

Ministère de l'enseignement supérieur et de la recherche scientifique

*Université Abderrahmane Mira de Béjaïa*

Faculté des sciences exactes

Département Mathématiques



جامعة بجاية  
Tasdawit n Bgayet  
Université de Béjaïa

## *Mémoire de fin d'étude*

**En vu d'obtention d'un Master académique en Mathématiques**

Option : Probabilités Statistiques et Applications

---

### *Thème*

## **Variables aléatoires mélangeantes et effets sur l'estimation**

---

*Réalisé par* : MOUHOUS Katia.

### *Devant le jury*

Mr. MAOUCHE Fouad	M.C.A	Président	Université de Béjaïa
Mme. TIMERIDJINE BELAIDE Karima	M.C.A	Rapporteur	Université de Béjaïa
Mme. TABTI Hadjila	M.C.B	Examinatrice	Université de Béjaïa
Mme. BARACHE Bahia	M.C.B	Examinatrice	Université de Béjaïa

# Dédicaces

---

*Je dédie ce modeste travail :*

*A ma chère mère,*

*A Mon cher père,*

*Qui n'ont jamais cessé de formuler des prières à mon égard, de me soutenir et de m'épauler pour que je puisse atteindre mes objectifs.*

*A mon cher frère khaled,*

*A ma chère souer wissam,*

*A mon cher Salas,*

*À qui je dois tout l'amour, avec tous mes voeux de les voir réussir dans leurs vies.*

*A ma chère tante touncia,*

*A mes amies et mes collègues : Sara, Nawal, Hanane, Saida, Sylia, nouara.*

*À qui je souhaite le succès, pour l'amitié qui nous à toujours unis.*

*A Tous ceux qui me sont chers*

***Katia***

## Remerciements

---

Tout d'abord, je remercie Dieu le tout-puissant qui m'a donné le courage, la force et la volonté pour mener ce travail.

Je tiens à exprimer tout ma reconnaissance à ma directrice de mémoire, **M<sup>me</sup>. TIMERIDJINE-BELAIDE Karima**. Je la remercie de m'avoir encadré, orienté, aidé et conseillé.

Mes remerciements vont également aux membres de jury Le président **M<sup>r</sup>. MAUCHE Bachir** et les deux examinatrices **M<sup>me</sup>. TABTI Hadjila** et **M<sup>me</sup>. IDRES Sonia** qui m'ont fait l'honneur de juger mon travail.

Je remercie mes très chers parents, qui ont toujours été là pour moi. Leur soutien inconditionnel et leurs encouragements ont été d'une grande aide.

Je remercie encore **M<sup>me</sup>. TABTI Hadjila** pour avoir eu la patience de répondre à mes innombrables questions.

Je remercie mes amis pour leurs aides et conseils qu'ils m'ont donné tout au long de mon travail.

Enfin, je remercie toutes les personnes ayant contribué de près ou de loin à la réalisation de ce travail.

# TABLE DES MATIÈRES

<b>Table des matières</b>	<b>II</b>
<b>Liste des figures</b>	<b>III</b>
<b>Liste des tableaux</b>	<b>IV</b>
<b>Liste des abréviations</b>	<b>V</b>
<b>Introduction générale</b>	<b>1</b>
<b>1 Indépendance en probabilité</b>	<b>3</b>
1.1 Indépendance en probabilité . . . . .	4
1.1.1 Indépendance d'événements . . . . .	4
1.1.2 Indépendance de tribus . . . . .	6
1.1.3 Indépendance de variables aléatoires . . . . .	6
1.2 Convergence presque complète . . . . .	8
1.3 Inégalités exponentielles . . . . .	9
1.4 Estimation dans le cas des variables aléatoires indépendantes . . . . .	12
1.4.1 Définition d'un estimateur . . . . .	12
1.4.2 La méthode de maximum de vraisemblance . . . . .	13
1.4.2.1 La fonction de vraisemblance . . . . .	13
1.4.2.2 L'estimateur de maximum de vraisemblance . . . . .	13
1.4.2.3 Propriétés des estimateurs de maximum de vraisemblance . . . . .	14
1.4.3 La méthode des moments . . . . .	15

1.4.3.1	L'estimateur des moments(EMM)	15
1.4.3.2	Propriétés des estimateurs des moments	16
1.4.4	Methode de moindres carrés ordinaires (MCO)	19
1.4.4.1	Modèle de régression linéaire simple	19
1.4.4.2	Modèle de régression linéaire multiple	21
<b>2</b>	<b>Les mélanges</b>	<b>24</b>
2.1	Tribus mélangeantes	24
2.1.1	$\alpha$ -mélange (mélange fort)	24
2.1.2	$\beta$ -mélange (régularisation absolue)	25
2.1.3	$\rho$ -mélange ( le coefficient de corrélations maximales)	25
2.1.4	$\phi$ -mélange (mélange uniforme)	25
2.1.5	$\psi$ -mélange	25
2.2	Processus mélangeants	30
2.2.1	Schéma pour les propriétés de mélange d'un processus	32
2.3	Variables aléatoires mélangeantes	32
2.3.1	Quelques statistique mélangeantes	33
2.4	Inégalités dans le cas des variables mélangeantes	34
2.4.1	Inégalités de covariance	34
2.4.2	Inégalités exponentielles	39
<b>3</b>	<b>Simulation</b>	<b>42</b>
3.1	Simulation	42
	<b>Conclusion générale</b>	<b>46</b>
	<b>Bibliographie</b>	<b>46</b>

TABLE DES FIGURES
-------------------

3.1	Les erreurs entre les moyennes théoriques et les moyennes empiriques. . . . .	44
3.2	Les erreurs entre les variances théoriques et les variances empiriques. . . . .	44

## LISTE DES TABLEAUX

3.1	Les erreurs absolues dans le cas d'un échantillon de taille 10 . . . . .	43
3.2	Les erreurs absolues dans le cas d'un échantillon de taille 100 . . . . .	43
3.3	Les erreurs absolues dans le cas d'un échantillon de taille 1000 . . . . .	44

## LISTE DES ABRÉVIATIONS

$\mathbb{E}(X)$	L'espérance mathématique de la variable aléatoire X.
$Var(X)$	La variance de la variable aléatoire X.
$Cov(X, Y)$	La covariance entre les variables aléatoires X et Y .
$\mathbf{1}_A$	La fonction indicatrice qui vaut 1 sur l'ensemble A et 0 ailleurs.
$\otimes$	Produit tensoriel.
$\perp\!\!\!\perp$	L'indépendance.
$L(\theta; x_1, \dots, x_n)$	Fonction de vraisemblance.
$EMV$	L'estimateur de maximum de vraisemblance.
$EMM$	L'estimateur des moment.
$MCO$	Moindres carrés ordinaire.
$SCR$	Somme des carrés des résidus.
$corr(X, Y)$	Correlation entre X et Y.
$cov(X, Y)$	Covariance entre X et Y.
$\ \cdot\ $	Semi norme
$\ X\ _\infty$	La norme infinie de X.
$\ X\ _p$	La norme p de X
$ess$	Essential supremum.
$\mathcal{F}_j^i$	La $\sigma$ -algèbre engendrée par $\{X_t, i \leq t \leq j\}$ .

## INTRODUCTION GÉNÉRALE

La notion d'indépendance pour les collection de variables aléatoires est un concept qui possède des propriétés intéressantes. La grande majorité des résultats qu'on connaît en théorie des probabilités et qui sont relatifs aux variables aléatoires, ne sont généralement valables que pour des variables aléatoires indépendantes. Et il est souvent difficile voire impossible de généraliser ces résultats aux variables aléatoires dépendantes. Cette dernière ne trouve pas beaucoup d'application en pratique à vrai dire, si l'on regarde profondément les choses, on s'apercevra que dans la majorité des cas, on ne peut pas échapper à la dépendances. Ainsi les phénomènes étudiés dans la physique, la chimie, la biologie, l'économie et la fiabilité étaient des sources principales dont l'émergence des modèles stochastiques et les modèles de variables aléatoires dépendantes[16].

Pour remédier à cette anomalie il serait intéressant d'introduire une structure probabiliste qui permettra de contrôler la dépendance au sein des variables de l'échantillon statistique ce qui reflétera l'évolution exacte d'un phénomène aléatoire.

C'est pour cette raison que, depuis quelques années la notion de dépendance à été un sujet d'intérêt et de préoccupation pour les probabilistes et les statisticiens, tel que elle a été largement utilisée et a fait l'objet d'une intense activité de recherche[15].

Plusieurs façons de contrôl de dépendance ont été introduites, parmi elles, on a l'association et le mélange.

Dans ce mémoire, nous nous intéressons particulièrement aux propriétés de mélange entre les différentes mesures (tribus, processus et variables aléatoires), et nous présentons ces différents types :

- Le mélange fort ( $\alpha$ -mélange) introduit par Rosemblat en 1956.

- Le mélange faible (ou mélange uniforme)(ou  $\phi$ -mélange) introduit par Ibragimov et Y.V.Linnik.
- Le coefficient de régularisation absolue ( $\beta$ -mélange) introduit par Kolmogrov et apparu dans le livre de Volkonski et Rozanov.
- Le coefficient de corrélations maximales ( $\rho$ -mélange) introduit par Hirschfeld(1935) et Gablein(1941).
- $\psi$ -mélange introduit par Blum, Hanson et Koopmas[20].

Ce mémoire comporte trois chapitres, le premier chapitre comporte une présentation de la notion d'indépendance en probabilité, on va parler sur l'indépendance des tribus ( $\sigma$ -algèbre), des événements et des variables aléatoires. Nous donnons quelques théorèmes, propriétés et exemples concernant l'estimation des variables aléatoires indépendantes et on va présenter quelques méthodes d'estimation, parmi elles la méthode des moments, la méthode du maximum de vraisemblance et la méthode des moindres carrés ordinaire (MCO). Nous introduisons aussi les inégalités exponentielle dans le cas d'idépendance.

On deuxième chapitre, nous présentons la notion de mélange entre les différentes propriétés (entre sous  $\sigma$ -algèbres, variables aléatoires ou processus), nous donnons les principales définitions, et nous présentons les différents mélanges (le mélange fort, le mélange faible, la régularité absolue et la régularité complète). On va donner quelques statistiques mélangeantes et aussi les inégalités de covariance et les inégalités exponentielles pour les mélanges.

Nous terminons ce mémoire par un troisième chapitre, dans lequel nous proposons une étude numérique à l'aide de logiciel Matlab.

# CHAPITRE 1

## INDÉPENDANCE EN PROBABILITÉ

### Introduction

L'indépendance est une notion probabiliste qualifiant de manière intuitive des différentes entités n'ayant aucune influence l'une sur l'autre. Il s'agit d'une notion très importante en théorie des probabilités et en statistique.

Ce chapitre se compose de quatre parties, la première partie est consacrée à la présentation des différents types d'indépendance, on a l'indépendance d'événements, l'indépendance de tribus et l'indépendance de variables aléatoires.

Dans la deuxième partie, nous donnons le lien entre les différents types d'indépendance. Dans la troisième partie, nous présentons les inégalités exponentielles dans le cas des variables aléatoires indépendantes et enfin, dans la quatrième partie nous rappelons d'estimation dans le cas des variables aléatoires indépendantes, on précisera les différentes méthodes d'estimation (méthode de maximum de vraisemblance, méthode des moments et méthode des moindres carrés).

## 1.1 Indépendance en probabilité

### 1.1.1 Indépendance d'événements

**Définition 1.1.1** [5] Deux événements  $A$  et  $B$  sur un même espace probabilisé  $(\Omega, \mathcal{A}, P)$  sont dits indépendants en probabilité si :

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B). \quad (1.1)$$

**Proposition 1.1.1** [5] Si les deux événements  $A$  et  $B$  sont indépendants, donc équivalent à dire que :

$$\mathbf{P}(B \setminus A) = \mathbf{P}(B).$$

$$\mathbf{P}(A \setminus B) = \mathbf{P}(A).$$

Telque :  $\mathbf{P}(A)$  et  $\mathbf{P}(B) > 0$  Dire que  $A$  et  $B$  sont indépendants, c'est dire que l'occurrence de  $B$  ou de  $A$  n'influence pas sur la probabilité de réalisation de l'autre événement. En d'autres termes, on ne peut rien inférer quant à la probabilité de l'un de ces événements lorsque l'on sait que l'autre s'est produit. par conséquent, si  $A$  et  $B$  sont indépendants, il devrait en aller de même pour  $A$  et  $B^c$ ,  $A^c$  et  $B$ , ainsi que pour  $A^c$  et  $B^c$  [3].

**Proposition 1.1.2** [5]

i) si  $A$  et  $B$  sont indépendants, alors :

- $A \perp\!\!\!\perp B^c$ .
- $A^c \perp\!\!\!\perp B$ .
- $A^c \perp\!\!\!\perp B^c$ .

ii) Si l'événement  $A$  tel que sa probabilité est nulle ou égale à 1, alors :

$$\forall B \in \mathcal{A} \Rightarrow A \perp\!\!\!\perp B.$$

**Preuve**

i) On a :

$$\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap B^c) = \mathbf{P}(A)\mathbf{P}(B) + \mathbf{P}(A \cap B^c).$$

D'où :

$$\mathbf{P}(A \cap B^c) = \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A)(1 - \mathbf{P}(B)) = \mathbf{P}(A)\mathbf{P}(B^c).$$

ce qui prouve l'indépendance entre  $A$  et  $B^c$ . les autres indépendance s'obtiennent ensuite facilement par symétrie.

ii) Si l'événement  $A$  telque  $\mathbf{P}(A) = 0$ , alors :

$$\begin{aligned} \mathbf{P}(A \cap B) &\leq \mathbf{P}(A) = 0 \\ &= \mathbf{P}(A)\mathbf{P}(B). \end{aligned}$$

Donc  $A \perp\!\!\!\perp B$  si à l'opposé  $\mathbf{P}(A) = 1$ , l'égalité  $\mathbf{P}(A^c) = 0$  est ce qui précède entraine alors :

$$B \perp\!\!\!\perp A^c, \forall B \in \mathcal{A}.$$

Alors :

$$B \perp\!\!\!\perp A, \forall B \in \mathcal{A}.$$

**Définition 1.1.2** Soit  $(A_i)_{i=1,\dots,n}$  une famille d'événements de  $\mathcal{A}$ . Ces événements sont mutuellement indépendants si  $\forall I \subset \{1, \dots, n\}$  :

$$\mathbf{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}(A_i). \tag{1.2}$$

**Remarque 1.1.1** L'indépendance mutuelle entraine clairement l'indépendance deux à deux mais la réciproque n'est pas nécessairement vraie.

**Exemple 1.1.1** [1] Considérons un lancer de deux dés et les événements :

**A** : les résultats des deux dés est paire.

**B** : le premier dé est pair.

**C** : le deuxième dé est pair.

Ces événements sont deux à deux indépendants, car :

$$\mathbf{P}(A) = \mathbf{P}(B) = \mathbf{P}(C) = 1/2 \quad \text{et} \quad \mathbf{P}(A \cap B) = \mathbf{P}(B \cap C) = \mathbf{P}(C \cap A) = 1/4.$$

En revanche :

$$\mathbf{P}(A \cap B \cap C) = \mathbf{P}(B \cap C) = 1/4.$$

Alors que :

$$\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = 1/8.$$

**Remarque 1.1.2** [5] *On peut généraliser cette notion d'indépendance mutuelle pour une famille non nécessairement finie d'éléments.*

**Définition 1.1.3** [5]  *$(A_i)_{i \in I}$  est une famille d'événements mutuellement indépendants si, pour tout ensemble d'indices  $K$  fini et dans  $I$ , la famille  $(A_i)_{i \in K}$  forme une famille d'événements mutuellement indépendants.*

### 1.1.2 Indépendance de tribus

**Définition 1.1.4** [5] *Une famille  $(\mathcal{A}_i)_{i=1, \dots, n}$  est dite indépendante si, quelque soit la famille d'événements  $(A_i)_{i=1, \dots, n}$  ou  $A_i \in \mathcal{A}_i$  pour tout  $i$  on a :*

$$\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbf{P}(A_i). \tag{1.3}$$

*On peut définir la notion d'indépendance pour toute famille non vide de parties de  $\mathcal{A}$ .*

*Une famille  $(C_i)_{i \in I}$  de parties de  $\mathcal{A}$  ( $C_i \subset \mathcal{A}$  pour tout  $i$  ou  $C_i$  n'est pas forcément une tribu) est indépendante si, quand  $I$  est fini, pour toute famille  $(A_i)_{i \in I}$  telque  $A_i \in C_i$  et dans  $C_i$  on a :*

$$\mathbf{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}(A_i). \tag{1.4}$$

**Remarque 1.1.3** [5] *Quand  $I$  est infini, toute sous-famille finie est indépendante.*

### 1.1.3 Indépendance de variables aléatoires

Soient  $X_1, X_2$  deux variables aléatoires définies sur un même espace de probabilité  $(\Omega, \mathcal{A}, \mathbf{P})$ .

**Définition 1.1.5** [23]  *$X_1$  et  $X_2$  indépendantes si, pour tout ensemble  $B$*

$$\mathbf{P}(X_1 \in B, X_2 \in B) = \mathbf{P}(X_1 \in B)\mathbf{P}(X_2 \in B). \tag{1.5}$$

Équivalent à :

$$\mathbf{P}_{(X_1, X_2)} = \mathbf{P}_{X_1} \otimes \mathbf{P}_{X_2}$$

**Définition 1.1.6** [23] Les variables aléatoires  $X_1, \dots, X_n$  sont mutuellement indépendantes si :

$$\mathbf{P}(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n \mathbf{P}(X_i \in B_i). \quad (1.6)$$

Pour tout ensemble mesurable  $B_1, \dots, B_n$  c'est-à-dire si :

$$\mathbf{P}_{(X_1, \dots, X_n)} = \mathbf{P}_{X_1} \otimes \dots \otimes \mathbf{P}_{X_n}.$$

**Remarque 1.1.4** i)  $X_1, \dots, X_n$  indépendantes  $\Rightarrow X_i, X_j$  indépendantes  $\forall i \neq j$ .

ii)  $X_i, X_j$  indépendantes  $\forall i \neq j \not\Rightarrow X_1, \dots, X_n$  indépendantes [23].

**Théorème 1.1.1** [5] Si pour tout  $i$ , les fonctions  $\varphi_i$  sont des fonctions mesurables de  $(E_i, \mathcal{B}_i)$  vers  $(E'_i, \mathcal{B}'_i)$ , alors l'indépendance des variables aléatoires  $(X_i)_{i=1, \dots, n}$  entraîne celle des  $(\varphi_i(X_i))_{i=1, \dots, n}$ .

**Preuve**

Pour toute famille  $(B'_i)_{i=1, \dots, n}$  ou  $B'_i \in \mathcal{B}'_i$ , pour tout  $i$ , on a :

$$\varphi_i^{-1}(B'_i) = B_i \in \mathcal{B}_i.$$

Par mentabilité des  $\varphi_i$ . On aura :

$$(\varphi_i(X_i))^{-1}(B'_i) = X_i^{-1}(\varphi_i^{-1}(B'_i)) = X_i^{-1}(B_i).$$

D'où :

$$\begin{aligned} \mathbf{P}\left(\bigcap_{i=1}^n \varphi_i(X_i) \in B'_i\right) &= \mathbf{P}\left(\bigcap_{i=1}^n (\varphi_i(X_i))^{-1}(B'_i)\right) = \mathbf{P}\left(\bigcap_{i=1}^n X_i^{-1}(B_i)\right) \\ &= \mathbf{P}\left(\bigcap_{i=1}^n X_i \in B_i\right) = \prod_{i=1}^n \mathbf{P}(X_i \in B_i) \\ &= \prod_{i=1}^n \mathbf{P}(\varphi_i(X_i) \in B'_i) \end{aligned}$$

Et les  $(\varphi_i(X_i))_{i=1, \dots, n}$  sont des variables aléatoires indépendantes.

### 1.1.4 Propositions[5]

Soit  $(\Omega, \mathcal{A}, \mathbf{P})$  un espace probabilisé et  $(\mathcal{A}_i)_{i=1, \dots, n}$  une famille de sous tribus de  $\mathcal{A}$ .

**Proposition 1.1.3** *On a l'équivalence entre les assertions suivantes :*

1.  $\prod_{i=1}^n \mathcal{A}_i$ .
2.  $\prod_{i=1}^n X_i$  , pour toute famille  $(X_i)_{i=1, \dots, n}$  de variables aléatoires ou  $X_i$  et  $\mathcal{A}_i$  mesurable.
3.  $\prod_{i=1}^n 1_{A_i}$ , pour tout famille  $(A_i)_{i=1, \dots, n}$  ou  $A_i \in \mathcal{A}_i$  pour tout  $i$ .
4.  $\prod_{i=1}^n \mathcal{A}_i$ , pour tout famille  $(\mathcal{A}_i)_{i=1, \dots, n}$  ou  $A_i \in \mathcal{A}_i$  pour tout  $i$ .

**Proposition 1.1.4** *Soit  $(A_i)_{i=1, \dots, n}$  une famille d'événements sur  $(\Omega, \mathcal{A}, \mathbf{P})$ .*

*On a :*

$$\prod_{i=1}^n A_i \Leftrightarrow \prod_{i=1}^n \delta(A_i) \Leftrightarrow \prod_{i=1}^n 1_{A_i} .$$

$\delta : \sigma - \text{algèbre}$

**Proposition 1.1.5** *Soit  $(X_i)_{i=1, \dots, n}$  une famille de variables aléatoires définies sur  $(\Omega, \mathcal{A}, \mathbf{P})$  et a valeurs respectivement dans  $(E_i, \mathcal{B}_i)_{i=1, \dots, n}$ .*

*On a :*

$$\prod_{i=1}^n X_i \Leftrightarrow \prod_{i=1}^n \delta(X_i) \Leftrightarrow \prod_{i=1}^n 1_{X_i \in B_i} .$$

*Pour  $B_i \in \mathcal{B}_i$ .*

## 1.2 Convergence presque complète

Dans la théorie des probabilités, il existe différentes notions de convergence, parmi elles (la convergence presque sûr, convergence presque complète, la convergence en probabilité, etc..).

Dans cette section nous présentons la convergence presque complète qui à été introduit par Hsu et Robbins (1947)[10].

**Définition 1.2.1** *Une suite de variables aléatoires réelles  $(Z_n)_{n \in \mathbb{N}}$  converge presque complètement vers une variable aléatoire  $Z$  lorsque  $n \rightarrow \infty$  (est on écrit  $\lim_{n \rightarrow \infty} Z_n = Z$  p.co) si et seulement si :*

$$\forall \epsilon > 0, \sum_{n \in \mathbb{N}} \mathbb{P} [| Z_n - Z | > \epsilon] < \infty .$$

**Définition 1.2.2** On dit que la vitesse de convergence presque complète de la suite de variables aléatoires réelles  $(Z_n)_{n \in \mathbb{N}}$  vers  $Z$  est d'ordre  $(U_n)$  (avec  $U_n$  est une suite numérique déterministe), et on note

$$Z_n = O(U_n)_{p.co}$$

si seulement si :

$$\exists \epsilon_0 > 0, \sum_{n \in \mathbb{N}} \mathbb{P} [| Z_n - Z | > \epsilon_0 U_n] < \infty.$$

Notons que la convergence presque complète entraîne à la fois la convergence presque sûre et la convergence en probabilité[8].

Si  $\lim_{n \rightarrow \infty} Z_n = Z$  p.co, alors  $Z_n$  converge en probabilité et presque sûrement vers  $Z$  [8].

## Preuve

La convergence en probabilité se déduit facilement de la convergence de la série suivant

$$\sum_{n \in \mathbb{N}} \mathbb{P} [| Z_n - Z | > \epsilon] < \infty$$

( $\mathbb{P} [| Z_n - Z | > \epsilon]$ , est le terme général d'une série convergente). Le lemme de Borel contelli implique que :

$$\forall \epsilon > 0, \mathbb{P} \left[ \lim_{n \rightarrow \infty} | Z_n - Z | > \epsilon \right] = 0 \text{ p.s}$$

De plus,  $\lim_{n \rightarrow \infty} Z_n(\omega) \neq Z(\omega)$ , implique l'existence de  $\epsilon > 0$ , telque

$$\lim_{n \rightarrow \infty} \sup | Z_n(\omega) - Z(\omega) | > \epsilon \text{ p.s}$$

On a alors  $\mathbb{P} \left( \lim_{n \rightarrow \infty} \sup | Z_n(\omega) - Z(\omega) | > \epsilon \right) = 0$ , c'est à dire  $Z_n \rightarrow Z$ , p.s.

## 1.3 Inégalités exponentielles

Dans cette section nous introduisons quelques inégalités exponentielles de type Bernstein (dans le cas des variables aléatoires indépendantes), qui nous seront utiles à la démonstrations des propriétés de la convergence presque complète.

Nous supposons que  $Z_1, Z_2, \dots, Z_n$  est une suite de variables aléatoires réelles indépendantes et centrées. Pour démontrer la convergence presque complète, nous avons besoin de trouver des

bornes supérieures pour certaines probabilités concernant des sommes de variables aléatoires réelles telles que :

$$\mathbb{P} \left( \left| \sum_{i=1}^n Z_i \right| > \epsilon \right).$$

Où le réel positif  $\epsilon$  diminue avec  $n$ .

Dans ce contexte, il existe de puissants outils probabilistes, appelés génériquement inégalités exponentielles.

La littérature contient divers versions d'inégalités exponentielles. Ces inégalités différentes selon les différentes hypothèses vérifiées par les variables  $Z_i$ .

Nous en présentons ici celles qu'on appelle inégalités de type Bernstein qui est la plus facile pour les développements théorique sur les statistiques fonctionnelles[8].

**Proposition 1.3.1** *on a l'inégalité de Cramer suivante : Si*

$$\forall m \geq 2, \left| \mathbb{E}(Z_i^m) \right| \leq \left(\frac{m!}{2}\right)(a_i)^2 b^{m-2}$$

Où :

$(a_i)_{1 \leq i \leq n}$  sont des réels positifs,  $b \in \mathbb{R}^+$  et  $A_n^2 = a_1^2 + a_2^2 + \dots + a_n^2$  Alors

$$\forall \epsilon \geq 0, \mathbb{P} \left[ \sum_{i=1}^n \left| Z_i \right| > \epsilon A_n \right] \leq 2 \exp \frac{-\epsilon^2}{2(1 + \frac{b\epsilon}{A_n})}.$$

## Preuve

**Corollaire 1.3.1 a)**  $\forall m \geq 2, \exists C_m > 0$  et une constante  $a$  positive, tels que :

$$\mathbb{E} \left| Z_1^m \right| \leq C_m a^{2(m-1)}.$$

On a

$$\forall \epsilon > 0, \mathbb{P} \left[ \left| \sum_{i=1}^n Z_i \right| > \epsilon n \right] \leq 2 \exp \frac{-\epsilon^2 n}{2a^2(1 + \epsilon)}.$$

**b)** *Supposons que les  $(Z_i)_{1 \leq i \leq n}$  dépendent de  $n$  ( c'est à dire que  $Z_i = Z_{i,n}$ ). Si  $\forall m \geq 2$ , il*

existe un réel  $C_m$  strictement positif et une suite  $(a_n)$  de réels positifs, tels que :

$$\mathbb{E} | Z_1^m | \leq C_m a_n^{2(m-1)}.$$

Et si  $U_n = n^{-1} a_n^2 \log n$ , vérifie  $\lim_{n \rightarrow \infty} U_n = 0$ , alors nous avons

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{U_n}) p.co.$$

### Démonstration

1. En remplace  $b = a^2$  et  $A_n = a\sqrt{n}$  dans la proposition précédente, on aboutit à **a**).
2. En posant  $\epsilon = \epsilon_0 \sqrt{U_n}$  dans a) et comme  $U_n$  tend vers zéro, pour une certaine constante  $C'$  nous avons :

$$\mathbb{P} \left[ \frac{1}{n} \left| \sum_{i=1}^n Z_i \right| > \epsilon_0 U_n \right] \leq 2 \exp \frac{-\epsilon_0^2 \log n}{2(1 + \epsilon_0 \sqrt{U_n})} \tag{1.7}$$

$$\leq 2n^{-C'} \epsilon_0^2. \tag{1.8}$$

D'où, pour un choix convenable de  $\epsilon_0$  nous déduisons que

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{U_n}).$$

**Corollaire 1.3.2 a)** S'il existe  $M < \infty$ , telle que :

$$| Z_1 | \leq M$$

Alors nous avons

$$\forall \epsilon > 0, \mathbb{P} \left[ \left| \sum_{i=1}^n Z_i \right| > \epsilon n \right] \leq 2 \exp \frac{-\epsilon^2 n}{2\sigma^2(1 + \frac{M\epsilon}{\sigma^2})}$$

Avec

$$\sigma^2 = \mathbb{E} Z_1^2.$$

**b)** Supposons que les  $(Z_i)_{1 \leq i \leq n}$  dépendent de  $n$  (c'est à dire  $Z_i = Z_{i,n}$ ) et que  $\sigma^2 = \mathbb{E} Z_i^2$ , s'il

existe  $M < \infty$  telle que :

$$|Z_1| \leq M$$

Et

$$\frac{M}{\sigma_n^2} \leq C < \infty$$

Et si

$U_n = n^{-1} \sigma_n^2 \log n$ , vérifie  $\lim_{n \rightarrow \infty} U_n = 0$ , alors nous avons,

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{U_n}) \cdot p.co[\delta]$$

## Démonstration

\* En appliquant la proposition (1.6.2) à  $a_i^2 = \sigma^2$ ,  $A_n^2$ ,  $b = M$  nous aboutissons à **a**).

\*\* Comme  $\frac{MU_n}{\sigma_n^2}$  tend vers zéro, il suffit de reprendre le résultat a) pour  $\epsilon = \epsilon_0 \sqrt{U_n}$ , nous arrivons donc à l'existence d'une constante  $C'$  telle que :

$$\begin{aligned} \mathbb{P} \left[ \frac{1}{n} \left| \sum_{i=1}^n Z_i \right| > \epsilon_0 U_n \right] &\leq 2 \exp \frac{-\epsilon_0^2 \log n}{2(1 + \epsilon_0 \sqrt{\frac{MU_n}{\sigma_n^2}})} \\ &\leq 2n^{-C'} \epsilon_0^2. \end{aligned}$$

## 1.4 Estimation dans le cas des variables aléatoires indépendantes

Parfois, on peut avoir besoin d'estimer les paramètres des lois de probabilités car ils sont inconnus, et ce à partir d'un échantillon de variables aléatoires. Et pour cela il existe de nombreuses méthodes.

Dans cette section nous allons présenter trois méthodes d'estimation les plus usuelles dans le cas des variables aléatoires indépendantes. On a la méthode du maximum de vraisemblance, la méthode des moments et la méthode des moindres carés ordinaires(MCO).

### 1.4.1 Définition d'un estimateur

Pour  $\theta$  on ne dispose que des données  $x_1, \dots, x_n$  donc une estimation de  $\theta$  sera une fonction de ces observations.

**Définition 1.4.1** Une statistique  $t$  est une fonction des observations  $x_1, \dots, x_n$

$$\begin{aligned} t : \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ (x_1, \dots, x_n) &\mapsto t(x_1, \dots, x_n) \end{aligned}$$

Puisque les observations  $x_1, \dots, x_n$  sont des réalisations des variables aléatoires  $X_1, \dots, X_n$ , la quantité calculable à partir des observations  $t(x_1, \dots, x_n)$  est une réalisation de la variable aléatoire  $t(X_1, \dots, X_n)$ . Et on retrouve par exemple le fait que  $\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$  est une réalisation de  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ , on note souvent  $t_n = t(x_1, \dots, x_n)$  et  $T_n = t(X_1, \dots, X_n)$ . Par abus, on donne le même nom de statistique aux deux quantités, mais dans une perspective d'estimation, on va nommer différemment  $t_n$  et  $T_n$ [11].

**Définition 1.4.2** [21] Un estimateur d'un paramètre  $\theta$  est une statistique  $T_n$  à valeurs dans l'ensemble des valeurs possibles de  $\theta$ . Une estimation de  $\theta$  est une réalisation  $t_n$  de l'estimateur  $T_n$ . Un estimateur est donc une variable aléatoire, alors qu'une estimation est une valeur déterministe.

## 1.4.2 La méthode de maximum de vraisemblance

### 1.4.2.1 La fonction de vraisemblance

On appelle fonction de vraisemblance de  $\theta$  pour une réalisation  $(X_1, \dots, X_n)$  d'un échantillon, la fonction de  $\theta$  :

$$L_{(x_1, \dots, x_n)}(\theta) = \prod_{i=1}^n f(x_i, \theta)[21]. \tag{1.9}$$

### 1.4.2.2 L'estimateur de maximum de vraisemblance

L'estimation de maximum de vraisemblance de  $\theta$  est la valeur  $\hat{\theta}$  de  $\theta$  qui rend maximale la fonction de vraisemblance  $L(\theta; x_1, \dots, x_n)$ . L'estimateur de maximum de vraisemblance de  $\theta$  est la variable aléatoire correspondante. La fonction de vraisemblance s'exprime comme un produit. Donc  $\hat{\theta}$  sera en général calculé en maximisant la log-vraisemblance :

$$\hat{\theta}_n = \arg \max \ln L(\theta; x_1, \dots, x_n).$$

Quand  $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$  et que toutes les dérivées partielles ci-dessus existent,  $\hat{\theta}_n$  est

solution du système d'équations appelées équations de vraisemblance :

$$\forall j \in 1, \dots, d, \frac{\partial}{\partial \theta_j} \ln L(\theta; x_1, \dots, x_n) = 0.$$

A priori, une solution de ce système d'équations pourrait être un minimum de la vraisemblance. Mais on peut montrer que la nature d'une fonction de vraisemblance fait que c'est bien un maximum que l'on obtient [11].

### 1.4.2.3 Propriétés des estimateurs de maximum de vraisemblance

Si les  $X_i$  sont indépendants et de même loi dépendant d'un paramètre réel  $\theta$ , cette loi vérifiant les conditions de régularité, on a :

- $\hat{\theta}_n$  converge presque sûrement vers  $\theta$ .
- $\sqrt{I_n(\theta)}(\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, 1)$ , donc quand  $n$  est grand,  $\hat{\theta}_n$  est approximativement de loi  $\mathcal{N}(\theta, \frac{1}{I_n(\theta)})$ .  
On en déduit que  $\hat{\theta}_n$  est asymptotiquement gaussien sans biais et efficace.
- Si  $\hat{\theta}_n$  est l'estimateur de maximum de vraisemblance de  $\theta$  alors  $\varphi(\hat{\theta}_n)$  est l'estimateur de maximum de vraisemblance de  $\varphi(\theta)$ [11].

#### Exemple 1.4.1 [21] (Loi exponentielle)

Si les  $X_i$  sont de loi  $\exp(\theta)$ , la fonction de vraisemblance est :

$$\begin{aligned} L(\theta; x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i} = \prod_{i=1}^n \theta \exp(-\theta x_i) \\ &= \theta^n \exp(-\theta \sum_{i=1}^n x_i) \end{aligned}$$

D'où :

$$\ln L(\theta; x_1, \dots, x_n) = n \ln \theta - \theta \sum_{i=1}^n x_i.$$

Qui s'annule pour :

$$\theta = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}_n}.$$

Par conséquent, l'EMV de  $\theta$  est :

$$\hat{\theta}_n = \frac{1}{\bar{X}_n}.$$

**Exemple 1.4.2 [11] (loi gaussienne)** Si les  $X_i$  sont de loi  $\mathcal{N}(m, \sigma^2)$ , la fonction de vraisem-

blance est :

$$\begin{aligned} L(m, \sigma^2; x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i}(x_i; m, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - m)^2}{2\sigma^2}\right] \\ &= \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2\right] \end{aligned}$$

On a :

$$\ln L(m, \sigma^2; x_1, \dots, x_n) = -\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2.$$

On doit annuler les dérivées partielles de ce logarithme par rapport à  $m$  et  $\sigma^2$  on a :

$$\frac{\partial}{\partial m} \ln L(m, \sigma^2; x_1, \dots, x_n) = -\frac{1}{2\sigma^2} \sum_{i=1}^n -2(x_i - m) = \frac{1}{\sigma^2} (\sum_{i=1}^n x_i - nm).$$

Qui s'annule pour :

$$m = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_n.$$

$$\frac{\partial}{\partial \sigma^2} \ln L(m, \sigma^2; x_1, \dots, x_n) = -\frac{n}{2\sigma^2} + \frac{1}{\sigma^4} \sum_{i=1}^n (x_i - m)^2.$$

Qui s'annule pour :

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2.$$

$\hat{m}_n$  et  $\hat{\sigma}_n^2$  sont les valeurs de  $m$  et  $\sigma^2$  qui vérifient les deux conditions en même temps.

On a donc :

$$\hat{m}_n = \bar{X}_n.$$

Et

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 = S_n^2$$

.

### 1.4.3 La méthode des moments

#### 1.4.3.1 L'estimateur des moments(EMM)

C'est la méthode la plus naturelle, que nous avons déjà utilisée sans la formaliser. L'idée de base est d'estimer une espérance mathématique par une moyenne empirique, une variance par

une variance empirique, etc... Si le paramètre à estimer est l'espérance de la loi des  $X_i$ , alors on peut l'estimer par la moyenne empirique de l'échantillon. Autrement dit, si  $\theta = E(X)$  alors l'estimateur de  $\theta$  par la méthode des moments est :

$$\tilde{\theta} = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i. \quad (1.10)$$

Plus généralement, pour  $\theta \in \mathbb{R}$ , si  $E(X) = \varphi(\theta)$ , ou  $\varphi$  est une fonction inversible, alors l'estimateur de  $\theta$  par la méthode des moments est :

$$\tilde{\theta}_n = \varphi^{-1}(\bar{X}_n).$$

De la même manière, on estime la variance de la loi des  $X_i$  par la variance empirique de l'échantillon :

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2$$

Plus généralement, si la loi des  $X_i$  a deux paramètres  $\theta_1$  et  $\theta_2$  tels que :

$$(E(X), Var(X)) = \varphi(\theta_1, \theta_2).$$

Ou  $\varphi$  est une fonction inversible, alors les estimateurs de  $\theta_1$  et  $\theta_2$  par la méthode des moments sont :

$$(\tilde{\theta}_{1n}, \tilde{\theta}_{2n}) = \varphi^{-1}(\bar{X}_n, \sigma_n^2).$$

Ce principe peut naturellement se généraliser aux moments de tous ordres, centrés ou non centrés  $E[(X - E(X))^k]$  et  $E(X^k)$ ,  $k \geq 1$  [11].

### 1.4.3.2 Propriétés des estimateurs des moments

Si  $\theta = \mathbb{E}(x)$ , alors l'EMM de  $\theta$  est  $\tilde{\theta}_n$ , la justification de cette méthode est la loi des grands nombres qui dit que  $\bar{x}_n$  converge presque sûrement vers  $\mathbb{E}(x)$ . Donc, si  $\theta = \mathbb{E}(x)$ ,  $\bar{x}_n$  est un estimateur de  $\theta$  convergent presque sûrement. Autrement dit, si on a beaucoup d'observations, on peut estimer une espérance par une moyenne empirique. On peut en fait montrer facilement que  $\bar{x}_n$  est un bon estimateur de  $\theta = \mathbb{E}(x)$ , sans utiliser la loi des grands nombres :

$$\mathbb{E}(\bar{x}_n) = \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n x_i \right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(x_i) = \frac{1}{n} n\theta = \theta.$$

Donc  $\bar{x}_n$  est un estimateur sans biais de  $\theta = \mathbb{E}(x)$ .

La variance de  $\bar{x}_n$  est :

$$\text{Var}(\bar{x}_n) = \text{Var}\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(x_i) = \frac{\text{Var}(x_i)}{n}.$$

$\text{Var}(\bar{x}_n)$  tend vers 0 quand n tend vers l'infini.

On considère que l'estimation de la variance de la loi des  $x_i$  par la variance empirique de l'échantillon :

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2.$$

$$\begin{aligned} \mathbb{E}S_n^2 &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2\right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i^2) - \mathbb{E}(\bar{X}_n^2) \\ &= \mathbb{E}(x^2) - \mathbb{E}(\bar{x}_n^2) \\ &= \text{Var}(x) + \mathbb{E}(x^2) - \text{Var}(\bar{X}_n) - \mathbb{E}(\bar{X}_n)^2 \\ &= \text{Var}(x) + \mathbb{E}(x^2) - \frac{\text{Var}(x)}{n} - \mathbb{E}(X)^2 \\ &= \frac{n-1}{n} \text{Var}(x) \neq \text{Var}(x). \end{aligned}$$

Donc la variance empirique  $S_n^2$  n'est pas un estimateur sans biais de  $\text{Var}(x)$

On voit que :

$$\mathbb{E}\left(\frac{n}{n-1} S_n^2\right) = \frac{n}{n-1} \mathbb{E}(S_n^2) = \text{Var}(x).$$

On pose donc :

$$\begin{aligned} S_n'^2 &= \frac{n}{n-1} S_n^2 \\ &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2. \end{aligned}$$

$S_n'^2$  est appelée variance estimée de l'échantillon (estimateur sans biais de  $\text{Var}(x)$ ).

On montre que :

$$\text{Var}(S_n'^2) = \frac{1}{n(n-1)} \left[ (n-1) \mathbb{E}[(x - \mathbb{E}(x))^4] - (n-3) \text{Var}(x)^2 \right].$$

qui tend vers l'infini, par conséquent la variance estimée :

$$S_n'^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X}_n)^2.$$

Est un estimateur sans biais et convergent en moyenne quadratique de  $Var(x)$ .

On peut montrer également toutes les deux presque sûrement vers  $Var(x)$ [11].

**Exemple 1.4.3** [21] (loi gaussienne)

On veut estimer la moyenne et la variance par la méthode des moments posons :

$$\theta = (m, \sigma^2) \in \mathbb{R} \times ]0; +\infty[$$

On a :

$m = E_\theta(X)$  et  $\theta^2 = Var_\theta(X)$ . On sait de plus que :  $E_\theta(X^2) = \sigma^2 + m^2$ , Ce qui donne alors :

$$\frac{1}{n} \sum_{i=1}^n X_i = \tilde{m}.$$

Et

$$\frac{1}{n} \sum_{i=1}^n X_i^2 = \tilde{\sigma}^2 + \tilde{m}^2.$$

On en déduit donc que  $\tilde{m}$  est la moyenne empirique et

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^2.$$

**Exemple 1.4.4** [11] (loi exponentielle)

Si  $X_1, \dots, X_n$  sont  $\perp\!\!\!\perp$  et de même loi  $\exp(\theta)$  alors :

$$E(X) = \frac{1}{\theta}$$

Et l'estimateur de  $\theta$  par la méthode des moments est :

$$\tilde{\theta}_n = \frac{1}{\bar{X}_n}.$$

### 1.4.4 Méthode de moindres carrés ordinaires (MCO)

La méthode des moindres carrés est très utilisée généralement pour estimer les paramètres d'une régression. Elle consiste à minimiser la distance entre une variable endogène et la partie certaine de la régression.

#### 1.4.4.1 Modèle de régression linéaire simple

##### Formulation analytique

Les  $x_i$  et  $y_i$  n'étant pas, dans l'immense majorité des cas, exactement liés de façon affine, on suppose qu'elles le sont en moyenne c'est-à-dire que

$$\mathbb{E}[y_i] = a + b\mathbb{E}[x_i] \text{ pour tout } i = 1, \dots, n.$$

Le modèle à estimer s'écrit :

$$y_i = a + bx_i + \varepsilon_i$$

Ou :

- $x_i$  est une variable aléatoire observée appelé variable explicative.
- $y_i$  est une variable aléatoire observée appelé variable expliquer.
- $a, b$  sont des paramètres de régression ou coefficients de régression.
- $\varepsilon_i$  sont des variables aléatoires indépendantes des  $x_i$ , non obserées, appelées erreurs ou bruits, aux quelles on impose certaines conditions complémentaires.

Les condétions standard imposées aux  $\varepsilon_i$  sont les suivantes :

- $\mathbb{E}[\varepsilon_i] = 0$  ,pour tout  $i = 1, \dots, n$  (centrage).
- $Cov(\varepsilon_i, \varepsilon_j) = 0$  , pour tout  $i \neq j$  (non correlation).
- $Var(\varepsilon_i) = \sigma^2$ , (inconnue) pour tout  $i = 1, \dots, n$  (homogécélasticité)[9].

Une fois que les coefficients sont estimés, le modèle va s'écrire :

$$\hat{y}_i = \hat{a} + \hat{b}x_i.$$

Ou encore :

$$y_i = \hat{a} + \hat{b}x_i + e_i$$

Ou  $\hat{a}$  et  $\hat{b}$  désignent les valeurs estimées des paramètres  $a$  et  $b$ ,  $e_i = y - \hat{y}$  est appelé le résidu de modèle.

$e_i$  est l'estimateur de l'erreur  $\varepsilon_i$  que l'on ne connaît pas.[13]

### **l'estimation par MCO [19]**

Le problème est de déterminer les paramètres estimés ( $\hat{a}$  et  $\hat{b}$ ) de telle sorte que l'ajustement  $\hat{y}_i = \hat{a} + \hat{b}x_i$ , soit aussi proche que possible de l'observation  $y_i$ , ou autrement dit que l'erreur (estimée)  $\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - \hat{a} - \hat{b}x_i$  soit aussi proche que possible de 0 et cela pour chaque  $i$ . La mesure de la proximité que l'on retient constitue le critère des moindres carrés ordinaires, c'est-à-dire qu'on retient les valeurs  $\hat{a}$  et  $\hat{b}$  qui minimisent la somme des carrés des résidus :

$$(\hat{a}, \hat{b}) = \arg \min \sum_{i=1}^n \varepsilon_i^2 = \arg \min \sum_{i=1}^n (y_i - a - bx_i)^2.$$

Les conditions du premier ordre, appelées équations normales, sont les suivantes :

$$\sum_{i=1}^n (y_i - a - bx_i) = 0.$$

$$\sum_{i=1}^n x_i (y_i - a - bx_i) = 0.$$

Ainsi, pour  $\hat{a}$  et  $\hat{b}$  solution de ce système, la première équation est  $\sum_{i=1}^n \hat{\varepsilon}_i = 0$  elle signifie que les paramètres doivent être tels que les résidus estimés sont de moyenne nulle. La seconde condition s'écrit quant à elle  $\sum_{i=1}^n x_i \hat{\varepsilon}_i = 0$ , elle implique que  $x_i$  et les résidus estimés sont orthogonaux après quelques réarrangements, on obtient :

$$\hat{a} = \bar{y} - \hat{b}\bar{x}.$$

$$\begin{aligned} \hat{b} &= \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n y_i x_i - n\bar{y}\bar{x}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \\ &= \frac{Cov(y_i, x_i)}{Var(x_i)}. \end{aligned}$$

Avec  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  et  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  tel que la droite passe par le point moyen du nuage  $(\bar{x}, \bar{y})$  et  $\hat{a}$  est donné par le rapport de la covariance entre  $x$  et  $y$  et la variance de  $x$ .

### Hypothèses et propriétés des estimateurs des MCO

Les hypothèses liées à l'erreur  $\varepsilon_i$  sont :

$H_1$  :  $x_i$  est une variable certaine (non aléatoire)  $\implies Cov(x_i, \varepsilon_i) = 0$   
 $\forall i$  (la variable explicative et l'erreur sont indépendantes).

$H_2$  :  $\mathbb{E}(\varepsilon_i) = 0 \forall i$  l'erreur est d'espérance nulle.

$H_3$  :  $Cov(\varepsilon_i, \varepsilon_j) = \mathbb{E}(\varepsilon_i, \varepsilon_j) - \mathbb{E}(\varepsilon_i)\mathbb{E}(\varepsilon_j) = 0$ . Car  $\mathbb{E}(\varepsilon_i) = 0 \forall i \neq j \implies$  les erreurs sont non corrélées [13].

Ces hypothèses permettent au x estimateurs d'obtenir les bonnes propriétés suivantes :

- Les estimateurs sont sans biais

$$\mathbb{E}(\hat{b}) = b$$

$$\mathbb{E}(\hat{a}) = a$$

- Les estimateurs sont convergents

$$\lim_{n \rightarrow \infty} Var(\hat{b}) = 0$$

$$\lim_{n \rightarrow \infty} Var(\hat{a}) = 0.$$

#### 1.4.4.2 Modèle de régression linéaire multiple

Le modèle de régression multiple est une généralisation du modèle de régression simple. Il comporte plusieurs variables explicatives.

Soit le modèle de régression multiple suivant qui comporte  $k$  variables explicatives :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_{k-1} x_{(k-1)i} + \varepsilon_i$$

pour  $i = 1, \dots, n$

Ou encore :

$$y = e\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{k-1} x_{(k-1)} + \varepsilon \Leftrightarrow y = x\beta + \varepsilon$$

Avec

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad x_j = \begin{pmatrix} x_{j1} \\ \vdots \\ x_{jn} \end{pmatrix}_{j=1,\dots,k} \quad x = [ex_1 \dots x_k] \quad \beta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}.$$

$y$  est un vecteur à  $n$  événements,  $x$  est une matrice à  $n$  lignes et  $k + 1$  colonnes,  $\beta$  est un vecteur à  $k + 1$  éléments, et  $\varepsilon$  est un vecteur à  $n$  éléments [19].

### l'estimation par MCO [13]

Le but de l'économètre est de déterminer la valeur du vecteur de paramètre  $\beta$  le mieux possible sur la base des observations  $i = 1, \dots, n$  dont il dispose. Il s'agit donc de construire un estimateur  $\hat{\beta}$  du vecteur  $\beta$  qui soit sans biais et convergent dans le modèle

$$y = \beta + \beta x + \varepsilon$$

On choisit  $\hat{\beta}$  de façon à ce que l'ajustement  $\hat{y}_i = x'_i \hat{\beta}$  soit aussi proche que possible des observations  $y_i$ , c'est-à-dire qui minimise la somme des carrés des résidus :

$$\min \sum \varepsilon_i^2 = \sum (y_i - x'_i \beta)^2 = \varepsilon' \varepsilon = (y - x\beta)'(y - x\beta).$$

La condition nécessaire s'écrit :

$$-2x'y + 2x'x\beta = 0.$$

La solution est alors donnée par :

$$\hat{\beta} = (x'x)^{-1}x'y.$$

### Propriétés de $\hat{\beta}$

$\hat{\beta}$  est un estimateur sans biais

$$\mathbb{E}(\hat{\beta}) = \mathbb{E}[(x'x)^{-1}x'(x\beta + \varepsilon)] = \beta + (x'x)^{-1}\mathbb{E}(x'\varepsilon) = \beta$$

La variance de  $\hat{\beta}$  est :

$$Var(\hat{\beta}) = (x'x)^{-1}x'Var(y)x(x'x)^{-1} = \sigma^2(x'x)^{-1}.$$

Montrons que l'estimateur sans biais de la variance des erreurs est donné par :

$$\hat{\sigma}^2 = \frac{1}{n - k - 1} \sum \hat{\varepsilon}_i^2 = \frac{SCR}{n - k - 1}.$$

Ou  $\hat{\varepsilon}_i = y_i - x_i' \hat{\beta}$  puisque  $\hat{\beta} = (x'x)^{-1} x'y$  s'écrit en fonction de  $y$ , et que  $y$  est supposé suivre une loi normale, alors  $\hat{\beta}$  suit aussi une loi normale d'espérance  $\mathbb{E}(\hat{\beta}) = \beta$  et de variance  $Var(\hat{\beta}) = \sigma^2(x'x)^{-1}$  puisque l'on doit estimer  $\sigma^2$ , alors  $\hat{\beta}$  suit une loi de student à  $n - k - 1$  degrés de liberté.

### Hypothèses et propriétés des estimateurs des MCO

$H_1$  :  $\mathbb{E}(\varepsilon) = 0 \Rightarrow \mathbb{E}(\varepsilon_i) = 0 \forall i$  l'erreur est d'espérance nulle.

$H_2$  :  $x$  est une matrice composée de variables certaines  $\Rightarrow Cov(x_{ji}, \varepsilon_i) = 0 \forall i$  et  $j$  pas de corrélation entre la variable explicative et l'erreur  $\varepsilon_i$ .

$H_3$  :  $\frac{1}{n}x'x$  tend vers une matrice finie non singulière.

Ces hypothèses permettent aux estimateurs d'obtenir les propriétés suivantes :

- Les estimateurs sont sans biais :

$$\mathbb{E}(\hat{\beta}) = \beta.$$

- Les estimateurs sont convergents :

$$\lim_{n \rightarrow \infty} Var(\hat{\beta}) = 0.$$

On a :

$$Var(\hat{\beta}) = \sigma_\varepsilon^2 (x'x)^{-1}.$$

Ou encore :

$$Var(\hat{\beta}) = \sigma_\varepsilon^2 \frac{1}{n} \left( \frac{1}{n} (x'x) \right)^{-1}.$$

## Introduction

La grande majorité des résultats qu'on connaît en théorie des probabilités et qui sont relatifs aux variables aléatoires ne sont généralement valables que pour des variables aléatoires indépendantes, et il est souvent difficile voire impossible de généraliser ces résultats aux 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, il y a les mélanges.

Dans ce chapitre, on traitera les différents types de mélanges qui sont reliés à des mesures de dépendances entre sous-tribus, variables aléatoires ou processus. Nous allons d'abord définir les différents mélanges entre tribus.

### 2.1 Tribus mélangeantes

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $\mathcal{A}$  et  $\mathcal{B}$  deux sous  $\sigma$ -algèbres de  $\mathcal{F}$ . Les différentes mesures de dépendance entre deux tribus  $\mathcal{A}$  et  $\mathcal{B}$  sont définies comme suit :

#### 2.1.1 $\alpha$ -mélange (mélange fort)

Le coefficient  $\alpha$ -mélange, a été introduit par Rosenblatt en 1956[12] :

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, A \in \mathcal{A}, B \in \mathcal{B}\}. \quad (2.1)$$

### 2.1.2 $\beta$ -mélange (régularisation absolue)

Le coefficient  $\beta$  est le coefficient de régularisation absolue introduit par Kolmogorov et apparu dans le livre de Volkonski et Rozanov, il s'écrit sous la forme :

$$\beta(\mathcal{A}, \mathcal{B}) = \frac{1}{2} \sup \sum_{i=1}^I \sum_{j=1}^J |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)|. \quad (2.2)$$

Avec le supremum est pris sur les partitions finies  $\{A_i, 1 \leq i \leq I\}$  et  $\{B_j, 1 \leq j \leq J\}$  de  $\Omega$  ne contenant respectivement que des éléments de  $\mathcal{A}$  et  $\mathcal{B}$ [12].

### 2.1.3 $\rho$ -mélange ( le coefficient de corrélations maximales)

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{|corr(X, Y) - 1|, X \in L^2(\mathcal{A}), Y \in L^2(\mathcal{B})\}. \quad (2.3)$$

introduit par Hirschfeld(1935) et Gablein(1941)[4].

### 2.1.4 $\phi$ -mélange (mélange uniforme)

$$\phi(\mathcal{A}, \mathcal{B}) = \sup \left\{ \left| \mathbb{P}(B) - \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \right|, A \in \mathcal{A}, \mathbb{P}(A) \neq 0, B \in \mathcal{B} \right\} \quad (2.4)$$

Ce coefficient dit  $\phi$  - *mélange* est le coefficient de mélange uniforme introduit par Ibragimov. Il s'écrit sous la forme suivante :

$$\phi(\mathcal{A}, \mathcal{B}) = \sup \{ |\mathbb{P}(B/A) - \mathbb{P}(B)|, A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P}(A) \neq 0 \} [12].$$

### 2.1.5 $\psi$ -mélange

$$\psi(\mathcal{A}, \mathcal{B}) = \sup \left\{ \left| \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} - 1 \right|, A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P}(A) \neq 0, \mathbb{P}(B) \neq 0 \right\}. \quad (2.5)$$

Il s'écrit aussi sous la forme :

$$\psi(\mathcal{A}, \mathcal{B}) = \sup \left\{ \frac{1}{\mathbb{P}(B)} |\mathbb{P}(B/A) - \mathbb{P}(B)|, A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P}(B) \neq 0 \right\}.$$

Ce coefficient est introduit par Blum, Hanson et Koopmas[16].

**Remarque 2.1.1** [6] Les plages des coefficients de dépendance précédents sont définies par les inégalités suivantes :

$$0 \leq \alpha(\mathcal{A}, \mathcal{B}) \leq \frac{1}{4}, \quad 0 \leq \beta(\mathcal{A}, \mathcal{B}) \leq 1, \quad 0 \leq \psi(\mathcal{A}, \mathcal{B}) \leq \infty, \quad 0 \leq \phi(\mathcal{A}, \mathcal{B}) \leq 1, \\ 0 \leq \rho(\mathcal{A}, \mathcal{B}) \leq 1.$$

**Proposition 2.1.1** [14]  $\mathcal{A}, \mathcal{B}$  sont deux sous- $\sigma$ -algèbre de  $\mathcal{F}$ , on a :

i)  $2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B}) \leq \phi(\mathcal{A}, \mathcal{B}) \leq \frac{1}{2}\psi(\mathcal{A}, \mathcal{B})$ .

ii)  $4\alpha(\mathcal{A}, \mathcal{B}) \leq \rho(\mathcal{A}, \mathcal{B}) \leq 2\sqrt{\phi(\mathcal{A}, \mathcal{B})}$ .

iii)  $\rho(\mathcal{A}, \mathcal{B}) \leq \psi(\mathcal{A}, \mathcal{B})$ .

### Démonstration

i) Soit  $(A_i)_{i \in I} \subset \mathcal{A}$  et  $(B_j)_{j \in J} \subset \mathcal{B}$  des partitions de  $\Omega$ . Soit  $A \in \mathcal{A}$  et soit  $B \in \mathcal{B}$ , On a :

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| + |\mathbb{P}(A^c \cap B) - \mathbb{P}(A^c)\mathbb{P}(B)| \\ + |\mathbb{P}(A \cap B^c) - \mathbb{P}(A)\mathbb{P}(B^c)| + |\mathbb{P}(A^c \cap B^c) - \mathbb{P}(A^c)\mathbb{P}(B^c)|.$$

Mais comme :

$$|\mathbb{P}(A^c \cap B) - \mathbb{P}(A^c)\mathbb{P}(B)| = |\mathbb{P}(B) - \mathbb{P}(A \cap B) - (1 - \mathbb{P}(A))\mathbb{P}(B)|.$$

Et

$$|\mathbb{P}(A \cap B^c) - \mathbb{P}(A)\mathbb{P}(B^c)| = |\mathbb{P}(A) - \mathbb{P}(A \cap B) - (1 - \mathbb{P}(B))\mathbb{P}(A)|.$$

Et

$$|\mathbb{P}(A^c \cap B^c) - \mathbb{P}(A^c)\mathbb{P}(B^c)| = |1 - \mathbb{P}(A \cup B) - (1 - \mathbb{P}(B))(1 - \mathbb{P}(A))|.$$

C'est-à-dire que :

$$|\mathbb{P}(A^c \cap B) - \mathbb{P}(A^c)\mathbb{P}(B)| = |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

Et

$$|\mathbb{P}(A \cap B^c) - \mathbb{P}(A)\mathbb{P}(B^c)| = |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$$

Et

$$|\mathbb{P}(A^c \cap B^c) - \mathbb{P}(A^c)\mathbb{P}(B^c)| = |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

On aura :

$$2|\mathbb{P}(A^c \cap B^c) - \mathbb{P}(A^c)\mathbb{P}(B^c)| \leq \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J (|\mathbb{P}(A_i)\mathbb{P}(B_j) - \mathbb{P}(A_i \cap B_j)|).$$

Ainsi

$$2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B}).$$

D'une autre part nous avons grâce aux propriétés de la probabilité et grâce au fait que les  $B_i$  constituent une partition de  $\Omega$ .

$$\sum_{i=1}^I \sum_{j=1}^J (|\mathbb{P}(A_i)\mathbb{P}(B_j) - \mathbb{P}(A_i \cap B_j)|) \leq 2 \sup_{A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P} \neq 0} \left( \left| \mathbb{P}(B) - \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \right| \right).$$

Donc :

$$\beta(\mathcal{A}, \mathcal{B}) \leq \phi(\mathcal{A}, \mathcal{B}).$$

Par conséquent :

$$\left| \frac{\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A \cap B)}{\mathbb{P}(A)} \right| \leq \left| \frac{\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} \right|, \mathbb{P}(A)\mathbb{P}(B) \neq 0.$$

Car  $0 < \mathbb{P}(B) \leq 1$ .

Finalement :

$$\phi(\mathcal{A}, \mathcal{B}) \leq \frac{1}{2}\psi(\mathcal{A}, \mathcal{B}).$$

Ce qui montre que :

$$2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B}) \leq \phi(\mathcal{A}, \mathcal{B}) \leq \frac{1}{2}\psi(\mathcal{A}, \mathcal{B})[4].$$

ii) Nous considérons  $X = \mathbf{1}_A$ ,  $Y = \mathbf{1}_B$  conduit à  $4\alpha(\mathcal{A}, \mathcal{B}) \leq \rho(\mathcal{A}, \mathcal{B})$

En effet, la variance de X est :

$$\mathbb{P}(A)(1 - \mathbb{P}(A)) \leq \frac{1}{4}.$$

$$4|\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A \cap B)| = 4|\text{cov}(\mathbf{1}_A, \mathbf{1}_B)| \leq |\text{corr}(\mathbf{1}_A, \mathbf{1}_B)|.$$

Car :

$$\begin{aligned} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| &= |\mathbb{P}(A \cap B) - \mathbb{P}(A \cap B)\mathbb{P}(B) - \mathbb{P}(A \cap B^c)\mathbb{P}(B)| \\ &= |\mathbb{P}(A \cap B)\mathbb{P}(B^c) - \mathbb{P}(A \cap B^c)\mathbb{P}(B)| \\ &\leq \mathbb{P}(B^c)\mathbb{P}(B) \leq \frac{1}{4}. \end{aligned}$$

Pour la première inégalité nous utilisons l'inégalité

$$||a| - |b|| \leq \max\{|a|, |b|\}.$$

notons que :

$$\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) = \text{cov}(\mathbf{1}_A, \mathbf{1}_B).$$

En utilisant l'inégalité de Cauchy-schwarz, on obtient :

$$4|\text{cov}(\mathbf{1}_A, \mathbf{1}_B)| \leq |\text{corr}(\mathbf{1}_A, \mathbf{1}_B)|$$

En prenant le sup des deux membres de l'inégalité précédente on obtient le résultat.

Maintenant l'inégalité  $\rho(\mathcal{A}, \mathcal{B}) \leq 2\phi^{\frac{1}{2}}(\mathcal{A}, \mathcal{B})$  est évidente d'après l'inégalité de covariance suivante :

$$|\text{cov}(X, Y)| \leq 2\phi^{\frac{1}{2}}(\mathcal{A}, \mathcal{B})\|X\|_p\|Y\|_q.$$

pour  $p = q = 2$ [14]

**Proposition 2.1.2** Soit  $\mathcal{A}, \mathcal{B}$  deux sous  $\sigma$ -algèbres de  $\mathcal{F}$ .

Il est clair que :

- i)  $0 \leq \alpha(\mathcal{A}, \mathcal{B}) \leq \frac{1}{4}$ .
- ii)  $\alpha(\mathcal{A}, \mathcal{A}) = \frac{1}{4}$  à condition que  $\mathcal{A}$  contient un événement d'une probabilité de  $\frac{1}{2}$ .
- iii)  $\alpha(\mathcal{A}, \mathcal{A}) = 0$  à condition que  $\mathcal{A}$  non triviale.
- iv)  $\alpha(\mathcal{A}, \mathcal{B}) \leq \alpha(\mathcal{A}', \mathcal{B}')$ , à condition que  $\mathcal{A}' \subset \mathcal{A}$  et  $\mathcal{B}' \subset \mathcal{B}$ .

**Démonstration** [4]

i) Soit  $A \in \mathcal{A}$  et  $B \in \mathcal{B}$  et grâce aux propriétés de la probabilité suivantes :

- 1)  $\mathbb{P}(U) = 1$ , ( $U$  est l'univers d'une expérience aléatoire).

$$2) \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

$$3) \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A).$$

On aura :

$$\max(\mathbb{P}(A) + \mathbb{P}(B) - 1, 0) \leq \mathbb{P}(A \cap B) \leq \min(\mathbb{P}(A), \mathbb{P}(B))$$

Grâce a la deuxième inégalité (c-à-d  $\mathbb{P}(A \cap B) \leq \min(\mathbb{P}(A), \mathbb{P}(B))$ ) on obtient :

$$\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \leq \min(\mathbb{P}(A), \mathbb{P}(B)) - \mathbb{P}(A)\mathbb{P}(B).$$

Sans perte de généralité on peut supposer que  $\mathbb{P}(A) \leq \mathbb{P}(B)$ , ce qui nous donne :

$$\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \leq \mathbb{P}(A)\mathbb{P}(A)^2.$$

Or

$$\max(x - x^2, x \in \mathbb{R}) = \frac{1}{4} \quad (2.6)$$

d'où le résultat :

$$\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \leq \frac{1}{4} \quad (2.7)$$

D'un autre côté, nous avons :

$$\max(\mathbb{P}(A) + \mathbb{P}(B) - 1, 0) - \mathbb{P}(A)\mathbb{P}(B) \leq \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)$$

On distingue ainsi deux cas :

$$\mathbb{P}(A) + \mathbb{P}(B) \leq 1 \text{ ou } \mathbb{P}(A) + \mathbb{P}(B) \geq 1.$$

Si  $\mathbb{P}(A) + \mathbb{P}(B) \leq 1$  alors

$$\mathbb{P}(A) \leq 1 - \mathbb{P}(B).$$

On obtient donc :

$$-\mathbb{P}(A)\mathbb{P}(B) \leq \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B).$$

Et

$$-\mathbb{P}(A \cap B) + \mathbb{P}(A)\mathbb{P}(B) \leq \mathbb{P}(A)\mathbb{P}(B).$$

Mais puisque

$$\mathbb{P}(A) \leq 1 - \mathbb{P}(B).$$

Alors :

$$-\mathbb{P}(A \cap B) + \mathbb{P}(A)\mathbb{P}(B) \leq (1 - \mathbb{P}(B))\mathbb{P}(B) = \mathbb{P}(B) - \mathbb{P}(B)^2.$$

Nous obtenons finalement grâce à (2.6) :

$$-\frac{1}{4} \leq \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \quad (2.8)$$

Dans le cas où :

$$\mathbb{P}(A) + \mathbb{P}(B) > 1.$$

On aura

$$\mathbb{P}(A) + \mathbb{P}(B) - 1 - \mathbb{P}(A)\mathbb{P}(B) \leq \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B).$$

Mais

$$\min[x + y - 1 - xy, (x, y) \in x + y > 1, 0 \leq x \leq 1, 0 \leq y \leq 1] = -\frac{1}{4}.$$

Ce qui nous donne alors :

$$-\frac{1}{4} \leq \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \quad (2.9)$$

De (2.7),(2.8)et(2.9) on obtient :

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \frac{1}{4}.$$

Par conséquent :

$$0 \leq \alpha(\mathcal{A}, \mathcal{B}) \leq \frac{1}{4}.$$

## 2.2 Processus mélangeants

Considérons un processus stochastique  $(X_t, t \in T)$  ou  $T = \mathbb{R}, \mathbb{Z}$  ou  $\mathbb{N}$  défini sur l'espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  et  $\mathcal{F}_{-\infty}^t$  et  $\mathcal{F}_{t \in k}^{+\infty}$  sont des sous tribus  $k > 0$ .

**Définition 2.2.1** [20]  $X = (X_t, t \in T)$  est dit fortement mélangeant (ou  $\alpha$ -mélangeant) si :

$$\alpha(k) = \sup_{t \in T} \alpha_t(k) = \sup_{t \in T} \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+k}^{+\infty}} | \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) | \xrightarrow{k \rightarrow +\infty} 0. \quad (2.10)$$

**Définition 2.2.2** [20] *Le processus  $X = (X_t, t \in T)$  est dit absolument régulier (ou  $\beta$ -mélangeant) si :*

$$\beta(k) = \sup_{t \in T} \beta_t(k) = \sup_{t \in T} E \left( \sup_{A \in \mathcal{F}_{t+k}^{+\infty}} | \mathbb{P}(A/\mathcal{F}_{-\infty}^t) - \mathbb{P}(A) | \right) \xrightarrow{k \rightarrow +\infty} 0. \quad (2.11)$$

**Définition 2.2.3** [20] *Un processus stochastique  $X = (X_t, t \in T)$  est dit complètement régulier (ou  $\rho$ -mélangeant) si :*

$$\rho(k) = \sup_{t \in T} \rho_t(k) \xrightarrow{k \rightarrow +\infty} 0. \quad (2.12)$$

Où  $\rho_t(k) = \mathbb{P}(X(s), s \leq t; X(s), s \geq t+k)$

**Définition 2.2.4** [20] *Le processus est dit faiblement mélangeant (ou  $\phi$ -mélangeant ou uniformément mélangeant) si :*

$$\phi(k) = \sup_{t \in T} \phi_t(k) = \sup_{t \in T} \sup_{A \in \mathcal{F}_{-\infty}^t} | \mathbb{P}(A/B) - \mathbb{P}(A) | \xrightarrow{k \rightarrow +\infty} 0. \quad (2.13)$$

## Propriétés [20]

Nous avons les inégalités suivantes entre les coefficients de mélange pour tout  $k > 0$  :

i)  $2\alpha(k) \leq \beta(k) \leq \phi(k), \alpha(k) \leq \frac{1}{4}$

ii) Dans le cas d'un processus réel nous avons :

$$2\alpha(k) \leq \rho(k) \leq 2\sqrt{\phi(k)}.$$

iii) Dans le cas d'un processus réel gaussien stationnaire on a :

$$\rho(k) \leq 2\pi\alpha(k).$$

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

### 2.2.1 Schéma pour les propriétés de mélange d'un processus

$$\psi\text{-mélangeant} \left\{ \begin{array}{l} \Rightarrow \\ \Leftarrow \end{array} \right\} \phi\text{-mélangeant} \left\{ \begin{array}{l} \Rightarrow \\ \Leftarrow \end{array} \right\} \left\{ \begin{array}{l} \beta\text{-mélangeant} \left\{ \begin{array}{l} \Rightarrow \\ \Leftarrow \end{array} \right\} \\ \Leftarrow \\ \rho\text{-mélangeant} \left\{ \begin{array}{l} \Rightarrow \\ \Leftarrow \end{array} \right\} \end{array} \right\} \alpha\text{-mélangeant}$$

## 2.3 Variables aléatoires mélangeantes

**Définition 2.3.1** [10] Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires définies sur le même espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Le coefficient de mélange fort ou d'alpha-mélange est donné par :

$$\alpha(n) = \sup_k \sup \{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, A \in \mathcal{F}_1^k(X), B \in \mathcal{F}_{k+n}^\infty(X), k \in \mathbb{N} \}. \quad (2.14)$$

Où  $F_i^k(X)$  désigne la tribu des événements engendrés par les  $X_j, i \leq j \leq k$ [10].

**Définition 2.3.2** On dit qu'une famille  $X_i, i \in \mathbb{N}^*$  est fortement mélangeante ou alpha-mélangeante, si l'on a :

$$\lim_{n \rightarrow \infty} \alpha(n) = 0 \quad (2.15)$$

On peut définir deux types de mélange fort :

**Définition 2.3.3** On dit que la famille  $X_i, i \in \mathbb{N}^*$  est algébriquement alpha-malangeante, s'il existe deux constantes positives non nulles  $C$  et  $a$  telles que les coefficients de mélange vérifient :

$$\alpha(n) = Cn^{-a}. \quad (2.16)$$

**Définition 2.3.4** S'il existe deux constantes  $C > 0$  et  $t \in ]0, 1[$  telles que les coefficients de mélange vérifient :

$$\alpha(n) = Ct^n. \quad (2.17)$$

On dit que la famille  $X_i, i \in \mathbb{N}^*$  est géométriquement alpha-mélangeante[14].

**Proposition 2.3.1** (Billingsly(1968)) Si  $(X_n)_{n \geq 1}$  est une suite  $\alpha$ -mélangeante, alors  $(f(X_n))_{n \geq 1}$  est aussi  $\alpha$ -mélangeante pour toute fonction mesurable  $f$ [15].

**Proposition 2.3.2** Supposons que  $\Omega'$  est un espace semi-normé, muni de la semi norme  $\|\cdot\|$ , on a :

- i)  $(X_n)_{n \in \mathbb{Z}}$  est  $\alpha$ -mélangeante  $\Rightarrow (\|X_n\|)_{n \in \mathbb{Z}}$  est  $\alpha$ -mélangeante.
- ii) Si  $(X_n)_{n \in \mathbb{Z}}$  est géométriquement  $\alpha$ -mélangeante (resp. arithmétiquement  $\alpha$ -mélangeante), alors  $(\|X_n\|)_{n \in \mathbb{Z}}$  est géométriquement  $\alpha$ -mélangeante (resp. arithmétiquement  $\alpha$ -mélangeante) de même ordre[8].

**Proposition 2.3.3** Supposons que  $\Omega'$  est un espace semi-métrique avec  $d$  semi-métrique et soit  $x$  un élément fixe de  $\Omega'$  mettons  $X_i = d(X_n, x)$  on a :

- i)  $(X_n)_{n \in \mathbb{Z}}$  est  $\alpha$ -mélangeante  $\Rightarrow (d(X_n, x))_{n \in \mathbb{Z}}$  est  $\alpha$ -mélangeante.
- ii) Si  $(X_n)_{n \in \mathbb{Z}}$  est géométriquement  $\alpha$ -mélangeante (resp. arithmétiquement  $\alpha$ -mélangeante), alors  $(d(X_n, x))_{n \in \mathbb{Z}}$  est géométriquement  $\alpha$ -mélangeante (resp. arithmétiquement  $\alpha$ -mélangeante) de même ordre[8].

### 2.3.1 Quelques statistique mélangeantes

Considérons un échantillon fonctionnel statistique  $X_1, \dots, X_n$  composé de variables fonctionnelles ayant chacune la même distribution qu'une variable fonctionnelle générique  $X$ . Cette variable fonctionnelle  $X$  prend des valeurs dans un espace semi-métrique  $(E, d)$ .  $X_1, \dots, X_n$  une suite de variables aléatoires de taille  $n$  consécutifs (du même loi que  $X$ ) d'une séquence satisfaisant à la condition de mélange fort ( $\alpha$ -mélange). On notera par  $\alpha(n)$  les coefficient de mélange associés a cette séquence de mélange fonctionnelle.[8]

**Définition 2.3.5** [8] Soit  $k$  une fonction du noyau,  $h$  étant un paramètre de lissage positif. On note par :

$$D_i = \frac{k\left(\frac{d(X_i, x)}{h}\right)}{E\left(k\left(\frac{d(X_i, x)}{h}\right)\right)}.$$

**Lemme 2.3.1** [8]  $(D_i)_{i=1, \dots, n}$  sont  $n$  termes consécutifs d'une suite de variables aléatoires  $\alpha$ -mélangeante si en plus, les coefficients de mélange de la suite de variables fonctionnelles  $(X_i)_{i=1, \dots, n}$  sont géométriques (resp. arithmétiques), alors ceux de  $(D_i)_{i=1, \dots, n}$  le sont aussi.

## 2.4 Inégalités dans le cas des variables mélangeantes

Les statistiques non paramétriques pour les processus de mélange à valeur réelle ont fait l'objet de beaucoup d'attention au cours des dernières décennies. et il s'avère que ces avancées statistiques ont été liées aux développements d'outils probabilistes. pour des suites de variables mélangeantes il existe essentiellement deux types d'outils utilisés à des fins non paramétriques :

- Les inégalités de covariance.
- Les inégalités exponentielle.

Ces inégalités seront utiles pour les variables évaluées dans des espaces semi-métriques. Commençons d'abord par l'inégalité de covariance.

### 2.4.1 Inégalités de covariance

Il existe de nombreux ouvrages sur les inégalités de covariance pour les variables mélangeantes. Ces inégalités diffèrent à la fois du type de condition de mélange introduit et du type d'hypothèses vérifiées par les variables. Nous présentons ici des inégalités nécessaire pour la covariance de mélange des variables aléatoires. Soit  $X$  et  $Y$  deux variables aléatoires mesurables par rapport  $\mathcal{A}$  et  $\mathcal{B}$  respectivement. Rappelons que  $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$  pour  $p < \infty$  et  $\|X\|_\infty = \text{ess - sup } |X|$  [14].

#### **Théorème 2.4.1** [14]

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

**Remarque 2.4.1** L'inégalité i est due à Wolkowski et Rozanov(1959) dans le cas  $p = q = \infty$  et à Davydov(1970) dans la forme actuelle. Davydov(1968) l'a prouvé sous sa forme la plus faible , où la constante 8 est remplacée par 12. L'inégalité ii, due à Ibragimov(1962), se retrouve dans Billingsley(1968). L'inégalité iii est due à Blum, Hanson et Koopmans(1963). L'inégalité iv est évidente compte tenu de la définition de  $\rho$ . [6]

**Preuves [6]**

i) Nous prouvons ce résultat en trois étapes :

(1) Premier étape :  $\|X\|_\infty < \infty$ ,  $\|Y\|_\infty < \infty$

**Lemme 2.4.2** *Inégalité de covariance d'Ibragimov(1962) :*

$$|\text{cov}(X, Y)| \leq 4\alpha(\mathcal{A}, \mathcal{B})\|X\|_\infty\|Y\|_\infty.$$

**Démonstration**

Soit  $A = \text{singe}(E(X/\mathcal{B}) - E(X))$  et  $B = \text{singe}(E(Y/\mathcal{A}) - E(Y))$  par la propriété de l'espérance conditionnelle on a :

$$\begin{aligned} |\text{cov}(X, Y)| &= |EXY - EXEY| = |E(XE(Y/\mathcal{A}) - E(Y))| \\ &\leq \|X\|_\infty E|E(Y/\mathcal{A}) - EY| \\ &\leq \|X\|_\infty E|B(Y/\mathcal{A}) - EY| \\ &\leq \|X\|_\infty |E(BY - EBEY)| \end{aligned}$$

L'argument similaire donne :

$$|\text{cov}(X, Y)| \leq \|X\|_\infty\|Y\|_\infty |EAB - EAEB|$$

Soit  $A^+ = \{a = 1\}$ ,  $A^- = \{a = -1\}$ ,  $B^+ = \{b = 1\}$ ,  $B^- = \{b = -1\}$  il est clair que  $A \in \mathcal{F}_{-\infty}^k$ ,  $B \in \mathcal{F}_{k+n}^\infty$

En utilisant la définition de  $\alpha$  - *mélange*, on obtient :

$$\begin{aligned} |EAB - EAEB| &= |[\mathbb{P}(A^+ \cap B^+) - \mathbb{P}(A^+)\mathbb{P}(B^+)] \\ &\quad + [\mathbb{P}(A^- \cap B^-) - \mathbb{P}(A^-)\mathbb{P}(B^-)] \\ &\quad + [\mathbb{P}(A^+ \cap B^-) - \mathbb{P}(A^+)\mathbb{P}(B^-)] \\ &\quad + [\mathbb{P}(A^- \cap B^+) - \mathbb{P}(A^-)\mathbb{P}(B^+)]|. \end{aligned}$$

D'où

$$|cov(X, Y)| \leq 4\alpha(\mathcal{A}, \mathcal{B})\|X\|_\infty\|Y\|_\infty$$

(2)  $\|X\|_p < \infty$ ,  $\|Y\|_\infty < \infty$ ,  $1 < p < \infty$

soit  $\bar{X} = X\mathbf{1}_{|x| \leq u}$ ,  $\underline{X} = X\mathbf{1}_{|x| > u}$  et  $X = \bar{X} + \underline{X}$

ainsi,

$$|cov(X, Y)| = |cov(\bar{X}, Y) + cov(\underline{X}, Y)| \leq 4\alpha(\mathcal{A}, \mathcal{B})u\|Y\|_\infty + 2\|Y\|_\infty\mathbb{E}|\underline{X}|$$

Maintenant, L'inégalité de Markov mène à :

$$\mathbb{E}|\underline{X}| \leq 2\frac{u\mathbb{E}|X|^p}{u^p}$$

Et :

$$\frac{\mathbb{E}|X|^p}{u^p} = \alpha(\mathcal{A}, \mathcal{B}).$$

Puis :

$$|cov(X, Y)| \leq 6\alpha^{1-1/p}(\mathcal{A}, \mathcal{B})\|Y\|_\infty\|X\|_p$$

(3)  $\|X\|_p < \infty$ ,  $\|Y\|_q < \infty$ ,  $\frac{1}{p} + \frac{1}{q} < 1$

Soit  $\bar{Y} = Y\mathbf{1}_{|y| \leq v}$ ,  $\underline{Y} = Y\mathbf{1}_{|y| > v}$  et  $Y = \bar{Y} + \underline{Y}$  par conséquent, de manière analogue :

$$|cov(X, Y)| \leq 6\alpha^{1-1/p}(\mathcal{A}, \mathcal{B})v\|X\|_p + 2v\mathbb{E}|\underline{Y}|$$

Donne :

$$\frac{\mathbb{E}|Y|^q}{v^q} = \alpha^{1-1/p}(\mathcal{A}, \mathcal{B})$$

$$|cov(X, Y)| \leq 8\alpha^{1/r}(\mathcal{A}, \mathcal{B})\|Y\|_q\|X\|_p$$

Preuve alternative pour (2) et (3) On aura :

$$\bar{X} = X\mathbf{1}_{|x| \leq u}$$

et

$$\bar{Y} = Y\mathbf{1}_{|y| \leq v}$$

. Et :

$$|cov(X, Y)| = |cov(\bar{X}, \bar{Y}) + cov(\bar{X}, \underline{Y}) + cov(\underline{X}, \bar{Y}) + cov(\underline{X}, \underline{Y})|$$

Qui découle des identités  $X = \bar{X} + \underline{X}$  et  $Y = \bar{Y} + \underline{Y}$

Maintenant, l'inégalité de Markov mène, avec  $a = p(1 - 1/r)$  et  $b = q(1 - 1/r)$  à :

$$cov(\bar{X}, \underline{Y}) \leq 2uv \frac{\mathbb{E}|Y|^q}{v^q}.$$

$$|cov(\underline{X}, \bar{Y})| \leq 2uv \frac{\mathbb{E}|X|^p}{u^p}.$$

et

$$|cov(\underline{X}, \underline{Y})| \leq 2uv \left( \frac{\|X\|_p}{u} \frac{\|Y\|_q}{v} \right)^{\frac{r}{r-1}}.$$

Par les inégalités

$$|cov(X, Y)| \leq 2\|X\|_a \|Y\|_b \text{ et } \mathbb{E}|X|^a \leq u^a \frac{\mathbb{E}|X|^p}{u^p}.$$

Par conséquent :

$$|cov(X, Y)| \leq 2uv \left( 2\alpha(\mathcal{A}, \mathcal{B}) + \frac{\mathbb{E}|Y|^q}{v^q} + \frac{\mathbb{E}|X|^p}{u^p} + \left\{ \frac{\|X\|_p}{u} \frac{\|Y\|_q}{v} \right\}^{\frac{r}{r-1}} \right).$$

Choisir la meilleure constante  $u$  et  $v$  dans cette expression conduirait à une constante inférieure à 8.

i) La valeur 8 obtenue pour

$$\frac{\mathbb{E}|Y|^q}{v^q} = \frac{\mathbb{E}|X|^p}{u^p} = 2\alpha(\mathcal{A}, \mathcal{B}).$$

ii) Les variables aléatoires sont d'abord supposées simples, par exemple écrire :

$$X = \sum_{i=1}^I \mathbf{1}_{A_i} \text{ et } Y = \sum_{j=1}^J \mathbf{1}_{B_j}.$$

$$\|X\|_p = \left( \sum_{i=1}^I \|X_i\|_p \mathbb{P}(A_i) \right)^{1/p}.$$

Et

$$cov(X, Y) = \sum_{i=1}^I \sum_{j=1}^J X_i Y_j \mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i) \mathbb{P}(B_j)$$

Ainsi

$$cov(X, Y) = \sum_{i=1}^I X_i (\mathbb{P}(A_i))^{1/p} \sum_{j=1}^J Y_j \mathbb{P}(B_j/A_i) - \mathbb{P}(B_j) (\mathbb{P}(A_i))^{1/q}.$$

à présent

$$\text{cov}(X, Y) \leq \|X\|_p \left( \sum_{i=1}^I \mathbb{P}(A_i) \left| \sum_{j=1}^J |Y_j| |\mathbb{P}(B_j/A_i) - \mathbb{P}(B_j)| \right|^q \right)^{1/q}$$

Mais

$$\begin{aligned} \left| \sum_{j=1}^J |Y_j| |\mathbb{P}(B_j/A_i)| \right|^q &= \left| \sum_{j=1}^J |Y_j| |\mathbb{P}(B_j/A_i) - \mathbb{P}(B_j)|^{1/q} |\mathbb{P}(B_j/A_i) - \mathbb{P}(B_j)|^{1/q} \right|^q \\ &\leq \sum_{j=1}^J |Y_j|^q |\mathbb{P}(B_j) + \mathbb{P}(B_j/A_i)| \left| \sum_{j=1}^J |\mathbb{P}(B_j) - \mathbb{P}(B_j/A_i)| \right|^{q/p} \end{aligned}$$

Par conséquent

$$|\text{cov}(X, Y)|^q \leq 2 \|X\|_p^q \|Y\|_q^q \left\{ \sup \left| \sum_{j=1}^J |\mathbb{P}(B_j) - \mathbb{P}(B_j/A_i)| \right| \right\}^{q/p}$$

Laisse maintenant :  $C_i^+$  (resp.  $C_i^-$ ) est l'union de  $\mathcal{B}$  – ensemble avec :

$$\mathbb{P}(B_j) - \mathbb{P}(B_j/A_i) \geq 0 \text{ (resp. } < 0)$$

Puis

$$\begin{aligned} \sum_{j=1}^J |\mathbb{P}(B_j) - \mathbb{P}(B_j/A_i)| &= [\mathbb{P}(C_i^+) - \mathbb{P}(C_i^+/A_i)] - [\mathbb{P}(C_i^-) - \mathbb{P}(C_i^-/A_i)] \\ &\leq 2\phi(\mathcal{A}, \mathcal{B}) \end{aligned}$$

les inégalités précédentes sont prouvées pour des variables aléatoires simples, ce qui permet de se rapprocher des variables aléatoires initiales dans  $L^p$  (resp.  $L^q$ ), par des erreurs simples, conduisant aux résultats souhaités. La preuve est analogue pour  $p = 1$  et  $q = \infty$ .

**iii)** On suppose d'abord que les variables aléatoires sont simples. laisser pour les partitions mesurables  $A_i B_j$  :

$$X = \sum_{i=1}^I \mathbf{1}_{A_i} \text{ et } Y = \sum_{j=1}^J \mathbf{1}_{B_j}.$$

Nous avons :

$$\|X\|_1 = \sum_{i=1}^I \mathbb{P}(A_i) \text{ et } |\text{cov}(X, Y)| = \left| \sum_{i=1}^I \sum_{j=1}^J X_i Y_j \mathbb{P}(A_i) \mathbb{P}(B_j) - \mathbb{P}(A_i \cap B_j) \right|.$$

Par conséquent :

$$|cov(X, Y)| \leq \psi(\mathcal{A}, \mathcal{B}) \sum_{i=1}^I \sum_{j=1}^J |X_i| |Y_j| \mathbb{P}(A_i) \mathbb{P}(B_j) = \psi(\mathcal{A}, \mathcal{B}) \|X\|_1 \|Y\|_1$$

Les inégalités précédentes sont prouvées pour des variables aléatoires simples. Le rapprochement des variables initiales dans  $L^1$  par des variables simples conduit au resultat souhaité.

## 2.4.2 Inégalités exponentielles

Présentons maintenant quelques inégalités exponentielles pour les somme partielles d'une suite de variables aléatoires mélangeantes  $(X_n)_{n \in \mathbb{Z}}$ . D'une certaine manière, même si les formules ne sont pas complètement comparables, les resultat présentés ci-dessous sont des extensions dépendantes. Au cours des vingt dernières années, la littérature sur les inégalités probabilistes exponentielles pour le mélange de séquences était directement liée aux avancées de la statistique non paramétrique pour les données dépendantes.

Introduisons la notation  $S_n^2 = \sum_{i=1}^n \sum_{j=1}^n |cov(X_i, Y_j)|$

**Proposition 2.4.1** [8] *Supposons que  $(X_n)_{n \in \mathbb{N}^*}$  sont identiques et soient arithmétiquement  $\alpha$  – mélange avec un taux  $a > 1$*

i) *S'il existe  $p > 2$  et  $M > 0$  tels que  $\forall t > M$ ,  $\mathbb{P}(|X_1| > t) \leq t^{-p}$ , alors nous avons pour tout  $r \geq 1$  et  $\epsilon > 0$  et pour certains  $c < \infty$  :*

$$\mathbb{P}\left(\left|\sum_{i=1}^n X_i\right| > \epsilon\right) \leq c \left\{ \left(1 + \frac{\epsilon^2}{r S_n^2}\right)^{-r/2} + nr^{-1} \left(\frac{r}{\epsilon}\right)^{(a+1)p/(a+p)} \right\}$$

ii) *S'il existe  $M < \infty$  tel que  $|X_1| \leq M$ , nous avons pour tout  $r \geq 1$  et pour tout  $c < \infty$  :*

$$\mathbb{P}\left(\left|\sum_{i=1}^n X_i\right| > \epsilon\right) \leq c \left\{ \left(1 + \frac{\epsilon^2}{r S_n^2}\right)^{-r/2} + nr^{-1} \left(\frac{r}{\epsilon}\right)^{a+1} \right\}$$

Nous proposons de nouvelles formulations de cette inégalité (corollaire (2.4.1), corollaire (2.4.2)) qui sont directement adapté aux applications statistiques non paramétriques. Les deux corollaire diffère par le type de coefficient de mélange : arithmétique ou géométrique.

**Corollaire 2.4.1** [8] *Supposons que les variables dépendent de  $n$  (c'est à dire  $X_i = X_{i,n}$ ) et que  $X_1, \dots, X_n$  sont  $n$  termes successifs d'une séquence de mélange avec des coefficients rithmétiques*

d'ordre  $a > 1$ .

Considérons la séquence déterministe  $U_n = n^{-2} S_n^2 \log n$ . Supposons également que l'une des deux hypothèses soit satisfaite :

i)  $\exists p > 2, \exists \theta > 2, \exists M = M_n < \infty$  tel que  $\forall t > M_n$  :

$$\mathbb{P}(|X_1| > t) \leq t^{-p}$$

et

$$S_n^{-\frac{(a+1)p}{a+p}} = o(n^{-\theta}).$$

Ou :

ii)  $\exists M = M_n < \infty$  et  $\exists \theta > 2$  tel que  $|X_1| \leq M_n S_n^{-(a+1)} = o(n^{-\theta})$  ensuite nous avons :

$$\frac{1}{n} \sum_{i=1}^n X_i = O_{a.co}(\sqrt{U_n}).$$

### Preuve

Notons d'abord que les conditions sur  $S_n$  assurent que  $\lim_{n \rightarrow \infty} U_n = 0$ . Nous allons prouver les deux assertions à la fois en utilisant la notation :

$$q = \begin{cases} (a+1)p/(a+p) & \text{dans les conditions de i} \\ a+1 & \text{dans les conditions de ii} \end{cases}$$

prendre  $r = (\log n)^2$  et appliquer la proposition (2.4.1) pour obtenir pour tout  $\epsilon_0 > 0$  :

$$\begin{aligned} \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i\right| > \epsilon_0 \sqrt{U_n}\right) &\leq c \left(1 + \frac{\epsilon_0^2 n^2 U_n}{(\log n)^2 S_n^2}\right)^{-\frac{(\log n)^2}{2}} + n(\log n)^{-2} \left(\frac{\log n}{n \epsilon_0 \sqrt{U_n}}\right)^q \\ &\leq c \left(1 + \frac{\epsilon_0^2}{\log n}\right)^{-\frac{(\log n)^2}{2}} + n(\log n)^{-2} \left(\frac{\sqrt{\log n}}{\epsilon_0 S_n}\right)^q \end{aligned}$$

En utilisant le fait que  $\log(1+x) = x - x^2/2 + o(x^2)$  quand  $x \rightarrow 0$ , on aura :

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i\right| > \epsilon_0 \sqrt{U_n}\right) \leq c \exp\left[\frac{-\epsilon_0^2 \log n}{2}\right] + n(\log n)^{-2+q/2} S_n^{-q} \epsilon_0^{-q}$$

Enfin, la condition sur  $S_n$  nous permet d'obtenir directement qu'il existe des  $\epsilon_0$  et des  $n_0 > 0$  tel que :

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n X_i\right| > \epsilon_0\sqrt{U_n}\right) \leq cn^{-1-n_0}.$$

**Corollaire 2.4.2** [8] *Supposons que les variables dépendent que de  $n$  (c-à-d  $X_i = X_{i,n}$ ) et que  $X_1, \dots, X_n$  sont  $n$  termes successifs d'une séquence de mélange avec des coefficients géométrique. Considérons la séquence déterministe  $U_n = n^{-2}S_n^2 \log n$ . Supposons également que l'une des deux hypothèses soit satisfaite :*

i)  $\exists p > 2, \exists \theta > 2/p, \exists M = M_n < \infty$  tel que  $\forall t > M_n$  :

$$\mathbb{P}(|X_1| > t) \leq t^{-p}$$

et

$$S_n^{-1} = o(n^{-\theta})$$

ii)  $\exists M = M_n < \infty$  et  $\exists \theta > 2$  tel que  $|X_1| \leq M_n S_n^{-1} = o(n^{-\theta})$  ensuite nous avons :

$$\frac{1}{n}\sum_{i=1}^n X_i = O_{a.co}(\sqrt{U_n}).$$

### Preuve

Cette preuve est évidente puisque la décroissance géométrique des coefficients nous permet d'appliquer le corollaire (2.4.1) pour toute valeur de  $a$ .

# CHAPITRE 3

## SIMULATION

### Introduction

Dans ce chapitre, nous proposons une étude numérique, à l'aide de logiciel Matlab, pour illustrer nos résultats précédents concernant les variables aléatoires mélangeantes et effets sur l'estimation.

### 3.1 Simulation

Afin d'examiner l'effet du mélange sur l'estimation on a généré :

1. Un échantillon aléatoire de taille  $n$  de la variable aléatoire  $X$  gaussienne ( $\mathcal{N}(0, 1)$ ).
2. Un échantillon  $\alpha$  – *mélangeant* de taille  $n$  de la variable aléatoire  $Y$ , de loi  $\mathcal{N}(0, 1)$  pour différentes tailles d'échantillon  $n \in \{10, 100, 1000\}$ .

Nous nous sommes intéressés à l'influence du mélange sur les estimateurs  $\bar{X}$  et  $S_n^2$  de la moyenne et variance respectivement :

$$\begin{cases} \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i. \\ S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2. \end{cases}$$

Dans les tableaux (3.1) à (3.3), nous avons noté par :

$\bar{X}$  : La moyenne empirique dans le cas des variables indépendantes.

$\bar{Y}$  : La moyenne empirique dans le cas des variables  $\alpha$  – *mélangeantes*.

$S$  : L'écart type empirique dans le cas des variables indépendantes.

$S'$  : L'écart type empirique dans le cas des variables  $\alpha$  – *mélangeantes*.

$\alpha$  : Le coefficient de mélange.

$\varepsilon$  : Vecteur des erreurs absolues :

$$\varepsilon_0 = \begin{pmatrix} |\mu - \bar{X}| \\ |\sigma - S| \end{pmatrix}, \quad \varepsilon_\alpha = \begin{pmatrix} |\mu - \bar{Y}| \\ |\sigma - S'| \end{pmatrix}$$

n=10	<i>ind</i>	$Mél_{\alpha=0.1}$	$Mél_{\alpha=0.15}$	$Mél_{\alpha=0.2}$	$\varepsilon_0$	$\varepsilon_{0.1}$	$\varepsilon_{0.15}$	$\varepsilon_{0.2}$
$\mu = 0$	$\bar{X} = -0.0801$	$\bar{Y} = 0.1150$	$\bar{Y} = 0.1198$	$\bar{Y} = 0.1262$	0.0801	0.1150	0.1198	0.1262
$\sigma = 1$	$S = 0.8429$	$S' = 0.3128$	$S' = 0.3125$	$S' = 0.3123$	0.1571	0.6872	0.6875	0.6877

TABLE 3.1 – Les erreurs absolues dans le cas d'un échantillon de taille 10

n=100	<i>ind</i>	$Mél_{\alpha=0.1}$	$Mél_{\alpha=0.15}$	$Mél_{\alpha=0.2}$	$\varepsilon_0$	$\varepsilon_{0.1}$	$\varepsilon_{0.15}$	$\varepsilon_{0.2}$
$\mu = 0$	$\bar{X} = -0.0091$	$\bar{Y} = 0.0098$	$\bar{Y} = 0.0121$	$\bar{Y} = 0.0124$	0.0091	0.0098	0.0121	0.0124
$\sigma = 1$	$S = 0.99219$	$S' = 0.1017$	$S' = 0.1014$	$S' = 0.110$	0.0078	0.8983	0.8986	0.89

TABLE 3.2 – Les erreurs absolues dans le cas d'un échantillon de taille 100

La variation de  $\alpha$  sur l'axe des abscisse.

La variation des erreurs entre les moyennes théoriques et les moyennes empiriques sur l'axe des ordonné.

La variation de  $\alpha$  sur l'axe des abscisse.

La variation des erreurs entre les écarts type théoriques et les écarts type empiriques sur l'axe des ordonné.

n=1000	<i>ind</i>	$Mél_{\alpha=0.1}$	$Mél_{\alpha=0.15}$	$Mél_{\alpha=0.2}$	$\varepsilon_0$	$\varepsilon_{0.1}$	$\varepsilon_{0.15}$	$\varepsilon_{0.2}$
$\mu = 0$	$\bar{X} = 0.00193$	$\bar{Y} = 0.00195$	$\bar{Y} = 0.00199$	$\bar{Y} = 0.0021$	0.00193	0.00195	0.00199	0.0021
$\sigma = 1$	S=0.98928	S'=0.0329	S'=0.0327	S'=0.0324	0.0107	0.9671	0.9673	0.9676

TABLE 3.3 – Les erreurs absolues dans le cas d’un échantillon de taille 1000

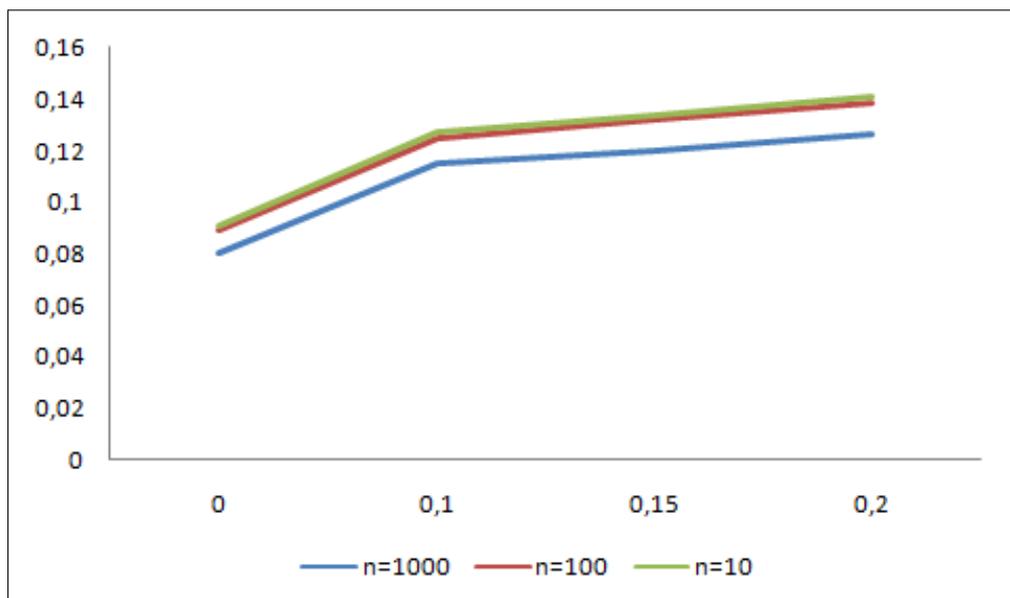


FIGURE 3.1 – Les erreurs entre les moyennes théoriques et les moyennes empiriques.

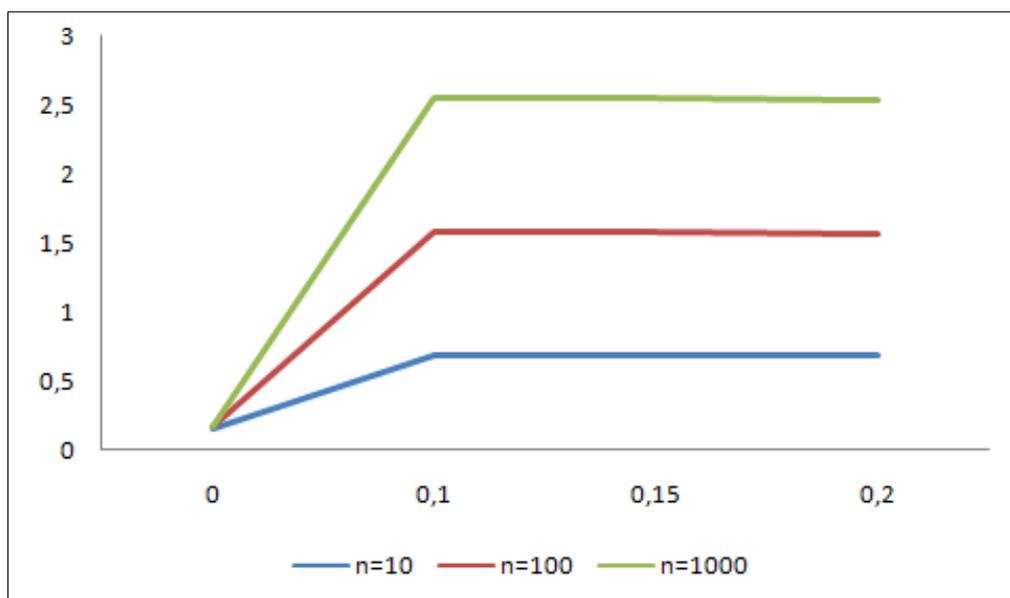


FIGURE 3.2 – Les erreurs entre les variances théoriques et les variances empiriques.

## Interpretations

D'après les trois tableaux (3.1), (3.2) et (3.3), On remarque que :

- Plus le  $\alpha$  (coefficient de mélange) est grand, plus la moyenne et la variance empiriques s'éloignent de la moyenne exacte  $\mu$  et de la variance  $\sigma^2$  respectivement (voir les figures (3.1) et (3.2)).
- Lorsque la taille de l'échantillon considéré est petite ( $n=10$ ) l'effet est significatif. Pour des grandes tailles ( $n=100$ ,  $n=1000$ ) l'influence de l'effet de mélange est moindres sur la moyenne empirique (voir la figure (3.1)). Par contre la variance empirique s'éloigne de la variance théorique (voir la figure (3.2)).
- Les estimateurs de maximum de vraisemblance de la moyenne et variance dans le cas normale centré-réduite sont sensibles aux mélanges.
- le mélange à plus d'effet sur l'estimation de la variance.

## CONCLUSION GÉNÉRALE

Dans ce memoire, nous avons rappelé la définition de la notion d'indépendance des variables aléatoires, nous avons donné quelques théorèmes, propriétés et exemples concernant l'estimation.

Nous avons introduit le concept de dépendence appelé le mélange, nous avons donné quelques définitions et propriétés, et nous avons aussi étudié les différents types de mélanges.

Enfin, nous avons réalisé une etude de simulation à l'aide de logiciel Matlab, concernant les variables aléatoires mélangeantes et effets sur l'estimation. On a constaté que les erreurs obtenues dans le cas d'un échantillon  $\alpha$  – *mélangeant* sont plus significatives par rapport aux erreurs obtenues dans le cas indépendant. On a constaté aussi que plus  $\alpha$  s'approche de zéro, plus la moyenne et la variance empirique s'approche de la moyenne théorique  $\mu$  et variance théorique  $\sigma$  respectivement, Plus la taille de l'échantillon est grand plus l'erreurs absolue entre la moyenne empirique et la moyenne théorique est petite dans les deux cas, indépendant et  $\alpha$ -mélangeant. Nous avons également constaté que le mélange à plus d'effet sur l'estimation de la variance.

## BIBLIOGRAPHIE

- [1] Appel,W.(2015). Probabilités :Pour les non-probabilistes. 2<sup>e</sup>éd.Paris,735p. ISBN 978-2-35141-326-5.
- [2] Bernstein, S.(1946). Probability Theory, 4th ed. (in russian), ed. M.L. Gostechizdat.
- [3] Bismans,F.(2016). Probabilités et statistique inférentielle :Prélud à l'économétrie. Ellipess, Paris. ISBN 9782340-010017.
- [4] Benslimane,S.(2017). Approximation stochastique :Inégalités exponentielles et etude numérique. Mémoire de master. Université A.MIRA. Bejaia.
- [5] Dauxois,J.(2013). Cours de probabilités.
- [6] Daukhan,P.(1994). Mixing :Properties and exemples. Springer-Verlag,France.ISBN-13 :978-0-387-94214-8.
- [7] Fuk, D.K, Nagaev, S.V.(1971). Probability inequalities for sums of independant random variables. Theory Probab. Appl., 16, 643-660.
- [8] Ferraty,F et Vieu,P.(2006). Nonparametric Functional Data Analysis : Teory and practice. Springer. ISBN-10 :0-387-30369-3.
- [9] Fromont,M. Modèles de régression linéaire. Master statistique appliquée, université Rennes.
- [10] Ferrani,Y.(2014). Sur l'estimation non paramétrique de la densité du mode dans le modèles des données incomplètes et associées. thèse de Doctorat USTHB Alger.
- [11] Gaudoin,O.(2002). Principes et methodes statistiques :Notes de cours. Ensimag-2éme année,145p.
- [12] Giraud,D.(2015). Théorèmes limites de lathéorie des probabilités dans les systèmes dynamiques. thèse de Doctorat. Université de Rouen .france.

- [13] Hamisultane,H.(2002). Econometrie. Licence,France.cel-01261163.
- [14] Kbaili,S.(2012). La notion de mélange dans les semi-groupes. Mémoire de Magister, Spécialité mathématique. Université A.MIRA.Bejaia .
- [15] Mechab,W.(2017). Analyse non paramétrique par régression relative ,thèse de Doctorat,Université Djillali Liabes Sidi Bel Abbes.
- [16] Mokkaedem,A.(1990) Propriétés de mélange des processus autorégressifs polynomiaux. Université du Paris-sud .France
- [17] Nagaev, S.(1997). Some refinements of probabilistic and moment inequalities. Teor :Veroyatnost. i Primenen (in russian), 42 (4), 832-838 .
- [18] Nagaev, S.(1998). Some refinements of probabilistic and moment inequalities (English-translation). Theory Probab. Appl., 42 (4), 707-713 .
- [19] Perraudin,C.(2004). Le modèle de régression linéaire.
- [20] Tra Terki,N.(2012). Mélange des processus linéaires et autorégressifs. Mémoire de master. Spécialité mathématique. Université A.MIRA.Bejaia.
- [21] Tugaut,J.(2017). Statistiques :Quelques méthodes d'estimations. Télécom Saint-étienne,32p.
- [22] Uspensky, J.(1937). Introduction to mathematical probability. McGraw-Hill, New York .
- [23] Villemonais, D.(2017). Probabilités :Indépendance de variables aléatoires. IECL-Ecole des Mines de Nancy.

# Résumé

---

L'objectif de ce mémoire est l'étude des variables aléatoires mélangeantes, et l'effet du mélange dans l'estimation.

Nous avons rappelé de la notion d'indépendance des variables aléatoires, nous avons donné quelques théorèmes, propriétés et exemples concernant l'estimation.

Nous avons présenté la notion de mélange, qui est un concept de dépendance, et donné quelques définitions et propriétés, ainsi que les différents mélanges.

Enfin, nous avons réalisé une étude de simulation à l'aide de logiciel Matlab,

Mots clés : l'indépendance, le mélange , estimation, inégalité exponentielle , inégalités de covariance, Simulation.

# Abstract

---

The objective of this memory is the study of mixing random variables, and the effect of the mixture in the estimation. We have recalled the notion of independence random variables, We have given some theorems, properties and examples concerning the estimation.

We have presented the concept of mixing, which is a concept of dependence, and given some definitions and properties, as well as the different mixtures.

Finally, we performed a simulation study using matlab software.

keywords : Dependence, mixing, estimation, exponential inequality, inequality of covariance, simulation.