REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mohammed Seddik Ben Yahia-Jijel

Faculté des Sciences Exactes et Informatique

Département de Mathématiques

N d'ordre : ....

N de séries : ....

Mémoire

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques Appliquée

Option : probabilités et statistique

Thème

Estimation non paramétrique à erreur- alpha- mélange

Présenté par

Fettane Rima

Boufnare Yassmina

Devant le jury composé de

Président :	M^{me} Madi Meriem ,	**(M.A.A)**	Université de Jijel
Encadreur :	M^{me} Abdi Zeyneb	**(M.A.A)**	Université de Jijel
Examinateur :	M^{me} Yaakoubi Fatima	**(M.A.A)**	Université de Jijel

Promotion 2019/2020

Remerciements

En préambule à ce mémoire nous remercions **ALLAH** qui nous a aidé et nous a donné la patience et le courage durant ces longues années d'étude.

Nous souhaitons adresser nos remerciements les plus sincères aux personnes qui nous ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de cette formidable année universitaire et particulièrement notre encadreur **Z.Abdi**. D'avoir voulu proposer et assurer la direction de ce mémoire, pour sa confiance et ses conseils judicieux et sa totale disponibilité.

Nous tenons à remercier sincèrement les membres de jury **yaakoubi fatima** et **Madi Meriem** qui ont accepté de jurer notre travail.

On n'oublie pas nos parents, membres de nos familles pour leur contribution, leur soutien et leur patience.

Enfin, nous adressons nos plus sincères remerciements à tous nos proches et amis, qui nous ont toujours encouragé au cours de la réalisation de ce mémoire.

Merci à tous et à toutes.

Dédicace

Je dédie ce travail de fin d'études.

À mes chèrs parents ⋆ nassima ⋆ qui m'ont accompagné

et soutenu durant toute ma vie.

À mes soeurs sara,iman.

À mes chères hadjer, yasmina, sara,et iman.

À mes amies de ma promotion.

À tous qui m'ont a aidé de prés ou de loin.

À tous mes enseignants depuis le primaire jusqu'à maintenant.

ÈEnfin, je dédie ce mémoire à ceux qui m'aiment et surtout ceux que j'aime.

Rima

Dédicace

Je dédie ce travail de fin d'études.

À la mémoire de mon pére, le destin ne m'a pas donné la chance de le connaitre et de lui exprimer tout mon amour et mon respect. Puisse Dieu le tout puissant lui accorde sa clémence, et l'accueille dans son vaste paradis. Fi El- Firdaws inchallah.

À ma mère qui grâce à elle, j'ai pu continuer, celle qui m'a encouragée par son amour, sa tendresse, ses larmes, celle qui donne toujours sans attendre en retour.

À mes très chers frères , Mouhcen, Mouhamed et Oussama.

À mon marie Ali.

À ma sœurs, Wafia.

À mes meilleurs amis Rima et Amina.

À mes amies de ma promotion.

A tous mes enseignants depuis le surtout ceux que j'aime.

Yassamina

TABLE DES MATIÈRES

Notations **6**

1 Estimation non paramétrique **9**

1.1 Introduction 9

1.2 Principe d'estimation non paramétrique 9

1.3 Le modèle non paramétrique 10

1.4 Méthode du Noyau 11

1.4.1 Principe de la méthode 11

1.4.2 La construction de l'estimateur 12

1.4.3 Propriétés de l'estimateur à noyau 15

1.5 Consistance : 16

1.6 Convergence 18

1.6.1 Convergence en probabilité 18

1.6.2 Convergence presque complète 18

1.6.3 Convergence en moyenne quadratique 21

1.6.4 Convergence presque sûre 22

1.6.5 Converge en loi 22

1.6.6 Liens entre les différents type de convergence 23

2 Les fonctions *alpha* mélangeants **24**

2.1 La dépendance 24

2.2 Inégalités exponentielles 28
2.3 Éléments sur les séries chronologiques 29
2.3.1 Séries chronologiques 29
2.3.2 Processus stochastiques 29
2.3.3 Fonctions d'autocovariance et d'autocorrélations 30
2.3.4 Les processus linéaires 31
2.4 Variables mélangeantes 33
2.4.1 Coefficients de dépendance 33
2.4.2 Inégalités de covariance 34
2.4.3 Champ mélangeant 35
2.4.4 Processus mélangeant 35

3 Grand déviation des estimations dans les processus 38
3.1 Introduction 38
3.2 Notation et hypothèses 38
3.3 Position du problème 39
3.4 Construction de l'estimateur 39
3.5 Déviation de l'estimateur 41
3.5.1 Inégalité exponentielle 41
3.5.2 Déviation de l'estimateur 43
3.5.3 Convergence presque complet de l'estimateur 44

Bibliographie 46

A Les théorème utilisès : 49

TABLE DES FIGURES

1.1 Les courbes des noyaux 15

RÉSUMÉ

Les inégalités exponentielles sont très importantes en théorie de l'estimation fonctionnelle. Elle permettent de donner une majoration de l'erreur de l'estimation qui quantifie la qualité de l'estimateur.
La thématique abordées dans ce travail est l'étude de grands déviations (notamment pour les processus) à partir de l'inégalité exponentielle et de la convergence presque complète, qui nous donne un bon sens pour la déviation de notre estimation dans le cas alpha mélange.
l'intérêt majeur de cas outils dans les études asymptotiques est la détermination des vitesse de convergence des estimateur étudiés.

NOTATIONS

$r(\cdot)$	Fonction de régression.
$K(\cdot)$	Fonction noyau.
$\mathbb{E}$	Espérance de probabilité.
$1(.)$	L'indicatrice de l'évènement.
$\mathbb{V}ar$	Variance d'un estimateur.
$f(.)$	Densitè marginale de X.
$\mathbb{C}ov(,)$	Covariance d'un estimateur.
$\hat{f}$	L'estimateur de la densitè marginale de X.
MSE	L'erreur moyenne quadratique.
$\mathbb{B}iais$	Biais d'un estimateur.
$MISE$	L'erreur moyenne quadratique integrèe.
F	Densitè rèpartition.
$C.p.c$	Converge présque complet
$C.p.s$	Converge présque sure
$C.p$	Converge en probabilité
$C.l$	Converge en loi
$Lips$	Lipshitziennes
$\|\|\cdot\|\|_\infty$	La norme
sup	superieur
$\sum$	la somme

INTRODUCTION

La théorie de l'estimation est une des branches les plus basiques de la statistique. Cette théorie est habituellement divisée en deux composantes principales : l'estimations paramétrique et l'estimations non paramétrique.
Les méthode non paramétrique ont connu un essor important depuis les travaux de bosq (1979), collomb (1980) et robinson(1983). L'estimation de fonction non paramétrique est un outil important pour l'analyse des données, à la fois à des fins d'inférence statistique et de visualisation graphique. Dans ce dernier contexte, il faut distinguer si une caractéristique de l'estimation de la courbe est uniquement due à des fluctuations aléatoires, ou si elle capture la structure pertinente de la courbe inconnue. À cette fin. des estimations d'intervalle sont fréquemment utilisées ; et des intervalles uniformes (c'est-à-dire des bandes de confiance) semblent plus appropriés que des intervalles de confiance uniformes.
Historiquement, le principe des méthodes non paramétriques remonte au 19 siècle selon Cleveland and Loader (1995), toute fois les premiers travaux modernes sur ce sujet datent des années 50. La première application que nous verrons relève de l'estimation de fonctions de densité par des méthodes d'opérateur à noyau (kernel) avec les travaux fondateurs de Rosenblatt (1956) et de Parzen (1962). Dans ce domaine, on identifie deux papiers fondateurs publiés la même année : Nadaraya (1964) et Watson (1964).

Un utile populaire aujourd'hui dans la classification réside dans le modèle de mélange, les modèles de mélanges apparu dans les travaux de pearson (1984) sont utilisée avec succès dans le bon nombre de discipline comme

économétrie et le domaine de la santé publique. nous nous tiendrons particulièrement à la notion de mélange fort appelée aussi α –mélangeantes. Ce choix a été motivé du fait que ce type de variables soit assez général et très présent dans la pratique statistique et des nombreux résultats dont on dispose et qui nous permettront d'étudier ce type de modélisation. On considère des échantillons de variables alpha-mélangeantes, qui est l'une des plus générales et notamment vérifiée par différents types de processus.
Dans ce mémoire nous exposerons le travail qui composé principalement de trois chapitres :

- ▷ Dans Le premiers chapitre, en présente la méthode de l'estimation par noyau nous présentons également les propriétés statistiques de cette méthode aussi nous présentons les types de convergences.
- ▷ Dans le deuxième chapitre nous présentons les fonctions α- mélangeant avec une liste des définitions permettant de fixer le vocabulaire utilisé dans la suite de document, cette liste contient entre Éléments sur les séries chronologiques et Variables mélangeantes .
- ▷ Finalement, dans le dernier chapitre nous présentons Grand déviation des estimations dans les processus

CHAPITRE 1

ESTIMATION NON PARAMÉTRIQUE

1.1 Introduction

L'utilisation des modèles paramétrique est très fréquente l'on fait appel à la régression afin d'analyser un jeu de donnés ou : il ya certains situation où ce modèle ne sont pas appropriés et où le choix d'un modèle non paramétrique est préférable.

Dans ce chapitre nous étudierons certains méthode d'estimation non paramétrique à savoir, particulièrement la méthode de noyaux, qui est une méthode très pratique lorsqu'on s'intéresse à la relation entre une variable réponse Y est une variable explicative x, mais que l'on ne veut supposer aucune forme particulière pour la relation entre ces deux variables. laissant ainsi aux données le choix exclusif de cette forme.

1.2 Principe d'estimation non paramétrique

Lorsque nous souhaitons décrire l'influence d'une variable quantitative sur un événement en faisant le moins d'hypothèses possible sur la forme de la relation, nous distinguons deux approches :

- L'approche de la régression paramétrique.
- L'approche de la régression non paramétrique.

Le but d'un modèle de régression consiste à déterminer la façon dont l'espérance d'une variable dépendante Y dépend d'un ensemble de variable explicative X, Supposons que $x \in \mathbb{R}$ le problème consiste donc à déterminer pour chaque réalisation x de la variable X ; la valeur de la fonction $f(x)$, dite fonction de lien ou fonction de régression.

Définition 1.2.1. *On appelle fonction de régression, la fonction $f(x)$ qui a pour toute réalisation x de la variable explicative X associe la quantité :*

$$E(Y/X = x) = f(x). \tag{1.1}$$

1.3 Le modèle non paramétrique

Nous pouvons retenir une approche non paramétrique dans la quelle on va estimer la relation entre le niveau moyen de Y et toutes les valeurs réalisées de X, nous ne supposons aucune forme spécifique sur la fonction de régression.[2]

Définition 1.3.1. *Dans un modèle de régression non paramétrique, la fonction de régression*

- *N'a pas de la forme explicité.*
- *Ne peut pas s'écrire en fonction d'un nombre réduit de paramètres*

$$\mathbb{E}(Y/X = x) = f(x).$$

Remarque 1.3.1.

- *Le principal avantage de cette approche est qu'elle ne nécessite aucune hypothèse a priori sur la forme du lien entre X et Y.*

Remarque 1.3.2. *Avec une approche non paramétrique, on obtient à :*

- *Une représentation graphique de la relation entre$X$ et Y,*
- *Il n'existe pas de forme analytique de la fonction de lien $f(x)$,*
- *Le modèle qui va nous intéresser est un modèle non paramétrique en ce sens que la seule condition que nous ferons sur la fonction est une condition de régularité*

$$\mathbb{E}(Y/X = x) = f(x).$$

Définition 1.3.2. *On dit qu'un estimateur $\hat{f}$ de f est sans biais si :*

$$\mathbb{E}(\hat{f}) = f(x).$$

Définition 1.3.3. *On dit qu'un estimateur $\hat{f}$ de f est asymptotiquement uniformément sans biais si :*

$$\lim_{n\to+\infty} sup_x|[\hat{f(x)} - f(x)]| = 0.$$

Définition 1.3.4. *L'erreur moyenne quadratique MSE :*

$$\begin{aligned}
MSE(f(x) - \hat{f}(x)) &= \mathbb{E}(f(x) - \hat{f}(x))^2 \\
&= \mathbb{E}(f(x))^2 + \mathbb{E}(\hat{f}(x) - 2\mathbb{E}(f(x)\hat{f}(x)) \\
&= \mathbb{E}(f(x))^2 + \mathbb{E}(\hat{f}(x))^2 - 2\mathbb{E}(f(x)\hat{f}(x) + [\mathbb{E}(\hat{f}(x))]^2 - 2[\mathbb{E}(\hat{f}(x))]^2 \\
&= (f(x))^2 + \mathbb{E}(\hat{(}f(x))^2 - 2f(x)\mathbb{E}(\hat{f}(x)) + [\mathbb{E}(\hat{f}(x))]^2 - [\mathbb{E}(\hat{f}(x))]^2 \\
&= [\mathbb{E}(\hat{f}(x) - f(x))]^2 + \mathbb{E}(\hat{f}(x))^2 - [\mathbb{E}(\hat{f}(x))]^2 \\
&= \mathbb{B}iais(\hat{f}(x))^2 + \mathbb{V}ar(\hat{f}(x)).
\end{aligned}$$

Définition 1.3.5. *L'erreur moyenne quadratique intégrée $MISE$:*

$$\begin{aligned}
MISE(\hat{f}, f) &= \int MSE(f(x), \hat{f}(x))dx \\
&= \int \mathbb{B}iais(\hat{f}(x))^2 dx + \int \mathbb{V}ar(\hat{f}(x))dx. \qquad (1.2)
\end{aligned}$$

Définition 1.3.6. *On dit qu'un estimateur $\hat{f}$ de f est ponctuellement consistant en moyenne quadratique si :*

$$\lim_{x\mapsto\infty} MSE(f(x), \hat{f}(x)) = 0.$$

Définition 1.3.7. *On dit qu'un estimateur $\hat{f}$ de f est uniformément en moyenne quadratique intégrée si :*

$$\lim_{x\mapsto\infty} MISE(f(x), \hat{f}(x)) = 0.$$

Définition 1.3.8. *On dit qu'un estimateur $\hat{f}$ de f est asymptotiquement normal si :*

$\hat{f} \longrightarrow \mathcal{N}(\mathbb{E}(\hat{f}), \mathbb{V}ar(\hat{f}))$ *en loi.*

1.4 Méthode du Noyau

1.4.1 Principe de la méthode

Le problème consiste à estimer la fonction de régression en tous points x_1, x_2,.....,x_n :

Le principe de la méthode du noyau repose en fait sur des méthodes de lissage, elle donne pour estimateur de

$$\mathbb{E}(Y/X = x),$$

Une moyenne pondérée des valeurs y_i pour les i dont le point x_i est proche du point d'estimation.

1.4.2 La construction de l'estimateur

Nous supposons que nous disposons des observations $(x_1, y_1), ..., (x_n; y_n)$ du couple $(X; Y)$. On se propose de construire un estimateur $\hat{f}(x)$ de la fonction de régression partir des couples d'observations. Le premier estimateur rencontré dans la littérature est l'estimateur à noyau de Nadaraye-Watson. Il est construit à partir d'une fonction $K(.)$ et d'une fenêtre h, de manière analogue à l'estimateur à noyau de la fonction de densité $f(.)$ introduit par [19] et [21]. Nous rappelons la définition de l'estimateur Parzen et Rosenblatt :

$$\begin{aligned}\hat{f}(x) &= \frac{1}{2nh_n}\sum_{i=1}^{n} 1_{[x-h,x+h]}(x_i)\\ &= \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h_n}\right).\end{aligned}$$

- Dans un premier temps, nous désignons par fenêtre une suite $\{h_n \geq 1\}$ de nombres strictement positifs vérifiant $h_n \to 0$, lorsque $n \to +\infty$.
- La fonction $K : \mathbb{R} \to \mathbb{R}$ sera supposée mesurable et satisfait certaines hypothèses basiques parmi celles énoncées ci-dessous :
  1- K est bornée ; i.e, $\sup_{u\in\mathbb{R}} \mid K(u) \mid \leq M < \infty$,
  2- $\lim\limits_{|u|\mapsto\infty} \mid u \mid K(u)=0$,
  3- $\displaystyle\int \mid K(u) \mid du < \infty$,
  4- $\displaystyle\int K(u)$ du $= 1$,
  5- $\forall u \in \mathbb{R},\ K(u)=K(-u)$,
  6- $\displaystyle\int_{-\infty}^{+\infty} u^2 K(u)du < \infty$,
  7- $\int_{\mathbb{R}} uK(u)du = 0$.

- Nous reprenons le modèle :

$$f(x) = E(Y/X = x) = \frac{f(x)}{r_X(x)}.$$

Nous avons

$$\begin{aligned} r(x) &= \int_R y f_{X,Y}(x,y)dy \\ &= \lim_{h\to 0} \frac{1}{2h} \int_{x-h}^{x+h} \int_R y F_{X,Y}(dx,dy) \\ &= \lim_{h\to 0} \frac{1}{2h} \mathbb{E}[Y 1_{(|x_i - x| \le h)}]. \end{aligned}$$

Ou $F_{X,Y}(.,.)$ est la fonction de répartition du (X,Y)

$$\hat{r}(x) = \frac{1}{n} \sum_{i=1}^{n} y_i \frac{1_{(|x_i - x| \le h)}}{2h}.$$

Donc $f(x)$ est estimée par :

$$\hat{f}(x) = \frac{\hat{r}(x)}{\hat{f}_X(x)} = \frac{\sum_{i=1}^n y_i 1_{(|x_i - x| \le h)}}{\sum_{i=1}^n 1_{(|x_i - x| \le h)}}.$$

Cet estimateur se présent sous la forme d'une moyenne locale pondérée des valeurs y_i, mais il présente le désavantage d'être discontinu. Sa généralisation naturelle est l'estimateur à noyau, défini comme suit :

$$\hat{f}(x) = \frac{\sum_{i=1}^n y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}, \qquad \forall x$$

Cet estimateur a été introduit par Nadaraya et Watson.

$$\hat{r}(x) = \frac{\sum_{i=1}^{n} y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)} 1_{\{K(\frac{x - x_i}{h}) \neq 0\}}, \qquad \forall x.$$

$$\hat{f}(x) = \begin{cases} \frac{\sum_{i=1}^n y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}, & si\{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) \neq 0\} \\ \frac{1}{n}\sum_{i=1}^n y_i, & sinon \end{cases}$$

Le noyau K détermine la forme du voisinage autour du point x et la fenêtre h contrôle la taille de ce voisinage, i.e, le nombre d'observations prises pour effectuer la moyenne locale.

Quelque fonctions noyaux usuelles

-i- Noyau rectangulaire (uniforme) :

$$K_1(u) = \begin{cases} 1/2, & |u| \leq 1 \\ 0, & |u| > 1 \end{cases}$$

-ii- Noyau triangulaire :

$$K_2(u) = \begin{cases} (1 - |u|), & |u| \leq 1 \\ 0, & |u| > 1 \end{cases}$$

-iii- Noyau parabolique ou d'Epanechnikov :

$$K_3(u) = \begin{cases} \frac{3}{4}(1 - u^2), & |u| \leq 1 \\ 0, & |u| > 1 \end{cases}$$

-iv- Noyau "biweight" quadratique :

$$K_4(u) = \begin{cases} \frac{15}{16}(1 - u^2), & |u| \leq 1 \\ 0, & |u| > 1 \end{cases}$$

-v- Noyau gaussien $K_5(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2})$

-vi- Noyau cubique :

$$K_6(u) = \begin{cases} \frac{35}{32}(1 - u^2)^3, & |u| \leq 1 \\ 0, & |u| > 1 \end{cases}$$

Les courbes de certains noyaux sont présentées ci-dessous :

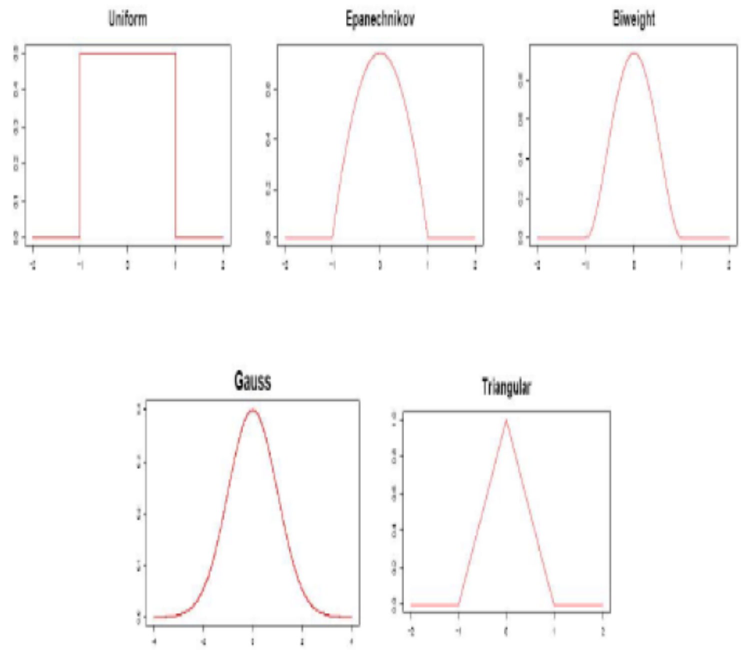

FIGURE 1.1 – Les courbes des noyaux

1.4.3 Propriétés de l'estimateur à noyau

Les propriétés de consistance et d'absence de biais sont reprises de Nadaraya (1989). Posons sur la fonction noyau $K : \mathbb{R} \longrightarrow \mathbb{R}$ les hypothèses suiantes :

$$\begin{cases} - \lim\limits_{|u| \mapsto \infty} \mid u \mid K(u) = 0, \\ - \displaystyle\int \mid K(u) \mid du < \infty, \\ - \displaystyle\int K(u) du = 1, \\ - \forall u \in \mathbb{R}, \quad K(u) = K(-u), \\ - \displaystyle\int_{-\infty}^{+\infty} u^2 K(u) du < \infty, \\ - \int_{\mathbb{R}} u K(u) du = 0. \end{cases}$$

Rappelons que :

$\hat{r}_n(x) = \dfrac{\hat{g}_n(x)}{\hat{f}_n(x)}$, où $\quad \hat{g}_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} y_i k\left(\frac{x_i - x}{h_n}\right)$ et $\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} k\left(\frac{x_i - x}{h_n}\right)$

1.5 Consistance :

L'estimateur à noyau de la densité dépend de deux paramètres la fenêtre h et le noyau K. Le noyau K établit l'aspect du voisinage de x et h contrôle la taille de ce voisinage, donc h est le paramètre prédominant pour avoir de bonnes propriétés asymptotiques, néanmoins le noyau K ne doit pas être négligé, comme le montre le travail de Parzen (1962) cité ci dessous sur la consistance de cet estimateur.Cette dernière est obtenue, en se basant sur l'étude asymptotique du biais, de la variance et de la décomposition suivante :

$$\mathbb{E}[f_n(x,h_n) - f(x)]^2 = \mathbb{V}ar[f_n(x,h_n)] + [Biais f_n(x,h_n)]^2$$

Théorème 1.5.1.
Supposons que le noyau K, $\mathbb{E}(y^2) < \infty$ *et* $f_X(x)$ *est strictement positive si* $h_n \longrightarrow 0$ *,* $nh_n \longrightarrow +\infty$ *(quand* $n \longrightarrow \infty$*), alors* $\hat{r}_n(x)$ *est un estimateur consistant de* $r(x)$*.*

Démonstration 1.5.2.
Nous déduisons du théorème de Bochner que, lorsque $h_n \longrightarrow 0$

$$\begin{aligned}
\mathbb{E}[\hat{f}_n(x)] &= \mathbb{E}\left[\frac{1}{nh_n}\sum_{i=1}^{n} k\left(\frac{x_i - x}{h_n}\right)\right] \\
&= \frac{1}{h_n}\mathbb{E}\left[k\left(\frac{x-X}{h_n}\right)\right] \\
&= \frac{1}{h_n}\int_R K\left(\frac{x-t}{h_n}\right) f_x(t)dt \longrightarrow f_X(x)\int_R K(t)dt \\
&\to f_X(x)
\end{aligned}$$

Donc $\hat{f}_n(x)$ est un estimateur asymptotiquement sans biais.
D'autre part, comme les X_i sont indépendantes et identiquement distribuées, il vient que :

$$\begin{aligned}
\mathbb{V}ar(\hat{f}(x)) &= \frac{1}{n}\mathbb{V}ar\left[\frac{1}{h_n}K\left(\frac{x-X}{h_n}\right)\right] \\
&\leq \frac{1}{h_n}\mathbb{E}\left[\frac{1}{h_n}K\left(\frac{x-X}{h_n}\right)\right]^2 \\
&\leq \frac{1}{h_n}\int_R \frac{1}{h_n}K^2\left(\frac{x-t}{h_n}\right) f_X(t)dt
\end{aligned}$$

D'après le théorème de Bochner :

$\int_R \frac{1}{h_n}K^2(\frac{x-t}{h_n})f_X(t)dt \longrightarrow f_X(x)\int_R K^2 dt \longrightarrow \infty$ quand $n \longrightarrow \infty$

Donc :

$\mathbb{V}ar(\hat{f}(x)) \longrightarrow 0$ quand $nh_n \longrightarrow \infty$

Puisque$\hat{f}(x)$ est un estimateur consistant de $f_X(x)$; il suffit donc de montrer que $g_n(x)$ est un estimateur consistant de :

$$g(x) = \int_R yf(x,y)dy$$

Nous avons :

$$\begin{aligned}
\mathbb{E}(\hat{g}_n(x)) &= \left[\frac{1}{nh_n}\sum_{i=1}^{n} Y_i k\left(\frac{x-X_i}{h_n}\right)\right] \\
&= \frac{1}{h_n}\mathbb{E}\left[YK\left(\frac{x-X}{h_n}\right)\right] \\
&= \frac{1}{h_n}\int_R \mathbb{E}\left(y/X=x\right)\frac{x-t}{h_n}f_X(t)dt \\
&= \frac{1}{h_n}\int_R r(t)K\left(\frac{x-t}{h_n}\right)f_X(t)dt \\
&= \frac{1}{h_n}\int_R K\left(\frac{x-t}{h_n}\right)r(t)f_X(t)dt \\
&= \frac{1}{h_n}\int_R K\left(\frac{x-t}{h_n}\right)g(t)dt \longrightarrow g(t)
\end{aligned}$$

par le théorème de Bochner :

De plus si : $m(x) = \int_R y^2 f(x,y)dy$,

$\mathbb{V}ar(\hat{g}_n(x)) = \mathbb{E}[\hat{g}_n(x)]^2 - \mathbb{E}^2[\hat{g}_n(x)] \sim \frac{1}{h_n}m(x)\int_R K2(t)dt \longrightarrow 0$ quand $n \longrightarrow \infty$

donc $\lim\limits_{n\longrightarrow\infty} \mathbb{V}ar(\hat{g}_n(x)) = 0$

Ceci implique que

$$\hat{g}_n(x) \longrightarrow g(x) \text{ en probabilité.}$$

D'où

$$\hat{r}_n(x) = \frac{\hat{g}_n(x)}{\hat{f}_n(x)} \longrightarrow \frac{f(x)}{g(x)}, \text{ en probabilité.}$$

1.6 Convergence

1.6.1 Convergence en probabilité

Définition 1.6.1. *la suite (X_n) Converge en probabilité vers X si :*

$$\forall \varepsilon > 0 \lim_{n \longrightarrow n} P\left(|X_n - X| < \varepsilon\right) = 0$$
$$X_n \xrightarrow{P} X$$

Alors, à chaque point de continuité de $f(x)$; $g_X(x)$ et $\sigma^2(x)$ avec

Proposition 1.6.1. ***{la loi des grand nombres}***
Si les variables aléatoire $X_n$sont deux à deux à deux non covariées, de même loi, d'espérance μ de variance σ^2, alors :

$$\tfrac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

1.6.2 Convergence presque complète

Pour étudier la convergence presque complète on peut utiliser les définitions suivantes :

Définition 1.6.2. *On dit que la suit $(X_n)_{n\in N}$ converge presque complètement vers X si :*

$$\forall \varepsilon > 0, \quad \sum_{n=0}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty$$

En on note $X_n \longrightarrow X$ en p.co

Définition 1.6.3. *On dit que la suite $X_n = O(Y_n)$ en p.co, s'il existe un $\varepsilon > 0$ vérifiant*

$$\sum_{n=0}^{\infty} \mathbb{P}(X_n > \varepsilon Y_n) < \infty$$

Corollaire 1.6.2. *Soient X_n , Y_n deux suites des variables aléatoire , la suite*

$$\left(\tfrac{X_n}{Y_n}\right)_{n\in N} \longrightarrow 0 \text{ en } p.co$$

Si

$$X_n \longrightarrow 0 \text{ en } p.co$$

Et

$$\exists\delta > 0, \sum_{n=0}^{\infty} (P|Y_n| < \delta) < \infty$$

Théorème 1.6.3. ***(Convergence presque complète ponctuelle sous condition de continuité [7])***
Considérons le modèle (1.1) *avec* $k = 0$ *et supposons que les conditions suivantes sont vérifiées*

1. $f_X(x) > 0$ *;*
2. $\lim\limits_{n\longrightarrow\infty} h = 0$ *et* $\lim\limits_{n\longrightarrow\infty} \frac{nh}{logn} = \infty$ *;*
3. *K bornée, intégrable et à support compact ;*
4. $|Y| < M < \infty$.

On a alors

$$\hat{f}(x) \longrightarrow f(x), \text{ presque complètement.}$$

Théorème 1.6.4. ***Convergence uniforme presque complète sous condition de continuité [7]***
Considérons le modèle de régression suivant

$$\inf_{x\in S} f_X(x) > \theta$$

supposons que les conditions suivantes sont vérifiées :

1. $\lim\limits_{n\to\infty} h = 0$, *et* $\lim\limits_{n\to\infty} \frac{nh}{logn} = +\infty$ *;*
2. *K bornée, intégrable et à support compact ;*
3. $|Y| < M < \infty$ *;*
4. $\inf_{x\in S} f_X(x) > \theta$ *;*
5. $\exists\beta > 0, \exists C > 0, \forall x \in S, \forall y \in S, \quad |k(x) - k(y)| \leq |x - y|^\beta$.

On a alors :

$$\sup_{x\in S} |\hat{f}(x) - f(x)| \longrightarrow 0 \text{ ; presque complètement.}$$

Maintenant nous allons énoncer les résultats sous l'hypothèse de type Lipschitz. Les conditions de Lipschitz sont des conditions ponctuelles en x fixé du type

$$\exists \beta > 0, \exists C < \infty, \exists \varepsilon > 0, \forall y \in]x-\varepsilon, x+\varepsilon[, \qquad |\phi(x)-\phi(y)| \leq C|x-y|^{\beta}. \tag{1.3}$$

soit des conditions uniformes sur un compact S du type

$$\exists \beta > 0, \exists C < \infty, \forall x \in S, \forall y \in S, \qquad |\phi(x)-\phi(y)| \leq C|x-y|^{\beta}. \tag{1.4}$$

où ϕ désigne indifféremment g_X ou f :

Théorème 1.6.5. ***Vitesse de convergence presque complète ponctuelle sous condition de Lipschitz [7]***
Considérons le modèle (1.4) et supposons que les conditions suivantes sont vérifiées

1. $g_X > 0$;
2. $\lim\limits_{n\to\infty} h = 0$,*et* $\lim\limits_{n\to\infty} \frac{nh}{logn} = +\infty$;
3. *K bornée, intégrable et à support compact ;*
4. $|Y| < M < \infty$.

On a alors :

$$\hat{f}(x) - f(x) = 0(h^{\beta}) + 0\left(\sqrt{\frac{\log n}{nh}}\right), \textit{ presque complètement.}$$

Théorème 1.6.6. ***Vitesse de convergence presque complète uniforme sous condition de Lipschitz [7]***
Considérons le modèle (1.1) et supposons que les conditions suivantes sont vérifies :

1. $\lim\limits_{n\to\infty} h = 0$, *et* $\lim\limits_{n\to\infty} \frac{nh}{logn} = +\infty$;
2. *K bornée, intégrable et à support compact ;*
3. $|Y| < M < \infty$;
4. $\inf_{x\in S} f_X(x) > \theta$;
5. $\exists \beta > 0, \exists C < \infty, \forall x \in S, \forall y \in S, \quad |K(x)-K(y)| \leq C|x-y|^{\beta}$

On a alors :

$$\sup_{x\in S} |\hat{f}(x) - f(x)| = 0(h)^{\beta} + 0\left(\sqrt{\frac{\log n}{nh}}\right), \textit{ presque complètement.}$$

1.6.3 Convergence en moyenne quadratique

Définition 1.6.4. *La suite* (X_n) *converge en moyenne quadratique vers* X *si :*

$$\lim_{n \longrightarrow \infty} \mathbb{E}\left(X_n - X\right)^2 = 0$$
$$X_n \overset{mq}{\longrightarrow} X$$

Proposition 1.6.7. (X_n) *converge en moyenne quadratique vers* X *si* $\lim\limits_{n \longrightarrow \infty} E(X_n) = E(X)$ *et* $\lim\limits_{n \longrightarrow \infty} Var(X_n - X) = 0$

***Erreur moyenne quadratique ponctuelle**

Théorème 1.6.8. ***Convergence en moyenne quadratique ponctuelle sous condition de continuité [7]***
Considérons le modèle (1.1) *avec* $k = 0$ *et supposons que les conditions suivantes sont vérifiées :*

1. $g_X(x) > 0$ *;*
2. $|Y| < M < \infty$ *;*
3. $\lim\limits_{n\to\infty} h = 0$, $\lim\limits_{n\to\infty} nh = \infty$ *;*
4. K *bornée, intégrable, positive, symétrique et à support compact ;*

On a alors :

$$\mathbb{E}[\hat{f}_n(x) - f(x)]^2 \underset{n\to\infty}{\longrightarrow} 0.$$

****Erreur moyenne quadratique intégrée**

[7] ont donné des versions uniformes sur un compact des deux théorèmes précédents. L'erreur moyenne quadratique intégrée est définie par :

$$MISE(\hat{f}(x)) = \mathbb{E}\left[\int \left(\hat{f}(x) - f(x)\right)^2 w(x)dx\right].$$

La fonction w est une fonction de poids vérifiant :

w est positive, bornée et à support compact S.

Cette fonction est fixée a priori, et sera souvent dans la pratique prise égale à une indicatrice sur un intervalle borné de $\mathbb{R}$, ou égale à un produit d'une telle indicatrice par la densité g_X de X [7].

Supposons aussi que

$$\mathbb{E}(Y^2|X = u) \text{ est continue autour de } S.$$

Théorème 1.6.9. ***Erreur moyenne quadratique intégrée sous condition de continuité [7]***

Considérons le modèle (1.3) avec $k = 0$ et supposons que les conditions suivantes sont vérifiées :

1. $\inf_{x\in S} f_X(x) > \theta, \theta > 0$ *;*
2. $\lim\limits_{n\to\infty} h = 0, \lim\limits_{n\to\infty} nh = \infty$ *;*
3. *K bornée, intégrable, positive, symétrique et à support compact ;*
4. *w est positive, bornée et à support compact S ;*

On a alors :

$$MISE(\hat{f}) \longrightarrow 0.$$

1.6.4 Convergence presque sûre

Définition 1.6.5. *la suite (X_n) converge presque sûrement vers X si :*

$$P\left(w/ \lim_{n\longrightarrow\infty} X_n(w) = X(w)\right) = 1.$$

Proposition 1.6.10. **{*Loi forte des grands nombres*}**
Si les variables aléatoire X_n sont mutuellement indépendantes de même loi d'espérance μ, alors

$$\tfrac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{Ps} \mu.$$

1.6.5 Converge en loi

Soit F_n la fonction de répartition de X_n et Fcelle de X

Définition 1.6.6. *La suite (X_n) Converge en loi vers X si pour tout x où F est continue,*

$$\lim_{n\longrightarrow\infty} F_n(x) = F(x)$$
$$X_n \xrightarrow{l} X$$

***Propriété de la convergence en loi**

- Si F est continue alors :$X_n \xrightarrow{\mathrm{l}} X$.
- Si X_n et X sont des variables aléatoires à valeur dans $\mathbb{N}$ alors$X_n \xrightarrow{\mathrm{l}} X$ si et seulement si : $\forall k \in \mathbb{N}, \lim\limits_{n \longrightarrow \infty} P(X_n = k) = P(X = k)$.
- Soit a et b deux réels.Si $X_n \xrightarrow{\mathrm{l}} X$ alors $aX_n + b$.

Théorème 1.6.11. **{*limite centrale*}**
Si les variables aléatoires. X_n sont mutuellement indépendant de méme loi ,d'espérance μ et d'ecart-type σ différent de 0 alors

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right) \xrightarrow{l} X$$

où X est une variable aléatoire de loi de laplace-Gauss centrée réduite.

1.6.6 Liens entre les différents type de convergence

6.1) Ils se résume de la façon suivante :
$X_n \xrightarrow{\mathrm{Pc}} X \Rightarrow X_n \xrightarrow{\mathrm{Ps}} X \Rightarrow X_n \xrightarrow{\mathrm{P}} X \Rightarrow X_n \xrightarrow{\mathrm{l}} X$
$X_n \xrightarrow{\mathrm{mq}} X \Rightarrow X_n \xrightarrow{\mathrm{P}} X \Rightarrow X_n \xrightarrow{\mathrm{l}} X$
la convergence en loi est la seule qui ne fait intervenir que les lois des variables aléatoires.

6.2) Dans le cas où X est une variables aléatoires égale à a ou presque sûrement égale à a

$$X_n \xrightarrow{\mathrm{P}} X \Leftrightarrow X_n \xrightarrow{\mathrm{l}} X$$

CHAPITRE 2

LES FONCTIONS *ALPHA* MÉLANGEANTS

La grande majorité des résultats qu'on connait en théorie des probabilités et qui sont relatifs au variables aléatoires, ne sont généralement valables que pour des variables aléatoires indépendants, et il est souvent difficile voir impossible de généraliser ces résultats au variables aléatoires dépendantes. Pour remédier à cette anomalie, des mesures de dépendance ont été créées pour savoir à quel point ces variables étaient dépendantes. Parmi ces mesures précédemment citées, il y a le mélange.

2.1 La dépendance

Dans les usages, est plus fréquent d'être exposé à des données de survie dépendantes.
une personne résidant dans une zone épargnée par un virus particulier aura moins de chance d'y être infectée ; contrairement à celle résidant dans une zone présentant plus de risques. Cependant, il est à remarquer que ces données présentent une structure de dépendance bien qu'elles soient asymptotiquement indépendantes. En effet, si l'on ordonne la provenance géographique d'individus issus d'un même échantillon, nous constaterons une dépendance entre deux individus proches géographiquement et une indépendance entre deux individus éloignés. Ceci est mathématiquement modélisé par des données mélangeantes. Il existe plusieurs formes de mélange, l'α-mélange est la forme la plus répandue et la moins restrictive car la plus faible. Ce type de dépendance est notamment utilisé dans les processus ARMA en particulier dans l'analyse des séries temporelles.

Le coefficient α est tel que $0 \leq \alpha \leq \frac{1}{4}$, il est plus faible que les autres coefficients de mélange (Voir Doukhan (1994), Rio (2000)). De multiples travaux existent dans la littérature dans le cadre de la structure de dépendance de type α-mélange pour des données tronquées, Sun et Zhou (2001) ont établi des résultats pour l'estimation de la f.d.r. Par ailleurs, Lemdani et al. (2005) se sont intéressés à la consistance de la fonction quantile pour des données tronquées à gauche. Ould-Saïd et Sadki (2008) ont établi la convergence presque sûre de l'estimateur à noyau du quantile conditionnel pour un modèle censuré à droite. D'autres travaux récents existent, il s'agit de ceux de Liang et de Uña -Álvarez (2011) qui donnent la convergence uniforme ainsi que la normalité asymptotique de l'estimateur du quantile conditionnel dans le cas α-mélangeant pour des données censurées.

Définition 2.1.1. ***(L'association)***
est une forme particulière de dépendance positive, elle a deux origines : la physique mathématique avec les inégalités FKG (Fortuin, Kasteleyn et Ginibre) appliquées en théorie de percolation, . . . et l'autre statistique (Esary, Proschan et Walkup) avec des applications en fiabilité.Pour illustrer ce type de dépendance.

Définition 2.1.2. *L'association négative a été introduite par [10] .Les v.a $X_1, ..., X_n$ sont dites négativement associées si pour tous sous-ensembles disjoints I et $J \in \{1...n\}$*

$$\mathbb{C}ov(\Psi_1(X_i, i \in I), \Psi_2(Y_j, j \in J)) \leq 0,$$

pour toutes fonctions non décroissantes Ψ_1 de $\mathbb{R}^I \longrightarrow \mathbb{R}$, Ψ_2 de $\mathbb{R}^J \longrightarrow \mathbb{R}$ pour lesquelles la covariance existe. Une suite infinie de variables aléatoires est négativement associée si toute sous suite finie est négativement associée. Pour ce type d'association, nous référons le lecteur aux travaux de Bozorgnia et al. (1996), *Patterson et Taylor* (1997).
Parmi les distributions multivariées, les lois gaussiennes négativement corrélées, les lois multinomiales et les lois de Dirichlet possèdent des propriétés de l'association négative.
Les variables associées et négativement associées vérifient l'équivalence entre la non corrélation et l'indépendance. Cette propriété provient de l'inégalité de covariance suivante

$$|\mathbb{C}ov(\Psi_1(X_i, i \in I), \Psi_2(Y_j, j \in J))| \leq \sum_{i \in I} \sum_{j \in J} \| \frac{\partial \Psi_1}{\partial x_i} \|_\infty \| \frac{\partial \Psi_2}{\partial x_j} \|_\infty |\mathbb{C}OV(X_i, X_j)| \tag{2.1}$$

pour I et $J \in \{1, ..., n\}$ et Ψ_1, Ψ_2 deux fonctions réelles à dérivées bornées.

Remarque 2.1.1.

• *L'inégalité* $(2,1)$ *a été établie par Birkel* (1988) *pour les variables associées, par Bulinski* (1996) *pour les variables négativement associées et par Doukhan et Louhichi* (1999) *pour les variables gaussiennes.*

Remarque 2.1.2.

• *Les v.a.* $X_1, ..., X_n$ *sont dites positivement (faiblement) associées si pour tous sous-ensembles disjoints* I *et* $J \in \{1, ..., n\}$

$$\mathbb{C}ov(\Psi_1(X_i, i \in I), \Psi_2(Y_j, j \in J)) \leq 0,$$

pour toutes fonctions non décroissantes Ψ_1 *de* $\mathbb{R}^I \longrightarrow \mathbb{R}$, Ψ_2 *de* $\mathbb{R}^j \longrightarrow \mathbb{R}$ *pour lesquelles la covariance existe. Une suite infinie de variables aléatoires est positivement associée si toute sous suite finie est positivement associée. Remarquons que la définition de l'association positive est très proche de l'association. Néanmoins, cette dernière est plus forte et implique l'association positive.*

Propriété 2.1.1. *On dire que* X *et* Y *sont associées implique que* $\mathbb{C}ov(\Psi_1(X), \Psi_2(Y)) \geq 0$, *ce qui implique que* $\mathbb{C}ov(X; Y) \geq 0$ *pour toutes fonctions non décroissantes* Ψ_1, Ψ_2.
Notons également qu'un vecteur $(X_1, X_2, ..., X_n)$ *de v.a. gaussiennes est associé si est seulement s'il était positivement corrélé (i.e.*$\mathbb{C}ov(X_i, X_j) \geq 0$*).*

Cette proposition a été démontrée par Pitt (1982).

Définition 2.1.3. *:(**la quasi-association**)*
on a une famille de v.a. réelles $X = \{X_n, n \in \mathbb{N}\}$ *telle que* $\mathbb{E}(X_j^2) < \infty$ *pour tout* $j \in \mathbb{N}$. *La famille* X *est dite quasi-associée si pour toutes fonctions lipschitziennes* $\Psi 1, \Psi 2$ *définies respectivement de* $\mathbb{R}^I \longrightarrow \mathbb{R}$ *et* $\mathbb{R}^J \longrightarrow \mathbb{R}$ *avec* I *et* J *deux sous ensembles finis disjoints de* $\mathbb{R}$, *on a*

$$|\mathbb{C}ov(\Psi_1(X_i, i \in I), \Psi_2(X_j, j \in J)| \leq \sum_{i \in I} \sum_{j \in J} lip_i(\Psi_1) lip_j(\Psi_2) |\mathbb{C}ov(X_i, Y_j) \tag{2.2}$$

Les constantes de Lipschitz $Lip_i(\Psi_1)$ *sont telles que pour* $x = (x_i, i \in I)$,$y = (y_j, j \in J)$ *dans* $\mathbb{R}^{|I|}$

$$|\Psi_1(x) - \Psi_1(y)| \leq \sum_{i \in I} lip_i \Psi_1 |x_i - y_j|$$

et

$$Lip_i = \sup_{x_i \neq x_i'} \frac{|\Psi(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_I) - \Psi(x_1, ..., x_{i-1}, x_i', x_{i+1}, ..., x_I)|}{|x_i - x'i|}$$

le sup étant pris pour $x_1, x_2, ..., x_{|I|}, x_i' \in \mathbb{R}$.

Théorème 2.1.2. ***Bulinski et Shashkin*** (2007)
Soit $\{X_t, t \in \mathbb{Z}\}$ une suite de v.a. associées, positivement associées ou négativement associées, telles que $\mathbb{E}(X_t^2) < \infty$ pour tout $t \in \mathbb{Z}$ soit I et J deux sous-ensembles finis de $\mathbb{Z}$. Dans l'association positive ou négative . I et J sont supposés disjoints. Pour toutes fonctions lipschitziennes $f : \mathbb{R}^{|I|} \longrightarrow \mathbb{R}$ et $g : \mathbb{R}^{|J|} \longrightarrow \mathbb{R}$, nous avons

$$|\mathbb{C}ov(f(X_i, i \in I), g(X_j, j \in J))| \leq \sum_{i \in I} \sum_{j \in j} Lip_i(f) Lip_j(g) |\mathbb{C}ov(X_i, X_j)|.$$

◦ *Pour la faible association, Doukhan et Louhichi* (1999) *choisissent de traiter la dépendance faible par une approche unifiée :*

Définition 2.1.4. *(**Doukhan et Louhichi** (1999))*
Si $\{X_t, t \in \mathbb{Z}\}$ est une suite de v.a.(γ, φ)-faiblement dépendante, s'il existe une suite $(\gamma_r)_{r \in \mathbb{N}}$ décroissante vers zéro à l'infini, et une fonction $\varphi : \mathbb{R}_+^2 \times \mathbb{N}^2 \longrightarrow \mathbb{R}$, telles que pour tous (u)-uplets $(i_1; , ..., i_u)$ et (v)-uplets $(j_1; , ..., j_v)$ avec

$$i_1 \leq .... \leq i_u < i_u + r \leq j_1 \leq ...... \leq j_v,$$

Alors

$$|\mathbb{C}ov(h(X_{i1}, ..., X_{iu}), k(X_{j1}, ..., X_{jv})| \leq \varphi(Lip(h), Lip(k), u, v)$$

pour toutes fonctions $h, k \in \mathcal{L}_1$ définies respectivement sur $\mathbb{R}^u$ et $\mathcal{R}^v$ ($\mathcal{L}$ étant l'ensemble de toutes les fonctions lipschitziennes définies sur $\mathbb{R}^n$, $n \in \mathbb{N}^*$ et $\mathcal{L}_1$) le sous-ensemble défini par $\mathcal{L}_1 = \{h \in \mathcal{L}; \|h\|_\infty \leq 1\}$.
$Lip(h)$ désigne le module de continuité lipschitzien de h, et

$$Lip(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{||x - y||_1}$$

avec $||x||_1 = |x|_1 + |x|_2 + ... + |x|_u$ pour $x = (x_1, ..., x_u)$.

2.2 Inégalités exponentielles

De multiples inégalités exponentielles existent dans l'estimation non paramétrique fonctionnelle, elles permettent de déterminer les vitesses de convergence des différents estimateurs. La construction des inégalités exponentielles s'inspire principalement du théorème central limite et de la loi forte des grands nombres. Cette construction requiert l'utilisation de l'inégalité exponentielle de Markov, il s'agit de définir une borne supérieure idéale de la fonction génératrice des moments.

Théorème 2.2.1. ***(Doukhan et Neumann (2007))***
Soient $X_1, ..., X_n$ des v.a. réelles de moyenne nulle définies sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et $\Psi : \mathbb{N}^2 \longrightarrow \mathbb{N}$ définie par l'une des fonctions suivantes :

-i- $\Psi(u, v) = 2v$ *;*

-ii- $\Psi(u, v) = u + v$ *;*

-iii- $\Psi(u, v) = uv$;

-iv- $\Psi(u, v) = a(u + v) + (1 - a)uv \quad \forall a \in (0, 1)$.

Soient les constantes $K, M, L_1, L_2 < \infty$,μ ,$\nu \geq 0$, et u ,$v \geq 0$ On suppose une suite non décroissante de coefficients réels $(\Phi(n))_{n\geq 0}$telle que pour tous u-uplets $(s_1, ..., s_u)$ et v-uplets $(t_1, ...t_v)$ avec $1 \leq s_1 \leq .... \leq s_u \leq t_1 \leq .... \leq t_v \leq n$ l'inégalité suivante est vérifiée :

$$|cov(X_{s1}, ..., X_{su}, X_1, ..., X_v)| \leq K^2 M^{u+v-2}((u + v)!)^{\nu} \Psi(u, v)\Phi(t_1 - s_u)$$

avec

$$\sum_{i=0}^{\infty} (s + 1)^k (\Phi(s)) \leq L_1 L_2^k (k!)^{\mu}, \forall k \geq 0$$

et

$$\mathbb{E}(|X_t|^k) \leq M^k (k!)^{\nu}, \forall k \geq 0$$

Alor,$\forall k \geq 0$,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i\right) \geq \exp\left(-\frac{t^2/2}{A_n + B_n^{1/(\mu+\nu+2)} t^{(2\mu+2\nu+3)(\mu+\nu+2)}}\right)$$

où $A_n \geq \sigma_n^2$ *avec* $\sigma_n^2 := Var\left(\sum_{i=1}^{n} X_i\right)$ *et*
$B_n = 2(K \vee M)L_2 \left(\frac{2^{4+\mu+\nu} n k^2 L_1}{A_n} \vee 1\right)$.

2.3 Éléments sur les séries chronologiques

2.3.1 Séries chronologiques

Une série temporelle ou série chronologique est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps . généralement pour comprendre son évolution passée et pour en prévoir le comportement futur . Une telle transposition mathématique utilisé le plus souvent des concepts de probabilité et de statistique [20].
Nous considérons une suite d'observation d'une famille de variables aléatoires réelles , nous la noterons

$$(X_t, t \in T) \text{ ou } (X_t)_{t\in T}$$

Où l'ensemble T est appelé espace des temps qui peut être discret ou continue. Dans l'étude d'une série chronologique, il est naturel de penser que la valeur de la série à la date t peut dépendre des valeurs prises aux dates dates précédentes :

$$X_t = f(X_{t-1}, X_{t-2}, ...)$$

Il n'est généralement pas nécessaire de prendre en compte tout le passé de la série et on peut le plus souvent se limiter à p valeurs :

$$X_t = f(X_{t-1}, X_{t-2}, ..., Xt - p)$$

2.3.2 Processus stochastiques

Un processus stochastique X est une famille de variables aléatoires réelles $(X_t)_{t\in T}$ où $T \subset \mathbb{R}$
Si $T \subset \mathbb{Z}$, le est dit à temps discret,
Si T est un intervalle de $\mathbb{R}$, le processus est dit à temps continu.
t est définies sur un même espace de probabilité (Ω, T, P)

Remarque 2.3.1.
Une séries chronologique est l'observation à n instants d'un processus stochastique

Opérateur de retard

On appelle opérateur retard L l'opérateur qui décale le processus d'une unité de temps vers passé :

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

Plus généralement,si on applique h fois l'opérateur de retard$L$ on aura :

$$L^h(X_t) = X_{t-h}, \quad t \in T \ , \ \ h \in N$$

C'est à dire le processus est décalé de h unité de temps. [11]

2.3.3 Fonctions d'autocovariance et d'autocorrélations

une manière de comprendre les liens d'interdépendance entre les termes d'un processus $(X_t)_{t\in T}$ est de considérer les fonctions d'autocovariance et d'autocorrélation qui mesurent respectivement la covariance et la corrélation entre les termes d'une série chronologique.

Définition 2.3.1. *Soit $(X_t)_{t\in T}$ un processus stochastique.*
On appelle fonction d'autocovariance la fonction γ définie de T dans $\mathbb{R}$ par

$$\forall t_1, t_2 \in T \quad \gamma(t_2 - t_1) = cov(X_{t_1}, X_{t_2})$$

Dans le cas stationnaire :
On appelle fonction d'autocovariance (resp. d'autocorrélation) la fonction γ (resp.ρ) , qui dépend uniquement de h, définie de T dans $\mathbb{R}$ par

$$\forall h, t \in T, \quad \gamma(h) := cov(X_t, X_{t+h})$$

$$(resp. \ \ \rho(h) := \frac{\gamma(h)}{\gamma(0)})$$

Avec la covaiance entre les variables X_t et X_{t+h}est donnée par :

$$cov(X_t, X_{t+h}) = \mathbb{E}[(x_t - E(X_t))(X_{t-h} - E(X_{t-h}))]$$

Le graphe de la fonction d'autocovariance est appelé *variogramme* tandis que celui de la fonction d'autocorrélation est appelé *correlograme.*

Proposition 2.3.1. *On a*

1. $\forall h \in T, \gamma(-h) = \gamma(h)$;
2. $\gamma(0) = var(X_t)$;
3. $|\gamma(h)| \leq \gamma(0), \forall h.$

2.3.4 Les processus linéaires

Les processus linéaire sont des combinaisons linéaires de bruit blanc, ils constituent alors le modèle le plus simple, l'intérêt de ses processus revient au comportement de leur fonction d'autocovariance, qui tend vers zéro lorsque l'on compare deux variable éloignées dans le temps :

$$\gamma(h) = Cov(X_t, X_{t+h}) \underset{h\to 0}{\longrightarrow} 0$$

Cella signifie que le mémoire du processus est contrôlée (la liaison du second ordre est plus importante pour deux variables proches dans le temps que pour des variables éloignées). Un processus$(X_t)_{t\in T}$ est dit linéaire μ s'il peut s'écrire sous la forme

$$X_t = \mu + \sum_{i=-\infty}^{+\infty} b_i \varepsilon_{t-i},$$

Où $(\varepsilon_t)_{t\in T}$ est un bruit blanc, de variance σ^2 et où b_i est telle que

$$\sum_{i=-\infty}^{+\infty} b_i < +\infty$$

Stationnarité d'un processus

Un processus est dit stationnaire si ses propriétés stochastique caractérisées par des espérances mathématique sont indépendantes du temps

Définition 2.3.2. *D'une manière générale un processus $X = (X_t)_{t\in T}$ est stationnaire si et seulement si*

-i- $E(X_t) = \mu, \forall \in T$

-ii- X_t *est de carré intégrable pour tout* $t \in T : E(X_t^2) < \infty$

-iii- $cov(X_s, X_{s+t}) = cov(X_{s-1}, X_{s-1+t}) = ... = cov(X_0, X_t), \quad \forall t, s \in T$

Processus inversible

Définition 2.3.3. *Un Processus $X = (X_t)_{t\in T}$ admet une représentation inversible s'il peut s'écrire comme combinaison linéaire des valeurs d'un autre processus, c'est à dire qu'il existe une suite $(\psi_i, i \in T)$ et un processus $(Y_t)_{t\in\mathbb{Z}}$ tels que :*

$$\forall t \in T \ , \ X_t = \sum_{i\in T} \psi_i Y_{t-i}$$

Processus causal

Définition 2.3.4. *Un processus $X = (X_t)_{t\in T}$ admet une représentation causale s'il peut s'écrire comme combinaison linéaire des valeurs passés d'un autre processus, c'est à dire qu'il existe une suite $(\varphi_i, i \in T)$ et un processus $(X_t)_{t\in Z}$ tels que :*

$$\forall t \in T, \quad Y_t = \sum_{i\in N} \varphi_i X_{t-i}$$

Processus autorégressifs AR et Moyen mobile MA

Définition 2.3.5. *On appelle processus moyenne mobile d'ordre q , noté $MA(q)$ ou $\{$Mouving Average$\}$ un processus $(X_t)_{t\in T}$ défini par :*

$$X_t = \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}, \quad \forall t \in T$$

Où les θ_i sont des réels ,$\theta_q \neq 0$ et $(\varepsilon_t)_{t\in T}$ est un processus bruit blanc de variance σ^2. [5]

Définition 2.3.6. *Soit $(\varepsilon_t)_{t\in T}$ un bruit blanc de variance σ^2. On dit qu'un processus $(X_t)t \in T$est un processus autorégressif ou encore processus AR d'ordre p,noté AR(p) [1], si*

$$\forall t \in T \quad X_t = \sum_{i=1}^{p} \phi_i Xt - i + \varepsilon_t$$

Où $(\phi_1, ..., \phi_p) \in \phi_p \neq 0$

Notation

On utilise généralement la notion suivante :

$$\Phi(L)X_t = \varepsilon_t$$

Ou

$$\Phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i,$$

L^i est un opérateur de retard d'ordre i

Exemple 2.1. *soit $(\varepsilon_t)_{t\in T}$ un bruit blanc de variance σ^2.*
Un modèle autorégressif d'ordre 1 noté $AR(1)$, prend une seule valeur du passé, il est donné par :

$$X_t = \phi_(1)X_{t-1} + \varepsilon_t$$

Où

$$(1 - \phi L)X_t = \varepsilon_t$$

avec ϕ_1 *est un paramètre réel du modèle, tel que* $\phi_1 \neq 0$ *De façon générale, un processus* $AR(p)$ *est un processus qui dépend linéairement des p valeur antérieures.[11]*

2.4 Variables mélangeantes

En exploitant les diverses caractérisations de l'indépendance, on peut construire plusieurs formes de dépendance, parmi lesquelles la notion de mélange(qui est un phénomène de dépendance faible) sous ses divers aspects, bénéficie d'un intérêt particulier.
Les différentes notions de mélanges sont reliées à des mesures de dépendances entre sous σ-algèbres, variables aléatoires ou processus . Plus précisément soit $(\Omega, \mathcal{F}, P)$ un espace probabilité et $\mathcal{A}$ et $\mathcal{B}$ deux sous σ-algèbres de $\mathcal{F}$, diverses mesures de dépendance enter $\mathcal{A}$ et $\mathcal{B}$ ont été définis . Nous allons d'abord, définir les différents mélanges entres σ-algèbres.

2.4.1 Coefficients de dépendance

Les différentes mesures de dépendance entre deux σ-algèbres $\mathcal{A}$ et $\mathcal{B}$ sont définies comme suit [23] :

1. α **mélange** ou **mélange fort**

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|P(A)P(B) - P(A \cap B)|, \quad A \in \mathcal{A} \ \ et \ \ B \in \mathcal{B}\} \tag{2.3}$$

le coefficient α est le coefficient de mélange fort introduit par Rosenblatt [22].

2. β **mélange**

$$\beta(\mathcal{A}, \mathcal{B}) = E \sup\{|P(B/A) - P(B)|, B \in \mathcal{B}\} \tag{2.4}$$

Le coefficient β est le coefficient de régularisation absolue introduit par Kolmogorov et Rosanov [16].Il s'écrié sous la forme

$$\beta(\mathcal{A}, \mathcal{B}) = \sup\{\frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}|P(A_i)P(B_i) - P(A_i \cap B_j)|, A_i \in \mathcal{A} \ \ B_j \in \mathcal{B}\} \tag{2.5}$$

3. ϕ **mélange**

$$\phi(\mathcal{A}, \mathcal{B}) = \sup\{|P(B) - \frac{P(A \cap B)}{P(A)}|, A \in \mathcal{A}, P(A) \neq 0 \ \ et \ \ B \in \mathcal{B}\} \tag{2.6}$$

Le coefficient ϕ est le coefficient de mélange uniforme introduit par Ibragimov [13]. Il s'écrié sous la forme suivante

$$\phi(\mathcal{A}, \mathcal{B}) = \sup\{|P(B/A) - P(B)|, B \in \mathcal{B}\} \tag{2.7}$$

4. $\psi(\mathcal{A}, \mathcal{B})$ **mélange**

$$\sup\{|1 - \frac{P(A \cap B)}{P(A)P(b)}|, A \in \mathcal{A}, P(A) \neq 0, B \in \mathcal{B} \ \ et \ \ P(B) \neq 0\} \tag{2.8}$$

Le coefficient ψest introduit par Blum, Hanson et Koopmans [4].Il s'écrié sous la forme suivante

$$\psi(\mathcal{A}, \mathcal{B}) = \sup\{\frac{1}{P(B)}|P(B/A) - P(B)|, B \in \mathcal{B} \ \ et \ \ P(B) \neq 0\} \tag{2.9}$$

5. ρ **mélange**

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{|corr(X, Y) - 1|, X \in L^2(\mathcal{A}) \ \ et \ \ Y \in L2(\mathcal{B})\} \tag{2.10}$$

Le coefficient ρ est le coefficient des corrélation maximales introduit par [12] et [15]

Remarque 2.4.1. *Les coefficients de mélange précédents sont définie par les inégalités suivantes :*

$$0 \leq \alpha(\mathcal{A}, \mathcal{B}) \leq \tfrac{1}{4}$$
$$0 \leq \phi(\mathcal{A}, \mathcal{B}) \leq 1$$
$$0 \leq \psi(\mathcal{A}, \mathcal{B}) \leq \infty$$
$$0 \leq \rho(\mathcal{A}, \mathcal{B}) \leq 1$$

De plus, si l'un d'eux est nulle, les α-algébres $\mathcal{A}$ et $\mathcal{B}$ sont indépendants.

Proposition 2.4.1. *Les coefficient de dépendance vérifient les inégalités suivantes :*
- $4\alpha(\mathcal{A}, \mathcal{B}) \leq 2\sqrt{\phi(\mathcal{A}, \mathcal{B})\phi(\mathcal{A}, \mathcal{B})} \leq \psi(\mathcal{A}, \mathcal{B})$
- $2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B} \leq \frac{1}{2}\psi(\mathcal{A}, \mathcal{B})$

Preuve 2.4.1. *Pour les démonstration on peut se référe a [6]*

2.4.2 Inégalités de covariance

Soient $\mathcal{A}$ et $\mathcal{B}$ deux tribus. Soient X une variable $\mathcal{A}$-mesurable et Y une variable $\mathcal{B}$-mesurable, alors [6]

- $|cov(X,Y)| \leq 8\alpha^{\frac{1}{r}}(\mathcal{A},\mathcal{B})\|X\|_p\|Y\|_q$ pour p,q et $r>1$ $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$
- $|cov(X,Y)| \leq 2\phi^{\frac{1}{2}}(\mathcal{A},\mathcal{B})\|X\|_p\|Y\|_q$ pour p,q et $r>1$ $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$
- $|cov(X,Y)| \leq \psi(\mathcal{A},\mathcal{B})\|X\|_1\|Y\|_1$
- $|cov(X,Y)| \leq \rho(\mathcal{A},\mathcal{B})\|X\|_2\|Y\|_2$

Avec $\|X\|_p = (E|X|^p)^{\frac{1}{p}}$

2.4.3 Champ mélangeant

Définition 2.4.1. *Un champ aléatoire est une suite de variable aléatoire (X_i) indexée par Z^d.Pour tout $I \subset Z^d$, on notera X_I pour la suite $(X_i)_{i\in I}$*
soit $X=(X_t)_{t\in T}$ un champ aléatoire, où T est un espace métrique. On note par $X_c=\{X_t, t\in C\}$ la C-martingale de X, et par χ_C la σ-algèbre engendrée par X_C pour $C\subset T$, $|C|$ représente le cardinale de C s'il est fini [5]. Et A et B deux sous ensemble de C.
On note par c le coefficient de mélange α,β,ϕ,ρ ou ψ

$$c_X(A,B) = c(\chi_A,\chi_B)$$

Et pour des entiers u et v ,

$$c_X(k,u,v) = \sup\{|c_X(A,B)|, d(A,B)\geq k, |A|<u \quad et \quad |B|<v\}$$

Définition 2.4.2. *On dit qu'un champ aléatoire $X=(X_t)_{t\in T}$ est c-mélangeant pour $c=\alpha,\beta,\phi,\rho$ ou ψ si $\lim\limits_{k\to\infty} c_X(k,u,v)=0$ pour u et v des entiers positifs ou nuls.*

Théorème 2.4.2. *si $X=(X_t)_{t\in^d_Z}$ est un champ aléatoire strictement stationnaire pour $d\neq 1$, alors*

$$\alpha_X(k,\infty,\infty)\leq \rho_X(k,\infty,\infty)\leq 2\pi\alpha_X(k,\infty,\infty)$$

2.4.4 Processus mélangeant

Si T est ordonné, exp $T\subset\mathbb{R}$ une autre notion de mélange peut être définit pour le processus $X=(X_t)_{t\in T}$ [6].soit

$$c_{X,k,u,v} = \sup_{A,B}\{c_X(A,B)\}$$

avec

$$|A| \leq u \; et \; \; |B| \leq v \; \; et \; \; a < b+k \; \; si \; \; a \in A \; , \; \; b \in B$$

le processus X est dit c-mélangent si

$$\lim_{k \to \infty} c_{X,k,u,v} = 0, \forall u, v \geq 0$$

Avec $c = \alpha, \beta, \phi, \rho$ ou ψ
En générale, le coefficient de mélange ne dépend pas de u ou v, on peut alors écrire $c_{x,k} = \sup\{c_{X,k,u,v}; u, v \geq 0\}$

Remarque 2.4.2. *[5]*

- *une suite aléatoire β-mélangeante est un champ α-mélangeant.*
- *une suite aléatoire ψ-mélangeante est un champ ϕ-mélangeant.*

Suites de mélange fort

Dans cette section, nous allons nous intéressé au mélange fort ou α-mélange de suites de variables aléatoires.

Définition 2.4.3. *une suite de variables aléatoires$(X_k)_{k \in Z}$ est dite α-mélangeante si*

$$\lim_{k \longrightarrow \infty} \alpha_k = 0$$

avec

$$\alpha(k) = \sup_{A \in Q_{\infty}^{u}, B \in Q_{u+k}^{+\infty}} |P(A \cap B) - P(A)P(B)|$$

On distingue principalement deux sous-classes de variables aléatoires α-mélangeantes [23].

Définition 2.4.4. *Une suite de variable aléatoires $(X_k)_{k \in Z}$ est dite arithmétiquement (ou algébriquement) α-mélangeante avec un taux a si*

$$\exists C > 0, \alpha_k \leq Ck^{-a} \tag{2.11}$$

Elle est dite *géométriquement* α-mélangeante si

$$\exists C > 0, \exists k \in [0,1], \alpha_k \leq Ck^{-a} \tag{2.12}$$

Proposition 2.4.3. *[23] On suppose que Ω' est un espace semi normé, muni de la semi norme $\| \ \|$ on a*

-i- $(\varepsilon_k)_{k\in Z}$ *est α-mélangeante* $\Longrightarrow$ $(\|\varepsilon_k\|)_{k\in Z}$ *mélangeante*

-ii- *si en plus $(\varepsilon_k)_{k\in Z}$ est géométriquement α-mélangeante (respectivement arithmétiquement) α-mélangeante alors $(\|\varepsilon_k\|)_{k\in Z}$ est géométriquement α-mélangeante (respectivement arithmétiqueme α-mélangeante) de même ordre*

Remarque 2.4.3. *[23] si $(X_n)_{n\geqslant 1}$ est une suite α-mélangeante alors$(f(X_n))_{n\geq 1}$ est aussi α-mélangeante pour tout fonction mesurable f.*

Remarque 2.4.4. *L'inégalité Davydov modifiée pour le cas mélangeant [6] est donnée par : pour $i \neq j$*

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

où c : une constante

Proposition 2.4.4. *[9] Nous avons les inégalités suivantes entre les coefficients de mélange : pour $k > 0$*

1.
$$2\alpha(k) \leq \beta(k) \leq \phi(k)$$

2. *Dans le cas d'un processus réel nous avons :*
$$2\alpha(k) \leq \rho(k) \leq 2\sqrt{\phi(k)}$$

3. *Dans le cas d'un processus réel gaussien stationnaire on montre que*
$$\rho(k) \leq 2\pi\alpha(k)$$

   Ainsi un processus réel stationnaire satisfait la condition de mélange fort si seulement s'il est complètement régulier.

Remarque 2.4.5. *On a les implication suivantes :*
ψ-mélangeant $\Rightarrow$ ϕ-mélangeant$\Rightarrow$ $\{\beta$-mélangeant, ρ-mélangeant$\}$$\Rightarrow$ α-mélangeant

CHAPITRE 3

GRAND DÉVIATION DES ESTIMATIONS DANS LES PROCESSUS

3.1 Introduction

La déviation des estimateurs est une étude très moderne est nécessaire pour évaluer notre estimateur
Dans ce chapitre, nous considérons une équation à opérateur de convolution de la forme

$$Y = A\theta(Z) + \varepsilon \tag{3.1}$$

Le but est d'estimer la fonctionnelle θ, lorsque la variable Z est contaminée d'erreurs de mesure, et de trouver une déviation pour notre estimateur .

3.2 Notation et hypothèses

soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $A : \mathbb{H} \longrightarrow \mathbb{H}$ un opérateur de convolution, linéaire compact donné par

$$A\theta_{(s)} = \int_{[0,1]} \psi(s-t)\theta(t)dt$$

où ψ est un fonction densité connue. On considère un échantillon de taille $n, (Y_1, Z_1), ..., (Y_n, Z_n)$ du couple de variables (Y, Z) vérifiant l'équation (3.1), supposées $\alpha-$ mélangeantes. On supposes que la suit des erreurs de régression $(\varepsilon_k)_{k\geq 1}$ sont des variables aléatoires identiquement distribuées de moyenne nulle et de variance σ^2 finie. Nous supposons que pour tout $k = 1, ..., n, z_k$ est telle que $X_k = z_k + \delta_k$ où δ_k est une erreur de contamination. ψ représente la fonction densité de $-\delta$. On note par $\varphi_(h)$ la transformée de Fourier de la fonction h. Dans le modèle de régression de déconvolution non paramétrique, la fonction θ est souvent supposée périodique [17], A est un opérateur de convolution sur $[0, 1]$ avec la fonction ψ périodique. Le problème $(3, 1)$ est très proche du problème de déconvolution de la densité [9].

3.3 Position du problème

Considérons le modèle donné par l'équation $(3, 1)$, où l'opérateur A de convolution avec la fonction densité ψ supposé connue. Le problème considéré est équivalent à

$$Y_k = \theta(X_k) + \varepsilon_k \quad et \quad X_k = Z_k + \delta_k \tag{3.2}$$

On suppose que les variables aléatoires δ_k,z_k et ε_k sont indépendantes entre elles. L'erreur de régression ε_k vérifie $\mathbb{E}(\varepsilon_k/X_k) = 0$. La fonction de régression θ est telle que $\theta_x = \mathbb{E}(Y_k/X_K = x)$ reliée à la fonction de régression observée $g(z) = \mathbb{E}(Y_k/Z_k = Z)$ par $g(z) = \theta * \psi(z)$. Où ψ est la fonction densité de $-\delta$. [3]

On pose $Z_k = z_k$

3.4 Construction de l'estimateur

Le modèle considéré est un modèle de régression non paramétrique classique avec déconvolution [17]. Nous proposons un estimateur de θ de type noyau, nous supposons que la transformée de fourier de la fonction densité ψ est tel que $\varphi_\psi(\omega) \neq 0$ pour tout $\omega \in \mathbb{R}$ et la transformée de Fourier du noyau K est à support compact. L'estimateur noyau de déconvolution de θ est donné par

$$\hat{\theta}_n(x) = \frac{1}{2\pi} \int\limits_R \exp(-iwx)\varphi_k(hw)\frac{\hat{\varphi}_g(w)}{\varphi_\psi(w)}dw \tag{3.3}$$

où h est un paramètre de lissage et$\hat{\varphi}_g(.)$ est la transformée de Fourier empirique de g donnée par

$$\hat{\varphi}_g(\omega) = \frac{1}{na_n}\sum_{k=1}^{n} Y_k \exp(i\omega z_k)$$

On a $g = \theta * \psi$ et son estimateur

$$\hat{g}_n(x) = \hat{\theta}_n * \psi(x)$$

Il est clair que

$$\hat{g}_n(x) = \frac{1}{nha_n}\sum_{k=1}^{n} Y_k k(\frac{x-z_k}{h}) \tag{3.4}$$

Dans ce cas, on peut facilement voir que l'estimateur $\hat{\theta}_n(x)$ de $\theta_(x)$ donné par (3.3) s'écrit sous la forme noyau comme

$$\hat{\theta}_n(x) = \frac{1}{nha_n}\sum_{k=1}^{n} Y_k k\left(\frac{x-z_k}{h}, h\right)$$

Où le noyau K est donné par

$$K(x,h) = \frac{1}{2\pi}\int\limits_R \exp(-iwx)\frac{\varphi_k(w)}{\varphi_\psi(\frac{w}{h})}dw \tag{3.5}$$

A partir de l'estimation de la fonction densité de déconvolution, il est bien connu que le taux optimal pour le quel on estime θ dépaend du lissage de θ et de ψ ou de maniere équivalente, sur les propriétés de la transformée de Fourier.
Supposons que

$$\varphi_\psi(\omega)\omega^\beta \to C_\varepsilon, \qquad \omega \to \infty \tag{3.6}$$

pour un certain $\beta \geq 0$ et $C_\varepsilon \in \mathbb{C}\backslash\{0\}$.
Notons que ceci implique que

$$\varphi_\psi(\omega)|\omega|^\beta \to \overline{C_\varepsilon}, \qquad \omega \to -\infty$$

Sous l'hypothèse (3.5) et par le théorème de convergence domine [9] on a une forme asymptotique du dé convolution (3.4) assez simple.

$$\begin{aligned}
h^\beta K(x,h) &= \frac{h^\beta}{2\pi}\int_0^\infty \exp(-iwx)\frac{\varphi_k(w)}{\varphi_\psi(\frac{\omega}{h})}d\omega + \frac{h^\beta}{2\pi}h^\beta K(x,h) + \frac{h^\beta}{2\pi}\int_{-\infty}^0 \exp(-iwx)\frac{\varphi_k(w)}{\varphi_\psi(\frac{\omega}{h})}d\omega \\
&= \frac{1}{2\pi}\int_0^\infty \exp(-i\omega x)\omega^\beta \frac{\varphi_k(\omega)}{(\frac{\omega}{h})^\beta \varphi_\psi(\frac{\omega}{h})}d\omega + \frac{1}{2\pi}\int_{-\infty}^0 \exp(-i\omega x)|\omega|^\beta \frac{\varphi_k(w)}{|\frac{\omega^\beta}{h}|\varphi_\psi(\frac{\omega}{h})}d\omega \\
&= \frac{1}{2\pi C_\varepsilon}\int_0^\infty \exp(-i\omega x)\omega^\beta \varphi_k(w)d\omega \int_{-\infty}^0 \exp(-i\omega x)|\omega|^\beta \varphi_k(\omega)d\omega \\
&= k(x)
\end{aligned}$$

Il est clair que $K(x) \in \mathbb{R}$. Ce résultat sera utilisé dans le calcul de la variance de $\hat{\theta}_n(x)$

Hypothèses

H1 :La transformée de Fourier φ_k de K est symétrique.
H2 :$\int_R K(z,h)|z|^{\frac{3}{2}}(loglog^+|z|)^{\frac{1}{2}}dz = o(h^\beta)$avec$loglog^+|z| = 0si|z| < e$ et $loglog^+|z| = loglog|z|$ ailleurs.
H3 :θ est j fois différentiable et $h^j = o\left(\frac{1}{\sqrt{n}h^{\beta+\frac{1}{2}}}\right)$ équivalent à $h = o\left(n^{-\frac{1}{2\beta+2j+1}}\right)$
H4 :$(Y_i)i \geq 1$est α-mélangeante.
H5 :$\exists M > 0$, tel que$|\epsilon K(x,h)| < M$.
H6 :$max\{\mathbb{E}(|Y_l Y_l|/K_k X_l)\} \leq C\ a.s.$
H7 :$\exists (u_n) \in \mathbb{N}^{\mathbb{N}}, o(u_n) + o\left(n[\alpha(u_n)]^{\frac{p-2}{p}}\right) \longrightarrow 0$ tq $\quad n \longrightarrow \infty$.
Le lemme suivant permet de donner l'expression asymptotique de la variance de $\hat{\theta_n}(x)$, dans le cas des variables fortement mélangeantes.

3.5 Déviation de l'estimateur

Pour étudier la déviation des estimateurs, et en plus pour vérifier la convergence presque complète. il faut tout abord introduire la notion de l'inégalité exponentielle pour les variables mélangeants.

3.5.1 Inégalité exponentielle

Nous allons établir deux inégalités exponentielle de la probabilité de la déviation de l'estimateur $\hat{\theta_n}(x)$ donné par $(3,2)$ de $\theta(x)$ à son espérance mathématique.
On pose

I)

$$Z_{1n}(x) = \frac{\sqrt{n}h^{1+\beta}a_n}{\sigma}\left[\hat{\theta_n}(x) - \mathbb{E}\left(\hat{\theta_n}(x)\right)\right]$$

II)

$$Z_{2n}(x) = \frac{\sqrt{n}h^{\frac{1}{2}+\beta}a_n}{g(x)}\left[\hat{\theta_n}(x) - \mathbb{E}\left(\hat{\theta_n}(x)\right)\right]$$

Théorème 3.5.1. *Sous les hypothèses H1 à H7, pour tout* $\varepsilon>0, \forall(n \geq 4), \forall k \in \{1, ..., [\frac{n}{2} - 1]\}$ $\forall \eta \in]0, ..., (4KMe)^{-1}[$*, on a*

I)

$$\mathbb{P}(|Z_{1n}(x)| > \varepsilon) \leq 2exp\left(-n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

II)

$$\mathbb{P}(|Z_{2n}(x)| > \varepsilon) \leq 2exp\left(-n\eta\left[g(x)\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

Démonstration 3.5.2.

I)

$$\begin{aligned}
Z_{1n}(x) &= \frac{\sqrt{n}h^{1+\beta}a_n}{\sigma}\left[\frac{1}{nha_n}\sum_{k=1}^{n}Y_kK\left(\frac{x-z_k}{h},h\right) - \frac{1}{nha_n}\sum_{k=1}^{n}\mathbb{E}(Y_k)K\left(\frac{x-z_k}{h},h\right)\right] \\
&= \frac{h^{\beta}}{\sqrt{n}\sigma}\sum_{k=1}^{n}(Y_k - \mathbb{E}(Y_k))K\left(\frac{x-z_k}{h},h\right) \\
&= \frac{h^{\beta}}{\sqrt{n}\sigma}\sum_{k=1}^{n}\varepsilon_kK\left(\frac{x-z_k}{h},h\right)
\end{aligned}$$

D'après l'inégalité de type exponentielle relative aux variables aléatoires mélangeantes l'inégalité de Carbon[18] appliquée aux variables $\alpha-$mélangeantes centrées, pour

$$\xi_k = \varepsilon_kK\left(\frac{x-z_k}{h},h\right)$$

on a bien $\mathbb{E}(\xi_k) = 0$, et

$$\mathbb{P}\left(\left|\sum_{i=1}^{n}\xi_i\right| > \varepsilon\right) \leq 2exp\left(-\eta\varepsilon + 6\eta^2e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right)n + 2\sqrt{e}(\alpha_k)^{\frac{2e}{3n}}\frac{n}{k}\right)$$

$$\mathbb{P}\left(|Z_{1n}(x)| > \varepsilon\frac{h^{\beta}}{\sqrt{n}\sigma}\right) \leq 2exp\left(-\eta\varepsilon + 6\eta^2e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right)n + 2\sqrt{e}(\alpha_k)^{\frac{2e}{3n}}\frac{n}{k}\right)$$

on pose $\varepsilon' = \varepsilon\frac{h^{\beta}}{\sqrt{n}\sigma}$, on obtient

$$\mathbb{P}(|Z_{1n}(x)| > \varepsilon') \leq 2exp\left(-n\eta\left[\frac{\sigma}{\sqrt{n}h^{\beta}}\varepsilon - 6\eta^2e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}(\alpha_k)^{\frac{2e}{3n}}\right]\right)$$

Démonstration 3.5.3.

II)

$$Z_{2n}(x) = \frac{\sqrt{n}h^{\frac{1}{2}+\beta}a_n}{g(x)}\left[\frac{1}{nha_n}\sum_{k=1}^{n}Y_kK\left(\frac{x-z_k}{h},h\right) - \frac{1}{nha_n}\sum_{k=1}^{n}\mathbb{E}(Y_k)K\left(\frac{x-z_k}{h},h\right)\right]$$
$$= \frac{h^{\beta}}{\sqrt{nh}g(x)}\sum_{k=1}^{n}(Y_k-\mathbb{E}(Y_k))K\left(\frac{x-z_k}{h},h\right)$$
$$= \frac{h^{\beta}}{\sqrt{nh}g(x)}\sum_{k=1}^{n}\varepsilon_kK\left(\frac{x-z_k}{h},h\right)$$

D'après l'inégalité de type exponentielle relative aux variables aléatoires mélangeantes l'inégalité de Carbon [18] appliquée aux variables $\alpha-$mélangeantes centrées, pour

$$\xi_k = \varepsilon_kK\left(\frac{x-z_k}{h},h\right)$$

on a bien $\mathbb{E}(\xi_k)=0$, et

$$\mathbb{P}\left(\left|\sum_{i=1}^{n}\xi_i\right| > \varepsilon\right) \leq 2exp\left(-\eta\varepsilon + 6\eta^2e\left(V+8M^2\sum_{l=1}^{k}\alpha_l\right)n + 2\sqrt{e}(\alpha_k)^{\frac{2e}{3n}}\frac{n}{k}\right)$$

$$\mathbb{P}\left(|Z_{2n}(x)| > \varepsilon\frac{h^{\beta}}{\sqrt{nh}g(x)}\right) \leq 2exp\left(-\eta\varepsilon + 6\eta^2e\left(V+8M^2\sum_{l=1}^{k}\alpha_l\right)n + 2\sqrt{e}(\alpha_k)^{\frac{2e}{3n}}\frac{n}{k}\right)$$

on pose $\varepsilon' = \varepsilon\frac{h^{\beta}}{\sqrt{nh}g(x)}$, on obtient

$$\mathbb{P}(|Z_{2n}(x)| > \varepsilon') \leq 2exp\left(-n\eta\left[\frac{g(x)}{\sqrt{n}h^{\beta}}\varepsilon - 6\eta^2e\left(V+8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}(\alpha_k)^{\frac{2e}{3n}}\right]\right)$$

3.5.2 Déviation de l'estimateur

D'après notre étude précédente on trouve que notre estimateur prend plusieurs déviation lorsqu'on change le Z_n Puisque :

I) Si

$$Z_{1n}(x) = \frac{\sqrt{n}h^{1+\beta}a_n}{\sigma}\left[\hat{\theta_n}(x) - \mathbb{E}\left(\hat{\theta_n}(x)\right)\right]$$

La déviation suit la fonction exponentielle suivante :

$$\exp(f_{n1}) = \exp - \left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V+8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

Avec

$f_{n1}(x)$=$\left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V+8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$

II) Si

$$Z_{2n}(x) = \frac{\sqrt{n}h^{\frac{1}{2}+\beta}a_n}{g(x)}\left[\hat{\theta_n}(x) - \mathbb{E}\left(\hat{\theta_n}(x)\right)\right]$$

La déviation suit la fonction exponentielle suivante :

$$\exp(f_{n1}) = \exp\left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

Avec

$$f_{n2}(x) = \left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

3.5.3 Convergence presque complet de l'estimateur

Pour étudier la convergence presque complète de l'estimateur, on a obligé d'introduite la limite :

I) Si :

$$f_{n1}(x) = \left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

- $\lim\limits_{n\longrightarrow\infty} f_{n1}(x) = \lim\limits_{n\longrightarrow\infty}\left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right) = \infty$

- $\lim\limits_{n\longrightarrow\infty} \exp(-f_{n1}(x)) = 0$

Donc

$$\mathbb{P}\left(|Z_{1n}| > \varepsilon'\right) = 0$$

$$\mathbb{P}\left(\frac{\sqrt{n}h^{1+\beta}a_n}{\sigma}\left[\hat{\theta_n}(x) - \mathbb{E}\left(\hat{\theta_n}(x)\right)\right] > \varepsilon'\right) = 0$$

Alors :

$$\hat{\theta}_n(x) \xrightarrow{\text{C.P.C}} \theta_n(x)$$

II) Si :

$$f_{n2}(x) = \left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)$$

- $\lim\limits_{n\longrightarrow\infty} f_{n2}(x) = \lim\limits_{n\longrightarrow\infty}\left(n\eta\left[\sigma\frac{\varepsilon}{h^{\beta}\sqrt{n}} - 6\eta e\left(V + 8M^2\sum_{l=1}^{k}\alpha_l\right) - \frac{2\sqrt{e}}{k\eta}\alpha_k^{\frac{2e}{3n}}\right]\right)=\infty$

- $\lim\limits_{n\longrightarrow\infty} \exp(-f_{n2}(x)) = 0$

Donc

$$\mathbb{P}\left(|Z_{2n}| > \varepsilon'\right) = 0$$

$$\mathbb{P}\left(\left|\frac{\sqrt{n}h^{\frac{1}{2}+\beta}a_n}{g(x)}\left[\hat{\theta_n}(x) - \mathbb{E}\left(\hat{\theta_n}(x)\right)\right]\right| > \varepsilon'\right) = 0$$

Alors

$$\hat{\theta}_{2n}(x) \xrightarrow{\text{C.P.C}} \theta_{2n}(x)$$

CONCLUSION

Ce mémoire porte sur l'étude des estimations non paramétrique à erreur alpha mélange, on a présenté la méthode de noyau qui permettre à effectuer l' estimateur dans la statistique non paramétrique, et la méthode de déconvolution dans la statistique non paramétrique à variables mélangeants .
Dans notre travail, on a présenté un modèle de régression, Et a partir de la méthode de déconvolution on a obtenu un estimateur donné par la formule $(3, 4)$, Et sous des hypothèses bien précis on a introduit la notion de alpha mélange pour étudier quelques caractéristiques qui concernent notre estimateur.
Puis nous avons utilisés un théorème qui nous donne quelques inégalités exponentielles pour étudier la convergence presque complète de notre estimateur par plusieurs types de $Z(x)$, qui est absolument donne des bons déviations pour notre estimateur.

BIBLIOGRAPHIE

[1] Adjemlan, S . (2008), Universite de Main le processus AR(p), gains and cepremap.pp. 1-19

[2] Amroun Sonia, Bejaia, 2011 Mémoire De Magister, Sur L'estimation De La Courbe De Régression De La Moyenne.

[3] Belaide Karima, 2013,these doctorat ,universite abderrahmane mira bejaia, resolution de problemes inverses lineaires par une methode iterative a erreurs gaussiennes

[4] Blum,J.R, Hanson, D .L. ; Koopmans (1963). On the strong law for large numbersfor a class of stochastic processes .Z.W.2 :1-11

[5] Davidson, R, James, G.M.1993. Estimation and inference in econometrics, New York, Oxford University Press, p.874.

[6] Doukhan, P.(1994). Mixing :Properties and examples , Lecture Notes in Statistics(V ols :85),springer.

[7] F. Ferraty, et P. Vieu, (2002/2003), Modeles Non parametrique de Regression.

[8] Ferraty, F,Vieu.P.(2006). Nomparametric functional data analysis theory and practice Springer.

[9] J.Fan(1996).Asymptotique Normality deconvolution kernel densite estimators. sankhya Ser .A 53 :97-110,a ;Bissantz et Holzaman (2008).Statistical inference for inverse problem. Inverse Problem 24,034009

[10] Joag-Dev et Proschan Association of random variable with application .Ann.Math.stat 38 :1446-1476

[11] Haddad Soraya, 2018/2019, doctorat,universite abderrahmane mira bejaia ,belaide karima /Processus de long Memoire a Errors Melangeantes.

[12] Hirschfeld, O.(1935).A connection between correlation and contengency. Math. Proc, of Cambridge phil . Soc.31 :520-524

[13] Ibrahimov, I.A.(1962). Some limit theorems for stationary processus. Th. Prob.Appl. 7 :349-382.

[14] Ibrahimov, I,Rozanov.Y.(1974). Processus aleatoires gaussiens . Mir .Moscou.

[15] Gebelein,H.1941. Des statistics problem der korelation variation und eigenwert-problem und sein zusammenhang mit der ausglleischnung . Z. Angew Math. Merch 21 :364-379

[16] Kolmogrov, A .N,Rosanov, Y.A.(1960). On the strong mixing conditions for stationary gaussien sequences .Th. pronb. 5 :204-207

[17] L.Cavalier et A.Tsybakov (2002). Sharp Adaptation for inverse problems with randam noise .Probab.Theor Relat Fieds 123 :323-354

[18] M. Carbon (1983). Inegalite de type exponentielle pour un processus fortement melanges.Application Pub. IRMA lille v 5,2,16p.

[19] Parzen, (1962). On the estimation of a probability density function and the mode. The Annals of Mathematical Statistics, 33, 10651076.

[20] Perraudin, C. (2004-2005). Universite Paris I. Series chronologiques, quelques elements du cours.

[21] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. The Annals of Mathematical Statistics

[22] Rosenbatt, M .(1956).A central limit theorem and a strong mixing condition, Proc. Nat. Acad.Sci.,42 :43-47,USA.

[23] Rio, E .2000. Theorie asymptoti des processus aleatoires faiblement dependants, Mathematique and Application, (Vol.31). Berlin :Springer-Verlag.

ANNEXE A

LES THÉORÈME UTILISÈS :

Définition A.0.1. *espace métrique*

Soit E un ensemble non vide. On appelle distance sur E toute application

$$d : E \times E \to \mathbb{R}^+$$

vérifiant les propriétés suivantes :

a) $\forall x, y \in \mathbb{E}, d(x, y) = 0$

b) $\forall x, y \in \mathbb{E}, d(x, y) = d(y, x)$

c) $\forall x, y, z \in \mathbb{E}, d(x, z) \leq d(x, y) + d(y, z)$ *(inégalité triangulaire).*

On appelle espace métrique tout ensemble non vide $\mathbb{E}$ muni d'une distance d et on le note$(\mathbb{E}, d)$.

Définition A.0.2. *Théorème de Bochner*

soit K : $(R^m, B^m) \longrightarrow (R, B)$ une fonction mesurable, où B^p est la tribu borélienne de $\mathbb{R}^p$, vérifiant

$$\forall z \in \mathbb{R}^m, |K(z)| \leq M$$

$$\int \mathbb{R}^m |K(z)| dz \leq \infty$$

Et

$$\parallel z \parallel^m |K(z)| \longrightarrow 0 : quand \parallel z \parallel \longrightarrow \infty$$

Par ailleurs, soit g :$(\mathbb{R}^m, \beta) \longrightarrow (\mathbb{R}, \beta) une fonction$

Théorème de convergence dominée

Soit $(f_n)_n$ une suite de fonction mesurable de (Ω, T) à valreurs dans $(\mathbb{R}, B(\mathbb{R}))$ si :
* La suite $(f_n)_n$ converge μ- p.p vers une fonction
$f(\Omega, T) \longrightarrow (\mathbb{R}, B\mathbb{R})$.
*Il existe une fonction g de $\mathcal{L}^1_\mu(\Omega, T, \mathbb{R}_+)$ tq :
$\forall n \in \mathbb{N}; |f_n| \leq g.\mu - p.p$

$$\int_\Omega f d\mu = \int_\Omega \lim_n f d\mu = \lim_n \int_\Omega f_n d\mu$$

Définition A.0.3. *On applle tribu (ou σ-algébre) sur E un sous-ensemble $\mathcal{A} \subset \mathcal{P}(E)$ tel que*

1. $E \in \mathcal{A}$
2. *$\mathcal{A}$ est stable par passage au complémentaire :*$A \in \mathcal{A} \Rightarrow A^c = E\ A \in \mathcal{A}$.
3. *$\mathcal{A}$ est stable par réunion dénombrable :*$A_n \in \mathcal{A} \forall n \in \mathbb{N} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Définition A.0.4. {*fonction mesurable*}
Soient (Ω_1, F_1) et (Ω_2, F_2) deux espaces mesurables. L'application $f : \Omega_1 \longrightarrow \Omega_2$ est dite F_1 ; F_2 mesurable
Si pour tout $B \in F_1$, $f^{-1}(B) \in F_1$, autrement dit si $f^{-1}(F_2) \subset F_1$.

Définition A.0.5. {*opérateur compact*} *Un opérateur T de X dans Y est dit compact lorsque T est continue et que toute partie bornée de X est envoyée une partie relativement compact de Y*

Définition A.0.6. {*La norme*}
Soit E un espace vectoriel sur $K = R$ ou c .Une norme sur E est une application N de E dans $\mathbb{R}^+$ vérifiant les conditions suivant pour tout x, y dans E :

1. $N(x) = 0$ *si et seulement si* $x = 0$.
2. $N(x + y) \leq N(x) + N(y)$
3. *Pour tout k dans $\mathbb{K}$,$N(kx) = |k|N(x)$ La deuxième relation s'appelle l'inégalité triangulaire. L'espace E,muni de la N s'appelle espace vectoriel normé.*

Définition A.0.7. *On dit que (M_n) est une martingale (resp. sur martingale ; sous-martingale) 'a temps discret par rapport 'a la filtration $(\mathcal{F})$ si :*

i. (M_n) *est adapté ,*

ii. $\forall n, \mathbb{E}(|M_n|) < \infty,$

iii. $\forall n, \mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n$