109-145.

[34] Granger, C. W., Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. Journal of time series analysis, 1(1), 15-29.

[35] Granger, C. W. (1982). Generating mechanisms, models, and causality. Advances in econometrics, 237-253.

[36] Granger, C. W., Ding, Z. (1996). Varieties of long memory models. Journal of econometrics, 73(1), 61-77.

[37] Gray, H. L., Zhang, N. F., Woodward, W. A. (1989). On generalized fractional processes. J. Time Series Anal. 10 233-257.

[38] Guta, M., and Kiukas, J. (2015). Equivalence classes and local asymptotic normality in system identification for quantum Markov chains. Communications in Mathematical Physics 335 (3): 1397-428.

[39] Hamilton, J., D. (1994). Time Series Analysis, Princeton N.J, Princeton University Press, 799 p.

[40] Haddad, S., Belaide, K. (2020). Local asymptotic normality for long-memory process with strong mixing noises. Communications in Statistics-Theory and Methods, 49(12), 2817-2830.

[41] Haddad, S. Processus de Long Mémoire à Erreurs Mélangeantes. université de Bejaia.

[42] Hajek, J. and Sidak, Z. (1967). Theory of rank tests. Academic Press, New York, second edition edition. .

[43] Heyde, C. C. (1980). Martingale limit theory and its application (No. 515.62 H3).

[44] Hallin, M., Serroukh, A. (1998). Adaptive Estimation of the Lag of a Long-memory Process. Statistical inference for stochastic processes, 1(2), 111-129.

[45] Hallin, M., and Puri, M. L. (1994). Aligned rank tests for linear models with autocorrelated error terms. Journal of Multivariate Analysis 50 (2):175-237.

[46] Hooker, R. H. (1901). Correlation of the marriage-rate with trade. Journal of the Royal Statistical Society, 64(3), 485-492.

[47] Hosking., J.R.M. (1981). Fractional differencing. Biometrika, 68:165-176.

[48] Huang, S.C. 2010. Return and Volatility Contagions of Financial Markets over Difference Time Scales. International Research Journal of Finance and Economics, 42, 140-148.

[49] Hurst, H. E. (1951). Long-term storage capacity of reservoirs. Transactions of the American society of civil engineers, 116(1), 770-799.

[50] Jeffreys, H. (1939). Random and systematic arrangements. Biometrika, 31(1/2), 1-8.

[51] King, W. I. (1912). The elements of statistical method. Macmillan.

[52] Doan, T., Litterman, R., Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. Econometric reviews, 3(1), 1-100.

[53] Kara-Terki, N., and Mourid, T. (2016). On local asymptotic normality for fractional autoregressive. processes. Journal of Multivariate Analysis 148:120-40.

[54] Karr. A. F. (1993). Probability. Springer-Verlag, New York.

[55] Kreiss, J. P. (1987). On adaptive estimation in autoregressive models when there are nuisance functions. Statistics and Risk Modeling, 5(1-2), 59-76.

[56] Kreiss, J. P. (1987). On adaptive estimation in stationary ARMA processes. The Annals of Statistics, 112-133.

[57] Le Cam., L. (1960). Locally asymptotically normal families of distributions; certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. University of California publications in statistics, v. 3, no. 2, Berkeley, university of california press edition.

[58] Le Cam., L. (1968) Théorie asymptotique de la décision statistique. Séminaire de Mathématiques Supérieures. Les Presses de l'Université de Montréal, Montréal, Québec.

[59] Le Cam., L. (1986) Asymptotic methods in statistical decision theory. Springer-Verlag, New York.

[60] Mandelbrot, B. B., Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. SIAM review, 10(4), 422-437.

[61] Newcomb, S. (1886). Generalized theory of the comination of observations soas to obtain the best result. American Journal of Mathematics, 8:343-66.

[62] Osborn, D.R., 1988, Seasonality and habit persistence in a life-cycle model of consumption, Journal of Applied Econometrics 3, 255-266.

[63] Park, K. I., Park, M. (2018). Fundamentals of probability and stochastic processes with applications to communications (pp. 51-72). Springer International Publishing.

[64] Peiris, M. S. (1984). Non-stationary first order autorgression:some multivariate extensions. Statistics Research Report No.92.

[65] Pillai, T. R., Shitan, M., Peiris, M. S. (2009). Time series properties of the class of first order autoregressive processes with generalized moving average errors. Journal of Statistics: Advances in theory and applications, 2(1), 71-92.

[66] Pillai, T. R., Shitan, M., Peiris, S. (2012). Some Properties of the Generalized Autoregressive Moving Average (GARMA (1, 1; d1, d2)) Model. Communications in Statistics-Theory and Methods, 41(4), 699-716.

[67] Pipiras, V., Taqqu, M. S. (2000). Integration questions related to fractional Brownian motion. Probability theory and related fields, 118(2), 251-291.

[68] Rosenblatt, M. (1956) A Cental limit theorem and a strong mixing condition. Proc Natl Acad Sci USA. doi: 10.1073/pnas.42.1.43

[69] Rio, E. (2000). Théorie asymptotique des processus aléatoires faiblement dépendants, mathéematiques and applications. Vol. 31. Berlin, Germany: Springer-Verlag.

[70] Sabzikar, F., Meerschaert, M. M., Chen, J. (2015). Tempered fractional calculus. Journal of Computational Physics, 293, 14-28.

[71] Serfling., R.J. (1980). Approximation Theorems of Mathematical Statistics. Wiley Series m Probability and Statistics, New York.

[72] Serroukh, A. (1996). Inférence asymptotique paramétrique et non paramétrique pour les modèles ARMA fractionnaires. PhD thesis, Institut Statist. Univ. Libre Bruxelles.

[73] Shitan, M., Peiris, S. (2011). Time series Properties of the class of generalized first-order autoregressive processes with moving average errors. Communications in Statistics-Theory and Methods, 40(13), 2259-2275.

[74] Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of econometrics, 53(1-3), 165-188.

[75] Slutsky, E. (1927). The summation of random causes as the source of cyclic processes. Econometrica 5 105-146.

[76] Smith, H. F. (1938). An empirical law describing heterogeneity in the yields of agricultural crops. The Journal of Agricultural Science, 28(1), 1-23.

[77] Stein, C. (1956). Efficient nonparametric testing and estimation. In Proceedings of the third Berkeley symposium on mathematical statistics and probability (Vol. 1, pp. 187-195).

[78] Student. (1927). Errors of routine analysis. Biometrika, 151-164.

[79] Swensen, A.R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. Journal of Multivariate Analysis, 16:54-70.

[80] Taqqu, M. S. (1978). A representation for self-similar processes. Stochastic Processes and their Applications, 7(1), 55-64.

[81] Tsitsika, E. V., Maravelias, C. D., Haralabous, J. (2007). Modeling and forecasting pelagic fish production using univariate and multivariate ARIMA models. Fisheries science, 73(5), 979-988.

[82] Van der Vaart, A. W. (2000). Asymptotic statistics (Vol. 3). Cambridge university press.

[83] Wald. A., (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, 54:426-482.

[84] Wegman, E. J. (1974). Some results on non stationary first order autoregression. Technometrics, 16(2), 321-322.

[85] Wold, H. (1938). A study in the analysis of stationary time series (Doctoral dissertation, Almqvist and Wiksell).

[86] Woodward, W. A., Cheng, Q. C. and Gray, H. L. (1998). A k-factor GARMA long-memory model. J. Time Series Anal. 19 485-504.

[87] Lin, H. C., Xirasagar, S. (2006). Seasonality of hip fractures and estimates of season-attributable effects: a multivariate ARIMA analysis of population-based data. Osteoporosis International, 17(6), 795-806.

[88] Yajima, Y. (1985). On estimation of long-memory time series models. Australian Journal of Statistics, 27(3), 303-320.

[89] Yule, G. U. (1926). Why do we sometimes get nonsensecorrelations between time-series? A study in sampling and the nature of time-series. J. Roy. Statist. Soc. 89 1-63.

[90] Zhou, Z., Lin, Z. (2016). Asymptotic normality of locally modelled regression estimator for functional data. Journal of Nonparametric Statistics, 28(1), 116-131.

Résumé

Dans cette thèse, nous avons étudié le processus Autoregressive moyenne mobile fractionnaire (ARFIMA), nous nous concentrons sur le modèle le plus simple Autoregressive fractionnaire d'ordre 1 FAR(1), ce modèle est souvent appliqué dans plusieurs domaines en hydrologie, économie, finance … etc. Dans notre cas, les erreurs étant supposées dépendantes, nous mettons en lumière les erreurs mélangeants notamment le coefficient de mélange fort . Nous avons établi diverses propriétés statistiques et probabilistes telles que la fonction d'autocorrélation et son comportement asymptotique ; on remarque l'effet du coefficient de mélange sur le comportement des propriétés probabilistes. Cela peut être clairement démontré dans une étude de simulation. De plus, plusieurs propriétés asymptotiques locales ont été principalement discutées, la normalité asymptotique locale (LAN) pour FAR(1) avec des erreurs mélange a été prouvée en utilisant les conditions de Swensen, puis nous avons traité la minimaxité asymptotique locale (LAM) et la linéarité asymptotique locale, ces propriétés sont très utiles pour étudier l'optimalité des estimateurs et des tests.

Mots Clés : Procesuss autoregressive fractionnaire, Autocorrélation, Erreurs mélangeants, Normalité Asymptotique locale, Normalité asymptotique minimax, normalié asymptotique linéaire, simulation.

Abstract

In this thesis, we have studied Fractional Autoregressive integrated Moving Average (ARFIMA) process, we focus on the simplest model which is Fractional Autoregressive process of order 1 FAR(1), this model has a large number of applications in hydrology, economic, finance… etc. In our case, the errors are assumed to be dependent, we shed the light on mixing errors especially strong mixing coefficient which is the strongest mixing coefficient. We have established various statistical and probabilistic properties such as autocorrelation function and its asymptotic behavior; we remark the effect of mixing coefficient on the behavior of the probabilistic properties. This can be shown clearly in a simulation study. Moreover, several local asymptotic properties have been mainly discussed, local asymptotic normality (LAN) for FAR(1) with strong mixing noises have been proved using Swensen's conditions, then we have dealt with local asymptotic minimaxity (LAM) and local asymptotic linearity, these properties are very useful to study the optimality of the estimators and tests.

Key words: Fractional Autoregressive process, Autocorreltion, mixing errors, Local asymptotic normality, Local asymptotic minimaxiy, local asymptotic linearity, simulation.

الملخص

في هذه الأطروحة، درسنا عملية المتوسط المتحرك الجزئي المتكامل،ونركز على أبسط نموذج وهو عملية الانحدار الذاتي الجزئي، هذا النموذج يحتوي على عدد كبير من التطبيقات في الري ،الاقتصاد ،المالية ... إلخ. في حالتنا,من المفترض أن تكون الأخطاء مرتبطة ، فقد قمنا بإلقاء الضوء على أخطاء الخلط وخاصة معامل الخلط القوي وهو أقوى معامل خلط. لقد أنشأنا العديد من الخصائص الإحصائية والاحتمالية مثل وظيفة الارتباط الذاتي وسلوكها المقارب ؛ نلاحظ تأثير معامل الخلط على سلوك الخواص الاحتمالية. يمكن أن يظهر هذا بوضوح في دراسة المحاكاة.علاوة على ذلك ، تمت مناقشة العديد من الخصائص المقاربة المحلية مع ضوضاء الخلط القوية باستخدام ظروف، ثم تعاملنا مع الحد لـ بشكل أساسي ، حيث تم إثبات الحالة الطبيعية المقاربة المحلية والخطية المقاربة المحلية و الحد الأدنى المقارب، هذه الخصائص مفيدة جدًا لدراسة أمثلية المقدرات والاختبارات .

الكلمات المفتاحية : نموذج الانحدار الذاتي, الارتباط الذاتي, الأخطاء المخلوطة, الحالة الطبيعية المقاربة المحلية, الحالة الخطية المقاربة المحلية, الحاد الأدنى المقارب المحلي