N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBÈS

THESE DE DOCTORAT

Présentée par :

CHEBBAB IKHLASSE

Domaine : Mathématiques Informatique
Filière : Mathématiques
Intitulé de la formation : Statistique, Mathématiques appliquées à l'économie et à la finance

Intitulée

Inégalité de concentration pour des suites des variables aléatoires, applications aux modèles statistiques

Soutenue le :04\03\2020

Devant le jury composé de :

Président :
Mr MECHAB Boubaker M.C. A à L'Université S.B.A
Examinateurs :
Mr AZZOUZI Badreddine M.C.A Ecole Supérieure de Management de Tlemcen
Mr CHOUAF Abdelhak M.C.A à L'Université S.B.A
Directeur de thèse :
Mr BENAISSA Samir Professeur à L'Université S.B.A

Année universitaire : 2019/2020

Acknowledgements

First, I want to thank ALLAH, the Almighty for giving us courage and will to accomplish this modest work.

I want to express my deep sense of respect and gratitude to my supervisor Mr. Benaissa Samir for his encouragement and guidance throughout the preparation of this work.

I would also like to thank Mr. MECHAB BOUBAKER for honoring me by presiding over the jury of this thesis. He always advised me and preciously help.

I warmly thank Mr. BADREDDINE AZZOUZI for the trust he has shown me in being part of this jury, I would also like to thank Mr. CHOUAF ABDELHAK for the interest, he was kind enough to grant to my work by accepting to participate in the jury.

I also wish to express my gratitude to all my teachers.
I thank the members of the Stochastic Statistics and Processes Laboratory from the Djillali Liabès University of Sidi Bel Abbès.

It is important for me to thank my family who have always been an inexhaustible source of encouragement.
Thank you to everyone who contributed directly or indirectly to the success of this work.

To my parents,
and to all those who are close to me.

Contents

List of Figures

Liste d'abréviation

AR:	Autoregressive.
MA:	Moving average.
ARMA:	Autoregressive moving average.
ARIMA:	Autoregressive integrated moving average.
ARH:	Autoregressive hilbertian.
ARB:	Autorégressif Banachique.
ARCH:	Autoregressive conditional heteroskedasticity.
GARCH:	Generilized autoregressive conditional heteroskedasticity.
EGARCH:	Exponential generilized autoregressive conditional heteroskedasticity.
TARCH:	Threshold autoregressive conditional heteroskedasticity.
TGARCH:	Threshold generilized autoregressive conditional heteroskedasticity.
END:	Extended negatively dependent.
UEND:	Extended negatively dependent.
LEND:	Extended negatively dependent.
WOD:	Widely orthant dependent.
WUOD:	Widely upper orthant dependent.

WUOD: Widely lower orthant dependent.

ACF: Autocorrelation function.

PACF: partiel autocorrelation function.

Publication and Communication

International Publication:
I.Chebbab. (2018): Probability Tail for Extended Negatively Dependent Random Variables of Partial Sums and Application to AR (1) Model Generated By END Errors. International Journal of Statistics and Economics, VL 19 No. 1, 1-9.

Ikhlasse Chebbab, Samir Benaissa. (2019): Complete convergence and Maximal inequalities for product sums of WOD sequences. International Journal of Statistics and Economics, VL 20 No. 1, 1-8.

Ikhlasse Chebbab, Samir Benaissa (2019): The Complete Convergence for the Parameter Estimator of the First-Order Autoregressive Process Created by WOD Errors. International Journal of Statistics and Economics, VL 20 No. 1, 80-87.

International Communication:
La Première Conférence Internationale sur «L'évolution des mathématiques contemporaines et leur impact sur les sciences et la technologie», 09-12 octobre 2017, Constantine, Algérie
Exponential inequalities for a LNQD autoregressive process.

National Communication:
Congrès des Mathématiciens Algériens CMA' 2018, 12-13 mai 2018, Boumerdès, Algérie
Probability inequalities for partial sums of ELNQD sequence and application to autoregressive process.

General Introduction

Despite our constant efforts to address the ideas that surround our vast world, there are usually questions that intrigue the researchers. Values in the future depend are usually stochastically, on the observations currently available. Such dependence must be taken into account when predicting the future of its past. In general, statisticians consider time series as an achievement of a stochastic process. There is an infinity of stochastic processes that can generate the same observed data, as the number of observations is always infinite. Random processes have provided models for analyzing many data. They have been used in astronomy (in the periodicity of sunspots, 1906), in meteorology (time-series regression of sea level on weather, 1968), in signal theory (Noise in FM receivers, 1963), in biology (the autocorrelation curves of schizophrenic brain waves and the power spectrum, 1960), in economics (time-series analysis of imports, exports and other economic variables, 1971)... and so on.

Modern techniques for the study of time series models have been started by Yule [44], and the most popular class of these linear models consists of autoregressive moving average models, including purely autoregressive models (AR). and purely mobile (MA) as special cases.

ARMA models are frequently used to model linear dynamic structures, to serve as vehicles for linear prediction. A particularly useful class of models contains autoregressive integrated moving average (ARIMA) models, which include stationary ARMA processes as a subclass.

Two articles in 1927 opened a study on autoregressive processes and moving averages: the article by Yule [44] and Slutsky [34]. Yule introduced the autoregressive models into the literature, the processes introduced by Yule will become the AR (p) processes and those introduced by Slutsky the MA(q) processes.

An autoregressive process is a regression model in which the series is explained by its past values rather than by other variables and as it plays an important role in predicting the problems leading to decision making are used to model time series in many fields, in biology, climatology, econometrics, finance, medicine, meteorology and many other fields. For example, in finance, we are interested in modeling the exchange rate of a currency. In meteorology, scientists for example model the temperature in the last month to predict the temperature it will do tomorrow. The idea is to take a sample of data and build the best model that adjusts that data. This model allows us to draw some conclusions about the series.

Autoregressive processes assume that each point can be predicted by the weighted sum of a set of previous points, plus a random error term.

The moving average processes assume that each point is a function of the errors in the previous points, plus its own error.

The analogy between the two processes will even be pushed further when it will be shown that the processes AR (p) and MA (q) are respectively MA (∞) and AR (∞) processes , in certain conditions. More generally one can prove that all AR (p) can have a representation MA(1) and

in a dual way, can also express any MA(q) as a AR(∞). the autoregressive moving average process (ARMA) is a tool for understanding and possibly predicting the future values of this series. The model is generally noted as ARMA(p, q), where p is the order of the AR part and q is the order of the MA part. This process is very used because of its simplicity.
The ARMA model only allows the so-called stationary series to be processed.
ARIMA models allow non-stationary series to be processed after determining the level of integration.
An ARIMA model is described as ARIMA model (p,d,q), in which :
p is the number of autoregressive terms,
d is the number of differences,
q is the number of moving averages.
The hypothesis of independent observations is often subjective, even erroneous, because it does not reflect the exact evolution of the random phenomenon. Indeed, the dependent observations are more adapted to reality. Dependent random variable concepts are very useful in reliability theory and applications. There are many concepts of dependence between them:
A notion of dependence is the so-called extended negatively dependent (END) introduced by Liu [27], the random variables are said to be END if they are at the same time upper extended negatively dependent (UEND) and lower extended negatively dependent(LEND).
The independent random variables and the NOD random variables are END, but the END random variables are much smaller than the independent random variables.
Another notion of dependence is this one called widely orthant dependent (WOD) was defined by Wang, Wang and Gao [37], the random variables are said to be WOD when both are widely upper orthant dependent (WUOD) and widely lower orthant dependent (WLOD). WOD random variables are lower than NA random variables, NSD random variables, NOD random variables and END random variables.

Inequalities of concentration are inequalities that limit the probability of deviation of a random variable from its mean by a certain value. Research on this topic has recently increased due to many applications in areas such as machine learning and random graphs.
Inequalities of concentration include inequalities in Bernstein [7], Bennett[4] and Hoeffding[23], and others.

Our thesis is presented in five chapters. In the first chapter we recall the notations and the tools used on the stochastic processes, and in particular, we recall the definition of the processes AR(p), MA(q), ARMA (p,q), ARIMA(p,d,q), ARCH(q) and GARCH(p,q).

In the second chapter, we consider the concentration inequalities of sums of random variables, independent case. We study the Bernstein's inequalities, Hoeffding's inequalities and Bennett's inequalities.

In The third chapter, we establish concentration inequalities for END random variables of partial sums. Using these inequalities, we show the complete convergence for parameter estimator θ_n of the first-order autoregressive process in the case where the error (white noise) ϵ_t are END. This work was the subject of an accepted publication in the International Journal of Statistics and Economics.

In the fourth chapter, we construct exponential inequalities for a new case of dependence WOD for the random variables of partial sums. Using these inequalities, we proof the complete convergence of the least square estimators θ_n of the first-order autoregressive process in the case where the error (white noise) ϵ_t are WOD. The results of this chapter are published in the International Journal of Statistics and Economics.

The last chapter is devoted to the study the complete convergence for the maximum of product sums and the moment inequality of Rosenthal type for sums of products, in the case widely orthant dependent. This work was the subject of an accepted publication in the International Journal of Statistics and Economics.

Chapter 1

Introduction

The main objective of this chapter is examined, all the important destinations of the analysis of the autoregressive processes and we present to ourselves an integrated view of the explanation of the frequently used AR, MA, ARMA, ARIMA and ARCH/GARCH models after we quote properties and important results.

1.1 Discrete operator

We consider the set of numerical sequence and on this set, we will define the backshift operators B, in advance F, difference r and summation S.

1. **Lag operators L**
   the lag operator (L) or backshift operator (B) is defined by

$$LX_t = X_{t-1}, \ \ \forall t$$

   or similarly

$$BX_t = X_{t-1}, \ \ \forall t$$

   or equivalently

$$X_t = LX_{t+1}, \ \ \forall t$$

   in which L is the lag operator. Sometimes the symbol B for backshift is used instead. notice that the lag operator may be raised to arbitrary integer powers so that $L^k X_t = X_{t-k}$.

2. **Advance operator F**
   The advance operator is defined by

$$FX_t = X_{t+1}, \ \ \forall t$$

3. **Difference operator ∇**
   defined by

$$\nabla X_t = X_t - X_{t-1}, \;\; \forall t$$
$$\nabla X_t = (1-L)X_t,$$

   for the k-th difference operator

$$\nabla^k X_t = (1-L)^k X_t.$$

4. **Summation operator S**
   It is defined by

$$SX_t = \sum_{-\infty}^{t} X_i,$$

   or $SX_t = (1 + L + L^2 + ...)X_t$.
   S is of course linear and we have:

$$SX_t = \sum_{j=k+1}^{t} X_j + SX_j,$$

   for $t \geq k+1$

$$S^2X_t = \sum_{i=k+1}^{t} \sum_{j=k+1}^{i} X_j + S^2X_k + (t-k)SX_k.$$

1.2 Notions on stochastic processes

Stochastic processes are modelling tools, and they are applied in many different scientific fields.

Definition 1. *A stochastic process or random process is a sequence of random variables $(X_t)_{t \in T}$. If $T = \mathbb{N}$ or $T = \mathbb{Z}$ the process is discrete time, and if $T = \mathbb{R}_+^*$ or $T = \mathbb{R}$ the process is a continuous time. It should be noted that the word time used here does not necessarily mean physical time. If the random variables that make up the process are discrete, we're talking about discrete value processes, if they're continuous we're talking about continuous value processes.*

1.2.1 Stationarity

Definition 2. *Let the process $(X_t)_{t \in T}$ be said strictly (or strongly) stationary if for all $\{t_1, t_2, ..., t_n\} \in T$ and for all $k > 0$ we have*

$$F_{t_1,t_2,...,t_n}(x_1, x_2, ..., x_n) = F_{t_1+k,t_2+k,...,t_n+k}(x_1, x_2, ..., x_n). \tag{1.1}$$

Definition 3. *A process $(X_t)_{t\in T}$ is said to be weakly stationary (stationary order 2), if*

1. $\mathbb{E}[X_t] = \mu < \infty$, $\forall t$
2. $\mathbb{E}[X_t^2] < \infty$, $\forall t$
3. $cov(X_t, X_s) = cov(X_{t+h}, X_{s+h})$, $\forall t, s, h$

1.2.2 Autocovariance function and autocorrelation function

Definition 4. *The autocovariance function of the process is defined by*

$$\gamma_X(h) = cov(X_t, X_{t+h}), \forall t, h$$

The autocorrelation function is defined by

$$\rho_X(h) = \gamma_X(h)/\gamma_X(0).$$

properties:

1. $\gamma_X(0) = \sigma_X^2$, $\gamma_X(0) \geq 0$,
2. $|\gamma_X(h)| \leq \gamma_X(0)$,
3. $\gamma_X(h) = \gamma_X(-h)$.

an autocorrelation function has all the properties of an autocovariance function, with $\rho_X(0) = 1$.

Definition 5. *Let $\{\zeta_t, t \geq 0\}$ a sequence of random variables constitute a weak white noise (resp strong) if*

1. $\mathbb{E}[\zeta_t] = 0, \quad \forall t \in \mathbb{N}$
2. $\mathbb{E}[\zeta_t^2] = \sigma^2$,
3. $cov(\zeta_t, \zeta_s) = 0, \quad \forall t \neq s$ *(resp ζ_t are iid).*

If $\{\zeta_t, t \geq 0\}$ is a weak white noise, we will denote by $\zeta_t \sim WN(0, \sigma^2)$.
If $\{\zeta_t, t \geq 0\}$ is a strong white noise, we will denote by $\zeta_t \sim iid(0, \sigma^2)$.
If $\{\zeta_t, t \geq 0\}$ is a Gaussian white noise, we will denote by $\zeta_t \sim N(0, \sigma^2)$.

1.2.3 Linear processes

The linear process is a stochastic process formed by a linear combination (no necessarily finite) of strong white noises, and when they are weak the linear process is general. For example, autoregressive processes (AR) and moving average processes (MA) belong to the class of linear processes.

Definition 6. *A process $(X_t)_{t\in\mathbb{Z}}$ is called linear, if it writes in the form*

$$X_t = \sum_{i=0}^{\infty} \Phi_i \zeta_{t-i}, \tag{1.2}$$

where $(\zeta)_{t\in\mathbb{Z}}$is a weak white noise of mean 0 and variance σ^2, and $\sum_{i=0}^{\infty} |\Phi_i| < \infty$.

Proposition 1. *Let $(X_t)_{t\in\mathbb{Z}}$ a stationary process solution of the following equation*

$$\Phi(B)X_t = Y_t, \quad \text{(B is the backshift operator)}$$

where $(Y_t)_{t\in\mathbb{Z}}$ is a stationary process, $\Phi(B) = \sum_{i=0}^{\infty} \Phi_i B^i$, then

1. *If Φ does not have a module root equal to 1, then there is an invertible representation of the process $(X_t)_{t\in\mathbb{Z}}$.*

2. *If all the roots of Φ are of module superior to 1, then there is a causal representation of the process $(X_t)_{t\in\mathbb{Z}}$.*

1.3 AR process

Definition 7. *An autoregressive process of order p, denoted AR (p) is defined by*

$$X_t = c + \theta_1 X_{t-1} + \theta_2 X_{t-2} + ... + \theta_p X_{t-p} + \epsilon_t, \tag{1.3}$$

where $\theta_1, \ldots, \theta_p$ are the parameters of the model, c is a constant and ϵ_t is the error associated to the process which is a random variable sequence, considered like a white noise (i.e: $\epsilon_t \backsim WN(0, \sigma^2)$) if the random variables are independent or dependent (see dependency concepts in the next chapter).
• *Using the backshift operator* **B**, *we can write it:*

$$\begin{aligned} (1 - \theta_1 B - ... - \theta_p B^p) X_t &= c + \epsilon_t, \\ \alpha(B) X_t &= c + \epsilon_t. \end{aligned}$$

Example 1. *An AR(1) process takes this form:*

$$X_t = c + \theta_1 X_{t-1} + \epsilon_t, \ \ \textit{or} \ \ \epsilon_t \backsim WN(0, \sigma^2).$$

Remark 1. *The AR (1) process can be written recursively with respect to the previous conditions*

$$\begin{aligned} X_t &= c + \theta X_{t-1} + \epsilon_t = c + \theta(c + \theta X_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= (1+\theta)c + \theta^2 X_{t-2} + \epsilon_t + \theta\epsilon_{t-1} \\ &. \\ &. \\ &. \\ &= c\sum_{i=0}^{\infty}\theta^i + \sum_{i=0}^{\infty}\theta^i\epsilon_{t-i}. \end{aligned}$$

It should be noted that the sums go here to infinity. This is because time series are often supposed to start from $t_0 = -\infty$ and not $t_0 = 0$.
Some authors, however, consider that the series starts with $t_0 = 0$ and then add the initial value X_0 in the formula.
We can see that X_t is the convoluted white noise with the kernel θ^k plus a constant mean. If the white noise is Gaussian, then X_t is also a normal process.

1.3.1 Causality and Invertibility

Definition 8. *It is said that the process is invertible if there is a real sequence d_k as $\sum_{k=0}^{\infty}|d_k| < \infty$ et*

$$\epsilon_k = \sum_{k=0}^{\infty} d_k X_{t-k}. \tag{1.4}$$

Another way of saying that a process is invertible if it has an AR (∞) representation.

Remark 2. *By this definition, any process AR(p) is invertible.*

Proposition 2. *The autoregressive process AR(p) is causal and stationary if only if its polynomial $\alpha(z)$ is such that*

$$\alpha(z) \neq 0 \ \textit{avec} \ z \in \mathbb{C} \ \textit{tel que} \ |z| \leq 1. \tag{1.5}$$

In other words, all the roots of $\alpha(z)$ are of norm greater than 1.
The proof of this proposition is in Brockwell and Davis [11].

1.3.2 Moments of an process AR(1)

To calculate the different moments of an AR (1) process, its expectancy, its variance, its autocovariance and its autocorrelation, we will assume that the white noise is independently and identically distributed, of zero expectancy and of variance σ^2 that we note ($\varepsilon_i \sim iid(0, \sigma^2)$)
Expectation:

$$\mathbb{E}[X_t] = \theta^t X_0 + c\sum_{i=0}^{t-1} \theta^i.$$

Proof 1. *(reasoning by recurrence)*
$P(0)$ *(initialization):*
$\mathbb{E}[X_0] = X_0$*, because* X_0 *is deterministic. The expression is:*

$$\theta^0 X_0 + c\sum_{i=0}^{-1} \theta^i = 1X_0 + 0 = X_0$$

$P(t+1)$ *(heredity):*
$\mathbb{E}[X_{t+1}] = \mathbb{E}[c + \theta X_t + \epsilon_t]$
Since $\mathbb{E}$ *is a linear operator:*
$\mathbb{E}[X_{t+1}] = c + \theta\mathbb{E}[X_t]$
With the induction hypothesis:

$$\begin{aligned}\mathbb{E}[X_{t+1}] &= c + \theta(\theta^t X_0 + c\sum_{i=0}^{t-1} \theta^i),\\ \mathbb{E}[X_{t+1}] &= c + \theta^{t+1} X_0 + c\sum_{i=0}^{t-1} \theta^{i+1}.\end{aligned}$$

By a change of variables in the sum, $i \to i-1$*:*

$$\mathbb{E}[X_{t+1}] = \theta^{t+1} X_0 + c + c\sum_{i=1}^{t} \theta^i$$

And, with $c = c\sum_{i=0}^{0} \theta^i$ *:*

$$\mathbb{E}[X_{t+1}] = \theta^{t+1} X_0 + c\sum_{i=0}^{t} \theta^i$$

Variance:

$$Var[X_t] = \sum_{i=0}^{\infty} \theta^{2i}\sigma^2.$$

Proof 2.

$$\begin{aligned}
Var[X_t] &= \mathbb{E}[X_t - \mathbb{E}(X_t)]^2 \\
&= \mathbb{E}\left[c\sum_{i=0}^{\infty}\theta^i + \sum_{i=0}^{\infty}\theta^i\epsilon_{t-i} - c\sum_{i=0}^{\infty}\theta^i\right]^2 \textit{(According to the results obtained in the previous page)} \\
&= \mathbb{E}\left[\sum_{i=0}^{\infty}\theta^i\epsilon_{t-i}\right]^2 \\
&= Var\left(\sum_{i=0}^{\infty}\theta^i\epsilon_{t-i}\right) (\textit{because } Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \textit{and by hypothesis } \mathbb{E}[\epsilon_t] = 0) \\
&= \sum_{i=0}^{\infty} Var(\theta^i\epsilon_{t-i}) \\
&= \sum_{i=0}^{\infty}\theta^{2i} Var(\epsilon_{t-i}) \\
&= \sum_{i=0}^{\infty}\theta^{2i}\sigma^2 \\
&= \sigma^2\sum_{i=0}^{\infty}\theta^{2i} \\
&= \sigma^2\frac{1}{1-\theta^2} \quad \textit{such that } |\theta| < 1 \quad \textit{the geometric series is convergent.}
\end{aligned}$$

Autocovariance:

$$Cov(X_t, X_{t-j}) = \theta^j\sum_{i=0}^{t-j}\theta^{2i}\sigma^2.$$

Proof 3.

$$\begin{aligned}
Cov(X_t, X_{t-j}) &= \mathbb{E}\left[[X_t - \mathbb{E}(X_t)]\left[X_{t-j} - \mathbb{E}(X_{t-j})\right]\right] \\
&= \mathbb{E}\left[\left(\sum_{i=0}^{\infty}\theta^i\epsilon_{t-i}\right)\left[\sum_{k=0}^{\infty}\theta^k\epsilon_{t-k-j}\right]\right] \\
&= \mathbb{E}\left[\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\theta^i\theta^k\epsilon_{t-i}\epsilon_{t-k-j}\right] \\
&= \sum_{i=0}^{\infty}\sum_{k=0,k+j\neq i}^{\infty}\theta^i\theta^k\mathbb{E}\left[\epsilon_{t-i}\epsilon_{t-k-j}\right] + \sum_{k=0}^{\infty}\theta^{2k+j}\mathbb{E}\left[\epsilon_{t-k-j}^2\right] \\
&= \sum_{k=0}^{\infty}\theta^{2k+j} Var\left(\epsilon_{t-k+j}\right) (\textit{by independence hypothesis of } \epsilon_m,\ \mathbb{E}[\epsilon_{t-i}\epsilon_{t-k+j}] = 0) \\
&= \theta^j\sum_{i=0}^{\infty}\theta^{2i}\sigma^2 \\
&= \theta^j\sigma^2\frac{1}{1-\theta^2}.
\end{aligned}$$

Autocorrelation:

$$Corr[X_t, X_{t-j}] = \frac{Cov(X_t, X_{t-j})}{Var(X_t)} = \theta^j.$$

Proof 4.

$$\begin{aligned} Corr[X_t, X_{t-j}] &= \frac{Cov(X_t, X_{t-j})}{Var(X_t)} \\ &= \frac{\theta^j \sum_{i=0}^{\infty} \theta^{2i}\sigma^2}{\sum_{i=0}^{\infty} \theta^{2i}\sigma^2} \\ &= \theta^j. \end{aligned}$$

1.3.3 Stationarity condition for AR(1)

We know that the parameter θ determines if the process $AR(1)$ is stationary or not, have considered the following cases:
case1: $|\theta| < 1$, the process is stationary.
case2: $|\theta| = 1$, the process is therefore non-stationary.
case3: $|\theta| > 1$, the process is explosive.

If $|\theta| < 1$, under the condition $X_0 = 0$ and according to the geometric series $\sum_{n=0}^{\infty} bq^n = \frac{b}{1-q}$ we obtain the following results,

$$\begin{aligned} \mathbb{E}[X_t] &= \frac{c}{1-\theta}. \\ Var(X_t) &= \frac{\sigma^2}{1-\theta^2}. \\ Cov(X_t) &= \frac{\theta^j}{1-\theta^2}\sigma^2 \\ Corr(X_t) &= \theta^j. \end{aligned}$$

it can be observed that the autocovariance function decreases with a rate of $\tau = -1/\ln(\theta)$. We see right here that expectancy and variance are constant and that autocovariance does not depend on time: the process is therefore stationary.

If $|\theta| = 1$, the process is written in the following form: $X_t = c + X_{t-1} + \epsilon_t$.

$$\begin{aligned} \mathbb{E}(X_t) &= ct + \mathbb{E}[X_0]. \\ Var(X_t) &= t\sigma^2. \\ Cov(X_t) &= (t-j)\sigma^2. \end{aligned}$$

Example 2. $\{AR(1)\}$
The polynomial of the backshift of a process $AR(1)$

$$X_t = \theta X_{t-1} + \epsilon_t,$$

s'écrit:

$$(1 - \theta B)X_t = \epsilon_t.$$

Its resolution (replacing the Backshift operator B by the simple value x) gives $1 - \theta x = 0 \Rightarrow x = \dfrac{1}{\theta}$.
The condition that the solution is greater than 1 is equivalent to $|\dfrac{1}{\theta}| > 1 \Rightarrow |\theta| < 1$

1.3.4 Moments of an AR(p) process

An AR (p) process is written: $X_t = c + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \ldots + \theta_p X_{t-p} + \epsilon_t$.
We will suppose that $\varepsilon_i \sim iid(0, \sigma^2)$.
The different moments of a stationary process are:
$\mathbb{E}[X_t] = \dfrac{c}{1 - \theta_1 - \theta_2 - \ldots - \theta_p}$.
$Var(X_t) = \theta_1\gamma_1 + \theta_2\gamma_2 + \ldots + \theta_p\gamma_p + \sigma^2$.
$Cov(X_t, X_{t-j}) = \theta_1\gamma_{j-1} + \theta_2\gamma_{j-2} + \ldots + \theta_p\gamma_{j-p}$.
The formulas of variance and covariance correspond to the so-called Yule and Walker equations.

1.3.5 Stationarity condition for AR(p)

Theorem 1. *the process AR (p) is stationary if the modulus of solutions (the roots) of its characteristic equation is every time strictly greater than 1 in absolute value.*

The condition is often formulated differently, according to which the roots must be outside of the complex unitary circle.

Example 3. $\{AR(2)\}$
The characteristic polynomial of the backshift of a process $AR(2)$

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \epsilon_t,$$

s'écrit:

$$(1 - \theta_1 B - \theta_2 B^2)X_t = \epsilon_t.$$

The resolution of the characteristic equation of the second degree $(1 - \theta_1 x - \theta_2 x^2)$ leads to the following conditions:

$$\begin{cases} \theta_1 + \theta_2 < 1, \\ \theta_2 - \theta_1 < 1, \\ |\theta_2| < 1 \end{cases}$$

1.3.6 Estimation for AR(1) process

We will focus on the estimation of the autoregressive parameter of the first order with Gaussian innovations. We consider the model

$$X_t = \theta X_{t-1} + \epsilon_t, \qquad t = 1, 2, ...$$

We assume that $X_0 = 0$ is a Gaussian variable.
Least Squares method:
To estimate the coefficient (unknown parameter) of a 1^{st} order autoregressive process, we use the least squares estimator. We will now minimize the quantity

$$Q(\theta) = \sum_{t=1}^{n} \epsilon_t^2 = (X_t - \theta X_{t-1})^2.$$

We derive the function $Q(\theta)$ with respect to θ:

$$\frac{\partial Q(\theta)}{\partial \theta} = 2\sum_{t=1}^{n}(-X_{t-1})(X_t - \theta X_{t-1}).$$

We find

$$2\sum_{t=1}^{n}(-X_{t-1})(X_t - \theta X_{t-1}) = 0,$$

$$\sum_{t=1}^{n}(-X_t X_{t-1} + \theta X_{t-1}^2) = 0,$$

$$-\sum_{t=1}^{n} X_t X_{t-1} + \theta \sum_{t=1}^{n} X_{t-1}^2 = 0,$$

$$\sum_{t=1}^{n} X_t X_{t-1} = \theta \sum_{t=1}^{n} X_{t-1}^2,$$

The estimator θ_n is given, for all $n \geq 1$, by

$$\theta_n = \frac{\sum_{t=1}^{n} X_{t-1} X_t}{\sum_{t=1}^{n} X_{t-1}^2}. \tag{1.6}$$

1.3.7 Estimation for AR(p) process

Yule-Walker method:
We have the autoregressive model of order p

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \ldots + \theta_p X_{t-p} + \epsilon_t.$$

Multiply both sides of the equation by X_{t-j} $(j > 0)$ and take the expectatione

$$\mathbb{E}[X_t X_{t-j}] = \theta_1 \mathbb{E}[X_{t-1}X_{t-j}] + \theta_2 \mathbb{E}[X_{t-2}X_{t-j}] + \ldots + \theta_p \mathbb{E}[X_{t-p}X_{t-j}] + \mathbb{E}[\epsilon_t X_{t-j}]. \quad (1.7)$$

White noise ϵ_t not correlated with X_{t-j}, where j is greater than zero.
For $j > 0, \quad \mathbb{E}[\epsilon_t X_{t-j}] = 0,$

For $j = 0,$

$$\mathbb{E}[\epsilon_t X_t] = \theta_1 \mathbb{E}[X_{t-1}\epsilon_t] + \theta_2 \mathbb{E}[X_{t-2}\epsilon_t] + \ldots + \theta_p \mathbb{E}[X_{t-p}\epsilon_t] + \mathbb{E}[\epsilon_t^2] = \sigma^2.$$

$$\begin{aligned} (1.7) \quad &\Rightarrow \quad \gamma_j = \theta_1 \gamma_{j-1} + \theta_2 \gamma_{j-2} + \ldots + \theta_p \gamma_{j-p} \\ &\Rightarrow \quad \rho_j = \theta_1 \rho_{j-1} + \theta_2 \rho_{j-2} + \ldots + \theta_p \rho_{j-p} \end{aligned}$$

$$\left.\begin{aligned} \rho_1 &= \theta_1 \rho_0 + \theta_2 \rho_1 + \ldots + \theta_p \rho_{p-1} \\ \rho_2 &= \theta_1 \rho_1 + \theta_0 \rho_1 + \ldots + \theta_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \theta_1 \rho_{p-1} + \theta_{p-2} \rho_1 + \ldots + \theta_p \rho_0. \end{aligned}\right\} \text{Yule walker equations.}$$

If ρ_j is known, the resolution of these p equations makes it possible to obtain the estimates of θ.

Variance estimation

For $j = 0, \mathbb{E}[X_t X_{t-j}] = \sigma^2,$

$\gamma_0 = \theta_1 \gamma_1 + \theta_2 \gamma_2 + \ldots + \theta_p \gamma_p + \sigma^2,$
$\Rightarrow \sigma^2 = \gamma_0 - \theta_1 \gamma_1 - \theta_2 \gamma_2 - \ldots - \theta_p \gamma_p.$
which gave the estimate of θ gives the estimate of σ^2

- However, in general, the ρ_j would be unknown and should be estimated for the sample.

- We have $\hat{\rho}_i = \hat{\gamma}_j / \hat{\gamma}_0,$

where

$$\hat{\gamma}_j = \frac{1}{n} \sum_{t=1}^{n-j} (X_t - \bar{X})(X_{t+j} - \bar{X}).$$

$$\begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \vdots \\ \hat{\rho}_{p-1} \\ \hat{\rho}_p \end{pmatrix} = \begin{pmatrix} \hat{\rho}_0 & \hat{\rho}_1 \cdots & \hat{\rho}_{p-1} \\ \hat{\rho}_1 & \hat{\rho}_0 \cdots & \hat{\rho}_{p-2} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \hat{\rho}_{p-1} & \hat{\rho}_{p-2} \cdots & \hat{\rho}_0 \end{pmatrix} \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_{p-1} \\ \hat{\theta}_p \end{pmatrix}$$

$$\hat{\rho}_p = \hat{\Gamma}_p \hat{\theta} \Rightarrow \hat{\theta} = \hat{\Gamma}_p^{-1} \hat{\rho}_p.$$

1.3.8 AR (p) processes: examples

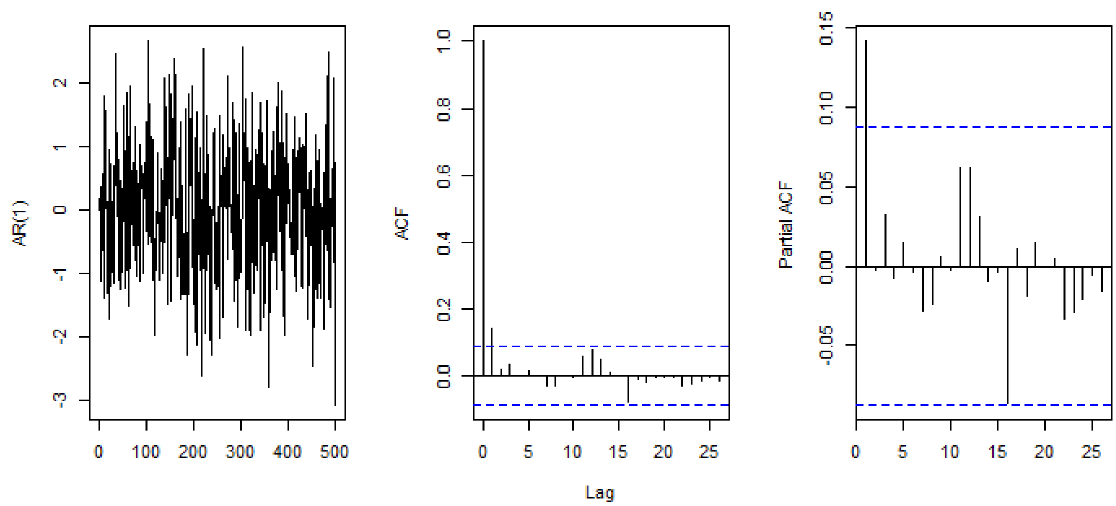

Figure 1.1: Path graph, correlogram and partial correlogram of the AR (1) process: $X_t = 0.1X_{t-1} + \epsilon_t$.

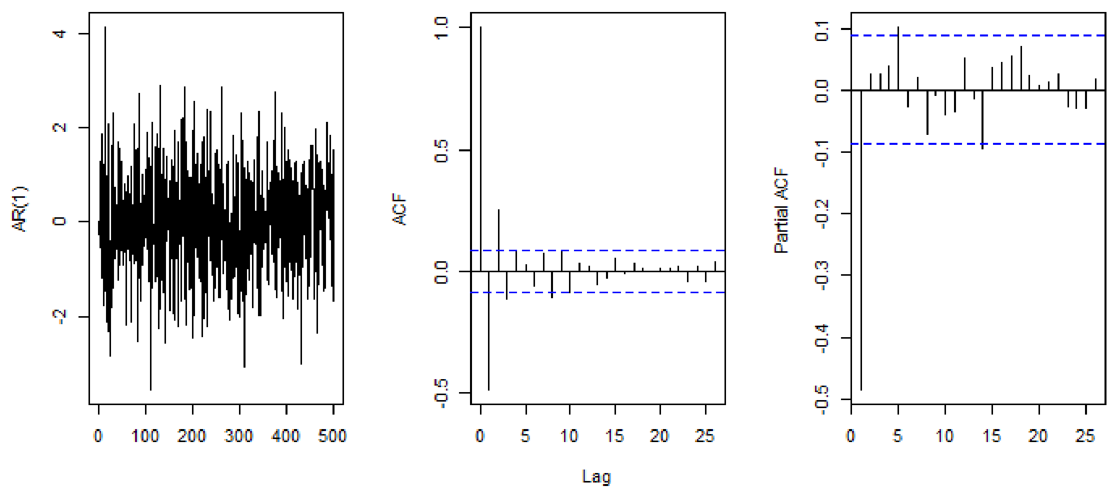

Figure 1.2: Path graph, correlogram and partial correlogram of the AR (1) process: $X_t = -0.5X_{t-1} + \epsilon_t$.

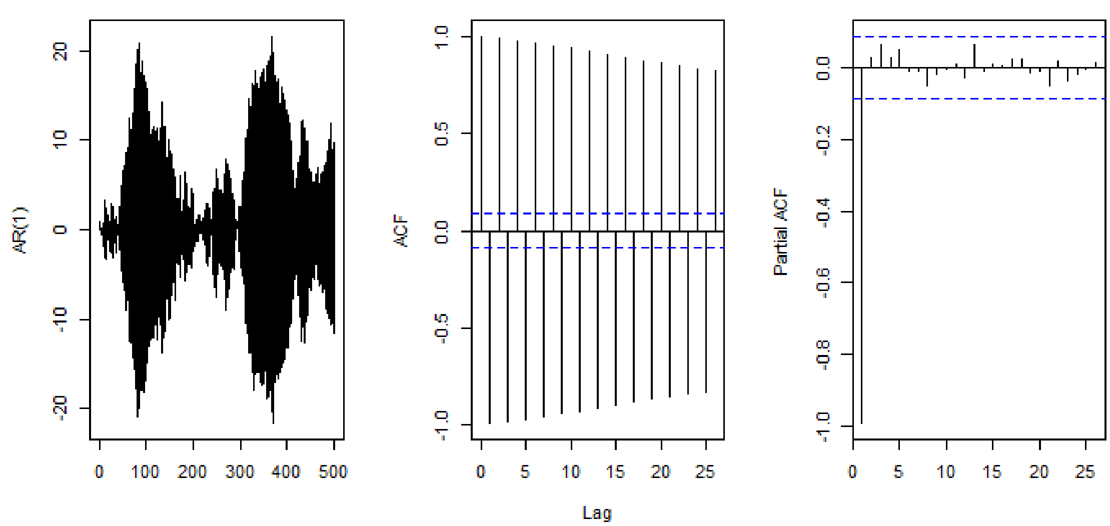

Figure 1.3: Path graph, correlogram and partial correlogram of the AR (1) process: $X_t = -1.001X_{t-1} + \epsilon_t$.

1.4 MA process

Definition 9. *An average moving process of order q, denoted MA (q) is defined by*

$$X_t = \delta_1 \epsilon_{t-1} + \delta_2 \epsilon_{t-2} + ... + \delta_q \epsilon_{t-q} + \epsilon_t, \tag{1.8}$$

where $\theta_1, \ldots, \theta_p$ are the parameters of the model and ϵ_t is a white noise of variance σ^2. as for autoregressive processes, this relationship is written:

$$\begin{aligned} (1 - \delta_1 B - ... - \delta_p B^p)\epsilon_t &= X_t, \\ \Theta(B)\epsilon_t &= X_t. \end{aligned}$$

Unlike the AR (p) the definition of MA (q) is explicit that the process X_t is stationary.

Example 4. *An MA(1) process takes this form:*

$$X_t = \delta_1 \epsilon_{t-1} + \epsilon_t, \ \ \textit{or} \ \ \epsilon_t \backsim WN(0, \sigma^2).$$

1.4.1 Causality and Invertibility

Definition 10. *It is said that the process is causal if there is a real c_k sequence such that $\sum_{k=0}^{\infty} |c_k| < \infty$ and*

$$X_t = \sum_{k=0}^{\infty} c_k \epsilon_{t-k}. \tag{1.9}$$

Sometimes, when we talk about a causal process, we say that it has a representation MA(∞).

Remark 3. *By this definition, any process MA (q) is causal.*

Proposition 3. *The moving average process MA (q) is invertible if and only if its polynomial $\Theta(z)$ is such that*

$$\Theta(z) \neq 0 \ \textit{avec} \ z \in \mathbb{C} \ \textit{tel que} \ |z| \leq 1. \tag{1.10}$$

The proof of this proposition is in Brockwell and Davis [11].

1.4.2 Estimation for MA process

The method of moments is not linear or effective for MA(q) processes.
Simple example: process estimation MA(1):
$\gamma_0 = (1+\delta^2)\sigma^2$, $\gamma_1 = \delta\sigma^2$,
$\Rightarrow \rho_1 = \dfrac{\gamma_1}{\gamma_0} = \dfrac{\delta}{1+\delta^2}$.
Moments method :
$\gamma_1 = \hat{\gamma}_1$, $\rho_1 = \hat{\rho}_1$, $\dfrac{\hat{\delta}}{1+\hat{\delta}^2} = \hat{\rho}_1$,
$\Rightarrow \hat{\rho}_1(1+\hat{\delta}^2) = \hat{\delta}$.

He can have no solution (if $|\hat{\rho}_1| > 1/2$) or two solutions (if $|\hat{\rho}_1| < 1/2$) from which we can choose which gives an invertible process.
Moreover, we can show that it is not effective (there are other simple estimators with asymptotically lower variance)

1.4.3 MA (q) processes: examples

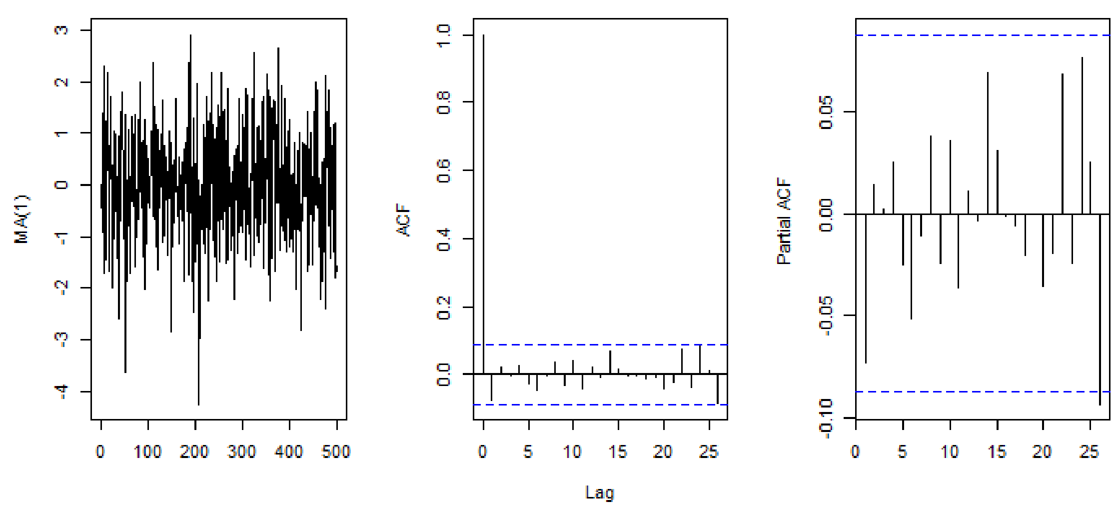

Figure 1.4: Path graph, correlogram and partial correlogram of the process MA(1): $X_t = -0.1\epsilon_{t-1} + \epsilon_t$.

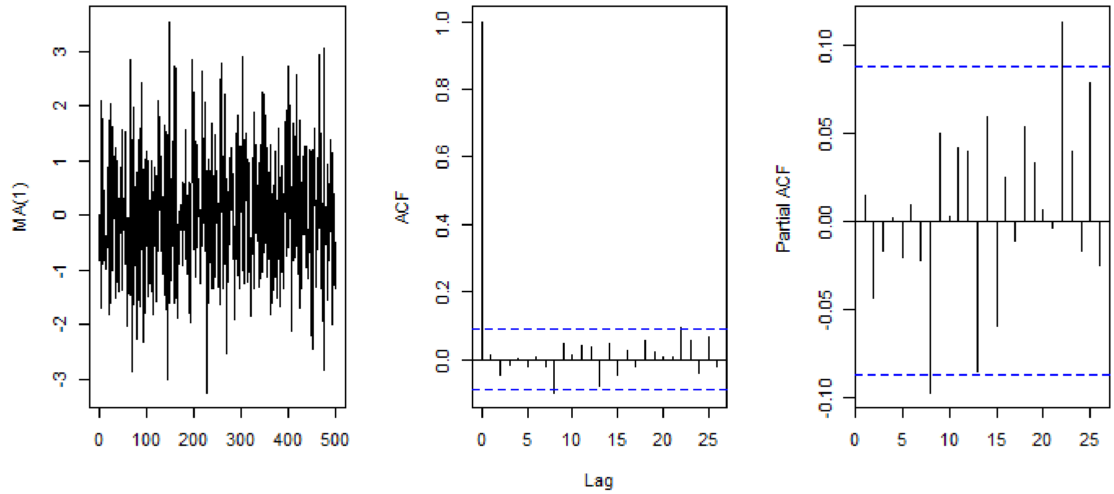

Figure 1.5: Path graph, correlogram and partial correlogram of the process MA(1): $X_t = 0.5\epsilon_{t-1} + \epsilon_t$.

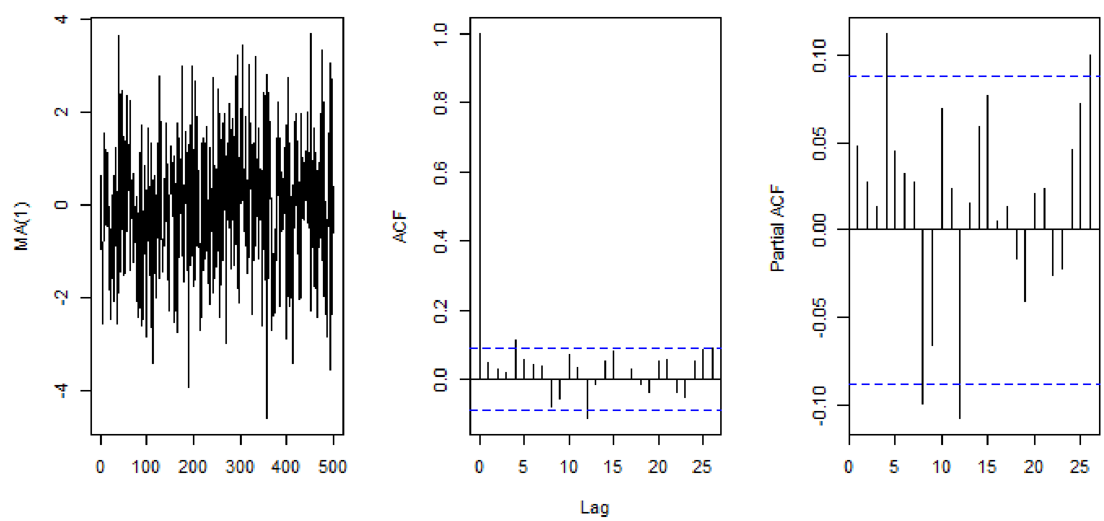

Figure 1.6: Path graph, correlogram and partial correlogram of the process MA(1): $X_t = -1.001\epsilon_{t-1} + \epsilon_t$.

1.5 The ARMA model(Mixed model)

AR and MA models can be ideal in some cases, but there may be a need to estimate a large number of parameters to adjust the model. If there are few observations, these estimates will tend to be unclear. In addition, if a model containing p parameters is suitable for the situation, it is not good to try to adjust a model that will contain more than p. The ARMA models (or Box-Jenkins model), are a mix of Autoregressive and Moving average models offered by Yule and Slutsky. They play an important role in specifying time series models for applications. As the solutions of stochastic difference equations with constant coefficients and those processes possess a linear structure. Herman Wold [42] showed that ARMA processes can be used to model any stationary series as long as the orders p and q are well chosen. Box and Jenkins [12] worked to develop a methodology for model estimation of a time series.

Definition 11. *An ARMA process of order (p, q) (discrete or continuous) is defined by*

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + ... + \theta_q X_{t-q} + \epsilon_t + \delta_1 \epsilon_{t-1} + \delta_2 \epsilon_{t-2} + ... + \delta_q \epsilon_{t-q}, \tag{1.11}$$

where $\theta_1, \theta_2, ..., \theta_p$ are the parameter of the model and the ϵ_t the error terms.

An autoregressive model AR (p) is a $ARMA(p, 0)$.
A moving average model MA(q) is a $ARMA(0, q)$.

1.5.1 Stationarity, Causality and Invertibility

Theorem 2. *If θ and δ don't have common factors, a (unique) stationary solution to $\alpha(B)X_t = \Theta(B)\epsilon_t$ exists if the roots of $\alpha(z)$ avoid unit circle:*

$$|z| = 1 \Rightarrow \alpha(z) = 1 - \theta_1 z - ... - \theta_p z^p \neq 0.$$

ARMA(p,q) process is causal if the roots of are outside unit circle:

$$|z| \leq 1 \Rightarrow \alpha(z) = 1 - \theta_1 z - ... - \theta_p z^p \neq 0.$$

is invertible if the root of $\Theta(z)$ are out side unit circle:

$$|z| \leq 1 \Rightarrow \Theta(z) = 1 + \theta_1 z + ... + \theta_q z^q \neq 0.$$

Definition 12. *(Causality) The ARMA(p,q) process is causal if there's a real c_k sequence such that $\sum_{k=0}^{\infty} |c_k| < \infty$ and*

$$X_t = \sum_{k=0}^{\infty} c_k \epsilon_{t-k}. \tag{1.12}$$

Causality means that an ARMA time series can be represented as a linear process.

Theorem 3. *Let $\{X_j, j \geq 1\}$ be an ARMA process such that the polynomials $\alpha(z)$ and $\Theta(z)$ don't have common zeroes. Then $\{X_j, j \geq 1\}$ is causal if $\Theta(z) \neq 0$ with $|z| \leq 1$. The coefficients c_k are determined by the power series expansion*

$$c(z) = \sum_{k=0}^{\infty} c_k z^k = \frac{\alpha(z)}{\Theta(z)}, \ \ |z| \leq 1. \tag{1.13}$$

An idea closely related to causality is invertibility.

Definition 13. *(Invertibility) The ARMA(p,q) process is invertible if there's a real sequence d_k as $\sum_{k=0}^{\infty} |d_k| < \infty$ et*

$$\epsilon_k = \sum_{k=0}^{\infty} d_k X_{t-k}. \tag{1.14}$$

Theorem 4. *Let $\{X_j, j \geq 1\}$ be an ARMA process such that the polynomials $\alpha(z)$ and $\Theta(z)$ don't have common zeros. Then $\{X_j, j \geq 1\}$ is invertible if $\Theta(z) \neq 0$ with $|z| \leq 1$. The coefficients d_k are determined by the power series expansion*

$$d(z) = \sum_{k=0}^{\infty} d_k z^k = \frac{\alpha(z)}{\Theta(z)}, \ \ |z| \leq 1. \tag{1.15}$$

Henceforth it's far assumed that each one ARMA sequences specified in the sequel are causal and invertible unless expressly provided otherwise.

1.5.2 Autocovariance of an ARMA(p,q)

Let the following ARMA process (p,q) be:

$$X_t - \theta_1 X_{t-1} ... - \theta_p X_{t-p} = \epsilon_t + \delta_1 \epsilon_t + ... + \delta_q \epsilon_t,$$

We obtain, by multiplying by X_{t-h} and taking the expectation:

$$\gamma_h - \sum_{j=1}^{p} \theta_j \gamma_{h-j} = \mathbb{E}[X_{t-h}\epsilon_t] - \sum_{j=1}^{q} \delta_j \mathbb{E}[X_{t-h}\epsilon_{t-j}],$$

Or $\mathbb{E}[X_{t-h}\epsilon_{t-j}] = 0$ if $t - h < t - i$ for $j = 1, 2, ..., q$.
Thus, for $h > q$, we find the p-order induction equation as in the case of the AR (p):

$$\gamma_h - \sum_{j=1}^{p} \theta_j \gamma_{h-j} = 0 \ \ h > 0$$

The p equations $h = q+1, ..., q+p$ are called Yule-Walker equations. They allow calculate the p coefficients γ_h as a function of p initial values $\gamma_q, ..., \gamma_{q-p+1}$. These equations do not make it possible to determine all the values γ_h since must have initial conditions.
The first values of γ_h, $h = 0, 1, ..., q$ are determined by:

$$\gamma_h - \sum_{j=1}^{p} \theta_j \gamma_{h-j} = \mathbb{E}[X_{t-h}\epsilon_t] - \sum_{j=1}^{q} \mathbb{E}[X_{t-h}\epsilon_{t-j}].$$

for $h = 0, 1, ..., q$. The terms on the right of equality are calculated from the expression MA(∞) of X_t.

1.5.3 Parameter Estimation

We will speak about more efficient estimators are provided by the maximum likelihood and least squares methods for ARMA(p,q) assuming that the orders p and q are known.

Method 1(Maximum Likelihood Estimation):
The method of Maximum Likelihood Estimation applies to any ARMA(p,q) model

$$X_t - \theta_1 X_{t-1} ... - \theta_p X_{t-p} = \epsilon_t + \delta_1 \epsilon_t + ... + \delta_q \epsilon_t.$$

The innovations algorithm applied to a causal ARMA(p,q) process $(X_t, t \geq 1)$ gives

$$\hat{X}_{i+1} = \sum_{j=1}^{i} \delta_{ij}(X_{i+1-j} - \hat{X}_{i+1-j}), \qquad 1 \leq i < \max(p,q)$$

$$\hat{X}_{i+1} = \sum_{j=1}^{p} \theta_j X_{i+1-j} + \sum_{j=1}^{q} \delta_{ij}(X_{i+1-j} - \hat{X}_{i+1-j}), \qquad i \geq \max(p,q)$$

with prediction error

$$K_{i+1} = \sigma^2 R_{i+1}.$$

where $R_p = \gamma_p / \gamma_0$
in the closing expression, σ^2 has been factored out due to reasons that becomes apparent from the form of the likelihood function to be mentioned beneath. recall that the series $(X_{i+1} - \hat{X}_{i+1} : i \geq 1)$ consists of uncorrelated random variables if the parameters are known. Assuming normality for the errors, we moreover achieve even independence. this could be exploited to define the Gaussian maximum likelihood estimation(MLE) procedure. Throughout, it is assumed that $(X_t, t \geq 1)$ has 0 mean ($\mu = 0$). The parameters of interest are collected inside the vectors $\beta = (\theta, \delta, \sigma^2)^t$ and $\beta' = (\theta, \delta)^t$, where $\theta = (\theta_1, ..., \theta_p)^t$ and $\delta = (\delta_1, ..., \delta_q)^t$. suppose finally that we've got observed the variables $X_1, ..., X_n$. Then, the Gaussian likelihood function for the innovations is

$$L(\beta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \left(\prod_{i=1}^{n} R_i^{1/2} \right) \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^{n} \frac{(X_i - \hat{X}_j)^2}{R_j} \right).$$

Taking the partial derivative of $\ln L(\beta)$ with respect to the variable σ^2 reveals that the Maximum Likelihood Estimation for σ^2 can be calculated from

$$\hat{\sigma}^2 = \frac{S(\hat{\theta}, \hat{\delta})}{n}, \qquad S(\hat{\theta}, \hat{\delta}) = \sum_{j=1}^{n} \frac{(X_j - \hat{X}_j)^2}{R_j}.$$

$\hat{\theta}$ and $\hat{\delta}$ are estimators of θ et δ by method of Maximum Likelihood Estimation obtained from minimizing the profile likelihood or reduced likelihood

$$l(\theta, \delta) = \ln\left(\frac{S(\theta, \delta)}{n}\right) + \frac{1}{n}\sum_{j=1}^{n} \ln(R_j).$$

Observe that the profile likelihood $l(\theta, \delta)$ can be computed using the innovations algorithm. the velocity of these computations depends upon heavily on the quality of initial estimates. these are frequently supplied by means of the non-optimal desirable Yule-Walker technique. For numerical methods, such as the Newton-Raphson and scoring algorithms, see phase 3.6 in Shumway and Stoffer (2006).

Proposition 4. *Let $\{X_j, j \geq 1\}$ be a causal and invertible ARMA(p,q) process defined with an i.i.d sequence $(\epsilon_t, t \geq 1)$ satisfying $\mathbb{E}[\epsilon_t] = 0$ and $\mathbb{E}[\epsilon_t^2] = 0$ Consider the Maximum Likelihood Estimation $\hat{\beta}'$ of β' that is initialized with the moment estimators of Method, Then*

$$\sqrt{n}(\hat{\beta}' - \beta') \xrightarrow{D} N(0, \sigma^2\Gamma_{p,q}^{-1}) \ \ (n \to \infty)$$

where the $(p+q) \times (p+q)$ dimensional matrix Γ_{p+q} depends on the model parameters.

A proof of this result is given in Section 8.10 of Brockwell and Davis [11](1991).
Method 2 (Least Squares Estimation): *The method of least squares for causal and invertible ARMA(p,q) processes, it is primarily based on minimizing the weighted sum of squares*

$$S(\theta, \delta) = \sum_{j=1}^{n} \frac{(X_j - \hat{X}_j)^2}{R_j}.$$

with respect to θ and δ , respectively. Assuming that $\tilde{\theta}$ and $\tilde{\delta}$ denote these Least Squares Estimates, the Least Squares Estimation for σ^2 is computed as

$$\tilde{\sigma}^2 = \frac{S(\tilde{\theta}, \tilde{\delta})}{n - p - q}.$$

The least squares procedure has the same asymptotics as the Maximum Likelihood Estimation.

Theorem 5. *The result of previous Theorem holds also if $\hat{\beta}'$ is replaced with $\bar{\beta}'$.*

1.5.4 ARMA(p,q) processes: examples

Figure1.7 shows realizations of three different autoregressive moving average time series based on independent, standard normally distributed ϵ_t. The left panel is an ARMA$(2, 2)$ process with parameter specifications $\theta_1 = 0.9$, $\theta_2 = -0.8$, $\delta_1 = -0.5$ and $\delta_1 = 0.8$. The center plot is obtained from an ARMA$(2, 1)$ process with parameters $\theta_1 = 0.9$, $\theta_2 = -0.8$, $\delta = 0.6$, while the right plot is from an ARMA$(1, 2)$ with parameters $\theta_1 = 0.6$, $\delta_1 = 0.9$ and $\delta_2 = -0.8$. as we can see these processes are still stationary.

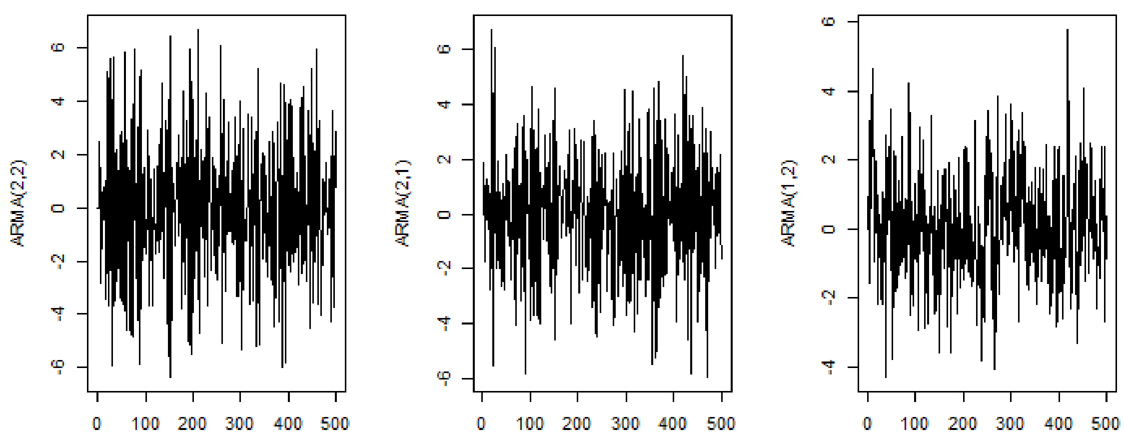

Figure 1.7: Realizations of three autoregressive moving average processes.

1.6 The ARIMA model

ARIMA models are, in theory, the foremost general category of models for predicting a time series which can be made to be stationary by differencing (if necessary), maybe in conjunction with nonlinear transformations like logging or deflating (if necessary). A random variable that's a time series is stationary if its statistical properties are all constant over time. A stationary series has no trend, its variations around its mean have a constant amplitude, and it wiggles in a very consistent fashion, i.e., its short random time patterns always look the same in a statistical sense. The latter condition implies that its autocorrelations (correlations with its own previous deviations from the mean)stay constant over time, or equivalently, that its power spectrum remains constant over time. A random variable of this type will be viewed (as usual) as a mixture of signal and noise, and therefore the signal (if one is apparent) might be a pattern of quick or slow mean reversion, or curved oscillation, or speedy alternation in sign, and it might even have a seasonal element.
An ARIMA model will be viewed as a filter that tries to separate the signal from the noise, and the signal is then extrapolated into the future to obtain forecasts.

The ARIMA forecasting equation for a stationary time series is a linear (i.e., regression-type) equation in which the predictors consist lags of the dependent variable quantity or lags of the forecast errors. That is:
The predicted value of $X = a$ constant or a weighted sum of 1 or more values of X or a weighted sum of 1 or more recent values of the errors.
If the predictors consist solely of lagged values of X, it's a pure autoregressive (self-regressed) model, that is simply a special case of a regression model and which could be fitted with standard regression software. as an example, a 1^{st} order autoregressive model for X is a simple regression model in which the independent variable is just X lagged by one period. If some of the predictors are lags of the errors, an ARIMA model it is not a linear regression model, as a result of there's no way to specify (last period's error) as an independent variable: the errors should be computed on a period to period basis when the model is fitted to the data. From a technical viewpoint, the problem with exploiting lagged errors as predictors is that the model's predictions don't seem to be linear functions of the coefficients, while they're linear functions of the past data. So, coefficients in ARIMA models that include lagged errors should be estimated by nonlinear optimisation methods (hill-climbing) instead of by just solving a system of equations.
The ARIMA abbreviation stands for Autoregressive Integrated Moving Average. Lags of the stationarized series in the forecasting equation are called autoregressive terms, lags of the forecast errors are called moving average terms, and a time series which needs to be differenced to be made stationary is said to be an integrated version of a stationary series. Random-walk and random-trend models, autoregressive models, and exponential smoothing models are all special cases of ARIMA models.

1.7 ARCH/GARCH Models

In order to overcome the inadequacies of the ARMA(p,q) representations for monetary and financial problems, Engle (1982)[17] propose a new class of models autoregressive conditional heteroskedasticity model (ARCH) able to capture the behavior of volatility over time. ARCH model are commonly used in modeling financial time series, which have variable volatilities, that is, restless periods followed by periods of relative calm. In these models, the conditional variance at time t is variable. It depends, for example, on the square of previous achievements of the process or the square of innovations.
GARCH is probably the most commonly utilized financial time series model and has inspired dozens of more sophisticated models.
ARCH models are based totally on an endogenous parameterization of the conditional variance. The family of ARCH models may be broken down into two subsets: linear ARCH models and nonlinear ARCH models.
The first is based totally on a quadratic specification of the conditional variance of perturbations: models ARCH(q), GARCH (p,q) and IGARCH (p,q). Nonlinear ARCH models are characterized by asymmetrical disturbance specifications. These are the models EGARCH (p, q), TARCH (q) and TGARCH (p, q). (Bresson and Pirotte, Time series)

1.7.1 Linear ARCH/GARCH Models

We will present the definition of models ARCH and GARCH. First, in order to be able to process series of non-zero average, the situation of certain yields, let us imagine that we observe a series Y_i that behaves like a yield, denoted X_i, at a constant.

$$Y_i = c + X_i,$$

where $\mathbb{E}[X_i] = 0$. Let us note σ_i^2 the variance of the conditional yield with the past.

$$\sigma_i^2 = Var(X_t/\widetilde{X}_{i-1}),$$

where $\widetilde{X}_{i-1}$ designates the past $i-1, i-2, ...$

Our observations lead us precisely to assume that σ_i^2 is a function of $X_{t-1}^2, X_{t-2}^2, ...$

Definition 14. *The process X_i satisfies an ARCH(1) representation if*

$$X_i = v_i h_i, \tag{1.16}$$

with

$$h_i = \sqrt{\alpha_0 + \alpha_1 X_{i-1}^2}$$

where v_i denotes a weak white noise such that $\mathbb{E}[v_i] = 0$, $\mathbb{E}[v_i^2] = \sigma_v^2$.

In general, v_i denotes a set of independent random variables, identically distributed, centered, reduced. The component h_i denotes a variable which, conditionally to the set of information of the past values of X_i, ie to $\widetilde{X}_{i-1} = \{X_{i-1}, X_{i-2}, ..., X_{i-j}, ...\}$, is deterministic and positive.
In this system, the X_i process is characterized by zero autocorrelations $\mathbb{E}[v_t v_s] = 0$ and a conditional variance that varies over time depending on the magnitude of past innovation.

we are able to set up interesting results by considering the autoregressive process on X_i^2. For simplicity, we are limited to the case of ARCH(1). In these conditions:

$$h_i^2 = \alpha_0 + \alpha_1 X_{i-1}^2 \Leftrightarrow X_i^2 = \alpha_0 + \alpha_1 X_{i-1}^2 + (X_i^2 - h_i^2), \tag{1.17}$$

is still

$$X_i^2 = \alpha_0 + \alpha_1 X_{i-1}^2 + u_i, \tag{1.18}$$

where $u_i = (X_i^2 - h_i^2)$ *verifying* $\mathbb{E}(u_i/\widetilde{X}_{i-1}) = 0$ *is an innovation process for* X_i^2.

Definition 15. *The process* X_i *satisfies an ARCH(q) representation if*

$$X_i = v_i h_i, \tag{1.19}$$

with $h_i = \sqrt{\alpha_0 + \sum_{k=1}^{q} \alpha_k X_{i-k}^2}$ *and where* v_i *denotes a weak white noise such that* $\mathbb{E}[v_i] = 0$, $\mathbb{E}[v_i^2] = \sigma_v^2$.

Definition 16. *The process* X_i *satisfies an GARCH(p,q) representation if*

$$X_i = v_i h_i, \tag{1.20}$$

with $h_i = \sqrt{\alpha_0 + \sum_{k=1}^{q} \alpha_k X_{i-k}^2 + \sum_{k=1}^{P} \beta_k h_{i-k}^2}$ *and where* v_i *denotes a weak white noise and where* $\alpha_0 > 0$, $\alpha_k \geq 0$, $k = 1, ..., q$ *and* $\beta_i \geq 0$, $k = 1, ..., p$.

Example 5. *Consider the case of a process GARCH(1,1):*

$$X_i = v_i h_i, \tag{1.21}$$

$$h_i = \sqrt{\alpha_0 + \alpha_1 X_{i-1}^2 + \beta_1 h_{i-1}^2}, \tag{1.22}$$

where $\alpha_0 > 0$, $\alpha_1 > 0$ and $\beta_1 > 0$.
In this model, the squares of the residues follow a ARMA(1,1) process,

$$X_i^2 = \alpha_0 + (\alpha_1 + \beta_1)X_{i-1}^2 + \beta_1 u_{i-1} + u_i,$$

It is stationary for $0 < \alpha_1 + \beta_1 < 1$, where $u_i = v_i^2 - h_i^2$ is a process of innovation for v_i^2. Under the condition of second order stationarity, the unconditional variance of process v_i exists and is constant over time. knowing that $Var(v_i) = \mathbb{E}[v_i^2]$, it's sufficient that from the form ARMA$(1,1)$ on v_i^2 define the variance of the process $Var(v_t) = \dfrac{\alpha_0}{1 - \alpha_1 - \beta_1}$.

According to BOLLERSLEV T. (1986) [8], kurtosis (Kurtosis measures the sharp or flat character of the cast of the series) exists if

$$3\alpha_1^2 + 2\alpha_1\beta_1 < 1,$$

and is given by

$$\begin{aligned} K_u &= \frac{\mathbb{E}[X_i^4]}{\mathbb{E}^2[X_i^2]} \\ &= 3\frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2}. \end{aligned}$$

She is always more than three. Consequently, if α_1 tends to zero, the energetic elasticity disappears and the value of kurtosis tends to three. Finally, it can be shown that for a GARCH process kurtosis is directly related to the conditional heteroskedasticity event.
Consider the case of kurtosis associated with unconditional law in a heteroskedasticity Gaussian GARCH process such as $v_i \sim N(0,1)$.

In this case, the conditional moments of order 2 and 4 of the process are linked

$$\mathbb{E}[X_i^4/\widetilde{X}_{t-1}] = 3(\mathbb{E}[X_i^2/\widetilde{X}_{i-1}])^2.$$

Indeed, we recall that if a centered variable z follows a Gaussian law,
$\mathbb{E}[z^4] = 3(Var(z))^2 = 3(\mathbb{E}[z^2])^2$.
If we apply the expectation on both sides of the previous equation, it becomes

$$\mathbb{E}[X_i^4] = \mathbb{E}(\mathbb{E}[X_i^4/\widetilde{X}_{i-1}]) = 3\mathbb{E}\left([\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]^2\right),$$

$$3\mathbb{E}\left([\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]^2\right) \geq 3\mathbb{E}\left([\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]\right)^2 = 3\left(\mathbb{E}[X_i^2]\right)^2.$$

The kurtosis can be calculated as follows:

$$\begin{aligned} K_u &= \frac{\mathbb{E}[X_i^4]}{\mathbb{E}^2[X_i^2]} \\ &= \frac{3\mathbb{E}\left([\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]^2\right)}{\mathbb{E}^2[X_i^2]} \\ &= 3\frac{\mathbb{E}^2[X_i^2]}{\mathbb{E}^2[X_i^2]} + \frac{3}{\mathbb{E}^2[X_i^2]}(\mathbb{E}[[\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]^2] - \mathbb{E}^2[X_i^2]) \\ &= 3 + \frac{3}{\mathbb{E}^2[X_i^2]}(\mathbb{E}[[\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]^2] - \mathbb{E}^2[\mathbb{E}[X_i^2/\widetilde{X}_{i-1}]]) \\ &= 3 + 3\frac{Var(\mathbb{E}[X_i^2/\widetilde{X}_{i-1}])}{\mathbb{E}[X_i^2]^2} > 3. \end{aligned}$$

Kurtosis is therefore linked to a measure of conditional heteroscedasticity.

1.8 Simulate processes

In this section, we explain how to simulate in R processes such as white noise, autoregressive, moving average, ARIMA and ARCH/GARCH. We will also describe the arima.sim function and garchSim function those generates respectively ARIMA processes and ARCH/GARCH processes.

1.8.1 Simulate an autoregressive process

To simulate simply T realizations of an autoregressive process of order p having the parameters $\theta_1, ..., \theta_p$ and whose variance of innovations is σ_ϵ^2, we can use the following algorithm:

1. Set p arbitrary real initial values $x_0, ..., x_{p-1}$.
2. Generate i.i.d. innovations $(\epsilon_t)_{t=1,...,T+T_0}$.
3. Calculate the values recursively $X_t = \sum_{j=1}^{p} \theta_j X_{t-j} + \epsilon_t$
4. Eliminate the first T_0 values thus generated.

Now, we will write the following command lines:
$n = 500$
$x = rep(0, n)$
$for(iin3 : n)x[i] = 0.1 * x[i-1] - 0.7 * x[i-2] + rnorm(1)$
$op = par(mfrow = c(3, 1), mar = c(2, 4, 2, 2)+0.1)plot(ts(x), xlab = "", ylab = "AR(2)")$
$acf(x, main = "", xlab = "")pacf(x, main = "", xlab = "")par(op)$

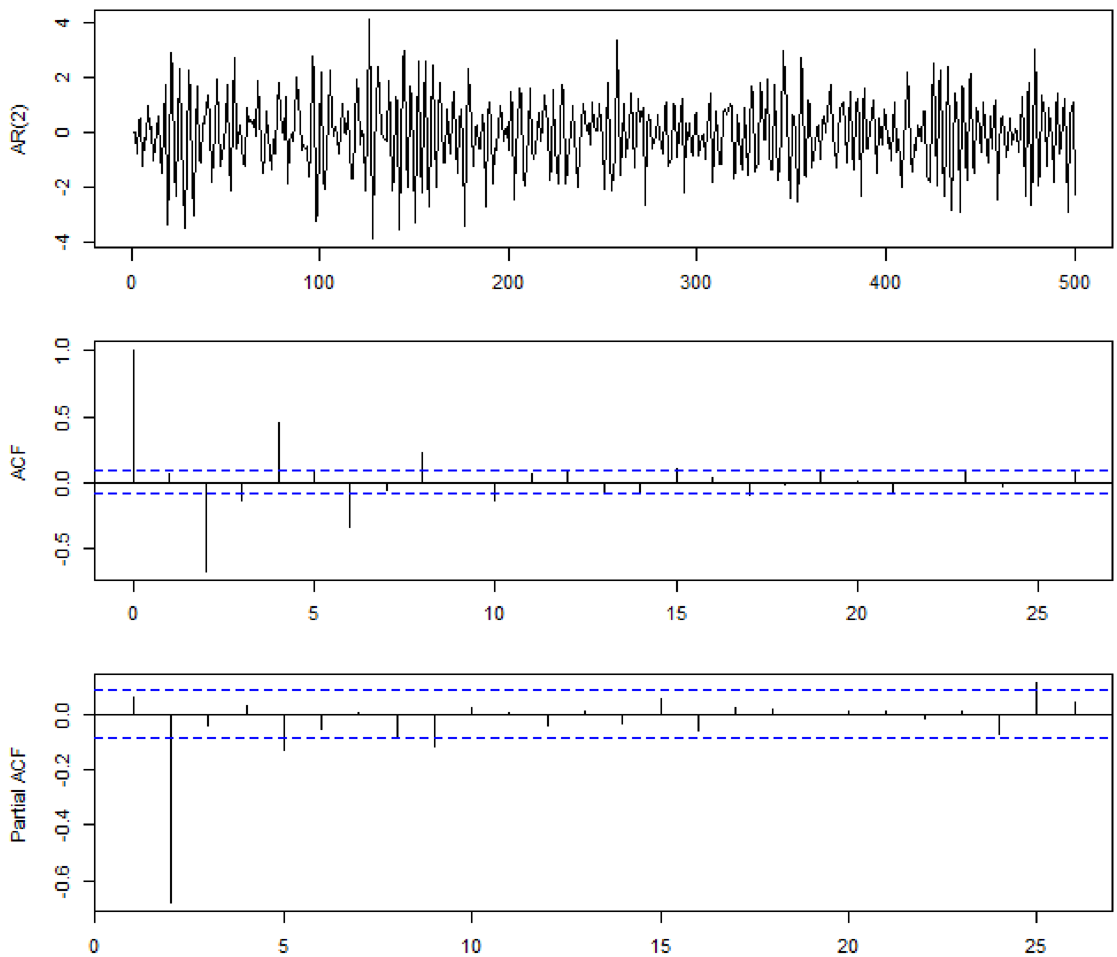

Figure 1.8: Simulation of AR(2) process.

1.8.2 Simulate a MA process

To simulate simply the realizations of a moving average process of order q having the Parameters $\theta_1, ..., \theta_p$ and whose variance of innovations is σ_ϵ^2, we can use the following algorithm:

1. Generate i.i.d. innovations $(\epsilon_t)_{t=1,...,T+T_0}$.
2. Calculate the values recursively

$$X_t = \sum_{j=1}^{p} \theta_j \epsilon_{t-j} + \epsilon_t.$$

Now, will write the following command lines:
$n = 500$
$x = rep(0, n)$
$epsilon = rnorm(n)$
$for(iin3 : n)$
$x[i] = 0.1 * epsilon[i - 1] - 0.7 * epsilon[i - 2] + rnorm(1)$
$op = par(mfrow = c(3, 1), mar = c(2, 4, 2, 2) + 0.1)$
$plot(ts(x), xlab = "", ylab = "MA(2)")$
$acf(x, main = "", xlab = "")$
$pacf(x, main = "", xlab = "")$
$par(op)$

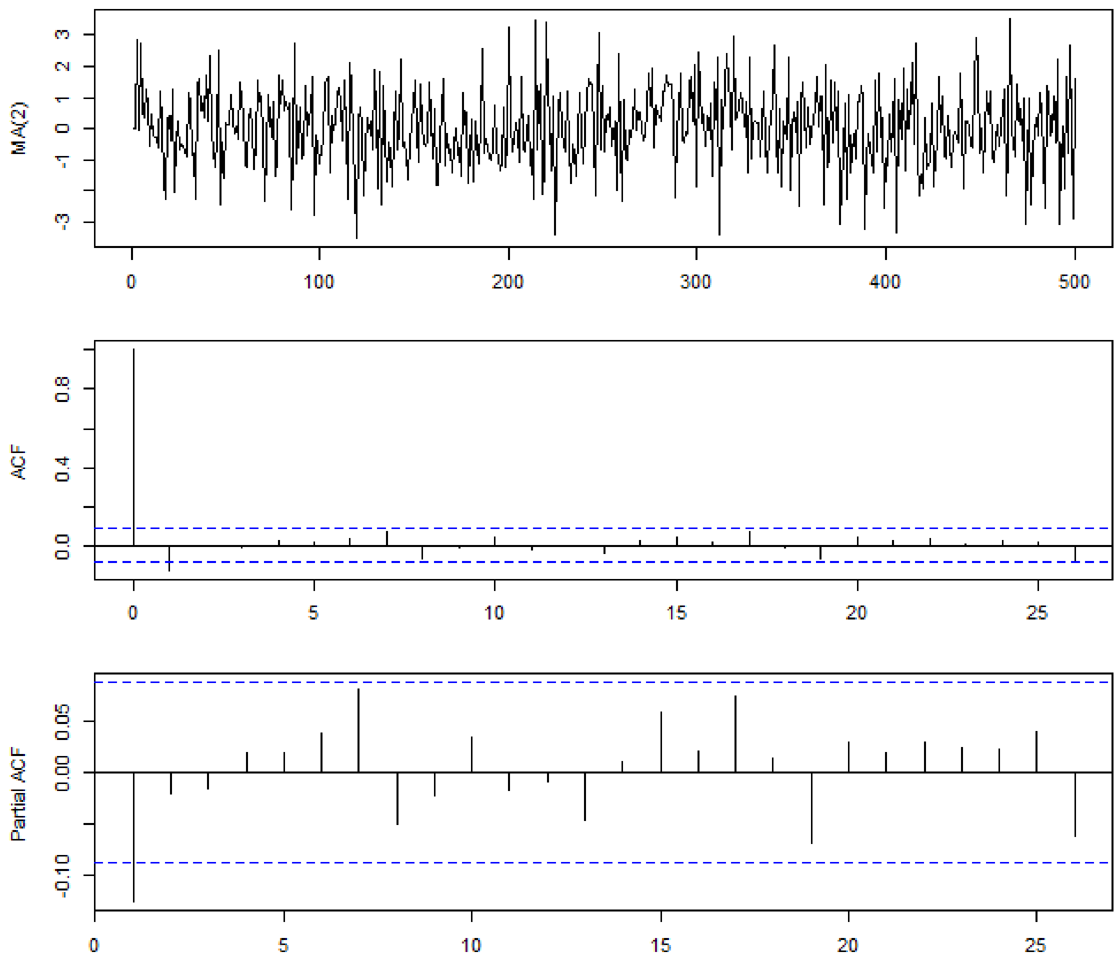

Figure 1.9: Simulation of MA(2) process.

1.8.3 Simulation of an ARIMA process

To directly simulate ARMA (or ARIMA) processes: We will now study the main parameters to use this function arima.sim to simulate an ARMA process, we need at least three elements: a model (a list with component **ar** and/or **ma** giving the AR and MA coefficients respectively. Optionally a component order can be used. An empty list offers an $ARIMA(0,0,0)$ model, that is white noise), a number of achievements and a process of innovation.
For example:
$n = 500$
$x = arima.sim(list(ar = c(0.11, -0.2, 0.2)), n)$
$op = par(mfrow = c(3,1), mar = c(2,4,2,2) + 0.1)$
$plot(ts(x), xlab = "", ylab = "ARIMA(3,0,0)")$
$acf(x, xlab = "", main = "")pacf(x, xlab = "", main = "")$
$par(op)$

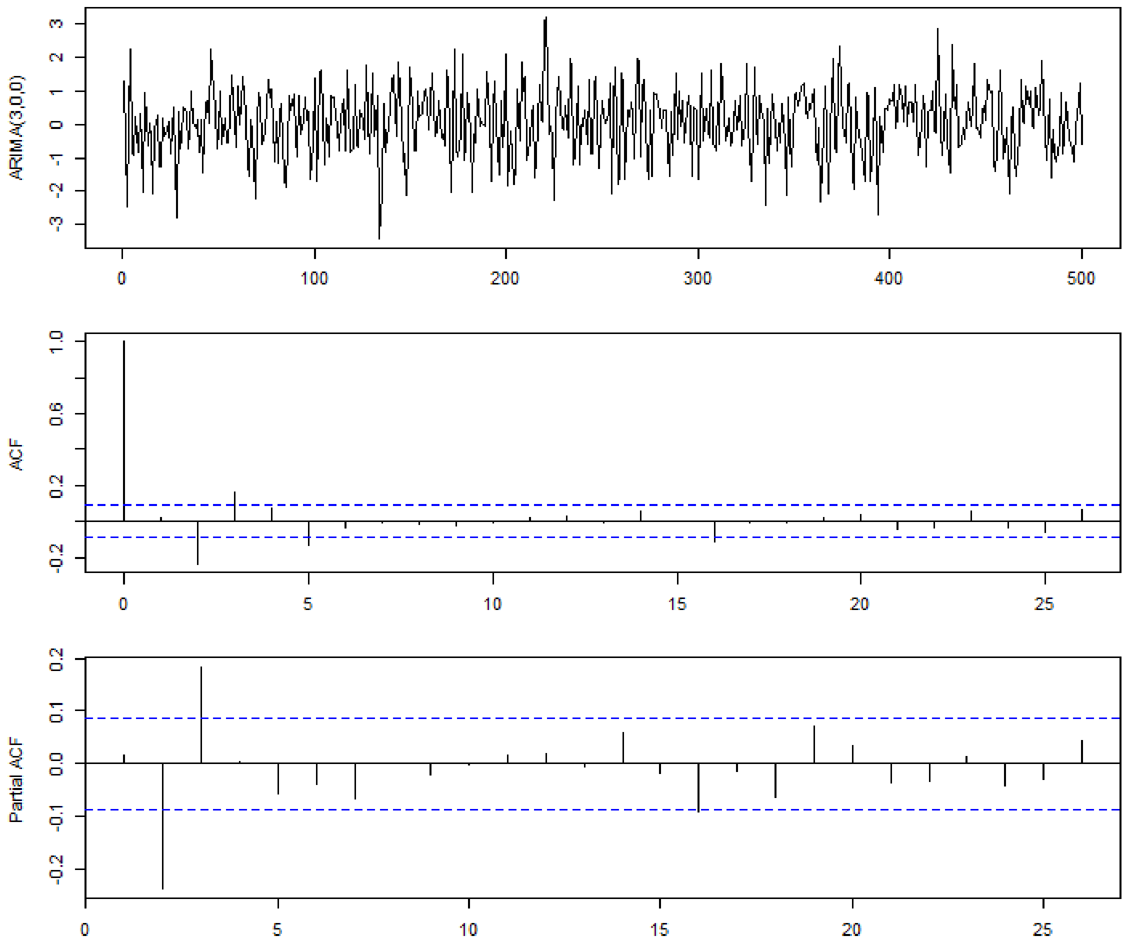

Figure 1.10: Simulation of ARIMA(3,0,0) process.

$n = 500$
$x = arima.sim(list(ma = c(0.9, -0.4, 0.2)), n)$
$op = par(mfrow = c(3, 1), mar = c(2, 4, 2, 2) + 0.1)$
$plot(ts(x), xlab = "", ylab = "ARIMA(0, 0, 3)")$
$acf(x, xlab = "", main = "")$
$pacf(x, xlab = "", main = "")$
$par(op)$

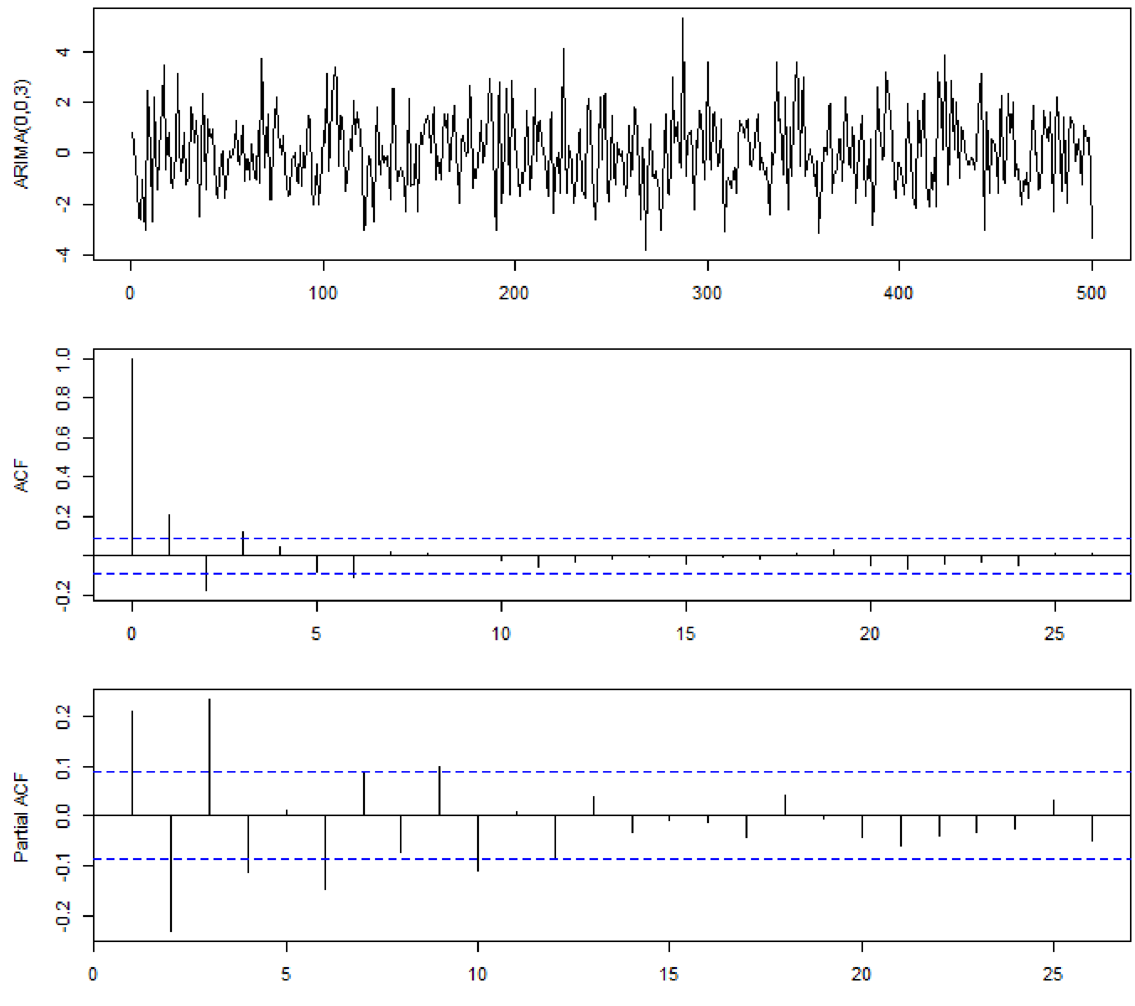

Figure 1.11: Simulation of ARIMA(0,0,3) process.

$n = 500$
$x = arima.sim(list(order = c(1, 1, 1), ar = 0.7, ma = 0.2), n)$
$op = par(mfrow = c(3, 1), mar = c(2, 4, 2, 2) + 0.1)$
$plot(ts(x), xlab = "", ylab = "ARIMA(1, 1, 1)")$
$acf(x, xlab = "", main = "")pacf(x, xlab = "", main = "")$
$par(op)$

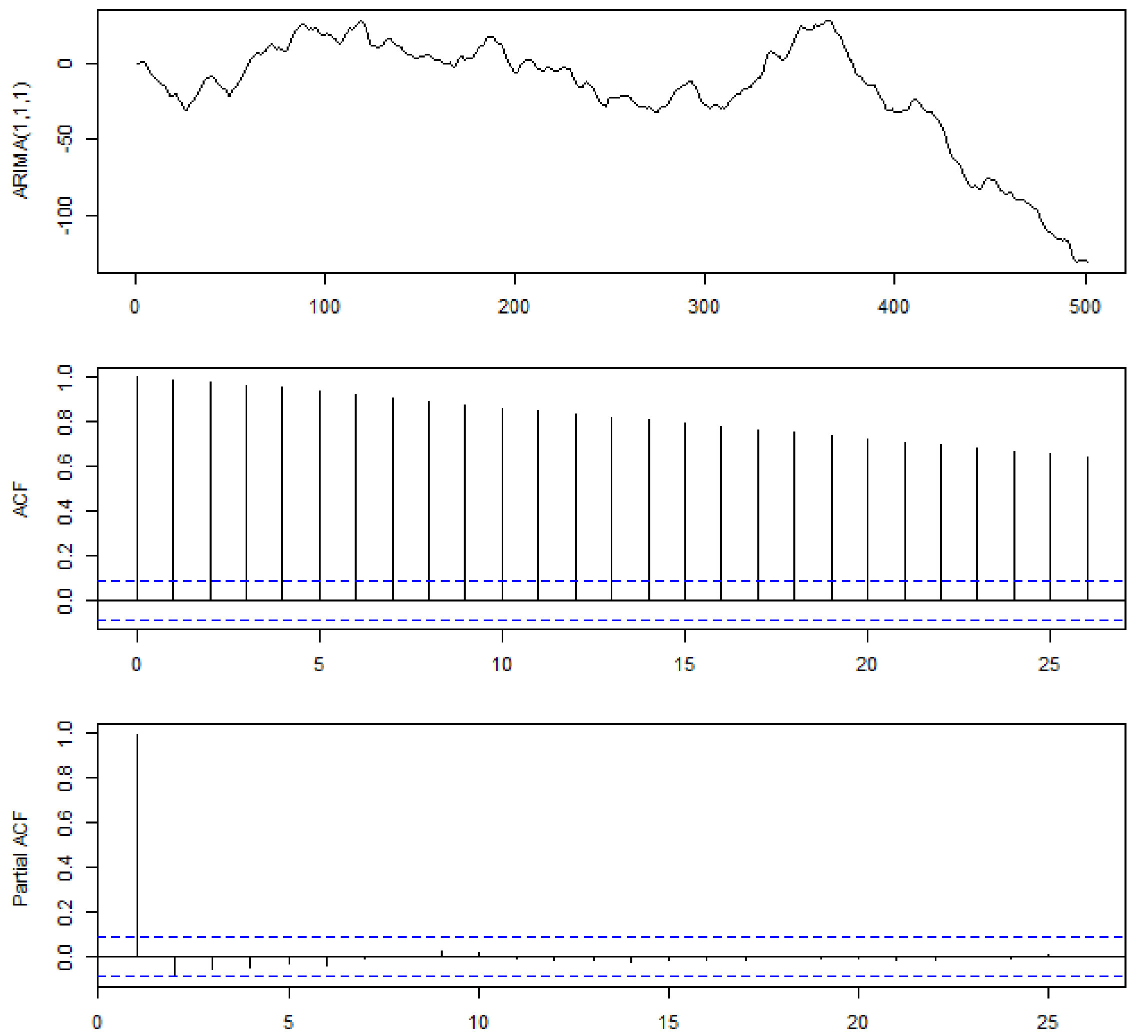

Figure 1.12: Simulation of ARIMA(1,1,1) process.

$n = 500$
$x = arima.sim(list(order = c(1, 0, 1), ar = -0.7, ma = -0.2), n)$
$op = par(mfrow = c(3, 1), mar = c(2, 4, 2, 2) + 0.1)$
$plot(ts(x), xlab = "", ylab = "ARIMA(1, 0, 1)")$
$acf(x, xlab = "", main = "")$
$pacf(x, xlab = "", main = "")$
$par(op)$

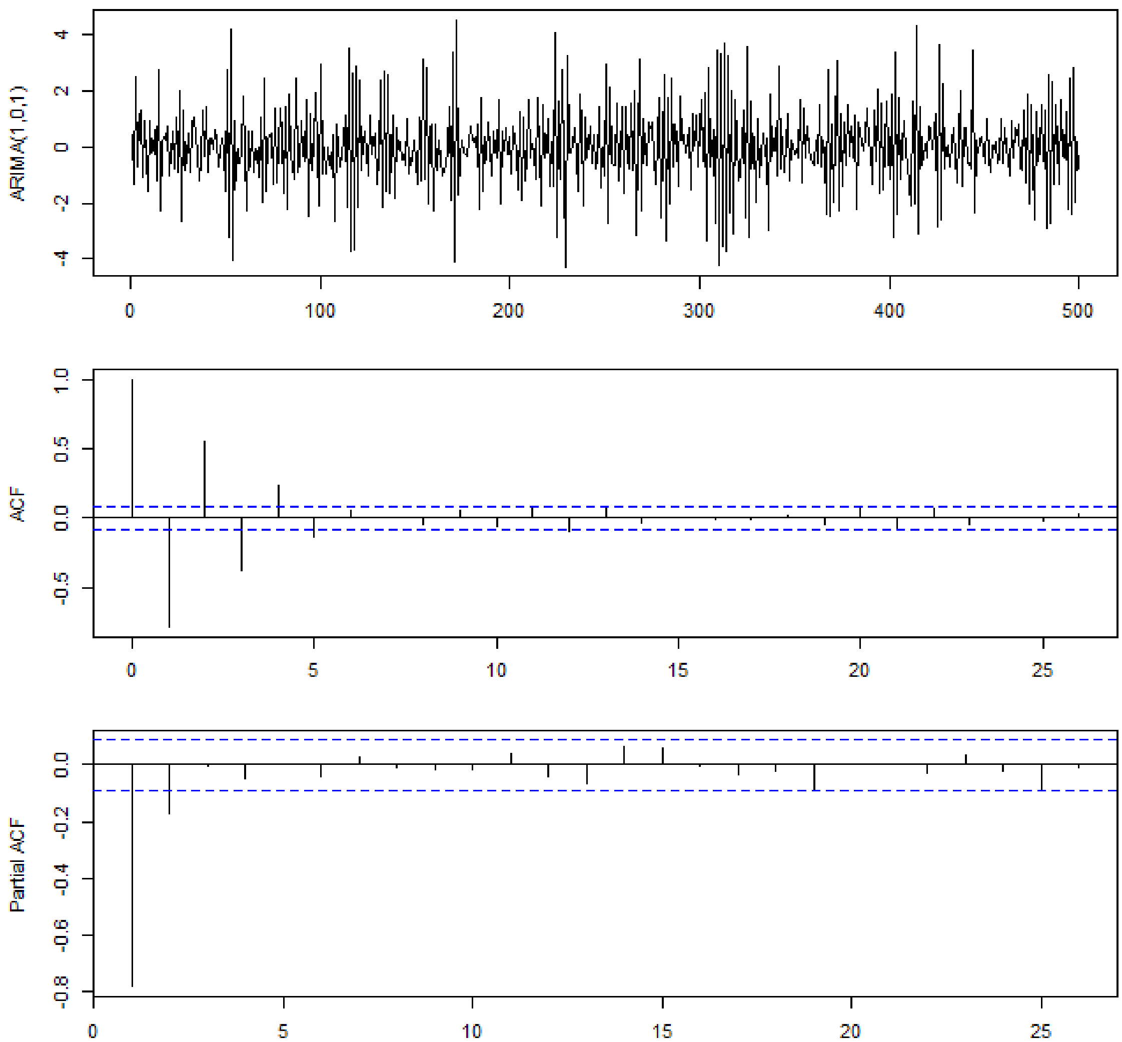

Figure 1.13: Simulation of ARIMA(1,1,1) process.

1.8.4 Simulation of an ARCH/GARCH processes

To simulate ARCH /GARCH models we will use garchSim from fGarch. This package is part of Rmetrics, cf. Wuertz and Rmetrics Foundation (2010), parallel to S + FinMetrics environment of which Zivot and Wang (2006).

Example 6. *We simulate a trajectory of an ARCH(1) following*

$$Y_i = 4 + X_i.$$
$$X_i = v_i h_i,$$
$$v_i \sim iidN(0, 1).$$
$$h_i^2 = 0.1 + 0.7X_{i-1}^2.$$

The model to be simulated is defined by garchSpec
$require(fGarch)$
$spec1 = garchSpec(model = list(mu = 4, omega = 0.1, alpha = 0.7, beta = 0))$
$arch.sim1 = garchSim(extended = TRUE, spec1, n = 500, n.start = 10)$

Since in the definition of the model v_t is a white noise Gaussian. Consider the simulated series arch.sim1[,1] (Figure 1.14) and the conditional standard deviation arch.sim1 [,2] (Figure 1.15).
$plot(arch.sim1[, 1])$
$plot(arch.sim1[, 2])$

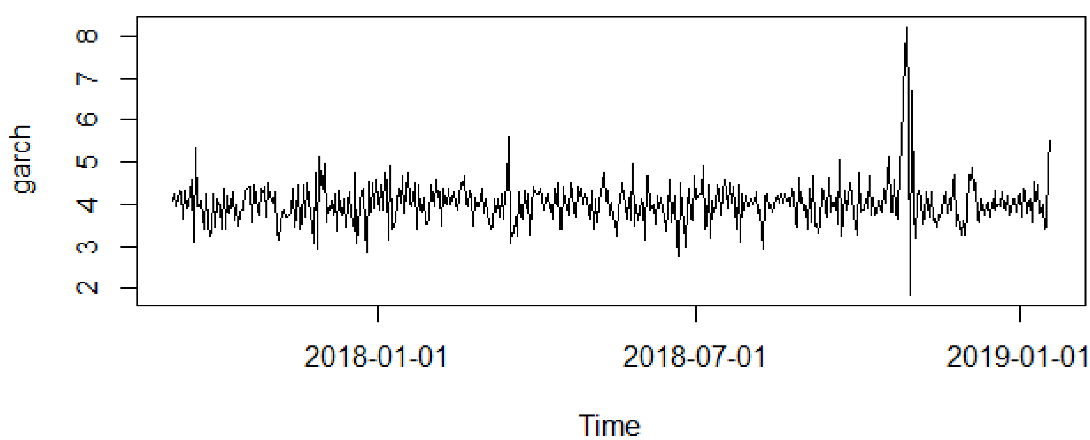

Figure 1.14: Simulated of series from ARCH(1) model.

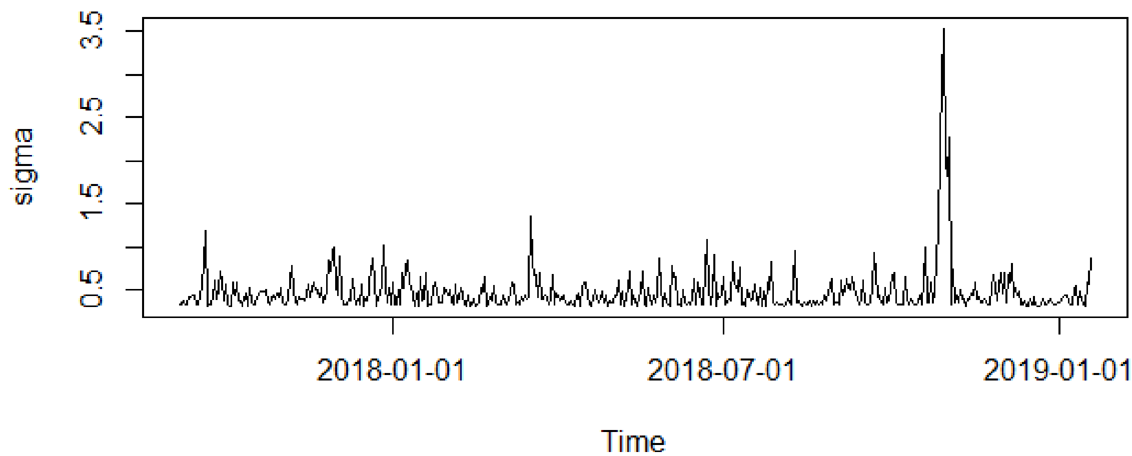

Figure 1.15: Simulated of conditional sigma from ARCH(1) model.

Example 7. *We simulate* 500 *observations according to the model*

$$Y_i = 2 + X_i.$$
$$X_i = v_i h_i.$$

$$h_i^2 = 0.01 + 0.7X_{i-1}^2 + 0.5X_{i-2}^2 + 0.3X_{i-3}^2 + 0.1h_{t-1}^2.$$

Now, we will write the following command lines:

$spec = garchSpec(model = list(mu = 2, omega = 0.01, alph = c(0.7, 0.5, 0.3), beta = 0.1))$
$y = garchSim(spec, n = 500, extended = TRUE)$
$y1 = y[1 : 500, 1]$
$y2 = y[1 : 500, 2]$
$plot.ts(y1)$
$plot.ts(y2)$

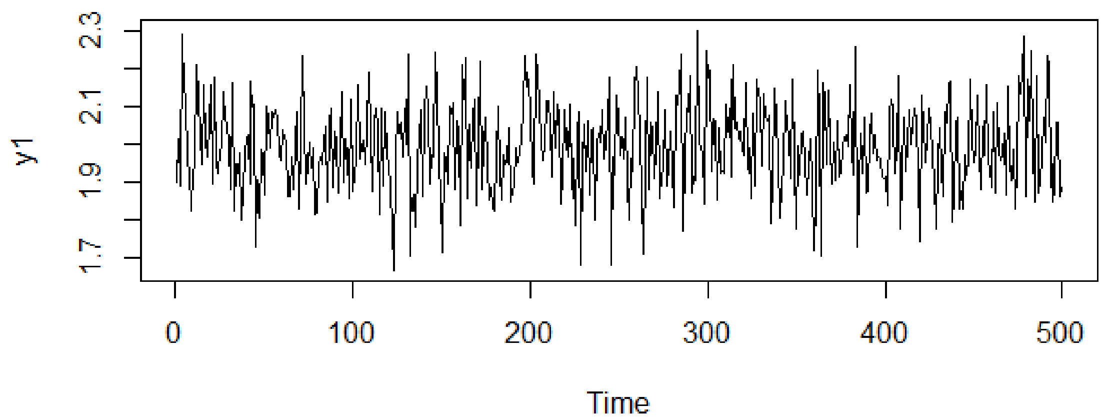

Figure 1.16: Simulated of series from GARCH(3,1) model.

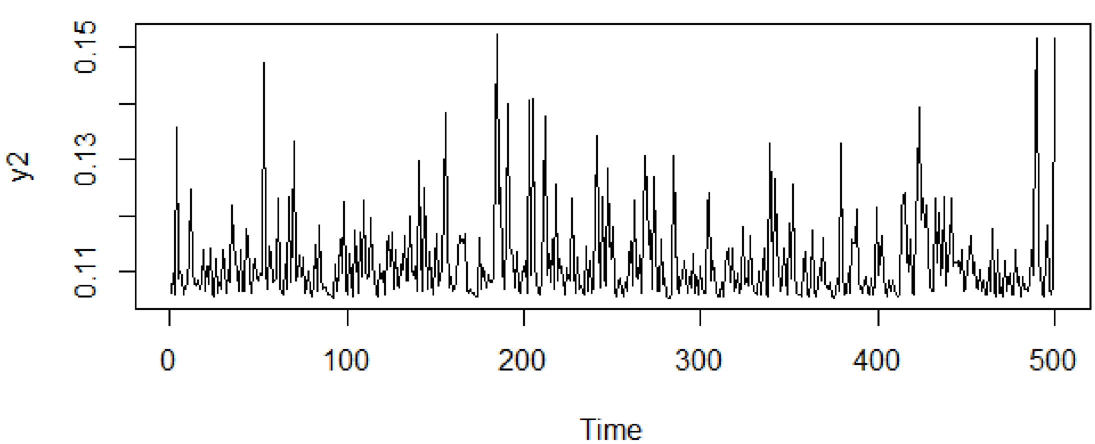

Figure 1.17: Simulated of conditional sigma from GARCH(3,1) model.

Chapter 2

Concentration Inequalities for Sums of Random Variables, Independent case

The objective of this chapter is to gives a presentation of the famous Bernstein inequality for the sums of independent and bounded random variables. We will then discuss the inequalities of hoeffding and Bennett.

2.1 Bernstein's inequality

The exponential inequality for sums of independent random variables was introduced by Sergei Bernstein [7]. In probability theory, Bernstein's inequality gives bounds on the probability that the sum of random variables deviates from their mean. Extensive studies of this inequality have been done in various fields such as model selection problem (Baraud, 2010 [2]), stochastic processes (Gao, Guillin and Wu 2014[22]). For example, Baraud (2010)[2] proposed a Bernstein type inequality for suprema of random processes with applications to model selection in non-Gaussian regression.
These inequalities, which can be primarily based on bounded independent random variables, are effective gear that can be applied in many areas such as laws of large numbers and asymptotics of inference problems. The importance of these inequalities have been demonstrated in lots of research of the asymptotic behavior of sums of independent bounded random variables, such as the laws (weak and strong) of large numbers and the probability of large deviations. One element that appears.

Theorem 6. *Let $X_1, ..., X_n$ be independent bounded random variables such that $\mathbb{E}[X_j] = 0$ and checking $|X_j| < b$ with probability 1 and $\sigma^2 = \frac{1}{n}\sum_{j=1}^{n} Var(X_j)$, then for all $d > 0$*

$$\mathbf{P}(\frac{1}{n}\sum_{j=1}^{n} X_j \geq t) \leq \exp\left(-\frac{nt^2}{2\sigma^2 + 2bt/3}\right). \tag{2.1}$$

Proof of Theorem 6.

Let $F_j = \sum_{r=2}^{\infty} \frac{\eta^{r-2}\mathbb{E}[X_j^r]}{r!\sigma_j^2}$,
where $\sigma_j^2 = \mathbb{E}[X_j^2]$.
Let $S_n = \sum_{j=1}^{n} X_j$ and by chernoff inequality, we have for any η positive

$$\begin{aligned}
\mathbf{P}(S_n \geq d) &= \mathbf{P}(S_n \geq d) \\
&\leq \exp(-\eta d)\mathbb{E}[\exp(\eta S_n)] \\
&\leq \exp(-\eta d)\prod_{j=1}^{n}\mathbb{E}[\exp(\eta X_j)],
\end{aligned}$$

or

$$\mathbb{E}[\exp(\eta X_j)] \leq 1 + \eta\mathbb{E}[X_j] + \sum_{r=2}^{\infty}\frac{\eta^r\mathbb{E}[X_j^r]}{r!}.$$

Since $\mathbb{E}[X_j] = 0$ we have,

$$\mathbb{E}[\exp(\eta X_j)] \leq 1 + F_j\eta^2\sigma_j^2 \leq \exp(F_j\eta^2\sigma_j^2).$$

Since expectation of a function is just the Lebesgue integral of the function with respect to probability measure, we have

$$\mathbb{E}[X_j^r] = \int_P X_j^{r-1}X_j \leq (\int_P |X_j^{r-1}|^2)^{1/2}(\int_P |X_j|^2)^{1/2} \leq \sigma_j(\int_P |X_j^{r-1}|^2)^{1/2},$$

applying Schwarz's inequality recursively n times, we obtain

$$\begin{aligned}
\mathbb{E}[X_j^r] &\leq \sigma_j^{1+\frac{1}{2}+\frac{1}{2}^2+\ldots+\frac{1}{2}^{n-1}}(\int_P |X_j^{(2^n r-2^{n+1}-1)}|)^{\frac{1}{2^n}} \\
&\leq \sigma_j^{2(1-\frac{1}{2}^n)}(\int_P |X_j^{(2^n r-2^{n+1}-1)}|)^{\frac{1}{2^n}},
\end{aligned}$$

we know that$|X_j| < b$. Therefore

$$\mathbb{E}[X_j^r] \leq \sigma_j^{2(1-\frac{1}{2}^n)}(b^{(2^n r-2^{n+1}-1)}|)^{\frac{1}{2^n}} \leq \sigma_j^{2(1-\frac{1}{2}^n)}(b^{(r-2-\frac{1}{2^n})}),$$

taking limit n to infinity we get

$$\mathbb{E}[X_j^r] \leq \sigma_j^2 b^{r-2}.$$

Therefore

$$\begin{aligned} F_j &= \sum_{r=2}^{\infty} \frac{\eta^{r-2}\mathbb{E}[X_j^r]}{r!\sigma_j^2} \\ &\leq \sum_{r=2}^{\infty} \frac{\eta^{r-2}\sigma_j^2 b^{r-2}}{r!\sigma_j^2} \\ &\leq \frac{1}{\eta^2 b^2}\sum_{r=2}^{\infty} \frac{\eta^r b^r}{r!} \\ &= \frac{1}{\eta^2 b^2}(e^{\eta b} - 1 - \eta b). \end{aligned}$$

Therefore

$$\mathbb{E}[X_j^r] \leq \exp\left(\eta^2\sigma_j^2 \frac{(e^{\eta b} - 1 - \eta b)}{\eta^2 b^2}\right),$$

we take $\sigma^2 = \frac{\sigma_j^2}{n}$

$$\mathbf{P}(S_n \geq d) \leq \exp(-\eta d)\exp\left(\eta^2 n\sigma^2 \frac{(e^{\eta b} - 1 - \eta b)}{\eta^2 b^2}\right),$$

the real η which minmising the second terme of the preceding inequality is

$$\eta = \frac{1}{b}\log(\frac{db}{n\sigma^2} + 1).$$

Therefore

$$\mathbf{P}(S_n \geq d) \leq \exp\left(\frac{n\sigma^2}{b^2}(\frac{db}{n\sigma^2} - \log(\frac{db}{n\sigma^2} + 1) - \frac{db}{n\sigma^2}log(\frac{db}{n\sigma^2} + 1))\right).$$

Let $H(x) = (1 + X)\log(1 + X) - X$, we get

$$\mathbf{P}(S_n \geq d) \leq \exp\left(\frac{-n\sigma^2}{b^2}H(\frac{db}{n\sigma^2})\right).$$

This is known as the Bennett's inequality (3.3), We can derive the Bernstien's inequality by further bounding the function $H(x)$.
Let $G(x) = \frac{3}{2}\frac{x^2}{x+3}$, $H(0) = G(0) = H'(0) = G'(0) = 0$.
$H''(x) = \frac{1}{x+1}$ and $G''(x) = \frac{27}{(x+3)^3}$.
Therefore $H''(0) \geq G''(0)$ and further if $f^n(x)$ of a function f represents the n^th derivative of the function then we have $H^n(0) \geq G^n(0)$ fo all $(n \geq 2)$. Consequently, according to the theorem Taylor's we have

$$H(x) \geq G(x)\forall x \geq 0.$$

Therefore

$$\begin{aligned} \mathbf{P}(S_n \geq d) &\leq \exp\left(\frac{-n\sigma^2}{b^2}G(\frac{db}{n\sigma^2})\right) \\ &\leq \exp\left(\frac{-d^2}{2(db + 3n\sigma^2)}\right), \end{aligned}$$

now let $d = nt$. Therefore

$$\begin{aligned} \mathbf{P}(\sum_{j=1}^{n} X_j \geq nt) &\leq \exp\left(\frac{-t^2n^2}{2(ntb + 3n\sigma^2)}\right) \\ &\leq \exp\left(-\frac{nt^2}{2\sigma^2 + 2bt/3}\right). \end{aligned}$$

2.2 Hoeffding's inequality

Hoeffding's inequality [23] is an inequality of concentration concerning the sums of independent and bounded random variables.

The Hoeffding's inequality offers an exponential bound on the probability of the deviation among the average of n independent bounded random variables and its mean. The study of this inequality has caused interesting applications in probability theory and statistics (Boucher, 2009 [10]; Yao and Jiang, 2012 [45]).

later on, in 2014, Hoeffding's inequalities for geometrically ergodic Markov chains on general state space have been proved by way of Miasojedow (2014)[29]. Recently, Tang (2007) proved an extension of Hoeffding's inequality in a category of ergodic time series. additionally, new extensions of this inequality for panel data have been proposed by Yao and Jiang (2012)[45].

Lemma 1. *(Hoeffding lemma). Let X be a bounded random variable with $a \leq X \leq b$. Then for any real ω,*

$$\mathbb{E}[\exp(\omega(X - \mathbb{E}[X]))] \leq \exp\left(\frac{\omega^2(b-a)^2}{8}\right). \tag{2.2}$$

Proof 5. *Let X' be an independent copy of X with the same distribution, in order that $a \leq X' \leq b$ and $\mathbb{E}[X'] = \mathbb{E}[X]$, however X and X' are independent. Then*

$$\mathbb{E}_X[\exp(\omega(X - \mathbb{E}_X[X]))] = \mathbb{E}_X[\exp(\omega(X - \mathbb{E}_{X'}[X']))] \leq \mathbb{E}_X[\mathbb{E}_{X'} \exp(\omega(X - X'))]$$

(by Jensen's inequality).

Now, we have

$$\mathbb{E}[\exp(\omega(X - \mathbb{E}[X]))] \leq \mathbb{E}[\exp(\omega(X - X'))].$$

We note a curious fact: the difference $X - X'$ is symmetric about zero, so that if $K \in \{-1, 1\}$ be a random sign variable, then $K(X\text{-}X')$ has precisely the same distribution as $X - X'$, we have

$$\begin{aligned} \mathbb{E}_{X,X'}[\exp(\omega(X - X'))] &= \mathbb{E}_{X,X',K}[\exp(\omega K(X - X'))] \\ &= \mathbb{E}_{X,X'}[\mathbb{E}_K[\exp(\omega K(X - X')) \setminus X, X']]. \end{aligned}$$

Use inequality $\mathbb{E}[e^{\omega K}] \leq \exp(\frac{\omega^2}{2})$ for all $\omega \in \mathbb{R}$ on the moment producing function of the *random sign, which offers that*

$$\mathbb{E}_K[\exp(\omega K(X-X')) \setminus X, X'] \leq \exp\left(\frac{\omega^2(X-X')^2}{2}\right).$$

By assumption we have $|X - X'| \leq (b-a)$*, so* $(X-X')^2 \leq (b-a)^2$.

$$\mathbb{E}_{X,X'}[\exp(\omega(X-X'))] \leq \exp\left(\frac{\omega^2(b-a)^2}{2}\right).$$

Theorem 7. *(Hoeffding's inequality) Let* $X_1, ..., X_n$ *be independent bounded random variables such that* $a \leq X_j \leq b$ *for all* $1 \leq j \leq n$ *where* $-\infty < a \leq b < \infty$*. Then*

$$\mathbf{P}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j - \mathbb{E}[X_j]) \geq t\right) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right),$$

and

$$\mathbf{P}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j - \mathbb{E}[X_j]) \leq -t\right) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right),$$

for all $t \geq 0$.

Proof of Theorem 7.
Using the Hoeffding lemma, and the Chernoff inequality, we have

$$\begin{aligned}
\mathbf{P}\left(\tfrac{1}{n}\sum_{j=1}^{n}(X_j - \mathbb{E}[X_j]) \geq t\right) &\leq \mathbf{P}\left(\sum_{j=1}^{n}(X_j - \mathbb{E}[X_j]) \geq nt\right) \\
&\leq \mathbb{E}\left[\left(\exp(\omega\sum_{j=1}^{n}(X_j - \mathbb{E}[X_j])\right)\right] e^{-\omega nt} \\
&= \left(\prod_{j=1}^{n}\mathbb{E}\left[e^{\omega(X_j - \mathbb{E}[X_j])}\right]\right) e^{-\omega nt} \\
&\leq \left(\prod_{j=1}^{n} e^{\dfrac{\omega^2(b-a)^2}{8}}\right) e^{-\omega nt}.
\end{aligned}$$

Minimise the second term of the preceding inequality for $\omega \geq 0$, we have

$$\mathbf{P}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j - \mathbb{E}[X_j]) \geq t\right) \leq \min\exp(\frac{n\omega^2(b-a)^2}{8} - \omega nt) = \exp(-\frac{2nt^2}{(b-a)^2}).$$

2.3 Bennett inequality

Bennett's inequality presents an upper bound on the probability that the sum of independent random variables deviates from its expected value.

Theorem 8. *Let $X_1, ..., X_n$ be a finite sequence of independent random variables, and assume that $\mathbb{E}[X_j] = 0$ and $\mathbb{E}[X_j^2] = \sigma_i^2$, $|X_j| < b$ almost surely. Then, for any $0 \leq d < nt$*

$$\mathbf{P}(\sum_{j=0}^{n} X_j \geq d) \leq \exp\left(-\frac{n\sigma^2}{b^2} H(\frac{db}{n\sigma^2})\right), \tag{2.3}$$

where $H(X) = (1 + X)\log(1 + X) - X$, $n\sigma^2 = \sum_{j=1}^{n} \sigma_i^2$.

Proof of Theorem 8.
see the proof of this theorem in the proof of theorem of inequality of Bernstein.

There are some efforts seeking to refine the Bennett's inequality. In Fan (2015a) [19], a missing factor of order $1/t$ is added to Bennett's inequality underneath the Bernstein's condition, in Pinelis (2014)[31], below the condition imposed to the third order moments, the author developed a pointy improvement to Bennett's inequality. However, they did not recall the variations a few of the variances of the random variables.

Chapter 3

Probability Tail for Extended Negatively Dependent Random Variables of Partial Sums and Application to AR(1) Model Generated By END Errors

Abstract 1. The exponential probability inequalities have been important tools in probability and statistics. In this paper, we establish exponential inequalities for END random variables of partial sums which enable us to build a confidence interval for the parameter of the first-order autoregressive process. In addition, Using these inequalities, We prove that the estimator complete converge to the unknown parameter θ.

3.1 Introduction

The autoregressive process takes an important part in predicting problems leading to decision making. The estimation of the unknown parameter θ of the order 1 autoregressive process are obtained by least square method.
Many research articles and text books have contributed to the expansion of linear and nonlinear autoregressive models (Belguerna and Benaissa [3]; Dahmani and Tari [16]; Galtchouk and Konev [21]).
In their fundamental work, (Chan and Wei [14]) they evaluated the limit in law of the least squares estimator in this instance the errors constitute a sequence of martingale differences.
The exponential inequalities are obtained by (Bondarev [9]) that have been used to construct a confidence interval for the unknown parameter θ_0 in the equation

$$\frac{dx}{dt} = \theta_0 f(t, x(t)) + \zeta'(t), \qquad x(0) = \zeta(0) = 0$$

where ζ' is a Gaussian noise with zero mean and known correlation function.
The study of The probability inequalities for extended negatively dependent random variables are derived originally by (Fakoor and Azarnoosh [18], Asadian and Fakoor [1]).

In this work we create a new exponential inequalities of Bernstein fréchet type for the parameter of the order 1 autoregressive process (AR(1)) in case of noise extended negatively dependent.
These inequality will allow us to build a confidence interval for this parameter and to show the complete convergence for this estimator.

3.2 Optimal Control

We will remember the definition of extended negatively dependent (END) sequences, and some lemmas.

Definition 17. *(Liu[27]). a sequence $\{\zeta_n, n \geq 1\}$ of random variables is said END if there exists a constant $M > 0$ such that*

$$\mathbf{P}(\zeta_1 > \epsilon_1, \zeta_2 > \epsilon_2, ..., \zeta_n > \epsilon_n) \leq M \prod_{i=1}^{n} \mathbf{P}(\zeta_i > \epsilon_i) \tag{3.1}$$

and

$$\mathbf{P}(\zeta_1 \leq \epsilon_1, \zeta_2 \leq \epsilon_2, ..., \zeta_n \leq \epsilon_n) \leq M \prod_{i=1}^{n} \mathbf{P}(\zeta_i \leq \epsilon_i) \tag{3.2}$$

if for all real numbers $\epsilon_1, \epsilon_2, ..., \epsilon_n$.

The extended negative dependence was introduced by (Liu [27]), it was found on some of the applications of the sequence END. (Liu [28]) examined the necessary and sufficient conditions of moderate deviations for dependent random variables with heavy tails. It can be clearly seen that NOD random variables and independent random variables are END. Also (Joag-Dev and Proschan [25]) have pointed out that negatively associeted (NA) random variables are negatively orthoant dependent (NOD), consequently NA random variables are END. The study of a limit patterns of END sequence is of interest.

Lemma 2. *(Liu[28]). Let $X_1, X_2, ..., X_n$ be the random variables END, then*

(i) *If $h_1, h_2, ..., h_n$ are all non decreasing (or non increasing) function, then random variables $h_1(X_1), h_2(X_2), ..., h_n(X_n)$ are END.*

(ii) *For every $n \geq 1$, there is a positive constant M, de note $X^+ = \max\{0, X\}$ such that*

$$\mathbf{E}(\prod_{j=1}^{n} X_j^+) \leq M \prod_{j=1}^{n} \mathbf{E}(X_j^+). \tag{3.3}$$

Lemma 3. *if $\{X_n, n \geq 1\}$ be an END sequence and $t > 0$ then for all $n \geq 1$, there is a positive constant M such that*

$$\mathbf{E}\left[\prod_{j=1}^{n}\exp(tX_j)\right] \leq M\prod_{j=1}^{n}\mathbf{E}\left[\exp(tX_j)\right]. \tag{3.4}$$

Lemma 4. *(Bentkus [5]). For $g \geq 0$ and $\dot{\eta} \geq 1$ we have*

$$\frac{1}{1+\dot{\eta}}\exp(-g\dot{\eta}) + \frac{\dot{\eta}}{1+\dot{\eta}}\exp(g) \leq \exp(\frac{g^2\dot{\eta}}{2}). \tag{3.5}$$

Throughout the paper, let random variables $X_1, X_2, ..., X_n$ defined on a fixed probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let M and C and L be positive constants, which can be different in various places.

3.2.1 Model and hypotheses

We propose the hypothesis that we use to be able to state exactly our result.

Let us consider the order 1 autoregressive process AR(1) defined by

$$X_k = \theta X_{k-1} + \zeta_k, \tag{3.6}$$

where θ is the autoregressive parameter with $|\theta| < 1$, and where $\{\zeta_k, k \geq 0\}$ is a sequence of identically distributed END random variables and finite variance, with $\zeta_0 = X_0 = 0$.

We can estimate the parameter θ by the method of least squares, and the estimator θ_n given, for all $n \geq 1$, by

$$\theta_n = \frac{\sum_{k=1}^{n} X_{k-1}X_k}{\sum_{k=1}^{n} X_{k-1}^2}, \tag{3.7}$$

and

$$\theta_n - \theta = \frac{\sum_{k=1}^{n} X_{k-1}\zeta_k}{\sum_{k=1}^{n} X_{k-1}^2}. \tag{3.8}$$

We introduce now the following hypothese:

H1: We assume that the sequences $(\zeta_n)_n$ and $(X_{n-1})_n$ are bounded. Then there exist positive constants L and C such that
$|X_{n-1}| \leq L, |\zeta_n| \leq C.$

H2: Suppose that $\mathbf{E}(\zeta_k) = 0$, for each $k \geq 1$.

H3: Assume that $\mathbf{E}(\phi^2(\zeta_1, ..., \zeta_n)) \leq B_2 < +\infty$ and $\mathbf{E}(\exp(\zeta_1^4)) \leq \frac{1}{3^{(n+1)}}, n \in \mathbb{N}^*$.

3.3 Main Result

Theorem 9. *Under hypotheses* $(H_1), (H_2)$ *and* (H_3)*, for any* $\widetilde{\epsilon} < \log(\frac{3}{2})$ *positive and for* W *rather large, we have*

$$\mathbf{P}(\sqrt{n}|\theta_n - \theta| > W) \leq M \exp(-\sqrt{n} W \widetilde{\epsilon} B_1) + B_2 M \sqrt{\frac{3}{2}} \exp\left(-\frac{n^2}{2}(\log(\frac{3}{2}) - \frac{1}{n}\widetilde{\epsilon})\right), \quad (3.9)$$

where
$B_1 = \frac{1}{2CL} arcsinh\left(\frac{C\sqrt{n}W\widetilde{\epsilon}}{2LD_n}\right)$, $\quad B_2$ *is a positive constant.*

Proof of Theorem 9.

In line with the equality (3.8), we have

$$\begin{aligned} \mathbf{P}(\sqrt{n}|\theta_n - \theta| > W) &= \mathbf{P}\left(\left|\frac{\frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_{k-1}\zeta_k}{\frac{1}{n}\sum_{k=1}^{n} X_{k-1}^2}\right| > W\right) \\ &\leq \mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_{k-1}\zeta_k| > \frac{W}{n}\sum_{k=1}^{n} X_{k-1}^2\right). \end{aligned}$$

By virtue of the probability properties, we have for any $\widetilde{\epsilon}$ positive

$$\mathbf{P}(\sqrt{n}|\theta_n - \theta| > W) \leq \mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{\mathbf{k=1}}^{n} X_{k-1}\zeta_k| > W\widetilde{\epsilon}\right) + \mathbf{P}\left(\frac{1}{n}\sum_{\mathbf{k=1}}^{n} X_{k-1}^2 \leq \widetilde{\epsilon}\right). \quad (3.10)$$

Let us now bound the first probability of the right hand-side of the inequality (3.10). Taking the chernoff inequality, we have for any λ positive

$$\begin{aligned}
\mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{\mathbf{k=1}}^{n}X_{k-1}\zeta_k|>W\widetilde{\epsilon}\right) &\leq \mathbf{P}\left(\tfrac{L}{\sqrt{n}}\sum_{\mathbf{k=1}}^{n}|\zeta_k|>W\widetilde{\epsilon}\right)\\
&\leq \mathbf{P}\left(L\sum_{\mathbf{k=1}}^{n}|\zeta_k|>\sqrt{n}W\widetilde{\epsilon}\right)\\
&\leq \mathbf{P}\left(\sum_{\mathbf{k=1}}^{n}|\zeta_k|>\frac{\sqrt{n}}{L}W\widetilde{\epsilon}\right)\\
&\leq \exp(\tfrac{-\lambda\sqrt{n}W\widetilde{\epsilon}}{L})\mathbf{E}(\exp(\lambda S_n)),
\end{aligned}$$

where $S_n=\sum_{k=1}^{n}\widetilde{\zeta_k}$, $\widetilde{\zeta_k}=|\zeta_k|$ and $D_n=\sum_{k=1}^{n}\mathbf{E}(\zeta_k^2)$ for each $n\geq 1$.

Thus, for any $\lambda>0$, $k=1,2,...,n$ and by the same argument in proof of theorem 1 in Pgokhorov [30] we have

$$\begin{aligned}
\mathbf{E}(\exp(\lambda\zeta_k)-1) &= \mathbf{E}(\exp(\lambda\zeta_k)-\lambda\zeta_k-1)\leq\mathbf{E}(\exp(\lambda\zeta_k)+\exp(-\lambda\zeta_k)-2)=2\mathbf{E}(\cosh\lambda\zeta_k-1)\\
&= 2\mathbf{E}(\cosh\lambda\widetilde{\zeta_k}-1) \qquad \leq\mathbf{E}(\lambda\widetilde{\zeta_k}\sinh\lambda\widetilde{\zeta_k})\\
&= \mathbf{E}(\lambda^2\zeta_k^2\frac{\sinh\lambda\widetilde{\zeta_k}}{\lambda\widetilde{\zeta_k}}) \qquad \leq\frac{\lambda\mathbf{E}(\zeta_k^2)}{C}\sinh\lambda C.
\end{aligned}$$

Using $z\leq\exp(z-1)$ for all $z\in\mathbb{R}$, We can write by **Lemma 2.** that

$$\mathbf{E}(\prod_{k=1}^{n}\exp(\lambda\zeta_k))\leq M\prod_{k=1}^{n}\mathbf{E}(\exp(\lambda\zeta_k))\leq M\prod_{k=1}^{n}\exp(\mathbf{E}(\exp(\lambda\zeta_k)-1))\leq M\exp(\lambda D_n\frac{\sinh\lambda C}{C}),$$

where C is positive constant.

We obtain

$$\mathbf{E}(\exp(\lambda S_n))\leq M\exp(\frac{\lambda D_n}{C}\sinh\lambda C).$$

The inequality

$$\begin{aligned}
\mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_{k-1}\zeta_k|>W\widetilde{\epsilon}\right) &\leq M\exp\left(\frac{-\lambda\sqrt{n}W\widetilde{\epsilon}}{L}+\frac{\lambda D_n}{C}\sinh\lambda C\right)\\
&\leq M\exp\left(\lambda(\frac{D_n}{C}\sinh\lambda C-\frac{\sqrt{n}W}{L}\widetilde{\epsilon})\right).
\end{aligned}$$

Choosing $\lambda=\frac{1}{C}\text{arcsinh}\left(\frac{C\sqrt{n}W\widetilde{\epsilon}}{2D_nL}\right)$ and we can see that $D_n\frac{\sinh\lambda C}{C}=\frac{\sqrt{n}W\widetilde{\epsilon}}{2L}$.

Then

$$\mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_{k-1}\zeta_k|>W\widetilde{\epsilon}\right)\leq M\exp\left(-\frac{1}{2C}\text{arcsinh}\left(\frac{C\sqrt{n}W\widetilde{\epsilon}}{2D_nL}\right)\frac{\sqrt{n}W\widetilde{\epsilon}}{L}\right).$$

Let us notice that for W rather large.

$$\left(\frac{1}{2CL}\text{arcsinh}\left(\frac{C\sqrt{n}W\widetilde{\epsilon}}{2LD_n}\right)\right) \text{ is a positive, finite and non nulle.}$$

Therefore

$$\mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_{k-1}\zeta_k| > W\widetilde{\epsilon}\right) \leq M\exp(-\sqrt{n}W\widetilde{\epsilon}B_1), \tag{3.11}$$

where $B_1 = \frac{1}{2CL}\text{arcsinh}\left(\frac{C\sqrt{n}W\widetilde{\epsilon}}{2LD_n}\right)$ is a positive constant.
Now, we limite the second probability of the right hand-side of the expression (3.10).

Acoording to chernoff inequality , for any $t > 0$

$$\mathbf{P}(\frac{1}{n}\sum_{k=1}^{n}X_{k-1}^2 \leq \widetilde{\epsilon}) \leq \exp(t\widetilde{\epsilon})\mathbf{E}\left(\exp(-\frac{t}{n}\sum_{k=1}^{n}X_{k-1}^2)\right). \tag{3.12}$$

Let g be a bounded and measurable function, defined on $\mathbb{R}^n$ and its values taken in $\mathbb{R}$ and let T be a bijective application changing the vector $(\zeta_1, \zeta_2, ..., \zeta_n)$ on the vector $(X_1, X_2, ..., X_n)$, we have then Guikhman and Skorokhod [15].

$$\int_{\mathbb{R}^n} g(x)\mu_x(dx) = \int_{\mathbb{R}^n} g(T(x))\mu_\zeta(dx).$$

where μ_X measure generated by $(X_1, X_2, ..., X_n)$ and μ_ζ is a measure generated by $(\zeta_1, \zeta_2, ..., \zeta_n)$
Changing variables by letting $T^{-1}X_k = \zeta_k$, we obtain

$$\int_{\mathbb{R}^n} g(x)\mu_x(dx) = \int_{\mathbb{R}^n} g(x)\phi(x)\left|\frac{DT^{-1}(x)}{Dx}\right|\mu_\zeta(dx),$$

where $\frac{DT^{-1}(x)}{Dx}$ is the Jacobian of the inverse application T^{-1}.
Using the definition of autoregressive process of first order, we have

$$\begin{aligned}
\zeta_1 &= X_1 \\
\zeta_2 &= X_2 - \theta X_1 \\
&. \\
&. \\
&. \\
\zeta_n &= X_n - \theta X_{n-1}.
\end{aligned}$$

In accordance with the properties of the absolute continuity of the measure μ_ζ when compared with the measure generated by $X_1, ..., X_n$, we have

$$\mathbf{E}\left(\exp(-\frac{t}{n}\sum_{k=1}^{n} X_{k-1}^2)\right) = \mathbf{E}\left(\phi(\zeta_1, \zeta_2, ..., \zeta_n)\exp(-\frac{t}{n}\sum_{k=1}^{n}\zeta_k^2)\right), \tag{3.13}$$

where ϕ as a density of $\zeta = (\zeta_1, \zeta_2, ..., \zeta_n)$.
We apply the inequality of the Cauchy-Shwarz in the last equation and by **Lemma 2.**, we write

$$\begin{aligned}\mathbf{E}\left(\exp(-\tfrac{t}{n}\sum_{k=1}^{n} X_{k-1}^2)\right) &= \mathbf{E}\left(\phi^2(\zeta_1, \zeta_2, ..., \zeta_n)\right)^{1/2}\mathbf{E}\left(\exp(-\tfrac{2t}{n}\sum_{k=1}^{n-1}\zeta_k^2)\right)^{1/2}\\ &\leq B_2 M\prod_{k=1}^{n-1}\mathbf{E}\left(\exp(-\frac{2t}{n}\zeta_k^2)\right)^{1/2}\\ &= B_2 M\prod_{k=1}^{n-1}\mathbf{E}\left((1+\frac{2t}{n})\frac{1}{1+\frac{2t}{n}}\exp(-\frac{2t}{n}\zeta_k^2)\right)^{1/2}.\end{aligned}$$

Accorging to **lemma 3.**, we obtain

$$\begin{aligned}\mathbf{E}\left(\exp(-\tfrac{t}{n}\sum_{k=1}^{n} X_{k-1}^2)\right) &\leq B_2 M\left(\prod_{k=1}^{n-1}(1+\frac{2t}{n})\mathbf{E}\left(\exp(\frac{2t}{n}\zeta_k^4) - \frac{\frac{2t}{n}}{1+\frac{2t}{n}}\exp(\zeta_k^2)\right)\right)^{1/2}\\ &= B_2 M\left(1+\tfrac{2t}{n}\right)^{\frac{n-1}{2}}\left(\prod_{k=1}^{n-1}\mathbf{E}\left(\exp(\frac{2t}{n}\zeta_k^4) - \frac{\frac{2t}{n}}{1+\frac{2t}{n}}\exp(\zeta_k^2)\right)\right)^{1/2}.\end{aligned}$$

Choosing $t = \frac{n}{2}$,

$$\mathbf{E}\left(\exp(-\frac{t}{n}\sum_{k=1}^{n} X_{k-1}^2)\right) \leq B_2 M 2^{\frac{n-1}{2}}\left(\prod_{k=1}^{n-1}\left(\mathbf{E}(\exp(\zeta_k^4) - \frac{1}{2}\exp(\zeta_k^2))\right)\right)^{1/2}.$$

Now, we apply Jensen's inequality to the term on the right

$$\begin{aligned}\mathbf{E}\left(\exp(-\tfrac{t}{n}\sum_{k=1}^{n} X_{k-1}^2)\right) &\leq B_2 M 2^{(n-1)/2}\left(\prod_{k=1}^{n-1}\left(\mathbf{E}(\exp(\zeta_k^4)) - \frac{1}{2}\exp(\mathbf{E}(\zeta_k^2))\right)\right)^{1/2}\\ &\leq B_2 M 2^{(n^2-1)/2}(\mathbf{E}(\exp(\zeta_1^4)))^{(n-1)/2}\\ &\leq B_2 M\left(\tfrac{2}{3}\right)^{(n^2-1)/2} = B_2 M\left(\tfrac{3}{2}\right)^{1/2}\exp(-\tfrac{n^2}{2}\log(\tfrac{3}{2})).\end{aligned} \tag{3.14}$$

Therefore

$$\mathbf{P}(\frac{1}{n}\sum_{k=1}^{n}X_{k-1}^2 \leq \widetilde{\epsilon}) \leq B_2M\sqrt{\frac{3}{2}}\exp(-\frac{n^2}{2}(\log\left(\frac{3}{2}\right) - \frac{1}{n}\widetilde{\epsilon})). \tag{3.15}$$

Considering the relations (3.11) and (3.15) together and taking in to account the expression (3.10) we obtain the result.

Corollary 1. *The sequence $(\theta_n)_{n\in\mathbb{N}}$ defined in (3.7) converges completely to the parameter θ of the autoregressive process of order 1.*

Proof 6. *The complete convergence follows from the inequalities (3.10).*
Indeed, applying Linearity property ($\sum_{n=1}^{+\infty} V_n = \sum_{n=1}^{+\infty}(U_n + Z_n) = \sum_{n=1}^{+\infty} U_n + \sum_{n=1}^{+\infty} Z_n$) on the positive réel term sequences V_n and using the integral test criteria for $\sum_{n=1}^{+\infty} U_n$ and the d'alembert rule for $\sum_{n=1}^{+\infty} Z_n$, where the general term is defined by

$$V_n = M\exp(-\sqrt{n}W\widetilde{\epsilon}B_1) + MB_2\sqrt{\frac{3}{2}}\exp(\frac{-n^2}{2}(\log(\frac{3}{2}) - \frac{1}{n}\widetilde{\epsilon})).$$

Its follows that

$$\sum_{n=1}^{+\infty}\mathbf{P}(\sqrt{n}|\theta_n - \theta| > W) < +\infty.$$

which yields to the result.

Remark 4. *The inequalities (3.10) give us the possibility to construct a confidence interval for the parameter θ of the first order autoregressive process.*
For large W, such as $W = \widetilde{\epsilon}\sqrt{n}$ its follows that

$$\lim_{n\to+\infty} V_n = \lim_{n\to+\infty}\left(M\exp(-n\widetilde{\epsilon}^2B_1) + MB_2\sqrt{\frac{3}{2}}\exp(\frac{-n^2}{2}(\log(\frac{3}{2}) - \frac{1}{n}\widetilde{\epsilon}))\right) = 0,$$

where $B_1 = \frac{1}{2CL}\text{arcsinh}\frac{Cn\widetilde{\epsilon}^2}{2LD_n} > 0$.
Which means, for a given level ω, we can found a natural integer n_ω such that

$$\forall n \geq n_\omega \Longrightarrow V_n \leq \omega.$$

Consequently

$$\mathbf{P}(|\theta_{n_\omega} - \theta| \leq \widetilde{\epsilon}) \geq 1 - \omega.$$

Which means that the parameter of the order 1 autoregressive process belongs to the inclusive interval of center θ_{n_ω} and radius $\widetilde{\epsilon}$ with a probability greater or equal to $1-\omega$.

3.4 Conclusion

Our paper consists in establishing new exponential inequalities of Bernstein Fréchet type for END that allowed us to construct a confidence interval for the parameter of the order 1 autoregressive process. Using these inequalities, we demonstrated the estimator complete converge to the unknown parameter θ.

Acknowledgment

The author is grateful to the editor and referee for carefully reading the manuscript and for providing detailed lists of remarks that considerably improved the presentation of the paper.

References

Asadian, N. and Fakoor, V. and Bozorgnia, A., 2006. Rosenthal's type inequalities for negatively orthant dependent random variables, *Jirss*, **5**: 69-75.

Belguerna, A. and Benaissa, S., 2014. The asymptotic convergence of the parameter estimator in the first order autoregressive model AR(1) under strong mixing errors, *I. J. Math. Stat*, **15**: 1-8.

Bentkus, V., 2003. An inequality for tail probabilities of martingales with differences bounded from one side, *J. Theoret. Probab*, **16**: 161-173.

Bondarev, B.V., 1991. On an astimator of the unknown drift coefficient parameter for an equation perturbed by Gaussian noise, *Theory probab. Math. Statist*, **42**: 1-12.

Chan, N.H. and Wei, C.Z., 1988. Limiting distribution of least squares estimates of unstable autoregressive process, *Ann. Statist*, **16**: 367-401.

Dahmani, A. and Tari, M., 2005. Bernstein-fréchet inequalities for the parameter of the first order autoregressive process, *Statistics. C. R. Acad. Sci. Paris*, **I340**: 309-314.

Fakoor, V. and Azarnoosh, H. 2005. Probability inequalities for sums of negatively dependent random variables,*Park J Stat*, **21**: 257–264.

Galtchouk, L. and Konev, V., 2003. On uniform asymptotic normality of sequantial least sequares estimators fot the parameters estimator in a stable AR(1), *Sequential Anal*, **22**: 31-54.

Guikhman, I. I. and Skorokhod, A. B., 1971. Theory of random process, *Naouka, Tome 1*.

Joag-Dev, K. and Proschan, F., 1983. Negative association of random variables with applications, *Ann. Statist*, **11**: 286-295.

Liu, L., 2009. Precise large deviations for dependent random variables with heavy tails, *Stat Probab Lett*, **79**: 1290-1298.

Liu, L.,2010. Necessary and sufficient conditions for moderate deviations of dependent random variables with heavy tails, *Sci China Ser A Math*, **53**: 1421-1434.

Pgokhorov, Yu, V., 1959. An extremal problem in probability theory, *Theory Prob. Applications*, **4**: 201-203.

Chapter 4

The Complete Convergence for the Parameter Estimator of the First-Order Autoregressive Process Created by WOD Errors

Abstract 2. The autoregressive process play an important role in predicting problems leading to decision making. In practise, to estimate the unknown parameter ϑ of the autoregressive model we use the least square method. We already saw that the least squares estimator ϑ_n complete converge to unknown parameter ϑ of the first-order autoregressive process generated by extended negatively dependent errors. In this paper, we examine the complete convergence of the estimator ϑ_n also under widely orthant dependent errors and we construct exponential inequalities of the coefficient of 1^{st} order autoregressive model which enable us to build a confidence interval.

4.1 Inroduction

Analysis regarding autoregressive model makes up one of fundamental problems carried by the statistical analysis of time series. Generally, the study of autoregressive models can facilitate the development of forecasting, establish controls and lead to the reduction of undesirable changes. The expansion of linear and nonlinear autoregressive models have investigated by many authors(Dahmani and Tari [16], Chebbab [13]).
(Chan and Wei [14]) evaluated the limit with law of the least squares estimator when the errors represent a sequence of martingale difference. The exponential inequalities has been used to construct a confidence interval for the unknown parameter ϑ_0 the first interesting results, about this subject, obtained by (Bondarev [9]) in the equation

$$\frac{dx}{dt} = \vartheta_0 f(t, x(t)) + \epsilon'(t), \qquad x(0) = \epsilon(0) = 0$$

where ϵ' is a Gaussian noise with zero mean and known correlation function.

In several statistical applications, the random variables are supposed to be independent. But which are often hypothetical is very unrealistic. Thus several statisticians extended this case to different dependence structures such as a weak dependence structure, i.e., WOD structure. The study of the limiting behavior of widely orthant dependent (WOD) random variables is of great importance, there are a lot of results checking the WOD random variables for example (Shen [35]) constructed the Bernstein-type inequality for widely dependent sequence and applied to nonparametric regression models, (Wang, Wu and Rosalsky [38]) obtained the complete convergence for arrays of rowwise widely orthant dependent random variables and their applications. In this work, we establish Bernstein-Fréchet type Inequality for the parameter of the autoregressive process of order 1 under widely orthant dependent errors. using these inequalities, we construct a confidence interval for this parameter and we prove that the estimator of the least squares complete convergence to the parameter of AR(1).

4.2 Optimal Control

The order 1 autoregressive process AR(1) defined by

$$X_j = \vartheta X_{j-1} + \zeta_j, \qquad \zeta_0 = X_0 = 0 \tag{4.1}$$

where $\{\zeta_j, j \geq 0\}$ is a sequence of identically distributed WOD random variables, with zero mean and finite variance, where ϑ is a parameter with $|\theta| < 1$.
By the least square method, we can get the estimator of ϑ, such that the estimator ϑ_n defined for $n \geq 1$ by

$$\vartheta_n = \frac{\sum_{j=1}^{n} X_{j-1} X_j}{\sum_{j=1}^{n} X_{j-1}^2}, \tag{4.2}$$

and

$$\vartheta_n - \vartheta = \frac{\sum_{j=1}^{n} X_{j-1} \zeta_j}{\sum_{j=1}^{n} X_{j-1}^2}. \tag{4.3}$$

We present now the definition of widely orthant dependent (WOD) sequences, and lemmas. The notion of WOD sequence was presented by (Wang, Wang and Gao [37]) as follows.

Definition 18. *The random variables $\zeta_1, \zeta_2, ..., \zeta_n$ are said to be*

(i) *widely upper orthant dependent (WUOD), if there is a sequence of positive numbers* $\{g_U(n), n \geq 1\}$*. Then for every* $n \geq 1$ *and for any* $\epsilon_j \in (-\infty, \infty)$*,* $1 \leq j \leq n$

$$\mathbf{P}(\zeta_1 > \epsilon_1, \zeta_2 > \epsilon_2, ..., \zeta_n > \epsilon_n) \leq g_U(n) \prod_{j=1}^{n} P(\zeta_j > \epsilon_j). \tag{4.4}$$

(ii) *widely lower orthant dependent (WLOD), if there is a sequence of positive numbers* $\{g_L(n), n \geq 1\}$*. Then, for every* $n \geq 1$ *and for any* $\epsilon_j \in (-\infty, \infty)$*,* $1 \leq j \leq n$

$$\mathbf{P}(\zeta_1 \leq \epsilon_1, \zeta_2 \leq \epsilon_2, ..., \zeta_n \leq \epsilon_n) \leq g_L(n) \prod_{j=1}^{n} P(\zeta_j \leq \epsilon_j). \tag{4.5}$$

(iii) *widely orthant dependent (WOD) if both (4.4) and (4.5) hold,*
where $g_U(n)$ *and* $g_L(n)$*,* $n \geq 1$*, these are called dominating coefficients and denote* $g(n) = \max\{g_U(n), g_L(n)\}$*.*

Remember that when $g_L(n) = g_U(n) = M$ *for a constant* $M > 0$*, then the random variables sequence* $\{\zeta_n, n \geq 1\}$ *are END (see, e.g., Liu [27]).*

When $g_L(n) = g_U(n) = 1$*, for all* $n \geq 1$*, then the random variables sequence* $\{\zeta_n, n \geq 1\}$ *are NOD (see, e.g., Joag-Dev and Proschan [25]; Lehmann [26]).*

Lemma 5. *(Wang,Wang and Gao[37])*

1. *Let random variables* $\zeta_1, \zeta_2, ..., \zeta_n$ *be WLOD(WUOD). If* $\phi_1, \phi_2, ..., \phi_n$ *are nondecreasing, then* $\phi_1(\zeta_1), \phi_1(\zeta_1), ..., \phi_n(\zeta_n)$ *are WLOD (WUOD).*
   If $\phi_1, \phi_2, ..., \phi_n$ *are nonincreasing, then* $\phi_1(\zeta_1), \phi_1(\zeta_1), ..., \phi_n(\zeta_n)$ *are WUOD (WLOD).*

2. *If random variables* $\zeta_1, \zeta_2, ..., \zeta_n$ *are nonnegative and WUOD, then for every* $n \geq 1$

$$\mathbf{E}\left[\prod_{j=1}^{n} \zeta_j\right] \leq g_U(n) \prod_{j=1}^{n} \mathbf{E}\left[\zeta_j\right]. \tag{4.6}$$

Particularly, if random variables $\zeta_1, \zeta_2, ..., \zeta_n$ *are WUOD, then for every* $n \geq 1$ *and any* $t > 0$

$$\mathbf{E}\left[\exp(t \sum_{j=1}^{n} \zeta_j)\right] \leq g_U(n) \prod_{j=1}^{n} \mathbf{E}\left[\exp(t\zeta_j)\right]. \tag{4.7}$$

By the Lemma 5., we can find the following Corollary immediately.

Corollary 2. *Let random variables $\{\zeta_n, n \geq 1\}$ an WOD sequence*

1. *Let random variables $\zeta_1, \zeta_2, ..., \zeta_n$ be WOD. If $\phi_1, \phi_2, ..., \phi_n$ are all nondecreasing (or nonincreasing) function, then $\phi_1(\zeta_1), \phi_2(\zeta_2), ..., \phi_n(\zeta_n)$ are WOD.*

2. *If $\{\zeta_n, n \geq 1\}$ are WOD random variable, then for each $n \geq 1$ and any $t \in \mathbb{R}$*

$$\mathbf{E}\left[\exp(t\sum_{j=1}^{n}\zeta_j)\right] \leq g(n)\prod_{j=1}^{n}\mathbf{E}\left[\exp(t\zeta_j)\right]. \tag{4.8}$$

4.3 Hypotheses

Let us present now the following hypotheses:

H1: We assume that the sequences $\{\zeta_n, n \geq 0\}$ and $\{X_{n-1}, n \geq 1\}$ are bounded. Then there exist a constants $R > 0$ and $K > 0$ such that $|X_{n-1}| \leq R, |\zeta_n| \leq K$.

H2: Assume that

$$\lim_{n\to+\infty} g(n)e^{-sn^r} = 0, \tag{4.9}$$

where s, r are positive finite constants.

4.4 Main Result

Theorem 10. *Under hypothese (H_1), for any positive $\widetilde{\varepsilon} < Q_2$ and for D rather large, we have for any $n > 0$*

$$\mathbf{P}(\sqrt{n}|\vartheta_n - \vartheta| > D) \leq 2g(n)\exp(-\sqrt{n}D\widetilde{\varepsilon}Q_1) + \exp(-\frac{n(\widetilde{\varepsilon} - Q_2)^2}{2Q_3}), \tag{4.10}$$

where $Q_1 = \dfrac{\sqrt{n}D\widetilde{\varepsilon}}{4R^2A + 2RK\sqrt{n}D\widetilde{\varepsilon}}$, $Q_2 = \mathbf{E}X_{j-1}^2 < \infty$, $Q_3 = \mathbf{E}X_{j-1}^4 < \infty$.

Proof of Theorem 10.

By the equality (4.3), we have

$$\begin{aligned}\mathbf{P}(\sqrt{n}|\vartheta_n - \vartheta| > D) &= \mathbf{P}\left(\left|\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_{j-1}\zeta_j}{\frac{1}{n}\sum_{j=1}^{n} X_{j-1}^2}\right| > D\right) \\ &\leq \mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_{j-1}\zeta_j| > \frac{D}{n}\sum_{j=1}^{n} X_{j-1}^2\right)\end{aligned}$$

According to the probability properties, we obtain that for any $\widetilde{\varepsilon} > 0$

$$\mathbf{P}(\sqrt{n}|\vartheta_n - \vartheta| > D) \leq \mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{\mathbf{j=1}}^{n} X_{j-1}\zeta_j| > D\widetilde{\varepsilon}\right) + \mathbf{P}\left(\frac{1}{n}\sum_{\mathbf{j=1}}^{n} X_{j-1}^2 \leq \widetilde{\varepsilon}\right). \tag{4.11}$$

We take
$I_1 = \mathbf{P}\left(|\frac{1}{\sqrt{n}}\sum_{\mathbf{j=1}}^{n} X_{j-1}\zeta_j| > D\widetilde{\varepsilon}\right)$ and $I_2 = \mathbf{P}\left(\frac{1}{n}\sum_{\mathbf{j=1}}^{n} X_{j-1}^2 \leq \widetilde{\varepsilon}\right)$.
For any $0 < z < \frac{1}{K}$ clearly, $|z\zeta_j| \leq 1$. Hence By Taylor's series, we have for $j = 1, ..., n$ and $l \geq 2$,

$$\begin{aligned}\exp(z\zeta_j) &= 1 + z\zeta_j + z^2\zeta_j^2 \sum_{l=2}^{+\infty} \frac{1}{l!}(z\zeta_j)^{l-2} \\ &\leq 1 + z\zeta_j + z^2\zeta_j^2 \sum_{l=2}^{+\infty} \frac{1}{l!} \\ &\leq 1 + z\zeta_j + z^2\zeta_j^2.\end{aligned} \tag{4.12}$$

Therefore, by the inequality $1 + x \leq e^x$ for $x \in \mathbb{R}$, we can write that

$$\mathbf{E}[\exp(z\zeta_j)] = 1 + z^2\mathbf{E}\zeta_j^2 \leq \exp(z^2\mathbf{E}\zeta_j^2). \tag{4.13}$$

By Chernoff inequality, Corollary 2., (4.13), we have

$$\begin{aligned}\mathbf{P}\left(\frac{1}{\sqrt{n}}\sum_{\mathbf{j=1}}^{n} X_{j-1}\zeta_j > D\widetilde{\varepsilon}\right) &\leq \mathbf{P}\left(\sum_{\mathbf{j=1}}^{n}\zeta_j > \frac{\sqrt{n}D\widetilde{\varepsilon}}{R}\right) \\ &\leq \exp(\frac{-z\sqrt{n}D\widetilde{\varepsilon}}{R})\mathbf{E}\left[\exp(z\sum_{j=1}^{n}\zeta_j)\right] \\ &\leq g(n)\exp(\frac{-z\sqrt{n}D\widetilde{\varepsilon}}{R})\prod_{j=1}^{n}\mathbf{E}\left[\exp(z\zeta_j)\right] \\ &\leq g(n)\exp(\frac{-z\sqrt{n}D\widetilde{\varepsilon}}{R} + z^2 A),\end{aligned} \tag{4.14}$$

where $A = \sum_{j=1}^{\infty}\mathbf{E}\zeta_{\mathbf{j}}^{\mathbf{2}} < \infty$.

Taking $z = \dfrac{\sqrt{n}D\widetilde{\varepsilon}}{2RA + K\sqrt{n}D\widetilde{\varepsilon}}$, then

$$\mathbf{P}\left(\frac{1}{\sqrt{n}}\sum_{\mathbf{j=1}}^{n}X_{j-1}\zeta_j > D\widetilde{\varepsilon}\right) \leq g(n)\exp(-\frac{(\sqrt{n}D\widetilde{\varepsilon})^2}{4R^2A+2RK\sqrt{n}D\widetilde{\varepsilon}}). \tag{4.15}$$

Now, we replace ζ_j by $-\zeta_j$ in (4.13), we obtain

$$\mathbf{P}\left(\frac{1}{\sqrt{n}}\sum_{\mathbf{j=1}}^{n}X_{j-1}\zeta_j < -D\widetilde{\varepsilon}\right) \leq \mathbf{P}\left(-\frac{1}{\sqrt{n}}\sum_{\mathbf{j=1}}^{n}X_{j-1}\zeta_j > D\widetilde{\varepsilon}\right) \leq g(n)\exp(-\frac{(\sqrt{n}D\widetilde{\varepsilon})^2}{4R^2A+2RK\sqrt{n}D\widetilde{\varepsilon}}). \tag{4.16}$$

By (4.15) and (4.16), we can get that

$$\begin{aligned}\mathbf{P}\left(|\tfrac{1}{\sqrt{n}}\sum_{j=1}^{n}X_{j-1}\zeta_j| > D\widetilde{\varepsilon}\right) &= \mathbf{P}\left(\tfrac{1}{\sqrt{n}}\sum_{j=1}^{n}X_{j-1}\zeta_j > D\widetilde{\varepsilon}\right) + \mathbf{P}\left(\tfrac{1}{\sqrt{n}}\sum_{j=1}^{n}X_{j-1}\zeta_j < -D\widetilde{\varepsilon}\right) \\ &\leq 2g(n)\exp(-\sqrt{n}D\widetilde{\varepsilon}Q_1),\end{aligned} \tag{4.17}$$

where $Q_1 = \dfrac{\sqrt{n}D\widetilde{\varepsilon}}{4R^2A+2RK\sqrt{n}D\widetilde{\varepsilon}}$ is a positive constant does not depend on n.

Now, we limite the probability I_2. Using the Markov Inequality we have for any $\omega > 0$ and for $j = 1, ..., n$

$$\begin{aligned} I_2 = \mathbf{P}\left(\tfrac{1}{n}\sum_{\mathbf{j=1}}^{n}X_{j-1}^2 \leq \widetilde{\varepsilon}\right) &\leq \exp(\omega n\widetilde{\varepsilon})\mathbf{E}\left[\exp(-\omega\sum_{j=1}^{n}X_{j-1}^2)\right] \\ &\leq \exp(\omega n\widetilde{\varepsilon})\prod_{j=1}^{n}\mathbf{E}\left[\exp(-\omega X_{j-1}^2)\right].\end{aligned}$$

We take the inequality

$$\exp(-x) \leq 1 - x + \frac{1}{2}x^2, \qquad x \geq 0. \tag{4.18}$$

To show this let the function

$$f(x) = \ln(1 - x + \frac{1}{2}x^2) + x,$$

we should prove that $f(x) \geq 0$ for every $x \geq 0$.
Take the derivative

$$\frac{\partial f(x)}{\partial x} = \frac{x^2}{2(1-x+\frac{1}{2}x^2)}.$$

Therefore, f is strictly increasing on $\mathbb{R}^+$.
By (4.18) and the inequality $1 + x \leq \exp(x)$ for $x \in R$, we can write that

$$\begin{aligned} \mathbf{E}\left[\exp(-\omega X_{j-1}^2)\right] &\leq 1 - \omega \mathbf{E} X_{j-1}^2 + \frac{\omega^2}{2}\mathbf{E} X_{j-1}^4 \\ &\leq \exp(-\omega \mathbf{E} X_{j-1}^2 + \frac{\omega^2}{2}\mathbf{E} X_{j-1}^4), \end{aligned} \tag{4.19}$$

where $Q_2 = \mathbf{E}X_{j-1}^2 < +\infty$, $Q_3 = \mathbf{E}X_{j-1}^4 < +\infty$, we have

$$\mathbf{P}\left(\frac{1}{n}\sum_{\mathbf{j=1}}^{n} X_{j-1}^2 \leq \widetilde{\varepsilon}\right) \leq \exp(\omega n\widetilde{\varepsilon} - n\omega Q_2 + \frac{n\omega^2}{2}Q_3).$$

Taking $\omega = \dfrac{-(\widetilde{\varepsilon} - Q_2)}{Q_3}$, then

$$\mathbf{P}\left(\frac{1}{n}\sum_{\mathbf{j=1}}^{n} X_{j-1}^2 \leq \widetilde{\varepsilon}\right) \leq \exp(-\frac{n(\widetilde{\varepsilon} - Q_2)^2}{2Q_3}). \tag{4.20}$$

The desired result (4.11) follows by (4.17) and (4.20) immediately.

Corollary 3. *The sequence $(\vartheta_n)_{n\in\mathbb{N}}$ defined in (4.2) converge complete to the parameter ϑ of the order 1 autoregressive process.*

Proof 7. *The complete convergence follows from (4.10). Indeed, using the convergence of serie and the equation (4.9) on the positive réel term sequences W_n where the general term is defined by*

$$W_n = 2g(n)\exp(-\sqrt{n}D\widetilde{\varepsilon}Q_1) + \exp(-\frac{n(\widetilde{\varepsilon} - Q_2)^2}{2Q_3}),$$

it follows that

$$\begin{aligned} \sum_{n=1}^{+\infty}\mathbf{P}(\sqrt{n}|\vartheta_n - \vartheta| > D) &\leq \sum_{n=1}^{+\infty} 2g(n)\exp(-\sqrt{n}D\widetilde{\varepsilon}Q_1) + \sum_{n=1}^{+\infty}\exp(-\frac{n(\widetilde{\varepsilon} - Q_2)^2}{2Q_3}) \\ &\leq C\sum_{n=1}^{+\infty}\exp(-\sqrt{n}\eta) + \sum_{n=1}^{+\infty}\exp(-\frac{n(\widetilde{\varepsilon} - Q_2)^2}{2Q_3}) < \infty. \end{aligned}$$

which gives the result. Here constant C is positive not depending on n and $\eta > 0$.

Remark 5. *the inequalities (4.10) allow us to build a confidence interval for the parameter ϑ_n of the order 1 autoregressive process. For large D, such that $D = \widetilde{\varepsilon}\sqrt{n}$ and the equation (4.9), we have*

$$\lim_{n\to+\infty}\left(2g(n)\exp(-n\widetilde{\varepsilon}^2 Q_1) + \exp(-\frac{n(\widetilde{\varepsilon} - Q_2)^2}{2Q_3})\right) = 0.$$

which means, for a given level α, we can find a naturel integer n_α such that

$$\forall n \geq n_\alpha \Longrightarrow W_n \leq \alpha.$$

Consequently,

$$\mathbf{P}(|\vartheta_{n_\alpha} - \vartheta| \leq \widetilde{\varepsilon}) \geq 1 - \alpha.$$

Meaning that the that the parameter ϑ of the 1^{st} order autoregressive process belongs to the inclusive interval of center ϑ_{n_α} and radius $\widetilde{\varepsilon}$ with a probability greater or equal to $1 - \alpha$.

References

Bondarev, B.V., 1991. On an astimator of the unknown drift coefficient parameter for an equation perturbed by Gaussian noise, *Theory probab. Math. Statist*, **42**: 1-12.

Chebbab, I., 2018. Probability Tail for Extended Negatively Dependent Random Variables of Partial Sums and Application to AR(1) Model Generated By END Errors, *I. J. Stat. Econ*, **19**, 1-9.

Chan, N.H. and Wei, C.Z., 1988. Limiting distribution of least squares estimates of unstable autoregressive process, *Ann. Statist*, **16**: 367-401.

Dahmani, A. and Rahmani, S., 2011. Exponential inequalities for a strongly mixing autoregressive process, *Kuwait J. Sci. Eng*, **38**: 1-19.

Joag-Dev, K. and Proschan, F., 1983. Negative association of random variables with applicationse, *Ann. Stat*, **11**: 286-295.

Lehmann, E.L., 1966. Some concepts of dependence, *Ann. Math. Stat*, **37**: 1137-1153.

Liu, L., 2009. Precise large deviations for dependent random variables with heavy tails, *Stat Probab Lett*, **79**: 1290-1298.

Shen, A., 2013. Bernstein-type inequality for widely dependent sequence and its application to nonparametric regression models, *Abstract and applied analysis*, **2013**: 9 pages.

Wang, K. and Wang, Y. and Gao, Q., 2013. Uniform asymptotic for the finite-time ruin probability of a new dependent risk model with a constant interest rate, *Methodology and Computing in Appl Prob*, **15**, 109-124.

Wang, X. and Wu, Y. and Rosalsky, A., 2017. Complete convergence for arrays of rowwise widely orthant dependent random variables and its applications, *Stochastics: An International journal of probability and stochastic processes*.

Chapter 5

Complete convergence and Maximal inequalities for product sums of WOD sequences

Abstract 3. in this paper, we first examine a complete convergence for product sums of widely orthant dependent (WOD) sequence, and we study the moment inequalities of Rosenthal-type, for the maximum of sums of products of WOD sequence of random variables. We can generalize the results we have obtained on some dependent random variables.

5.1 Introduction

The notion of complete convergence was presented by (Hsu and Robbins [24]) as follows. The sequence of random variables $(V_n)_{n\geq 1}$ converges completly to the constant ϑ. Then for all $\zeta > 0$,

$$\sum_{n=1}^{\infty} \mathbf{P}(|V_n - \vartheta| > \zeta) < \infty.$$

(Hsu, P. and Robbins[24]) has shown if the variance of the summands is finite, so the sequence of arithmetic means of random variables independent identically distributed converges completely to the mathematical expectation.

Suppose $(Y_j)_{j\geq 1}$ is a WOD sequence of random variables, denote

$$\widetilde{S_n} = \sum_{1\leq j_1<\ldots<j_m\leq n} \prod_{l=1}^{m} Y_{j_l}.$$

Many researchers have strived the product sums in the last years. In 1998, Gadidov studied the moment inequality for product sums of sequences of independent identically distributed random variables. (Qui and Chen [33]) obtained the complete convergence for product sums of extended negatively dependent sequence.
In this work, we will examine the complete convergence for the maximum of product sums and the moment inequality of rosenthal type for product sums, in the case widely orthant dependent.

5.2 Optimal Control

We will recall the definition of widely orthant dependent (WOD) sequences, and few lemmas.

Proposition 5. *(Gadidov [20])(Rosenthal-type inequality for sums of products).*
Let $q \geq 2$, m be a positive integer and $(Y_j)_{j\geq 1}$ a sequences of independent identically distributed symmetric random variables, Then

$$\mathbf{E}\left|\max_{m\leq k\leq n}|\widetilde{S_k}|\right|^q \leq \left(n^{q/2}(\mathbf{E}|Y|^2)^{q/2}+n\mathbf{E}|Y|^q\right)^m.$$

The notion of WOD sequence was given by (Wang, Wang and Gao [37]) as follows.

Definition 19. *The sequence $(Y_n)_{n\geq 1}$ of random variables are said widely upper orthant dependent (WUOD), if there is a sequence $\{g_U(n), n\geq 1\}$ of finite real numbers where for each $n\geq 1$ and for any $y_j \in (-\infty,\infty)$, $1\leq j\leq n$*

$$\mathbf{P}(Y_1>y_1, Y_2>y_2, ..., Y_n>y_n) \leq g_U(n)\prod_{j=1}^n P(Y_j>y_j).$$

The sequence $(Y_n)_{n\geq 1}$ of random variables are said widely lower orthant dependent (WLOD), if there is a sequence $\{g_L(n), n\geq 1\}$ of finite real numbers where for each $n\geq 1$ and for any $y_j\in(-\infty,\infty)$, $1\leq j\leq n$

$$\mathbf{P}(Y_1\leq y_1, Y_2\leq y_2, ..., Y_n\leq y_n) \leq g_L(n)\prod_{j=1}^n P(Y_j\leq y_j).$$

The sequence $(Y_n)_{n\geq 1}$ of random variables are said widely orthant dependent (WOD) if $(Y_n)_{n\geq 1}$ are both WUOD and WLOD, with $g_U(n)$ and $g_L(n)$ for any $n\geq 1$, these are known as dominating coefficients.
We call random variables $\{Y_{nj}: j=1...,n; n\geq 1\}$ array of rowwise WOD, if for each $n\geq 1$, the sequence $\{Y_{nj}: i=1,...,n\}$ be WOD random variables.

Remember that if $g_L(n)=g_U(n)=M$ for a positive constant M, then the random variables $(Y_n)_{n\geq 1}$ are END which were presented by (Liu [27]). There are plenty of research paper and textbooks have investigated the END random variables(Qui and Chen [33]; Wu, Wang, Hu and Volodin[43]; Chebbab [13]).

If $g_L(n)=g_U(n)=1$, for all $n\geq 1$, then the random variables $(Y_n)_{n\geq 1}$ are NOD which were presented by (Joag-Dev and Proschan [25]; Lehmann [26]). More research on NOD random variables, please see (Qui, Wu and Chen [32]; Wang and Hu[39]).

Lemma 6. *(Wang, Xu, Hu, Volodin and Hu [40]). Let random variables $Y_1, Y_2, ..., Y_n$ be WOD. If $g_1, g_2, ..., g_n$ are all non decreasing (or non increasing) function.*
Then random variables $g_1(Y_1), g_2(Y_2), ..., g_n(Y_n)$ are WOD.

Lemma 7. *(Wang, Xu, Hu, Volodin and Hu [40]). Let random variables $(Y_n)_{n\geq 1}$ be an WOD sequence. Then as $t \geq 2$, there exist constants $C_1(t) > 0$ and $C_2(t) > 0$ only depends on t. Moreover, suppose that $\mathbf{E}(Y_n) = 0$ for each $n \geq 1$, then for all $n \geq 1$ we have*

$$\mathbf{E}\left|\sum_{j=1}^{n} Y_j\right|^t \leq C_1(t)\sum_{j=1}^{n}\mathbf{E}|Y_j|^t + C_2(t)g(n)\left(\sum_{j=1}^{n}\mathbf{E}|Y_j|^2\right)^{t/2}.$$

According to the Lemma 7. and using the same argument as the Theorem 2.3.1 in (Stout [36]), The Lemma 8. holds.

Lemma 8. *Let random variables $(Y_n)_{n\geq 1}$ be an WOD sequence. Then as $t \geq 2$, there exist constants $C_1(t) > 0$ and $C_2(t) > 0$ only depends on t. Moreover, suppose that $\mathbf{E}(Y_n) = 0$ for each $n \geq 1$, then for all $n \geq 1$ we have*

$$\mathbf{E}\max_{1\leq k\leq n}\left|\sum_{j=1}^{k} Y_j\right|^t \leq \left(\frac{log(4n)}{log 2}\right)^t\left[C_1(t)\sum_{j=1}^{n}\mathbf{E}|Y_j|^t + C_2(t)g(n)\left(\sum_{j=1}^{n}\mathbf{E}|Y_j|^2\right)^{t/2}\right].$$

Lemma 9. *(Wang, Yan, Cheng and Cai [41]) The sequence $(Y_n)_{n\geq 1}$ be real numbers and m, n are positive integers such as $1 \leq m \leq n$. Then*

$$\sum_{1\leq i_1<i_2<...<i_m\leq n}\prod_{l=1}^{m} Y_{i_l} = \sum_{\sum_{k=1}^{m} s_l t_l = m} A(m, s_l, t_l : l = 1, ..., m)\prod_{l=1}^{m}\left(\sum_{j=1}^{n} Y_j^{s_l}\right)^{t_l},$$

where $A(m, s_l, t_l : l = 1, ..., m)$ are constants , s_l, t_l are positive integers depends only on m.

Lemma 10. *(Qui and Chen [33]) Let Z, W be random variables. there exist two constants a and b such that*
$a + b = 1$. Then for any $\epsilon > 0$

$$(|Z + W| > \epsilon) \subseteq (|Z| > a\epsilon) \cup (|W| > b\epsilon), \quad (|ZW| > \epsilon) \subseteq (|Z| > \epsilon^a) \cup (|W| > \epsilon^b).$$

Throught the paper, the symbol $\sharp B$ signifies the number of element in the set B.

5.3 Main Result

Theorem 11. *Let $(Y_j)_{j\geq 1}$ is a WOD sequence of random variables. $\alpha > 1/2$ and $\alpha p > 1$, m be positive integers. Moreover, additionally suppose that $\mathbf{E}(Y_n) = 0$ for each $n \geq 1$, if $\mathbf{E}|Y|^p < \infty$. Then for all $\zeta > 0$*

$$\sum_{n=m}^{\infty} n^{\alpha p-2}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > n^{m\alpha}\zeta\right) < \infty. \tag{5.1}$$

$$\sum_{n=m}^{\infty} n^{\alpha p-2}\mathbf{P}\left(\sup_{k\geq n} k^{-m\alpha}|\widetilde{S_k}| > \zeta\right) < \infty. \tag{5.2}$$

Theorem 12. *Let* $(Y_j)_{j\geq 1}$ *is a WOD sequence of random variables.* $q > 0$ *and* $t \geq 2$*, m, u be positive integers, there is a constant* $C > 0$*. Suppose that* $\mathbf{E}(Y_n) = 0$ *for each* $n \geq 1$*.*
if $u = 1$ *Then for all* $n \geq m$

$$\mathbf{E}\left(\max_{m\leq k\leq n}|\widetilde{S_k}|\right)_+^q \leq C\left(\frac{log(4n)}{\log 2}\right)^t n^{mq-t}\left(C_1(t)\sum_{j=1}^n \mathbf{E}|Y_j|^t + C_2(t)g(n)(\sum_{j=1}^n \mathbf{E}|Y_j|^2)^{t/2}\right),$$

as $1 < u \leq m$

$$\mathbf{E}\left(\max_{m\leq k\leq n}|\widetilde{S_k}|\right)_+^q \leq Cn^{mq-ut}\left(C_1(t)\sum_{j}^n \mathbf{E}|Y_j|^{ut} + C_2(t)g(n)(\sum_{j=1}^n \mathbf{E}|Y_j|^{2u})^{t/2}\right),$$

where $C_1(t)$ *and* $C_2(t)$ *positive constants only depends on t.*

Proof of Theorem 11.
We prove (5.1). According to the **Lemma 9.**, **Lemma 10.** and the Jensen inequality, in order to show (5.1), it is enought to prove that for any $\zeta > 0$

$$\sum_{n=m}^{\infty} n^{\alpha p-2}\mathbf{P}\left(\max_{m\leq k\leq n}|\sum_{i=1}^k Y_i| > n^{\alpha}\zeta\right) < \infty, \tag{5.3}$$

and

$$\sum_{n=m}^{\infty} n^{\alpha p-2}\mathbf{P}\left(\max_{m\leq k\leq n}|\sum_{i=1}^k Y_i^2| > n^{2\alpha}\zeta\right) < \infty, \tag{5.4}$$

We can prove (5.3), by using the same method of proof as in Theorem 2.1 in (Qui et al.,[32]).

To prove (5.4), We take $Y_n^2 = Y_n^2 I(Y_n < 0) + Y_n^2 I(Y_n \geq 0)$, thus, we suppose that $Y_n \geq 0$ for each $n \geq 1$. De note
$(Y_n^2)_{n\geq 1} \prec Y^2$,
$\mathbf{E}(Y^2)^{p/2} = \mathbf{E}|Y|^p < \infty$, and $\alpha > 1/2$.

Hence, according to the **Lemma 6.** and (5.3), we obtain (5.4). So (5.1) holds.

Now, we will prove (5.2), for all m positive integer, there existe a positive integer j_0 such that $2^{j_0-1} \leq m \leq 2^{j_0}$. According to (5.1) and $\alpha > 1/p$, we have

$$\begin{aligned}
&\sum_{n=2^{j_0}}^{\infty} n^{\alpha p-2}\mathbf{P}\left(\sup_{k\geq n} k^{\frac{-m}{p}}|\widetilde{S_k}| > \zeta\right) \\
&= \sum_{j=j_0}^{\infty}\sum_{n=2^j}^{2^{j+1}-1} n^{\alpha p-2}\mathbf{P}\left(\sup_{k\geq n} k^{\frac{-m}{p}}|\widetilde{S_k}| > \zeta\right) \\
&\leq C\sum_{j=j_0}^{\infty}\mathbf{P}\left(\sup_{k\geq 2^j} k^{\frac{-m}{p}}|\widetilde{S_k}| > \zeta\right)\sum_{n=2^j}^{2^{j+1}-1} 2^{j(\alpha p-2)} \\
&= C\sum_{j=j_0}^{\infty} 2^{j(\alpha p-1)}\mathbf{P}\left(\sup_{v\geq j}\max_{2^v\leq k\leq 2^{v+1}} k^{\frac{-m}{p}}|\widetilde{S_k}| > \zeta\right) \\
&\leq C\sum_{j=j_0}^{\infty} 2^{j(\alpha p-1)}\sum_{v=j}^{\infty}\mathbf{P}\left(\max_{2^v\leq k\leq 2^{v+1}} k^{\frac{-m}{p}}|\widetilde{S_k}| > \zeta\right) \\
&\leq C\sum_{v=j_0}^{\infty} 2^{v(\alpha p-1)}\mathbf{P}\left(\max_{m\leq k\leq 2^{v+1}}|\widetilde{S_k}| > 2^{\frac{m}{p}v}\zeta\right) \\
&\leq C\sum_{v=j_0}^{\infty}\sum_{n=2^v}^{2^{v+1}-1} n^{\alpha p-2}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > 2^{\frac{m}{p}(v+1)}2^{\frac{-m}{p}}\zeta\right) \\
&\leq C\sum_{n=m}^{\infty} n^{\alpha p-2}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > n^{\frac{m}{p}}\zeta_0\right) < \infty. \qquad (where\ \ \zeta_0 = 2^{\frac{-m}{p}}\zeta)
\end{aligned}$$

the proof is completed.

Proof of Theorem 12.

$$\begin{aligned}
\mathbf{E}\left(\max_{m\leq k\leq n}|\widetilde{S_k}|\right)_+^q &= \int_0^{+\infty}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > r^{\frac{1}{q}}\right)dr \\
&= \int_0^{n^{mq}}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > r^{\frac{1}{q}}\right)dr + \int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > r^{\frac{1}{q}}\right)dr \\
&= n^{mq}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > n^m\right) + \int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\leq k\leq n}|\widetilde{S_k}| > r^{\frac{1}{q}}\right)dr.
\end{aligned}$$

According to the **Lemma 9.** and **Lemma 10.**, we have

$$\begin{aligned}
\mathbf{E}\left(\max_{m\le k\le n}|\widetilde{S_k}|\right)_+^q &\le n^{mq}\sum_{\sum_{l=1}^m s_lt_l=m}\mathbf{P}\left(\max_{m\le k\le n}|\prod_{l=1}^m\sum_{j=1}^k (Y_j^{s_l})^{t_l}|>C_m^{-1}n^m\right)\\
&+\sum_{\sum_{l=1}^m s_lt_l=m}\int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\le k\le n}|\prod_{l=1}^m\sum_{j=1}^k (Y_j^{s_l})^{t_l}|>C_m^{-1}r^{\frac{1}{q}}\right)dr\\
&\le n^{mq}\sum_{\sum_{l=1}^m s_lt_l=m}\sum_{l=1}^m\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j^{s_l}|^{t_l}>C_m^{-\frac{s_lt_l}{m}}n^{s_lt_l}\right)\\
&+\sum_{\sum_{l=1}^m s_lt_l=m}\sum_{l=1}^m\int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j^{s_l}|^{t_l}>C_m^{-\frac{s_lt_l}{m}}r^{\frac{s_lt_l}{mq}}\right)dr\\
&= n^{mq}\sum_{\sum_{l=1}^m s_lt_l=m}\sum_{l=1}^m\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j^{s_l}|>C_m^{-\frac{s_l}{m}}n^{s_l}\right)\\
&+\sum_{\sum_{l=1}^m s_lt_l=m}\sum_{l=1}^m\int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j^{s_l}|>C_m^{-\frac{s_l}{m}}r^{\frac{s_l}{mq}}\right)dr,
\end{aligned}$$

where $C_m = dA(m, s_l, t_l : l = 1, ..., m)$ constants, which only depends on m, such that

$d = \sharp\{s_l, t_l : l = 1, ..., m : \sum_{l=1}^m r_l s_l = m\}$ is a constant depend only on m. Therefore, for any integers $n \ge m$ positive and $1 \le u \le m$,we have

$$\begin{aligned}
\mathbf{E}\left(\max_{m\le k\le n}|\widetilde{S_k}|\right)_+^q &\le n^{mq}\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j^u|>\xi n^u\right)\\
&+\int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j^u|>\xi r^{\frac{u}{mq}}\right)dr \qquad \forall\xi>0.
\end{aligned}$$

We will consider two cases :
case 1: $u = 1$.

$$\begin{aligned}
\mathbf{E}\left(\max_{m\le k\le n}|\widetilde{S_k}|\right)_+^q &\le n^{mq}\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j|>\xi n\right)\\
&+\int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\le k\le n}|\sum_{j=1}^k Y_j|>\xi r^{\frac{1}{mq}}\right)dr = I_1.
\end{aligned}$$

We take t such that $t \ge 2$, according to the markov inequality, **Lemma 6.**, **Lemma 8.**, and the jensen inequality, we have

$$\begin{aligned}
I_1 &\le \left(\frac{log(4n)}{\log 2}\right)^t n^{mq-t}\left(C_1(t)\sum_{j=1}^n\mathbf{E}|Y_j|^t + C_2(t)g(n)(\sum_{j=1}^n\mathbf{E}|Y_j|^2)^{t/2}\right)\\
&+\left(\frac{log(4n)}{\log 2}\right)^t\left(C_1(t)\sum_{j=1}^n\mathbf{E}|Y_j|^t + C_2(t)g(n)(\sum_{j=1}^n\mathbf{E}|Y_j|^2)^{t/2}\right)\int_{n^{mq}}^{+\infty}r^{-\frac{t}{mq}}dr\\
&\le C\left(\frac{log(4n)}{\log 2}\right)^t n^{mq-t}\left(C_1(t)\sum_{j=1}^n\mathbf{E}|Y_j|^t + C_2(t)g(n)(\sum_{j=1}^n\mathbf{E}|Y_j|^2)^{t/2}\right).
\end{aligned}$$

case 2: $1 < u \leq m$.

$$\begin{aligned}\mathbf{E}\left(\max_{m\leq k\leq n}|\widetilde{S_k}|\right)_+^q &\leq n^{mq}\mathbf{P}\left(\max_{m\leq k\leq n}|\sum_{j=1}^{k}Y_j^u| > \xi n^u\right)\\ &+ \int_{n^{mq}}^{+\infty}\mathbf{P}\left(\max_{m\leq k\leq n}|\sum_{j=1}^{k}Y_j^u| > \xi r^{\frac{u}{mq}}\right)dr = I_2.\end{aligned}$$

We have then (Qui and Chen [33])

$$I_2 \leq n^{mq}\mathbf{P}\left(\sum_{j=1}^{k}|Y_j|^u > \xi n^u\right) + \int_{n^{mq}}^{+\infty}\mathbf{P}\left(\sum_{j=1}^{k}|Y_j|^u > \xi r^{\frac{u}{mq}}\right)dr.$$

We take t such that $t \geq 2$, using the Markov inequality, **Lemma 6.**, **Lemma 7.**, and the jensen inequality, we have

$$\begin{aligned}I_2 &\leq n^{mq-ut}\left(C_1(t)\sum_{j=1}^{n}\mathbf{E}|Y_j|^{ut} + C_2(t)g(n)(\sum_{j=1}^{n}\mathbf{E}|Y_j|^{2u})^{t/2}\right)\\ &+ \left(C_1(t)\sum_{j=1}^{n}\mathbf{E}|Y_j|^{ut} + C_2(t)g(n)(\sum_{j=1}^{n}\mathbf{E}|Y_j|^{2u})^{t/2}\right)\int_{n^{mq}}^{+\infty}r^{-\frac{ut}{mq}}dr\\ &\leq Cn^{mq-ut}\left(C_1(t)\sum_{j=1}^{n}\mathbf{E}|Y_j|^{ut} + C_2(t)g(n)(\sum_{j=1}^{n}\mathbf{E}|Y_j|^{2u})^{t/2}\right).\end{aligned}$$

Then the proof is completed.

Acknowledgment

The authors wish to thank the Editor-in-Chief Kaushal K. Srivastava and the referees for their very valuable remarks which helped to improve an earlier version of the manuscript.

References

Chebbab, I., 2018. Probability Tail for Extended Negatively Dependent Random Variables of Partial Sums and Application to AR(1) Model Generated By END Errors, *I. J. Stat. Econ*, **19**: 1-9.

Hsu, P. and Robbins, H., 1947. Complete convergence and the law of large numbers, *Proc. Natl. Acad. Sci. USA*, **33**: 25-31.

Qui, D. and Wu, Q. and Chen, P., 2014. Complete convergence for negatively orthant dependent random variables, *J. Inequal. Appl*, **2014**: 145.

Qui, D. and Chen, P., 2015. Complete moment convergence for product sums of sequence of extended negatively dependent random variables, *J. Inequal. Appl*, **2015**: 212.

Gadidov, A., 1998. Strong law of large numbers for multilinear forms, *Ann. Probab*, **26(2)**: 902-923.

Lehmann, E.L., 1966. Some concepts of dependence, *Ann. Math. Stat*, **37**: 1137–1153.

Joag-Dev, K. and Proschan, F., 1983. Negative association of random variables with applicationse, *Ann. Stat*, **11**:286-295.

Liu, L., 2009. Precise large deviations for dependent random variables with heavy tails, *Stat Probab Lett*, **79**: 1290-1298.

Stout, WF., 1947. Almost Sure Convergence, *Academic Press, New York*.

Wang, K. and Wang, Y. and Gao, Q., 2013. Uniform asymptotic for the finite-time ruin probability of a new dependent risk model with a constant interest rate, *Methodology and Computing in Appl Prob*, **15**: 109-124.

Wang, X.J. and Xu, C. and Hu, T-Ch. and Volodin, A. and Hu, S.H., 2014. On complete convergence for widely orthant dependent random variables and its applications in nonparametric regression models, *Test*, **23**: 607-629.

Wang, X. and Hu, Sh., 2017. On consistency of least square estimators in the simple linear EV model with negatively orthant dependent errors, *E. J. Stat*, **11**:1434–1463.

Wang, YB. and Yan, J. and Cheng, FY. and Cai, XZ., 2011. On the strong stability for Jamison type weighted product sums of pairwise NQD series with different distribution, *Chin. Ann. Math*, **22**: 701-706.

Wu, Y. and Wang, X. and Hu, T-Ch. and Volodin, A, 2017. Complete f-moment convergence for extended negatively dependent random variables, *RACSAM*.

Conclusion and Perspectives

Conclusion

We are interested in this thesis to establish an exponential concentration inequality and the complete convergence of partial sums of random variables with application to the AR (1) model generated by the errors in the dependent cases (Extended negatively dependent, Widely orthant dependent). Then we study the complete convergence and the maximum inequality for product sums of Widely orthant dependent sequences.

Perspectives

In this section, we draw some perspectives for potential future researches.

For Chapter 3

1. See what are the conditions to get a similar result of an autoregressive process of order p $(p > 1)$, the same for ARH(1) and ARB(1).

2. Study the case of a process with $\varphi-$ mixing functional variables.

For Chapter 4

1. Study the cases of ARMA and GARCH models.

For Chapter 5

1. It is possible to study the complete moment convergence for product sums of Negatively superadditive dependence sequences, the same for widely orthant dependent.

Bibliography

[1] Asadian, N. and Fakoor, V. and Bozorgnia, A., 2006. Rosenthal's type inequalities for negatively orthant dependent random variables, *Jirss*, **5**: 69-75.

[2] Baraud, Y., (2010): A Bernstein-type inequality for suprema of random processes with applications to model selection in non-Gaussian regression, *Bernoulli*, **16(4)**: 1064–1085.

[3] Belguerna, A. and Benaissa, S., 2014. The asymptotic convergence of the parameter estimator in the first order autoregressive model AR(1) under strong mixing errors, *I. J. Math. Stat*, **15**: 1-8.

[4] Bennett, G. (1962): Probability inequalities for the sum of independent random variables, Journal of the American Statistical Association, **57(297)**, 33–45.

[5] Bentkus, V., 2003. An inequality for tail probabilities of martingales with differences bounded from one side, *J. Theoret. Probab*, **16**: 161-173.

[6] Bercu, B., Delyon, B. and Rio,E., 2015. Concentration Inequalities for Sums and Martingales, *Springer Briefs in Math*.

[7] Bernstein, S. N. (1946). Theory of Probability, *Gastehizdat Publishing House Moscow*.

[8] Bollerslev, T., 1986. Generalized Autoregressive Conditional Heterosckedasticity. *J. of Econometrics*: 307-327.

[9] Bondarev, B.V., 1991. On an astimator of the unknown drift coefficient parameter for an equation perturbed by Gaussian noise, *Theory probab. Math. Statist*, **42**: 1-12.

[10] Boucher, T. R., (2009): A Hoeffding inequality for Markov chains using a generalized inverse. *Stat Probab Lett*, **79**: 1105–1107.

[11] Brockwell P.J and Davis R.A., (1991). Time series: theory and methods. *Springer Series in Statistics. Springer Verlag, New York, Second Edition*.

[12] Box, G., Jenkins, G., 1970. Time series analysis. *Forecasting and Daviscontrol*. Holden Day.

[13] Chebbab, I., 2018. Probability Tail for Extended Negatively Dependent Random Variables of Partial Sums and Application to AR(1) Model Generated By END Errors, *I. J. Stat. Econ*, **19**: 1-9.

[14] Chan, N.H. and Wei, C.Z., 1988. Limiting distribution of least squares estimates of unstable autoregressive process, *Ann. Statist*, **16**: 367-401.

[15] Guikhman, I. I. and Skorokhod, A. B., 1971. Theory of random process, *Naouka, Tome 1*.

[16] Dahmani, A. and Rahmani, S., 2011. Exponential inequalities for a strongly mixing autoregressive process, *Kuwait J. Sci. Eng*, **38**: 1-19.

[17] Engle, B., 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Ináation. *Econometrica*: 987-1008.

[18] Fakoor, V. and Azarnoosh, H. 2005. Probability inequalities for sums of negatively dependent random variables,*Park J Stat*, **21**: 257–264.

[19] Fan, X. Q., Grama, I. and Liu, Q. S., 2015a. Sharp large deviation results for sums of independent random variables. *Science China Mathematics*, **58 (9)**: 1939-1958.

[20] Gadidov, A., 1998. Strong law of large numbers for multilinear forms, *Ann. Probab*, **26(2)**: 902-923.

[21] Galtchouk, L. and Konev, V., 2003. On uniform asymptotic normality of sequantial least sequares estimators fot the parameters estimator in a stable AR(1), *Sequential Anal*, **22**: 31-54.

[22] Gao, F., Guillin, A., and Wu, L., (2014): Bernstein-type concentration inequalities for symmetric Markov processes, *Theory of Probab and Its Appl*, **58(3)**: 358–382.

[23] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc*. *58*: 13–30.

[24] Hsu, P. and Robbins, H., 1947. Complete convergence and the law of large numbers, *Proc. Natl. Acad. Sci. USA*, **33**: 25-31.

[25] Joag-Dev, K. and Proschan, F., 1983. Negative association of random variables with applications, *Ann. Statist*, **11**: 286-295.

[26] Lehmann, E.L., 1966. Some concepts of dependence, *Ann. Math. Stat*, **37**: 1137-1153.

[27] Liu, L., 2009. Precise large deviations for dependent random variables with heavy tails, *Stat Probab Lett*, **79**: 1290-1298.

[28] Liu, L.,2010. Necessary and sufficient conditions for moderate deviations of dependent random variables with heavy tails, *Sci China Ser A Math*, **53**: 1421-1434.

[29] Miasojedow, B., (2014). Hoeffding's inequalities for geometrically ergodic Markov chains on general state space, *Stat Proba Lett*, **87**: 115–120.

[30] Pgokhorov, Yu, V., 1959. An extremal problem in probability theory, *Theory Prob. Applications*, **4**: 201-203.

[31] Pinelis, I. On the Bennett-Hoeffding inequality, *Ann. Inst. H. Poincaré Probab. Statist*, **50(1)**: 15-27.

[32] Qui, D. and Wu, Q. and Chen, P., 2014. Complete convergence for negatively orthant dependent random variables, *J. Inequal. Appl*, **2014**: 145.

[33] Qui, D. and Chen, P., 2015. Complete moment convergence for product sums of sequence of extended negatively dependent random variables, *J. Inequal. Appl*, **2015**: 212.

[34] Slutsky E. (1937). The Summation of Random Causes as the Source of Cyclical Processes, *Econometrica*, **5**, 105-146.

[35] Shen, A., 2013. Bernstein-type inequality for widely dependent sequence and its application to nonparametric regression models, *Abstract and applied analysis*, **2013**: 9 pages.

[36] Stout, WF., 1974. Almost Sure Convergence, *Academic Press, New York*.

[37] Wang, K. and Wang, Y. and Gao, Q., 2013. Uniform asymptotic for the finite-time ruin probability of a new dependent risk model with a constant interest rate, *Methodology and Computing in Appl Prob*, **15**, 109-124.

[38] Wang, X. and Wu, Y. and Rosalsky, A., 2017. Complete convergence for arrays of rowwise widely orthant dependent random variables and its applications, *Stochastics: An International journal of probability and stochastic processes*.

[39] Wang, X. and Hu, Sh., 2017. On consistency of least square estimators in the simple linear EV model with negatively orthant dependent errors, *E. J. Stat*, **11**:1434–1463.

[40] Wang, X.J. and Xu, C. and Hu, T-Ch. and Volodin, A. and Hu, S.H., 2014. On complete convergence for widely orthant dependent random variables and its applications in nonparametric regression models, *Test*, **23**: 607-629.

[41] Wang, YB. and Yan, J. and Cheng, FY. and Cai, XZ., 2011. On the strong stability for Jamison type weighted product sums of pairwise NQD series with different distribution, *Chin. Ann. Math*, **22**: 701-706.

[42] Wold, H., 1938. A study in the Analysis of Stationary Time Series. *Almgrist and Wiksell*.

[43] Wu, Y. and Wang, X. and Hu, T-Ch. and Volodin, A, 2017. Complete f-moment convergence for extended negatively dependent random variables, *RACSAM*.

[44] Yule G.U., (1927). On a Method of Investigating Periodicities in Disturbed Series with Special Reference to Wolfer's Sunspot Numbers, *Philosophical Transactions of the Royal Society, Series A*, **226**: 267-298.

[45] Yao, L. and Jiang, W., (2012): On extensions of Hoeffding's inequality for panel data. *Stat Proba Lett*, **82**: 446–454.

[46] Yves A., (2011). Series temporelles avec R [electronic resource] : Methodes et cas, *Springer*.

Summary

In this thesis, we are primarily interested in studying the inequality of concentration (exponential type) of the sequence of random variables with application to the autoregressive model. We have identified the main objective, that the study of the complete convergence of the estimator of first-order autoregressive process in the case where the error is dependent (Extended negatively dependent (END), widely orthant dependent (WOD)).

Next, we examine the complete convergence and maximal inequalities for product sums of widely orthant dependent (WOD) sequences, directly resulting from the works of Dehua Qiu, Pingyan Chen.

الملخص

في هذه الأطروحة ، نحن مهتمون بشكل أساسي بدراسة عدم المساواة في التركيز (النوع الأسّي) لتسلسل المتغيرات العشوائية مع تطبيقها على نموذج الانحدار الذاتي. هدفنا الرئيسي هو دراسة التقارب الكامل للمقدر عملية الانحدار الذاتي في حالة وجود أخطاء مرتبطة (المرتبطة والممتدة بشكل سلبي، المرتبطة على نطاق واسع).

ثم ندرس التقارب الكامل وأوجه التفاوت القصوى لمجموع المنتجات من السلاسل مرتبطة على نطاق واسع، المستمدة مباشرة من أعمال ده هوا تشيو، بينغيان تشن.

Résumé

Dans cette thèse, nous intéressons principalement à l’étude de l’inégalité de concentration (type exponentiel) d'une suite de variables aléatoires avec application au modèle autorégressif. Nous fixons comme objectif principal, l’étude de la convergence complète de l'estimateur d’un processus autorégressif du premier ordre dans le cas où l'erreur est dépendante (Étendue négativement dépendant, Largement orthant dépendant).

Ensuite, nous examinons la convergence complète et l'inégalités maximales des sommes des produits de WOD, directement issu des travaux de Dehua Qiu, Pingyan Chen.