N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBÈS

THESE DE DOCTORAT

Présentée par

BERKANE KHEIRA

Domaine : Mathématiques Informatique
Filière : Mathématiques
Intitulé de la formation : Statistique, Mathématiques

Intitulée

Théorèmes limites et estimation récursive des modèles linéaires, applications aux processus autorégressifs

Soutenue le 23 Juin 2020
Devant le jury composé de :

Président :
Mr CHOUAF Abdelhak — MCA — à L'Université S.B.A.
Examinateur :
Mr AZZOUZI BADREDDINE — M.C.A — à ESN TLEMCEN
Directeur de thèse :
Mr BENAISSA Samir — Professeur — à L'Université S.B.A
Co-Directeur de thèse :
Mr MECHAB Boubaker — M.C. A — à L'Université S.B.A.
Année universitaire : 2019/2020

Acknowledgements

First of all, I would like to thank God, the almighty and merciful who gave me strength and the patience of accomplish this modest work.

The first person I want to thank is my director of thesis Professor **BENAISSA Samir** for orientation, trust, patience and especially his wise advice which have made a considerable contribution without which this work could not have done to the good port. That he finds in this work a living tribute to his high personality.

I like to thank Mister **MECHAB Boubaker** for his precious help, especially in this trying time that is the last straight line.

I would like also to thank each of the members of the jury for having been kind enough to participate in the defense.

I sincerly thank you Mister **CHOUAF Abdelhak** for the honor he gives me by presiding the jury. He always encouraged me and heavily advised.

I warmly thank Mister **BADREDDINE Azzouzi** for a careful reading and his relevant remarks, i am very honored by his presence.

I thank the members of the laboratory of statistics and stochastic processes of university of Djillali liabÃÍs of sidi bel abbÃÍs. I have always found support and encouragement.

I would like to express my appreciation to my friends and colleagues who told me provided their moral and intellectual support throughout my approach.

Finally, I thank all those who from near of far contributed to the realization of this work..

This thesis is dedicated to my parents and my brothers and my sister…

My husband and my daughter…

And to all my family and best friends…

Abbreviation list

AR	Autoregressive
MA	Moving average
ARMA	Autoregressive moving average
ARIMA	Autoregressive integrated moving average
END	Extended negatively dependent
UEND	Upper extended negatively dependent
LEND	Lower extended negatively dependent
WOD	Widely orthant dependent
WUOD	Widely upper orthant dependent
WLOD	Widely lower orthant dependent
$X_n = O(V_n)$	X_n is dominated by V_n
$X_n = o(V_n)$	X_n is negligible compared to V_n

Publication and Communication

International Publication:

Kheira Berkane, Samir Bennaissa. (2019): Complete convergence for widely orthant dependent random variables and its applications in autoregressives AR(1) models. International Journal of Mathematics and Computation, VL 30 NO. 2, 1-12.

International Communication:

The eight edition of "International Workshop on PerspectivesOn high dimensional Data Analysis(HDDA-VIII-2018)", Marakesh, Morocco. April 9-13, 2018
On the rate of strong convergence for a recursive probability density estimator of END and its applications.

National Communication:

Congrès des Mathématiciens Algériens CMA'2018, 12-13 mai 2018, Boumerdès, Algérie
Distributions asymptotiques des estimateurs de densité d'erreur dans les modèles autorégressifs du premier ordre.

La première édition des doctoriales nationales de mathématiques du 28 au 31 octobre 2017, Ecole normale supérieure Assia DJEBAR-Constantine, Algérie
Dérivation fractionnaire et applications.

Contents

General Introduction **8**

1 Time series **11**

1.1 Some fundamental concepts 11

1.1.1 The stochastic process 11

1.1.2 Stationarity 11

1.2 Description of a time series 14

1.3 Main goals 18

1.4 Schematic description of the complete study of a chronological series 19

1.4.1 Correction of data 20

1.5 white noise 23

1.6 Operators defined on a time series 25

1.6.1 Delay operator : 25

1.6.2 Difference operator order of d: 25

1.7 Types of model 25

1.7.1 Autoregressive process $AR(p)$: 26

1.7.2 Moving average process $MA(q)$: 26

1.7.3 $ARMA(p, q)$ model 28

1.8 Nonparametric estimate of the probability density 28

1.8.1 Estimation of the probability density by the kernel method 29

1.8.2 Asymptotic Expressions of Bias and Variance 29

1.8.3 Optimal theoretical choice of the smoothing parameter 30

1.9 Recursive kernel estimator 31

1.9.1 Some convergence results for the Wolverton-Wagner estimator 32

1.9.2 Some convergence results for the Deheuvels estimator 33

1.10 Some notions about convergence 33

1.11 Some properties on probabilities 34

2 Asymptotic distribution of the END random error in first-order autoregressive processes using the recursive kernel estimator 37

2.1 Introduction 37

2.2 Some auxiliary results 39

2.3 Asymptotic Distribution of $\hat{f}_n$ at a fixd point 40

3 Complete convergence for recursive probability density estimator of LNQD orthant dependent variables 48

3.1 Introduction 48

3.2 Main results 50

3.3 Some lemmas 51

3.4 Proof of Main Results 53

4 Complete convergence for widely orthant dependent random variables and its applications in autoregressive $AR(1)$ models 59

4.1 Introduction 59

4.2 Some lemmas 60

4.3 Main Results and proofs 63

4.4 Application results in the first-order autoregressive $AR(1)$ model 67

4.4.1 The $AR(1)$ model 67

Conclusion et Perspectives 74

List of Figures

1.1 Air passengers 12

1.2 Monthly index of consumer prices I_t 14

1.3 Passenger traffic of the SNCF in second class 15

1.4 Relative monthly increase in the price index 15

1.5 Medium-term evolution of the relative monthly increase in the price index 16

1.6 United States population 16

1.7 Number of strikes united state, 1951-1980 17

1.8 All star Baseball games, 1933-1980 17

1.9 Monthly number of accidental deaths in the USA, 1973-1978 18

1.10 Quarterly sales of sunscreens, trend, seasonal factors and irregular fluctuations 23

General Introduction

In statistics, we designate by chronological series, or time series, the modeling of a series of random and sequentially observed events, generally on a temporal scale. Many real data flows admit such a representation, we can think for example of the consumption of electricity, the evolution of a stock market action, the gross domestic product, the tide cycle or the progress of a process chemical. Although astronomers and meteorologists can claim paternity, the main advances in the formalization of time series seem to be the task of statisticians and econometricians. Today we find the legacy in the fields of economics and finance of course, but also in engineering, signal processing or the natural and social sciences. It's from the publication of the pioneering work of Box and Jenkins in 1970 that was born the theory of time series as we know it today. The latter was motivated by the weak predictive power of the structural models of Keynesianism in force at the time and proposes a completely innovative approach. The ARMA process, forming the heart of the work and on which we will have the opportunity to return in detail later, relies exclusively on the past observations of a curve to intuit a global linear behavior disturbed by a noise. random center. The main feature of a time series is the strong dependence and considerable practical interest linking two consecutive observations, at the origin of the dynamics of the models. The latter are therefore all the more apt to reflect actual flows and natural events in which the chronological evolution is manifest. Among the range of applications from the theory, it is mainly modeling and prediction that focus all attention. While we are looking from one side to build the best model for a dataset without prejudging the events taking place outside the study interval, we focus on the other on optimal learning of observations to infer future behavior.

From the ARMA linear process came a series of derived models, more and more general, with as a guideline the formalization of an extended class of real phenomena. The latter take into account non-stationarity, heteroscedasticity, and even non-linearity. Although this Since this thesis focuses solely on a particular subset of processes, it seemed essential to give in the first chapter the succinct, and therefore not exhaustive, history of the time series currently in force. It will also be necessary to explain all the concepts that we approach - in italics - in this introduction, before presenting in a descriptive way the model of Box and Jenkins, its properties then its major evolutions. Once the framework is well defined, we will discuss the autoregressive process, linear model on which our entire study will be focused. The rigorous analysis of linear time series has expanded considerably at the end of the last century. While Harvey in 1991 and Kitagawa and Gersch in 1996 apprehend the time series by a Bayesian approach based mainly on state-space models treated with a Kalman filter, it is on the frequentist approach that we will focus our attention. In this respect, we can cite the work of Jenkins and Watts in 1968, Hannan in 1970, Bloomfield in 1976, Priestley and Brillinger in 1981, and Fuller in 1995. We will particularly emphasize the work of Brockwell and Davis during the 1990s on which we have largely relied upon, source of our theoretical reminders.

Random shocks in the autoregressive process are usually treated as white noise, in any case one of the usual working hypotheses guaranteeing both the consistency of the estimators and their asymptotic normality. In the mid-twentieth century, Durbin and Watson adopted as a subject of study the standard linear regression model, the random perturbation of which results from a first-order autoregressive process, abandoning in fact the residual whiteness, but also the consistency of the estimate. They suggest then a statistic still used today in the field of econometrics, at the origin of a relatively elementary test procedure allowing quite often to reject a hypothesis of absence of residual correlation. In 1970, Durbin, confronted with the inferential consequences of the presence of a residual autocorrelation and placing himself in the chronological framework, proposes a subtle revision of the procedure better known today under the name of H-test. Thus.

Our thesis is presented in four chapters. In the first chapter we recall the notations and the tools used on the stochastic processes, and in particular, we recall the definition of the processes AR(p), MA(q), ARMA(p,q), and the non parametric estimate of the probability density function by kernel method especially recursive kernel estimator.

In the seconde chapter, we consider the asymptotic distributions of the error estimators in the first order autoregressive model by using the recursive density estimator of the probability density function for a sequence of extended negatively dependent random variables.

In the third chapter, we establish the complete convergence for a kind of hazard rate function estimator by considerate a kind of recursive density estimator of the probability density function for a sequence of linear negatively quadrant dependent random variables .

The last chapter is devoted to prove new exponential inequality for a new case of dependence WOD for the distributions of sums of widely orthant deprndent (WOD, in short) random variables. Using these inequality. The results are applied to the first-order autoregressive processes AR(1) model. This work was published in International Journal of Mathematics and Computation.

Chapter 1

Time series

1.1 Some fundamental concepts

1.1.1 The stochastic process

Consider a probability space $(\Omega, \mathcal{F}, P)$, a set of indices $\mathit{\Gamma}$ and a metric space S provided with the Borel tribe $\mathcal{B}(S)$.

Definition 1.1 *We call "stochastic process" a family of random variables X_t set to $(\Omega, \mathcal{F}, P)$, indexed by $t \in T$ and with values in S. For any realization $\omega \in \Omega$ the family $(X_t = X_t(\omega))$ is a "trajectory" of the process.*

It is in the trajectory of a process observed on a subset of $\mathit{\Gamma}$ that we will associate the notion of "time series".

1.1.2 Stationarity

Stationarity reflects the ability of a process not to depend on the index temporal. The latter is therefore fully described by its stationary law which, by definition, no longer changes over time. It is thus understood that such a property is certainly of practical interest considerable, but also has a strong theoretical impact since it is found as an assumption underlying many results. We generally distinguish stationarity in the strict sense of stationarity in the weak sense. To define them, consider a process X_t set to $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, the set of integrable square random variables.

Definition 1.2 *We say that the process X_t is "strictly stationar" if for all $k \in \mathbb{N}^*$ and all time shift $h \in \mathbb{Z}$ the law of the vector $(X_1, \ldots, X_{1+k})$ is the same as that of the vector $(X_{1+h}, \ldots, X_{k+h})$.*

Strict stationarity is a very strong working hypothesis, necessarily difficult to verify in practice when the process is not gaussian. That's why we introduced a notion of less restrictive stationarity.

Definition 1.3 *We say that the process X_t is "weakly stationary" if for any time offset $h \in \mathbb{Z}$, $\mathbb{E}(X_0)$ and $Var(X_0)$ are constant, and $Cov(X_0, X_h)$ depends only on h.*

It is also said that the process is "second order stationary", in relation to stabilization of its variance. It is this property of stationarity that we will implicitly refer to by the following. Note that strict stationarity naturally implies weak stationarity. The example the most trivial of stationary process is white noise.

Definition 1.4 *A time series comes from the realization of a family of random variable $\{X_t, t \in I\}$, or the set I is a time interval that can be discrete or continuous.*
for this thesis we use the whole $I = \{0, 1, \dots, T\}$, or T the total number of observations.

Example 1.1 *The figure represents the global total of air passengers per month between* 1949 *and* 1960*. Note that the points are connected by lines (which are there to look pretty and do not have of special significance). The data (Air Passengers) are available in* $\mathbb{R}$.

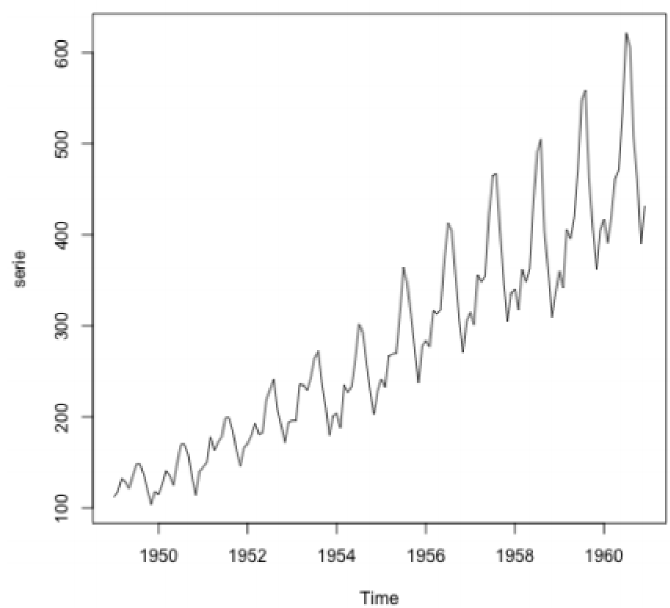

Figure 1.1: Air passengers

The objective of the time series study is to make predictions about the evolution of the series. Here is a non-exhaustive list of mathematical models that can be used :

- Regression. It is assumed that X_t is polynomial in t, for example $X_t = \alpha_2 t^2 + \alpha_1 t + \alpha_0 + \epsilon_t$ (with ϵ_t a random noise). The coefficients are estimated by $\hat{\alpha}_2, \hat{\alpha}_1, \hat{\alpha}_0$ (from the values $X_1, \dots, X_n$). So with the data of $X_1, \dots, X_n$, we will make the prediction $\hat{X}_{n+1} = \hat{\alpha}_2(n+1)^2 + \hat{\alpha}_1(n+1) + \hat{\alpha}_0$ value of X_{n+1},
- Exponential smoothing,
- ARMA models, to remove trends and seasonality from the series (periodicity). These models are numerically heavier, but more efficient.
  Challenges (in order):

- Define a model with a finite number of parameters.
- Estimate the model parameters.
- Check the quality of the fit of the model, compare different models (we will be able to cut the data into a learning sample and a test sample).
- Make predictions.

Example 1.2 *We can think for example of the number of travelers using the train, to the relative monthly increase in the price index or to the occurrence of a natural phenomenon (such as the number of sunspots).*
This series of observations of a family of real random variables noted $(X_t)_{t\in\Theta}$ is called serie chronological (or temporal). In the continuation of this course, we will note it

$$(X_t)_{t\in\Theta} \quad or \quad \{X_t, t \in \Theta\},$$

where the set Θ is called the time space that can be

- ***Discreet*** *(number of SNCF passengers daily, maximum temperature ...). in that case, $\Theta \in \mathbb{Z}$. The observation dates are most often equidistant, for example, monthly statements, quarterly... These equidistant dates are then indexed by integers: $t = 1, 2, \dots, T$ and T is the number of observations. Therefore we have observations of the variables $X_1, X_2, \dots, X_T$ from the family $(X_t)_{t\in\Theta}$ or $\Theta \subset \mathbb{Z}$ (most of the time $\Theta = \mathbb{Z}$). So if h is the time interval separating two observations and t_0 the moment of the first observation, we have the following schema:*

$$\begin{array}{cccc}
t_0 & t_0+h & \dots & t_0+(T-1)h \\
\downarrow & \downarrow & \dots & \downarrow \\
X_{t_0} & X_{t_0+h} & \dots & X_{t_0+(T-1)h} \\
\downarrow & \downarrow & \dots & \downarrow \\
X_1 & X_2 & \dots & X_T
\end{array}$$

- ***continued***_(radio signal, result of an electrocardiogram ...). The time indexe is at values in an interval of_ $\mathbb{R}$ _and one has (at least potentially) an infinity of observations from of a process_ $(X_t)_{t\in\Theta}$ _or_ Θ _is an interval of_ $\mathbb{R}$_. Such a process is said to be continuous. The methods presented in this context are different from those for chronological time series discreet and presented in the sequel._

1.2 Description of a time series

We consider that a time series X_t is the result of different fundamental components:

- the trend (or trend) Z_t represents the long-term evolution of the studied series. It translates the "average" behavior of the seerie.
  Examples of time series :
  For example, the seerie of the figure 1 trend to increase linearly.

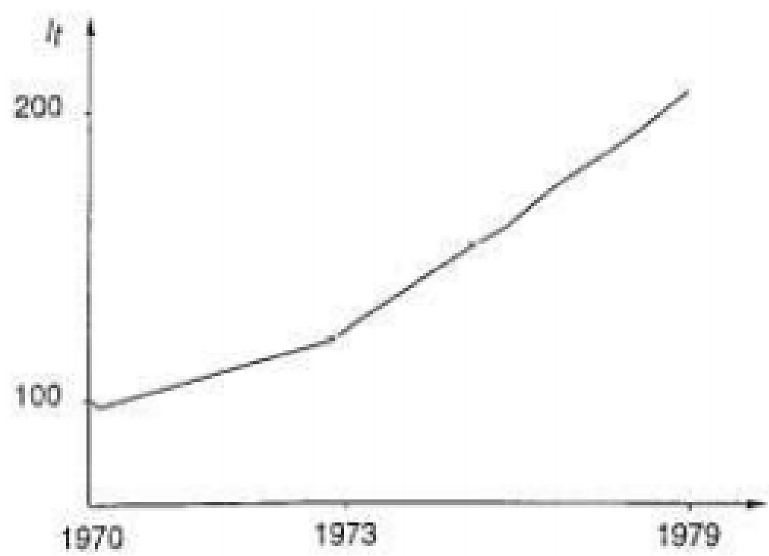

Figure 1.2: Monthly index of consumer prices I_t

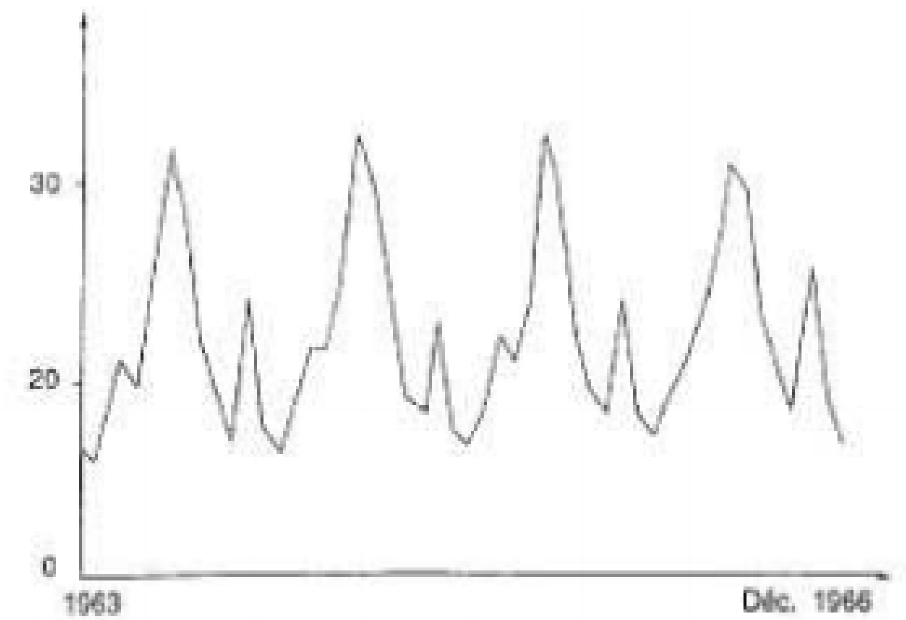

Figure 1.3: Passenger traffic of the SNCF in second class

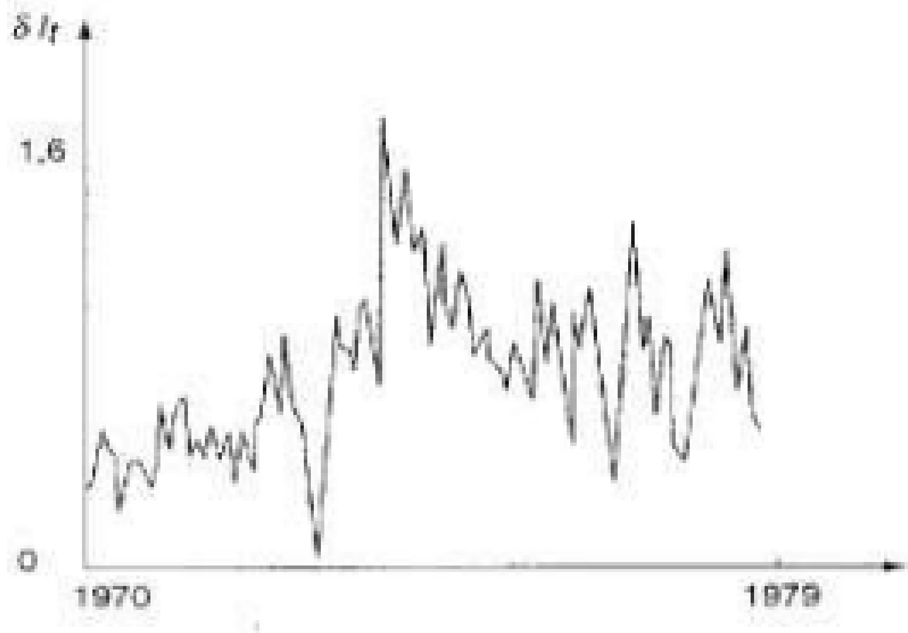

Figure 1.4: Relative monthly increase in the price index

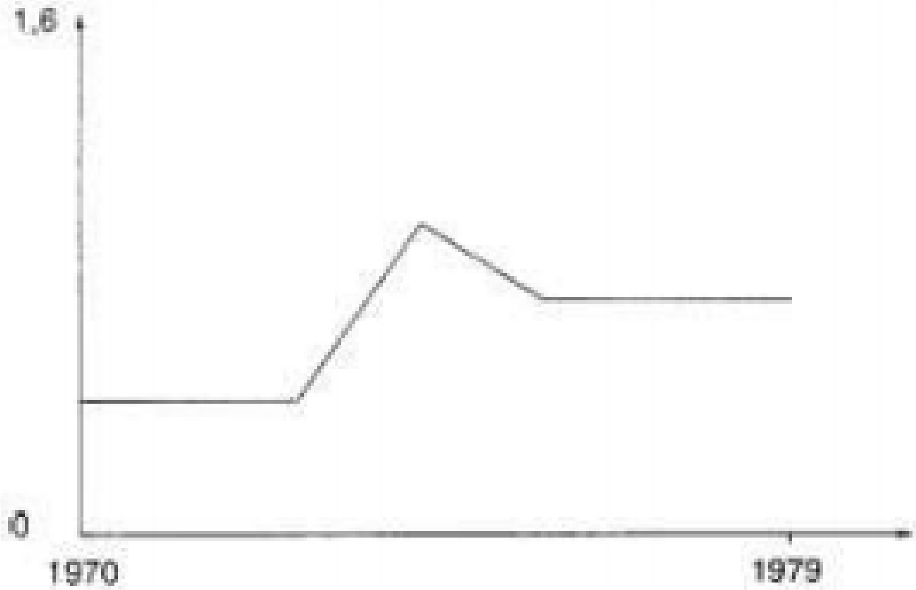

Figure 1.5: Medium-term evolution of the relative monthly increase in the price index

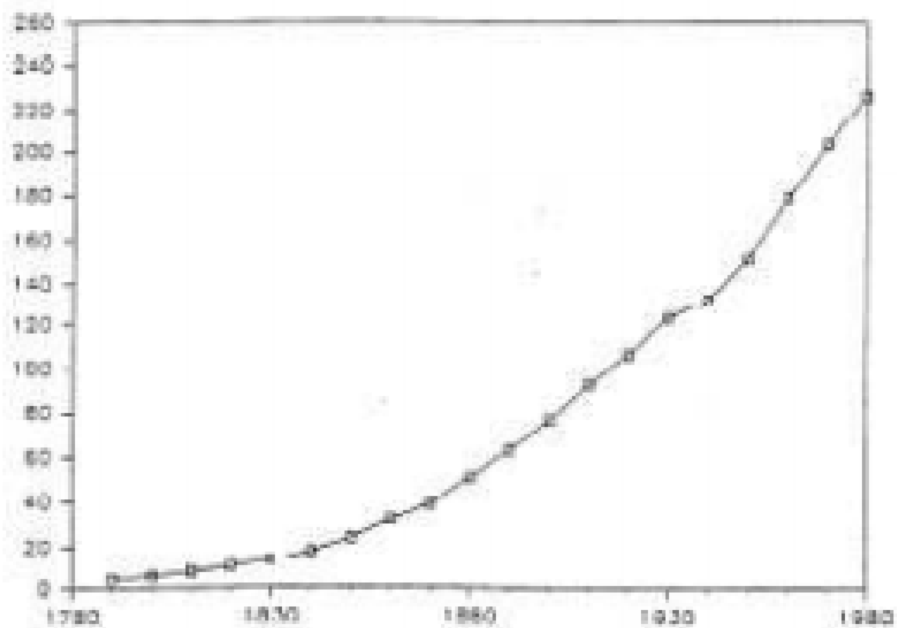

Figure 1.6: United States population

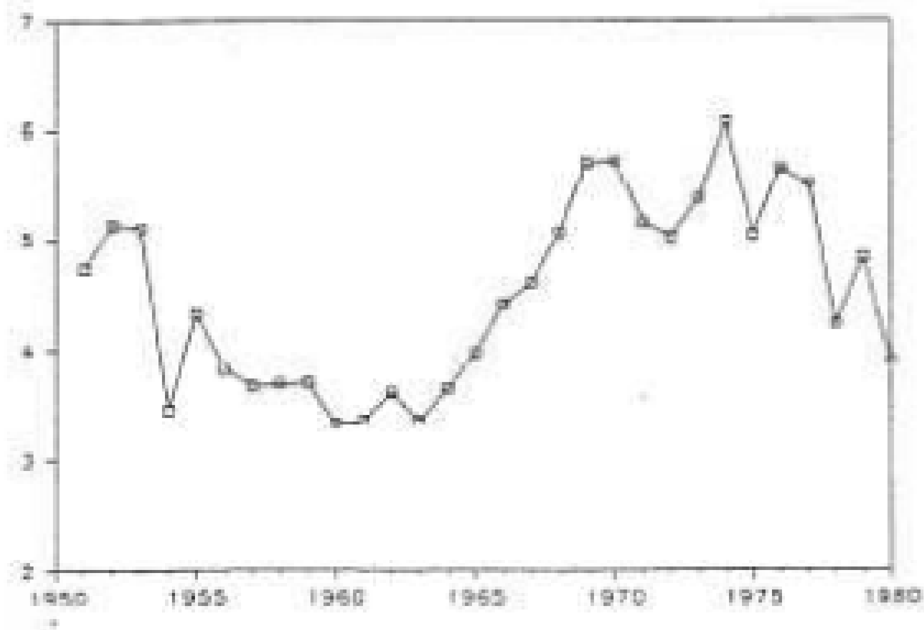

Figure 1.7: Number of strikes united state, 1951-1980

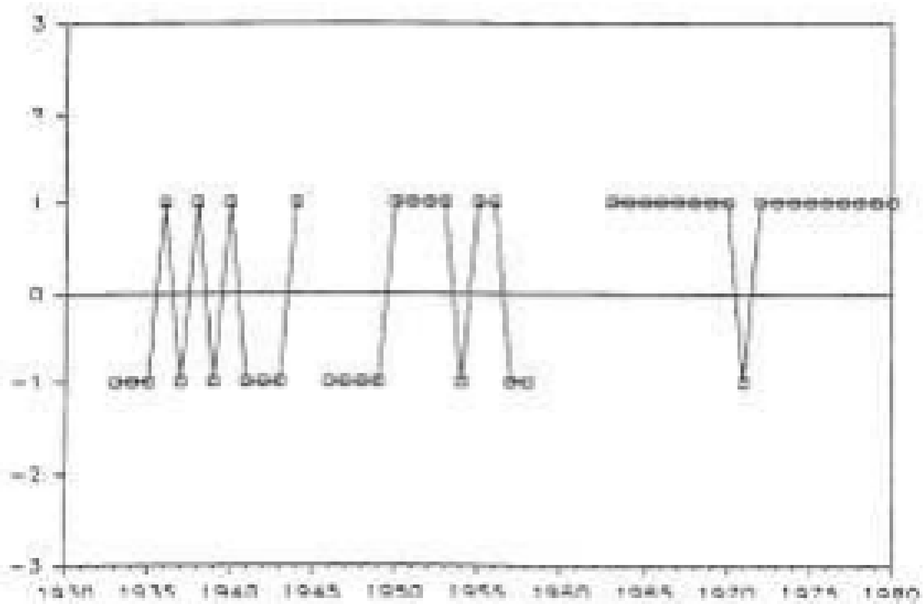

Figure 1.8: All star Baseball games, 1933-1980

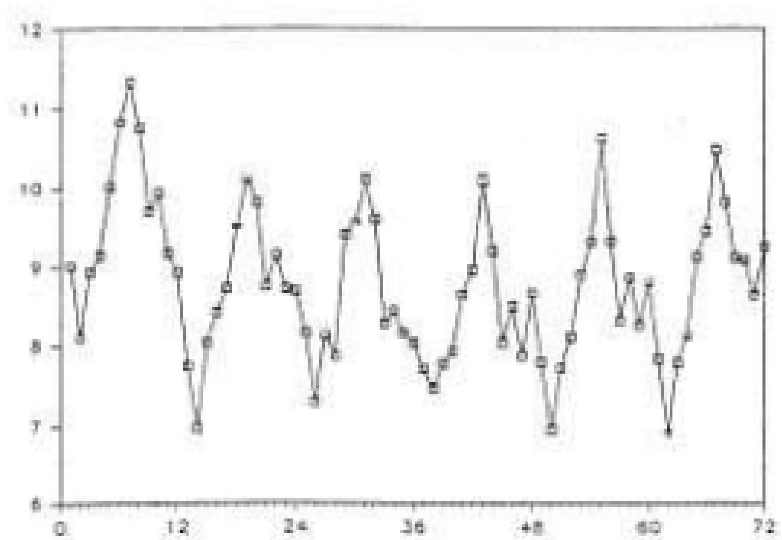

Figure 1.9: Monthly number of accidental deaths in the USA, 1973-1978

- **The seasonal component** (seasonality) S_t corresponds to a phenomenon which is repeated at regular (periodic) intervals of time. In general, it is a seasonal phenomenon from which the term of seasonal variations.
  For example, The series of the figure 1.3. presents regular cycles over time and likewise amplitude.
- **The residual component** (or noise or residue) ϵ_t corresponds to irregular fluctuations, in general low intensity but random in nature. We also talk about aleas.
  For example, The series of the figure 1.4. rather irregular behavior: there is like a kind of low amplitude noise that disturbs the data.
- **Accidental phenomena** (strikes, exceptional weather conditions, financial crash) may intervene in particular.
  For example, the series of the figure 1.5 has two breaks.
- Another component sometimes studied in a specific way relates to **cyclical phenomenon**: this is often the case in climatology and economics (example: recession and expansion...). This is a phenomenon but in contrast to seasonality over periods that are not fixed and generally long. Without specific information, it is generally very difficult to dissociate trend and cycle.

1.3 Main goals

The study of a chronological series allows to analyze, to describe and explain a phenomenon during time and draw conclusions for decision-making (marketing...).

This study also makes it possible to control, for example for the management of stocks, the control of a chemical process...More generally, we can already pose some problems when studying a chronological window.

But one of the main objectives of the study of a chronological series is the projection of predict future values $X_{T+h}(h = 1, 2, 3, \ldots)$ the chronological sequence from its observed values until the time $T = X_1, X_2, \ldots, X_T$. The prediction of the time series at time $T + h$ is noted $\hat{X}_T(h)$ and in general is different from the real value X_{T+h} what does the seerie take at time $T + h$. To measure this difference, we will define the prediction error by the difference $\hat{X}_T(h) - X_{T+h}$ "on average" with the idea that plus h is big. bigger is the mistake. Precision interval, defined by the values $\hat{X}_T^{(1)}(h)$ and $\hat{X}_T^{(2)}(h)$ is likely to contain the unknown value X_{T+h}. The quality of the prediction can be measured based on 80% of observations, then simulating a prediction on the 20% of observations remaining. This technique is also useful for :

- The seeries that contain "holes",
- Measure the effect of an accidental phenomenon(error,...).

There are many other immediate objectives for the study of time series. For example, if two series are observed we may wonder what influence they exert on each other. Noting X_t and Y_t the two series in question. for example, if there are relations of the type

$$Y_t = a_1 X_{t+1} + a_3 X_{t+3}.$$

1.4 Schematic description of the complete study of a chronological series

One of the main objectives of the study of a chronological series is the forecast future values of this series. For that, we need to know or at least to model the production mechanism of the chronological series.

Note that the variables X_t are most often not independent (we can expect indeed that observations relatively close in time are linked) nor identically distributed (in most of the cases, the phenomenon evolves, changes over time, which means that variables that describe it are not equidistributed). This requires statistical methods of treatment and specific modeling since especially in a standard setting (that of the description of a sample), classical statistical methods are based on assumptions of independence.

Schematically, the main stages of treatment of a chronological sequence are the following:

1. Correction of data,
2. Observation of the series,
3. Modelization (with a finite number of parameters),
4. Analysis of the series from its components,
5. Diagnosis of the model, adjustment of the model,
6. prediction (forecast).

1.4.1 Correction of data

Before embarking on the study of a chronological series, it is often necessary to treat, modify the raw data. For example,

- Evaluation of missing data, accidental data replacement, ...,
- Breaking into sub-series,
- Standardization to reduce to fixed-length intervals. For example, for data monthly, we go back to the standard month by calculating the daily average over the month (total observations on the month divided by the number of days in the month),
- data transformation: for various reasons, sometimes we may have to use transformed data. For example in economy, we use the transform family of BoxCox:

$$Y_t = \frac{1}{\lambda}[(X_t)^\lambda - 1].$$

Observation of the series

A general rule in Descriptive Statistics is to start by looking at the data before making any calculations. So, once the set is corrected and pre-adjusted, we trace his graph it's a tell the coordinate curve (t, X_t).

Modelization

A model is a simplified image of the reality of translating the operating mechanisms of the phenomenon studied and makes it easier to understand them. One model may be better than another to describe the reality.
There are mainly two types of models:

- **The determinist models**: These models belong to Descriptive Statistics, they only involve the calculation of probabilities in an underlying way and assume that the observation from date to date t which is a function of time and a variable ϵ_t centered as an error to the model, representing the difference between the reality and the proposed model:

$$X_t = f(t, \epsilon_t),$$

We assume furthermore that the ϵ_t are decorrelated.
The two most commonly used models of this type are:

1. **Additive model:** It is the "classic model of decomposition" in the treatment of adjustment models, The variable X_t is written as the sum of three terms:

$$X_t = Z_t + S_t + \epsilon_t,$$

or Z_t represents the trend (determinist), S_t is the seasonality (also determinist) and ϵ_t the random components i.i.d (errors in the model).

2. **Multiplicative model:** The variable X_t is written at the end of error almost like the product of the trend and a seasonality component.

$$X_t = Z_t(1 + S_t)(1 + \epsilon_t),$$

the adjustment here is multiplicative and intervenes in the models $ARCH$.

- **Stochastic models:** They are of the same type as the models of determinists to this close noise variables ϵ_t are not i.i.d but have a non-zero correlation structure.

$$\epsilon_t = g(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \eta_t) \text{or} \eta_t \text{is the error term.}$$

The class of the most commonly used models of this type is model classes $SARIMA$ and its sub-models $ARMA, ARIMA, \dots$.

The particular case where the functional relation g is linear is very important and widely used, it leads to linear regressive models, for example a model of order 2 with coefficients autoregressive a_1, a_2 is given by:

$$\epsilon_t = a_1 X_{t-1} + a_2 X_{t-2} + \eta_t,$$

or η_t is a white noise that is an unrelated zero average random variable.

The two types of models above induce very particular forecasting techniques. schematically, we are interested first of all in the trend and the seasonality of the eventuality that we isolate first.

Then we try to model them, estimate them and finally we eliminate them from the series, these two operations are called the detendancialisation and the seasonal adjustment of the series, Once these components are removed we get the random serie ϵ_t:

- For deterministic models, this series will be considered decorrelated and there is nothing left to do,
- For stochastic models, we obtain a stationary series which means that the successive observations of the series are identically distributed but not necessarily independent that is for modeling.

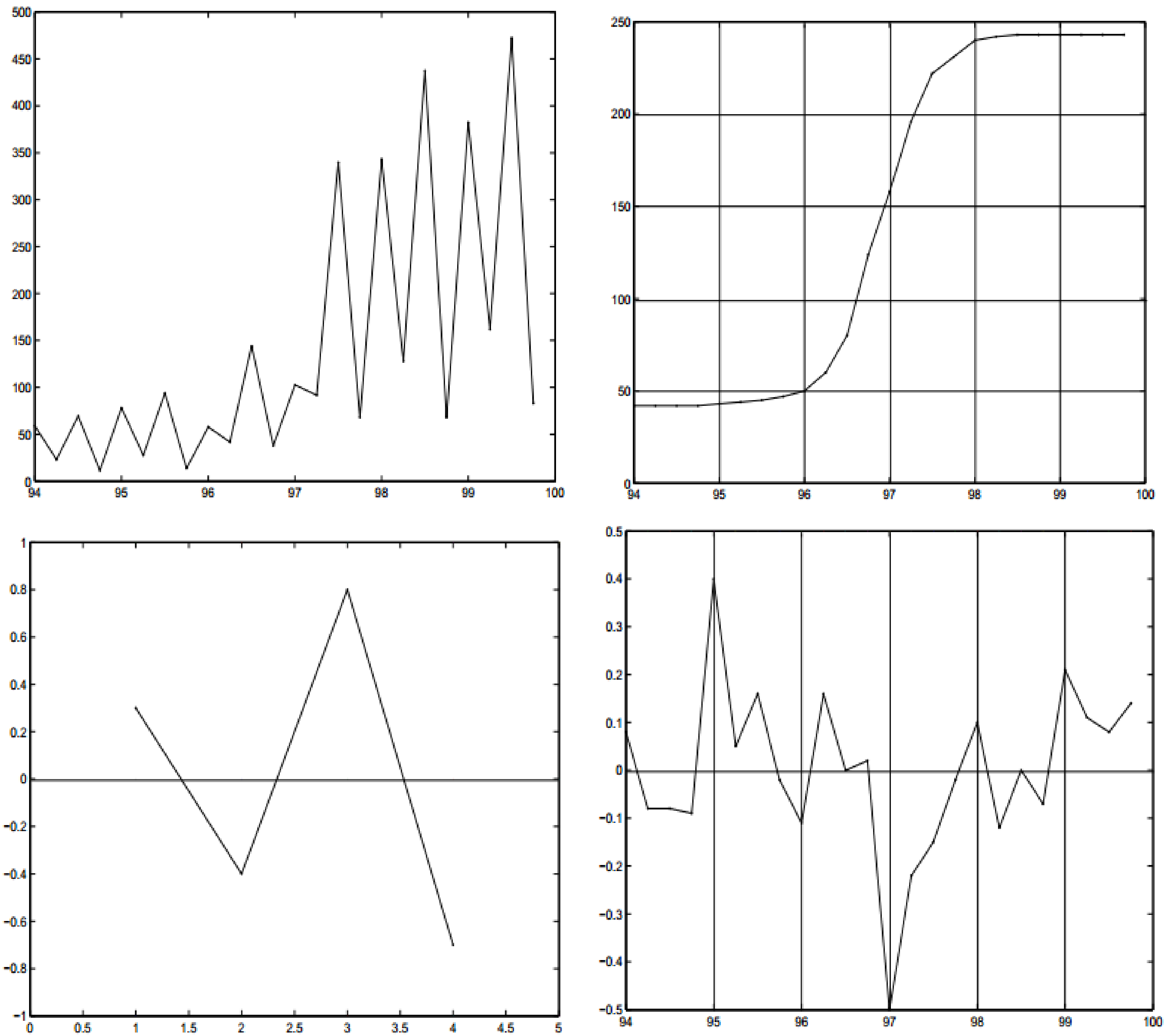

Figure 1.10: Quarterly sales of sunscreens, trend, seasonal factors and irregular fluctuations

1.5 white noise

Definition 1.5 *we say that the sequence of the random variables $\{\epsilon_t\}$ is a weak white noise if it has the following properties :*

- $\mathbb{E}(\epsilon_t) = 0$, *for* $t \in \mathbb{Z}$,
- $\mathbb{E}(\epsilon_t^2) = \sigma^2 \neq 0$ *and constant,*

- *$cov(\epsilon_s, \epsilon_t) = 0$ if $t \neq s$.*

In other words, the random variables $\{\epsilon_t\}$ are of zero mean, of constant variance and uncorrelated.

We say that $\{\epsilon_t\}$ is a strong white noise if it is a weak white noise and that the random variables $\{\epsilon_t\}$ are i.i.d.

Notation

1. If $\{\epsilon_t\}$ low white noise, so $\{\epsilon_t\} \sim \mathrm{WN}(0, \sigma_\epsilon^2)$,
2. If $\{\epsilon_t\}$ loud white noise, so$\{\epsilon_t\} \sim \mathrm{I.I.D}(0, \sigma_\epsilon^2)$.

stationarity

An important property of time series is stationarity. This property is necessary to apply some theorems on causality. The following definition presents the type of stationarity most used.

Definition 1.6 *A sequence $\{X_t, t \geq 0\}$ of random variables is said to be stationary of the second order if it checks the following properties :*

- $\mathbb{E}(X_t) = \mu < \infty$,
- $\mathbb{E}(X_t^2) < \infty$*, $Cov(X_s, X_{s+t})$=$Cov(X_{s-1}, X_{s-1+t})$=$\ldots$, $Cov\{X_0, X_t\}$ for all $s, t \in \mathbb{N}$.*

Remark 1.1 *we call $Cov\{X_t, X_{t+h}\}$ the autocovariance (ACV) on the horizon h and we note it :*

$$r^X(h) = Cov(X_t, X_{t+h}).$$

we can take this function and divide it by variance of $\{X_t\}$ to obtain a new function that we will call autocorrelation. we note it :

$$\rho^X(h) = \frac{r^X(h)}{r^X(0)},$$

with the following properties:

1. *$-1 < \rho^X(h) \leq 1$,*
2. *$\rho^X(h) = 0$ means that observations X_t and X_{t+h} are uncorrelated,*
3. *$\rho^X(h) = \pm 1$ means that the correlation is perfect (negative or positive).*

1.6 Operators defined on a time series

1.6.1 Delay operator :

Definition 1.7 *the delay operator B is defined as follows :*

$$B(X_t) = X_{t-1}.$$

Remark 1.2

$$B^n(X_t) = X_{t-n}, \text{ for } \forall n \in \mathbb{N}.$$

1.6.2 Difference operator order of d:

Definition 1.8 *we define $\triangle_d$ the order difference operator d as the linear operator such that :*

$$\triangle_d = X_t - X_{t-d} = (1 - B^d)X_t.$$

We can also take the order operator one and apply it several times :

Example 1.3

$$\begin{aligned} \triangle^2(X_t) &= \triangle(\triangle(X_t)), \\ &= \triangle(X_t - X_{t-1}) = (1 - B)(X_t - X_{t-1}), \\ &= X_t - 2X_{t-1} + X_{t-2}. \end{aligned}$$

these operators can be used to transform a non-zero mean process into a zero-average process. It can also be used to remove the season's song of the series. In the next section, we use them to better represent the models.

1.7 Types of model

One-dimensional case :

Some frequently used models are considered for a time series $\{X_t\}$. We start with two simple models: the autoregressive model and the moving average model. We continue with the ARMA model that combines these two models.

1.7.1 Autoregressive process $AR(p)$:

the first self-reinforcing processes were introduced by George Udny yule. Yule uses the first autoregressive model to model the sunspot time series rather than the Schuster periodogram method. An autoregressive process is a process where an observation is written at time t as a linear combination of past observations plus some white noise.

Definition 1.9 *The sequence $\{X_t, t \geq O\}$ is a autoregressive process of order p $(p > 0)$ if it can be written in the following form :*

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + \epsilon_t, \ \text{or} \ \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2).$$

The $\phi_k(k = 1, \dots, p)$ are the parameters of the model.
In this case, we write $\{X_t \sim AR(p)\}$, in the same way, we can rewrite a process $AR(p)$ with a polynomial $\phi(B)$ which will multiply X_t this time :

$$\phi(B)X_t = \epsilon_t \ \text{with} \ \phi(B) = 1 - \phi_1 B - \phi_2 B^2 \dots \phi_p B^p.$$

Example 1.4 *A process $AR(1)$ take the following form :*

$$X_t = \phi X_{t-1} + \epsilon_t, \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2).$$

1.7.2 Moving average process $MA(q)$:

It is Eugen Slutzky who in his article introduced for the first time moving average processes. The following definition presents this process.

Definition 1.10 *We say that the sequence $\{X_t, t \geq 0\}$ is a moving average process of order q if it can be written in the following form :*

$$X_t = \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t, \ \text{or} \ \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2),$$

or the $\theta_k(k = 1, 2, \dots, q)$ are the parameters of the model, in this case, we note $\{X_t\} \sim MA(q)$.

Example 1.5 *A process $MA(1)$ takes the following form :*

$$X_t = \theta \epsilon_{t-1} + \epsilon_t, \ \text{or} \ \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2).$$

We can use the delay operator B to write this program in another form, so we will have a polynomial in B that will multiply ϵ_t :

$$X_t = \theta(B)\epsilon_t \ \text{with} \ \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q.$$

Remark 1.3 *If $\{X_t\} \sim MA(q)$, so :*

1. *X_t is stationary,*
2. *$\rho^X(h) = 0$ for $\forall h > q$.*

Definition 1.11 *A process is called causal if there is a real $\{a_k\}$ sequence such as $\sum_{k=0}^{\infty}|a_k| < \infty$ and that :*

$$X_t = \sum_{k=0}^{\infty} a_k \epsilon_{t-k}.$$

Sometimes, when we speak of a causal process, we say that it has a representation $MA(\infty)$.

Remark 1.4 *any process $MA(q)$ is causal.*

Definition 1.12 *A process is said to be invertible if there is a real sequence $\{b_k\}$ real such that $\sum_{k=0}^{\infty}|b_k| < \infty$ and :*

$$\epsilon_t = \sum_{k=0}^{\infty} b_k X_{t-k}.$$

Another way of saying that a process is invertible is to say that it has a representation $AR(\infty)$.

Remark 1.5 *With this definition, any process $AR(p)$ is invertible.*

Theorem 1.1 *An autoregressive process $AR(p)$ is causal and stationary if and only if its polynomial $\phi(Z)$ is such that :*

$$\phi(Z) \neq 0 \text{ with } Z \in \mathbb{C} \text{ such that } |Z| \leq 1.$$

In other words, all roots of $\phi(Z)$ are greater than 1 standard.

Example 1.6 *The following process $AR(2)$ is stationary and causal :*

$$X_t = -\frac{1}{4}X_{t-1} + \frac{1}{8}X_{t-2} + \epsilon_t, \text{ or } \epsilon_t \sim WN(0, \sigma_\epsilon^2).$$

Indeed, we can rewrite it as :

$$\epsilon_t = X_t(1 + \frac{1}{4}B - \frac{1}{8}B^2),$$

so, we have that :

$$\phi(Z) = -\frac{1}{8}Z^2 + \frac{1}{4}Z + 1 = Z^2 - 2Z - 8 = (Z+2)(Z-4),$$

the roots are outside the unit circle, so the process is stationary and causal.

Theorem 1.2 *A moving average process $MA(q)$ is invertible if and only if its polynomial $\theta(Z)$ is such that :*

$$\theta(Z) \neq 0 \text{ with } Z \in \mathbb{C} \text{ such that } |Z| \leq 1.$$

We note the resemblance of this statement with the stationarity and causality theorem for autoregressive processes.

1.7.3 $ARMA(p, q)$ model

the AR and MA models may be perfect in some cases, but it may be necessary to estimate a large number of parameters to adjust the model. these estimates will tend to be unclear. Moreover, if a model containing p parameters is appropriate for the situation. it is not good to try to fit a model that will contain more than p.
$ARMA$ models consist of having an autoregressive part and a moving average part. Herman Wold showed that the ARMA processes could be used to model any stationary series as long as the orders p and q are well chosen. Bosc and Jinkins worked to develop a methodology for estimating the model of a time series.
It can be easy to check the causality and invertibility of the **ARMA** model.

Definition 1.13 *A process is said $ARMA(p, q)$ if there are real sequences $\{\phi_k\}$ and $\{\theta_k\}$ such that :*

$$X_t - \sum_{k=1}^{p} \phi_k X_{t-k} = \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}, \text{ with } \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2).$$

We can also use polynomials $\phi(B)$ and $\theta(B)$ to rewrite this model in the form :

$$\phi(B)X_t = \theta(B)\epsilon_t,$$

with $\phi(B) = 1 - \phi_1 B - C\phi_2 B^2 - \ldots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q$.
We notice that $\{X_t\} \sim ARMA(p, q)$.

Remark 1.6 *We note some properties for $ARMA(p, q)$ models :*

1. *If $p = q = 0$, we have $X_t = \epsilon_t \sim WN(0, \sigma_\epsilon^2)$,*
2. *If $p = 0$ and $q \neq 0$, we have $\{X_t \sim MA(q)\}$,*
3. *If $q = 0$ and $p \neq 0$, we have $\{X_t \sim AR(p)\}$.*

1.8 Nonparametric estimate of the probability density

In many applications, the density f is unknown and there is a n sample, $X_1, X_2, \ldots, X_n$ random variables that are independent and identically distributed, admitting f as density. The problem of

the statistician is then to use this sample to build an estimator that is as close as possible to the density f. Several estimators of probability density have been proposed since Rosenblatt, Cencov and Parzen.
The vast majority of them fit into a very large class of contributors by estimators built from a nucleus (the nouyau method and the estimation by histogram).

1.8.1 Estimation of the probability density by the kernel method

This is the most popular estimator. It is adapted to continuous random variables. Let $X_1, X_2, \ldots, X_n$ a sample of independent random variables and identically distributed with F distribution function and a density f.

The kernel density estimator, denoted $\hat{f}(x)$, is defined by

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} k(\frac{x - X_i}{h_n}) \tag{1.1}$$

Where k is called the weight or kernel function, and h_n is called the smoothing parameter or window.

Usual kernels:

Kernels	$K(u)$
Uniform	$\frac{1}{2}, \vert u\vert \leq 1$
Epanechncov	$\frac{3}{4}(1-u^2), \vert u\vert \in \mathbb{R}$
Triangular	$(1-\vert u\vert), \vert u\vert \in \mathbb{R}$
Normal	$\frac{1}{\sqrt{2\pi}}\exp(\frac{-u^2}{2}), u \in \mathbb{R}$

Properties of a kernel estimator:

1. $\int_{\mathbb{R}} k(u)\mathrm{du} = 1$ and $k(u) \geq 0$ (the kernel estimator is a density function).
2. f has the same properties of continuity and differentiability as k:
   - If k is continuous, f will be a continuous function,
   - If k is differentiable, f will be a different function,
   - If k can take negative values, then f can also take negative values.

1.8.2 Asymptotic Expressions of Bias and Variance

An asymptotic approximation of the esperance of the estimator $f(x)$ is given under following conditions on f, h_n and k.

1. The seconde derivative $f''(x)$ is continuous, square integral and monotone on $]-\infty, -M]$ and $[M, +\infty[$ for $M > 0$,

2. $\lim_{n\to+\infty} h_n = 0$ and $\lim_{n\to+\infty} nh_n = 0$,

3. For $\hat{f}(x)$ to be a density, we suppose that $k(u) \geq 0$ and $\int k(u)\mathrm{du} = 1$, The kernel function is supposed to be symmetric around zero ($\int k(u)\mathrm{du} = 0$) and possesses a finite second order moment ($\int u^2 k(u)\mathrm{du} < \infty$).

We have established that

$$\mathrm{Biais}(\hat{\mathrm{f}}(\mathrm{x})) = \frac{h_n}{2} f''(x)\mu_2 + O(h_n^2),$$

and

$$\mathrm{Var}(\hat{\mathrm{f}}_\mathrm{n}(\mathrm{x})) = \frac{f(x)}{n} R(k) + O(\frac{1}{nh_n}),$$

where $\mu_2 = \int u^2 k(u)\mathrm{du}$ and $R(g) = \int g^2(u)\mathrm{du}$ for g a function of integral square.
We must therefore try to choose an h_n that makes a compromise between the bias_2 and variance.
Asymptotic expressions of bias and variance allow us to find asymptotic expressions for the MSE and le MISE.
These expressions were obtained under condition (3) on k and assuming that the density of probability f had all the necessary (continuous) derivations.
The following asymptotic approximations can easily be obtained for the MSE and the MISE.

$$\mathrm{MSE}(\hat{f}_n(x)) = \frac{h^4}{4}(f''(x))^2\mu^2 + \frac{f(x)}{nh_n}R(k) + O(\frac{1}{nh_n}) + O(h_n^4),$$

and

$$\mathrm{MISE}(\hat{f}_n(x)) = \frac{h_n^4}{4}\mu^2 \int_{\mathbb{R}} (f''(x))^2\mathrm{dx} + \frac{1}{nh_n}R(k) + O(h^4 + \frac{1}{nh_n}),$$

Under appropriate conditions of integrity of f and its derivative.
On note l'approximation asymptotique de le MSE par

$$\mathrm{AMSE}(\hat{f}_n(x)) = \frac{h_n^4}{4}(f''(x))^2\mu_2^2 + \frac{f(x)}{nh_n}R(k),$$

and the asymptotic approximation of the MISE by

$$\mathrm{AMISE}(\hat{f}_n(x)) = \frac{h_n^4}{4}\mu_2^2 \int_{\mathrm{R}} (f''(x))^2\mathrm{dx} + \frac{R(k)}{nh_n}.$$

1.8.3 Optimal theoretical choice of the smoothing parameter

For the smoothing parameter, we distinguish between the constant smoothing parameter h_n, (or global), and $h_n(x)$ variable (local) smoothing parameter.

These different choices of the smoothing parameter result in the following kernel estimators:

$$\hat{f}_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n} k(\frac{x-X_i}{h_n}),$$

and

$$\hat{f}_n(x) = \frac{1}{nh_n(x)}\sum_{i=1}^{n} k(\frac{x-X_i}{h_n(x)}).$$

We will describe the optimal theoretical choices of the smoothing parameters h_n and $h_n(x)$.
An appropriate criterion for selecting a constant smoothing parameter h_n and the MISE.
The optimal smoothing parameter is the value of h_n that minimizes the MISE. Note this value h_{MISE}.
An asymptotic approximation of h_{MISE} is given by h_{AMISE}, the value of h which minimizes AMISE$\hat{f}_n(x)$.
It is easy to verify that

$$h_{AMISE} = \{\frac{R(k)}{\mu^2 R(f'')}\}^{\frac{1}{5}} n^{\frac{-1}{5}},$$

and

$$h_{MISE} \approx \{\frac{R(k)}{\mu^2 R(f'')}\}^{\frac{1}{5}} n^{\frac{-1}{5}}$$

, that is to say

$$\lim_{n\to\infty}\frac{h_{\mathrm{MISE}}}{h_{\mathrm{AMISE}}} = 1.$$

An appropriate criterion for selecting a variable (local) smoothing parameter $h_n(x)$ is the local performance measure MSE $f_{n,L(x)}$. . We introduce the following notations:

$$h_{\mathrm{MISE}} = \mathrm{argmin}_{h_n} MSE(\hat{f}_{n,L(x)}),$$

and

$$h_{\mathrm{AMISE}} = \mathrm{argmin}_{h_n} AMSE(\hat{f}_{n,L(x)}),$$

Under condition that $f^{''}(x) \neq 0$. The h_{AMISE} and $h_{\mathrm{AMSE}}(x)$ choices are theoretical choices, that are not usable in practice because it depends on unknown quantities f and $f^{''}$. By substituting h_{AMISE} for the expression of AMISE, we show that for the kernel estimator

$$n^{\frac{4}{5}}\mathrm{AMISE}\hat{f}_{h_{\mathrm{AMISE}}} = O(1).$$

1.9 Recursive kernel estimator

A kernel recursive estimator of the probability density can be preferred f. A recursive kernel estimator is given by the following recursive relationship:

$$\hat{f}_n(x) = \frac{n-1}{n}\hat{f}_{n-1}(x) + \frac{1}{nh_n}k(\frac{x-X_i}{h_n}),$$

where $\frac{1}{n}$ and h_n are two positive series of positive ones that tend to 0. Here are two examples for recursive estimators.
Wolverton-Wagner estimator:

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i}k(\frac{x-X_i}{h_i}).$$

This recursive estimator was introduced by Wolverton and Wagner (1969) and studied by Yamato (1971).
Deheuvels estimtor (1973):

$$\hat{f}_n(x) = \frac{1}{\sum_{i=1}^{n}h_i}\sum_{i=1}^{n}k(\frac{x-X_i}{h_i}).$$

1.9.1 Some convergence results for the Wolverton-Wagner estimator

Yamato (1971) demonstrated in particular that for any pair $(K(y),(h_n))$ belonging to $\mathbb{C}_2$ subset of $\mathbb{C}_1$ for which the sequences (h_n) are decreasing, $\hat{f}_n(x)$ converges in quadratic mean to $f(x)$. If more $(K(y),(h_n))$ belonging to $\mathbb{C}_2^{'}$ subset of $\mathbb{C}_2$ for which the suites (h_n) verify

$$\frac{h_n}{n}\sum_{i=1}^{n}\frac{1}{h_i} \to \alpha,$$

with $\alpha \in]0,1[$, so he proved that

$$\lim_{n\to\infty} nh_n \mathrm{Var}(\hat{f}_n(x)) = \alpha\tau^2 f(x).$$

The following theorem gives us almost sure convergence and the asymptotic normality of $\hat{f}_n(x)$.

Theorem 1.3 *If f is out of bounded and if the window is such that $h_n = n^{-\alpha}$ with $0 < \alpha < 1$, we have for all $x \in \mathbb{R}$,*

$$\lim_{n\to\infty}\hat{f}_n(x) = f(x), a.s$$

Furthermore, $\frac{1}{5} < \alpha < 1$, we have

$$\sqrt{nh_n}(\hat{f}_n(x) - f(x) \to N\left(0, \frac{\tau^2 f(x)}{1+\alpha}\right).$$

The central limit theorem was obtained by Duflo (1997) with $\frac{1}{3} < \alpha < 1$, and by Bercu and Chafai (2007) with $\frac{1}{5} < \alpha < 1$.

1.9.2 Some convergence results for the Deheuvels estimator

Among many other results, Deheuvels has shown that $\hat{f}_n(x)$ converges to quadratic average to $f(x)$ and in particular Deheuvels (1973):

$$\mathrm{Var}(\hat{f}_n(x)) \sim \frac{1}{\displaystyle\sum_{i=1}^{n} h_i} \tau^2 f(x).$$

The following theorem gives us almost sure convergence.

Theorem 1.4 *If f is out of bounded and if the window is such that $h_n = n^{-\alpha}$ with $0 < \alpha < 1$, then we have for all $x \in \mathbb{R}$,*

$$\lim_{n\to\infty} \hat{f}_n(x) = f(x), as.$$

Furthermore, if $\frac{1}{5} < \alpha < 1$, we have that

$$\sqrt{nh_n}(\hat{f}_n(x) - f(x) \to N(0, \tau^2 f(x)).$$

Using the martingales as in the previous theorem, we can obtain even a central limit theorem for $\hat{f}(x)$.

Choice of the kernel and the window

These choices can only be made by the use of certain criteria. Without going into all the details, it turns out that the choice the kernel has no major influence if it is chosen in a reasonable class. In however, the choice of the h_n window is crucial. In general, hn is obtained by cross validation techniques.

1.10 Some notions about convergence

Definition 1.14 *A sequence of random variable $\{X_n, n \in \mathbb{N}\}$ converges stochastically (or in probability) to the random variable X ($P\{\lim X_n\} = X$) if for all $\epsilon > 0$*

$$P(|X_n - X| > \epsilon) \to 0 \text{ when } n \to \infty.$$

Definition 1.15 *(Markov inequality) If $f(x) > 0$ is a non-decreasing monotonic function ($x \geq 0$) then for $a > 0$*

$$P(|X| > a) \leq \frac{\mathbb{E}f(|X|)}{f(b)}.$$

Definition 1.16 *(Jensen inequality) If g is a continuous convex function on $]a,b[, -\infty \leq a < b \leq +\infty$ and $a < X < b$, so*

$$g(\mathbb{E}X) \leq \mathbb{E}g(X).$$

we recall that a function g is convex on $]a,b[$ if

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x) + (1-\lambda)g(x_2).$$

Definition 1.17 *(Hölder inequality) For all random variables X and Y we have*

$$\mathbb{E}(|XY|) \leq (\mathbb{E}|X|^p)^{\frac{1}{p}}(\mathbb{E}|Y|^q)^{\frac{1}{q}} \text{ with } \frac{1}{p} + \frac{1}{q} = 1, p > 1.$$

Corollary 1.1

$$(|\mathbb{E}X|)^p \leq \mathbb{E}(|X|^p) \text{ for } p > 1 \text{ and,}$$

$$(\mathbb{E}|X|^p)^{\frac{1}{p}} \leq (\mathbb{E}|X|^q)^{\frac{1}{q}} \text{ for } 0 < p < q.$$

1.11 Some properties on probabilities

let $(\Omega, \mathcal{F}, P)$ be a probability space

- For any event $B \in \mathcal{F}$, we have:

$$P(B) = P(A \cap B) + P(\bar{A} \cap B),$$

  or $\bar{A}$ denotes the negation of A.

- Let X and Y be two random variables and $\epsilon > 0$, so:

$$P(X + Y > \epsilon) \leq P(X > \frac{\epsilon}{2}) + P(Y > \frac{\epsilon}{2}),$$

$$P(Z > sY) \leq P(Z > s\epsilon) + P(X \leq \epsilon), \forall s \in \mathbb{R}.$$

Definition 1.18 *Let $(\Omega, \mathcal{F})$ is a measurable space and $\mu : \mathcal{F} \to \bar{\mathbb{R}}_+$ an application.*

We say that μ is a positive measure on $(\Omega, \mathcal{F})$ if:

1. $\mu(\varnothing)$,
2. $\forall (B_n)_{n\geq 1} \in \mathcal{F}$ *a sequence of disjointed events, we have:*

$$\mu(\bigcup_{n\geq 1} B_n) = \sum_{n\geq 1} \mu(B_n).$$

Definition 1.19 *let μ and λ be two bounded measurements on $(\Omega, \mathcal{F})$. we say that λ is absolutely continuous with respect to μ and we note $\lambda \ll \mu$ if:*

$$\forall C \in \mathcal{F} \text{ such that } \mu(C) = 0 \Rightarrow \lambda(C) = 0.$$

A measure μ is said to be finite or bounded on $(\Omega, \mathcal{F})$ if $\mu(\Omega) < +\infty$.

Theorem 1.5 *(Radon-Nicodym theorem) let μ and λ be two bounded measures on $(\Omega, \mathcal{F})$.*

If $\lambda \ll \mu$ so there exists g a non negative function, μ integrable and $\mathcal{F}$ measurable such that $\lambda = g\mu$.

The function g is called the density of λ with respect to μ, we denote it:

$$g = \frac{\mathrm{d}\lambda}{\mathrm{d}\mu},$$

and also

$$\lambda(L) = \int_L g \mathrm{d}\mu, \quad \forall \mu \in \mathcal{F}.$$

Definition 1.20 *(Almost complete convergence) We say θ_n converges almost completely to θ if the series converges:*

$$\sum_{n=0}^{\infty} P(|\theta_n - \theta| > \epsilon) < +\infty.$$

Definition 1.21 *(Bienayami-Tchebychev inequality) Let X random variable of esperance $\mathbb{E}(X)$ and variance σ^2.*
For all α strictly positive, we have

$$\mathbb{P}(|X - \mathbb{E}(X)| > \alpha) \leq \frac{\sigma^2}{\alpha^2}.$$

Properties on conditional esperance

Let $(\Omega, \mathcal{F}, P)$ a probabilized space, $\mathcal{B}$ a sub-tribe of $\mathcal{F}$ and M and N any two real random variables.

1. If M is $\mathcal{B}$ measurable and integrable then

$$\mathbb{E}(M|\mathcal{B}) = M \text{ almost surely }.$$

2. If M is integrable and N is $\mathcal{B}$ measurable and bounded then

$$\mathbb{E}(MN|\mathcal{B}) = N\mathbb{E}(M|\mathcal{B}) \text{ almost surely }.$$

3. If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two sub-tribes of $\mathcal{F}$ such that $\mathcal{B}_1 \subset \mathbb{B}_2$ we have then for all random variables M integrable,
$$\mathbb{E}(\mathbb{E}(M|\mathcal{B}_2)|\mathcal{B}_1) = \mathbb{E}(M|\mathcal{B}_1).$$
Especially if $\mathcal{B}_1$ is the trivial tribe, $\mathcal{B}_1 = \{\varnothing, \Omega\}$ so
$$\mathbb{E}(\mathbb{E}(M|\mathcal{B}_1)) = \mathbb{E}(M).$$

Chapter 2

Asymptotic distribution of the END random error in first-order autoregressive processes using the recursive kernel estimator

Abstract

The purpose of this paper is to consider the asymptotique distributions of the error density estimators in first order autoregressive models (based on the true error) using a kind of recursive density estimator of the probability density function for a sequence of extended negatively dependent random variables.

2.1 Introduction

In the parametric regression and autoregressive models, several authors have studied the properties of estimators for distributions of the errors. The weak convergence of the empirical processus based on residuals in parametric regression models is discussed in Koul (1970, 1977, 1992, 1996), Loynes (1980), Portnoy (1986), and Mammen (1996), while Boldin (1982), Koul (1991), and Koul and Ossiander (1994) discuss this for parametric autoregressive models. Then uniform consistency of error density in these models is discussed in Koul(1992). Lee and Na (2002) extended the asympototic result (based on the L_2-norm) in Bickel and Rosenblatt (1973) to the error density estimator in the first-order autoregressive models while Horvath and Zitikis (2003) consider asymptotics of the L_p-norms of the estimators.

In this paper, we first obtain asymptotic distribution of the kernel error density estimators for a sequence of extended negatively dependent random variables.

Definition 2.1 *Random variables $X_1, X_2, \ldots, X_n$ are said to be extended negatively dependent (END)* *if there exists a constant $M > 0$ such that for each $n \geq 2$, the following two inequalities hold:*

$$\mathbb{P}\{X_1 \leq x_1, \ldots, X_n \leq x_n\} \leq M \prod_{i=1}^{n} \mathbb{P}\{X_i \leq x_i\} \tag{2.1}$$

and

$$\mathbb{P}\{X_1 > x_1, \ldots, X_n > x_n\} \leq M \prod_{i=1}^{n} \mathbb{P}\{X_i > x_i\} \tag{2.2}$$

for each $n \geq 1$ and all real numbers $x_1, x_2, \ldots, x_n$.

In the case $M = 1$ the notion of END random variables reduces to the well known notion of so-called negatively dependent (ND) random variables which was introduced by Lehmann (1966) (cf. also Joag-Dev and Proschan, 1983). Not looking that the notion of END seems to be a straightforward generalization of the notion of negative dependence, the extended negative dependence structure is substantially more comprehensive. As is mentioned in Liu (2009), the END structure can reflect not only a negative dependence structure but also a positive one (inequalities from the definition of ND random variables hold both in reverse direction), to some extend.

Throughout the paper, let $\varepsilon_1, \ldots, \varepsilon_n$ be a strictly stationary sequence of END random variables with the unknown marginal probability density function $f(x)$. We consider the following recursive kernel estimator of f(x)

$$f_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_i} \mathrm{k} \left(\frac{x - \varepsilon_i}{h_i} \right). \tag{2.3}$$

Where $0 < h_n \downarrow 0$ are bandwidths, k(.) is the kernel density function. The recursive kernel density estimator (2.3) was introduced by Wolverton and Wagner (1969). Note that (2.3) can be computed recursively by

$$f_n(x) = \frac{n-1}{n} f_{n-1}(x) + (nh_n)^{-1} \mathrm{k} \left(\frac{x - \varepsilon_n}{h_n} \right). \tag{2.4}$$

This propery of (2.3) is particulary useful in large sample since $f_n(x)$ can be easily updated with each additional observation. Liang and Baek(2004) discussed the point asymptotic normality for $f_n(x)$ under negatively associated random variables.
Since END random variables are much weaker than independent random variables, and NA random variables, studying the large sample character of the kernel density estimate for END sequence is of interest. In this article, we will discuss the asymptotic distribution of error using the recursive density estimates in $AR(1)$ models.
In the sequel, let $\mathcal{C}^2$ stand for set in where the second order derivative f'' exists and is bounded and continuous. All limits are taken as the sample size n tends to ∞, $M_1, M_2, \ldots$ and k_0, k_1 denote positive constants whose values may changefrom from one place to another, unless specified otherwise.

Suppose that the sequence X_i satisfies the first order autoregressive model

$$X_i = \theta X_{i-1} + \varepsilon_i, \tag{2.5}$$

where the ε_i's are extended negatively dependent with mean 0, variance σ^2, finite fourth moment and unknown density f. We also assume that $|\theta| < 1$ and the sequence X_i is stationary. Then we have the representation

$$X_i = \sum_{j=0}^{\infty} \theta^j \varepsilon_{i-j}. \tag{2.6}$$

Assume that we observe $X_0, X_1, \dots, X_n$. Let $\hat{\theta}$ be an estimator of θ with the following property:

$$\sqrt{n}(\hat{\theta}_n - \theta) = O(1). \tag{2.7}$$

Property (2.6) is natural since the least square estimator satisfies it.
Let

$$\hat{\varepsilon_i} = X_i - \hat{\theta}_n X_{i-1}, \quad 1 \le i \le n. \tag{2.8}$$

Denote the residuals. Based on these residuals, we construct a estimator of the error density $f(x)$ as follow:

$$\hat{f}_n(x) := \frac{1}{n} \sum_{i=1}^{n} K_{h_i}(x - \hat{\varepsilon}_i), \tag{2.9}$$

and $x \in \mathbb{R}$ with $K_{h_i}(x) = \frac{1}{h_i} K(\frac{x}{h_i})$. We shall show that $\hat{f}_n(x)$ is asymptotically normal at fixed point.

ASSUMPTIONS

(A1) $\int_{-\infty}^{+\infty} K(u)\mathrm{d}u = 1, \int_{-\infty}^{+\infty} uK(u)\mathrm{d}u = 0, \int_{-\infty}^{+\infty} u^2 K(u)\mathrm{d}u < \infty,$

(A2) $K(.), K' \in L_1, K'''$ exists on the real line and $K''(.) \in L_1$(bounded),

(A3) $\mathbb{E}X_{i-1}^4 < \infty,$

(A4) $h_n \downarrow 0,$

(A5) $\int_{\mathbb{R}} |u|K(u)\mathrm{du} < \infty,\ \int_{\mathbb{R}} |u|K^2(u)\mathrm{du} < \infty,$

(A6) $\int_{\mathbb{R}} |uK'(u)|\mathrm{du} < \infty$ and $\int_{\mathbb{R}} |u||K'(u)|^2\mathrm{du} < \infty.$

2.2 Some auxiliary results

In this section, we will present some important lemmas which will be used to prove the above main results.

Lemma 2.1 *Let random variables $X_1, \dots, X_n$ be END.*

(i) If $f_1, f_2, \dots, f_n$ are are all non decreasing (or non increasing) functions, then random variables $f_1(X_1), f_2(X_2), \dots, f_n(X_n)$ 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 2.2 *Let $\{X_n, n \geq 1\}$ be a sequence of END random variables, then for each $n \geq 1$ and $\lambda \in \mathbb{R}$ there exists a constant $M > 0$ such that*

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

Lemma 2.3 *Let $\{X_n, n \geq 1\}$ be a sequence of END random variables with $\mathbb{E}X_n = 0$ and $|X_n| \leq d_n$ a.s. For each $n \geq 1$, where $\{d_n, n \geq 1\}$ is a sequence of positive constants. Assume that $t > 0$ such that $t.\max_{1\leq i\leq n} d_i \leq 1$. Then for any $\varepsilon > 0$, there exists a constant $M > 0$ such that*

$$\mathbb{P}\left(|\sum_{i=1}^{n} X_i| > \varepsilon\right) \leq 2M \;\; exp\left\{-t\varepsilon + t^2 \sum_{i=1}^{n} \mathbb{E}X_i^2\right\}.$$

2.3 Asymptotic Distribution of $\hat{f}_n$ at a fixd point

In this section, we consider the asymptotic distribution of $\hat{f}_n(x)$, we have the asymptotic normality result follow

Theorem 2.1 *Suppose that at a fixed $t \in \mathbb{R}$, there existe a constant $0 < r < \infty$ such that f satisfies*

$$|f(x) - f(x-y)| \leq r|y|, \; \textit{for all } y \in \mathbb{R}. \tag{2.10}$$

Assume that h_n satisfies $nh_n \to \infty$. Then, under assumptions (A2), (A4) and(2.7) we have that in distribution

$$\frac{\hat{f}_n(x) - \int K(u) f(x - h_i u) \mathrm{du}}{\sqrt{Var(f_n(x))}} \to N(0,1). \tag{2.11}$$

Now, we start prove the auxiliary lemma before proving the Theorem (2.1).

Lemma 2.4 *Suppose that (2.7), (2.10) and the assumptions (A2), (A4) and (A6) hold. Then we have that*

$$|\hat{f}_n(x) - f_n(x)| = O\left(\frac{(n\mathrm{log}n(\mathrm{loglog}n)^l)^{\frac{1}{2}}}{h_n^2}\right). \tag{2.12}$$

Proof of Lemma(2.4): We have by (2.5) and (2.8)

$$\hat{\varepsilon}_i - \varepsilon_i = -(\hat{\theta}_n - \theta)X_{i-1} \tag{2.13}$$

Using $(A2)$ and (2.13), we can obtain

$$\begin{aligned}
\left|\hat{f}_n(x) - f_n(x)\right| &= \left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i}\left[K\left(\frac{x-\hat{\varepsilon}_i}{h_i}\right) - K\left(\frac{x-\varepsilon_i}{h_i}\right)\right]\right|, \\
&= \left|\frac{1}{n}\sum_{i=1}^{n}\left[\frac{(\hat{\theta}_n-\theta)X_{i-1}}{h_i^2}K'\left(\frac{x-\varepsilon_i}{h_i}\right) + \frac{(\hat{\theta}_n-\theta)^2X_{i-1}^2}{2h_n^3}K''(\zeta_{ix})\right]\right|, \\
&\leq \frac{|\hat{\theta}_n-\theta|}{n}\left|\sum_{i=1}^{n}\frac{X_{i-1}}{h_i^2}\left[K'\left(\frac{x-\varepsilon_i}{h_i}\right) - \mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right) + \mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right)\right]\right| \\
&+ \frac{c_1|\hat{\theta}_n-\theta|^2}{2n}\sum_{i=1}^{n}\frac{X_{i-1}^2}{h_i^3}, \\
&\leq \frac{|\hat{\theta}_n-\theta|}{n}\left|\sum_{i=1}^{n}\frac{X_{i-1}}{h_i^2}\left[K'\left(\frac{x-\varepsilon_i}{h_i}\right) - \mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right)\right]\right| \\
&+ \frac{|\hat{\theta}_n-\theta|}{n}\left|\sum_{i=1}^{n}\frac{X_{i-1}}{h_i^2}\mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right)\right| + \frac{c_1|\hat{\theta}_n-\theta|^2}{2n}\sum_{i=1}^{n}\frac{X_{i-1}^2}{h_i^3}, \\
&= \frac{|\hat{\theta}_n-\theta|}{n}|S_{n1}| + \frac{|\hat{\theta}_n-\theta|}{n}|S_{n2}| + \frac{c_1|\hat{\theta}_n-\theta|^2}{2n}S_{n3}
\end{aligned}$$

where $c_1(0 < c_1 < \infty)$ is an upper bound for $|K''|$, $(|K''(u)| \leq c_1, \forall u \in \mathbb{R})$, and ζ_{ix} is a number between $\frac{x-\hat{\varepsilon}_i}{h_i}$ and $\frac{x-\varepsilon_i}{h_i}$, more ever that $S_{n1} = \frac{1}{n}\sum_{i=1}^{n}Y_i$ where

$$Y_i = \frac{X_{i-1}}{h_i^2}\left[K'\left(\frac{x-\varepsilon_i}{h_i}\right) - \mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right)\right].$$

By lemma 2.1, we see that $Y_1, \ldots, Y_n$ still END random variables with $\mathbb{E}Y_i = 0$ for $i = 1, \ldots, n$.

$$\mathbb{E}Y_i^2 = \mathbb{E}\left(\frac{X_{i-1}}{h_i^2}\left[K'\left(\frac{x-\varepsilon_i}{h_i}\right) - \mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right)\right]\right)^2.$$

Since X_i and $(\varepsilon_{i+1}, \varepsilon_{i+2}, \ldots)$ are independent by (2.6), we have

$$\begin{aligned}
\mathbb{E}Y_i^2 &= \mathbb{E}(X_{i-1}^2)\mathbb{E}\frac{1}{h_i^4}\left[K'\left(\frac{x-\varepsilon_i}{h_i}\right) - \mathbb{E}\left(K'\left(\frac{x-\varepsilon_i}{h_i}\right)\right)\right]^2, \\
&= \mathbb{E}(X_{i-1}^2)\mathbb{E}\frac{1}{h_i^4}K'\left(\frac{x-\varepsilon_i}{h_i}\right)^2.
\end{aligned}$$

Using $\mathbb{E}X_i^2 = \mathbb{E}X_0^2 < \infty$ and $K(x)$ is bounded function, which implies that

$$\mathbb{E}Y_i^2 = \mathbb{E}(X_{i-1}^2)\mathbb{E}\frac{1}{h_i^4}K'\left(\frac{x-\varepsilon_i}{h_i}\right)^2 \leq C_1\frac{1}{h_i^4} < \infty \quad \text{for} \quad i = 1, \ldots, n.$$

By $h_n = n^{-\sigma} \downarrow 0$ with $0 < \sigma < 1$ we can obtain

$$\sum_{i=1}^{n} Var(Y_i) \leq C_1 \sum_{i=1}^{n} \frac{1}{h_i^4} \leq C_1 \frac{n}{h_n^4}.$$

Set $\lambda(n) = \left[\frac{nh_n^4}{(\text{log}n(\text{loglog}n)^l)}\right]^{\frac{1}{2}}$, and taking $t = \frac{\varepsilon h_n^4}{2C_1\lambda(n)}$, them by Lemma 4.2 we get

$$\begin{aligned}
\mathbb{P}\left(\lambda(n)\left|\frac{1}{n}S_{n1}\right| > \varepsilon\right) &= P\left(|S_{n1}| > \frac{n\varepsilon}{\lambda(n)}\right), \\
&\leq 2M_1 \text{exp}\left\{-\frac{\varepsilon nt}{\lambda(n)} + \frac{C_1 nt^2}{h_n^4}\right\}, \\
&= 2M_1 \text{exp}\left\{-\frac{\varepsilon^2 nh_n^4}{2C_1\lambda^2} + \frac{\varepsilon^2 nh_n^4}{4C_1\lambda^2(n)}\right\}, \\
&= 2M_1 \text{exp}\left\{-\frac{\varepsilon^2 nh_n^4}{4C_1\lambda^2(n)}\right\}, \\
&= 2M_1 \text{exp}\left\{-\frac{\varepsilon^2 \text{log}n(\text{loglog}n)^l}{4C_1}\right\}, \\
&\leq 2M_1 n^{-(1+k_0)}.
\end{aligned}$$

Therefore,

$$\sum_{n=1}^{\infty} P\left(|S_{n1}| > \varepsilon \frac{(n\text{log}n(\text{loglog}n)^l)^{\frac{1}{2}}}{h_n^2}\right) \leq 2M_1 \sum_{n=1}^{\infty} n^{-(1+k_0)} < \infty. \tag{2.14}$$

From (2.14), it follows that $S_{n1} = O\left(\frac{(n\text{log}n(\text{loglog}n)^l)^{\frac{1}{2}}}{h_n^2}\right)$ complete convergence.

Similarly, we proof S_{n2} such that $S_{n2} = \sum_{i=1}^{n} Z_i$ with

$$Z_i = \frac{X_{i-1}}{h_i^2} \mathbb{E}K'\left(\frac{x - \varepsilon_i}{h_i}\right)$$

After the same calculations we get

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|S_{n2}| > \varepsilon \frac{(n\text{log}n(\text{loglog}n)^l)^{\frac{1}{2}}}{h_n^2}\right) \leq 2M_2 \sum_{n=1}^{\infty} n^{-(1+k_1)} < \infty.$$

So $S_{n2} = O\left(\frac{(n\text{log}n(\text{loglog}n)^l)^{\frac{1}{2}}}{h_n^2}\right)$ complete convergence

It remains to be seen that S_{n3} converges almost complete We have

$$S_{n3} = \sum_{i=1}^{n} \frac{X_{i-1}^2}{h_i^3}$$

We know that

$$\left\{ \left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - EX_{i-1}^2 + \mathbb{E}X_{i-1}^2 \right] \right| > \frac{n\varepsilon}{\lambda'(n)} \right\} \subset$$
$$\left\{ \left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 \right] \right| + \left| \sum_{i=1}^{n} \frac{1}{h_i^3} EX_{i-1}^2 \right| > \frac{n\varepsilon}{\lambda'(n)} \right\}$$

By using the triangular inequality we obtains that

$$\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 + \mathbb{E}X_{i-1}^2 \right] \right| > \frac{n\varepsilon}{\lambda'(n)} \right) \leq$$
$$\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 \right] \right| + \left| \sum_{i=1}^{n} \frac{1}{h_i^3} \mathbb{E}X_{i-1}^2 \right| > \frac{n\varepsilon}{\lambda'(n)} \right)$$

Then

$$\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 \right] \right| + \left| \sum_{i=1}^{n} \frac{1}{h_i^3} \mathbb{E}X_{i-1}^2 \right| > \frac{n\varepsilon}{\lambda'(n)} \right) \leq$$
$$\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 \right] \right| > \frac{n\varepsilon}{\lambda'(n)} - \left| \sum_{i=1}^{n} \frac{1}{h_i^3} \mathbb{E}X_{i-1}^2 \right| \right)$$

Set $\mathbb{E}X_{i-1}^2 = a < \infty$ and using $h_n \downarrow 0$ we obtain that

$$\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 \right] \right| + \left| \sum_{i=1}^{n} \frac{1}{h_i^3} \mathbb{E}X_{i-1}^2 \right| > \frac{n\varepsilon}{\lambda'(n)} \right) \leq$$
$$\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - EX_{i-1}^2 \right] \right| > \frac{n\varepsilon}{\lambda'(n)} - \frac{na}{h_n^3} \right)$$

Set$\lambda'(n) = \left[\frac{nh_n^3}{(\log n(\log\log n)^l)} \right]^{\frac{1}{2}}$, and taking $t' = \frac{\varepsilon h_n^3}{2C_1\lambda(n)}$, by Lemma 4.2 we get

$$\begin{aligned}
\mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - \mathbb{E}X_{i-1}^2 \right] \right| > \frac{n\varepsilon}{\lambda'(n)} - \frac{na}{h_n^3} \right) &\leq 2M_2 \exp\left\{ -t' \left(\frac{n\varepsilon}{\lambda'(n)} - \frac{na}{h_n^3} \right) + t'^2 \frac{nC}{h_n^3} \right\}, \\
&\leq 2M_2 \exp\left\{ -\frac{n\varepsilon^2 h_n^3}{2C_1\lambda'^2(n)} + \frac{na\varepsilon}{2\lambda'(n)} + \frac{n\varepsilon^2 h_n^3}{4C_1\lambda'^2(n)} \right\}, \\
&\leq 2M \exp\left\{ -\frac{\varepsilon^2 \log n(\log\log n)^l}{4C_1} + \frac{na\varepsilon}{2\lambda'(n)} \right\}, \\
&\leq 2Mn^{-(1+k_2)}.
\end{aligned}$$

Therefore

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left| \sum_{i=1}^{n} \frac{1}{h_i^3} \left[X_{i-1}^2 - EX_{i-1}^2 \right] \right| > \frac{n\varepsilon}{\lambda'(n)} - \frac{na}{h_n^3} \right) \leq 2M_1 \sum_{n=1}^{\infty} n^{-(1+k_2)} < \infty.$$

So $S_{n3} = O\left(\frac{(n\log n(\log\log n)^l)^{\frac{1}{2}}}{h_n^3} - \frac{na}{h_n^3}\right)$ complete convergence

From the above results it can be concluded that

$$|\hat{f}_n(x) - f_n(x)| = O\left(\frac{(n\log n(\log\log n)^l)^{\frac{1}{2}}}{h_n^2}\right).$$

Proof of Theorem (2.1):
On a

$$\mathbb{E}(f_n(x)) = \int_{\mathbb{R}} K(u)f(x - h_i u)\mathrm{du},$$

we write

$$\hat{f}_n(x) - \int_{\mathbb{R}} K(u)f(x - h_i u)\mathrm{du} = [\hat{f}_n(x) - f_n(x)] + [f_n(x) - \mathbb{E}(f_n(x))]. \tag{2.15}$$

By (2.10),we find that

$$|\int_{\mathbb{R}} K(u)f(x - h_i u)\mathrm{du} - f(x)\int_{\mathbb{R}} K(u)\mathrm{du}| \leq rh_n \int_{\mathbb{R}} |u|K(u)\mathrm{du},$$

the same bound when we replaced K by K^2. Morover, as combinig these bounds with assumptions $(A5)$, we have

$$\int_{\mathbb{R}} K(u)f(x - h_i u)\mathrm{du} = f(x)\int_{\mathbb{R}} K(u)\mathrm{du} + O(h_n),$$

The same expression when we replaced K by K^2. Therefore, we have

$$\begin{aligned} Var(f_n(x)) &= \frac{1}{nh_n^2} Var\left(K\left(\frac{x - \epsilon_1}{h_i}\right)\right), \qquad (2.16)\\ &= \frac{1}{nh_n}\left[\int_{\mathbb{R}} K^2(x)f(x - h_i u)\mathrm{du} - h_i\left(\int_{\mathbb{R}} K(u)f(x - h_n u)\mathrm{du}\right)^2\right], \\ &= \frac{1}{nh_n}\left[f(x)\int_{\mathbb{R}} K^2(u)\mathrm{du} + O(h_n)\right]. \end{aligned}$$

Therefore, we have

$$\frac{f_n(x) - \mathbb{E}(f_n(x))}{\sqrt{Var(f_n(x))}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \frac{K(\frac{x-\epsilon_i}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))}{\sqrt{Var(K(\frac{x-\epsilon_1}{h_i}))}}.$$

We shall useLindeberg Feller's Central Limit Theorem to prove the asymptotic normality, we must verify the Lindeberg Feller's condition as follows.

Fixing any $\epsilon > 0$. Using the boundness of K, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{\left(K(\frac{x-\epsilon_i}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))\right)^2}{Var(K(\frac{x-\epsilon_1}{h_i}))} I\left\{\left|\frac{K(\frac{x-\epsilon_i}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))}{\sqrt{Var(K(\frac{x-\epsilon_1}{h_i}))}}\right| > \epsilon\sqrt{n}\right\}\right] &= \\
O\left(\frac{1}{Var(K(\frac{x-\epsilon_1}{h_i}))}\mathbb{E}\left[I\left\{\left|\frac{K(\frac{x-\epsilon_1}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))}{\sqrt{Var(K(\frac{x-\epsilon_1}{h_i}))}}\right| > \epsilon\sqrt{n}\right\}\right]\right) &= \\
O\left(\frac{\mathbb{P}\left(\left|K(\frac{x-\epsilon_1}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))\right| > \epsilon\sqrt{nVar(K(\frac{x-\epsilon_1}{h_i}))}\right)}{Var(K(\frac{x-\epsilon_1}{h_i}))}\right)&,
\end{aligned}$$

we find by Chebyshev's inequality

$$\mathbb{P}\left(\left\{\left|K(\frac{x-\epsilon_1}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))\right| > \epsilon\sqrt{nVar(K(\frac{x-\epsilon_1}{h_i}))}\right\}\right) \leq \frac{1}{e^2 n}.$$

So, we obtain that

$$\begin{aligned}
\mathbb{E}\left[\frac{\left(K(\frac{x-\epsilon_i}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))\right)^2}{Var(K(\frac{x-\epsilon_1}{h_i}))} I\left\{\left|\frac{K(\frac{x-\epsilon_i}{h_i}) - \mathbb{E}(K(\frac{x-\epsilon_1}{h_i}))}{\sqrt{Var(K(\frac{x-\epsilon_1}{h_i}))}}\right| > \epsilon\sqrt{n}\right\}\right] &= O\left(\frac{1}{\epsilon^2 nVar(K(\frac{x-\epsilon_1}{h_n}))}\right), \\
&= O\left(\frac{1}{\epsilon^2 nh_n[f(x)\int_{\mathbb{R}} K^2(u)\mathrm{du} + O(h_n)}\right. \\
&\to 0,
\end{aligned}$$

when we use (2.16), $f(x) > 0$ and $nh_n \to \infty$, $h_n \to 0$. By the Lindeberg Feller's Central Limit Theorem (e.g, Theorem 4.12 in Kallenberg (1997)), we obtain that in distribution

$$\frac{f_n(x) - \mathbb{E}(f_n(x))}{\sqrt{Var(f_n(x))}} \to N(0,1). \tag{2.17}$$

Combining this result with (2.15), in order to show (2.11) we see that it suffices to prove

$$\frac{|\hat{f}_n(x) - f_n(x)|}{\sqrt{Var(f_n(x))}} = O(1).$$

By (2.16), $h_n \to 0$ and $f(x) > 0$, it follows that $1/\sqrt{Var(f_n(x))} = O(\sqrt{nh_n})$. So, in order to show (2.17), it suffices to prove that

$$\sqrt{nh_n}|\hat{f}_n(x) - f_n(x)| = O(1),$$

this follows from (2.12) that

$$\sqrt{nh_n}|\hat{f}_n(x) - f_n(x)| = O\left(\frac{(\log n(\log\log n)^l)^{\frac{1}{2}}}{h_n^{\frac{5}{2}}}\right).$$

Bibliography

[1] Fuxia Cheng. 2005. Asymptotic distributions of error density estimators in first-order autoregressive models, The Indian Journal of Statistics. **67**: 553-567.

[2] H. Y. Liang, J. Baek. 2004. Asymptotic normality of recursive density density estimates under some dependence assumptions, Metrika. **60**: 155-166.

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

[4] Li, Y. F, Yang, S. C. 2005. Uniformly asymptotic normality of the smooth estimation of the distribution function under associated samples, Acta Math Appl Sinica. **28**: 639-651.

[5] Li, Y. M, Yang, S. C. 2003. Asymptotic normality of the recursive kernel estimate of a probability density function under negatively associated, Chin J Appl Probab Statist. **19**: 383-393.

[6] Liang, H. Y, Jing, B. Y. 2005. Asymptotic properties for estimate of nonparametric regression models based on negatively associated sequences, J Multivariate Anal. **95**: 227-245.

[7] Loynes, R.M. 1980. The empirical d.f. of residuals from generalized regression, Ann Statist. **8**: 285-298.

[8] Portnoy, S. 1986. Asymptotic behavior of the empirical distribution of M-estimated residuals from a regression model with many parameters, Ann Statist. **14**: 1152-1170.

[9] Roussas, G. G. 1994. Asymptotic normality of random fields of positively and negatively associated processes, J Multivariate Anal. **50**: 152-173.

[10] Roussas, G. G. 1995. Asymptotic normality of a smooth estimate of a random field distribution under association, Statist Probab Lett. **24**: 77-90.

[11] S. C. Yang. 2003. Consistency of nearest neighbor estimator of density function for negative associated samples, Acta Math Appl Sinica. **26**: 385-395.

[12] X. J.Wang, L. L. Zheng, C. Xu, S. H. Hu. 2015. Complete consistency for the estimator of non parametric regression models based on extended negatively dependent errors. Statistics: A journal of Theoretical and Applied Statistics. **49**: 396-407.

[13] Y. M. Li, S. C. Yang. 2005. Strong convergence rate of recursive probability density estimator for NA sequences, Chin J Engineering Mathematics. **22**: 659-665.

[14] Yang, S. C. 2005. Uniform asymptotic normality of the regression weighted estimator for positively associated samples, Chin J Appl Probab Statist. **21**: 150-160.

[15] Yongfeng Wu And Mei Guan. 2012. Convergence propertes of the sums for sequences of END random variables, Korean mathematical society. **49**: 1097-1110.

[16] Yongming Li. 2017. On the rate of strong convergence for a recursive probability density estimator of END samples and its applications, J Math Inequal. **11**: 335-343.

Chapter 3

Complete convergence for recursive probability density estimator of LNQD orthant dependent variables

Abstract

Recursive estimation of the probability density function $f(x)$ for stationary processes X_t is considered. Quadratic-mean convergence and asymptotic normality for density estimators $f_n(x)$ are established for strong mixing and for asymptotically uncorrelated processes X_t. Recent results for nonrecursive density estimators are extended to the recursive case.

3.1 Introduction

Lehmann (1966) introduced a definition of negative dependence: two random variables X and Y are said to be negatively quadrant dependent (NQD, for short) if, for all real x and y, $\mathbb{P}(X > x, Y > y) \leq \mathbb{P}(X > x)\mathbb{P}(Y > y)$. Note that two random variables X and Y are NQD if and only if $\mathrm{Cov}(f(X), g(Y)) \leq 0$ for all real valued nondecreasing functions f and g (such that $f(X)$ and $g(Y)$ have finite variances).

A sequence $\{X_i, 1 \leq X_i \leq n\}$ of random variables is said to be linearly negative quadrant dependent (LNQD, for short) if, for any disjoint finite subsets A, B of $\{1, 2, \dots, n\}$ and any positive real numbers rjs, $\sum_{i \in A} r_i X_i$ and $\sum_{j \in B} r_j X_j$ are NQD. The concept of LNQD is due to Newman [2]. It is obvious that LNQD implies NQD. A sequence $\{X_i, 1 \leq X_i \leq n\}$ of random variables is said to be negatively lower orthant dependent (NLOD) if

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) \leq \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i) \tag{3.1}$$

and it is said to be negatively upper orthant dependent (NUOD) if

$$\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) \leq \prod_{i=1}^{n} \mathbb{P}(X_i > x_i) \tag{3.2}$$

A sequence $\{X_i, 1 \leq X_i \leq n\}$ of random variables is said to be negatively orthant dependent (NOD) if it is both NUOD and NLOD.

Definition 3.1 *Two random variables X and Y are said to be negatively orthant quadrant dependent (NQD) if for all real numbers x, y we have that,*

$$\mathbb{P}(X \leq x, Y \leq y) \leq \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) \tag{3.3}$$

or

$$\mathbb{P}(X > x, Y > y) \leq \mathbb{P}(X > x)\mathbb{P}(Y > y) \tag{3.4}$$

The estimation of a probability density function is a fundamental problem. Throughout this paper, let $X_1, \ldots, X_n$ be a sequence of random variables with the unknown probability density function $f(x)$, and distribution function $F(x)$. The hazard rate $r(x) = \frac{f(x)}{1-F(x)}$. Parzen and Rosenblatt introducing the classical kernel estimator of $f(x)$

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K(\frac{x - X_i}{h_n}).$$

Wolverton and Wagner introduced the following recursive kernel estimator of $f(x)$:

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_i} K(\frac{x - X_i}{h_i}), \tag{3.5}$$

where $0 < h_n \downarrow 0$ are bandwidths and K is some kernel function. Note that (3.5) can be computed recursively by

$$f_n(x) = \frac{n-1}{n} f_{n-1}(x) + (nh_n)^{-1} K(\frac{x - h_n}{h_n}). \tag{3.6}$$

This property of (3.6) useful in large sample size since $f_n(x)$ can be easily updated with each additional observation. Liang and Baek discussed the point asymptotic normality for $f_n(x)$ under negatively associated random variables. Li and Yang studied the strong covergence rate of recursive probability density estimator based NA random variables. Li et al discuss the asymptotic bias, quadratic-mean convergence and establish the pointwise asymptotic normality of $f_n(x)$ for a stationary sequence of negatively associated sequences. furthermore, the estimator can be applied in estimating the hazard rate function, which is defined as $r(x) = f(x)/(1 - F(x))$, where $f(x)$ is the unknown marginal probability density function and $F(x)$ is the distribution function. The general hazard rate estimator of $r(x)$ is

$$\hat{r}_n(x) = \frac{\hat{f}_n(x)}{1 - F_n(x)}, \tag{3.7}$$

where $F_n(x)$ is the empirical distribution of $X_1, \dots, X_n$.

Therefore, the properties of $\hat{f}_n(x)$ are extensively discussed by some authors. For example, Liang and Baek discussed the point asymptotic normality for $\hat{f}_n(x)$ under NA random variables. Masry obtained the quadratic mean convergence and asymptotic normality of the recursive estimator under various assumptions on the dependence of X_i; Li et al. discuss the asymptotic bias, quadratic-mean convergence and establish the pointwise asymtotic normality of $\hat{f}_n(x)$ for a stationary sequence of NA sequences.Li and Yang studied the strong convergence rate of recursive probability density estimator $\hat{f}_n(x)$ based on NA random variables. Li extend the results of Li and Yang from NA random variables to END random variables.

In this paper, we will consider the complete convergence rate probability density estimator of (3.5) under strictly stationary LNQD random variables.

3.2 Main results

In this section, we will present the complete convergence for the recursive kernel estimator $\hat{f}_n(x)$. We adopte the following assumptions which were also used in Li and Yang and Li.

(A_1) $\int_{\mathbb{R}} K(u)\mathrm{du} = 1, \int_{\mathbb{R}} uK(u)\mathrm{du} = 0, \int_{\mathbb{R}} u^2K(u)\mathrm{du} < \infty, K(.) \in L_1$;

(A_2) The sequence of bandwiths h_n satisfies the requirement $0 < h_n \downarrow 0$ and $nh_n \to \infty$ as $n \to \infty$.

Theorem 3.1 *Let $\{X_n; n \geq 1\}$ be LNQD sequence. and let assumptions (A_1) and (A_2) hold. Suppose tat the kernel $K(.)$ is a bounded monotone density function and the bandwith $h_n = O(\frac{\log n}{n\beta})$. Then for any $x \in C^2(f)$,*

$$|\hat{f}_n(x) - f(x)| = O\left([\log n/\beta n h_n]^{1/2}\right), completely. \tag{3.8}$$

Theorem 3.2 *Let $\{X_n; n \geq 1\}$ be LNQD sequence, and let assumptions (A_1) and (A_2) hold. Suppose that the kernel $K(.)$ is a bounded monotone density function and the bandwith $h_n = O(\frac{\log n}{n\beta}) \to 0$. Then for any $x \in C^2(f)$,*

$$\hat{f}_n(x) - f(x) \to 0, completely. \tag{3.9}$$

Theorem 3.3 *Let $\{X_n; n \geq 1\}$ be LNQD sequence and $[\log n/n\beta h_n]^{\frac{1}{2}} \to 0$. Suppose that the assumptions (A_1) and (A_2) hold with the kernel $K(.)$ is a bounded monotone density function and the bandwith satisfies that $h_n = O([\log n/n\beta])$. If there exists a point x_0 such that $F(x_0) < 1$, then for any $x \leq x_0$ and $x \in C^2(f)$,*

$$|\hat{r}_n(x) - r(x)| = O\left([\log n/\beta n h_n]^{1/2}\right), completely. \tag{3.10}$$

Theorem 3.4 *Let $\{X_n; n \geq 1\}$ be LNQD sequence. Suppose that the kernel $K(.)$ is a bounded monotone density function and $[\log n/\beta n h_n]^{1/2} \to 0$. IF there exists a point x_0 such that $F(x_0) < 1$, then for any $x \leq x_0$ and $x \in C^2(f)$,*

$$\hat{r}_n(x) - r(x) \to 0, completely. \tag{3.11}$$

3.3 Some lemmas

In this section, we will present some important lemmas which will be used to prove the above main results.

Lemma 3.1 *Let X and Y two random variables be LNQD, then*

(i) if f and g are both nondecreasing or both nonincreasing functions, then $f(X)$ and $g(Y)$ are LNQD.

(ii) if X and Y are nonnegative random variables, then $\mathbb{E}(XY) \leq \mathbb{E}X\mathbb{E}Y$,

(iii) especially, for any real number h, $\mathbb{E}(e^{h(X+Y)}) \leq \mathbb{E}(e^{hX}\mathbb{E}(e^{hY}))$.

Lemma 3.2 *Let $\{X_n, n \geq 1\}$ be a sequence of LNQD random variables and $t > 0$. Then for each $n \geq 1$, we have*

$$\mathbb{E}(\prod_{i=1}^{n} e^{tX_i}) \leq \prod_{i=1}^{n}(\mathbb{E}e^{tX_i}). \tag{3.12}$$

Lemma 3.3 *Let $\alpha > 0$ constants and $0 < \beta \leq \frac{\alpha^2}{e^\alpha - 1 - \alpha}$. Then for all $0 \leq x \leq \alpha$,*

$$\exp(x) \leq 1 + x + \frac{x^2}{\beta}.$$

Lemma 3.4 *Let $\{X_n, n \geq 1\}$ be a sequence of LNQD random variables with $\mathbb{E}X_n = 0$, suppose that there exists some $t > 0$. Then for any $\epsilon > 0$,*

$$\mathbb{P}(|\sum_{i=1}^{n} X_i| > \epsilon) \leq \exp\{-t\epsilon + \frac{t^2}{\beta}\sum_{i=1}^{n}\mathbb{E}X_i^2\} \tag{3.13}$$

Proof.

For each $i \geq 1$ and $\mathbb{E}X_i = 0$, by the fact that $1 + x \leq e^x$ we have

$$\mathbb{E}(e^{tX_i}) \leq 1 + \frac{t^2}{\beta}\mathbb{E}X_i^2 \leq \exp\{\frac{t^2}{\beta}\mathbb{E}X_i^2\}$$

Then, by lemma (3.2)

$$\mathbb{E}(e^{t\sum_{i=1}^{n} X_i}) = \mathbb{E}(\prod_{i=1}^{n} e^{tX_i}) \leq \prod_{i=1}^{n}\mathbb{E}(e^{tX_i}) \leq \exp\{\frac{t^2}{\beta}\sum_{i=1}^{n}\mathbb{E}X_i^2\}. \tag{3.14}$$

By Markov's inequality, Lemma (3.2) and (3.4) we see that

$$\begin{aligned} \mathbb{P}(\sum_{i=1}^{n} X_i > \epsilon) &\leq e^{-t\epsilon}\mathbb{E}\exp\{t\sum_{i=1}^{n} X_i\} \\ &\leq e^{-t\epsilon}\prod_{i=1}^{n}\mathbb{E}\exp\{tX_i\} \\ &\leq \exp\{-t\epsilon + \frac{t^2}{\beta}\sum_{i=1}^{n}\mathbb{E}X_i^2\}. \end{aligned} \tag{3.15}$$

From $\{-X_n, n \geq 1\}$ still LNQD, by rempalcing X_i by $-X_i$ in (3.15) we obtain the desired result. Hence the proof is complete

Lemma 3.5 *Suppose that (A_1) holds, then for all $x \in C^2(f)$,*

$$\lim_{n\to 0}\int_{\mathbb{R}} k(u)f(x-hu)\mathrm{d}u = f(x)$$

Lemma 3.6 *Suppose that (A_1) holds, then for all $x \in C^2(f)$,*

$$\left(\frac{1}{n}\sum_{i=1}^{n} h_i^2\right)^{-1} |\mathbb{E}\hat{f}_n(x) - f(x)| \leq C < \infty.$$

Lemma 3.7 *Let $\{X_n, n \geq 1\}$ be a sequence of LNQD random variables with known distribution $F(x)$ and probability function $f(x)$. Let $f(x)$ be the empirical distribution function. If $\lambda_n = [\log n/n\beta]^{1/2} \to 0$, then*

$$\sup_x |F_n(x) - F(x)| = O(\lambda_n), \textit{completely}. \tag{3.16}$$

Proof. For $n \geq 1$ and $1 \leq i \leq n-1$. Let $F_{ni}(x) = \frac{i}{n}$. By Lemma 2 in Yang we have

$$\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \leq \max_{1\leq j\leq n-1}|F_n(x_{nj} - F_{x_{nj}})| + \frac{2}{n}. \tag{3.17}$$

Set $n\lambda_n \to \infty$, then for any constant $C_1 > 0$, we have that $\frac{2}{n} < \frac{C_1\lambda_n}{2}$ for all n large enough. Then from (3.17) it follows that

$$\begin{aligned} \mathbb{P}\left(\sup_{-\infty<x<\infty}|F_n(x) - F(x)| > C_1\lambda_n\right) &\leq \mathbb{P}\left(\max_{1\leq j\leq n-1}|F_n(x_{nj}) - F(x_{nj})| > C_1\frac{1}{n}\right) \\ &\leq \sum_{j=1}^{n-1}\mathbb{P}(|F_n(x_{nj}) - F(x_{nj})| > C_1\frac{\lambda_n}{2}). \end{aligned} \tag{3.18}$$

Now. Let $\xi_i = I(X_i < x_{nj}) - \mathbb{E}I(X_i < x_{nj})$. By Lemma (3.2) ,we have that $\{\xi_i, i \geq 1\}$ is still a sequence of LNQD random variables with $\mathbb{E}\xi_i = 0$, $|\xi_i| \leq 2$ and $E\xi_i^2 \leq 1$. By choosing $t = \frac{C_1\lambda_n}{4}\beta$ in Lemma (3.4), for all n large enough we have that

$$\begin{aligned}
\mathbb{P}(|F_n(x_{nj}) - F(x_{nj})| > \frac{C_1\lambda_n}{2}) &= \mathbb{P}\left(|\sum_{i=1}^{n}\xi_i| > \frac{C_1 n\lambda_n}{2}\right) \\
&\leq \exp\{\frac{t^2}{\beta}\sum_{i=1}^{n}E\xi_i^2 - \frac{C_1 n\lambda_n t}{2}\} \\
&\leq \exp\{\frac{t^2}{\beta}n - \frac{C_1 n\lambda_n t}{2}\} \\
&\leq \exp\{-\frac{C_1^2 n\lambda_n^2}{16}\beta\} \\
&\leq \exp\{-\frac{C_1^2 \log n}{16}\} \\
&\leq n^{-\frac{C_1^2}{16}}.
\end{aligned}$$

Taking C_1 sufficiently large such that $C_1 > 4\sqrt{2}$. by (3.6) and (3.7) we have that

$$\begin{aligned}
\sum_{n=1}^{\infty}\mathbb{P}\left(\sup_{-\infty<x<\infty}|F_n(x) - F(x)| > C_1\lambda_n\right) &\leq \sum_{n=1}^{\infty}\sum_{j=1}^{n-1} n^{-\frac{C_1^2}{16}} \\
&\leq \sum_{n=1}^{\infty} n^{1-\frac{C_1^2}{16}} \\
&\leq \infty
\end{aligned}$$

Hence. The proof of lemma is complete.

3.4 Proof of Main Results

Proof of Theorem 3.1. Set $A_i = h_i^{-1}[K(\frac{x-X_i}{h_i}) - \mathbb{E}K(\frac{x-X_i}{h_i})]$ for $1 \leq i \leq n$. Since $K(.)$ is bounded and monotone, then the $A_i, i \geq 1$ is still a sequence of LNQD random variables. Morever, it follows from $0 < h_n \downarrow 0$ that there exists some positive constant M_1 such that $\max_{1\leq i\leq n}|A_i| \leq \frac{M_1}{h_n}$. By Lemma (3.4) we have that

$$\begin{aligned}
\sum_{i=1}^{n}\mathbb{E}A_i^2 &\leq \sum_{i=1}^{n} h_i^{-2}\mathbb{E}K^2(\frac{x-X_i}{h_i}) \\
&= \sum_{i=1}^{n} h_i^{-2}\int_R K^2(\frac{x-u}{h_i})f(u)\mathrm{du} \\
&= \sum_{i=1}^{n} h_i^{-1}\int_R K^2(u)f(x-h_iu)\mathrm{du} \\
&\leq M_2 n h_n^{-1}.
\end{aligned}$$

Set $\lambda_n = [\log n/n]^{1/2}$. Applying Lemma 3.2 with $t = [N_2\lambda_n h_n\beta/2M_2]$, where N_2 is some positive constant which will be specified later. It is easy to check that $t.\frac{M_1}{h_n} \leq 1$ for all n large enough, then we get that

$$\begin{aligned}
\mathbb{P}(|\hat{f}_n(x) - \mathbb{E}\hat{f}_n(x)| > N_2\lambda_n) &= \mathbb{P}(|\sum_{i=1}^{n} A_i| > nN_2\lambda_n) \\
&\leq \exp\{-nN_2\lambda_n t + \frac{t^2}{\beta}\sum_{i=1}^{n}\mathbb{E}A_i^2\} \\
&\leq \exp\{-nN_2\lambda_n t + \frac{t^2 nM_2}{\beta h_n}\} \\
&\leq \exp\{-\frac{nN_2^2\lambda_n^2\beta}{2M_2} + \frac{nN_2^{2h_n\beta}}{2M_2} + \frac{nN_2^2h_n^2M_2\beta^2}{4M_2^2\beta h_n} \\
&\leq \exp\{-\frac{nN_2^2\lambda_n^2 h_n\beta}{4M_2}\} \\
&\leq \exp\{-\frac{N_2^2\log n}{4M_2}\} \\
&\leq n^{-\frac{N_2^2}{4M_2}}
\end{aligned}$$

Taking N_2 large enough such that $N_2 > 2\sqrt{M_2}$, then we have that

$$\sum_{n=1}^{\infty}\mathbb{P}(|\hat{f}_n(x) - \mathbb{E}\hat{f}_n(x)| > N_2\lambda_n) < \infty,$$

that is

$$|\hat{f}_n(x) - \mathbb{E}\hat{f}_n(x)| = O([\log n/(n\beta h_n)]^{1/2}), \text{completely.} \tag{3.19}$$

On the other hand, noting that $h_n = O([\log n/n\beta])$, we have by Lemma (3.6) that

$$\begin{aligned}
\left(\frac{\log n}{n\beta h_n}\right)^{-\frac{1}{2}}|\mathbb{E}\hat{f}_n(x) - f(x)| &\leq C\left(\frac{\log n}{n\beta h_n}\right)^{-\frac{1}{2}}\frac{1}{n}\sum_{i=1}^{n}h_i^2 \\
&\leq C\left(\frac{h_n\beta}{n\log n}\right)^{-\frac{1}{2}}\sum_{i=1}^{n}h_i^2 \\
&\leq C\frac{1}{n\beta^2}\sum_{i=1}^{n}\frac{\log^2 i}{i^2} \\
&\leq C,
\end{aligned}$$

which implies that

$$|\mathbb{E}\hat{f}_n(x) - f(x)| = O\left([\log n/n\beta h_n])^{1/2}\right). \tag{3.20}$$

Note that

$$|\hat{f}_n(x) - f(x)| \leq |\hat{f}_n(x) - \mathbb{E}\hat{f}_n(x)| + |\mathbb{E}\hat{f}_n(x) - f(x)|. \tag{3.21}$$

Therefore, the desired result follows immediately by 3.19 and 3.21. The proof is completed.

Proof of Theorem 3.2. By Theorem 3.1 and $[\log n/n\beta h_n]^{1/2} \to 0$ we have that

$$\hat{f}_n(x) - \mathbb{E}\hat{f}_n(x) \to 0, \text{ completely.} \tag{3.22}$$

Therefore, we only need to show that

$$|\mathbb{E}\hat{f}_n(x) - f(x)| \to 0 \tag{3.23}$$

without using in Theorem 3.1 the condition $h_n = O([\log n/n\beta])$, by Lemma (3.6) and Stolz's Theorem we have

$$\lim_{n\to\infty} |\mathbb{E}\hat{f}_n(x) - f(x)| \le C \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} h_i^2 = C \lim_{n\to\infty} h_n^2 = 0.$$

Hence, the proof of the theorem 3.2 is completed.

Since the proof of Theorem 3.4 is similair, we only present the proof of Theorem 3.3 as follows.

Proof of Theorem 3.3. Set $\hat{F}_n(x) = 1 - F_n(x)$ and $\hat{F}(x) = 1 - F(x)$. It follows from (3.7) that

$$|\hat{r}_n(x) - r(x)| \le \frac{\hat{F}(x)|\hat{f}_n(x) - f(x)| + |F_n(x) - F(x)|f(x)}{\hat{F}_n(x)\hat{F}(x)}. \tag{3.24}$$

From $0 \le F(x) \le F(x_0) < 1$ for all $x \le x_0$, $\sup_x f(x) \le C < \infty$ and 3.8 one has that for $x \le x_0$ and all n large enough,

$$\hat{F}_n(x) \ge \frac{\hat{F}(x)}{2} \ge \frac{\hat{F}(x_0)}{2} > 0. \tag{3.25}$$

Consequently, the desired result folows from . The proof is completed.

Bibliography

[1] D. H. Qiu, T. C. HU, M. O. CABRERA, and A. VOLODIN. 2012. Complete convergence for weighted sums of arrays of banach-space-valued random elements, J Lithanian Math. **52**: 316-325.

[2] H. C. Kim. 2011. Exponential inequalities for ELNQD random variables with applications, J Chungcheong Math Society. **24**: 784-793.

[3] J. I. Baek and S.T. Park. 2010. Convergence of weighted sums for arrays of negatively dependent random variables and its applications, J Theor Probab. **23**: 362-377.

[4] Kheira Berkane, Samir Bennaissa. (2019): Complete convergence for widely orthant dependent random variables and its applications in autoregressives AR(1) models. International Journal of Mathematics and Computation. 30(2): 1-12.

[5] Ko, M.-H, Choi, Y.-K, and Choi, Y.-S. 2007. Exponential probability inequality for linearly negative quadrant dependent random variables. Communications of Korean Math Soc. **22**: 137-143.

[6] Masry, E. 1986. Recursive porobability density estimation for weakly dependent stationary processes, IEEE Trans Inform Theor. **32**: 254-267.

[7] Sung S. H. 2011. On the exponetial inequalities for negatively dependent random variables, J Math Anal and Appl. **381**: 538-545.

[8] S. H. Sung and A. I. Volodin. 2006. On the rate of complete convergence for weighted sums of arrays of random elements, J Korean Math Soc. **43**: 815-828.

[9] T. S. Kim. 2007. On the exponential inequality for negative dependent sequence, Communications of Korean Math Society. **22**: 315-321.

[10] Yongming Li, Chengdong Wei,Shanchao Yang. 2010. The recursive kernel distibution function estimator based on negatively and positively associated sequences, Communications in statistics-Theory and Methods. **39**: 3585-3595.

[11] Yongming Li . 2017. On the rate of strong convergence for a recursive probability density estimator of END samples and its applications, J Mathematical Inequalities. **11**: 335-343.

[12] Y. Q. CHen, A. Y. CHen, W. NG Kai. 2010. The strong law of large numbers for extended negatively dependent random variables, J Appl Prob. **47**: 908-922.

[13] Y. F. Wu, M. Guan. 2012. Convergence properties of the partial sums for sequences of END random variables, J Korean Math Soc. **49**: 1097-1110.

This article is published in the International Journal of Mathematics and Computation

Chapter 4

Complete convergence for widely orthant dependent random variables and its applications in autoregressive $AR(1)$ models

Abstract

Exponential inequalities have be an important tool in probability and statistics. Version of Bernstein type inequalities have proved for independent and for some dependence structure. We prove an new exponential inequality for the distributions of sums of widely orthant dependent(WOD, in short)random variables. The results are applied to first-order autoregressive processes AR(1), if it satisfies the following relation:

$$X_n = \theta X_{n-1} + \zeta_n.$$

With (ζ_n) is a sequence of the widely orthant dependent (WOD, in short) random variables.

4.1 Introduction

The probability limit theorem and its applications for independent random variables have been studied by many authors, while the assumption of independence is not reasonable in real practice. If the independent case is classical in the literature, the treatment of dependent random variables is more recent. One of the important dependence structure is the wide dependence structure, which was introduced by (Wang, Wang and Gao, 2013) as follows.

Definition 4.1 *For the random variables $\{X_n, n \geq 1\}$, if there exists a finite real sequence $\{g_U(n), n \geq 1\}$ satisfying for each n*

$$\mathbb{P}(X_1 > x_1, X_2 > x_2, \ldots, X_n > x_n) \leq g_U(n) \prod_{i=1}^{n} \mathbb{P}(X_i > x_i)$$

then we say that the $\{X_n, n \geq 1\}$ are widely upper orthant dependent (WUOD, in short); if there exists a finite real sequence $\{g_L(n), n \geq 1\}$ satisfying for each $n \geq 1$ and for all $x_i \in]-\infty, +\infty[, 1 \leq i \leq n$,

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \leq g_L(n) \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i)$$

then we say that the $\{X_n, n \geq 1\}$ are widely lower orthant dependent (WLOD, in short); if they are both WUOD and WLOD, then we say that the $\{X_n, n \geq 1\}$ are widely orthant dependent (WOD, in short), and $g_U(n), g_L(n), n \geq 1$, are called dominating coefficients.
An array $\{X_{ni}, i \geq 1, n \geq 1\}$ of random variables is called row-wise WOD if forevery $n \geq 1$, $\{X_{ni}, i \geq 1\}$ is a sequence of WOD random variables.
Recall that when $g_L(n) = g_U(n) = N$ for some constant N, 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 (ENOD, in short). The concept of general extended negative dependence was proposed by (Liu 2009), (Liu 2010) and further promoted by (Chen, Chen and Ng,2010), (Chen, Yuen and Ng, 2011), (Shen, 2011), (Shen, 2011), (Shen, 2013a), (Wang Y, 2011), (Wang Y, 2012), 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.
The concept of WOD random variables was introduced by (Wang et al, 2013) and many applications have been found subsequently. See, for example, (Wang et al., 2013) 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 asymptotics for the finite-time ruin probability of a new dependent risk model with a constant interest rate. (Wang Y, 2011) presented some basic renewal theorems for a random walk with widely dependent increments and gave some applications. (Wang Y, 2012) studied the asymptotics of the finite-time ruin probability for a generalized renewal risk model with independent strong subexponential claim sizes and widely lower orthant dependent inter-occurrence times.)Liu, Gao and Wang, 2012) gave the asymptotically equivalent formula for the finite-time ruin probability under a dependent risk model with constant interest rate. (He and Cheng, 2013) provided the asymptotic lower bounds of precise large deviations with nonnegative and dependent random variables. (Chen, Wang and Wang, 2013) considered uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional renewal risk models with constant interest forces and diffusion generated by Brownian motions. (Shen, 2013b) established the Bernstein type inequality for WOD random variables and gave some applications, and so forth.

4.2 Some lemmas

To prove the main results of the paper, we need the following important lemmas. The first one is a basic property for WOD random variables, which was obtained by(Wang et al., 2013)

Lemma 4.1

(i) Let $\{X_n, n \geq 1\}$ be WLOD (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* $g_L(n), n \geq 1 (g_U(n), n \geq 1)$*;*
- *if* $\{f_n(.), n \geq 1\}$ *are nonincreasing, then* $\{f_n(X_n), n \geq 1\}$ *are WUOD (WLOD) with dominating coefficients* $g_L(n), n \geq 1 (g_U(n), n \geq 1)$*.*

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

$$\mathbb{E}\prod_{i=1}^{n} X_i \leq g_U(n)\prod_{i=1}^{n}\mathbb{E}X_i.$$

In particular, 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* $\lambda > 0$*,*

$$\mathbb{E}\exp\{\lambda\sum_{i=1}^{n} X_i\} \leq g_U(n)\prod_{i=1}^{n}\mathbb{E}\exp\{\lambda X_i\}.$$

By Lemma 4.1*, we can get the following corollary immediately.*

Corollary 4.1 *Let* $\{X_n, n \geq 1\}$ *be a sequence of WOD random variables.*

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

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

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

Proof. *For* $\lambda > 0$*, it is easy to see that* λX_i *and* $\lambda\sum_{j=i+1}^{n} X_j$ *are WOD by the definition. Which implies that* $\exp(\lambda X_i)$ *and* $\exp(\lambda\sum_{j=i+1}^{n} X_j)$ *are also WOD for* $i = 1, 2, \ldots, n-1$*, by Lemma* 4.1

(i). It follows from Lemma 4.1 and Definition 4.1 that

$$\begin{aligned}
\mathbb{E}(\prod_{i=1}^{n} e^{\lambda X_i}) &= \mathbb{E}(\exp(\lambda X_1)\exp(\lambda \sum_{j=2}^{n} X_j)), \\
&\leq g_1(n)\mathbb{E}[\exp(\lambda X_1)]\mathbb{E}[\exp(\lambda \sum_{j=2}^{n} X_j)], \\
&= g_1(n)\mathbb{E}[\exp(\lambda X_1)]\mathbb{E}[\exp(\lambda X_2)\exp(\lambda \sum_{j=3}^{n} X_j)], \\
&\leq g_1(n)g_2(n)\mathbb{E}[\exp(\lambda X_1)]\mathbb{E}[\exp(\lambda X_2)]\mathbb{E}[\exp(\lambda \sum_{j=3}^{n} X_j)], \\
&\leq \prod_{i=1}^{n-1} g_i(n) \prod_{i=1}^{n}(\mathbb{E}e^{\lambda X_i}), \\
&= g(n)\prod_{i=1}^{n}(\mathbb{E}e^{\lambda X_i}),
\end{aligned}$$

where $g(n) = \prod_{i=1}^{n-1} g_i(n)$.

Lemma 4.2 *(see (Boulenoir, 2018; Chebbab, 2018)) Let $\alpha > 0$ constants and $0 < \beta \leq \frac{\alpha^2}{e^{\alpha}-1-\alpha}$. Then for all $0 \leq x \leq \alpha$,*

$$\exp(x) \leq 1 + x + \frac{x^2}{\beta}.$$

Proof. *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 $0 < \beta \geq \frac{\alpha^2}{e^{\alpha}-1-\alpha}$ and $0 \leq x \leq \alpha$. Take the derivative

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

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

Lemma 4.3 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $EX_i = 0$ for each $n \geq 1$, for all $0 \leq X_i \leq \frac{\alpha}{\lambda}$ with take α and λ constants positives . Then for any $\epsilon > 0$*

$$\mathbb{P}(|S_n| \geq n\epsilon) \leq 2g(n)\exp\{-\frac{n^2\epsilon^2}{4B_n}\beta\}, \tag{4.1}$$

where $B_n = \sum_{i=1}^{n} \mathbb{E}X_i^2$

Proof. For $0 \leq X_i \leq \frac{\alpha}{\lambda}$ and $\mathbb{E}X_i = 0$, $i \geq 1$ and the fact that $1 + x \leq e^x$, therefore, by Markov's inequality and (4.1) it follows that

$$\begin{aligned} \mathbb{P}(S_n > n\epsilon) &= \mathbb{P}(\lambda S_n > \lambda n\epsilon), \\ &= \mathbb{P}(e^{\lambda S_n} > e^{\lambda n\epsilon}), \\ &\leq e^{-\lambda n\epsilon}\mathbb{E}(e^{\lambda S_n}), \\ &\leq g(n)e^{-\lambda n\epsilon}\prod_{i=1}^{n}\mathbb{E}[\exp(\lambda X_i)], \\ &\leq g(n)e^{-\lambda n\epsilon}\prod_{i=1}^{n}(1 + \frac{\lambda^2}{\beta}\mathbb{E}X_i^2), \\ &\leq g(n)e^{-\lambda n\epsilon}\prod_{i=1}^{n}(e^{\frac{\lambda^2}{\beta}\mathbb{E}X_i^2}), \\ &\leq g(n)\exp(\frac{\lambda^2}{\beta}B_n - \lambda n\epsilon). \end{aligned} \tag{4.2}$$

By minimizing (respect to λ) the right-hande side of (4.2) we obtain

$$\mathbb{P}(S_n > n\epsilon) \leq g(n)\exp\{-\frac{n^2\epsilon^2}{4B_n}\beta\}. \tag{4.3}$$

We know that

$$\mathbb{P}(|S_n| \geq n\epsilon) = \mathbb{P}(S_n > n\epsilon) + \mathbb{P}(S_n \leq -n\epsilon)$$

Since $\{-X_n, n \geq 1\}$ is also WOD we obtain by (4.3) that

$$\mathbb{P}(S_n \leq -n\epsilon) = \mathbb{P}(-S_n > n\epsilon) \leq g(n)\exp\{-\frac{n^2\epsilon^2}{4B_n}\beta\}. \tag{4.4}$$

By (4.3) and (4.4) the result (4.1) follows.

4.3 Main Results and proofs

Theorem 4.1 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $\mathbb{E}X_i = 0$. If there exists a positive constants α, λ such that $0 \leq X_i \leq \frac{\alpha}{\lambda}, i \geq 1$,then for any $\lambda > 0$, there exists a finite sequence $g(n)$ such that*

$$\mathbb{E}(e^{\lambda S_n}) \leq g(n)\exp(\frac{\lambda^2}{2\beta}B_n). \tag{4.5}$$

Proof. From conditions $\mathbb{E}X_i = 0$ and using the Lemma (4.2), $1 + x \leq e^x$, we have

$$\begin{aligned} \mathbb{E}e^{\lambda X_i} &\leq 1 + \frac{\lambda^2}{2}(\frac{X_i^2}{\beta}), \\ &\leq 1 + \frac{\lambda^2}{2\beta}\mathbb{E}X_i^2, \\ &\leq \exp(\frac{\lambda^2}{2\beta}\mathbb{E}X_i^2), \end{aligned} \tag{4.6}$$

for any $\lambda > 0$. By Corrolary (4.1) and (4.6) we get

$$\mathbb{E}\exp(\lambda \sum_{i=1}^{n} X_i) \leq g(n) \prod_{i=1}^{n} \mathbb{E}(\exp(\lambda X_i)) \leq g(n)\exp(\frac{\lambda^2}{2\beta} \sum_{i=1}^{n} \mathbb{E}X_i^2) = g(n)\exp(\frac{\lambda^2}{2\beta} B_n).$$

Theorem 4.2 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^2 < \infty$ such that $0 \leq X_i \leq \frac{\alpha}{\lambda}$ for each $1 \leq i \leq n$, $n \geq 1$ where α and λ are positive constants and $B_n = \sum_{i=1}^{n} \mathbb{E}X_i^2$, then for any $\epsilon > 0$ and $n \geq 1$, there exists a finite sequence $g(n)$ such that*

$$\mathbb{P}(S_n/B_n \geq \epsilon) \leq g(n)\exp\{-\frac{\epsilon^2 \beta B_n}{4}\}. \tag{4.7}$$

Proof. For $\lambda > 0$ we have $0 \leq X_i \leq \frac{\alpha}{\lambda}$. Then from Markov's inequality it follows that

$$\begin{aligned} \mathbb{P}(S_n/B_n \geq \epsilon) &= \mathbb{P}(e^{\lambda S_n} \geq e^{\lambda \epsilon B_n}), \\ &\leq e^{-\lambda \epsilon B_n} \mathbb{E}(\prod_{i=1}^{n} e^{\lambda X_i}), \\ &\leq g(n)\exp\{\frac{\lambda^2}{\beta} B_n - \lambda \epsilon B_n\}. \end{aligned} \tag{4.8}$$

By taking $\lambda = \frac{\epsilon \beta}{2}$ we obtain (4.7) from (4.8).

Theorem 4.3 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $\mathbb{E}X_i = 0$. If there existe a positive constant α, λ such that $0 \leq X_i \leq \frac{\alpha}{\lambda}$ for each $n \geq 1$, then for any $\epsilon > 0$, there exists a finite sequence $g(n)$ such that*

$$\mathbb{P}(|S_n| \geq \epsilon) \leq 2g(n)\exp(-\frac{\epsilon^2 \beta}{4B_n}), \tag{4.9}$$

where $B_n = \sum_{i=1}^{n} \mathbb{E}X_i^2$.

Proof. From conditions $\mathbb{E}X_i = 0$ and $0 \leq X_i \leq \frac{\alpha}{\lambda}$, $i \geq 1$.

By Markov's inequality and Corrolary (4.1), Lemma (4.2) with the fact that $1 + x \leq e^x$, then

$$\begin{aligned} \mathbb{P}(S_n > \epsilon) &\leq e^{-\lambda \epsilon} \mathbb{E}(e^{\lambda \epsilon}), \\ &\leq g(n) e^{-\lambda \epsilon} \prod_{i=1}^{n} (e^{\frac{\lambda^2}{\beta} \mathbb{E}x_i^2}), \\ &\leq g(n)\exp\{\frac{\lambda^2}{\beta} B_n - \lambda \epsilon\}. \end{aligned} \tag{4.10}$$

Take $\lambda = \frac{\epsilon\beta}{2B_n}$. Hence from (4.10) it follows that

$$\mathbb{P}(S_n > \epsilon) \leq g(n)\exp\{-\frac{\epsilon^2\beta}{4B_n}\}. \tag{4.11}$$

Since $\{-X_n, n \geq 1\}$ is also sequence of WOD random variables, from (4.11) it follows that

$$\mathbb{P}(S_n \leq -\epsilon) = \mathbb{P}(-S_n \geq \epsilon) \leq g(n)\exp(-\frac{\epsilon^2\beta}{4B_n}). \tag{4.12}$$

From (4.11) and (4.12) we obtain

$$\mathbb{P}(|S_n| \geq \epsilon) = \mathbb{P}(S_n \geq \epsilon) + \mathbb{P}(S_n \leq -\epsilon) \leq 2g(n)\exp(-\frac{\epsilon^2\beta}{4B_n}).$$

Theorem 4.4 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables such that $0 \leq X_i \leq \frac{\alpha}{\lambda}, i \leq n$, where α and λ are positives constants. Then for any $\epsilon > 0$ we have*

$$\mathbb{P}(|S_n - \mathbb{E}S_n| \geq \epsilon) \leq 2g(n)\exp(-\frac{\epsilon^2\beta}{4B_n}). \tag{4.13}$$

Proof. Let $B_n = \sum_{i=1}^{n} \mathbb{E}X_i^2$. By Markov's inequality and lemma (4.2) and the fact that $1 + x \leq e^x$

$$\begin{aligned}
\mathbb{P}(S_n - \mathbb{E}S_n \geq \epsilon) &\leq e^{-\lambda\epsilon}\mathbb{E}(e^{\lambda(S_n - \mathbb{E}S_n)}), \\
&\leq e^{-\lambda\epsilon}\mathbb{E}(\prod_{i=1}^{n} e^{\lambda(X_i - \mathbb{E}X_i)}), \\
&\leq g(n)e^{-\lambda\epsilon}\prod_{i=1}^{n}\mathbb{E}(e^{\lambda(X_i - \mathbb{E}X_i)}),
\end{aligned} \tag{4.14}$$

on the other hands we have

$$\begin{aligned}
\mathbb{E}(e^{\lambda(X_i - \mathbb{E}X_i)}) &\leq 1 + \mathbb{E}(X_i - \mathbb{E}X_i) + \frac{\lambda^2\mathbb{E}(X_i - \mathbb{E}X_i)^2}{\beta}, \\
&\leq 1 + \frac{\lambda^2\mathbb{E}X_i^2}{\beta}, \\
&\leq \exp(\frac{\lambda^2 EX_i^2}{\beta})
\end{aligned}$$

We compensate for this result in the right-hand side of (4.14) we get

$$\begin{aligned}
\mathbb{P}(S_n - \mathbb{E}S_n \geq \epsilon) &\leq g(n)e^{-\lambda\epsilon}\prod_{i=1}^{n} e^{\frac{\lambda^2\mathbb{E}X_i^2}{\beta}}, \\
&\leq g(n)\exp(\frac{\lambda^2 B_n}{\beta} - \lambda\epsilon).
\end{aligned} \tag{4.15}$$

By taking $\lambda = \frac{\epsilon\beta}{2B_n}$. Hence from (4.15) it follows that

$$\mathbb{P}(S_n - \mathbb{E}S_n \geq \epsilon) \leq f(n)\exp(-\frac{\epsilon^2\beta}{4B_n}). \tag{4.16}$$

Since $\{-X_n, n \geq 1\}$ is also a sequence of WOD random variables we also have

$$\mathbb{P}(S_n - \mathbb{E}S_n \leq -\epsilon) = \mathbb{P}(-(S_n - \mathbb{E}S_n) > \epsilon) \leq g(n)\exp(-\frac{\epsilon^2\beta}{4B_n}), \tag{4.17}$$

by (4.16) and (4.17) we get (4.13).

Corollary 4.2 *Let $\{X_n, n \geq 1\}$ be a sequence of identically distributed WOD random variables. Assume that there exists a postive integer n_0 such that $0 \leq X_i \leq \frac{\alpha}{\lambda}$, for each $1 \leq i \leq n, n \geq n_0$, where α and λ are positives constants. Then for any $\epsilon > 0$ such that $\epsilon > 0$ we have*

$$\mathbb{P}(|S_n - \mathbb{E}S_n| \geq n\epsilon) \leq 2g(n)\exp(-\frac{n^2\epsilon^2}{4B_n}) \tag{4.18}$$

Theorem 4.5 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $\mathbb{E}X_i = 0$ and $0 \leq X_i \leq \frac{\alpha}{\lambda}$, for each $i \geq 1$, where α and β are positives constants. Then for any $s > 0$*

$$\sum_{n=1}^{\infty} \mathbb{P}(|S_n| > n^s\epsilon) < \infty. \tag{4.19}$$

Proof. Let $B = \sum_{n=1}^{\infty} \mathbb{E}X_n^2 < \infty$. For any $\epsilon > 0$, from Theorem (4.3) it follows that for a finite sequence $g(n)$ we have

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(|S_n| > n^s\epsilon) &\leq 2\sum_{n=1}^{\infty} g(n)\exp(-\frac{n^{2s}\epsilon^2\beta}{4B_n}), \\ &\leq 2C\sum_{n=1}^{\infty} [\exp(-\frac{\epsilon^2\beta}{4B_n})]^{n^{2s}}, \\ &\leq 2C\sum_{n=1}^{\infty} [\exp(-m)]^{n^{2s}} < \infty, \end{aligned} \tag{4.20}$$

where m is a positive number not depending on n.

Using the inequality $e^{-x} \leq (\frac{ae^{-1}}{x})^a$ such that $a > 0, x > 0$. Then the right-hand side of (4.20) become

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(|S_n| > n^s\epsilon) &\leq 2C\sum_{n=1}^{\infty} (ae^{-1})^a \frac{1}{m^a n^{2sa}}, \\ &\leq \frac{2C(ae^{-1})^a}{m^a} \sum_{n=1}^{\infty} \frac{1}{n^{2sa}}, \\ &\leq \frac{2C(ae^{-1})^a}{m^a} \sum_{n=1}^{\infty} \frac{1}{n^2}, \\ &\leq \frac{2C(ae^{-1})^a}{m^a} \frac{\pi^2}{6}, \\ &\leq \infty \end{aligned}$$

by suppose that $a = \frac{1}{s}$. Hence the proof is complete.

Theorem 4.6 *Let $\{X_n, n \geq 1\}$ be a sequence of identically distributed WOD random variables. Assume that there exists a positive integer n_0 such that $0 \leq X_i \leq \frac{\alpha}{\lambda}$, for each $1 \leq i \leq n, n \geq n_0$ where α and β are positives constants, Then*

$$\sum_{n=1}^{\infty} \mathbb{P}(|S_n - \mathbb{E}S_n| \geq n\epsilon_n) < \infty. \tag{4.21}$$

Theorem 4.7 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $0 \leq X_i \leq \frac{\alpha}{\lambda} < \infty$, for each $i \geq 1$, such that α and λ are positive constants. Then for any $s > 0$*

$$\sum_{n=1}^{\infty} \mathbb{P}(|S_n - \mathbb{E}S_n| > n^s\epsilon) < \infty. \tag{4.22}$$

Proof. We have from Theorem (4.4), for any $\epsilon > 0$, there exists a sequence function $g(n)$ such that

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(|S_n - \mathbb{E}S_n| > n^s\epsilon) &\leq 2\sum_{n=1}^{\infty} g(n)[\exp(-\frac{(n^s\epsilon)^2\beta}{4B_n})], \\ &\leq 2C\sum_{n=1}^{\infty}[\exp(-\frac{\epsilon^2\beta}{4B_n})]^{n^{2s}} < \infty. \end{aligned}$$

After this result we get (4.22).

4.4 Application results in the first-order autoregressive $AR(1)$ model

4.4.1 The $AR(1)$ model

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

$$X_n = \theta X_{n-1} + \xi_i, \quad i = 1, 2, \ldots, \tag{4.23}$$

where $\{\xi_i, i \geq 0\}$ is a sequence of WOD random variables with $\xi_0 = X_0 = 0$, $0 < \mathbb{E}\xi_k^4 < \infty$, $k = 1, 2, \ldots$ and θ is a parameter with $|\theta| < 1$. Hence (4.23) as follows:

$$X_i = \sum_{j=0}^{\infty} \theta^j \xi_{i-j}. \tag{4.24}$$

The coefficient θ is fitted least squares, giving the estimator

$$\hat{\theta}_n = \frac{\sum_{j=1}^{n} X_j X_{j-1}}{\sum_{j=1}^{n} X_{j-1}^2}, \tag{4.25}$$

from (4.23) and (4.25) we obtain that

$$\hat{\theta}_n - \theta = \frac{\sum_{j=1}^{n} X_j X_{j-1}}{\sum_{j=1}^{n} X_{j-1}^2}. \tag{4.26}$$

Theorem 4.8 *let the conditions of theorem (4.4) be satisfied then for any* $\frac{(\mathbb{E}X_1^2)^{\frac{1}{2}}}{\rho} < \xi$ *positive, we have*

$$\mathbb{P}(\sqrt{n}|\hat{\theta}_n - \theta| > \rho) \leq g(n)\left[\exp\{-\frac{n^2(\rho^2\xi^2 - \mathbb{E}X_1)}{4B_n}\} + \exp\{-\frac{\mathbb{E}X_{j-1}^2 - n\xi^2}{4\mathbb{E}X_{j-1}^4}\}\right],$$

where $\mathbb{E}X_j^2 < \infty$ *and* $\mathbb{E}X_j^4 < \infty$.

Proof. From (4.26) it follows that

$$\mathbb{P}(\sqrt{n}|\hat{\theta}_n - \theta| > \rho) = \mathbb{P}\left(\left|\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_j X_{j-1}}{\frac{1}{n}\sum_{j=1}^{n} X_{j-1}^2}\right| > \rho\right).$$

Therefore, by virtue if the probability properties and Hölder's inequalities, we have that for any $\xi > 0$

$$\begin{aligned}
\mathbb{P}(\sqrt{n}|\hat{\theta}_n - \theta| > \rho) &\leq \mathbb{P}(\frac{1}{n}\sum_{j=1}^{n} X_j \geq \xi^2\rho^2) + \mathbb{P}(\frac{1}{n^2}\sum_{j=1}^{n} X_{j-1}^2 \leq \rho^2), \\
&= \mathbb{P}(\sum_{j=1}^{n} X_j \geq (\rho^2\xi^2)n) + \mathbb{P}(\sum_{j=1}^{n} X_{j-1}^2 \leq n^2\xi^2), \\
&= J_{n1} + J_{n2}.
\end{aligned}$$

Now, we start by estimate J_{n1} then estimate J_{n2}. We have that

$$\begin{aligned}
J_{n1} &= \mathbb{P}(\sum_{j=1}^{n} X_j \geq (\rho^2\xi^2)n), \qquad (4.27)\\
&= \mathbb{P}(\sum_{j=1}^{n} (X_j - \mathbb{E}X_j + \mathbb{E}X_j) \geq (\rho^2\xi^2)n), \\
&= \mathbb{P}(\sum_{j=1}^{n} (X_j - \mathbb{E}X_j) \geq (\rho^2\xi^2 - \mathbb{E}X_1)n), \\
&\leq \mathbb{P}(|\sum_{j=1}^{n} (X_j - \mathbb{E}X_j)| \geq (\rho^2\xi^2 - \mathbb{E}X_1)n).
\end{aligned}$$

By using the Theorem (4.4) the right hand side of (4.27) become

$$J_{n1} = \mathbb{P}(\sum_{j=1}^{n} X_j \geq (\rho^2\xi^2)n) \leq g(n)\exp\{-\frac{n^2(\rho^2\xi^2 - \mathbb{E}X_1)}{4B_n}\}. \tag{4.28}$$

Then we shall bound the right hande side of J_{n2}. By Markov's inequality, for any $\lambda > 0$ it follows that

$$\begin{aligned} J_{n2} &= \mathbb{P}(\sum_{j=1}^{n} X_{j-1}^2 \geq n^2\xi^2, \quad (4.29)\\ &= \mathbb{P}(n^2\xi^2 - \sum_{j=1}^{n} X_{j-1}^2 \geq 0),\\ &= \mathbb{E}(\amalg_{\{}n^2\xi^2 - \sum_{j=1}^{n} X_{j-1}^2 \geq 0\}),\\ &\leq \mathbb{E}(\exp\{\lambda(n^2\xi^2 - \sum_{j=1}^{n} X_{j-1}^2)\}),\\ &\leq e^{\lambda n^2\xi^2}\mathbb{E}(e^{-\lambda\sum_{j=1}^{n} X_{j-1}^2}),\\ &\leq e^{\lambda n^2\xi^2}\mathbb{E}(\prod_{j=1}^{n} e^{-\lambda X_{j-1}^2}). \end{aligned}$$

By using Corollary (4.1) and Lemma (4.2) the right hand side of the expression J_{n2} become

$$\begin{aligned} J_{n2} &\leq g(n)e^{\lambda n^2\xi^2}\prod_{i=1}^{n}\mathbb{E}(e^{-\lambda X_{j-1}^2}),\\ &\leq g(n)e^{\lambda n^2\xi^2}\prod_{i=1}^{n}\mathbb{E}\left(1 - \lambda X_{j-1}^2 + \frac{\lambda^2 X_{j-1}^4}{\beta}\right),\\ &\leq g(n)e^{\lambda n^2\xi^2}\left(1 - \lambda\mathbb{E}X_{j-1}^2 + \frac{\lambda^2\mathbb{E}X_{j-1}^4}{\beta}\right)^n,\\ &\leq g(n)\exp\{\lambda n^2\xi^2 - n(\lambda\mathbb{E}X_{j-1}^2 + \frac{\lambda^2\mathbb{E}X_{j-1}^4}{\beta})\},\\ &\leq g(n)\exp\{-\frac{\mathbb{E}X_{j-1}^2 - n\xi^2}{4\mathbb{E}X_{j-1}^4}\}. \end{aligned}$$

By taking $\lambda = \frac{(\mathbb{E}X_{j-1}^2 - n\xi^2)\beta}{2\mathbb{E}X_{j-1}^4}$. Then for any $\rho > 0$

$$\mathbb{P}(\sqrt{n}|\hat{\theta}_n - \theta| > \rho) \leq g(n)\left[\exp\{-\frac{n^2(\rho^2\xi^2 - \mathbb{E}X_1)}{4B_n}\} + \exp\{-\frac{\mathbb{E}X_{j-1}^2 - n\xi^2}{4\mathbb{E}X_{j-1}^4}\}\right]$$

Corollary 4.3 *The sequence $(\hat{\theta}_n)_{n\in\mathbb{N}}$ is completely converges to the parameter θ of autoregressive process $AR(1)$ model. Then we have*

$$\sum_{n=1}^{\infty}\mathbb{P}(\sqrt{n}|\hat{\theta}_n - \theta| > \rho) < \infty \quad (4.30)$$

Proof. By using Theorem (4.7) and $\mathbb{E}X_j^2 < \infty$, $\mathbb{E}X_j^4 < \infty$ we get the result of (4.30) immediately.

Bibliography

[1] Boulenoir, Z. 2018. 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(2)**: 1-6.

[2] Chebbab, I. 2018. Probability tail for extended negatively dependent random variables of partial sums and application to AR(1) model generated by END errors, International journal of statistics and economics. **19(1)**: 1-9.

[3] Chen, Y., Chen, A. and Ng, K. 2010. The strong law of large numbers for extend negatively dependent random variables., J App Probab **47**: 908-922.

[4] Chen, Y., Wang, L. and Wang, Y. 2013. Uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional risk models, Methodol Comput Appl Probab. **401**: 114-129.

[5] Chen, Y., Yuen, K. and Ng, K. 2011. Precise large deviations of random sums in presence of negative dependence and consistent variation, Methodol Comput Appl Probab. **13**: 821-833.

[6] He, W. and Cheng, DY. and Wang, Y. 2013. Asymptotic lower bounds of precise large deviations ith nonnegatve and dependent random variables, Stat Probab Lett. **83**: 331-338.

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

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

[9] Lui, X., Gao, Q. and Wang, Y. 2012. Anote on a dependent risk model with constant interest rate, Stat Probab Lett. **82**:707-712.

[10] Shen, A. 2011.Probability inequalities for END sequence and their applications, J Inequal Appl. **98**: 1-12.

[11] Shen, A. 2013a. On asymptotic approximation of inverse moments for a class of nonnegative random variables, Statistics. **48**: 1371-1379.

[12] Shen, A. 2013b.Bernstein-type inequality for widely dependnet sequence and its application to nonparametric regression models, Abstract Appl Anal. **9**: 1-9.

[13] Wang, K., Wang, Y. and Gao, Q.2013. Uniform asymptotics for the finite-time ruin probability of a new dependent risk model with a constant interest rate, Methodol Comput Appl Probab. **15**: 109-124.

[14] Wang Y, C. D. 2011. Basic renewal theorems for a random walk with widely dependent increments and their applications, J Math Anal Appl. **384**: 597-606.

[15] Wang Y, Cui Z, W. K. M. X. 2012. Uniform asymptotics of the finite-time ruin probability for all times, J Math Anal Appl. **390**: 208-223.

[16] D. H. Qiu, T. C. HU, M. O. CABRERA, and A. VOLODIN. 2012. Complete convergence for weighted sums of arrays of banach-space-valued random elements, J Lithanian Math. **52**: 316-325.

[17] H. C. Kim. 2011. Exponential inequalities for ELNQD random variables with applications, J Chungcheong Math Society. **24**: 784-793.

[18] J. I. Baek and S.T. Park. 2010. Convergence of weighted sums for arrays of negatively dependent random variables and its applications, J Theor Probab. **23**: 362-377.

[19] Ko, M.-H, Choi, Y.-K, and Choi, Y.-S. 2007. Exponential probability inequality for linearly negative quadrant dependent random variables. Communications of Korean Math Soc. **22**: 137-143.

[20] Masry, E. 1986. Recursive porobability density estimation for weakly dependent stationary processes, IEEE Trans Inform Theor. **32**: 254-267.

[21] Sung S. H. 2011. On the exponetial inequalities for negatively dependent random variables, J Math Anal and Appl. **381**: 538-545.

[22] S. H. Sung and A. I. Volodin. 2006. On the rate of complete convergence for weighted sums of arrays of random elements, J Korean Math Soc. **43**: 815-828.

[23] T. S. Kim. 2007. On the exponential inequality for negative dependent sequence, Communications of Korean Math Society. **22**: 315-321.

[24] Yongming Li, Chengdong Wei,Shanchao Yang. 2010. The recursive kernel distribution function estimator based on negatively and positively associated sequences, Communications in statistics-Theory and Methods. **39**: 3585-3595.

[25] Yongming Li . 2017. On the rate of strong convergence for a recursive probability density estimator of END samples and its applications, J Mathematical Inequalities. **11**: 335-343.

[26] Y. Q. Chen, A. Y. Chen, W. NG Kai. 2010. The strong law of large numbers for extended negatively dependent random variables, J Appl Prob. **47**: 908-922.

[27] Y. F. Wu, M. Guan. 2012. Convergence properties of the partial sums for sequences of END random variables, J Korean Math Soc. **49**: 1097-1110.

Bibliography

[1] Boulenoir, Z. 2018. 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(2)**: 1-6.

[2] Chebbab, I. 2018. Probability tail for extended negatively dependent random variables of partial sums and application to AR(1) model generated by END errors, International journal of statistics and economics. **19(1)**: 1-9.

[3] Chen, Y., Chen, A. and Ng, K. 2010. The strong law of large numbers for extend negatively dependent random variables., J App Probab **47**: 908-922.

[4] Chen, Y., Wang, L. and Wang, Y. 2013. Uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional risk models, Methodol Comput Appl Probab. **401**: 114-129.

[5] Chen, Y., Yuen, K. and Ng, K. 2011. Precise large deviations of random sums in presence of negative dependence and consistent variation, Methodol Comput Appl Probab. **13**: 821-833.

[6] D. H. Qiu, T. C. HU, M. O. CABRERA, and A. VOLODIN. 2012. Complete convergence for weighted sums of arrays of banach-space-valued random elements, J Lithanian Math. **52**: 316-325.

[7] Fuxia Cheng. 2005. Asymptotic distributions of error density estimators in first-order autoregressive models, The Indian Journal of Statistics. **67**: 553-567.

[8] H. Y. Liang, J. Baek. 2004. Asymptotic normality of recursive density density estimates under some dependence assumptions, Metrika. **60**: 155-166.

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

[10] H. C. Kim. 2011. Exponential inequalities for ELNQD random variables with applications, J Chungcheong Math Society. **24**: 784-793.

[11] He, W. and Cheng, DY. and Wang, Y. 2013. Asymptotic lower bounds of precise large deviations ith nonnegatve and dependent random variables, Stat Probab Lett. **83**: 331-338.

[12] J. I. Baek and S.T. Park. 2010. Convergence of weighted sums for arrays of negatively dependent random variables and its applications, J Theor Probab. **23**: 362-377.

[13] Kheira Berkane, Samir Bennaissa. (2019): Complete convergence for widely orthant dependent random variables and its applications in autoregressives AR(1) models. International Journal of Mathematics and Computation. 30(2): 1-12.

[14] Ko, M.-H, Choi, Y.-K, and Choi, Y.-S. 2007. Exponential probability inequality for linearly negative quadrant dependent random variables. Communications of Korean Math Soc. **22**: 137-143.

[15] Li, Y. F, Yang, S. C. 2005. Uniformly asymptotic normality of the smooth estimation of the distribution function under associated samples, Acta Math Appl Sinica. **28**: 639-651.

[16] Li, Y. M, Yang, S. C. 2003. Asymptotic normality of the recursive kernel estimate of a probability density function under negatively associated, Chin J Appl Probab Statist. **19**: 383-393.

[17] Liang, H. Y, Jing, B. Y. 2005. Asymptotic properties for estimate of nonparametric regression models based on negatively associated sequences, J Multivariate Anal. **95**: 227-245.

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

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

[20] Lui, X, Gao, Q. and Wang, Y. 2012. Anote on a dependent risk model with constant interest rate, Stat Probab Lett. **82**:707-712.

[21] Loynes, R.M. 1980. The empirical d.f. of residuals from generalized regression, Ann Statist. **8**: 285-298.

[22] Masry, E. 1986. Recursive porobability density estimation for weakly dependent stationary processes, IEEE Trans Inform Theor. **32**: 254-267.

[23] Portnoy, S. 1986. Asymptotic behavior of the empirical distribution of M-estimated residuals from a regression model with many parameters, Ann Statist. **14**: 1152-1170.

[24] Roussas, G. G. 1994. Asymptotic normality of random fields of positively and negatively associated processes, J Multivariate Anal. **50**: 152-173.

[25] Roussas, G. G. 1995. Asymptotic normality of a smooth estimate of a random field distribution under association, Statist Probab Lett. **24**: 77-90.

[26] S. C. Yang. 2003. Consistency of nearest neighbor estimator of density function for negative associated samples, Acta Math Appl Sinica. **26**: 385-395.

[27] Shen, A. 2011. Probability inequalities for END sequence and their applications, J Inequal Appl. **98**: 1-12.

[28] Shen, A. 2013a. On asymptotic approximation of inverse moments for a class of nonnegative random variables, Statistics. **48**: 1371-1379.

[29] Shen, A. 2013b. Bernstein-type inequality for widely dependnet sequence and its application to non-parametric regression models, Abstract Appl Anal. **9**: 1-9.

[30] Sung S. H. 2011. On the exponetial inequalities for negatively dependent random variables, J Math Anal and Appl. **381**: 538-545.

[31] S. H. Sung and A. I. Volodin. 2006. On the rate of complete convergence for weighted sums of arrays of random elements, J Korean Math Soc. **43**: 815-828.

[32] T. S. Kim. 2007. On the exponential inequality for negative dependent sequence, Communications of Korean Math Society. **22**: 315-321.

[33] Wang, K., Wang, Y. and Gao, Q.2013. Uniform asymptotics for the finite-time ruin probability of a new dependent risk model with a constant interest rate, Methodol Comput Appl Probab. **15**: 109-124.

[34] Wang Y, C. D. 2011. Basic renewal theorems for a random walk with widely dependent increments and their applications, J Math Anal Appl. **384**: 597-606.

[35] Wang Y, Cui Z, W. K. M. X. 2012. Uniform asymptotics of the finite-time ruin probability for all times, J Math Anal Appl. **390**: 208-223.

[36] X. J. Wang, L. L. Zheng, C. Xu, S. H. Hu. 2015. Complete consistency for the estimator of non parametric regression models based on extended negatively dependent errors. Statistics: A journal of Theoretical and Applied Statistics. **49**: 396-407.

[37] Y. M. Li, S. C. Yang. 2005. Strong convergence rate of recursive probability density estimator for NA sequences, Chin J Engineering Mathematics. **22**: 659-665.

[38] Yang, S. C. 2005. Uniform asymptotic normality of the regression weighted estimator for positively associated samples, Chin J Appl Probab Statist. **21**: 150-160.

[39] Yongfeng Wu And Mei Guan. 2012. Convergence propertes of the sums for sequences of END random variables, Korean mathematical society. **49**: 1097-1110.

[40] Yongming Li. 2017. On the rate of strong convergence for a recursive probability density estimator of END samples and its applications, J Math Inequal. **11**: 335-343.

[41] Yongming Li, Chengdong Wei,Shanchao Yang. 2010. The recursive kernel distibution function estimator based on negatively and positively associated sequences, Communications in statistics-Theory and Methods. **39**: 3585-3595.

[42] Y. Q. Chen, A. Y. Chen, W. NG Kai. 2010. The strong law of large numbers for extended negatively dependent random variables, J Appl Prob. **47**: 908-922.

[43] Y. F. Wu, M. Guan. 2012. Convergence properties of the partial sums for sequences of END random variables, J Korean Math Soc. **49**: 1097-1110.

Conclusion

In this thesis we establish asymptotic distribution of the END random error in first-order autoregressive processes using the recursive kernel estimator and complete convergence for recursive probability density estimator of LNQD orthant dependent variables. Then we study the complete convergence for widely orthant dependent random variables and its applications in autoregressive $AR(1)$ models.

Perspectives

In this section, we write some perspectives for future researches.

- Consider the asymptotique distributions of the error density estimators in p order autoregressive models .
- We establich recursive estimation of the probability density function $f(x)$ for stationary processes X_t in Banach space.
- It is possible to study the complete convergence for m acceptable random variables and its applications in autoregressive $AR(p)$ models.

Summary

The objective of this thesis is to study the non-parametric estimation by recursive nuclei as well as their applications to the linear model.

Then, we study its asymptotic properties (the almost complete convergence of recursive estimation).

Finally, we use the results obtained in the autoregressive processes.

الملخص

الهدف من هذه الأطروحة هو دراسة التقدير غير المعياري بواسطة أنوية متكررة وكذلك تطبيقاتها على النموذج الخطي.

بعد ذلك، ندرس خواصه الفرعية (التقارب الشبه الكامل للمقدر المتكرر).

أخيرًا، نستخدم النتائج التي تم الحصول عليها في عمليات الرد التلقائي.

Résumé

L'obejectif de cette thèse est d'étudier l'estimation non-paramétrique par noyaux récursifs ainsi qu'à leurs applications aux modèle linéaire.

Ensuite, nous étudions ses propriétés asymptotiques (la convergence presque complète de l'estimation récursifs).

Nous utilisons enfin les résultats obtenus aux processus autorégressifs