REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES
SIDI BEL-ABBÈS

BP 89 SBA 22000 –ALGERIE- TEL/FAX 048-54-43-44

THESE

Présentée par:

BOULENOIR ZOUAOUIA

Pour obtenir le Diplôme de Doctorat

Spécialité : Mathématiques
Option : Probabilités-Statistique

Intitulée

Convergence presque complète des suites de variables aléatoires dépendantes, Applications aux modèles autorégressifs non linéaires.

Thèse soutenue le 17- 09- 2018
Devant le jury composé de :

Président : *Mr MECHAB BOUBAKER* *M.C.A à U.D.L- S.B.A*

Directeur de thèse : *Mr BENAISSA SAMIR Professeur à M.C.A à U.D.L- S.B.A*

Examinateurs :
Mr KANDOUCI Abdeldjebar *Professeur à L'Université de Saida*

Mr BELGUERA ABDERRAHMANE *M.C.A à L'Université de Naima*

Acknowledgements

First of all I would like to thank Mr. Boubaker Mechab who did me the honor to chair the jury of my defense and to judge my work.

I give my sincere thanks to the members of the jury : Mr. Abdeldjebar Kandouci and Mr. Abderrahmane Belguerna who have accepted to examine this work.

I thank my thesis supervisor Mr. Samir Benaissa for proposing this subject to me and for guiding and counselling me during the making of this thesis. I am grateful for his patience during all these years.

I thank everyone who has contributed directly or indirectly to the accomplishment of this thesis.

I would also like to express my deep gratitude for my family : my dear parents who were always here for me, my sisters, my brothers and their families.

Table des matières

Table des figures

General Introduction

The general framework presented in this thesis is the estimation of the parameters of an autoregressive process. We divided our whole work into four chapters.

The first chapter explains the basic notions and highlights some of the objectives of time series analysis, it also recalls the estimation of the auto-covariance function, and then it discusses autoregressive moving average processes ($ARMA$, in short) along with autoregressive processes in Hilbert spaces. Finally, a framework has been provided that may potentially be useful when facing the problem of analyzing a data set in practice.

In the second chapter, a new concentration inequality and complete convergence of weighted sums for arrays of row-wise linearly negative quadrant dependent ($LNQD$, in short) random variables has been established, we also obtained a result dealing with complete convergence of first-order autoregressive processes with identically distributed $LNQD$ innovations which has been introduced in the article entitled "Tail probabilities and complete convergence for weighted sequences of $LNQD$ random variables with application to first-order autoregressive processes model" (8).

In the third chapter we demonstrate almost complete convergence of dependant random variables sequences with application to non-linear autoregressive processes model which is defined by

$$X_i = g_\theta(X_{i-1}, \cdots, X_{i-p}) + \zeta_i, \quad i \geq 1,$$

where $\zeta = (\zeta_t,\ t \in Z)$ is an extended negatively dependent error(END, in short).

The last chapter is concerned with the almost complete convergence of the value of the process of autoregressive Hilbertian of order one ($ARH(1)$), which directly stems from works of Serge Guillas, Denis Bosq, that is defined by

$$X_t = \rho(X_{t-1}) + \zeta_t; t \in Z$$

where the random variables are all Hilbertian, ρ is a linear operator on a space of separable Hilbert and ζ_t which constitute a widley orthant dependent error (WOD, in short) after recalling some results on the finite-dimensional model of this type, we introduce the mathematical and statistical tools which will be used afterwards. Then we build an estimator of the operator and we establish its asymptotic properties.

Chapitre 1

Introduction

A time series can be defined as an ordered sequence of values of a variable at time intervals which are equally spaced.

1.1 Basic Concepts in Time Series

1.1.1 Introduction and Examples

The first definition clarifies the notion of time series analysis.

Definition 1.1.1 ***(Time Series)****. Let $T \neq \emptyset$ be an index set, conveniently being thought of as "time". A family $(X_t, t \in T)$ of random variables (random functions) is called a stochastic process. A realization of $(X_t, t \in T)$ is called a time series. We will series use the notation $(x_t, t \in T)$ in the discourse.*

The most common choices for the index set T include the integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$, the positive integers $\mathbb{N} = \{1, 2, \cdots\}$, the nonnegative integers $\mathbb{N}_0 = \{0, 1, 2, \cdots\}$, the real numbers $\mathbb{R} = (-\infty, +\infty)$ and the positive halfline $\mathbb{R}^+ = [0, \infty)$. This class is mainly concerned with the first three cases which are subsumed under the notion discrete of time series analysis.

Oftentimes the stochastic process $(X_t, t \in T)$ is itself described to as a time series, in the sense that a realization is identified with the probabilistic generating mechanism. The objective of time series analysis is to gain knowledge of this underlying random phenomenon through examining one (and typically only one) realization. This separates time series analysis from, say, regression analysis for independent data.

In the following a number of examples are given emphasizing the multitude of possible applications of time series analysis in various scientific fields.

Example 1.1.2 ***(Wölfer's sunspot numbers)*** *In figure 1.1, the number of sunspots (that is, dark spots visible on the surface of the sun) observed annually are plotted against time. The horizontal axis labels time in years, while the vertical axis represents the observed values x_t of the random variable*

$$X_t = \sharp of \ sunspots \ at \ time \ t, t = 1700, \cdots, 1994.$$

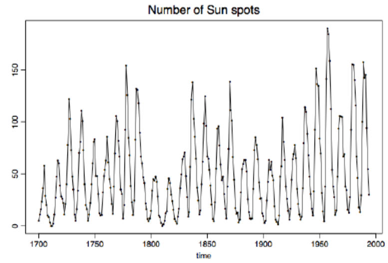


FIGURE 1.1 – Wölfer's sunspot from 1700 to 1994.

The figure is called a time series plot. It is a useful device for a preliminary analysis. Sunspot numbers are used to explain magnetic oscillations on the sun surface.

Example 1.1.3 *(**Canadian lynx data**). The time series plot in Figure (1.2) comes from a biological data set. It contains the annual returns of lynx at auction in London by the Hudson Bay Company from* 1821 − 1934 *(on a* $\log_{10}$ *scale). These are viewed as observations of the stochastic process*

$$X_t = \log_{10} (number \ of \ lynx \ trapped \ at \ time \ 1820 + t), t = 1, \cdots, 114.$$

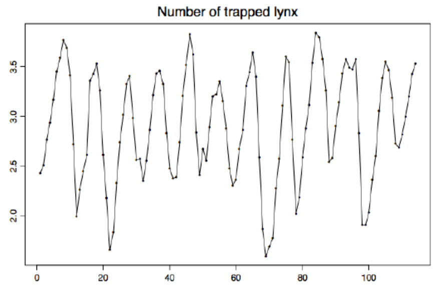


FIGURE 1.2 – Number of lynx trapped in the MacKenzie River district between 1821 and 1934.

The data is used as an estimate for the number of all lynx trapped along the MacKenzie River in Canada. This estimate, in turn, is often taken as a proxy for the true population size of the lynx.

Example 1.1.4 *(**Treasury bills**). Another important field of application for time series analysis lies in the area of finance. To hedge the risks of portfolios, investors commonly use short-term risk-free interest rates such as the yields of three-month, six-month, and twelve-month Treasury bills plotted in Figure 1.3. The (multivariate) data displayed consists of* 2, 386 *weekly observations from July* 17, 1959, *to December* 31, 1999. *Here,*

$$X_t = (X_{t,1}, X_{t,2}, X_{t,3}), \quad t = 1, \cdots, 2386,$$

where $X_{t,1}, X_{t,2}$ and $X_{t,3}$ denote the three-month, six-month, and twelve-month yields at time t, respectively. It can be seen from the graph that all three Treasury bills are moving very similarly over time, implying a high correlation between the components of X_t.

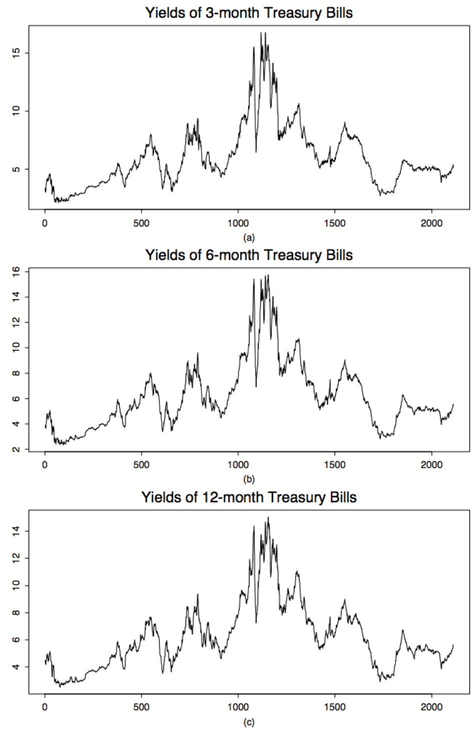


FIGURE 1.3 – Yields of Treasury bills from July 17, 1959, to December 31, 1999.

Example 1.1.5 *(***S&P500***). The Standard and Poor's* 500 *index (S&P500) is a value-weighted index based on the prices of* 500 *stocks that account for approximately* 70% *of the U.S. equity market capitalization. It is a leading economic indicator and is also used to hedge market portfolios. Figure 1.4 contains the* 7,076 *daily S&P500 closing prices from January* 3, 1972*, to December* 31, 1999*, on a natural logarithm scale. It is consequently the time series plot of the process*

$$X_t = \ln(\textit{ closing price of} S\&P500 \textit{ at time } t),\ \ t = 1, \cdots, 7076.$$

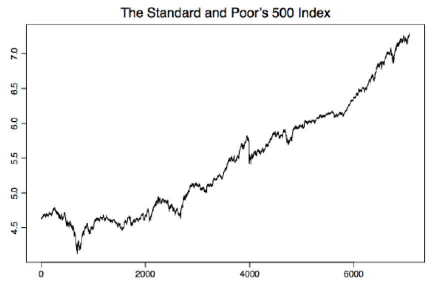


FIGURE 1.4 – S & P 500 from January 3, 1972, to December 31, 1999.

Note that the logarithm transform has been applied to make the returns directly comparable to the percentage of investment return.

There are countless other examples from all areas of science. To develop a theory capable of handling broad applications, the statistician needs to rely on a mathematical framework that can explain phenomena such as

- trends (apparent in Example (1.1.5)) ;
- seasonal or cyclical effects (apparent in Examples (1.1.2) and (1.1.3)) ;
- random fluctuations (all Examples) ;
- dependence (all Examples).

The classical approach taken in time series analysis is to postulate that the stochastic process $(X_t : t \in T)$ under investigation can be divided into deterministic trend and seasonal components plus a centered random component, giving rise to the model

$$X_t = m_t + s_t + Y_t, \quad t \in T \tag{1.1}$$

where $(m_t, t \in T)$ denotes the trend function ("mean component"), $(s_t, t \in T)$ the seasonal effects and $(Y_t, t \in T)$ a (zero mean) stochastic process. After an appropriate model has been chosen, the statistician may aim at

- estimating the model parameters for a better understanding of the time series ;
- predicting future values, for example, to develop investing strategies ;
- checking the goodness of fit to the data to confirm that the chosen model is appropriate.

Estimation procedures and prediction techniques are dealt with in detail in later sections of the notes.

1.1.2 Stationary Time Series

Our goal is to introduce a concept that keeps some of the desirable properties of independent and identically distributed random variables ("regularity"), but that also considerably enlarges the class of stochastic processes to choose from by allowing dependence as well as varying distributions. Dependence between two random variables X and Y is usually measured in terms of the covariance function

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

and the correlation function

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

With these notations at hand, the classes of strictly and weakly dependent stochastic processes can be introduced.

Definition 1.1.6 *(Strict Stationarity). A stochastic process $(X_t : t \in T)$ is called strictly stationary if, for all $t_1, \cdots, t_n \in T$ and h such that $t_1 + h, \cdots, t_n + h \in T$, it holds that*

$$(X_{t_1}, \cdots, X_{t_n}) \stackrel{\mathcal{D}}{=} (X_{t_1+h}, \cdots, X_{t_n+h})$$

That is, the so-called finite-dimensional distributions of the process are invariant under time shifts. Here $=^{\mathcal{D}}$ *indicates equality in distribution.*

The definition in terms of the finite-dimensional distribution can be reformulated equivalently in terms of the cumulative joint distribution function equalities

$$P(X_{t_1} \leq x_1, \cdots, X_{t_n} \leq x_n) = P(X_{t_1+h} \leq x_1, \cdots, X_{t_n+h} \leq x_n)$$

holding true for all $x_1, \cdots, x_n \in \mathbb{R}$, $t_1, \cdots, t_n \in T$ *and* h *such that* $t_1+h, \cdots, t_n+h \in T$. *This can be quite difficult to check for a given time series, especially if the generating mechanism of a time series is far from simple, since too many model parameters have to be estimated from the available data, rendering concise statistical statements impossible. A possible exception is provided by the case of independent and identically distributed random variables.*

To get around these difficulties, a time series analyst will commonly only specify the first- and second-order moments of the joint distributions. Doing so then leads to the notion of weak stationarity.

Definition 1.1.7 *(Weak Stationarity). A stochastic process* $(X_t; t \in T)$ *is called weakly stationary if*

- *the second moments are finite :* $\mathbb{E}[X_t^2] < \infty$ *for all* $t \in T$ *;*
- *the means are constant :* $\mathbb{E}[X_t] = m$ *for all* $t \in T$ *;*
- *the covariance of* X_t *and* X_{t+h} *depends on* h *only :*

$$\gamma(h) = \gamma_X(h) = Cov(X_t, X_{t+h}), \; h \in T \text{ such that } t+h \in T,$$

is independent of $t \in T$ *and is called the auto-covariance function (ACVF). Moreover,*

$$\rho(h) = \rho_X(h) = \frac{\gamma(h)}{\gamma(0)}, \quad h \in T,$$

is called the autocorrelation function (ACF).

Remark 1.1.8 *If* $(X_t : t \in T)$ *is a strictly stationary stochastic process with finite second moments, then it is also weakly stationary. The converse is not necessarily true. If* $(X_t : t \in T)$*, however, is weakly stationary and Gaussian, then it is also strictly stationary. Recall that a stochastic process is called Gaussian if, for any* $t_1, \cdots, t_n \in T$*, the random vector* $(X_{t_1}, \cdots, X_{t_n})$ *is multivariate normally distributed.*

This subsection is concluded with examples of stationary and non-stationary stochastic processes.

Example 1.1.9 (White Noise)*. Let* $(Z_t, t \in \mathbb{Z})$ *be a sequence of real-valued, pairwise uncorrelated variables with* $\mathbb{E}[Z_t] = 0$ *and* $0 < Var(Z_t) = \sigma^2 < \infty$ *for all* $t \in \mathbb{Z}$. *Then* $(Z_t : t \in \mathbb{Z})$ *is called white noise, abbreviated by* $(Z_t : t \in \mathbb{Z}) \sim WN(0, \sigma^2)$. *It defines a centered, weakly stationary process with ACVF and ACF given by*

$$\gamma(h) = \begin{cases} \sigma^2, & h = 0, \\ 0, & h \neq 0, \end{cases} \quad \text{and} \quad \rho(h) = \begin{cases} 1, & h = 0, \\ 0, & h \neq 0, \end{cases}$$

respectively. If the $(Z_t : t \in \mathbb{Z})$ *are moreover independent and identically distributed, they are called iid noise, shortly* $(Z_t : t \in \mathbb{Z}) \sim IID(0, \sigma^2)$*. The left panel of Figure 1.5 displays* 1000 *observations of an i.i.d. noise sequence* $(Z_t : t \in \mathbb{Z})$ *based on standard normal random variables.*

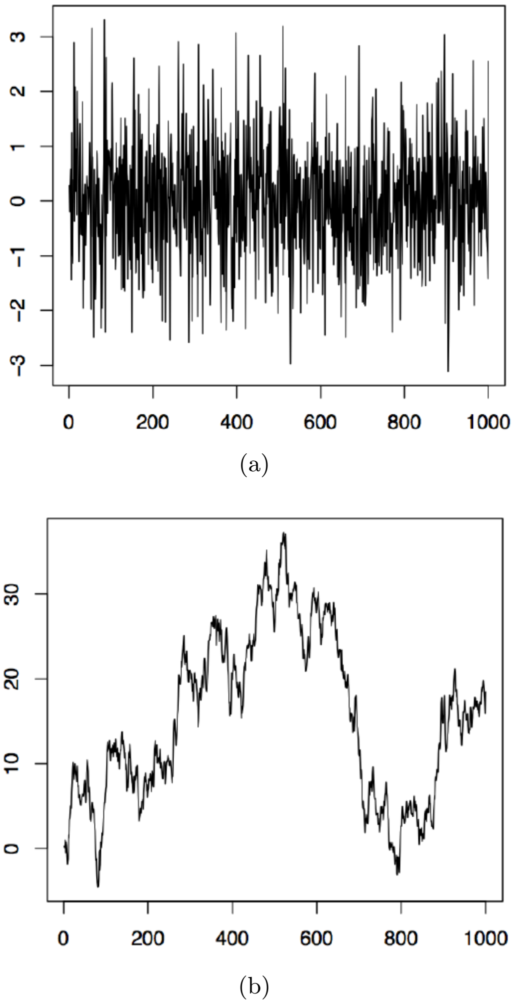


FIGURE 1.5 – 1000 simulated values of i.i.d $\mathcal{N}(0, 1)$ noise (left panel) and a random walk with i.i.d $\mathcal{N}(0, 1)$ innovations (right panel).

Example 1.1.10 (Cyclical Time Series). *Let* A *and* B *be uncorrelated random variables with zero mean and variances* $Var(A) = Var(B) = \sigma^2$*, and let* $\lambda \in \mathbb{R}$ *be a frequency parameter. Define*

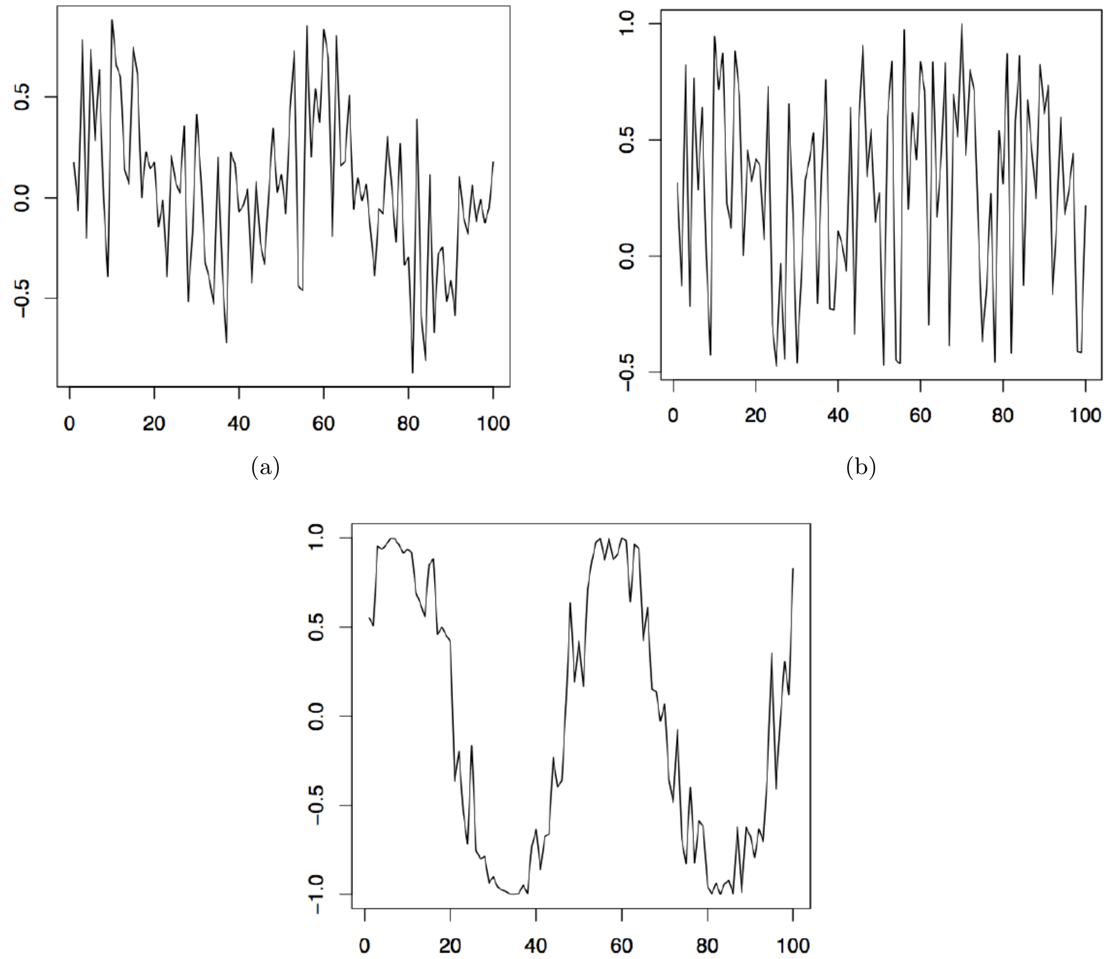
$$X_t = A\cos(\lambda t) + B\sin(\lambda t), \quad t \in \mathbb{R}.$$

The resulting stochastic process $(X_t : t \in \mathbb{R})$ *is then weakly stationary. Since* $\sin(\lambda t + \varphi) = \sin(\varphi)\cos(\lambda t) + \cos(\varphi)\sin(\lambda t)$*, the process can be represented as*

$$X_t = R\sin(\lambda t + \varphi), \quad t \in \mathbb{R},$$

so that R *is the stochastic amplitude and* $\varphi \in [-\pi, \pi]$ *the stochastic phase of a sinusoid. Some computations show that one must have* $A = R\sin(\varphi)$ *and* $B =$

$R\cos(\varphi)$. *In the left panel of Figure 1.6, 100 observed values of a series* $(X_t)_{t\in\mathbb{Z}}$ *are displayed. Therein,* $\lambda = \pi/25$ *was used, while* R *and* φ *were random variables uniformly distributed on the interval* $(-.5, 1)$ *and* $(0, 1)$*, respectively. The middle panel shows the realization of* R*, the right panel the realization of* $\sin(\lambda t + \varphi)$*. Using cyclical time series bears great advantages when seasonal effects, such as annually recurrent phenomena, have to be modeled.*



(a) (b) (c)

FIGURE 1.6 – 100 simulated values of the cyclical time series (left panel), the stochastic amplitude (middle panel), and the sine part (right panel).

Example 1.1.11 ***(Random Walk)****. Let* $(Z_t : t \in \mathbb{N}) \sim WN(0, \sigma^2)$*. Let* $S_0 = 0$ *and*

$$S_t = Z_1 + \cdots + Z_t, \quad t \in \mathbb{N}.$$

The resulting stochastic process $(S_t : t \in \mathbb{N}_0)$ *is called a random walk and is the most important non-stationary time series. Indeed, it holds here that, for* $h > 0$,

$$Cov(S_t, S_{t+h}) = Cov(S_t, S_t + R_{t,h}) = t\sigma^2,$$

where $R_{t,h} = Z_{t+1} + \cdots + Z_{t+h}$*, and the ACVF obviously depends on* t*.*

Section 1.3 discusses in detail so-called autoregressive moving average processes which have become a central building block in time series analysis. They are constructed from white noise sequences by an application of a set of stochastic difference equations similar to the ones defining the random walk $(S_t : t \in \mathbb{N}_0)$ of Example (1.1.11).
In general, the true parameters of a stationary stochastic process $(X_t : t \in T)$ are unknown to the statistician. Therefore, they have to be estimated from a realization $x_1, \cdots, x_n$. The following set of estimators will be used here. The sample mean of $x_1, \cdots, x_n$ is defined as

$$\overline{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

The sample auto-covariance function (sample ACVF) is given by

$$\widehat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \overline{x})(x_t - \overline{x}), \;\; h = 0, 1, \cdots, n-1. \tag{1.2}$$

Finally, the sample autocorrelation function (sample ACF) is

$$\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}, \qquad h = 0, 1, \cdots, n-1. \tag{1.3}$$

Example 1.1.12 *Let $(Z_t : t \in \mathbb{Z})$ be a sequence of independent standard normally distributed random variables (see the left panel of Figure 1.5 for a typical realization of size $n = 1,000$). Then, clearly, $\gamma(0) = \rho(0) = 1$ and $\gamma(h) = \rho(h) = 0$ whenever $h \neq 0$. Table 1.1 gives the corresponding estimated values $\widehat{\gamma}(h)$ and $\widehat{\rho}(h)$ for $h = 0, 1, \cdots, 5$.*

h	0	1	2	3	4	5
$\widehat{\gamma}(h)$	1.069632	0.072996	−0.000046	−0.000119	0.024282	0.0013409
$\widehat{\rho}(h)$	1.000000	0.068244	−0.000043	−0.000111	0.022700	0.0012529

TABLE 1.1 – Estimate $ACVF$ and ACF for selected values of h.

The estimated values are all very close to the true ones, indicating that the estimators work reasonably well for $n = 1,000$. Indeed it can be shown that they are asymptotically unbiased and consistent. Moreover, the sample autocorrelations $\widehat{\rho}(h)$ are approximately normal with zero mean and variance $1/1000$. See also Theorem 1.1.13 below.

Theorem 1.1.13 *Let $(Z_t : t \in \mathbb{Z}) \sim WN(0, \sigma^2)$ and let $h \neq 0$. Under a general set of conditions, it holds that the sample ACF at lag h, $\widehat{\rho}(h)$, is for large n approximately normally distributed with zero mean and variance $1/n$.*

Theorem 1.1.13 and Example 1.1.12 suggest a first method to assess whether or not a given data set can be modeled conveniently by a white noise sequence : for a white noise sequence, approximately 95% of the sample ACFs should be within the the confidence interval $\pm 2/\sqrt{n}$. Using the data files on the course webpage, one can compute with R the corresponding sample ACFs to check for whiteness of the underlying time series. The properties of the sample ACF are revisited in section 1.2.

1.1.3 Eliminating Trend Components

In this subsection three different methods are developed to estimate the trend of a time series model. It is assumed that it makes sense to postulate the model (1.1) with $s_t = 0$ for all $t \in T$, that is,

$$X_t = m_t + Y_t, \quad t \in T \tag{1.4}$$

where (without loss of generality) $\mathbb{E}[Y_t] = 0$. In particular, three different methods are discussed, (1) the least squares estimation of m_t, (2) smoothing by means of moving averages and (3) differencing.

Method 1 (**Least squares estimation**) It is often useful to assume that a trend component can be modeled appropriately by a polynomial,

$$m_t = b_0 + b_1 t + \cdots + b_p\, t^p, \qquad p \in \mathbb{N}_0.$$

In this case, the unknown parameters $b_0, \cdots, b_p$ can be estimated by the least squares method. Combined, they yield the estimated polynomial trend

$$\widehat{m}_t = \widehat{b}_0 + \widehat{b}_1 t + \cdots + \widehat{b}_p t^p, \qquad t \in T,$$

where $\widehat{b}_0, \cdots, \widehat{b}_p$ denote the corresponding least squares estimates. Note that the order p is not estimated. It has to be selected by the statistician – for example, by inspecting the time series plot. The residuals $\widehat{Y}_t$ can be obtained as

$$\widehat{Y}_t = X_t - \widehat{m}_t = X_t - \widehat{b}_0 + \widehat{b}_1\, t + \cdots + \widehat{b}_p\, t^p, \qquad t \in T,$$

How to assess the goodness of fit of the fitted trend will be subject of subsection 1.1.5 below.

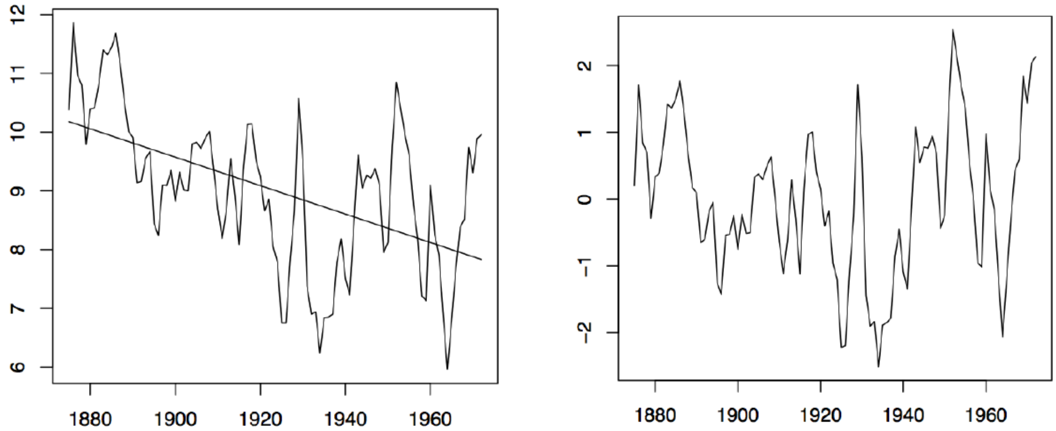
Example 1.1.14 *(**Level of Lake Huron**). The left panel of Figure 1.7 contains the time series of the annual average water levels in feet (reduced by* 570*) of Lake Huron from* 1875 *to* 1972*. It is a realization of the process*

$$X_t = (\textit{Average water level of Lake Huron in the year } 1874{+}t) - 570, \qquad t = 1, \cdots, 98.$$

There seems to be a linear decline in the water level and it is therefore reasonable to fit a polynomial of order one to the data. Evaluating the least squares estimators provides us with the values

$$\widehat{b}_0 = 10.202 \quad \textit{and} \quad \widehat{b}_1 = -0.0242$$

for the intercept and the slope, respectively. The resulting observed residuals $\widehat{y}_t = \widehat{Y}_t(\omega)$ *are plotted against time in the right panel of Figure 1.7. There is no apparent trend left in the data. On the other hand, the plot does not strongly support the stationarity of the residuals. Additionally, there is evidence of dependence in the data.*



(a)

(b)

FIGURE 1.7 – Annual water levels of Lake Huron (left panel) and the residual plot obtained from fitting a linear trend to the data (right panel).

Method 2 (**Smoothing with Moving Averages**) Let $(X_t : t \in \mathbb{Z})$ be a stochastic process following model 1.4. Choose $q \in \mathbb{N}_0$ and define the two-sided moving average

$$W_t = \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t+j}, \quad t \in \mathbb{Z}. \tag{1.5}$$

The random variables W_t can be utilized to estimate the trend component m_t in the following way. First note that

$$W_t = \frac{1}{2q+1} \sum_{j=-q}^{q} m_{t+j} + \frac{1}{2q+1} \sum_{j=-q}^{q} Y_{t+j} \approx m_t,$$

assuming that the trend is locally approximately linear and that the average of the Y_t over the interval $[t-q, t+q]$ is close to zero. Therefore, m_t can be estimated by

$$\widehat{m}_t = W_t, \quad t = q+1, \cdots, n-q.$$

Notice that there is no possibility of estimating the first q and last $n-q$ drift terms due to the two-sided nature of the moving averages. In contrast, one can also define one-sided moving averages by letting

$$\widehat{m}_1 = X_1, \quad \widehat{m}_t = aX_t + (1-a)\widehat{m}_{t-1}, \quad t = 2, \cdots, n.$$

Figure 1.8 contains estimators $\widehat{m}_t$ based on the two-sided moving averages for the Lake Huron data of Example 1.1.14. for selected choices of q (upper panel) and the corresponding estimated residuals (lower panel).
More general versions of the moving average smoothers can be obtained in the following way. Observe that in the case of the two-sided version W_t each variable $X_{t-q}, \cdots, X_{t+q}$ obtains a "weight" $a_j = (2q+1)^{-1}$. The sum of all weights thus equals one. The same is true for the one-sided moving averages

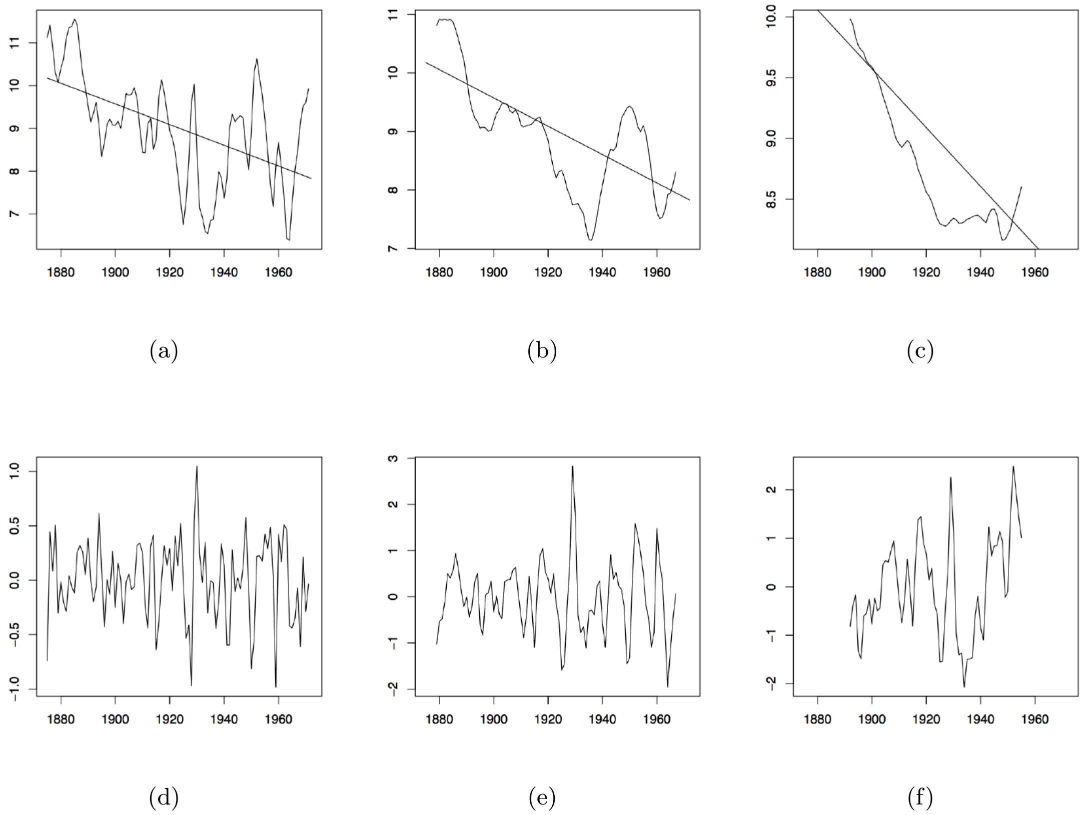


FIGURE 1.8 – The two-sided moving average filters W_t for the Lake Huron data (upper panel) and their residuals (lower panel) with bandwidth $q = 2$ (left), $q = 10$ (middle) and $q = 35$ (right).

with weights a and $1-a$. Generally, one can hence define a smoother by letting

$$\widehat{m}_t = \sum_{j=-q}^{q} a_j X_{t+j}, \quad t = q+1, \cdots, n-q, \tag{1.6}$$

where $a_{-q} + \cdots + a_q = 1$. These general moving averages (two-sided and one-sided) are commonly referred to as linear filters. There are countless choices for the weights. The one here, $a_j = (2q+1)^{-1}$, has the advantage that linear trends pass undistorted. In the next example, a filter is introduced which passes cubic trends without distortion.

Example 1.1.15 ***(Spencer's** 15**-point moving average)**. *Suppose that the filter in display 1.6 is defined by weights satisfying $a_j = 0$ if $|\, j \,| > 7$, $\;a_j = a_{-j}$ and*

$$(a_0, a_1, \cdots, a_7) = \frac{1}{320}(74, 67, 46, 21, 3, -5, -6, -3).$$

Then, the corresponding filters passes cubic trends $m_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3$ undistorted. To see this, observe that

$$\sum_{j=-7}^{7} a_j = 1 \quad \textit{and} \quad \sum_{j=-7}^{7} j^r a_j = 0, \;\; r = 1, 2, 3.$$

Now apply Proposition 1.1.16 below to arrive at the conclusion.

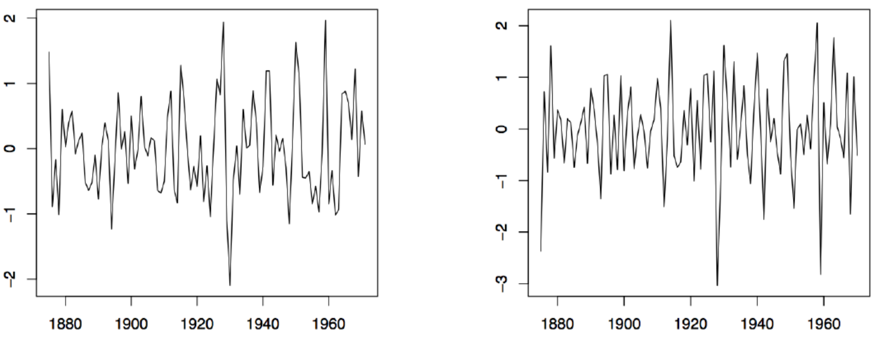
Proposition 1.1.16 *A linear filter (1.6) passes a polynomial of degree p if and only if*

$$\sum_j a_j = 1 \quad \textit{and} \quad \sum_j j^r a_j = 0, \; r = 1, \cdots, p.$$

Proof 1 *It suffices to show that* $\sum_j a_j (t+j)^r = t^r$ *for* $r = 0, \cdots, p$. *Using the binomial theorem, write*

$$\begin{aligned} \sum_j a_j (t+j)^r &= \sum_j a_j \sum_{k=0}^{r} \binom{r}{k} t^k j^{r-k} \\ &= \sum_{k=0}^{r} \binom{r}{k} t^k \left(\sum_j a_j j^{r-k} \right) \\ &= t^r \end{aligned}$$

for any $r = 0, \cdots, p$ *if and only if the above conditions hold.*



(a) (b)

FIGURE 1.9 – Time series plots of the observed sequences (∇x_t) in the left panel and $(\nabla 2 x_t)$ in the right panel of the differenced Lake Huron data described in Example 1.1.14

Method 3 (**Differencing**) A third possibility to remove drift terms from a given time series is differencing. To this end, introduce the difference operator ∇ as

$$\nabla X_t = X_t - X_{t-1} = (1 - B) X_t, \quad t \in T,$$

where B denotes the backshift operator $BX_t = X_{t-1}$. Repeated application of ∇ is defined in the intuitive way :

$$\nabla^2 X_t = \nabla(\nabla X_t) = \nabla(X_t - X_{t-1}) = X_t - 2X_{t-1} + X_{t-2}$$

and, recursively, the representations follow also for higher powers of ∇. Suppose that the difference operator is applied to the linear trend $m_t = b_0 + b_1 t$, then

$$\nabla m_t = m_t - m_{t-1} = b_0 + b_1 t - b_0 - b_1(t-1) = b_1$$

which is a constant. Inductively, this leads to the conclusion that for a polynomial drift of degree p, namely $m_t = \sum_{j=0}^{p} b_j\, t^j$, $\nabla^p m_t = p! b_p$ and thus constant. Applying this technique to a stochastic process of the form (1.1.14) with a polynomial drift m_t, yields then

$$\nabla^p X_t = p! b_p + \nabla^p Y_t, \quad t \in T.$$

This is a stationary process with mean $p!\, b_p$. The plots in Figure 1.9 contain the first and second differences for the Lake Huron data.

The next example shows that the difference operator can also be applied to a random walk to create stationary data.

Example 1.1.17 *Let $(S_t : t \in \mathbb{N}_0)$ be the random walk of Example 1.1.15. If the difference operator ∇ is applied to this stochastic process, then*

$$\nabla S_t = S_t - S_{t-1} = Z_t, \quad t \in \mathbb{N}.$$

In other words, ∇ does nothing else but recover the original white noise sequence that was used to build the random walk.

1.1.4 Eliminating Trend and Seasonal Components

Recall the classical decomposition (1.1),

$$X_t = m_t + s_t + Y_t, \quad t \in T,$$

with $E[Y_t] = 0$. In this subsection, three methods are discussed that aim at estimating both the trend and seasonal components in the data. As additional requirement on $(s_t : t \in T)$, it is assumed that

$$s_{t+d} = s_t, \quad \sum_{j=1}^{d} s_j = 0,$$

where d denotes the period of the seasonal component. (If dealing with yearly data sampled monthly, then obviously $d = 12$). It is convenient to relabel the observations $x_1, \cdots, x_n$ in terms of the seasonal period d as $x_{j,k} = x_{k+d(j-1)}$.

In the case of yearly data, observation $x_{j,k}$ thus represents the data point observed for the k^{th} month of the j^{th} year. For convenience the data is always referred to in this fashion even if the actual period is something other than 12.

Method 1 (**Small trend method**) If the changes in the drift term appear to be small, then it is reasonable to assume that the drift in year j, say, m_j is constant. As a natural estimator one can therefore apply

$$\widehat{m}_j = \frac{1}{d} \sum_{k=1}^{d} x_{j,k}.$$

To estimate the seasonality in the data, one can in a second step utilize the quantities

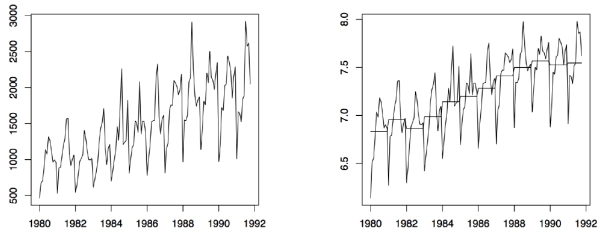
$$\widehat{s}_k = \frac{1}{N}\sum_{j=1}^{N}(x_{j,k} - \widehat{m}_j),$$

where N is determined by the equation $n = Nd$, provided that data has been collected over N full cycles. Direct calculations show that these estimators possess the property $\widehat{s}_1 + \cdots + \widehat{s}_d = 0$ (as in the case of the true seasonal components s_t). To further assess the quality of the fit, one needs to analyze the observed residuals

$$\widehat{y}_{j,k} = x_{j,k} - \widehat{m}^j - \widehat{s}_k.$$

Note that due to the relabeling of the observations and the assumption of a slowly changing trend, the drift component is solely described by the "annual" subscript j, while the seasonal component only contains the "monthly" subscript k.

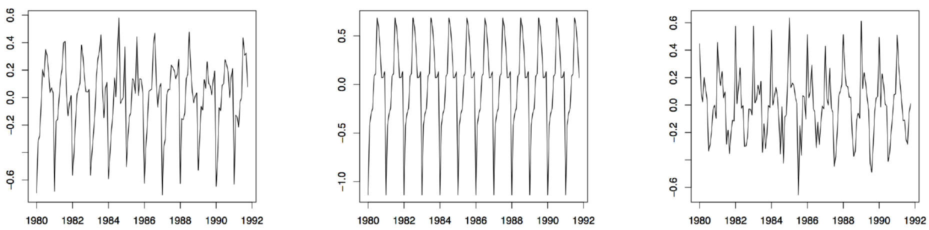
Example 1.1.18 *(**Australian Wine Sales**). The left panel of Figure 1.10 shows the monthly sales of red wine (in kiloliters) in Australia from January* 1980 *to October* 1991*. Since there is an apparent increase in the fluctuations over time, the right panel of the same figure shows the natural logarithm transform of the data. There is clear evidence of both trend and seasonality. In the following, the log transformed data is studied. Using the small trend method as described above, the annual means are estimated first. They are already incorporated in the right time series plot of Figure 1.10. Note that there are only ten months of data available for the year 1991, so that the estimation has to be adjusted accordingly. The detrended data is shown in the left panel of Figure 1.11. The middle plot in the same figure shows the estimated seasonal component, while the right panel displays the residuals. Even though the assumption of small changes in the drift is somewhat questionable, the residuals appear to look quite nice. They indicate that there is dependence in the data (see Subsection 1.1.5 below for more on this subject).*



(a) (b)

FIGURE 1.10 – Time series plots of the red wine sales in Australia from January 1980 to October 1991 (left) and its log transformation with yearly mean estimates (right).

Method 2 (**Moving average estimation**) This method is to be preferred over the first one whenever the underlying trend component cannot be assumed



(a) (b) (c)

FIGURE 1.11 – The detrended log series (left), the estimated seasonal component (center) and the corresponding residuals series (right) of the Australian red wine sales data.

constant. Three steps are to be applied to the data.

- **1st Step** : Trend estimation. At first, focus on the removal of the trend component with the linear filters discussed in the previous subsection. If the period d is odd, then one can directly use $\widehat{m}_t = W_t$ as in (1.5) with q specified by the equation $d = 2q + 1$. If the period $d = 2q$ is even, then slightly modify W_t and use

$$\widehat{m}_t = \frac{1}{d}(0.5x_{t-q} + x_{t-q+1} + \cdots + x_{t+q-1} + 0.5x_{t+q}), \quad t = q+1, \cdots, n-q.$$

- **2nd Step** : Seasonality estimation. To estimate the seasonal component, let

$$\mu_k = \frac{1}{N-1}\sum_{j=2}^{N}(x_{k+d(j-1)} - \widehat{m}_{k+d(j-1)}), \quad k = 1, \cdots, q,$$

$$\mu_k = \frac{1}{N-1}\sum_{j=1}^{N-1}(x_{k+d(j-1)} - \widehat{m}_{k+d(j-1)}), \quad k = q+1, \cdots, d.$$

Define now

$$\widehat{s}_k = \mu_k - \frac{1}{d}\sum_{l=1}^{d}\mu_l, \quad k = 1, \cdots, d,$$

and set $\widehat{s}_k = \widehat{s}_{k-d}$ whenever $k > d$. This will provide us with deseasonalized data which can be examined further. In the final step, any remaining trend can be removed from the data.

- **3rd Step** : Trend Re-estimation. Apply any of the methods from Subsection 1.1.3.

Method 3 (**Differencing at** *lag* d) Introducing the $lag-d$ difference operator ∇_d, defined by letting

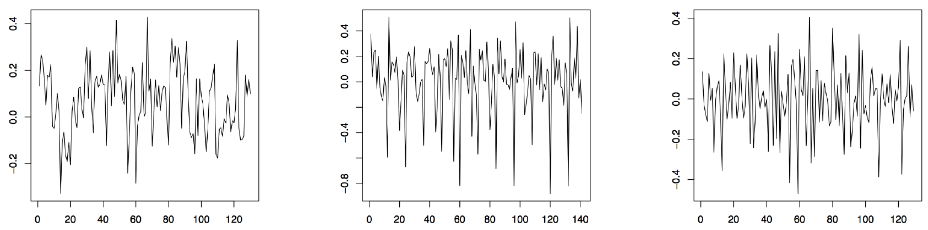
$$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d)X_t, \quad t = d+1, \cdots, n,$$

and assuming model (1.1), one arrives at the transformed random variables

$$\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}, \quad t = d+1, \cdots, n.$$

Note that the seasonality is removed, since $s_t = s_{t-d}$. The remaining noise variables $Y_t - Y_{t-d}$ are stationary and have zero mean. The new trend component $m_t - m_{t-d}$ can be eliminated using any of the methods developed in subsection 1.1.3.

Example 1.1.19 *(**Australian wine sales**). Revisit the Australian red wine sales data of Example 1.1.18 and apply the differencing techniques just established. The left plot of Figure 1.12 shows the the data after an application of the operator ∇_{12}. If the remaining trend in the data is estimated with the differencing method from Subsection 1.1.3, the residual plot given in the right panel of Figure 1.12 is obtained. Note that the order of application does not change the residuals, that is, $\nabla\nabla_{12}x_t = \nabla_{12}\nabla x_t$. The middle panel of Figure 1.12 displays the differenced data which still contains the seasonal component.*



(a) (b) (c)

FIGURE 1.12 – The differenced observed series $\nabla_{12}x_t$ (left), ∇_{x_t} (middle) and $\nabla\nabla_{12}x_t = \nabla_{12}\nabla x_t$ (right) for the Australian red wine sales data.

1.1.5 Assessing the Residuals

In this subsection, several goodness-of-fit tests are introduced to further analyze the residuals obtained after the elimination of trend and seasonal components. The main objective is to determine whether or not these residuals can be regarded as obtained from a sequence of independent, identically distributed random variables or if there is dependence in the data.
Throughout $Y_1, \cdots, Y_n$ denote the residuals and $y_1, \cdots, y_n$ a typical realization.

Method 1 (**The sample** *ACF*) It could be seen in Example 1.1.12 that, for $j \neq 0$, the estimators $\widehat{\rho}(j)$ of the *ACF* $\rho(j)$ are asymptotically independent and normally distributed with mean zero and variance n^{-1}, provided the underlying residuals are independent and identically distributed with a finite variance. Therefore, plotting the sample *ACF* for a certain number of *lags*, say h, it is expected that approximately 95% of these values are within the bounds $\pm 1.96/\sqrt{n}$. (See Theorem 1.1.13).

Method 2 (**The Portmanteau test**) The Portmanteau test is based on the test statistic

$$Q = n \sum_{j=1}^{h} \widehat{\rho}^2(j).$$

Using the fact that the variables $\sqrt{n}\widehat{\rho}(j)$ are asymptotically standard normal, it becomes apparent that Q itself can be approximated with a chi-squared distribution possessing h degrees of freedom. The hypothesis of independent and identically distributed residuals is rejected at the level α if $Q > \chi^2_{1-\alpha}(h)$, where $\chi^2_{1-\alpha}(h)$ is the $1-\alpha$ quantile of the chi-squared distribution with h degrees of freedom. Several refinements of the original Portmanteau test have been established in the literature. We refer here only to the papers Ljung and Box (1978), and McLeod and Li (1983) for further information.

Method 3 (**The rank test**) This test is very useful for finding linear trends. Denote by

$$\Pi = \sharp\{(i,j) : Y_i > Y_j, i > j, \quad i = 2, \cdots, n\}$$

the random number of pairs (i,j) satisfying the conditions $Y_i > Y_j$ and $i > j$. There are $\begin{pmatrix} n \\ 2 \end{pmatrix} = \frac{1}{2}n(n-1)$ pairs (i,j) such that $i > j$. If $Y_1, \cdots, Y_n$ are independent and identically distributed, then $P(Y_i > Y_j) = 1/2$ (assuming a continuous distribution). Now it follows that $\mu_\Pi = \mathbb{E}[\Pi] = \frac{1}{4}n(n-1)$ and, similarly, $\sigma^2_\Pi = Var(\Pi) = \frac{1}{72}n(n-1)(2n+5)$. Moreover, for large enough sample sizes n, Π has an approximate normal distribution with mean μ_Π and variance σ^2_Π. Consequently, the hypothesis of independent, identically distributed data would be rejected at the level α if

$$P = \frac{\mid \Pi - \mu_\Pi \mid}{\sigma_\Pi} > z_{1-\alpha/2},$$

where $z_{1-\alpha/2}$ denotes the $1-\alpha/2$ quantile of the standard normal distribution.

Method 4 (**Tests for normality**) If there is evidence that the data are generated by Gaussian random variables, one can create the *qq* plot to check for normality. It is based on a visual inspection of the data. To this end, denote by $Y_{(1)} < \cdots < Y_{(n)}$ the order statistics of the residuals $Y_1, \cdots, Y_n$ which are normally distributed with expected value μ and variance σ^2. It holds that

$$\mathbb{E}[Y_{(j)}] = \mu + \sigma\mathbb{E}[X_{(j)}], \tag{1.7}$$

where $X_{(1)} < \cdots < X_{(n)}$ are the order statistics of a standard normal distribution. According to display (1.7), the resulting graph will be approximately linear with the squared correlation R^2 of the points being close to 1. The assumption of normality will thus be rejected if R^2 is "too" small. It is common to approximate $\mathbb{E}[X_{(j)}] \approx \Phi_j = \Phi^{-1}((j-0.5)/n)$ (Φ being the distribution function of the standard normal distribution). The previous statement is made precise by letting

$$R^2 = \frac{\left[\sum_{j=1}^{n}(Y_{(j)} - \overline{Y})\Phi_j\right]^2}{\sum_{j=1}^{n}(Y_{(j)} - \overline{Y})^2 \sum_{j=1}^{n}\Phi_j^2},$$

where $\overline{Y} = \frac{1}{n}(Y_1 + \cdots + Y_n)$. The critical values for R^2 are tabulated and can be found, for example in Shapiro and Francia (1972).

1.2 The Estimation of Covariances

In this brief second section, some results concerning asymptotic properties of the sample $ACVF$ are collected. Throughout, $(X_t : t \in \mathbb{Z})$ denotes a weakly stationary stochastic process with $ACVF \gamma$. In Subsection 1.2.1 it was shown that such a process is completely characterized by these two quantities. The mean $\frac{1}{4}$ was estimated by the sample mean $\overline{x}$, and the $ACVF \gamma$ by the sample $ACVF \widehat{\gamma}$ is defined in Subsection (1.2.1). In the following, some properties of these estimators are discussed in more detail.

1.2.1 Estimation of the Auto-covariance Function

This section deals with the estimation of the $ACVF$ and ACF at $lag\ h$. Recall from equation (1.2) that the estimator

$$\widehat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \overline{X}_n)(X_t - \overline{X}_n), \quad h = 0, \pm 1, \cdots, \pm(n-1), \tag{1.8}$$

may be utilized as a proxy for the unknown $\gamma(h)$. As estimator for the ACF $\rho(h)$,

$$\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}, \quad h = 0, \pm 1, \cdots, \pm(n-1), \tag{1.9}$$

was identified. Some of the theoretical properties of $\widehat{\rho}(h)$ are briefly collected in the following. They are not as obvious to derive as in the case of the sample mean, and all proofs are omitted. Note also that similar statements hold for $\widehat{\gamma}(h)$ as well.

- The estimator $\widehat{\rho}(h)$ is generally biased, that is, $E[\widehat{\rho}(h)] \neq \rho(h)$. It holds, however, under non-restrictive assumptions that

$$\mathbb{E}[\widehat{\rho}(h)] \to \rho(h) \quad (n \to \infty). \tag{1.10}$$

  This property is called asymptotic unbiasedness.

- The estimator $\widehat{\rho}(h)$ is consistent for $\rho(h)$ under an appropriate set of assumptions, that is, $Var(\widehat{\rho}(h) - \rho(h)) \to 0$ as $n \to \infty$.

It was already established in Subsection 1.1.5 how the sample ACF $\widehat{\rho}$ can be used to test if residuals consist of white noise variables. For more general statistical inference, one needs to know the sampling distribution of $\widehat{\rho}$. Since the estimation of $\rho(h)$ is based on only a few observations for h close to the sample size n, estimates tend to be unreliable. As a rule of thumb, given by Box and Jenkins (1976), n should at least be 50 and h less than or equal to $n/4$.

Theorem 1.2.1 *For $m \geq 1$, let $\rho_m = (\rho(1), \cdots, \rho(m))^T$ and $\widehat{\rho}_m = (\widehat{\rho}(1), \cdots, \widehat{\rho}(m))^T$, where ^{T} denotes the transpose of a vector. Under a set of suitable assumptions, it holds that*

$$\sqrt{n}(\widehat{\rho}_m - \rho_m) \sim AN(0, \Sigma) \quad (n \to \infty), \tag{1.11}$$

where $\sim AN(0, \Sigma)$ stands for approximately normally distributed with mean vector 0 and covariance matrix $\Sigma = (\sigma_{ij})$ given by Bartlett's formula

$$\sigma_{ij} = \sum_{k=1}^{\infty} [\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)][\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)]. \quad (1.12)$$

The subsection is concluded with two examples. The first one recollects the results already known for independent, identically distributed random variables, the second deals with the autoregressive process of Example (1.2.2).

Example 1.2.2 *Let $(X_t : t \in \mathbb{Z}) \sim IID(0, \sigma^2)$. Then, $\rho(0) = 1$ and $\rho(h) = 0$ for all $h \neq 0$. The covariance matrix Σ is therefore given by*

$$\sigma_{ij} = 1 \ if \ i = j \quad and \quad \sigma_{ij} = 0 \ if \ i \neq j. \quad (1.13)$$

*This means that Σ is a diagonal matrix. In view of Theorem (1.2.1) it holds thus that the estimators $\widehat{\rho}(1), \cdots, \widehat{\rho}(k)$ are approximately independent and identically distributed normal random variables with mean 0 and variance $1/n$. This was the basis for Methods **1** and **2** in subsection 1.2.1 (see also Theorem 1.1.13).*

Example 1.2.3 *Reconsider the autoregressive process $(X_t : t \in \mathbb{Z})$ defined by :*

$$X_t = \phi(X_{t-1}) + Z_t.$$

Dividing $\gamma(h)$ by $\gamma(0)$ yields that

$$\rho(h) = \phi^{|h|}, \quad h \in \mathbb{Z}. \quad (1.14)$$

Now the diagonal entries of Σ are computed as

$$\begin{aligned}
\sigma_{ii} &= \sum [\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)]^2 \\
&= \sum_{k=1}^{i} \phi^{2i}(\phi^{-k} - \phi^{k})^2 + \sum_{k=i+1}^{\infty} \phi^{2k}(\phi^{-i} - \phi^{i})^2 \\
&= (1 - \phi^{2i})(1 + \phi^2)(1 - \phi^2)^{-1} - 2i\phi^{2i}.
\end{aligned}$$

1.3 *ARMA* processes

In this section autoregressive moving average processes are discussed. They play a crucial role in specifying time series models for applications. As the solutions of stochastic difference equations with constant coefficients and these processes possess a linear structure.

1.3.1 Introduction to ARMA

Definition 1.3.1 *(***ARMA** ***processes****)*

(a) *A weakly stationary process $X_t : t \in \mathbb{Z}$ is called an autoregressive moving average time series of order p, q, abbreviated by $ARMA(p, q)$, if it satisfies the difference equations*

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \quad t \in \mathbb{Z}, \quad (1.15)$$

where $\phi_1, \cdots, \phi_p$ and $\theta_1, \cdots, \theta_q$ are real constants, $\phi_p \neq 0 \neq \theta_q$, and $(Z_t : t \in \mathbb{Z}) \sim WN(0, \sigma^2)$.

(b) *A weakly stationary stochastic process $X_t : t \in \mathbb{Z}$ is called an $ARMA(p, q)$ time series with mean μ if the process $X_t - \mu : t \in \mathbb{Z}$ satisfies the equation system.*

A more concise representation of (1.15) can be obtained with the use of the backshift operator B. To this end, define the autoregressive polynomial and the moving average polynomial by

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p, z \in \mathbb{C},$$

and

$$\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q, z \in \mathbb{C},$$

respectively, where $\mathbb{C}$ denotes the set of complex numbers. Inserting the backshift operator into these polynomials, the equations in (1.15) become

$$\phi(B)X_t = \theta(B)Z_t, \quad t \in \mathbb{Z}. \tag{1.16}$$

Example 1.3.2 *Figure 1.13 displays realizations of three different autoregressive moving average time series based on independent, standard normally distributed $(Z_t : t \in \mathbb{Z})$. The left panel is an $ARMA(2, 2)$ process with parameter specifications $\phi_1 = 0.2, \phi_2 = -0.3, \theta_1 = -0.5$ and $\theta_2 = 0.3$. The middle plot is obtained from an $ARMA(1, 4)$ process with parameters $\phi 1 = .3, \theta_1 = -0.2, \theta_2 = -0.3, \theta_3 = 0.5$, and $\theta_4 = 0.2$, while the right plot is from an $ARMA(4, 1)$ with parameters $\phi_1 = -0.2, \phi_2 = -0.3, \phi_3 = 0.5$ and $\phi_4 = 0.2$ and $\theta_1 = 0.6$. The plots indicate that $ARMA$ models can provide a flexible tool for modeling diverse residual sequences. It will turn out in the next subsection that all three realizations here come from (strictly) stationary processes.*

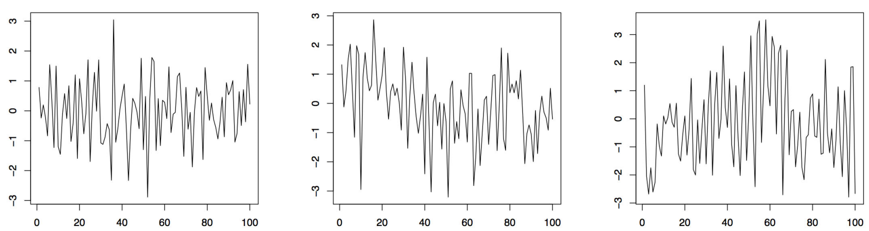


FIGURE 1.13 – Realizations of three autoregressive moving average processes.

Some special cases covered in the following two examples have particular relevance in time series analysis.

Example 1.3.3 *(**AR Processes**) If the moving average polynomial in (1.16) is equal to one, that is, if $\theta(z) \equiv 1$, then the resulting $(X_t : t \in \mathbb{Z})$ is referred to as autoregressive process of order p, $AR(p)$. These time series interpret the value of the current variable X_t as a linear combination of p previous variables $X_{t-1}, \cdots, X_{t-p}$ plus an additional distortion by the white noise Z_t. Figure (1.14) displays two $AR(1)$ processes with respective parameters $\phi_1 = -0.9$ (left) and $\phi_1 = 0.8$ (middle) as well as an $AR(2)$ process with parameters $\phi_1 = -0.5$ and $\phi_2 = 0.3$.*

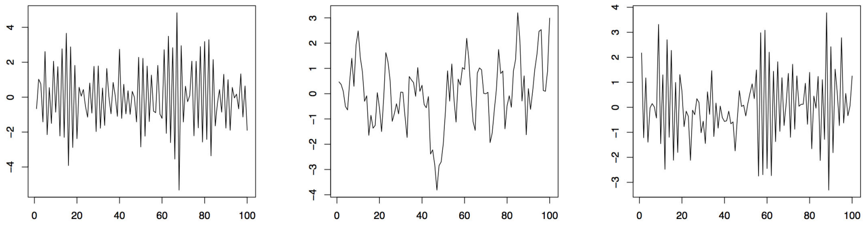


FIGURE 1.14 – Realizations of three autoregressive processes.

Example 1.3.4 *(***MA** ***Processes****) If the autoregressive polynomial in (1.16) is equal to one, that is, if $\phi(z) \equiv 1$, then the resulting $(X_t : t \in \mathbb{Z})$ is referred to as moving average process of order q, $MA(q)$. Here the present variable X_t is obtained as superposition of q white noise terms $Z_t, \cdots, Z_{t-q}$. Figure (1.15) shows two $MA(1)$ processes with respective parameters $\theta_1 = 0.5$ (left) and $\theta_1 = -0.8$ (middle). The right plot is observed from an $MA(2)$ process with parameters $\theta_1 = -0.5$ and $\theta_2 = 0.3$.*

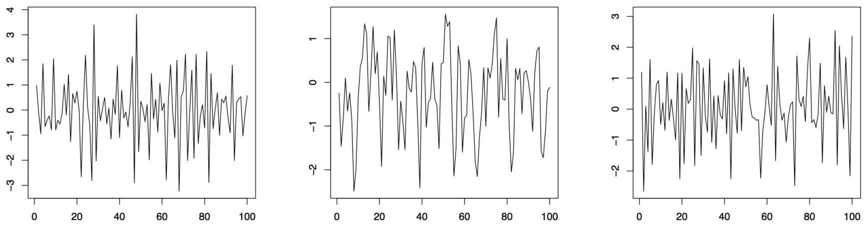


FIGURE 1.15 – Realizations of three moving average processes.

For the analysis upcoming in the next sections, we now introduce moving average processes of infinite order ($q = \infty$). They are an important tool for determining stationary solutions to the difference equations (1.15).

Definition 1.3.5 Linear processes

A stochastic process $(X_t : t \in \mathbb{Z})$ is called linear process or $MA(\infty)$ time series if there is a sequence $(\psi_j : j \in \mathbb{N}_0)$ with $\sum_{j=0}^{\infty} | \psi_j | < \infty$ such that

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z}, \tag{1.17}$$

where $(Z_t : t \in \mathbb{Z}) \sim WN(0, \sigma^2)$.

Moving average time series of any order q are special cases of linear processes. Just pick $\psi_j = \theta_j$ for $j = 1, \cdots, q$ and set $\psi_j = 0$ if $j > q$. It is common to introduce *the power series*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j, \quad z \in \mathbb{C},$$

to express a linear process in terms of the backshift operator. Display (1.17) can now be rewritten in the compact form

$$X_t = \psi(B) Z_t, \quad t \in \mathbb{Z}.$$

With the definitions of this subsection at hand, properties of ARMA processes, such as stationarity and invertibility, are investigated in the next subsection. The current subsection is closed giving meaning to the notation $X_t = \psi(B)Z_t$*. Note that one is possibly dealing with an infinite sum of random variables. For completeness and later use, in the following example the mean and ACVF of a linear process are derived.*

Example 1.3.6 Mean and ACVF of a linear process
Let $(X_t : t \in \mathbb{Z})$ *be a linear process according to Definition (1.3.5). Then, it holds that*

$$\mathbb{E}[X_t] = \mathbb{E}\left[\sum_{j=0}^{\infty} \psi_j Z_{t-j}\right] = \sum_{j=0}^{\infty} \psi_j \mathbb{E}\left[Z_{t-j}\right] = 0, \ t \in \mathbb{Z}.$$

Next observe also that

$$\begin{aligned} \gamma(h) &= Cov(X_{t+h}, X_t) \\ &= \mathbb{E}\left[\sum_{j=0}^{\infty} \psi_j Z_{t+h-j} \sum_{k=0}^{\infty} \psi_k Z_{t-k}\right] \\ &= \sigma^2 \sum_{k=0}^{\infty} \psi_{k+h} \psi_k < \infty \end{aligned}$$

by assumption on the sequence $(\psi_j : j \in \mathbb{N}_0)$ *.*

1.3.2 Causality and Invertibility

While a moving average process of order q will always be stationary without conditions on the coefficients $\theta_1, \cdots, \theta_q$*, some deeper thoughts are required in the case of* $AR(p)$ *and* $ARMA(p,q)$ *processes. For simplicity, we start by investigating the autoregressive process of order one, which is given by the equations* $X_t = \phi X_{t-1} + Z_t$ *(writing* $\phi = \phi_1$*). Repeated iterations yield that*

$$X_t = \phi X_{t-1} + Z_t = \phi^2 X_{t-2} + Z_t + \phi Z_{t-1} = \cdots = \phi^N X_{t-N} + \sum_{j=0}^{N-1} \phi^j Z_{t-j}. \tag{1.18}$$

Letting $N \to \infty$*, it could now be shown that, with probability one,*

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j} \tag{1.19}$$

is the weakly stationary solution to the $AR(1)$ *equations, provided that* $\mid \phi \mid < 1$*. These calculations would indicate moreover, that an autoregressive process of order one can be represented as linear process with coefficients* $\psi_j = \phi^j$*.*

Example 1.3.7 Mean and** $ACVF$ **of an** $AR(1)$ **process

Since an autoregressive process of order one has been identified as an example of a linear process, one can easily determine its expected value as

$$\mathbb{E}[X_t] = \sum_{j=0}^{\infty} \phi_j \mathbb{E}[Z_{t-j}] = 0, \quad t \in \mathbb{Z}. \tag{1.20}$$

For the $ACVF$, it is obtained that

$$\begin{aligned} \gamma(h) &= Cov(X_{t+h}, X_t) \\ &= \mathbb{E}\left[\sum_{j=0}^{\infty} \phi^j Z_{t+h-j} \sum_{k=0}^{\infty} \phi^k Z_{t-k}\right] \\ &= \sigma^2 \sum_{k=0}^{\infty} \phi^{k+h} \phi^k = \sigma^2 \phi^h \sum_{k=0}^{\infty} \phi^{2k} = \frac{\sigma^2 \phi^h}{1-\phi^2}, \end{aligned}$$

where $h \geq 0$. This determines the $ACVF$ for all h using that $\gamma(-h) = \gamma(h)$. It is also immediate that the ACF satisfies $\rho(h) = \phi^h$. See also Example 1.3.2 for comparison.

Example 1.3.8 Non Stationary** $AR(1)$ **Processes

In Example 1.1.11 we have introduced the random walk as a non-stationary time series. It can also be viewed as a non-stationary $AR(1)$ process with parameter $\phi = 1$. In general, autoregressive processes of order one with coefficients $\mid \phi \mid > 1$ are called explosive for they do not admit a weakly stationary solution that could be expressed as a linear process. However, one may proceed as follows. Rewrite the defining equations of an $AR(1)$ process as

$$X_t = -\phi^{-1} Z_{t+1} + \phi^{-1} X_{t+1}, \quad t \in \mathbb{Z}. \tag{1.21}$$

Apply now the same iterations as before to arrive at

$$X_t = \phi^{-N} X_{t+N} - \sum_{j=1}^{N} \phi^{-j} Z_{t+j}, \quad t \in \mathbb{Z}. \tag{1.22}$$

Note that in the weakly stationary case, the present observation has been described in terms of past innovations. The representation in the last equation however contains only future observations with time lags larger than the present time t. From a statistical point of view this does not make much sense, even though by identical arguments as above we may obtain

$$X_t = -\sum_{j=1}^{\infty} \phi^{-j} Z_{t+j}, \; t \in \mathbb{Z}, \tag{1.23}$$

as the weakly stationary solution in the explosive case.

The result of the previous example leads to the notion of causality which means that the process $(X_t : t \in \mathbb{Z})$ has a representation in terms of the white noise $(Z_s : s \leq t)$ and that is hence uncorrelated with the future as given by $(Z_s : s > t)$. We give the definition for the general $ARMA$ case.

Definition 1.3.9 Causality

An $ARMA(p,q)$ process given by (1.15) is causal if there is a sequence $(\psi_j : j \in \mathbb{N}_0)$ such that $\sum_{j=0}^{\infty} | \psi_j | < \infty$ and

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z}. \tag{1.24}$$

Causality means that an ARMA time series can be represented as a linear process. It was seen earlier in this subsection how an $AR(1)$ process whose coefficient satisfies the condition $| \phi | < 1$ can be converted into a linear process. It was also shown that this is impossible if $| \phi | > 1$. The conditions on the autoregressive parameter ϕ can be restated in terms of the corresponding autoregressive polynomial $\phi(z) = 1 - \phi z$ as follows. It holds that

$$\begin{array}{llll} | \phi | < 1 & if\ and\ only\ if & \phi(z) \neq 0 & for\ all \quad | z | \leq 1, \\ | \phi | > 1 & if\ and\ only\ if & \phi(z) \neq 0 & for\ all \quad | z | \geq 1 \end{array}$$

It turns out that the characterization in terms of the zeroes of the autoregressive polynomials carries over from the $AR(1)$ case to the general $ARMA(p,q)$ case. Moreover, the ψ-weights of the resulting linear process have an easy representation in terms of the polynomials $\phi(z)$ and $\theta(z)$. The result is summarized in the next theorem.

Theorem 1.3.10 *Let $(X_t : t \in \mathbb{Z})$ be an $ARMA(p,q)$ process such that the polynomials $\phi(z)$ and $\theta(z)$ have no common zeroes. Then $(X_t : t \in \mathbb{Z})$ is causal if and only if $\theta(z) \neq 0$ for all $z \in \mathbb{C}$ with $| z | \leq 1$. The coefficients $(\psi_j : j \in \mathbb{N}_0)$ are determined by the power series expansion*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad | z | \leq 1. \tag{1.25}$$

A concept closely related to causality is invertibility. This notion is motivated with the following example that studies properties of a moving average time series of order 1.

Example 1.3.11 *Let $(X_t : t \in \mathbb{N})$ be an $MA(1)$ process with parameter $\theta = \theta_1$. It is an easy exercise to compute the ACVF and the ACF as*

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma^2, & h = 0 \\ \theta\sigma^2, & h = 1 \\ 0 & h > 1, \end{cases} \qquad \rho = \begin{cases} 1 & h = 0. \\ \theta(1+\theta^2)^{-1} & h = 1. \\ 0 & h > 1. \end{cases} \tag{1.26}$$

These results lead to the conclusion that $\rho(h)$ does not change if the parameter θ is replaced with θ^{-1}. Moreover, there exist pairs (θ, σ^2) that lead to the same ACVF, for example $(5,1)$ and $(1/5, 25)$. Consequently, we arrive at the fact that the two $MA(1)$ models

$$X_t = Z_t + \frac{1}{5} Z_{t-1}, \quad t \in \mathbb{Z}, \quad (Z_t : t \in \mathbb{Z}) \sim iid\mathcal{N}(0, 25) \tag{1.27}$$

and

$$X_t = \widehat{Z}_t + 5\widehat{Z}_{t-1}, \quad t \in \mathbb{Z}, \quad (\widehat{Z}_t : t \in \mathbb{Z}) \sim iid\mathcal{N}(0,1) \tag{1.28}$$

are indistinguishable because we only observe X_t *but not the noise variables* Z_t *and* $\widehat{Z}_t$.

For convenience, the statistician will pick the model which satisfies the invertibility criterion which is to be defined next. It specifies that the noise sequence can be represented as a linear process in the observations.

Definition 1.3.12 Invertibility

An $ARMA(p,q)$ *process given by 1.15 is invertible if there is a sequence* $(\pi_j : j \in \mathbb{N}_0)$ *such that* $\sum_{j=0}^{\infty} \mid \pi_j \mid < \infty$ *and*

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad t \in \mathbb{Z}. \tag{1.29}$$

Theorem 1.3.13 *Let* $(X_t : t \in \mathbb{Z})$ *be an* $ARMA(p,q)$ *process such that the polynomials* $\phi(z)$ *and* $\theta(z)$ *have no common zeroes. Then* $(X_t : t \in \mathbb{Z})$ *is invertible if and only if* $\theta(z) \neq 0$ *for all* $z \in \mathbb{C}$ *with* $\mid z \mid \leq 1$. *The coefficients* $(\pi_j)_{j\in\mathbb{N}_0}$ *are determined by the power series expansion*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad \mid z \mid \leq 1. \tag{1.30}$$

From now on it is assumed that all ARMA sequences specified in the sequel are causal and invertible unless explicitly stated otherwise. The final example of this subsection highlights the usefulness of the established theory. It deals with parameter redundancy and the calculation of the causality and invertibility sequences $(\psi_j : j \in \mathbb{N}_0)$ *and* $(\pi_j : j \in \mathbb{N}_0)$.

Example 1.3.14 Parameter Redundancy

Consider the ARMA equations

$$X_t = 0.4X_{t-1} + 0.21X_{t-2} + Z_t + 0.6Z_{t-1} + 0.09Z_{t-2}, \tag{1.31}$$

which seem to generate an $ARMA(2,2)$ *sequence. However, the autoregressive and moving average polynomials have a common zero :*

$$\begin{aligned} \widetilde{\phi}(z) &= 1 - 0.4z - 0.21z^2 &= (1-0.7z)(1+0.3z), \\ \widetilde{\theta}(z) &= 1 + 0.6z + 0.09z^2 &= (1+0.3z)^2. \end{aligned}$$

Therefore, one can reset the ARMA equations to a sequence of order $(1,1)$ *and obtain*

$$X_t = 0.7X_{t-1} + Z_t + 0.3Z_{t-1}. \tag{1.32}$$

Now, the corresponding polynomials have no common roots. Note that the roots of $\phi(z) = 1 - 0.7z$ *and* $\theta(z) = 1 + 0.3z$ *are* $10/7 > 1$ *and* $-10/3 < -1$, *respectively.*

Thus theorems 1.3.10 and 1.3.13 imply that causal and invertible solutions exist. In the following, the corresponding coefficients in the expansions

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \ \ \textit{and}\ Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},\ t \in \mathbb{Z}, \tag{1.33}$$

are calculated. Starting with the causality sequence $(\psi_j : j \in \mathbb{N}_0)$. *Writing, for* $\mid z \mid \leq 1$,

$$\sum_{j=0}^{\infty} \psi_j z^j = \psi(z) = \frac{\theta(z)}{\phi(z)} = \frac{1+0.3z}{1-0.7z} = (1+0.3z)\sum_{j=0}^{\infty}(0.7z)^j, \tag{1.34}$$

it can be obtained from a comparison of coefficients that

$$\psi_0 = 1 \ \ \textit{and}\ \ \psi_j = (0.7+0.3)(0.7)^{j-1} = (0.7)^{j-1},\ j \in \mathbb{N}. \tag{1.35}$$

Similarly one computes the invertibility coefficients $(\pi_j : j \in \mathbb{N}_0)$ *from the equation*

$$\sum_{j=0}^{\infty} \pi_j z^j = \pi(z) = \frac{\phi(z)}{\theta(z)} = \frac{1-0.7z}{1+0.3z} = (1-0.7z)\sum_{j=0}^{\infty}(-0.3z)^j \tag{1.36}$$

$(\mid z \mid \leq 1)$ *as*

$$\pi_0 = 1 \ \ \textit{and}\ \ \pi_j = (-1)^j(0.3+0.7)(0.3)^{j-1} = (-1)^j(0.3)^{j-1}. \tag{1.37}$$

Together, the previous calculations yield to the explicit representations

$$X_t = Z_t + \sum_{j=1}^{\infty}(0.7)^{j-1} Z_{t-j} \ \ \textit{and}\ \ Z_t = X_t + \sum_{j=1}^{\infty}(-1)^j(0.3)^{j-1} X_{t-j}. \tag{1.38}$$

In the remainder of this subsection, a general way is provided to determine the weights $(\psi_j : j \geq 1)$ *for a causal* $ARMA(p,q)$ *process given by* $\phi(B)X_t = \theta(B)Z_t$, *where* $\phi(z) \neq 0$ *for all* $z \in \mathbb{C}$ *such that* $\mid z \mid \leq 1$. *Since* $\psi(z) = \theta(z)/\phi(z)$ *for these* z, *the weight* ψ_j *can be computed by matching the corresponding coefficients in the equation* $\psi(z)\phi(z) = \theta(z)$, *that is,*

$$(\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots)(1 - \phi_1 z - \cdots - \phi_p z^p) = 1 + \theta_1 z + \cdots + \theta_q z^q. \tag{1.39}$$

Recursively solving for $\psi_0, \psi_1, \psi_2, \cdots$ *gives*

$$\begin{aligned} \psi_0 &= 1, \\ \psi_1 - \phi_1\psi_0 &= \theta_1, \\ \psi_2 - \phi_1\psi_1 - \phi_2\psi_0 &= \theta_2, \end{aligned}$$

and so on as long as $j < \max\{p, q+1\}$. *The general solution can be stated as*

$$\psi_j - \sum_{k=1}^{j} \phi_k \psi_{j-k} = \theta_j, \quad 0 \leq j < \max\{p, q+1\}, \tag{1.40}$$

$$\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = 0, \quad j \geq \max\{p, q+1\}, \tag{1.41}$$

if we define $\phi_j = 0$ *if* $j > p$ *and* $\theta_j = 0$ *if* $j > q$. *To obtain the coefficients* ψ_j *one therefore has to solve the homogeneous linear difference equation (1.41) subject to the initial conditions specified by (1.40).*

For more this subject, see (23) section 3.6 *and (27) section* (3.3).

1.3.3 Parameter Estimation

Let $(X_t : t \in \mathbb{Z})$ be a causal and invertible $ARMA(p,q)$ process with known orders p and q, possibly with mean μ. This section is concerned with estimation procedures for the unknown parameter vector

$$\beta = (\mu, \phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q, \sigma^2)^T. \tag{1.42}$$

To simplify the estimation procedure, it is assumed that the data has already been adjusted by subtraction of the mean and the discussion is therefore restricted to zero mean ARMA models.

In the following, three estimation methods are introduced. The method of moments works best in case of pure AR processes, while it does not lead to optimal estimation procedures for general ARMA processes. For the latter, more efficient estimators are provided by the maximum likelihood and least squares methods which will be discussed subsequently.

Method 1 *(**Method of Moments**) Since this method is only efficient in their case, the presentation here is restricted to $AR(p)$ processes*

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t, \quad t \in \mathbb{Z}, \tag{1.43}$$

where $(Z_t : t \in \mathbb{Z}) \sim WN(0, \sigma^2)$. The parameter vector β consequently reduces to $(\phi, \sigma^2)^T$ with $\phi = (\phi_1, \cdots, \phi_p)^T$ and can be estimated using the Yule-Walker equations

$$\Gamma_p \phi = \gamma_p \quad and \quad \sigma^2 = \gamma(0) - \phi^T \gamma_p, \tag{1.44}$$

where $\Gamma_p = (\gamma(k-j))_{k,j=1,\cdots,p}$ and $\gamma_p = (\gamma(1), \cdots, \cdots (p))^T$. Observe that the equations are obtained by the same arguments applied to derive the Durbin-Levinson algorithm in the previous subsection. The method of moments suggests to replace every quantity in the Yule-Walker equations with their estimated counterparts, which yields the Yule-Walker estimators

$$\widehat{\phi} = \widehat{\Gamma}_p^{-1}\ \widehat{\gamma}_p = \widehat{R}_p^{-1} \widehat{\rho}_p \tag{1.45}$$

$$\widehat{\sigma}^2 = \widehat{\gamma}(0) - \ \widehat{\gamma}_p^T\ \widehat{\Gamma}_p^{-1} \widehat{\gamma}_p = \widehat{\gamma}(0)[1 - \ \widehat{\rho}_p^T\ \widehat{R}_p^{-1} \widehat{\rho}_p]. \tag{1.46}$$

Therein, $\widehat{R}_p = \widehat{\gamma}(0)^{-1} \widehat{\gamma}_p$ and $\widehat{\rho}_p = \widehat{\gamma}(0)^{-1} \widehat{\Gamma}_p$ with $\widehat{\gamma}_p(h)$ defined as in (1.2). Using $\widehat{\gamma}(h)$ as estimator for the ACVF at lag h, a dependence on the sample size n is obtained in an implicit way. This dependence is suppressed in the notation used here. The following theorem contains the limit behavior of the Yule-Walker estimators as n tends to infinity.

Theorem 1.3.15 *If $(X_t : t \in \mathbb{Z})$ is a causal $AR(p)$ process, then*

$$\sqrt{n}(\widehat{\phi} - \phi) \to_{\mathcal{D}} N(0, \sigma^2 \Gamma_p^{-1}) \quad and \quad \widehat{\sigma}^2 \to_{\mathcal{P}} \sigma^2 \tag{1.47}$$

as $n \to \infty$, where $\longrightarrow_{\mathcal{P}}$ indicates convergence in probability.

Corollary 1.3.16 *If $(X_t : t \in \mathbb{Z})$ is a causal $AR(p)$ process, then*

$$\sqrt{n} \widehat{\phi}_{hh} \longrightarrow_{\mathcal{D}} Z \qquad \to (n \to \infty) \tag{1.48}$$

for all $h > p$, where Z stands for a standard normal random variable.

Example 1.3.17 *(**Yule-Walker estimates for** $AR(2)$ **processes**). Suppose that $n = 144$ values of the autoregressive process $X_t = 1.5X_{t-1} - .75X_{t-2} + Z_t$ have been observed, where $(Z_t : t \in \mathbb{Z})$ is a sequence of independent standard normal variate. Assume further that $\widehat{\gamma}(0) = 8.434, \widehat{\rho}(1) = 0.834$ and $\widehat{\rho}(2) = 0.476$ have been calculated from the data. The Yule-Walker estimators for the parameters are then given by*

$$\widehat{\phi} = \begin{pmatrix} \widehat{\phi}_1 \\ \widehat{\phi}_2 \end{pmatrix} = \begin{pmatrix} 1.000 & 0.834 \\ 0.834 & 1.000 \end{pmatrix}^{-1} \begin{pmatrix} 0.834 \\ 0.476 \end{pmatrix} = \begin{pmatrix} 1.439 \\ -0.725 \end{pmatrix} \tag{1.49}$$

and

$$\widehat{\sigma}^2 = 8.434 \left[1 - \begin{pmatrix} 0.834 & 0.476 \end{pmatrix} \begin{pmatrix} 1.439 \\ -0.725 \end{pmatrix} \right] = 1.215. \tag{1.50}$$

To construct asymptotic confidence intervals using Theorem 1.3.15, the unknown limiting covariance matrix $\sigma^2\Gamma_p^{-1}$ needs to be estimated. This can be done using the estimator

$$\frac{\widehat{\rho}^2\widehat{\Gamma}_p^{-1}}{n} = \frac{1}{144}\frac{1.215}{8.434} \begin{pmatrix} 1.000 & 0.834 \\ 0.834 & 1.000 \end{pmatrix}^{-1} = \begin{pmatrix} 0.057^2 & -0.003 \\ -0.003 & 0.057^2 \end{pmatrix}. \tag{1.51}$$

Then, the $1-\alpha$ level confidence interval for the parameters ϕ_1 and ϕ_2 are computed as

$$1.439 \pm 0.057 z_{1-\alpha/2} \quad \text{and} \quad -0.725 \pm 0.057 z_{1-\alpha/2}, \tag{1.52}$$

respectively, where $z_{1-\alpha/2}$ is the corresponding normal quantile.

Method 2 *(**Maximum Likelihood Estimation**) The innovations algorithm of the previous subsection applied to a causal $ARMA(p, q)$ process $(X_t : t \in \mathbb{Z})$ gives*

$$\widehat{X}_{i+1} = \sum_{j=1}^{i} \theta_{ij}(X_{i+1-j} - \widehat{X}_{i+1-j})b, \quad 1 \le i < \max\{p, q\}, \tag{1.53}$$

$$\widehat{X}_{i+1} = \sum_{j=1}^{p} \phi_j X_{i+1-j} + \sum_{j=1}^{q} \theta_{ij}(X_{i+1-j} - \widehat{X}_{i+1-j}), i \ge \max\{p, q\}, \tag{1.54}$$

with prediction error

$$P_{i+1} = \sigma^2 R_{i+1}. \tag{1.55}$$

In the last expression, σ^2 has been factored out due to reasons that will become apparent from the form of the likelihood function to be discussed below. Recall that the sequence $(X_{i+1} - \widehat{X}_{i+1} : i \in \mathbb{Z})$ consists of uncorrelated random variables if the parameters are known. Assuming normality for the errors, we moreover obtain even independence. This can be exploited to define the Gaussian maximum likelihood estimation (MLE) procedure. Throughout, it is assumed that $(X_t : t \in \mathbb{Z})$ has zero mean ($\mu = 0$). The parameters of interest are collected in the vectors $\beta = (\phi, \theta, \sigma^2)^T$ and $\beta' = (\phi, \ \theta)^T$, where

$\phi = (\phi_1, \cdots, \phi_p)^T$ and $\theta = (\theta_1, \cdots, \theta_q)^T$. Assume finally that we have observed the variables $X_1, \cdots, X_n$. Then, the Gaussian likelihood function for the innovations is

$$L(\beta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \left(\prod_{i=1}^{n} R_i^{1/2} \right) \exp \left(-\frac{1}{2\sigma^2} \sum_{j=1}^{n} \frac{(X_j - \widehat{X}_j)^2}{R_j} \right). \tag{1.56}$$

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

$$\widehat{\sigma}^2 = \frac{S(\widehat{\phi}, \widehat{\theta})}{n}, \quad S(\widehat{\phi}, \widehat{\theta}) = \sum_{j=1}^{n} \frac{(X_j - \widehat{X}_j)^2}{R_j}. \tag{1.57}$$

Therein, $\widehat{\phi}$ and $\widehat{\theta}$ denote the MLEs of ϕ and θ obtained from minimizing the profile likelihood or reduced likelihood

$$\ell(\phi, \theta) = \ln \left(\frac{S(\phi, \theta)}{n} \right) + \frac{1}{n} \sum_{j=1}^{n} \ln(R_j). \tag{1.58}$$

Observe that the profile likelihood $\ell(\phi, \theta)$ can be computed using the innovations algorithm. The speed of these computations depends heavily on the quality of initial estimates. These are often provided by the non-optimal Yule-Walker procedure. For numerical methods, such as the Newton-Raphson and scoring algorithms, see (27).

The limit distribution of the MLE procedure is given as the following theorem. Its proof can be found in (23) Section 8.8.

Theorem 1.3.18 *. Let $(X_t : t \in \mathbb{Z})$ be a causal and invertible $ARMA(p, q)$ process defined with an i.i.d sequence $(Z_t : t \in \mathbb{Z})$ satisfying $\mathbb{E}[Z_t] = 0$ and $\mathbb{E}[Z_t^2] = \sigma^2$. Consider the MLE $\widehat{\beta}'$ of β' that is initialized with the moment estimators of Method 1. Then,*

$$\sqrt{n}(\widehat{\beta}' - \beta') \to^{\mathcal{D}} N(0, \sigma^2 \Gamma_{p,q}^{-1}) \qquad (n \to \infty). \tag{1.59}$$

The result is optimal. The covariance matrix $\Gamma_{p,q}$ is in block form and can be evaluated in terms of covariances of various autoregressive processes.

Method 3 *(**Least Squares Estimation**) An alternative to the method of moments and the MLE is provided by the least squares estimation (LSE). For causal and invertible $ARMA(p, q)$ processes, it is based on minimizing the weighted sum of squares*

$$S(\phi, \theta) = \sum_{j=1}^{n} \frac{(X_j - \widehat{X}_j)^2}{R_j} \tag{1.60}$$

with respect to ϕ and θ, respectively. Assuming that $\widetilde{\phi}$ and $\widetilde{\theta}$ denote these LSEs, the LSE for σ^2 is computed as

$$\widetilde{\sigma}^2 = \frac{S(\widetilde{\phi}, \widetilde{\theta})}{n - p - q}. \tag{1.61}$$

The least squares procedure has the same asymptotic as the MLE.

Theorem 1.3.19 *The result of Theorem 1.3.18. holds also if $\widehat{\beta}'$ is replaced with $\widetilde{\beta}'$.*

1.4 Autoregressive Hilbertian Processes

In this section, we note some results of functional autoregressive processes theory introduced by Bosq (1991) (7). We are interested in estimating autocorrelation operator defining autoregressive processes structure. First, we refer the estimation results in case where the operator is defined in Hilbert space.
Let H Hilbert real separable space muni with scalar product $<,>$ related to the norm $\| . \|$, and its borelien tribe $\mathcal{B}$. Let ρ linear bounded operator on H such that

$$\| \rho \|^{j_0} < 1 \ for \ some \ j_0 \geq 1,$$

and $\zeta = (\zeta_t, \ t \in \mathbb{Z})$ sequence of independents random variables identically distributed to value in H, such that

$$0 < \mathbb{E} \| \zeta_t \|^2 = \sigma^2 < +\infty \ and \ \mathbb{E}(\zeta_t) = 0.$$

Hilbertian stable autoregressive processes of first order, note that $ARH(1)$, is unique stable solution of equation :

$$X_t = \rho(X_{t-1}) + \zeta_t, \ t \in \mathbb{Z} \tag{1.62}$$

The general method to estimate ρ consists to use covariance operators and cross-covariance of processes.
Covariance operator of X_0 is a symmetric positive atomic (nuclear)operator from H to H defined by :

$$C(x) = \mathbb{E}[< X_0, x > X_0], \ x \in H.$$

Cross-covariance operator is :

$$D(x) = \mathbb{E}[< X_0, x > X_1], \ x \in H.$$

The operators C and D satisfy

$$D = \rho C \tag{1.63}$$

Then, to estimate ρ in view of observations $(X_1, \cdots, X_n)$ we start by estimating C and D we put :

$$C_n(x) = \frac{1}{n} \sum_{i=1}^{n} < X_i, x > X_i \ ;$$

and

$$D_n(x) = \frac{1}{n-1} \sum_{i=1}^{n-1} < X_i, x > X_{i+1}.$$

Since C_n is not reversible in general, we are brought to project the observations on space generated by k_n eigenvectors of C, or if they are unknown, on space generated by k_n eigenvectors of C_n.
Let (λ_j) set of eigenvectors of C and H_{k_n} space engendered by $v_1, \cdots, v_{k_n}$ the eigenvectors C, where (k_n) is an integer sequence such that $k_n \leq n, n \geq 1$, and $k_n \longrightarrow \infty$. Assume that :

(i) $\mathbb{E} \mid X_0 \mid^4 < \infty$,
(ii) $\lambda_j > 0$, *for all* $j \geq 1$,
(iii) $\mathbb{P}(< X_0; v_j >= 0) = 0$, *for all* $j \geq 1$.
we distinct two cases :
(a) *If the* (v_j) *are known, we can replace* C_n *by :*

$$\widehat{C}_n = \sum_{j=1}^{\infty} \widehat{\lambda}_{jn} < v_j, . > v_j$$

where $\widehat{\lambda}_{jn} = \frac{1}{n}\sum_{i=1}^{n} < X_i, v_j >^2, j \geq 1, n \geq 1,$
with $\sum_{j=1}^{\infty} \widehat{\lambda}_{jn} = \frac{1}{n}\sum_{j=1}^{\infty}\sum_{i=1}^{n} < X_i, v_j >^2 = \frac{1}{n}\sum_{i=1}^{n} \|X_i\|^2 < \infty.$
which proves that $\widehat{C}_n$ *is nuclear.*
Now $\widehat{C}_n$ *reversible in* H_{k_n}, *then the estimator of* ρ *is :*

$$\widehat{\rho}_n(x) = (\pi^{k_n} D_n \widehat{C}_n^{-1} \pi^{k_n})(x), \ x \in H$$

where $\widehat{C}_n^{-1} = \sum_{j=1}^{k_n} \widehat{\lambda}_{jn}^{-1} < ., v_j > v_j$ *and* π^{k_n} *denotes the orthogonal projector on* H_{k_n}.

(b) *If the* (v_j) *are unknown,* H_{k_n} *is replaced by* $\widetilde{H}_{k_n}$ *space engendered by* $v_{1n}, \cdots, v_{k_n n}$ *the eigenvectors of* C_n. *In this case, we put the following assumptions*
(i_1) $\lambda_1 > \lambda_2 > \cdots > \lambda_j > \cdots > 0$,
(i_2) $\lambda_{k_n n} > 0, n \geq 1$ *(a.s.).*
Now define $\widetilde{C}_n = \widetilde{\pi}^{k_n} C_n = \sum_{j=1}^{k_n} \lambda_{jn} < v_{jn}, . > v_{jn}$ *Then the estimator of* ρ *is written :*

$$\widetilde{\rho}_n(x) = (\widetilde{\pi}^{k_n} D_n \widetilde{C}_n^{-1} \widetilde{\pi}^{k_n})(x), \ x \in H \tag{1.64}$$

where $\widetilde{C}_n^{-1} = \sum_{j=1}^{k_n} \lambda_{jn}^{-1} < ., v_{jn} > v_{jn}$ *and* $\widetilde{\pi}^{k_n}$ *denotes the orthogonal projector of* $\widetilde{H}_{k_n}$.

1.4.1 Asymptotic normality and convergence of operator estimator of ARH(1) : in case $\zeta = (\zeta_t, \ t \in \mathbb{Z})$ are i.i.d.

Let $(X_t, \ t \in \mathbb{Z})$ *an* $ARH(1)$. *We consider the general case where the* (v_j) *are unknown. Bosq has shown the convergence a.s. of estimator (1.64) in linear norm by considering the following notations*

$$A_1 = 2\sqrt{2}(\lambda_1 - \lambda_2) \ if \ \lambda_1 \neq \lambda_2 \ ;$$

and

$$A_j = 2\sqrt{2}\max[(\lambda_{j-1} - \lambda_j)^{-1}, (\lambda_j - \lambda_{j+1})^{-1}].$$

with the following assumption :

(H_1) *X is a standard $ARH(1)$ with $dim H < \infty$,*

(H_2) $\mathbb{E} \parallel X_0 \parallel^4 < \infty$,

(H_3) *(ζ_t) is a strong white noise,*

(H_4) *$C = C_{X_0}$ is a reversible operator,*

(H_5) *C_n is an almost sure reversible operator for $n \geq k$.*

D. Bosq *has found that the convergence is almost sure of the estimator $\widehat{\rho}_n$ of parameter ρ of $ARH(1)$.*

Theorem 1.4.1 ***(D. Bosq*** *(7)) If (H_1), (H_2), (H_3), (H_4) and (H_5) are holds, then*

$$\frac{n^{\frac{1}{4}}}{(\ln n)^{\beta}} \parallel \widehat{\rho}_n - \rho \parallel_{\mathcal{L}} \longrightarrow 0 \ a.s, \ \ \beta > \frac{1}{2} \tag{1.65}$$

(with $\parallel_{\mathcal{L}} = \sup_{\|x\|\leq 1} \parallel l(x) \parallel, \ l \in \mathcal{L} =$ class the bounded operators).

Theorem 1.4.2 ***(D. Bosq*** *(7)) If (H_1), (H_2), (H_3), (H_4), (H_5) holds and $\parallel X_0 \parallel$ is bounded then, for all $\varepsilon_1 > 0$,*

$$\mathbb{P}\left\{\parallel \widehat{\rho}_n - \rho \parallel_{\mathfrak{L}} \geq \varepsilon_1\right\} \leq 16 \exp\left(-\frac{n\lambda_k^2\varepsilon_2^2}{a + b\lambda_k\varepsilon_2}\right) \tag{1.66}$$

with

$$\varepsilon_2 = \min\left(\varepsilon_1, \frac{\varepsilon_1}{\parallel D \parallel_{\mathcal{L}} \parallel C^{-1} \parallel_{\mathcal{L}}}, 2\right), \ a > 0, \ b > 0,$$

depend only on the distribution of the process X.

Theorem 1.4.3 *Assume that $(i), (i_1)$ and (i_2) hold and ρ is Hilbert Schmidt operator. Then if for $\beta > 1$*

$$\lambda_{k_n}^{-1} \sum_{j=1}^{k_n} A_j = O(n^{1/4}(\log n)^{-\beta}).$$

we have :

$$\parallel \widehat{\rho}_n - \rho \parallel_{\mathcal{L}} \longrightarrow 0 \ a.s.$$

Moreover if $\parallel X_0 \parallel$ is bounded, then

$$\mathbb{P}(\parallel \widehat{\rho}_n - \rho \parallel_{\mathcal{L}} \geq \eta) \leq c_1(\eta) \exp\left(-c_2(\eta) n \lambda_{k_n}^2 (\sum_{j=1}^{k_n} A_j)^{-2}\right),$$

where $\eta > 0, n \geq \eta(n)$ and $c_1(\eta), c_2(\eta)$ are positive constants.
Thus

$$\frac{n\lambda_{k_n}^2}{\ln n(\sum_{j=1}^{k_n} A_j)^2} \longrightarrow \infty \Longrightarrow \parallel \widehat{\rho}_n - \rho \parallel_{\mathcal{L}} \longrightarrow 0 \ a.s.$$

Guillas (2001) *has introduced a slight changes on estimator $\widehat{\rho}_n$ to give fast convergence of $\mathbb{E} \parallel \widehat{\rho}_n - \rho \parallel_{\mathcal{L}}^2$. He considered the following assumption*

(H) **:** *There exist a sequence* (A_n) *which satisfies :*

$$\exists 0 < \beta < 1, 0 < A_n \leq \beta \lambda_{k_n}, n \in \mathbb{N}.$$

The sequence (A_n) *allows knowing better the variance of eigenvalues of* C_n *. In fact, if the eigenvalues get close fastly to 0, it becomes difficult to master behavior of* C_n^{-1} *and the speed of convergence degrades. By assumption* (H) *then we use estimator :*

$$\widehat{\rho}_n(x) = (\widetilde{\pi}_{k_n} D_n \widetilde{C}_{n,A}^{-1} \widetilde{\pi}_{k_n}), \ \ x \in H \tag{1.67}$$

where

$$\widetilde{C}_{n,A} = \sum_{j=1}^{\infty} \max(\widetilde{\lambda}_{jn}, A_n) < v_{jn}, . > v_{jn}.$$

Theorem 1.4.4 Guillas (2001) *: Assume that* $(i), (ii)$ *and* (i_1) *hold, there exist* $\alpha > 0, 0 < \beta < 1, \varepsilon < 1/2$ *and* $\tau \geq 1$ *such that*

$$\alpha \frac{\lambda_{k_n}^{\tau}}{n^{\varepsilon}} \leq A_n \leq \beta \lambda_{k_n}$$

then

$$\mathbb{E} \parallel \widehat{\rho}_n - \rho \parallel_{\mathcal{L}}^2 = O\left(\frac{\Lambda_{k_n}^2}{n^{(1-2\varepsilon)} \lambda_{k_n}^{2(1+\tau)}} \right) + O(\lambda_{k_n}^2)$$

where $\Lambda_{k_n} = \sup_{j=1,\ldots,k_n} \frac{1}{\lambda_j - \lambda_{j+1}}$.

Mas (1999) *has established the following result on asymptotic normality estimator of* ρ *:*

Theorem 1.4.5 *Suppose that :*

- (i), (i_1) *and* (i_2) *hold,*
- C_n^{-1} *exist over* $\widetilde{H}_{k_n}$,
- $\mathbb{E} \parallel C^{-1}(\zeta_0) \parallel^2 < \infty$,
- $n\lambda_{k_n}^4 \longrightarrow \infty$, $n^{-1} \sum_{j=1}^{k_n} A_j \lambda_j^{-2} < \infty$,
- $\lambda_j \lambda_{jn}^{-1}$ *in probability is bounded for each* j.

Then :

$$\sqrt{n}(\widehat{\rho}_n - \widetilde{\pi}^{k_n} \rho) \xrightarrow{\mathcal{D}} N,$$

where the limit is taken in $\mathcal{S}$ *and* N *is Gaussian random* $\mathcal{S}$*-valued operator (*$\mathcal{S}$ *: space of Hilbert-Schmidt operators over* H*).*

Chapitre 2

Tail probabilities and complete convergence for weighted sequences of LNQD random variables with application to first-order autoregressive processes model

Abstract

In this chapter, we establish a new concentration inequality and complete convergence of weighted sums for arrays of row-wise linearly negative quadrant dependent (LNQD, in short) random variables and obtain a result dealing with complete convergence of first-order autoregressive processes with identically distributed LNQD innovations.

2.1 Introduction and preliminaries

The concept of complete convergence of a sequence of random variables was introduced by Hsu and Robbins (24) as follows. A sequence $\{X_n, n \geq 1\}$ *of random variables converges completely to the constant* C *if*

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - C| > \varepsilon) < \infty \text{ for all } \varepsilon > 0.$$

By the Borel-Cantelli lemma, this implies $X_n \to C$ *almost surely (a.s.), and the converse implication is true if the* $\{X_n, n \geq 1\}$ *are independent. Hsu and Robbins (24) proved that the sequence of arithmetic means of independent and identically distributed (i.i.d.) random variables converges completely to the expected value if the variance of the summands is finite. Erdö (11) proved the converse. The Hsu and Robbins- Erdös [(11), (24)] result may be formulated as follows. This result has been generalized and extended in several directions and carefully studied by many authors (see, Pruitt (25), Rohatgi (29), Gut (12), Wang and el. (14), Kuczmaszewska and Szynal.(2), Ghosal and Chandra (30), Hu and al (38), Ahmed and al (31)).*

Complete convergence for sequence of random variables plays a central role in the area of limit theorems in probability theory and mathematical statistics. Conditions of independence and identical distribution of random variables are basic in historic results due to Bernoulli, Borel and Kolmogorov. Since then, serious attempts have been made to relax these strong conditions. For example, independence has been relaxed to pairwise independence or pairwise negative quadrant dependence or, even replaced by conditions of dependence such as mixing or martingale. In particular, many authors showed that many results could be obtained by replacing i.i.d. condition by uniformly bounded condition. We recall that an array $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ *of random variables is said to be stochastically dominated by a nonnegative random variable* X *(write* $\{X_{ni} \prec X\}$*) if there exists a constant* $C > 0$ *such that*

$$\mathbb{P}(|X_{ni}| > t) \leq C\mathbb{P}(X > t) \ \forall \ t > 0 \ , \ n \geq 1, \ 1 \leq i \leq n. \tag{2.1}$$

The main purpose of this paper, is to discuss the complete convergence for sums of rowwise linearly negative quadrant dependent (LNQD, in short) random variables under suitable conditions, since independent and identically random variables are a special case of linearly negative quadrant dependent random variables. The exponential inequality plays an important role in various proofs of limit theorems. In particular, it provides a measure of the complete convergence for partial sums. The exponential inequality for negatively associated (NA, in short) random variables has been studied by many authors; see, for example, (12),(24), (38), (15), (35), and so forth. The main purpose of this work is to extend the exponential inequality for NA random variables to the case of LNQD random variables. In addition, we obtain the complete convergence for $S_n = \sum_{i=1}^{n} X_i$*, which improves on the corresponding ones of (12),(24) and (38). Lehmann (17) introduced a simple and natural definition of negative dependence : A sequence* $\{\zeta_i, 1 \leq i \leq n\}$ *of random variables is said to be pairwise negative quadrant dependent (pairwise NQD) if for any real* ϵ_i, ϵ_j *and* $i \neq j, \mathbb{P}(\zeta_i > \epsilon_i, \zeta_j > \epsilon_j) \leq \mathbb{P}(\zeta_i > \epsilon_i)\mathbb{P}(\zeta_j > \epsilon_j)$ *: Much stronger concept than NQD was considered by Joag-Dev and Proschan (15) : A sequence* $\{\zeta_i, 1 \leq i \leq n\}$ *is said to be negatively associated(NA) if for any disjoint subsets,* $A, B \subset \{1, 2, \cdots, n\}$ *and any real coordinatewise increasing functions* f *on* $\mathbb{R}^A$ *and* g *on* $\mathbb{R}^B$*,* $Cov(f(\zeta_i, i \in A), g(\zeta_i, i \in B)) \leq 0$*. Instead of negative association, Newman (22) noticed that his method of proof yielding the central limit theorem for negatively associated sequence requires only that positive linear combinations of the random variables are NQD, i.e., the random variables are linearly negative quadrant dependent (LNQD). This notion of negative dependence was formulated by Newman (22) as follows :* $\{\zeta_i, i \in \mathbb{N}\}$ *is a sequence of LNQD random variables if for any disjoint subsets* A*,* B *of* $\mathbb{N}$ *and positive* r_i*, the random vector* $(\sum_{i\in A} r_i\zeta_i; \sum_{i\in B} r_i\zeta_i)$ *is NQD. Negatively associated sequences are LNQD and LNQD sequences are not necessarily NA, as it can be seen from examples in Newman (22) or Joag-Dev (15).*
We note also that negative association and its weaker concepts are of considerable use in probability and statistics (cf. Joag-Dev and Proschan (15), Newman (22) and the references there in). Newman(22) was first to establish a central limit theorem for LNQD random variables, Zhang (53) proved a functional central limit theorem for LNQD random fields and Kim and al (36) derived a general central limit theorem for weighted sum of LNQD random variables.

Firstly, we will recall the definitions of negatively associated, negative quadrant dependent and linearly negative quadrant dependent sequence.

Definition 2.1.1 *(Cf. Joag-Dev and Prochan (15)). A finite collection of random variables $\zeta_1, \zeta_2, \cdots, \zeta_n$ is said to be negatively associated (NA, in short) if for every pair of disjoint subsets A_1, A_2 of $\{1, 2, \cdots, n\}$*

$$Cov(f(\zeta_i : i \in A_1), g(\zeta_j : j \in A_2)) \leq 0,$$

whenever f and g are coordinatewise nondecreasing such that this covariance exists. An infinite sequence $\{\zeta_n, n \geq 1\}$ is NA if every finite subcollection is NA.

Definition 2.1.2 *(Cf. Lehmann (17)). Two random variables ζ_1 and ζ_2 are said to be negative quadrant dependent (NQD, in short) if for any $\epsilon_1, \epsilon_2 \in \mathbb{R}$,*

$$\mathbb{P}(\zeta_1 < \epsilon_1, \zeta_2 < \epsilon_2) \leq \mathbb{P}(\zeta_1 < \epsilon_1)\mathbb{P}(\zeta_2 < \epsilon_2). \tag{2.2}$$

A sequence $\{\zeta_n, n \geq 1\}$ of random variables is said to be pairwise NQD if ζ_i and ζ_j are NQD for all $i, j \in \mathbb{N}^+$ and $i \neq j$.

Definition 2.1.3 *(Cf. Newman (22)). A sequence $\{\zeta_n, n \geq 1\}$ of random variables is said to be linearly negative quadrant dependent (LNQD, in short) if for any disjoint subsets $A, B \subset \mathbb{Z}$ and positive $r_j^{'}s$,*

$$\sum_{k \in A} r_k \zeta_k \text{ and } \sum_{j \in B} r_j \zeta_j \text{ are NQD.}$$

Remark 2.1.4 *It is easily seen that if $\{\zeta_n, n \geq 1\}$ is a sequence of LNQD random variables, then $\{a\zeta_n + b, n \geq 1\}$ is still a sequence of LNQD random variables, where a and b are real numbers.*

Lemma 2.1.5 *(Cf. Lehmann (17)). Let random variables X and Y be NQD. Then*

(i) $\mathbb{E}(XY) \leq \mathbb{E}(X)\mathbb{E}(Y)$ *;*

(ii) $\mathbb{P}(X > x, Y > y) \leq \mathbb{P}(X > x)\mathbb{P}(Y > y)$ *;*

(iii) If f and g are both nondecreasing (or both nonincreasing) functions, then $f(X)$ and $g(Y)$ are NQD.

The following lemmas play an essential role in our main result.

Lemma 2.1.6 *(Cf. Wang and Zhang (41), Lemma 3.4). Suppose that $\{X_n, n \geq 1\}$ is a LNQD sequence of random variables with $\mathbb{E}(X_n) = 0$. Then for any $p > 1$, there exists a positive constant D such that*

$$\mathbb{E}\left|\sum_{i=1}^{n} X_i\right|^p \leq D\mathbb{E}\left(\sum_{i=1}^{n} X_i^2\right)^{p/2}, \quad n \geq 1 \tag{2.3}$$

Lemma 2.1.7 *Let $\{X_n, n \geq 1\}$ be a sequence of LNQD random variables and $t > 0$, then for each $n \geq 1$,*

$$\mathbb{E}\left[\prod_{i=1}^{n} \exp(tX_i)\right] \leq \prod_{i=1}^{n} \mathbb{E}(\exp(tX_i)) \tag{2.4}$$

Proof 2

For $t > 0$, it is easy to see that tX_i and $t\sum_{j=i+1}^{n} X_j$ are NQD by the definition of LNQD, which implies that $\exp(tX_i)$ and $\exp(t\sum_{j=i+1}^{n} X_j)$ are also NQD for $i = 1, 2, \cdots, n-1$ by Lemma2.1.5(iii). It follows from Lemma2.1.5(i) and induction that

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{n} \exp(tX_i)\right] &= \mathbb{E}\left[\exp(tX_1)\exp\left(\sum_{i=2}^{n} tX_i\right)\right] \\
&\leq \mathbb{E}\left[\exp(tX_1)\right]\mathbb{E}\left[\exp\left(\sum_{i=2}^{n} tX_i\right)\right] \\
&= \mathbb{E}\left[\exp(tX_1)\right]\mathbb{E}\left[\exp(tX_2)\exp\left(\sum_{i=3}^{n} tX_i\right)\right] \\
&\leq \mathbb{E}\left[\exp(tX_1)\right]\mathbb{E}\left[\exp(tX_2)\right]\mathbb{E}\left[\exp\left(\sum_{i=3}^{n} tX_i\right)\right] \\
&\leq \prod_{i=1}^{n} \mathbb{E}(\exp(tX_i)).
\end{aligned}$$

This completes the proof of the lemma.
Throughout the paper, let $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ be a sequence of random variables defined on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote $S_n = \sum_{i=1}^{n} X_{ni}$ and $B_n^2 = \sum_{i=1}^{n} \mathbb{E}(X_{ni}^2)$ for each $1 \leq i \leq n$ and $n \geq 1$.

2.2 Main results and their proofs

With the preliminary lemmas, we now state and prove our main result. In this paper we consider arrays of random variables.
Let $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ be a sequence of random variables and some strictly increasing sequence $\{c_n, n \geq 1\}$ of constants positive. Define for $1 \leq i \leq n, n \geq 1$

Lemma 2.2.1 *Let $\alpha > 0$ constants and $0 < \beta \leq \dfrac{\alpha^2}{e^\alpha - 1 - \alpha}$. Then*

$$\exp(x) - 1 - x \leq \frac{x^2}{\beta} \tag{2.5}$$

for all $0 \leq x \leq \alpha$

Proof 3

Consider the function

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

We need to prove that $\Psi(x, \beta) \geq 0$ *for all* $0 < \beta \leq \dfrac{\alpha^2}{e^\alpha - 1 - \alpha}$ *and* $0 \leq x \leq \alpha$.
Take the derivative

$$\frac{\partial \Psi(x, \beta)}{\partial x} = -\frac{x(x - (2 - \beta))}{\beta(1 + x + \dfrac{x^2}{\beta})}.$$

Hence, Ψ *is increasing in* x *on the interval* $(0, 2 - \beta)$ *and decreasing on the interval* $(2 - \beta, \alpha)$. *Note that* $\Psi(0, \beta) = 0$ *and* $\Psi(\alpha, \beta) \geq 0$ *since* $0 < \beta \leq \dfrac{\alpha^2}{e^\alpha - 1 - \alpha}$

Let

$$\begin{aligned} X_{1,ni} &= -a_n \mathbb{I}_{\{X_{ni} < -a_n\}} + X_{ni} \mathbb{I}_{\{|X_{ni}| \leq a_n\}} + a_n \mathbb{I}_{\{X_{ni} > a_n\}}, \\ X_{2,ni} &= (X_{ni} - a_n) \mathbb{I}_{\{X_{ni} > a_n\}}, \ X_{3,ni} = (X_{ni} + a_n) \mathbb{I}_{\{X_{ni} < -a_n\}}. \end{aligned} \tag{2.6}$$

Here, and in the sequel, $\mathbb{I}_A$ *denotes the indicator function of the* A *set in the braces, that is, it takes value* 1 *or* 0 *according to whether or not the sample point belongs to the set.*
It is easy to check that $X_{1,ni} + X_{2,ni} + X_{3,ni} = X_{ni}$ *for* $1 \leq i \leq n, n \geq 1$ *and* $X_{1,n1}, X_{1,n2}, \cdots, X_{1,nn}$ *are bounded by* a_n *for each fixed* $n \geq 1$.
If $\{X_{ni}, n \geq 1\}$ *are LNQD random variables, then* $\{X_{p,ni}, 1 \leq i \leq n\}$, $p = 1, 2, 3$ *are also LNQD random variables for each fixed* $n \geq 1$

Theorem 2.2.2 *Let* $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ *be a sequence of row-wise identically distributed LNQD random variables and* $\{X_{p,ni}, 1 \leq i \leq n, n \geq 1\}$, $p = 2, 3$, *be defined by (2.6). Assume that there exists a* $\tau > 0$ *satisfying* $\sup_{|\mu| \leq \tau} \mathbb{E}(e^{\mu X_{11}}) \leq A_\tau < \infty$, *where* A_τ *is a positive constant depending only on* τ, $\sum_{i=1}^{n} b_{ni}^2 = O((\log n)^{-1})$.
Then for any $\varepsilon > 0$ *and* $\mu \in (0, \tau]$,

$$\mathbb{P}\left(|\sum_{i=1}^{n} b_{ni}(X_{p,ni} - \mathbb{E}X_{p,ni})| \geq \varepsilon \right) \leq \Phi(\mu, \varepsilon, \tau, a) \frac{1}{n^{a/2} \log n}, \ p = 2, 3. \tag{2.7}$$

Where $\Phi(\mu, \varepsilon, \tau, a) = \dfrac{2^{a+1} a^a e^{-a} D D' A_\tau}{\mu^{2+a} K_1^a (\mathbb{E}(X_{11})^2)^{a/2} \varepsilon^2}$, *choosing* $a > 2$

Proof 4

For $p=2$*, by Markov's inequality and lemma*2.1.6*, we can see that*

$$\begin{aligned}
\mathbb{P}\left(|\sum_{i=1}^{n} b_{ni}(X_{2,ni}-\mathbb{E}X_{2,ni})|\geq\varepsilon\right) &\leq \frac{1}{\varepsilon^2}\mathbb{E}|\sum_{i=1}^{n} b_{ni}(X_{2,ni}-\mathbb{E}X_{2,ni})|^2 \\
&\leq \frac{D}{\varepsilon^2}\mathbb{E}(\sum_{i=1}^{n} b_{ni}^2(X_{2,ni}-\mathbb{E}X_{2,ni})^2) \\
&= \frac{D}{\varepsilon^2}\sum_{i=1}^{n}\mathbb{E}(b_{ni}(X_{2,ni}-\mathbb{E}X_{2,ni}))^2 \\
&= \frac{D}{\varepsilon^2}\sum_{i=1}^{n} b_{ni}^2(\mathbb{E}(X_{2,ni})^2-(\mathbb{E}X_{2,ni})^2) \\
&\leq \frac{D}{\varepsilon^2}\sum_{i=1}^{n} b_{ni}^2\mathbb{E}(X_{2,ni})^2 \\
&= \frac{D}{\varepsilon^2}\mathbb{E}\left(X_{11}-\frac{K_1B_n}{2}\right)^2\mathbb{I}_{\{X_{11}>\frac{K_1B_n}{2}\}}\left(\frac{D'}{\log n}\right) \\
&= \frac{D}{\varepsilon^2}\left[-\int_{\frac{K_1B_n}{2}}^{+\infty}\left(x_{11}-\frac{K_1B_n}{2}\right)^2\, d\overline{F}_{X_{11}}(x_{11})\right]\left(\frac{D'}{\log n}\right) \\
&,\quad \overline{F}_{X_{11}}(x_{11})=\mathbb{P}(X_{11}>x_{11}) \\
&= -\frac{D}{\varepsilon^2}\left|(x_{11}-\frac{K_1B_n}{2})^2\overline{F}_{X_{11}}(x_{11})\right|_{\frac{K_1B_n}{2}}^{+\infty}\left(\frac{D'}{\log n}\right) \\
&+ \frac{D}{\varepsilon^2}\int_{\frac{K_1B_n}{2}}^{+\infty} 2\left(x_{11}-\frac{K_1B_n}{2}\right)\overline{F}_{X_{11}}(x_{11})\, dx_{11}\left(\frac{D'}{\log n}\right) \\
&\quad \textit{(integrating by parts)} \\
&\leq \frac{D}{\varepsilon^2}\int_{\frac{K_1B_n}{2}}^{+\infty} 2\left(x_{11}-\frac{K_1B_n}{2}\right)\overline{F}_{X_{11}}(x_{11})\, dx_{11}\left(\frac{D'}{\log n}\right) \\
&\leq \frac{2DA_\tau}{\varepsilon^2\mu}\int_{\frac{K_1B_n}{2}}^{+\infty}\left(x_{11}-\frac{K_1B_n}{2}\right)e^{-\mu x_{11}}\, dx_{11}\left(\frac{D'}{\log n}\right) \\
&\quad \textit{(by Markov's inequality)} \\
&= \frac{2DA_\tau}{\varepsilon^2\mu}\int_{\frac{K_1B_n}{2}}^{+\infty} e^{-\mu x_{11}}\, dx_{11}\left(\frac{D'}{\log n}\right) \\
&= \frac{2DA_\tau}{\varepsilon^2\mu^2}e^{-\mu\frac{K_1\sqrt{n}(\mathbb{E}(X_{11}^2))^{\frac{1}{2}}}{2}}\left(\frac{D'}{\log n}\right) \\
&\leq \frac{2^{a+1}a^a e^{-a}DD'A_\tau}{\mu^{2+a}K_1^a(\mathbb{E}(X_{11})^2)^{a/2}\varepsilon^2}\frac{1}{n^{a/2}\log n}
\end{aligned}$$

(using the elementary inequality, $\exp(-y)\leq\left(\frac{a}{ey}\right)^a$ *for all* $a>0$*, choosing* $a>2$*).*
For $p=3$*, the proof is similar to the case* $p=2$ *and is omitted.*

Lemma 2.2.3 *If* $\{X_n, n \geq 1\}$ *is a sequence of LNQD random variables and* $\{f_n, n \geq 1\}$ *is a sequence of Borel functions all of which are monotone increasing (or all monotone decreasing), then* $\{f_n(X_n), n \geq 1\}$ *is a sequence of LNQD random variables.*

Lemma 2.2.4 *Let* $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ *be an array of rowwise LNQD random variables with* $\mathbb{E}X_{ni} = 0$*, and* $\{a_n, n \geq 1\}$ *a sequence of positive constants. Suppose that*

(i) $\sum_{n=1}^{\infty} \exp\{-\frac{\beta\varepsilon^2}{4a_n}\} < \infty$ *for some* $0 < \beta \leq \frac{\alpha^2}{e^\alpha - 1 - \alpha}$ *and* $|X_{ni}| \leq \alpha$.

(ii) $\sum_{i=1}^{n} \mathbb{E}(X_{ni}^2) = O(a_n)$,

Then $\sum_{i=1}^{n} X_{ni}$ *converges completely to zero.*

Proof 5

From the inequality $\exp(x) \leq 1 + x + \frac{x^2}{\beta}$ *for all* $0 \leq x \leq \alpha$ *and* $0 < \beta \leq \frac{\alpha^2}{e^\alpha - 1 - \alpha}$ *(see lemma 2.2.1), we have by (i) that for any* $\lambda > 0$

$$\begin{aligned} \mathbb{E}\exp(\lambda X_{ni}) &\leq \mathbb{E}\left\{1 + \lambda X_{ni} + \frac{1}{\beta}\lambda^2|X_{ni}|^2\right\} \\ &= 1 + \frac{1}{\beta}\lambda^2\mathbb{E}|X_{ni}|^2 \\ &\leq \exp\left\{\frac{1}{\beta}\lambda^2\mathbb{E}|X_{ni}|^2\right\} \end{aligned}$$

The second inequality follows by the fact that $1 + t \leq e^t$ *for all real number* t*. It follows by Markov's inequality, Lemma 2.1.7, and (i) that for any* $\lambda > 0$,

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{n} X_{ni} > \varepsilon\right) &\leq e^{-\lambda\varepsilon}\mathbb{E}\exp(\lambda\sum_{i=1}^{n} X_{ni}) \\ &\leq e^{-\lambda\varepsilon}\prod_{i=1}^{n}\mathbb{E}\exp(\lambda X_{ni}) \\ &\leq e^{-\lambda\varepsilon}\exp\left\{\frac{1}{\beta}\lambda^2\sum_{i=1}^{n}\mathbb{E}|X_{ni}|^2\right\} \\ &\leq e^{-\lambda\varepsilon}\exp\left\{\frac{1}{\beta}\lambda^2 O(a_n)\right\} \\ &= \exp\left\{-\lambda\varepsilon + \frac{1}{\beta}\lambda^2 O(a_n)\right\}. \end{aligned}$$

Choosing $\lambda = \dfrac{\varepsilon\beta}{2O(a_n)}$*, we have that for all large* n*,*

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{n} X_{ni} > \varepsilon\right) &\leq \exp\left\{-\varepsilon^2 \frac{\beta}{2O(a_n)} + \frac{\varepsilon^2\beta^2}{4\beta(O(a_n))^2} O(a_n)\right\} \\ &= \exp\left\{-\varepsilon^2 \frac{\beta}{4O(a_n)}\right\} \\ &\leq \exp\left\{-\varepsilon^2 \frac{\beta}{4a_n}\right\}.\end{aligned}$$

Thus by (i)

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_{ni} > \varepsilon\right) < \infty. \tag{2.8}$$

Since $\{-X_{ni}, 1 \leq i \leq n, n \geq 1\}$ *is still an array of row-wise LNQD random variables by Lemma 2.2.3, we can replace* X_{ni} *by* $-X_{ni}$ *from the above statement. That is,*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_{ni} < -\varepsilon\right) < \infty. \tag{2.9}$$

The result follows by (2.8) and (2.9).
Now we state and prove our main result.

Theorem 2.2.5 *Let* $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ *be an array of row-wise identically distributed LNQD random variables such that* $\mathbb{E}X_{ni} = 0$ *satisfying* $\mathbb{E}|X_{11}|^{\gamma+1} < \infty$ *for some* $\gamma \geq 1$*. Assume that* $\{b_{ni}, 1 \leq i \leq n, n \geq 1\}$ *is an array of constants satisfying*

$$\max_{1\leq i\leq n} |b_{ni}| = O(c_n^{-\delta}), 0 < c_n \uparrow \infty, \text{ for all any } \delta > 0 \text{ and } a_n c_n^{-\delta} \leq 1, \tag{2.10}$$

$$\sum_{i=1}^{n} b_{ni}^2 = O\left(\frac{1}{\log n}\right). \tag{2.11}$$

Then $\sum_{i=1}^{n} b_{ni}X_{ni}$ *converges completely to zero.*

Proof 6

Note that $b_{ni} = b_{ni}^+ - b_{ni}^-$ *, where* $b_{ni}^+ = \max\{0, b_{ni}\}$ *and* $b_{ni}^- = \max\{0, -b_{ni}\}$*. To prove the result, it suffices to show that, for any* $\varepsilon > 0$*,*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\Big|\sum_{i=1}^{n} b_{ni}^+ X_{ni}\Big| > \varepsilon\right) < \infty$$

and

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\Big|\sum_{i=1}^{n} b_{ni}^- X_{ni}\Big| > \varepsilon\right) < \infty.$$

So, without loss of generality, we may assume that $b_{ni} > 0$*. Moreover, we may assume that* $\max_{1\le i\le n} b_{ni} \le c_n^{-\delta}$*. By Remark 2.1.4 and (2.6) , for any* $\varepsilon > 0$ *and* $\delta > 0$*, we have*

$$\begin{aligned}\sum_{n=1}^{\infty}\mathbb{P}\left(\Big|\sum_{i=1}^{n}b_{ni}X_{ni}\Big|>\varepsilon\right) &\le \sum_{n=1}^{\infty}\mathbb{P}\left(\Big|\sum_{i=1}^{n}b_{ni}X_{1,ni}\Big|>\varepsilon/3\right)+\sum_{n=1}^{\infty}\mathbb{P}\left(\Big|\sum_{i=1}^{n}b_{ni}X_{2,ni}\Big|>\varepsilon/3\right)\\ &+ \sum_{n=1}^{\infty}\mathbb{P}\left(\Big|\sum_{i=1}^{n}b_{ni}X_{3,ni}\Big|>\varepsilon/3\right)\\ &= I_1+I_2+I_3.\end{aligned}$$

Since $\{X_{1,ni}, 1 \le i \le n\}$ *are monotone transformations of* $\{X_{ni}, 1 \le i \le n\}$, $\{X_{1,ni}, 1 \le i \le n\}$ *is an array of row-wise LNQD random variables by Lemma 2.2.3.Moreover,* $\{b_{ni}X_{1,ni}, 1 \le i \le n\}$ *is also an array of rowwise LNQD random variables, since* $b_{ni} > 0$*. We first show that* $I_1 < \infty$*. To do this, we will apply Lemma 2.2.4 to the array* $\{b_{ni}(X_{1,ni} - \mathbb{E}X_{1,ni}), 1 \le i \le n, n \ge 1\}$ *and sequence* $\{1/\log n, n \ge 1\}$*. Noting that* $|b_{ni}(X_{1,ni} - \mathbb{E}X_{1,ni})| \le 2\dfrac{a_n}{c_n^{\delta}} \le 2$ *since* $a_n c_n^{-\delta} \le 1$*, condition (i) holds trivially by choosing* $\alpha \ge 2$ *and* $\varepsilon > \dfrac{2}{\sqrt{\beta}}$*. By Remark 2.1.4 and (2.11),*

$$\begin{aligned}\mathbb{E}\Big|\sum_{i=1}^{n}b_{ni}(X_{1,ni}-\mathbb{E}X_{1,ni})\Big|^2 &\le \mathbb{E}\left[\sum_{i=1}^{n}(b_{ni}(X_{1,ni}-\mathbb{E}X_{1,ni}))^2\right]\\ &+ 2\sum_{1\le i<j\le n} b_{ni}b_{nj}\mathbb{E}[(X_{1,ni}-\mathbb{E}X_{1,ni})(X_{1,nj}-\mathbb{E}X_{1,nj})]\\ &\le \sum_{i=1}^{n}\mathbb{E}(b_{ni}(X_{1,ni}-\mathbb{E}X_{1,ni}))^2\\ &+ 2\sum_{1\le i<j\le n}\mathbb{E}(b_{ni}(X_{1,ni}-\mathbb{E}X_{1,ni}))\mathbb{E}(b_{nj}(X_{1,nj}-\mathbb{E}X_{1,nj}))\\ &= \sum_{i=1}^{n}\mathbb{E}(b_{ni}(X_{1,ni}-\mathbb{E}X_{1,ni}))^2\\ &\le \sum_{i=1}^{n}b_{ni}^2\mathbb{E}(X_{1,ni}^2)\\ &= \sum_{i=1}^{n}b_{ni}^2\mathbb{E}(X_{11}^2)\\ &\le \mathbb{E}(X_{11}^2)D'\left(\frac{1}{\log n}\right).\end{aligned}$$

Hence condition (ii) holds. Thus by Lemma 2.2.4

$$\sum_{n=1}^{\infty}\mathbb{P}\left(\Big|\sum_{i=1}^{n}b_{ni}(X_{1,ni}-\mathbb{E}X_{1,ni})\Big|>\frac{\varepsilon}{3}\right)<\infty.$$

To show that $I_1 < \infty$, it remains to show that

$$\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{1,ni} \to 0 \ \text{as } n \to \infty. \tag{2.12}$$

In view of $\mathbb{E}X_{ni} = 0$,

$$\mathbb{E}|X_{ni}|^{\gamma+1}\mathbb{I}_{\left\{|X_{ni}|>\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}} \geq \left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)^{\gamma}\mathbb{E}|X_{ni}|\mathbb{I}_{\left\{|X_{ni}|>\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}}$$

and Markov's inequality, we obtain

$$\begin{aligned}
\left|\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{1,ni}\right| &= \Big|\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{1,ni}\mathbb{I}_{\left\{|X_{ni}|\leq\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}} \\
&+ \sum_{i=1}^{n} b_{ni}\mathbb{E}X_{1,ni}\mathbb{I}_{\left\{|X_{ni}|>\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}}\Big| \\
&\leq \Big|\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{1,ni}\mathbb{I}_{\left\{|X_{ni}|\leq\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}}\Big| \\
&+ \Big|\sum_{i=1}^{n} b_{ni}\left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{ni}^2}}{2}\right)\mathbb{E}\ \mathbb{I}_{\left\{|X_{ni}|>\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}}\Big| \\
&= \Big|\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{1,ni}\mathbb{I}_{\left\{|X_{ni}|\leq\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}}\Big| \\
&+ \left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)\sum_{i=1}^{n}|b_{ni}|\mathbb{P}\left(|X_{ni}| > \frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right) \\
&\leq \frac{1}{\left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)^{\gamma}}\sum_{i=1}^{n}|b_{ni}|\mathbb{E}|X_{ni}|^{\gamma+1}\mathbb{I}_{\left\{|X_{ni}|>\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right\}} \\
&+ \left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)\frac{1}{\left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)^{\gamma+1}}\sum_{i=1}^{n}|b_{ni}|\mathbb{E}|X_{ni}|^{\gamma+1}
\end{aligned}$$

$$\leq \frac{2}{\left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)^{\gamma}} \sum_{i=1}^{n} |b_{ni}|\mathbb{E}|X_{ni}|^{\gamma+1}$$

$$\leq \frac{2}{\left(\frac{K_1\sqrt{n}\sqrt{\mathbb{E}X_{11}^2}}{2}\right)^{\gamma}} \mathbb{E}|X_{11}|^{\gamma+1}(\sum_{i=1}^{n} b_{ni}^2)^{1/2} n^{1/2} \text{ (by Hölder's inequality)}$$

$$\leq \frac{2^{\gamma+1}C\mathbb{E}|X_{11}|^{\gamma+1}}{K_1^{\gamma}(\mathbb{E}X_{11}^2)^{\gamma/2}} \frac{1}{n^{\gamma/2-1/2}(\log n)^{1/2}} \to 0$$

as $n \to \infty$, since $\gamma \geq 1$. Hence (2.12) holds.
Finally, we show that $I_2 < \infty$ and $I_3 < \infty$. Thus by theorem 2.2.2

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\sum_{i=1}^{n} b_{ni}(X_{2,ni} - \mathbb{E}X_{2,ni})\right| > \frac{\varepsilon}{3}\right) < \infty.$$

To show that $I_2 < \infty$, it remains to show that

$$\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{2,ni} \to 0 \text{ as } n \to \infty. \tag{2.13}$$

In view of elementary inequality

$$\left|\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{2,ni}\right| \leq \sum_{i=1}^{n} |b_{ni}|\mathbb{E}|X_{2,ni}|$$

and Cauchy-Schwarz's inequality

$$\begin{aligned} \mathbb{E}|X_{2,ni}| &\leq (\mathbb{E}|X_{2,ni}|^2)^{1/2} \\ &\leq \mathbb{E}X_{2,ni}^2. \end{aligned}$$

We obtain

$$\begin{aligned} \left|\sum_{i=1}^{n} b_{ni}\mathbb{E}X_{2,ni}\right| &\leq \sum_{i=1}^{n} |b_{ni}| D A_\tau e^{-\frac{\mu\sqrt{n}(\mathbb{E}X_{11}^2)^{1/2}K_1}{2}} \\ &\leq \max_{1\leq i\leq n} |b_{ni}| n D A_\tau e^{-\frac{\mu\sqrt{n}(\mathbb{E}X_{11}^2)^{1/2}K_1}{2}} \\ &\leq K_2 c_n^{-\delta} e^{-\frac{\mu\sqrt{n}(\mathbb{E}X_{11}^2)^{1/2}K_1}{2}} \\ &\leq \frac{K_2 4^6 n D A_\tau}{c_n^{\delta}\mu^4 e^4 n^2(\mathbb{E}X_{11}^2)^2 K_1^4} \to 0 \quad \text{as } n \to \infty, \end{aligned}$$

(by the inequality $e^{-y} \leq \left(\frac{a}{ey}\right)^a$, choosing $a = 4$), since $\delta > 0$. Hence (2.13) holds.
Similar to I_2, we have $I_3 < \infty$.
We complete the proof of the theorem.

2.3 Applications to the results to AR(1) model

The basic object of this section is applying the results to first-order autoregressive processes(AR(1)).

2.3.1 The AR(1) model

We consider an autoregressive time series of first order AR(1) defined by

$$X_{n+1} = \theta X_n + \zeta_{n+1}, \;\; n = 1, 2, \cdots, \tag{2.14}$$

where $\{\zeta_n, n \geq 0\}$ is a sequence of identically distributed LNQD random variables with $\zeta_0 = X_0 = 0$, $0 < \mathbb{E}\zeta_k^4 < \infty$, $k = 1, 2, \cdots$ and where θ is a parameter with $|\theta| < 1$. Here, we can rewrite X_{n+1} in (2.14) as follows :

$$X_{n+1} = \theta^{n+1} X_0 + \theta^n \zeta_1 + \theta^{n-1}\zeta_2 + \cdots + \zeta_{n+1}. \tag{2.15}$$

The coefficient θ is fitted least squares, giving the estimator

$$\widehat{\theta}_n = \frac{\sum_{j=1}^{n} X_j X_{j-1}}{\sum_{j=1}^{n} X_{j-1}^2} \tag{2.16}$$

It immediately follows from (2.14) and (2.16) that

$$\widehat{\theta}_n - \theta = \frac{\sum_{j=1}^{n} \zeta_j X_{j-1}}{\sum_{j=1}^{n} X_{j-1}^2} \tag{2.17}$$

We start with the following basic lemma.

Lemma 2.3.1 *If $\{\zeta_n, n \geq 1\}$ is a sequence of identically distributed LNQD random variables such that $|\zeta_1|^4 < \alpha$, then for any $R > 0$ real, $\widetilde{\varepsilon} > \dfrac{\mathbb{E}\zeta_1^2}{R^2}$ and $0 < \beta < \dfrac{\alpha}{e^{\alpha} - \alpha - 1}$*

$$\mathbb{P}\Big(\Big|\sum_{j=1}^{n}(\zeta_j^2 - \mathbb{E}\zeta_j^2)\Big| \geq (R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)n\Big) \leq 2\exp\left\{-\beta\frac{(R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)^2 n}{36}\right\} + 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon}, \tau, a)}{n^{a/2+1}}. \tag{2.18}$$

Where

$$\widetilde{\Phi}(\widetilde{\varepsilon}, \tau, a) = 9\frac{2^{a+1}a^a e^{-a} D D' A_\tau}{\mu^{a+2} K_1^a (\mathbb{E}\zeta_1^4)^{a/2} (R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)^2}$$

Proof 7

$X_{nj} = \zeta_j^2$. Then $\mathbb{E}|X_{nj}|^2 < \alpha^2 < +\infty$. Further let $b_{nj} = \frac{1}{n}$ for $n \geq 1$. Then $\{b_{nj}\}$ satisfies conditions (2.10) and (2.11). So the result follows by theorem 2.2.2 and lemma 2.2.4, since

$$\begin{aligned}
\mathbb{P}\Big(\Big|\sum_{j=1}^{n}(\zeta_j^2 - \mathbb{E}\zeta_j^2)\Big| \geq (R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)n\Big) &\leq \mathbb{P}\Big(\Big|\sum_{j=1}^{n}(\zeta_{1,j}^2 - \mathbb{E}\zeta_{1,j}^2)\Big| \geq \frac{(R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)n}{3}\Big) \\
&+ \mathbb{P}\Big(\Big|\sum_{j=1}^{n}(\zeta_{2,j}^2 - \mathbb{E}\zeta_{2,j}^2)\Big| \geq \frac{(R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)n}{3}\Big) \\
&+ \mathbb{P}\Big(\Big|\sum_{j=1}^{n}(\zeta_{3,j}^2 - \mathbb{E}\zeta_{3,j}^2)\Big| \geq \frac{(R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)n}{3}\Big) \\
&\leq 2\mathbb{P}\Big(\frac{1}{n}\sum_{j=1}^{n}\widetilde{\zeta_{1,j}} \geq \frac{(R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)}{3}\Big) + 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon},\tau,a)}{n^{a/2+1}} \\
&\quad \text{where } \widetilde{\zeta_{1,j}} = \zeta_{1,j}^2 - \mathbb{E}\zeta_{1,j}^2 \\
&\leq 2\exp\{-\beta\frac{(R^2\widetilde{\varepsilon} - \mathbb{E}\zeta_1^2)^2 n}{36}\} + 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon},\tau,a)}{n^{a/2+1}}.
\end{aligned}$$

Theorem 2.3.2 *Let the conditions of lemma 2.3.1 be satisfied then for any $\frac{(\mathbb{E}\zeta_1^2)^{1/2}}{R^2} < \widetilde{\varepsilon}$ positive, we have*

$$\begin{aligned}
\mathbb{P}(\sqrt{n}|\widehat{\theta}_n - \theta| > R) &\leq 2\exp\{-\beta\frac{(R^2\widetilde{\varepsilon}^2 - \mathbb{E}\zeta_1^2)^2 n}{36}\} + 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon}^2,\tau,a)}{n^{a/2+1}} \\
&+ \exp\{-\frac{1}{2}n\frac{(T_1 - n\widetilde{\varepsilon}^2)^2}{T_2}\}
\end{aligned} \tag{2.19}$$

where $T_1 = \mathbb{E}(X_i^2) < \infty, T_2 = \mathbb{E}(X_i^4) < \infty$.

Proof 8

Firstly, we notice that :

$$\widehat{\theta}_n - \theta = \frac{\sum_{j=1}^{n}\zeta_j X_{j-1}}{\sum_{j=1}^{n}X_{j-1}^2}.$$

It follows that

$$\mathbb{P}(\sqrt{n}|\widehat{\theta}_n - \theta| > R) = \mathbb{P}\left(\left|\frac{1/\sqrt{n}\sum_{j=1}^{n}\zeta_j X_{j-1}}{1/n\sum_{j=1}^{n}X_{j-1}^2}\right| > R\right)$$

By virtue of the probability properties and Hölder's inequality, we have for any $\widetilde{\varepsilon}$ positive

$$
\begin{aligned}
\mathbb{P}(\sqrt{n}|\widehat{\theta}_n - \theta| > R) &\leq \mathbb{P}(\frac{1}{n}\sum_{j=1}^{n}\zeta_j^2 \geq R^2\widetilde{\varepsilon}^2) + \mathbb{P}(\frac{1}{n^2}\sum_{j=1}^{n}X_{j-1}^2 \leq \widetilde{\varepsilon}^2) \\
&= \mathbb{P}(\sum_{j=1}^{n}\zeta_j^2 \geq (R^2\widetilde{\varepsilon}^2)n) + \mathbb{P}(\sum_{j=1}^{n}X_{j-1}^2 \leq n^2\widetilde{\varepsilon}^2) \\
&= I_{1n} + I_{2n}.
\end{aligned}
$$

Next we estimate I_{1n} and I_{2n}.

$$
\begin{aligned}
I_{1n} &= \mathbb{P}(\sum_{j=1}^{n}\zeta_j^2 \geq (R^2\widetilde{\varepsilon}^2)n) \\
&= \mathbb{P}(\sum_{j=1}^{n}(\zeta_j^2 - \mathbb{E}\zeta_j^2 + \mathbb{E}\zeta_j^2) \geq (R^2\widetilde{\varepsilon}^2)n) \\
&= \mathbb{P}(\sum_{j=1}^{n}(\zeta_j^2 - \mathbb{E}\zeta_j^2) \geq (R^2\widetilde{\varepsilon}^2 - \mathbb{E}\zeta_1^2)n) \\
&\leq \mathbb{P}(\Big|\sum_{j=1}^{n}(\zeta_j^2 - \mathbb{E}\zeta_j^2)\Big| \geq (R^2\widetilde{\varepsilon}^2 - \mathbb{E}\zeta_1^2)n) \\
&\leq 2\exp\left\{-\beta\frac{(R^2\widetilde{\varepsilon}^2 - \mathbb{E}\zeta_1^2)^2 n}{36}\right\} \\
&+ 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon}^2, \tau, a)}{n^{a/2+1}} \text{ (by lemma 2.3.1)}. \qquad (2.20)
\end{aligned}
$$

We will bound now, the second probability of the right-hand side of the expression I_{2n}.According to the Markov's inequality, it follows for any λ positive

$$
\begin{aligned}
I_{2n} &= \mathbb{P}\left(\frac{1}{n^2}\sum_{j=1}^{n}X_{j-1}^2 \leq \widetilde{\varepsilon}^2\right) \\
&= \mathbb{P}\left(n^2\widetilde{\varepsilon}^2 - \sum_{j=1}^{n}X_{j-1}^2 \geq 0\right) \\
&= \mathbb{E}\left(\mathbb{I}_{\{n\epsilon^2 - \sum_{j=1}^{n}X_{j-1}^2 \geq 0\}}\right) \\
&\leq \mathbb{E}\left(\exp\lambda\left(n^2\widetilde{\varepsilon}^2 - \sum_{j=1}^{n}X_{j-1}^2\right)\right) \qquad (\lambda > 0) \\
&\leq e^{\lambda n^2\widetilde{\varepsilon}^2}\mathbb{E}\left(\exp -\lambda\sum_{j=1}^{n}X_{j-1}^2\right) \\
&\leq e^{\lambda n^2\widetilde{\varepsilon}^2}\prod_{j=1}^{n}\mathbb{E}\left(\exp -\lambda X_{j-1}^2\right).
\end{aligned}
$$

Since
$$I_{2n} \le e^{\lambda n^2 \widetilde{\varepsilon}^2} \prod_{j=1}^{n} \mathbb{E}\left(\exp -\lambda X_{j-1}^2\right).$$
we put
$$\widetilde{X} = \lambda X_{j-1}^2,$$
we first claim that for $x \ge 0$
$$e^{-x} \le 1 - x + \frac{1}{2}x^2. \tag{2.21}$$
To see this let $\psi(x) = e^{-x}$ *and* $\phi(x) = 1 - x + \frac{1}{2}x^2$, $(\psi'(x) = -e^{-x})$ *and recall that for every* x
$$e^x \ge 1 + x \quad \forall x, \tag{2.22}$$
so that $\psi'(x) = -e^{-x} \le -1 + x = \phi'(x)$. *Since* $\psi(0) = 1 = \phi(0)$ *this implies* $\psi(x) \le \phi(x)$ *for all* $x \ge 0$ *and* (2.21) *is claimed.*
From (2.21) *and* (2.22) *it follows that for* $\lambda > 0$
$$\begin{aligned} e^{\lambda n \epsilon^2} \prod_{j=1}^{n} \mathbb{E}\left(\exp(-\lambda X_{j-1}^2)\right) &\le e^{\lambda n^2 \widetilde{\varepsilon}^2} \left(1 - \lambda T_1 + \frac{\lambda^2}{2} T_2\right)^n \\ &\le e^{\lambda n^2 \widetilde{\varepsilon}^2} \left(\exp\left(-\lambda T_1 + \frac{\lambda^2}{2} T_2\right)\right)^n \\ &\le e^{\lambda n^2 \widetilde{\varepsilon}^2} \exp\left(-n\lambda T_1 + \frac{\lambda^2}{2} n T_2\right) \end{aligned}$$
where $T_1 = \mathbb{E}(X_j^2) < \infty, T_2 = \mathbb{E}(X_j^4) < \infty$.
Hence
$$I_{2n} = \mathbb{P}\left(\sum_{i=1}^{n} X_{j-1}^2 \le n^2 \widetilde{\varepsilon}^2\right) \le \exp\left[\lambda\left(n^2 \widetilde{\varepsilon}^2 - nT_1\right) + \frac{n\lambda^2 T_2}{2}\right]. \tag{2.23}$$

With $h(\lambda) = \lambda(n^2\widetilde{\varepsilon}^2 - nT_1) + \frac{n\lambda^2 T_2}{2}$ *and* $\lambda > 0$, *the equation* $h'(\lambda) = 0$ *has the unique solution* $\lambda = \dfrac{T_1 - n\widetilde{\varepsilon}^2}{T_2}$ *which minimize* $h(\lambda)$. *Hence*
$$\mathbb{P}\left(\sum_{j=1}^{n} X_{j-1}^2 \le n^2 \widetilde{\varepsilon}^2\right) \le \exp\left\{-\frac{1}{2} n \frac{(T_1 - n\widetilde{\varepsilon}^2)^2}{T_2}\right\} \tag{2.24}$$
Then for every $R > 0$, $T_1 < \infty$, $T_2 < \infty$, *and by the assumption*
$$\begin{aligned} \mathbb{P}(\sqrt{n}|\widehat{\theta}_n - \theta| > R) &\le 2\exp\{-\beta \frac{(R^2\widetilde{\varepsilon}^2 - \mathbb{E}\zeta_1^2)^2 n}{36}\} + 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon}^2, \tau, a)}{n^{a/2+1}} \\ &+ \exp\{-\frac{1}{2} n \frac{(T_1 - n\widetilde{\varepsilon}^2)^2}{T_2}\}. \end{aligned} \tag{2.25}$$
These complete the proof.

Corollary 2.3.3 *The sequence* $(\widehat{\theta}_n)_{n\in\mathbb{N}}$ *defined in (2.16) completely converges to the parameter* θ *of the first-order autoregressive process.*

Proof 9

The complete convergence follows from the inequality (2.19). Indeed, applying the Cauchy's rule on the positive real term sequences U_n *where the general term is defined by*

$$U_n = 2\exp\{-\beta\frac{(R^2\widetilde{\varepsilon}^2 - \mathbb{E}\zeta_1^2)^2 n}{36}\} + 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon}^2,\tau,a)}{n^{a/2+1}} + \exp\{-\frac{1}{2}n\frac{(T_1 - n\widetilde{\varepsilon}^2)^2}{T_2}\}.$$

By using the elementary inequality $\exp\{-y\} \leq \dfrac{4e^{-2}}{y^2}$ *for all* $y > 0$*, it follows that*

$$\sum_{n=}^{+\infty}\mathbb{P}(\sqrt{n}|\widehat{\theta}_n - \theta| > R) < +\infty\ . \tag{2.26}$$

which yields to the result.

Remark 2.3.4 *The inequalities (2.19) give us the possibility to construct a confidence interval for the parameter of the first-order autoregressive process. For large* R*, such as* $R = \widetilde{\varepsilon}\sqrt{n}$*, it follows*

$$\begin{aligned}\lim_{n\to+\infty} U_n &= 2\exp\{-\beta\frac{(n\widetilde{\varepsilon}^4 - \mathbb{E}\zeta_1^2)^2 n}{36}\} \\ &+ 2\frac{\widetilde{\Phi}(\widetilde{\varepsilon}^2,\tau,a)}{n^{a/2+1}} + \exp\{-\frac{1}{2}n\frac{(T_1 - n\widetilde{\varepsilon}^2)^2}{T_2}\} = 0\end{aligned}$$

which means, for a given level ϱ*, we can found a natural integer* n_ϱ *such as*

$$\forall n \geq n_\varrho \text{ we have } U_n \leq \varrho. \tag{2.27}$$

Consequently,

$$\mathbb{P}(|\widetilde{\theta}_{n_\varrho} - \theta| < \widetilde{\varepsilon}) \geq 1 - \varrho. \tag{2.28}$$

Chapitre 3

Almost complete convergence of sequences of dependent random variables, non-linear auto-regressive models applications.

3.1 Introduction

It is well known that the probability inequality plays an important role in various proofs of limit theorems. In particular, it provides a measure of convergence rate for the strong law of large numbers. The main purpose of the article is to provide some probability inequalities for extended negatively dependent (END) sequence, which contains independent sequence, NA sequence, and NOD sequence as special cases. These probability inequalities for END random variables are mainly inspired by Fakoor and Azarnoosh (39) and Asadian et al. (20). Using the probability inequalities, we can further study the moment inequalities and asymptotic approximation of inverse moment for END sequence.
First, we will recall the definitions of NOD and END sequences.

Definition 3.1.1 *(cf. Joag-Dev and Proschan (15)). A finite collection of random variables $X_1, X_2, \cdots, X_n$ is said to be negatively upper orthant dependent (NUOD) if for all real numbers $x_1, x_2, \cdots, x_n$,*

$$\mathbb{P}\left(X_i > x_i, i = 1, 2, \cdots, n\right) \leq \prod_{i=1}^{n} \mathbb{P}\left(\{X_i > x_i\}\right), \tag{3.1}$$

and negatively lower orthant dependent (NLOD) if for all real numbers $x_1, x_2, \cdots, x_n$,

$$\mathbb{P}\left(X_i \leq x_i, i = 1, 2, \cdots, n\right) \leq \prod_{i=1}^{n} \mathbb{P}\left(\{X_i \leq x_i\}\right), \tag{3.2}$$

A finite collection of random variables $X_1, X_2, \cdots, X_n$ is said to be negatively orthant dependent (NOD) if they are both NUOD and NLOD.
An infinite sequence $\{X_n, n \geq 1\}$ is said to be NOD if every finite sub-collection is NOD.

Definition 3.1.2 *(cf.Liu (18)). We call random variables $\{X_n, n \geq 1\}$ END if there exists a constant $M > 0$ such that both*

$$\mathbb{P}\left(X_1 > x_1, X_2 > x_2, \cdots, X_n > x_n\right) \leq M \prod_{i=1}^{n} \mathbb{P}\left(\{X_i > x_i\}\right), \tag{3.3}$$

and

$$\mathbb{P}\left(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n\right) \leq M \prod_{i=1}^{n} \mathbb{P}\left(\{X_i \leq x_i\}\right), \tag{3.4}$$

Lemma 3.1.3 *(cf. Liu (19)) Let random variables$X_1, X_2, \cdots, X_n$ be END.*

(i) Let random variables $\{f_1, f_2 \cdots, f_n\}$ re all nondecreasing (or non-increasing) functions, then random variables $f_1(X_1), f_2(X_2), \cdots, f_n(X_n)$, are END.

(ii) For each $n \geq 1$, there exists a constant $M > 0$ such that

$$\mathbb{E}\left(\prod_{i=1}^{n} X_i^+\right) \leq M \prod_{i=1}^{n} \mathbb{E}\{X_i^+\}.$$

Lemma 3.1.4 *Let $\{X_n, n \geq 1\}$ be a sequence of END random variables and $\{t_n, n \geq 1\}$ be a sequence of nonnegative numbers (or nonpositive numbers), then for each $n \geq 1$, there exists a constant $M > 0$ such that.*

$$\mathbb{E}\left(\prod_{i=1}^{n} \exp\{t_i X_i\}\right) \leq M \prod_{i=1}^{n} \mathbb{E} \exp\{t_i X_i\}.$$

As a by product, for any $t \in \mathbb{R}$,

$$\mathbb{E}\left(\prod_{i=1}^{n} \exp\{t X_i\}\right) \leq M \prod_{i=1}^{n} \mathbb{E} \exp\{X_i\}.$$

Suppose that $\{X_i, i = 0, \pm 1, \pm 2, \cdots\}$ is a sequence of strictly stationary real random variables satisfying real random variables satisfying the nonlinear autoregressive model of order p.

$$X_i = g_\theta(X_{i-1}, \cdots, X_{i-p}) + \zeta_i, \quad i \geq 1, \tag{3.5}$$

for some $\theta = (\theta_1, \cdots, \theta_q)' \in \Theta \subset \mathfrak{R}^q$, where $g_\theta, \theta \in \Theta$, is a family of known measurable functions from $\mathfrak{R}^q \to \mathfrak{R}$. Also the ζ_i''s are i.i.d. random variables with mean zero, finite variance σ^2 and common density f. Moreover, we assume that $X_{i-1}, \cdots, X_{i-p}$ are independent of $\{\zeta_i, i = 1, 2, \cdots\}$.
Let $\widehat{\theta} = (\widehat{\theta}_1, \cdots, \widehat{\theta}_q)'$ be an estimator for θ, and let

$$\widehat{\zeta}_i = X_i - g_{\widehat{\theta}}(X_{i-1}, \cdots, X_{i-p}) + \zeta_i, \quad i \geq 1, \tag{3.6}$$

denote the residuals. Based on these residuals, we construct a kernel type estimator of the error density f as follows :

$$\widehat{f}_n(t) := \frac{1}{n} \sum_{i=1}^{n} K_{h_n}(t - \widehat{\zeta}_i), \quad t \in \mathfrak{R},$$

where $K_h(t) = \frac{1}{h}K(\frac{1}{h})$, h_n *is a sequence of positive numbers tending to zero, and* K *is the kernel density function. We also need to define the kernel error based on the true errors (which we can not observe)* $\zeta_1, \zeta_2, \cdots, \zeta_n$ *:*

$$f_n(t) := \frac{1}{n}\sum_{i=1}^{n} K_{h_n}(t-\zeta_i), \quad t \in \Re.$$

Notations 3.1.5 • $A_i = \sum_{j=1}^{q}(Y_{ij}^2 - E(Y_{ij}^2))$, $B_i = \sum_{j=1}^{q}\sum_{k=1}^{q}(Z_{ijk}^2 - E(Z_{ijk}^2))$,

- $C_i = \sum_{j=1}^{q}(Y_{ij}^4 - E(Y_{ij}^4))$, $D_i = \sum_{j=1}^{q}\sum_{k=1}^{q}(Z_{ijk}^4 - E(Z_{ijk}^4))$,
  for each $n \geq 1$.
- $A_n = \sum_{i=1}^{n}\mathbb{E}(A_i^2) = \sum_{j=1}^{q}\sum_{i=1}^{n}V(Y_{ij}^2)$, $B_n = \sum_{i=1}^{n}\mathbb{E}(B_i^2) = \sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n}V(Z_{ijk}^2)$,
- $C_n = \sum_{i=1}^{n}\mathbb{E}(C_i^2) = \sum_{i=1}^{q}\sum_{i=1}^{n}V(Y_{ij}^4)$, $D_n = \sum_{i=1}^{n}\mathbb{E}(D_i^2) = \sum_{i=1}^{q}\sum_{i=1}^{n}V(Z_{ijk}^4)$.
- $\Phi_1 = \frac{nh_n^2\epsilon}{a(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}} - \sum_{j=1}^{q}\sum_{i=1}^{n}\mathbb{E}(Y_{ij}^2)$,
- $\Phi_2 = \frac{n^{1/2}h_n^{3/2}\epsilon}{2a^2(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}} - \sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n}\mathbb{E}(Z_{ijk})^2$,
- $\Phi_3 = \frac{n^{1/2}h_n^{5/2}\epsilon}{4a^2q(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}} - \sum_{j=1}^{q}\sum_{i=1}^{n}\mathbb{E}(Y_{ij}^4)$,
- $\Phi_4 = \frac{n^{1/2}h_n^{5/2}\epsilon}{a^4q^2(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}} - \sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n}\mathbb{E}(Z_{ijk}^4)$.

Basic assumptions 3.1.6 $(H1)$ $\mathbb{E}(A_{ij}) = \mathbb{E}(B_{ijk}) = \mathbb{E}(C_{ij}) = \mathbb{E}(D_{ijk}) = 0$,

$(H2)$ $|A_{ij}| \leq \alpha, |B_{ijk}| \leq \beta, |C_{ij}| < \gamma, |D_{ijk}| < \delta$ *for each* $i \geq 1$, *where* $\alpha, \beta, \gamma, \delta$ *are a positive constant.*

$(H3)$ $A_n = B_n = C_n = D_n = O(n)$,

$(H4)$ *The estimator* $\widehat{\theta} = (\widehat{\theta}_1, \cdots, \widehat{\theta}_q)'$ *for* θ *satisfies that there exists a constant* $C_1 (0 < C_1 < \infty)$ *such that*

$$\limsup_{n\to\infty}\sqrt{\frac{n}{\log\log n}}\left|\widehat{\theta}-\theta\right| \leq C_1 \quad a.s.,$$

where $\left|\widehat{\theta}-\theta\right| = \sqrt{\sum_{j=1}^{q}(\widehat{\theta}_j - \theta_j)^2}$.

(*H*5) *K is a continuous bounded symmetric kernel, K and K' have a compact support ; K''' exists and is bounded ; $(K')^2$, $(K'')^2$ are integrable.*

(*H*6) $h_n \to 0, \qquad \frac{n^{\frac{1}{2}} h_n^{5/2}}{\log\log n} \to \infty.$

3.2 Main result and its proof

Theorem 3.2.1 *Assume that (H1) and (H2) hold. Then, for any $\Phi_l > 0, l = 1, \cdots, 4$, there exist positive constants M_l such that.*

$$P\left(\sqrt{nh_n}\left|\widehat{f}_n(t) - f_n(t)\right| > \epsilon\right) \leq M_1 \exp\left\{-\left(\frac{\Phi_1}{2\alpha}\right)\operatorname{arcsin} h\left(\frac{\alpha\Phi_1}{2A_n}\right)\right\} + M_2 \exp\left\{-\left(\frac{\Phi_2}{2\beta}\right)\operatorname{arcsin} h\left(\frac{\beta\Phi_2}{2B_n}\right)\right\} + M_3 \exp\left\{-\left(\frac{\Phi_3}{2\gamma}\right)\operatorname{arcsin} h\left(\frac{\gamma\Phi_3}{2C_n}\right)\right\} + M_4 \exp\left\{-\left(\frac{\Phi_4}{2\delta}\right)\operatorname{arcsin} h\left(\frac{\delta\Phi_4}{2D_n}\right)\right\} \tag{3.7}$$

Proof 10 *By 3.5 , 3.6 and Taylor's expansion, it follows that*

$$\begin{aligned}
\widehat{\zeta}_i - \zeta_i &= g_{\widehat{\theta}}(X_{i-1}, \cdots, X_{i-p}) - g_\theta(X_{i-1}, \cdots, X_{i-p}) \\
&= \sum_{j=1}^{q} (\widehat{\theta}_j - \theta_j) Y_{ij} + \frac{1}{2}(\widehat{\theta} - \theta)' Z_i (\widehat{\theta} - \theta) \\
&= \sum_{j=1}^{q} (\widehat{\theta}_j - \theta_j) Y_{ij} + \frac{1}{2}\sum_{j=1}^{q}\sum_{k=1}^{q} (\widehat{\theta}_j - \theta_j)(\widehat{\theta}_k - \theta_k) Z_{ijk}.
\end{aligned}$$

where $Y_{ij} := \frac{\partial}{\partial\theta_j} g_\theta(X_{i-1}, \cdots, X_{i-p})$ and Z_i is a $q \times q$ matrix with the j^{th} row and k^{th} column element $Z_{ijk} := \frac{\partial^2}{\partial\theta_j\partial\theta_k} g_\theta(X_{i-1}, \cdots, X_{i-p})$ and evaluated at $\theta^* := \theta + \lambda(\widehat{\theta} - \theta), \lambda \in (0, 1)$, i.e., $\widehat{\theta}$ takes place of θ in 10

Under the assumptions and the fact that $(a_1 + a_2 + \cdots + a_m)^2 \leq m(a_1^2 + a_2^2 + \cdots + a_m^2)$ for any positive integer m and real numbers $a_1, a_2, \cdots, a_m$, we have that

$$P\left(\sqrt{nh_n}\left|\widehat{f}_n(t)-f_n(t)\right|>\epsilon\right)=P\left(\left|\tfrac{1}{nh_n}\sum_{i=1}^{n}\left[K\left(\frac{t-\widehat{\zeta}_i}{h_n}\right)-K\left(\frac{t-\zeta_i}{h_n}\right)\right]\right|>\tfrac{\epsilon}{\sqrt{nh_n}}\right)$$

$$=P\left(\left|\tfrac{1}{nh_n}\sum_{i=1}^{n}\left[\frac{\zeta_i-\widehat{\zeta}_i}{h_n}K'\left(\frac{t-\zeta_i}{h_n}\right)+\frac{(\zeta_i-\widehat{\zeta}_i)^2}{2h_n^2}K''(\eta_i(t))\right]\right|>\tfrac{\epsilon}{\sqrt{nh_n}}\right)$$

$$\leq P\left(\frac{\left|\sum_{j=1}^{q}(\widehat{\theta}_j-\theta_j)\sum_{i=1}^{n}Y_{ij}K'\left(\frac{t-\zeta_i}{h_n}\right)\right|}{nh_n^2}>\tfrac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$+P\left(\frac{\left|\sum_{j=1}^{q}\sum_{k=1}^{q}(\widehat{\theta}_j-\theta_j)(\widehat{\theta}_k-\theta_k)\sum_{i=1}^{n}Z_{ijk}K'\left(\frac{t-\zeta_i}{h_n}\right)\right|}{2nh_n^2}>\tfrac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$+P\left(\frac{q\sum_{j=1}^{q}(\widehat{\theta}_j-\theta_j)2\sum_{i=1}^{n}Y_{ij}^2\left|K''(\eta_i(t))\right|}{nh_n^3}>\tfrac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$+P\left(\frac{q^2\sum_{j=1}^{q}\sum_{k=1}^{q}(\widehat{\theta}_j-\theta_j)^2(\widehat{\theta}_k-\theta_k)^2\sum_{i=1}^{n}Z_{ijk}^2\left|K''(\eta_i(t))\right|}{4nh_n^3}>\tfrac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$=I+II+III+IV,$$

where $\eta_i(t)$ is a random quantity between $\frac{t-\widehat{\zeta}_i}{h_n}$ and $\frac{t-\zeta_i}{h_n}$. Now we begin to deal with them, respectively. As to I, notice that :

$$P\left(\frac{\left|\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)\sum_{i=1}^{n}Y_{ij}k'(\frac{t-\epsilon_i}{h_n})\right|}{nh_n^2}>\tfrac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$\leq P\left(\sum_{j=1}^{q}\left|\hat{\theta}_j-\theta_j\right|\left[(\sum_{i=1}^{n}Y_{ij}^2)^{\frac{1}{2}}(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}\right]>\frac{n^{1/2}h_n^{3/2}}{4}\right)$$

$$\leq P\left((\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2)^{\frac{1}{2}}\left[(\sum_{i=1}^{n}Y_{ij}^2)^{\frac{1}{2}}(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}\right]>\frac{n^{1/2}h_n^{3/2}}{4}\right)$$

$$\leq P\left(\sum_{j=1}^{q}(\sum_{i=1}^{n}Y_{ij}^2)^{\frac{1}{2}}>\frac{nh_n^2\epsilon}{a(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}}\right)\leq P\left(\sum_{j=1}^{q}\sum_{i=1}^{n}Y_{ij}^2>\frac{n^{1/2}h_n^{3/2}\epsilon}{4a(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}}\right)$$

$$=P\left(\underbrace{\sum_{i=1}^{n}\underbrace{\sum_{j=1}^{q}\left[Y_{ij}^2-E(Y_{ij}^2)\right]}_{A_i}}_{S_{n1}}>\underbrace{\frac{nh_n^2\epsilon}{a(\sum_{i=1}^{n}k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}}-\sum_{j=1}^{q}\sum_{i=1}^{n}E(Y_{ij}^2)}_{\Phi_1}\right)=P\left(S_{n1}>\Phi_1\right)$$

For all, $\nu_1 > 0$ *and* $\Phi_1 > 0$*, we have*

$$P\left(S_{n1} > \Phi_1\right) \leq e^{-\nu_1 \Phi_1} \mathbb{E}(e^{\nu_1 \S_{n1}}) = M_1 \exp\left\{\nu_1 \left(A_n \frac{\sinh(\nu_1 \alpha)}{\alpha} - \Phi_1\right)\right\} \tag{3.8}$$

Taking $\nu_1 = \frac{1}{\alpha} \operatorname{arcsin} h\left(\frac{\alpha \Phi_1}{2A_n}\right)$ *in the right-hand of 3.8, we can see that* $A_n \frac{\nu_1 \alpha}{\alpha} = \frac{\Phi_1}{2}$ *and*

$$P\left(S_{n1} > \Phi_1\right) \leq M_1 \exp\left\{-\left(\frac{\Phi_1}{2\alpha}\right) \operatorname{arcsin} h(\frac{\alpha \Phi_1}{2A_n})\right\}$$

For II*, note that*

$$P\left(\frac{\left|\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_k - \theta_k)\sum_{i=1}^{n} Z_{ijk} k'(\frac{t-\epsilon_i}{h_n})\right|}{2nh_n^2} > \frac{\epsilon}{4\sqrt{nh_n}}\right)$$
$$\leq P\left(\sum_{j=1}^{q}\sum_{k=1}^{q}\left|\hat{\theta}_j - \theta_j\right|\left|\hat{\theta}_k - \theta_k\right|\left[(\sum_{i=1}^{n} Z_{ijk}^2)^{\frac{1}{2}}(\sum_{i=1}^{n} k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}\right] > \frac{n^{1/2}h_n^{3/2}\epsilon}{2}\epsilon\right)$$
$$\leq P\left(\sum_{j=1}^{q}\sum_{k=1}^{q}(\sum_{i=1}^{n} Z_{ijk}^2)^{\frac{1}{2}} > \frac{n^{1/2}h_n^{3/2}\epsilon}{2a^2(\sum_{i=1}^{n} k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}}\right)$$
$$\leq P\left(\sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n} Z_{ijk}^2 > \frac{n^{1/2}h_n^{3/2}\epsilon}{2a^2(\sum_{i=1}^{n} k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}}\right)$$
$$= P\left(\underbrace{\sum_{i=1}^{n}\underbrace{\sum_{j=1}^{q}\sum_{k=1}^{q}\left[Z_{ijk}^2 - E(Z_{ijk}^2)\right]}_{B_i}}_{S_{n2}} > \underbrace{\frac{n^{1/2}h_n^{3/2}\epsilon}{2a^2(\sum_{i=1}^{n} k'^2(\frac{t-\epsilon_i}{h_n}))^{\frac{1}{2}}} - \sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n} E(Z_{ijk})^2}_{\Phi_2}\right)$$
$$= P\left(S_{n2} > \Phi_2\right) \leq e^{-\nu_1 \Phi_1} \mathbb{E}(e^{\nu_1 \S_{n1}}) \leq M_2 \exp\left\{\nu_2 \left(B_n \frac{\sinh(\nu_2 \beta)}{\beta} - \Phi_2\right)\right\}$$
$$\leq M_2 \exp\left\{-\left(\frac{\Phi_2}{2\beta}\right) \operatorname{arcsin} h(\frac{\beta \Phi_2}{2B_n})\right\}.$$

For III

$$P\left(\frac{q\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\sum_{i=1}^{n}Y_{ij}^2\left|k''(\eta_i(t))\right|}{nh_n^3}>\frac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$\leq P\left(\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\left[(\sum_{i=1}^{n}Y_{ij}^4)^{\frac{1}{2}}(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}\right]>\frac{n^{1/2}h_n^{5/2}\epsilon}{4q}\right)$$

$$\leq P\left(\sum_{j=1}^{q}(\sum_{i=1}^{n}Y_{ij}^4)^{\frac{1}{2}}>\frac{n^{1/2}h_n^{5/2}\epsilon}{4a^2q(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}}\right)\leq P\left(\sum_{j=1}^{q}\sum_{i=1}^{n}Y_{ij}^4>\frac{n^{1/2}h_n^{5/2}\epsilon}{4a^2q(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}}\right)$$

$$=P\left(\underbrace{\sum_{i=1}^{n}\underbrace{\sum_{j=1}^{q}\left[Y_{ij}^4-E(Y_{ij}^4)\right]}_{C_i}}_{S_{n3}}>\underbrace{\frac{n^{1/2}h_n^{5/2}\epsilon}{4a^2q(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}}-\sum_{j=1}^{q}\sum_{i=1}^{n}E(Y_{ij}^4)}_{\Phi_3}\right)$$

$$=P(S_{n3}>\Phi_3)\leq e^{-\nu_3\Phi_3}\mathbb{E}(e^{\nu_3 \S_{n3}})\leq M_1\exp\left\{\nu_3\left(C_n\frac{\sinh(\nu_3\gamma)}{\gamma}-\Phi_3\right)\right\}$$

$$\leq M_3\exp\left\{-\left(\frac{\Phi_3}{2\gamma}\right)\arcsin h\left(\frac{\gamma\Phi_3}{2C_n}\right)\right\}.$$

For IV

$$P\left(\frac{q^2\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j-\theta_j)^2(\hat{\theta}_k-\theta_k)^2\sum_{i=1}^{n}Z_{ijk}^2\left|k''(\eta_i(t))\right|}{4nh_n^3}>\frac{\epsilon}{4\sqrt{nh_n}}\right)$$

$$\leq P\left(\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j-\theta_j)^2(\hat{\theta}_k-\theta_k)^2(\sum_{i=1}^{n}Z_{ijk}^4)^{\frac{1}{2}}(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}>\frac{n^{1/2}h_n^{5/2}\epsilon}{q^2}\right)$$

$$\leq P\left(\sum_{j=1}^{q}\sum_{k=1}^{q}(\sum_{i=1}^{n}Z_{ijk}^4)^{\frac{1}{2}}>\frac{n^{1/2}h_n^{5/2}\epsilon}{a^4q^2(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}}\right)\leq P\left(\sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n}Z_{ijk}^4>\frac{n^{1/2}h_n^{5/2}\epsilon}{a^4q^2(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}}\right)$$

$$\leq P\left(\underbrace{\sum_{i=1}^{n}\underbrace{\sum_{j=1}^{q}\sum_{k=1}^{q}\left[Z_{ijk}^4-E(Z_{ijk}^4)\right]}_{D_i}}_{S_{n4}}>\underbrace{\frac{n^{1/2}h_n^{5/2}\epsilon}{a^4q^2(\sum_{i=1}^{n}k''^2(\eta_i(t)))^{\frac{1}{2}}}-\sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{i=1}^{n}E(Z_{ijk}^4)}_{\Phi_4}\right)$$

$$=P\left(S_{n4}>\Phi_4\right)\leq e^{-\nu_4\Phi_4}\mathbb{E}(e^{\nu_4 \S_{n4}})\leq M_4\exp\left\{\nu_4\left(D_n\frac{\sinh(\nu_4\delta)}{\delta}-\Phi_1\right)\right\}$$

$$\leq M_4\exp\left\{-\left(\frac{\Phi_4}{2\delta}\right)\arcsin h\left(\frac{\delta\Phi_4}{2D_n}\right)\right\}.$$

Then for $(I),(II),(III)$ *and* (IV)

$$
\begin{aligned}
P\left(\left|\widehat{f}_n(t)-f_n(t)\right|>\epsilon\right) \;\leq\; & M_1\exp\left\{-\left(\tfrac{\Phi_1}{2\alpha}\right)\operatorname{arcsin}h\left(\tfrac{\alpha\Phi_1}{2A_n}\right)\right\} \\
& +M_2\exp\left\{-\left(\tfrac{\Phi_2}{2\beta}\right)\operatorname{arcsin}h\left(\tfrac{\beta\Phi_2}{2B_n}\right)\right\} \\
& +M_3\exp\left\{-\left(\tfrac{\Phi_3}{2\gamma}\right)\operatorname{arcsin}h\left(\tfrac{\gamma\Phi_3}{2C_n}\right)\right\} \\
& +M_4\exp\left\{-\left(\tfrac{\Phi_4}{2\delta}\right)\operatorname{arcsin}h\left(\tfrac{\delta\Phi_4}{2D_n}\right)\right\}
\end{aligned}
$$

Theorem 3.2.2 *Assume that* $(H1)$, $(H2)$ *and* $(H3)$ *hold. If* $(H3)$ *is satisfied then* $\sqrt{nh_n}\left|\widehat{f}_n(t)-f_n(t)\right|\to 0$ *completely and in consequence* $\sqrt{nh_n}\left|\widehat{f}_n(t)-f_n(t)\right|\to 0$ *a.s.for each* $n\geq 1$.

Proof 11 *For any* $\epsilon>0$, *we have by Theorem 3.2.1 that*

$$
\begin{aligned}
& P\left(\left|\widehat{f}_n(t)-f_n(t)\right|>\tfrac{\epsilon}{\sqrt{nh_n}}\right) \\
& \leq M_1\exp\left\{-\left(\tfrac{\Phi_1}{2\alpha}\right)\operatorname{arcsin}h\left(\tfrac{\alpha\Phi_1}{2A_n}\right)\right\}+M_2\exp\left\{-\left(\tfrac{\Phi_2}{2\beta}\right)\operatorname{arcsin}h\left(\tfrac{\beta\Phi_2}{2B_n}\right)\right\} \\
& +M_3\exp\left\{-\left(\tfrac{\Phi_3}{2\gamma}\right)\operatorname{arcsin}h\left(\tfrac{\gamma\Phi_3}{2C_n}\right)\right\}+M_4\exp\left\{-\left(\tfrac{\Phi_4}{2\delta}\right)\operatorname{arcsin}h\left(\tfrac{\delta\Phi_4}{2D_n}\right)\right\} \\
& \leq M_1\exp\{-nL_1\}+M_2\exp\{-nL_2\}+M_3\exp\{-nL_3\}+M_4\exp\{-nL_4\} \\
& \leq M\exp\{-nL\}
\end{aligned}
$$

where $L=\min\{L_1,L_2,L_3,L_4\}$ *(*$L_i(i=1,\cdots,4)$ *are positive constants),* $M=\max\{M_1,M_2,M_3,M_4\}$ *and* $a=O_p(\sqrt{\frac{\log\log n}{n}})$. *Therefore,*

$$
\sum_{n=1}^{\infty}P\left(\left|\widehat{f}_n(t)-f_n(t)\right|>\frac{\epsilon}{\sqrt{nh_n}}\right)<\infty,
$$

which implies that $\sqrt{nh_n}\left|\widehat{f}_n(t)-f_n(t)\right|\to 0$ *completely and in consequence* $\sqrt{nh_n}\left|\widehat{f}_n(t)-f_n(t)\right|\to 0$ *a.s. by Borel-Cantelli Lemma. The proof is completed.*

Chapitre 4

Complete convergence of the operator estimator of first-order autoregressive in Hilbert space generated by WOD errors

4.1 Introduction

Let $\{X_n, n \geq 1\}$ be a sequence of random variables that is defined on a fixed probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. As known that there are many results on probability limit theorems for independent random variables. In fact, the independence assumption is not always appropriate in applications.

So many authors introduced some dependent structure and mixing structure. Widely orthant dependent structure was one of the newest dependence structure that has attracted the interest of probabilists and statisticians, this structure contains most of negatively dependent random variables, some positively dependent random variables and some other random variables. The concept of widely orthant dependent random variables was introduced by Wang et al. (16) as follows.

Definition 4.1.1 *For $\{X_n, n \geq 1\}$ a sequence of random variables :*

(i) *if there exists a sequence of real numbers $\{g_u(n), n \geq 1\}$ such that for each $n \geq 1$ and for all $x_i \in (-\infty, \infty), 1 \leq i \leq n$;*

$$\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i > x_i\}\right) \leq g_u(n) \prod_{i=1}^{n} \mathbb{P}\left(\{X_i > x_i\}\right),$$

Then we say that the random variables $\{X_n, n \geq 1\}$ are widely upper orthant dependent (WUOD) with dominating coefficients $g_u(n)$, $n \geq 1$;

(ii) *if there exists a sequence of real numbers $\{g_l(n), n \geq 1\}$ such that for each $n \geq 1$ and for all*

$$\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i \leq x_i\}\right) \leq g_l(n) \prod_{i=1}^{n} \mathbb{P}\left(\{X_i \leq x_i\}\right),$$

Then we say that the random variables $\{X_n, n \geq 1\}$ are widely lower orthant dependent (WLOD) with dominating coefficients $g_l(n), n \geq 1$;

(iii) *If* $\{X_n, n \geq 1\}$ *are both* $WUOD$ *and* $WLOD$*, then we say that the random variables* $\{X_n, n \geq 1\}$ *are widely orthant dependent (*WOD*) with dominating coefficients* $g_u(n)$ *and* $g_l(n), n \geq 1$.

(iv) *An array* $\{X_{ni}, i \geq 1, n \geq 1\}$ *is said row-wise* WOD *random variables, if for each* $n \geq 1$, $\{X_{ni}, i \geq 1\}$ *is a sequence of* WOD *random variables.*

Recall that when $g_l(n) = g_u(n) = M$ *for some positive constant* M*, the random variables* $\{X_n, n \geq 1\}$ *are called extended negatively upper orthant dependent (*$ENUOD$*, in short) and extended negatively lower orthant dependent (*$ENLOD$*, in short), respectively. If they are both* $ENUOD$ *and* $ENLOD$*, then we say that the random variables* $\{X_n, n \geq 1\}$ *are extended negatively orthant dependent (*END*, in short). The concept of* END *random variables was proposed by Liu (18), and further promoted by Chen et al. (48), Shen (32), Wang and Wang (33), Wu and Guan (52), Qiu et al. (9), Wang et al. (44)-(45), and so forth. When* $g_l(n) = g_u(n) = 1$ *for any* $n \geq 1$*, the random variables* $\{X_n, n \geq 1\}$ *are called negatively upper orthant dependent (*$NUOD$*, in short) and negatively lower orthant dependent (*$NLOD$*,in short), respectively. If they are both* $NUOD$ *and* $NLOD$*, then we say that the random variables* $\{X_n, n \geq 1\}$ *are negatively orthant dependent (*NOD*, in short). The concept of* NOD *random variables was introduced by Ebrahimi and Ghosh (21) and carefully studied by Joag and Proschan (15), Bozorgnia and al. (1), Taylor and al. (28), Wang and al. (47), Sung (34), Qiu and al. (10), Wu (43), Wu and Jiang (26), Shen ((4),(5)), and so on. Joag and Proschan (15) pointed out that* NA *random variables are* NOD*. Hu (13) introduced the concept of negatively super-additive dependence (*NSD*, in short) and pointed out that* NSD *implies* NOD*. Christofides and Vaggelatou (37) indicated that* NA *implies* NSD*. By the statements above, we can see that the class of* WOD *random variables contains* END *random variables,* NOD *random variables,* NSD *random variables,* NA *random variables and independent random variables as special cases. Hence, studying the probability limiting behavior of* WOD *random variables is a great interest.*

The concept of WOD *random variables was introduced by Wang et al. (16) and many applications have been found subsequently. See, for example, Wang et al. (16) provided some examples which showed that the class of* WOD *random variables contains some common negatively dependent random variables, some positively dependent random variables and some others ; in addition, they studied the uniform asymptotic for the finite-time ruin probability of a new dependent risk model with a constant interest rate. Wang and Cheng (50) presented some basic renewal theorems for a random walk with widely dependent increments and gave some applications. Wang and al. (51) studied the asymptotic of the finite-time ruin probability for a generalized renewal risk model with independent strong sub-exponential claim sizes and widely lower orthant dependent inter-occurrence times. Liu (42) gave the asymptotically equivalent formula for the finite-time ruin probability under a dependent risk model with constant interest rate. Chen et al. (49) considered uniform asymptotic for the finite-time ruin probabilities of two kinds of nonstandard bi-dimensional renewal risk models with constant interest forces and diffusion generated by Brownian motions. Shen (3) established the Bernstein type inequality for* WOD *random variables and gave some applications, Wang et al. (46) studied the complete convergence for*

WOD random variables and gave its applications in nonparametric regression models. Yang et al. (40) established the Bahadur representation of sample quantiles for WOD random variables under some mild conditions.

Definition 4.1.2 *Let an array $\{X_{ni}, i \geq 1, n \geq 1\}$ of random variables is said to be stochastically dominated by a random variable X if there exists a positive constant C so that*

$$\mathbb{P}(\mid X_{ni} \mid) \quad \leq \quad C\mathbb{P}(\mid X \mid) \tag{4.1}$$

$$\textit{for all} \quad x \geq 0, i \geq 1 \quad \textit{and} \qquad n \geq 1.$$

Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with dominating coefficients $g_u(n), g_l(n), n \geq 1$ and $\{X_{ni}, i \geq 1, n \geq 1\}$ be an array of row-wise WOD random variables with dominating coefficients $g_u(n), g_l(n), n \geq 1$ in each row. Denote $g(n) = \max\{g_u(n), g_l(n)\}$. $I()$ denotes the indicator function, with C a positive constant, which value may be different places.

Lemma 4.1.3 **(i)** *Let random variables $\{X_n, n \geq 1\}$ be WOLD (WUOD) with dominating coefficients $g_l(n), n \geq 1$ $(g_u(n), n \geq 1)$. If $\{f_n(.), n \geq 1\}$ are nondecreasing, then $\{f_n(X_n), n \geq 1\}$ are still WLOD (WUOD) with dominating coefficients $g_l(n), n \geq 1$ $(g_l(n)), n \geq 1)$; if $\{f_n(.), n \geq 1\}$ are non-increasing, then $\{f_n(X_n), n \geq 1\}$ are WUOD (WLOD) with dominating coefficients $g_l(n), n \geq 1$ $(g_l(n), n \geq 1)$.*

(ii) *If $\{X_n, n \geq 1\}$ are nonnegative WUOD with dominating coefficients $g_u(n), n \geq 1$, then for each $n \geq 1$,*

$$\mathbb{E}\exp\left(t\sum_{i=1}^{n} X_i\right) \leq g_u(n)\prod_{i=1}^{n}\mathbb{E}\exp\{tX_i\}.$$

Particularly, if $\{X_n, n \geq 1\}$ are WUOD with dominating coefficients $g_u(n), n \geq 1$, then for each $n \geq 1$ and any $t > 0$

$$\mathbb{E}\exp\left(t\sum_{i=1}^{n} X_i\right) \leq g_u(n)\prod_{i=1}^{n}\mathbb{E}\exp\{tX_i\}.$$

Wang et al. (46) obtained the following corollary by Lemma 4.1.3.

Corollary 4.1.4 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables.*

(i) *If $f_n(.)$ are all nondecreasing (or non-increasing), then $\{f_n(X_n), n \geq 1\}$ are still WOD.*

(ii) *For each $n \geq 1$ and any $t \in \mathbb{R}$,*

$$\mathbb{E}\exp\left(t\sum_{i=1}^{n} X_i\right) \leq g(n)\prod_{i=1}^{n}\mathbb{E}\exp\{tX_i\}.$$

4.2 Principal results

Theorem 4.2.1 *(See Avanissian (6), pp* $314-315$*) Let* $(v_j)_{j\in I}$ *orthonormal family in Hilbert space* H *; the following conditions are equivalents :*

(a) *The family* $(v_j)_{j\in I}$ *is maximal (i.e. is hilbertian base).*

(b) *All* $x \in H$ *is written in unique way*

$$x = \sum_{j\in I} x_j v_j \qquad \text{where} \quad x_j =< x, v_j >$$

(c) *For any* $x \in H$ *the following statement is hold*

$$\| x \| = \sum_{j\in I} | x_j |^2 \text{ .(Parseval relation)} \tag{4.2}$$

Theorem 4.2.2 *Let* $(\zeta_t, t \in \mathbb{Z})$ *sequence of identically distributed widely orthant dependent (WOD, in short) random variables, of mean zero (i.e* $\mathbb{E}\zeta_t = 0, \forall t \in \mathbb{Z}$*) and hold* $\sup_{j\in\mathbb{Z}} \| \zeta_t \| \leq b < \infty$*, indeed if, for any* $\epsilon < \mathbb{E} \mid \zeta_n \mid \forall n \in \mathbb{Z}$*. Then*

$$\mathbb{P}\left\{ \mid \frac{\zeta_1 + \ldots + \zeta_n}{n} \mid > \epsilon \right\} \quad \leq \quad g(n) e^{-(A_2 - A_1)n}$$

where $A_1 = \frac{\mathbb{E}|\zeta|}{b}\left(\frac{\epsilon}{\mathbb{E}|\zeta|} - 1\right)$ *and* $A_2 = \frac{\epsilon}{b}\log\left(\frac{\epsilon}{\mathbb{E}|\zeta|}\right)$.

Proof 12 *we notice that*

$$\left\{ \left| \frac{\zeta_1 + \ldots + \zeta_n}{n} \right| \right\} \subseteq \left\{ \frac{| \zeta_1 | + \ldots + | \zeta_n |}{n} \right\}$$

we obtain the following inequality valid for each $\lambda, \epsilon > 0$ *:*

$$\begin{aligned}
\widetilde{I} = \mathbb{P}\left\{ \left| \frac{\zeta_1 + \ldots + \zeta_n}{n} \right| > \epsilon \right\} \quad &\leq \quad \mathbb{E}\left[\exp\left(\lambda \left(\sum_{j=1}^{n} | \zeta_j | - n\epsilon \right) \right) \right] \\
&= \quad e^{-\lambda n \epsilon} \mathbb{E}\left[\exp\left(\lambda \sum_{j=1}^{n} | \zeta_j | \right) \right] \\
&= \quad e^{-\lambda n \epsilon} \mathbb{E}\left[\prod_{j=1}^{n} e^{\lambda |\zeta_j|} \right]
\end{aligned}$$

We use lemma 4.1.3 and the corollary 4.1.4 we have

$$\widetilde{I} \quad \leq \quad e^{-\lambda n \epsilon} g(n) \prod_{j=1}^{n} \mathbb{E}\left[e^{\lambda |\zeta_j|} \right]$$

assumption $\sup_{j\in\mathbb{Z}} \| \zeta_j \| \leq b < \infty$ *of theorem and the elementary inequality* $e^{xt} \leq 1 + x(e^t - 1), 0 \leq x \leq 1, t > 0$ *then allow to bounded* $\widetilde{I}$ *by :*

$$e^{-\lambda n \epsilon} g(n) \prod_{j=1}^{n} \mathbb{E}\left[1 + \frac{| \zeta_j |}{b}(e^{\lambda b} - 1) \right]. \tag{4.3}$$

On another hand, for $\epsilon, \lambda > 0$, *we have :*

$$\begin{aligned}\widetilde{I} &\leq e^{-\lambda n\epsilon} g(n) \prod_{j=1}^{n} \left[1 + \mathbb{E} \mid \zeta_j \mid \left(\frac{e^{\lambda b} - 1}{b}\right)\right] \\ &\leq e^{-\lambda n\epsilon} g(n) \prod_{j=1}^{n} \exp\left(\mathbb{E} \mid \zeta_j \mid \left(\frac{e^{\lambda b} - 1}{b}\right)\right) \text{(using} 1+ x \leq e^x, \forall x \in \mathbb{R})\end{aligned}$$

(4.5)

$$\begin{aligned} &= e^{-\lambda n\epsilon} g(n) \exp\left(n\mathbb{E} \mid \zeta_j \mid \left(\frac{e^{\lambda b} - 1}{b}\right)\right) \\ &= g(n) \exp\left(n\mathbb{E} \mid \zeta_j \mid \left(\frac{e^{\lambda b} - 1}{b}\right) - \lambda n\epsilon\right).\end{aligned}$$

The equation $\frac{\partial \Lambda(\lambda)}{\partial \lambda} = \frac{\partial\left(n\mathbb{E}|\zeta_j|\left(\frac{e^{\lambda b}-1}{b}\right)-\lambda n\epsilon\right)}{\partial \lambda} = 0$ *has a unique solution* $\lambda = \frac{\ln\left(\frac{\epsilon}{\mathbb{E}|\zeta_j|}\right)}{b}$ *that minimizes* $\Lambda(\lambda)$. *Then*

$$\mathbb{P}\left\{\left| \frac{\zeta_1 + \ldots + \zeta_n}{n}\right| > \epsilon\right\} \leq g(n) e^{-n(A_2 - A_1)}$$

where $A_1 = \frac{\mathbb{E}|\zeta|}{b}\left(\frac{\epsilon}{\mathbb{E}|\zeta|} - 1\right)$ *and* $A_2 = \frac{\epsilon}{b}\ln\left(\frac{\epsilon}{\mathbb{E}|\zeta|}\right)$. *that concludes our Theorem.*

Remark 4.2.3 **(a)** $e^{-n(A_2 - A_1)} \leq \frac{4e^{-2}}{(A_2 - A_1)^2 n^2}$ *(using the inequality :* $e^{-x} \leq \frac{4e^{-2}}{x^2}, \forall x > 0$*).*

(b) $\mathbb{P}\left\{\mid \frac{\zeta_1 + \ldots + \zeta_n}{n} \mid > \epsilon\right\} \leq A_3 \sum_{n=1}^{+\infty} \frac{n^\gamma}{n^2} < +\infty$

where $A_3 = \frac{4e^{-2}}{(A_2 - A_1)^2}$ *and* $g(n) = O(n^\gamma), (0 \leq \gamma < 1)$.

Lemma 4.2.4 (t_1) $\sum_{j=1}^{r} j^{-1} \leq 1 + \ln r$,

(t_2) $\sum_{j=1}^{r} j^{\kappa - 1} \leq \kappa^{-1} r^\kappa, 0 < \kappa < 1$,

(t_3) $\sum_{j=r+1}^{+\infty} j^{-1-\omega} \leq \omega^{-1} r^{-\omega}$.

Proof 13 • *With reference to* (τ_1) *: if* $0 < j \leq k$, *then*

$$k^{-1} \leq j^{-1}$$

and

$$\int_{k-1}^{k} k^{-1} dj \leq \int_{k-1}^{k} j^{-1} d\, j = \ln k - \ln(k-1) \text{ so } \sum_{j=1}^{t} k^{-1} \leq 1 + \ln r.$$

- *As to* (τ_2) *: if* $0 < j \leq k$*, then*

$$k^{\kappa-1} \leq j^{\kappa-1}$$

and

$$\int_{k-1}^{k} K^{\kappa-1} dj \leq \int_{k-1}^{k} j^{\kappa-1} dj = \kappa^{-1}(k^{\kappa} - (k-1)^{\kappa}) so \sum_{j=1}^{r} k^{\kappa-1} \leq \kappa^{-1} r^{\kappa}, 0 < \kappa < 1.$$

- *As to* (t_3) *: if* $0 < j \leq k$*, then*

$$k^{-1-\omega} \leq j^{-1-\omega}$$

and

$$\int_{k-1}^{k} K^{-1-\omega} dj \leq \int_{k-1}^{k} j^{-1-\omega} dj = -\omega^{-1}(k^{-\omega} - (k-1)^{-\omega})$$

therefor $\displaystyle\sum_{j=r+1}^{+\infty} k^{-1-\omega} \leq \omega^{-1} r^{-\omega}, 0 < k < 1.$

Theorem 4.2.5 *Let* $X = (X_t, t \in \mathbb{Z})$ *hilbertian autoregressive processes of first order* $(ARH(1))$*. Assume that* $\sup_{t\in\mathbb{Z}} \parallel \zeta_t \parallel \leq b < \infty$ *where* (ζ_t) *are widely orthant dependent noise (WOD) and* $(\lambda_j, j \geq 1)$ *are eigenvalues of operator* $C = C_{x_0}$*. Then*

$$\mathbb{P}\left\{\parallel \frac{S_n}{n} \parallel \ > \epsilon\right\} \ \leq \ \eta(t+1)g(n)e^{-(A_2^{'}-A_1^{'})n} + \frac{g(n)}{\delta\epsilon^2}\sum_{j>\eta}\lambda_j. \qquad (4.5)$$

where

$$A_2^{'} \ = \ \frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}b} \ln\left[\frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}\mathbb{E}\mid<\Phi_{t,k}, v_j>\mid}\right]$$

and

$$A_1^{'} \ = \ \frac{\mathbb{E}\mid<\Phi_{t,k}, v_j>\mid}{b}\left(\frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}\mathbb{E}\mid<\Phi_{t,k}, v_j>\mid} - 1\right), 1 \leq j \leq \eta.$$

Proof 14 *Let* $(v_j, j \geq 1)$ *eigenvectors of operator* C *that form an orthonormal base of* H*.*
For each $\delta \in]0,1[$ *we can write*

$$\begin{aligned}\overline{I} = \mathbb{P}\left\{\parallel \frac{S_n}{n} \parallel \ > \epsilon\right\} \ &= \ \mathbb{P}\left\{\parallel \frac{S_n}{n} \parallel^2 > \epsilon^2\right\} \\ &= \ \mathbb{P}\left\{\sum_{j\geq 1}\mid \ < \frac{S_n}{n}, v_j >^2 \mid \ > \epsilon^2\right\}.\end{aligned}$$

Using $\parallel \frac{S_n}{n} \parallel^2 = \sum_{j\geq 1} \mid x_j \mid^2$*,* $\frac{S_n}{n} = \sum_{j\geq 1} x_j v_j$*,* $x_j = < \frac{S_n}{n}, v_j >$ *we deduce*

$$\overline{I} \ \leq \ \mathbb{P}\left\{\sum_{j=1}^{\eta} < \frac{S_n}{n}, v_j >^2 \geq (1-\delta)\epsilon^2\right\} + \mathbb{P}\left\{\sum_{j>\eta} < \frac{S_n}{n}, v_j >^2 \geq \delta\epsilon^2\right\},$$

we obtain the following bounded-up

$$I_1 = \mathbb{P}\left\{\sum_{j=1}^{\eta} < \frac{S_n}{n}, v_j >^2 > (1-\delta)\epsilon^2\right\}$$

$$I_1 \leq \sum_{j=1}^{\eta} \mathbb{P}\left\{< \frac{S_n}{n}, v_j >\geq \epsilon \frac{\sqrt{1-\delta}}{\sqrt{\eta}}\right\} \tag{4.6}$$

we have $\mathbb{P}\{T_1^2 + \ldots + T_\eta^2 \geq K\} \leq \mathbb{P}\{T_1^2 \geq \frac{K}{\eta}\} + \ldots + \mathbb{P}\{T_\eta^2 \geq \frac{K}{\eta}\}$.

We put $I_1' = \mathbb{P}\left\{< \frac{S_n}{n}, v_j >\geq \epsilon \frac{\sqrt{1-\delta}}{\sqrt{\eta}}\right\}$ *where* $\frac{S_n}{n} = \frac{\sum_{t=1}^{n} X_t}{n}$,

$X_t = \zeta_t + \rho(\zeta_{t-1}) + \ldots + \rho^{t-1}(\zeta_1)$, $\parallel \rho \parallel < 1$ *and* (ζ_t) *are* $H-W.O.D.$

So

$$\begin{aligned}
I_1' &= \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n} < X_t, v_j >\geq \epsilon \frac{\sqrt{1-\delta}}{\sqrt{\eta}}\right\} \\
&= \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n} < \sum_{k=1}^{t+1} \rho^{k-1}(\zeta_{t-(k-1)}), v_j >\geq \epsilon \frac{\sqrt{1-\delta}}{\sqrt{\eta}}\right\} \\
&= \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n}\sum_{k=1}^{t+1} < \rho^{k-1}(\zeta_{t-(k-1)}), v_j >\geq \epsilon \frac{\sqrt{1-\delta}}{\sqrt{\eta}}\right\} \\
&\leq \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n} < \zeta_t, v_j >\geq \frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}}\right\} + \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n} < \rho(\zeta_{t-1}), v_j >\geq \frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}}\right\} + \\
&\quad + \ldots + \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n} < \rho^t(X_0), v_j >\geq \frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}}\right\}
\end{aligned}$$

we put

$$I'_{k,1} = \mathbb{P}\left\{\frac{1}{n}\sum_{t=1}^{n} < \underbrace{\rho^{k-1}(\zeta_{t-(k-1)})}_{\Phi_{t,k}}, v_j >\geq \frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}}\right\}.$$

By applying lemma 4.2.2 on $(< \Phi_{t,k}, v_j >, t \in \mathbb{Z})$ *we achieve*

$$I'_{k,1} \leq g(n)e^{-(A'_2 - A'_1)n}, \qquad \forall 1 \leq k \leq t+1,$$

thus we conclude that

$$I'_1 \leq (t+1)I'_{k,1} \leq g(n)(t+1)e^{-(A'_2-A'_1)n}, \forall 1 \leq k \leq t+1 \tag{4.7}$$

From (4.7) and (4.6) we find :

$$I_1 \leq \sum_{j=1}^{\eta} I'_1 \leq \sum_{j=1}^{\eta}(t+1)I'_{k,1} \leq \eta(t+1)g(n)e^{-(A'_2-A'_1)n} \tag{4.8}$$

where

$$A_2^{'} = \frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}b} \ln\left[\frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}\mathbb{E}\,|<\Phi_{t,k}, v_j>|}\right]$$

and

$$A_1^{'} = \frac{\mathbb{E}\,|<\Phi_{t,k}, v_j>|}{b}\left(\frac{\epsilon\sqrt{1-\delta}}{(t+1)\sqrt{\eta}\mathbb{E}\,|<\Phi_{t,k}, v_j>|} - 1\right), 1 \le j \le \eta.$$

On another hand, from Markov inequality

$$\mathbb{P}\left\{\sum_{j>\eta} < \frac{S_n}{n}, v_j >^2 \ge \delta\epsilon^2\right\} \le \frac{g(n)}{\delta\epsilon^2}\sum_{j>\eta}\left[\mathbb{E} < \frac{S_n}{n}, v_j >^2\right].$$

Since

$$\left(\frac{1}{n}\sum_{t=1}^{\eta} < X_t, v_j >\right)^2 \le \frac{1}{n}\sum_{t=1}^{\eta} < X_t, v_j >^2,$$

we obtain

$$\begin{aligned}\mathbb{P}\left\{\sum_{j>\eta} < \frac{S_n}{n}, v_j >^2 \ge \delta\epsilon^2\right\} &\le \frac{g(n)}{\delta\epsilon^2}\sum_{j>\eta}\left[\mathbb{E} < X_0, v_j >^2\right]. \\ &\le \frac{g(n)}{\delta\epsilon^2}\sum_{j>\eta}\lambda_j. \qquad (4.9)\end{aligned}$$

Finally, from (4.8) and (4.9), we have :

$$\overline{I} = \mathbb{P}\left\{\| \frac{S_n}{n} \| > \epsilon\right\} \le \eta(t+1)g(n)e^{-(A_2^{'}-A_1^{'})n} + \frac{g(n)}{\delta\epsilon^2}\sum_{j>\eta}\lambda_j.$$

Thus the result.

Corollary 4.2.6 *Indeed, if* $\exists a > 0$*,* $\beta > 1$ *and* $g(n) = O(n^\gamma)(0 \le \gamma < 1)$ *such that*

$$\sum_{j>\eta}\lambda_j \le \sum_{j>n}\lambda_j \quad (n > \eta) \qquad (4.10)$$

and

$$\lambda_j \le aj^{-1-\beta} \quad \forall j \ge 1. \qquad (4.11)$$

Then

$$\sum_{n=1}^{+\infty}\mathbb{P}\left\{\| \frac{S_n}{n} \| > \epsilon\right\} \le \eta(t+1)\sum_{n=1}^{+\infty} n^\gamma e^{-(A_2^{'}-A_1^{'})n} + \frac{a}{\beta\delta\epsilon^2}\sum_{n=1}^{+\infty} n^{\gamma-\beta} < \infty. \qquad (4.12)$$

(i.e. $\frac{S_n}{n} \to 0$ *completely almost when* $n \to +\infty$ *with* $\frac{S_n}{n}$ *is an element of Hilbert space* H*).*

Proof 15 *From theorem we have that*

$$\mathbb{P}\left\{\| \frac{S_n}{n} \|> \epsilon\right\} \;\leq\; \eta(t+1)g(n)e^{-(A_2'-A_1')n} + \underbrace{\frac{g(n)}{\delta\epsilon^2}\sum_{j>\eta}\lambda_j}_{\widetilde{I}}.$$

The expression $\widetilde{I}$ *be bounded up by* $\frac{1}{\delta\epsilon^2}a\beta^{-1}n^{\gamma}n^{-\beta}$ *using lemma 4.2.4*
Finally, from (4.10) and (4.11) we deduce

$$\sum_{n=1}^{+\infty}\mathbb{P}\left\{\| \frac{S_n}{n} \|> \epsilon\right\} \;\leq\; \eta(t+1)\sum_{n=1}^{+\infty}n^{\gamma}e^{-(A_2'-A_1')n} + \frac{a}{\beta\delta\epsilon^2}\sum_{n=1}^{+\infty}n^{\gamma-\beta} < \infty.$$

That is to say that

$\frac{S_n}{n}$ *converges almost completely to* 0 *when* n *tends to* ∞.

Perspectives

In this section, we sketch some perspectives for possible future researches.

For chapter 2

1. see what are the conditions for obtaining a similar result for :
   (a) autoregressive processes of order p (AR(p), $p > 1$).
   (b) autoregressive Hilbertian processes $ARH(1)$.
   (b) autoregressive processes in Banach spaces $ARB(1)$.
2. study the cases of the models $ARMA$ and $GARCH$.

For chapter 4

All the results have not been stated since we replace the Hilbert space by the space of continuous functions on a compact ($[0, 1]$).

Bibliographie

[1] R.L. Taylor. A. Bozorgnia., R.F. Patterson. Limit theorems for dependent random variables. World Congress Nonlinear Analysts., 92 :1639–1650, 1996.

[2] D. Szynal. A. Kuczmaszewska. On complete convergence in a banach space. Internat. J. Math. Math. Sci., 17 :1–14, 1994.

[3] A.Shen. Bernstein-type inequality for widely dependent sequence and its application to nonparametric regression models. Abstract and Applied Analysis., Article ID 862602 :9, 2013.

[4] A.Shen. On the strong convergence rate for weighted sums of arrays of rowwise negatively orthant dependent random variables. RACSAM., 107(2) :257–271., 2013.

[5] A.Shen. On asymptotic approximation of inverse moments for a class of nonnegative random variables statistics. A Journal of Theoretical and Applied Statistics., 48(6) :1371–1379., 2014.

[6] V. Avanissian. Initiation à l'analyse fonctionnelle. 1^{re} édition. Presses Universitaires de France., 1996.

[7] D. Bosq. Linear processes in function spaces. Lecture Notes in Statistiques, springer., 2000.

[8] Z. Boulenoir. Tail probabilities and complete convergence for weighted sequences of *lnqd* random variables with application to first-order autoregressive processes model. International Journal of Statistics and Economics., 19 :1–16, 2018.

[9] R.G. Antonini. A. Volodin. D.H. Qiu., P.Y. Chen. On the complete convergence for arrays of rowwise extended negatively dependent random variables. Journal of the Korean Mathematical Society., 50(2) :379–392, 2013.

[10] R.G. Antoninic. A. Volodin. D.H. Qiu., K.C. Chang. On the strong rates of convergence for arrays of rowwise negatively dependent random variables. Stochastic Analysis and Applications., 29 :375–385, 2011.

[11] P. Erdös. On a theorem of hsu and robbins. Ann. Math. Statist., 20 :286–291, 1949.

[12] A. Gut. Complete convergence for arrays. Period. Math. Hungar., 25 :51–75, 1992.

[13] T.Z. Hu. Negatively superadditive dependence of random variables with applications. Chinese Journal of Applied Probability and Statistics., 16 :133–144., 2000.

[14] L. X. Zhang. J. F. Wang. A berry-esseen theorem for weakly negatively dependent random variables and its applications. Acta Mathematica Hungarica., 110 :293–308, 2006.

[15] F. Proschan. K. Joag-Dev. Negative association of random variables with applications. Annals of Applied Statistics., 11 :286–295, 1983.

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

[17] E.L. Lehmann. Some concepts of dependence. Ann. Math. Statist., 37 :1137–1153, 1966.

[18] L. Liu. Precise large deviations for dependent random variables with heavy tails. statistics and probability letters. Journal of Applied Probability., 79(9) :1290–1298, 2009.

[19] L. Liu. Necessary and sufficient conditions for moderate deviations of dependent random variables with heavy tails. Sci China Ser A Math., 53 (6) :1421–1434, 2010.

[20] A. Bozorgnia. N. Asadian., V. Fakoor. Rosenthal's type inequalities for negatively orthant dependent random variables. JIRSS., 5 :69–75, 2006.

[21] M. Ghosh. N. Ebrahimi. Multivariate negative dependence. Communications in Statistics-theory and Methods., 10 :307–337, 1981.

[22] C. M. Newman. Asymptotic independence and limit theorems for positively and negatively dependent random variables. Inequalities in Statistics and Probability, IMS Lectures Notes-Monograph Series., 5 :127–140, 1984.

[23] R.A. Davis. P.J. Brockwell. Time series : theory and methods. Springer Series in Statistics. Springer Verlag, New York., 2^{nd} Edition, 1991.

[24] H. Robbins. P.L. Hsu. Complete convergence and the law of large numbers. Proceedings of the National Academy of Sciences, USA., 33 :25–31, 1947.

[25] W. E. Pruitt. Summability of independent random variables. J. Math. Mech., 15 :769–776, 1966.

[26] Y.Y. Jiang. Q.Y. Wu. The strong consistency of m estimator in a linear model for negatively dependent random samples. Communications in Statistics-Theory and Methods., 40 :467–491., 2011.

[27] D.S. Stoffer. R.H. Shumway. Time series analysis and its applications with r exmples. Springer Series in Statistics. Springer Verlag, New York., 4^{th} Edition, 2016.

[28] A. Bozorgnia. R.L. Taylor., R.F. Patterson. A strong law of large numbers for arrays of rowwise negatively dependent random variables. Stochastic Analysis and Applications., 20 :643–656, 2002.

[29] V. K. Rohatgi. Convergence of weighted sums of independent random variables. Proc. Cambridge Philos. Soc., 69 :305–307, 1971.

[30] T.K. Chandra. S. Ghosal. Complete convergence of martingale arrays. J. Theo. Probab., 11 :621–631, 1998.

[31] A. Volodin. S.E. Ahmed., R.G. Antonini. On the rate of complete convergence for weighted sums of arrays of banach space valued random elements with application to moving average processes. Statist. Probab. Lett., 58 :185–194, 2002.

[32] A. Shen. Probability inequalities for *end* sequence and their applications. Journal of Inequalities and Applications., article ID98 :12, 2011.

[33] X.J. Wang. S.J. Wang. Precise large deviations for random sums of *end* real-valued random variables with consistent variation. Journal of Mathematical Analysis and Applications., 402 :660–667, 2013.

[34] S.H. Sung. On the exponential inequalities for negatively dependent random variables. Journal of Mathematical Analysis and Applications., 381 :538–545, 2011.

[35] H. C. Kim. T. S. Kim. On the exponential inequality for negative dependent sequence. Communications of the Korean Mathematical Society., 22 :315–321, 2007.

[36] M. H. Ko. T. S. Kim., D. H. Ryu. A central limit theorem for general weighted sums of *lnqd* random variables and its applications. Rocky Mountain J. of Math., 37 :259–268, 2007.

[37] E. Vaggelatou. T.C. Christofides. A connection between supermodular ordering and positive/negative association. Journal of Multivariate Analysis., 88 :138–151., 2004.

[38] A. Rosalsky. A. Volodin. T.C. Hu., D. Li. On the rate of complete convergence for weighted sums of arrays of banach space valued random elements. Theory Probab. Appl., 47 :455–468, 2001.

[39] HA. Azarnoosh. V.Fakoor. Probability inequalities for sums of negatively dependent random variables. Park J Stat., 21 (3) :257–264, 2005.

[40] X.J. Wang. S.H. Hu. W.Z. Yang., T.T. Liu. On the bahadur representation of sample quantiles for widely orthant dependent sequences. Filomat., 28(7) :1333–1343, 2014.

[41] X. Yang. X. Wang., M.B. Rao. Convergence rates on strong laws of large numbers for arrays of rowwise independent elements. Stochastic Anal. Appl., 11 :105–132, 1993.

[42] Y.B. Wang. X.J. Liu., Q.W. Gao. A note on a dependent risk model with constant interest rate. Statistics and Probability Letters., 82 :707–712, 2012.

[43] A. Volodin. S.H. Hu. X.J. Wang., T.C. Hu. A strong limit theorem for weighted sums of sequences of negatively dependent random variables. Journal of Inequalities and Applications., Article ID 383805 :8, 2010.

[44] A. Volodin. S.H. Hu. X.J. Wang., T.C. Hu. Complete convergence for weighted sums and arrays of rowwise extended negatively dependent random variables. Journal of the Korean Mathematical Society., 42 :2391–2401, 2013.

[45] C. Xu. S.H. Hu. X.J. Wang., L.L. Zheng. Complete consistency for the estimator of nonparametric regression models based on extended negatively dependent errors. Statistics : A Journal of Theoretical and Applied Statistics., 49(2) :396–407, 2015.

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

[47] W.Z. Yang. N.X. Ling. X.J. Wang., S.H. Hu. Exponential inequalities and inverse moment for *nod* sequence. Statistics and Probability Letters., 80 :452–461, 2010.

[48] W. NG. Kai. Y. Chen., A. Chen. The strong law of large numbers for extend negatively dependent random variables. Journal of Applied Probability., 47 :908–922, 2010.

[49] Y. Wang. Y. Chen., L. Wang. Uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional risk models. Journal of Mathematical Analysis and Applications., 401 :114–129, 2013.

[50] D. Cheng. Y. Wang. Basic renewal theorems for a random walk with widely dependent increments and their applications. Journal of Mathematical Analysis and Applications., 384 :597–606., 2011.

[51] K. Wang. X. Ma. Y. Wang., Z. Cui. Uniform asymptotic of the finite-time ruin probability for all times. Journal of Mathematical Analysis and Applications., 390 :208–223, 2012.

[52] M. Guan. Y.F. Wu. Convergence properties of the partial sums for sequences of end random variables. Journal of the Korean Mathematical Society., 49(6) :1097–1110, 2012.

[53] L. X. Zhang. A functional central limit theorem for asymptotically negatively associated dependent random variables. Acta Mathematica Hungarica., 86 :237–259, 2000.

Abstract

The dissertation is composed of four chapters. In the first chapter, we explain the basic notions and highlight some of the objectives of time series analysis. In chapter two, we study a a new concentration inequality and complete convergence of weighted sums for arrays of row-wise linearly negative quadrant dependent random variables, then in chapter three we demonstrate almost complete convergence of dependant random variables sequences with application to non-linear autoregressive processes model. Regarding the fourth and last chapter, we discuss the almost complete convergence of the value of the process of autoregressive Hilbertian of order one.

Keywords : autoregressive process, complete convergence, estimation, tail probabilities, linearly negative quadrant dependent ($LNQD$) sequence, random variables, weighted sums, nonlinear autoregressive models, extended negatively dependent (END) sequence, exponential inequality, autoregressive Hilbertian process, widely orthant dependent (WOD) random variables,