



**UNIVERSITE DE OUAGADOUGOU**

**THESE**

*Présentée à l'Unité de Formation et de Recherche  
En Sciences Exactes et Appliquées  
Pour obtenir le grade*

**Docteur d'Etat Es Sciences  
Spécialité : Mathématiques Appliquées  
Option : Statistique**

Par

**Dembo GADIAGA**

*Sujet de la thèse*

**TESTS FONCTIONNELS D'AJUSTEMENT ET DE NON  
INFLUENCE POUR DES VARIABLES ALEATOIRES  
DEPENDANTES**

*Soutenu le 10 Novembre 2003 devant le Jury composé de :*

|                  |                             |            |                           |
|------------------|-----------------------------|------------|---------------------------|
| <b>Président</b> | <b>: Denis BOSQ,</b>        | Professeur | Université Paris VI       |
| <b>Membres</b>   | <b>: Akry COULIBALY,</b>    | Professeur | Université de Ouagadougou |
|                  | <b>. Moussa OUATTARA,</b>   | Professeur | Université de Ouagadougou |
|                  | <b>. Albert OUEDRAOGO,</b>  | Professeur | Université de Ouagadougou |
|                  | <b>. Fortunato PESARIN,</b> | Professeur | Université de Padoue      |

*Je dédie ce mémoire*

*A la mémoire de mon père **Diaoudi GADIAGA** et ma mère*

***Laya GADIAGA**, que je n'ai malheureusement pas eu la chance de bien connaître car ayant disparu très tôt.*

*A la mémoire de ma fille **Djénéba**.*

## **REMERCIEMENTS**

Le Professeur Albert Ouédraogo m'a accepté dans son équipe de recherche et, depuis, il n'a cessé de m'encourager et de me prodiguer des conseils; qu'il me soit permis de lui exprimer toute ma reconnaissance et ma gratitude.

Le Professeur Denis Bosq m'a proposé le sujet de cette thèse et, malgré la distance, a su motiver mon intérêt pour ce travail et ma persévérance durant son cheminement. Il a su me communiquer par son enthousiasme, son sens aigu de la recherche et sa disponibilité, le goût de la recherche en statistique. Ma sincère reconnaissance et ma profonde estime lui sont adressées à travers l'aboutissement de ce travail. De plus, il me fait l'honneur de présider le jury de cette thèse; qu'il trouve ici l'expression de mes remerciements.

Je souhaite adresser des sincères remerciements aux professeurs Michel Broniatowski de l'Université de Reims, Marc Hallin de l'Université Libre de Bruxelles, Fortunato Pesarin de l'Université de Padoue pour m'avoir fait l'honneur de faire un rapport sur cette thèse. Ils ont réalisé cette tâche ingrate malgré de nombreuses activités; qu'ils trouvent ici l'expression de ma reconnaissance.

Les professeurs Akry Coulibaly, Moussa Ouattara, Fortunato Pesarin ont accepté de participer à ce jury; qu'ils veuillent accepter mes sincères remerciements.

Cette recherche a été finalisée au sein du Laboratoire de Statistique Théorique et Appliquée ( LSTA) de l'Université Paris VI, dirigé par le Professeur Paul Deheuvels, qu'il me soit permis de lui adresser mes remerciements. Je saisis également l'occasion pour dire combien la collaboration avec les chercheurs, notamment les doctorants du bureau 8A27 et le personnel( Mme Louise Lamart et Mr Pascal Epron ) de ce laboratoire a été fructueuse. Mme Lydia Giroud, secrétaire de l'ISUP nous a assisté dans la saisie et la reproduction de nos manuscrits, qu'elle en soit remerciée.

Je tiens à souligner la collaboration dynamique avec Mme Rosalba Ignaccolo de l'Université de Turin avec qui, nous avons rédigé le

chapitre V lors de notre séjour de recherche au LSTA. De plus, nous lui devons la réalisation des simulations dans les chapitres III et V; qu'elle trouve ici l'expression de mes amitiés.

Je tiens à remercier le collègue et ami, le professeur Taladidia Thiombiano qui par ces encouragements et son soutien pendant les moments difficiles, a contribué à l'aboutissement de ce travail.

Je ne saurais oublier de remercier les collègues Marcel Bonkian, Aboubakary Seynou, Boubakar Yobi, Bernard Bonzi et le professeur Dorosso Abdoul Samaté pour le soutien amical que j'ai toujours trouvé auprès d'eux. Mes remerciements vont également aux collègues de l'UFR/SEG et du département de mathématiques(UFR/SEA) pour les encouragements dont j'ai bénéficié.

Les collègues Françoise Ouédraogo, Côme Béré et Stanislas Ouaro ont su rendre agréable la présentation de ce document, nous leur sommes très reconnaissant.

Mes remerciements vont également à tous ceux là (parents et amis) qui m'ont manifesté leur compréhension et encouragements en particulier Mr Amadou Gadiaga qui m'a toujours marqué son appui. Mr Abdoul Zoromé s'est intéressé à l'évolution de notre travail en nous apportant un soutien amical; qu'il trouve ici l'expression de ma reconnaissance.

Le Programme de Troisième Cycle Interuniversitaire(PTCI) de la Conférence des Institutions d'Enseignement et de Recherche Economique et de Gestion en Afrique(CIERA).

Je ne saurais terminer sans souligner l'appui financier du projet RUG/UFR-SEG qui m'a permis d'effectuer un séjour de recherche au LSTA pour finaliser cette thèse.

# Table des matières

|   |    |
|---|----|
| Introduction Générale   | 1  |
| <b>CHAPITRE I</b>   |    |
| <b>Une synthèse sur les tests hilbertiens<br/>pour des variables aléatoires<br/>indépendantes</b> |    |
| I.1 Introduction  | 5  |
| I.2 Tests d'ajustement hilbertien   | 7  |
| I.2.1 Les données du problème   | 7  |
| I.2.2 Test d'ajustement utilisant un noyau de dimension finie                                     | 7  |
| I.2.3 Cas où le noyau est de dimension finie qui varie avec la taille de l'échantillon            | 11 |
| I.2.4 Cas d'un noyau fixe de dimension infinie  | 13 |
| I.3 Tests hilbertiens sous des hypothèses adjacentes  | 16 |
| I.3.1 Les données du problème   | 16 |
| I.3.2 Cas d'un noyau fixe de dimension finie $k$  | 17 |
| I.3.3 Comparaison avec le test de Neyman-Pearson  | 19 |
| I.3.4 Cas d'un noyau variable de dimension finie  | 20 |
| I.3.5 Cas d'un noyau de dimension infinie   | 22 |
| I.3.6 vitesse de convergence et efficacité du test  | 25 |
| I.3.7 Efficacité asymptotique de BAHADUR  | 26 |
| I.4 Tests du $\chi^2$ généralisés   | 29 |
| I.4.1 Tests hilbertiens   | 29 |
| I.4.2 Choix des cellules dans les tests du $\chi^2$   | 31 |
| I.4.3 Choix du nombre de polynômes dans le test de Legendre                                       | 31 |

|  |    |
|--|----|
| 1.4.1 Comparaison empirique des tests de Legendre et du $\chi^2$ . . . . .                         | 32 |
| 1.5 Appendices . . . . .   | 33 |
| 1.5.1 Appendice 1 : le théorème de Sazonov . . . . .   | 33 |
| 1.5.2 Appendice 2 : Courbes de puissances pour différentes hypothèses alternatives $H_a$ . . . . . | 33 |

## Chapitre II

### Tests hilbertiens pour un processus stationnaire et mélangant

|   |    |
|---|----|
| II.1. Introduction . . . . .  | 39 |
| II.2. Notations et hypothèses générales. . . . .  | 40 |
| II.3. Loi limite de $\ S_n(\cdot)\ ^2$ dans le cas d'un noyau de dimension finie $k$ . . . . .      | 41 |
| II.3.1. Loi limite de $S_n(\cdot)$ . . . . .  | 41 |
| II.3.2. Comportement asymptotique de $\ S_n(\cdot)\ ^2$ . . . . .                                   | 47 |
| II.4. <b>Vitesse de convergence</b> . . . . .   | 48 |
| II.4.1. Cas du noyau de dimension $k$ . . . . .   | 48 |
| II.4.2. Cas d'un noyau de dimension finie $k_n$ qui varie avec la taille de l'échantillon . . . . . | 50 |
| II.4.3. Cas où le noyau est de dimension infinie . . . . .  | 51 |
| II.5. Construction et propriétés asymptotiques du test . . . . .                                    | 56 |
| II.5.1. Construction du test. . . . .   | 56 |
| II.5.2. Propriétés asymptotiques du test : convergence du test . . . . .                            | 57 |

## Chapitre III

### Tests du $\chi^2$ Généralisé

#### Application au processus autoregressif

|   |    |
|---|----|
| III.1. Introduction . . . . .   | 65 |
| III.2. Estimation de $\Sigma = (\sigma_{jj'})_{2 \leq j, j' \leq k}$ . . . . .                  | 66 |
| III.2.1 Hypothèses et notations . . . . .   | 66 |
| III.2.2 Estimation . . . . .  | 67 |
| III.3. Approximation de la loi limite de $\ S(\cdot)\ ^2$ . . . . .                             | 69 |
| III.3.1. Loi limite de $\ S_n(\cdot)\ ^2$ en fonction des valeurs propres de $\Sigma$ . . . . . | 69 |
| III.3.2. Estimation des valeurs propres $\lambda_j$ ; $j=2, \dots, k$ . . . . .                 | 70 |
| III.3.3. Approximation de la loi limite de $\ S_n(\cdot)\ ^2$ . . . . .                         | 71 |

|   |    |
|---|----|
| III.4. Simulation . . . . .                                       | 72 |
| III.4.1. Introduction . . . . .                                   | 72 |
| III.4.2. Test d'hypothèse . . . . .                               | 72 |
| III.4.3. Appendice : Tableaux des puissances empiriques . . . . . | 73 |

## CHAPITRE IV

### Tests Banachiques : le cas de $C[0,1]$

|   |    |
|---|----|
| IV.1. Introduction . . . . .  | 77 |
| IV.2. Notations et définitions . . . . .                                    | 77 |
| IV.3. Loi limite de $S_n(\cdot)$ . . . . .                                  | 78 |
| IV.4. Étude du niveau du test . . . . .                                     | 81 |
| IV.4.1. Introduction . . . . .  | 81 |
| IV.4.3. Cas où le système orthonormal est la base trigonométrique . . . . . | 82 |
| IV.4.4. Cas où les $e_j$ sont lipschitziennes . . . . .                     | 86 |
| <u>Conclusion</u> . . . . .   | 87 |

## CHAPITRE V Tests de non-influence

|   |     |
|---|-----|
| V.1 Introduction et hypothèses générales . . . . .                                    | 89  |
| V.1.1 Introduction . . . . .  | 89  |
| V.1.2 Estimation de $r(x)$ . . . . .  | 91  |
| V.1.3 Définition de la statistique du test . . . . .                                  | 91  |
| V.2 Cas des observations indépendantes . . . . .                                      | 92  |
| V.2.1 Comportement asymptotique de $\ R_n\ ^2$ . . . . .                              | 92  |
| V.2.2 Construction et consistance du test . . . . .                                   | 96  |
| V.3 Cas des observations corrélées . . . . .  | 99  |
| V.3.1 Comportement asymptotique de $\ R_n\ ^2$ . . . . .                              | 99  |
| V.3.2 Construction et consistance du test . . . . .                                   | 112 |
| V.4 SIMULATIONS . . . . .   | 115 |
| V.4.1 INTRODUCTION . . . . .  | 115 |
| V.4.2 NIVEAU EMPIRIQUE . . . . .  | 115 |
| V.4.3 PUISSANCE EMPIRIQUE . . . . .   | 116 |
| V.5 APPENDICE . . . . .   | 118 |
| V.5.1 APPENDICE1 : Théorèmes limites pour un processus $\alpha$ -mélangeant . . . . . | 118 |
| V.5.2 APPENDICE2 : TABLEAUX : Niveau empirique - Puissance empirique . . . . .        | 119 |

# Introduction Générale

Une manière naturelle de construire un test d'adéquation consiste à utiliser un estimateur de la loi dont on teste l'ajustement. Ainsi, le célèbre test du  $\chi^2$  de Karl Pearson est fondé sur un estimateur élémentaire de la densité, l'histogramme des fréquences. Or, il est bien connu que les estimateurs lisses de la densité approchent la loi des observations beaucoup mieux que l'histogramme. C'est le cas notamment des estimateurs à noyau et des estimateurs par projection (Bosq - Lecoutre 1987). Aussi, peut-on espérer que les tests associés à ces estimateurs soient plus performants que le test du  $\chi^2$  usuel. Bosq (1989) a comparé par simulation le test de Legendre (c'est-à-dire le test hilbertien basé sur les polynômes de Legendre en tant que système orthonormal) et le test du  $\chi^2$  : il constate la supériorité du test de Legendre lorsque la variance est plus grande sous l'alternative que sous l'hypothèse nulle. Dans le cas contraire on peut considérer que les deux tests sont au pire équivalents. Du point de vue théorique, Gregory (1977-1980) a étudié l'efficacité et l'optimalité des tests d'adéquation. On trouve également des résultats sur les tests basés sur les estimateurs par projection (tests hilbertiens) dans Bosq (1978, 1980, 1981, 1983, 1989, 2000) quand les variables aléatoires utilisées dans la statistique hilbertienne sont indépendantes. Cette classe de tests contient le fameux test du  $\chi^2$ . Il a obtenu des résultats sur la convergence et l'optimalité asymptotique de ces tests sous diverses alternatives. Gadiaga (1982, 1983, 2002), Ignaccolo (2002) et Gadiaga et Ignaccolo (2002) ont généralisé certains de ces résultats au cas des variables aléatoires dépendantes. Pour effectuer un test d'ajustement, de nombreuses méthodes ont été introduites et étudiées (fonction de répartition empirique, estimation de la densité,...). De façon générale, on ne peut pas dire qu'il existe une méthode meilleure que les autres, ainsi de nouveaux tests sont proposés, soit plus généraux soit pour des familles précises de lois. Actuellement de nombreuses méthodes d'estimation fonctionnelle de la densité sont disponibles (voir par exemple D. Bosq et



J.P. Lecoutre [24]), comme la méthode du noyau, la méthode des "splines" ou des séries orthogonales. Dans cette variété de méthodes, nous avons choisi d'utiliser la méthode d'estimation par projection ou des séries orthogonales, tout d'abord par ce que l'histogramme des fréquences en est un cas particulier et donc, le test construit à partir d'un estimateur par projection généralise le test classique du  $\chi^2$  de Pearson. L'idée de construire un test d'ajustement associé à un estimateur par projection d'un paramètre fonctionnel peut être aussi étendue à la fonction de régression.

Le travail que nous présentons rentre dans le cadre général de ces tests hilbertiens. Ainsi, le premier chapitre fait une synthèse de l'étude de ces tests au cas des observations indépendantes. L'efficacité asymptotique, ainsi que l'optimalité des tests construits sont étudiés en introduisant les notions d'adjacence des hypothèses et la pente de Bahadur. On compare notamment les tests hilbertiens obtenus à partir du système orthonormal constitué par les polynômes de Legendre et le test du  $\chi^2$ .

Les deuxième et troisième chapitres sont une synthèse remaniée des résultats obtenus dans notre thèse 3ème cycle. Ces résultats ont fait l'objet de la soumission de deux articles dont un a été accepté et le deuxième, suite aux remarques des referees a été révisé et il reste à le compléter par des simulations. On examine dans ces chapitres les tests sous diverses alternatives pour un processus stationnaire et mélangeant à temps discret. Des lois limites ainsi que des vitesses de convergence qui permettent d'étudier la convergence du test construit sont obtenues. Comme application, nous faisons des simulations sur un processus autoregressif d'ordre un.

Au quatrième chapitre nous obtenons des théorèmes limites dans le cas où les observations sont indépendantes et dans un espace de Banach et plus particulièrement l'espace  $C[0, 1]$ . Les lois limites des statistiques considérées sont étudiées dans  $C[0, 1]$  en utilisant les travaux de Jain et Marcus (1975) ainsi que le comportement du niveau du test construit.

Le chapitre cinq est une prépublication (preprint) que nous avons rédigée avec R. Ignaccolo et porte sur tests de "non-influence". Après avoir considéré l'estimateur de la fonction de régression  $r(x), r(x) = E\{Y/X = x\}$  nous étudions le test d'hypothèse nulle " $r(x) = cste$ ". C'est-à-dire que la variable aléatoire  $X$  n'a pas d'effet en moyenne sur la variable aléatoire  $Y$ . Des lois limites sous diverses alternatives sont obtenues ainsi que des conditions nécessaires et suffisantes de convergence du test. Des simulations sont indiquées pour illustrer certains résultats.

Enfin, nous étudions au chapitre six les tests non paramétriques de la den-

sité marginale d'un processus stationnaire à temps continu. Nous rappelons les résultats obtenus par A.Frenay (2001) sur l'estimation à temps continu par projection orthogonale et nous construisons des tests sous diverses alternatives.

# CHAPITRE I

## Une synthèse sur les tests hilbertiens pour des variables aléatoires indépendantes

### I.1 Introduction

Le test du  $\chi^2$  de K.PEARSON [67] est très employé par les praticiens qui veulent ajuster des observations à une loi donnée. L'alternative d'un tel test reste vague dans l'esprit de l'utilisateur. Pourtant la statistique utilisée correspond à un choix bien précis des hypothèses  $H_0$  et  $H_1$  ; aussi, en faisant varier ces hypothèses, on peut construire toute une classe de tests du même type que le test du  $\chi^2$ . Les tests de cette classe seront appelés "tests hilbertiens".

Pour honorer le centenaire du fameux article de Karl Pearson, un colloque international ("International workshop on Goodness of fit Tests and validity of models : GOF 2000"), auquel nous avons assisté, a été organisé à Paris les 29, 30 et 31 mai 2000. Ce colloque a permis de célébrer la publication de l'article de K.Pearson (1857-1936) et faire le point sur les recherches portant sur les méthodes statistiques liées à la modélisation et à la validation des modèles (plus de 80 communications ont été faites).

L'article initial constitue une méthode proposée par Pearson dans les années 1894-1900 pour rendre compte des données de toute nature. Après avoir choisi la forme fonctionnelle générale de la "courbe" de fréquence des données dont il s'agit, et déterminé les "constantes" dont cette courbe dépend, il convient de juger la qualité de son ajustement à la nature des

# CHAPITRE I

## Une synthèse sur les tests hilbertiens pour des variables aléatoires indépendantes

### I.1 Introduction

Le test du  $\chi^2$  de K.PEARSON [67] est très employé par les praticiens qui veulent ajuster des observations à une loi donnée. L'alternative d'un tel test reste vague dans l'esprit de l'utilisateur. Pourtant la statistique utilisée correspond à un choix bien précis des hypothèses  $H_0$  et  $H_1$ ; aussi, en faisant varier ces hypothèses, on peut construire toute une classe de tests du même type que le test du  $\chi^2$ . Les tests de cette classe seront appelés "tests hilbertiens".

Pour honorer le centenaire du fameux article de Karl Pearson, un colloque international ("International workshop on Goodness of fit 'Tests and validity of models : GOF 2000'"), auquel nous avons assisté, a été organisé à Paris les 29, 30 et 31 mai 2000. Ce colloque a permis de célébrer la publication de l'article de K.Pearson (1857-1936) et faire le point sur les recherches portant sur les méthodes statistiques liées à la modélisation et à la validation des modèles (plus de 80 communications ont été faites).

L'article initial constitue une méthode proposée par Pearson dans les années 1894-1900 pour rendre compte des données de toute nature. Après avoir choisi la forme fonctionnelle générale de la "courbe" de fréquence des données dont il s'agit, et déterminé les "constantes" dont cette courbe dépend, il convient de juger la qualité de son ajustement à la nature des

observations considérées. Certains des tests de la classe mentionnée plus haut ont été étudiés par J.Neyman [65] et E.Nadaraja [62]. Des tests analogues ont été également étudiés par Bickel et Rosenblatt [9].

Plus récemment, G.Gregory [46,47] a étudié la théorie des grands échantillons pour les U-statistiques et tests d'ajustement, ainsi que sur l'efficacité et l'optimalité des tests quadratiques. Les résultats obtenus sont utilisés pour déterminer l'efficacité asymptotique relative pour les tests statistiques portant sur deux mesures. De plus, les résultats de Gregory[46] permettent d'obtenir des tests asymptotiquement optimaux. La plupart des résultats établis pour les tests hilbertiens sont dus à D. Bosq[14,15,16,17,18,19,23].

Dans [14,15,16] on étudie une classe de tests non paramétriques basés sur des statistiques qui sont des normes hilbertiennes. Le fameux test du  $\chi^2$  fait partie de cette classe. Des lois limites sont fournies ainsi que des vitesses de convergence qui précisent le comportement asymptotique des statistiques utilisées. Ces vitesses de convergence permettent d'étudier le comportement asymptotique des tests construits.

Dans [17,18], les résultats obtenus dans [14] sont utilisés pour étudier la loi limite de la statistique de test pour les hypothèses adjacentes à l'hypothèse nulle.

Dans [19], le cas des tests du  $\chi^2$  généralisés est envisagé et on les compare avec le test du  $\chi^2$  par des simulations.

Dans [23] est abordé le problème de l'efficacité locale des tests d'ajustement par l'utilisation de la pente de Bahadur(cf. Nikitin 1995).

Ainsi, dans ce chapitre, nous présentons une synthèse sur les tests hilbertiens en nous appuyant sur les études ci-dessus mentionnées.

Le plan de la suite de ce chapitre est le suivant :

L'étude générale des tests hilbertiens est effectuée dans la section 2.

Dans la 3<sup>e</sup> section nous étudions le cas où les hypothèses sont adjacentes.

La section 4 est consacrée à une comparaison par simulation avec le test du  $\chi^2$  classique. Enfin dans l'appendice nous rappelons un résultat fondamental pour l'étude asymptotique des tests hilbertiens : le théorème de Sazonov.

## 1.2 Tests d'ajustement hilbertien

### 1.2.1 Les données du problème

Soit  $(X_n)_{n \geq 1}$  une suite de v.a indépendantes et de même loi, définies sur le même espace probabilisé  $(\Omega, \mathcal{A}, P)$  et à valeurs dans l'espace mesurable  $(E, \mathcal{B})$ ; on cherche à tester l'hypothèse  $H_0$  : "la loi commune des  $X_j$  est  $\mu$ ". Ce problème est étudié sous diverses alternatives.

### 1.2.2 Test d'ajustement utilisant un noyau de dimension finie

#### 1.2.2.1 Loi limite et vitesse de convergence d'une statistique hilbertienne.

L'espace  $L^2(\mu)$  étant supposé de dimension au moins égale à  $k$  ( $k \geq 2$ ), considérons un système orthonormal  $(e_1, \dots, e_k)$  dans  $L^2(\mu)$ , avec  $e_1 = 1$ , et posons :

$$K(x, t) = \sum_{j=1}^k e_j(x)e_j(t), \quad (x, t) \in E^2$$

$K(\dots)$  étant le noyau de l'espace vectoriel  $E_k$  engendré par  $(e_1, \dots, e_k)$ .  
Considérons maintenant la statistique

$$T_n(\cdot) = \sum_{i=1}^n K(X_i, \cdot)$$

Sous  $H_0$ , on a

$$T_n(\cdot) = n(1 + \sum_{j=2}^k \hat{a}_{jn} e_j(\cdot))$$

où

$$\hat{a}_{jn} = \frac{1}{n} \sum_{i=1}^n e_j(X_i), \quad j = 1, \dots, k$$

Nous avons d'abord les résultats suivants :

**Proposition 1.** Si la loi de  $X_i$  est  $\mu$ , alors

1. La v.a

$$S_n(\cdot) = n^{-1/2}(T_n(\cdot) - n)$$

tend en loi dans  $L^2(\mu)$  vers une v.a  $Z$  gaussienne, centrée et de même opérateur de covariance que :

$$K(X_1, \cdot) - E\{K(X_1, \cdot)\}$$

2. La v.a  $Z$  peut s'écrire sous la forme

$$Z = \sum_{j=2}^k Z_j e_j(\cdot)$$

Où les  $Z_j$  suivent des lois normales centrées réduites indépendantes.

3.  $\|S_n(\cdot)\|^2$  converge en loi vers  $Q$ , où  $Q$  suit une loi du  $\chi^2$  à  $k-1$  d.l

4. Si de plus

$$\int \left[ \sum_{j=2}^n |e_j|^3 \right] d\mu < \infty$$

alors, pour tout  $n \geq 1$  :

$$\sup_{a \in \mathbb{R}^+} |P(\|S_n(\cdot)\|^2 \leq a) - P(Q \leq a)|$$

$$\leq c_0(k-1)^3 \times \int \left[ \sum_{j=2}^n |e_j|^3 \right] d\mu^{1/2}$$

où  $c_0$  est une constante universelle.

**Démonstration.** La loi limite est une conséquence directe du théorème central limite dans un espace de Hilbert, et du fait que la transformation par une fonction continue préserve la convergence en loi. ■

Pour la vitesse de convergence, on a appliqué le théorème de Sazonov (voir appendice).

**Remarque :**

Si la loi commune des  $X_i$  est une probabilité  $\nu$  telle qu'il existe  $j \in \{2, \dots, k\}$  pour lequel  $e_j$  soit  $\nu$ -intégrable et d'intégrale non nulle, alors

$$\|S_n(\cdot)\|^2 \xrightarrow{p.s} +\infty$$

en effet la loi forte des grands nombres permet de dire que

$$\hat{a}_{j,n} \xrightarrow{p.s} \int e_j d\nu \neq 0, p.s$$

Donc

$$\|S_n(\cdot)\|^2 = n \sum_{j=2}^k \hat{a}_{j,n}^2 \xrightarrow{p.s} +\infty$$

Nous pouvons maintenant construire un test basé sur  $S_n(\cdot)$ .

### 1.2.2.2 Construction du test

On veut tester :

$H_0$  : "la loi des  $X_i$  est  $\mu$ "

contre  $H_1(k)$  : "la loi des  $X_i$  est une probabilité  $\nu$  telle que  $\int e_j d\nu$  existe et est non nulle pour au moins un  $j \in \{2, \dots, k\}$ ".

Nous avons vu que

Sous  $H_0$  :

$$\|S_n(\cdot)\|^2 = n \sum_{j=2}^k \hat{a}_{j,n}^2 \xrightarrow{p.s} \chi_{k-1}^2$$

Sous  $H_1$  :

$$\|S_n(\cdot)\|^2 = n \sum_{j=2}^k \hat{a}_{j,n}^2 \xrightarrow{p.s} +\infty$$

D'où un test de la forme

$$n \sum_{j=2}^k \hat{a}_{j,n}^2 > w$$



Soit alors  $\alpha \in ]0, 1[$  et  $w$  tel que  $P(Q > w) = \alpha$ . Alors le test de la forme ci dessus est asymptotiquement de niveau  $\alpha$ .  
De plus, sous l'hypothèse nulle on a :

$$|P(n \sum_{j=2}^k \hat{a}_{j,k}^2 > w) - \alpha| = O(1/\sqrt{n}) \quad (\text{Proposition 1})$$

**Remarques**

1. Considérons une partition  $\{A_1, \dots, A_k\}$  de  $E$  telle que  $p_j = P(A_j) > 0$ , et le système orthonormal dans  $L^2(\mu)$  défini par :

$$\{f_j = p_j^{-1/2} 1_{A_j}, j = 1, 2, \dots, k\}$$

qui engendre un espace vectoriel de dimension  $k$  et contenant les constantes. Alors

$$\begin{aligned} \|S_n(\cdot)\|^2 &= \frac{1}{n} \left\| \sum_{i=1}^n K(X_i, \cdot) - n \right\|^2 \\ &= \sum_{j=1}^k 1/n p_j \left[ \sum_{i=1}^n 1_{A_j}(X_i) - n p_j \right]^2 \end{aligned}$$

C'est la statistique utilisée dans le test du  $\chi^2$ .

2. On établit que le test  $n \sum_{j=2}^k \hat{a}_{j,n}^2 > w_n$  converge si et seulement si  $w_n \rightarrow +\infty$  et  $w_n/n \rightarrow 0$ .

Rappelons qu'un test est dit convergent si son niveau  $\alpha_n$  tend vers 0, et sa puissance  $\beta_n(\nu)$  tend vers 1 pour toute loi  $\nu$  de l'alternative.

$$\alpha_n = P_\mu[\|S_n(\cdot)\|^2 > w_n]$$

$$\beta_n(\nu) = P_\nu[\|S_n(\cdot)\|^2 > w_n]$$

## 1.2.3 Cas où le noyau est de dimension finie qui varie avec la taille de l'échantillon

### 1.2.3.1 Vitesse de convergence

On considère une suite de systèmes orthonormaux  $(e_{jn}, j \geq 1)$  dans  $L^2(\mu)$  telle que  $e_{1n} = 1, n \geq 1$ .

Le noyau  $K_n$  sera défini par :

$$K_n(x, t) = \sum_{j=1}^{k_n} e_{jn}(x)e_{jn}(t); (x, t) \in E \times E, n \geq 1$$

où

$$\lim_{n \rightarrow \infty} k_n = +\infty$$

et on pose

$$S_n(\cdot) = \sqrt{n} \left| \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right|$$

et

$$\hat{a}_{jn} = \frac{1}{n} \sum_{i=1}^n e_{jn}(X_i), j \geq 1, n \geq 1$$

On a le résultat :

**Proposition 2.** Soit  $\mu$  la loi des  $X_i$  et soit  $Q_n$  une suite de v.a.r qui suivent des lois du  $\chi^2$  dont les degrés de liberté respectifs valent  $(k_n - 1)$ . Alors

$$\sup_{a \in \mathbb{R}^+} | [P(|S_n|^2 \leq a) - P(Q_n \leq a)] | \leq c_0(k_n - 1)^3 / \sqrt{n} \int \sum_{j=1}^{k_n} |e_{jn}|^3 d\mu$$

*Démonstration.* La preuve utilise le théorème de Zazonov. ■

**Remarques**

1. Pour tester l'hypothèse  $H_0$  "La loi des  $X_j$  est  $\mu$ , on peut utiliser encore un test de la forme

$$n \sum \hat{a}_{jn}^2 > w_n$$

L'alternative  $H_1(k_n)$  sera, par exemple, l'ensemble des lois  $\nu$  telles qu'il existe une suite d'entiers  $j(n)$  pour laquelle :

$$0 < \inf_n \left| \int c_{j(n),n} d\nu \right| < +\infty$$

2. Si  $c_{jn} = c_j$ ,  $n \geq 1$ ,  $j \geq 1$ , l'alternative sera l'ensemble des lois  $\nu$  telles que, pour au moins un  $j \geq 2$ ,  $\int c_j d\nu$  existe et soit non nulle.  $H_1(k_n)$  contient alors toutes les lois (différente de  $\mu$ ) qui admettent une densité appartenant à  $L^2(\mu)$ .

### 1.2.3.2 Convergence du test

On suppose dans cette partie que :

$c_{jn} = c_j$ ,  $j \geq 1$ ,  $n \geq 1$ , et que les  $c_j$  sont uniformément bornés.

$k_n \rightarrow +\infty$  et

$$k_n^4/\sqrt{n} \rightarrow 0$$

Alors on a la proposition suivante

**Proposition 3.** 1. Pour que le test converge, il est nécessaire et suffisant que l'on ait :

$$\frac{w_n - k_n}{\sqrt{k_n}} \rightarrow +\infty$$

et

$$w_n/n \rightarrow 0$$

2. Si  $H \subset H_1(k_n)$  est telle que, pour  $n$  assez grand, on ait :

$$\inf_{\nu \in H} \max_{1 \leq j \leq k_n} \left| \int e_j d\nu \right| \geq m \sqrt{\frac{w_n}{n}}$$

où  $m$  est un entier donné supérieur à 1, alors la convergence de la puissance est uniforme sur  $H$  dès que  $w_n \rightarrow +\infty$ . De plus :

$$\sup_{\nu \in H} |1 - p_n(\nu)| = O(1/w_n)$$

Démonstration. Voir [14] p 29-32. ■

## 1.2.4 Cas d'un noyau fixe de dimension infinie

### 1.2.4.1 Les données du problème

On pose

$$K(x, t) = \sum_{j=1}^{+\infty} \lambda_j e_j(x) e_j(t)$$

avec

$$\lambda_1 = 1, e_1 = 1, \sum \lambda_j^2 < +\infty$$

et

$$\forall j \geq 1, 1 \geq |\lambda_j| \geq |\lambda_{j+1}| > 0$$

Les  $(e_j)$  forment un système orthonormal dans  $L^2(\mu)$ , et la convergence de  $\sum_{j>1} \lambda_j e_j \otimes e_j$  a lieu au sens de  $L^2(\mu \otimes \mu)$

Enfin, on suppose que

$$\sup_{x \in E, j \geq 1} |e_j(x)| = M < \infty$$

### 1.2.4.2 Vitesse de convergence

**Proposition 4.** a) Si la loi des  $X_i$  est  $\mu$ , alors :

1. La suite de v.a. hilbertiennes :

$$S_n(\cdot) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right], \quad n \geq 1$$

converge en loi vers une v.a.  $Z$  à valeurs dans  $L^2(\mu)$ , gaussienne, dont l'opérateur de covariance est celui de  $K(X_1, \cdot) - 1$ .  $Z$  admet une représentation de la forme :

$$Z = \sum_{j=2}^{+\infty} \lambda_j Z_j e_j(\cdot)$$

où les  $Z_j$  sont des variables gaussiennes centrées réduites indépendantes.

2.  $\|S_n(\cdot)\|^2$  admet une loi limite dont la fonction caractéristique est :

$$\phi(t) = \prod_{j=2}^{+\infty} (1 - 2i\lambda_j^2 t)^{-1/2}$$

$$\|S_n(\cdot)\|^2 \longrightarrow \|Z\|^2 = \sum_{j=2}^{+\infty} \lambda_j^2 Z_j^2$$

3.

$$\begin{aligned} & \sup_{n \in \mathbb{N}^+} | [P(\|S_n\|^2 \leq a) - P(\|Z\|^2 \leq a)] | \\ & \leq K_1 \lambda_{r(n)}^{2/3} + K_2 \frac{r(n)^4}{\sqrt{n}} \end{aligned}$$

où

$$K_1 = M^2 + 6 + 2\sqrt{\frac{2}{\pi}} + \frac{\sqrt{2}}{\pi |\lambda_4|^3}$$

$$K_2 = 3C_0 M$$

$$\lambda_{r(n)} = \sum_{j=r(n)+1}^{+\infty} \lambda_j^2$$

et

$$\lim_{n \rightarrow +\infty} r(n) = +\infty$$

b) Si les  $X_i$  suivent une loi  $\nu$  telle que  $\int e_j d\nu$  existe et est non nulle pour au moins un entier  $j \geq 2$ , alors

$$\|S_n(\cdot)\|^2 \rightarrow +\infty, \text{ p.s.}$$

*Démonstration.* Elle est basée sur l'utilisation du théorème de Sazonov. Voir [14] p 13-21. ■

**Remarque :**

La construction du test est analogue à celle du cas précédent.

### 1.2.4.3 Convergence du test

**Proposition 5.** Le test défini par

$$\|S_n(\cdot)\|^2 > w_n$$

est convergent si et seulement si

$$w_n \rightarrow +\infty \text{ et } \frac{w_n}{n} \rightarrow 0$$

*Démonstration.* On utilise la proposition 4. On choisit  $r(n)$  de telle sorte que les deux membres de l'inégalité tendent vers 0. Dans ces conditions :

$$\alpha_n \rightarrow 0 \Leftrightarrow P(\|Z\|^2 > w_n) \rightarrow 0 \Leftrightarrow w_n \rightarrow +\infty$$

Pour l'étude de la puissance, soit  $\nu \in H_1(\lambda_n)$ , et soit  $j_0 \geq 2$  tel que

$$\int e_{j_0} d\nu \neq 0.$$

Comme

$$\|S_n(\cdot)\|^2 \leq w_n \Rightarrow n\lambda_{j_0}^2 \hat{a}_{j_0 n} \leq w_n$$

on a aussi

$$P(\|S_n(\cdot)\|^2 \leq w_n) \leq P(|\hat{a}_{j_0 n}| \leq \sqrt{\frac{w_n}{n}} \lambda_{j_0}^{-1})$$

Si  $\frac{w_n}{n} \rightarrow 0$  on aura, pour  $n$  assez grand,

$$\frac{w_n}{n} \leq \frac{a_{j_0}^2 \lambda_{j_0}^2}{m^2}$$

( $m \rightarrow 1$  donné). D'où

$$\begin{aligned} P(\|S_n(\cdot)\|^2 \leq w_n) &\leq P(|\hat{a}_{j_0 n}| < \frac{|a_{j_0}|}{m}) \\ &\leq P(|\hat{a}_{j_0 n} - a_{j_0}| > |a_{j_0}| \frac{m-1}{m}) \\ &< \frac{\lambda^2}{a_{j_0}^2} \left(\frac{m}{m-1}\right)^2 \frac{1}{n} \end{aligned}$$

par conséquent  $p_n(\nu) \rightarrow 1$ . ■

**Remarques :**

En utilisant la proposition 4, on voit que le niveau du test tend vers 0, et plus précisément que

$$\alpha_n = O\left(\frac{1}{w_n} + \lambda_{r(n)}^{2/3} + \frac{r(n)^4}{\sqrt{n}}\right)$$

Par exemple, pour

$$\lambda_j = O(j^{-\gamma}), \quad \gamma > 1/2,$$

$$r(n) = O(n^{3/8\gamma+2}), \text{ et } w_n = O(n^{\frac{2\gamma-1}{4\gamma+1}})$$

on obtient

$$\alpha_n = O\left(\frac{1}{n^{\frac{2\gamma-1}{4\gamma+1}}}\right)$$

Concernant la puissance, on a  $1 - \beta_n = O(1/n)$ .

## I.3 Tests hilbertiens sous des hypothèses adjacentes

### I.3.1 Les données du problème

Soit  $\mathcal{M}$  une famille de probabilités sur  $(E, \mathcal{B})$ , et  $\mu \in \mathcal{M}$ , on s'intéresse au test

$$"H_0 = \mu" \text{ contre } "H_1 = \mathcal{M} - \mu"$$

Pour cela, on se donne une suite  $\nu_n \in \mathcal{M}, n \geq 1$  "adjacente" à  $\mu$ , en un sens qui sera précisé ultérieurement.

Soit également  $(\{X_{1n}, \dots, X_{kn}, n \geq 1\})$  une suite de séquences de variables aléatoires indépendantes et de loi  $\nu_n$ , définies sur l'espace probabilisé  $(\Omega, \mathcal{A}, P)$ .

Le comportement asymptotique de  $S_n(\cdot)$  a été étudié précédemment lorsque la loi des  $X_i$  est  $\mu$  ou une loi fixe de la contre hypothèse.

Ici, on s'intéressera à l'étude des propriétés asymptotiques locales au voisinage de l'hypothèse nulle.

### 1.3.2 Cas d'un noyau fixe de dimension finie $k$

#### 1.3.2.1 Adjacence de $\nu_n$

$\nu_n$  est dite adjacente à  $\mu$  si les deux conditions suivantes sont vérifiées :

**A1 :**

$$\sqrt{n} \left| \int K(x, \cdot) d\nu_n - 1 \right|$$

converge dans  $E'$ , le sous espace de  $E$  orthogonal à 1, avec

$$\sqrt{n} \left| \int K(x, \cdot) d\nu_n - 1 \right| \rightsquigarrow g, \quad \|g\|^2 = \lambda^2 > 0$$

**A2 :** L'opérateur de covariance de  $\nu_n$ , noté  $\Gamma_{\nu_n}$ , converge simplement vers l'identité de  $E'$ , c'est à dire :

$$\forall 1 \leq j, l \leq k, \Gamma_{\nu_n}(j, l) = \int c_j c_l d\nu_n - \int c_l d\nu_n \int c_j d\nu_n \rightsquigarrow \delta_{j,l}$$

Où  $\delta_{j,l}$  désigne le symbole de Kronecker.

#### 1.3.2.2 Loi limite de $\|S_n(\cdot)\|^2$

**Proposition 6.** Si A1 et A2 sont vérifiées, alors on a :

$$n^{-1/2} \left( \sum_{i=1}^n c_1(X_{in}), \dots, \sum_{i=1}^n c_k(X_{in}) \right) \xrightarrow{\mathcal{L}} (U_1 + v_1, \dots, U_k + v_k)$$



Où  $(U_1, \dots, U_k)$  est un échantillon de  $N(0, 1)$ , et  $(v_1, \dots, v_k)$  est la limite de  $V_n$ , avec

$$V_n = \sqrt{n} \left( \int c_1 d\nu_n, \dots, \int c_k d\nu_n \right)$$

et par la suite :

$$\|S_n(\cdot)\|^2 \xrightarrow{\mathcal{L}} \sum_{j=1}^k (U_j + v_j)^2$$

**Démonstration.** Nous donnons les grandes lignes de cette preuve. On utilise successivement les étapes suivantes :

- 1) A1 et A2 impliquent la convergence de  $C_{\nu_n}$  (matrice de variances-covariances de  $U_n = (U_{1n}, \dots, U_{kn})$ ) vers  $I_k$ .
- 2)  $U_n \rightarrow U = (U_1, \dots, U_k)$  qui est un vecteur de matrice de covariance  $I_k$ .
- 3) La convergence en loi se conserve par continuité.  
Pour les détails, voir [18] p.7-8

■

**Remarque :**

$\|S_n(\cdot)\|^2$  converge vers une loi du  $\chi^2$  décentrée à  $k$  degrés de liberté, dont le paramètre de décentrage vaut

$$\lambda^2 = \|v\|^2 = \lim_{n \rightarrow \infty} n \left\| \int K(x, \cdot) d\nu_n - 1 \right\|^2$$

### 1.3.2.3 Efficacité asymptotique du test

Soit  $\alpha \in ]0, 1[$ . Sous A1 et A2, la puissance du test de niveau asymptotique  $\alpha$  s'écrit :

$$P_n(k) = P(\chi^2(k, \lambda^2) > \chi_\alpha^2)$$

où  $\chi_\alpha^2$  est défini par la relation

$$\alpha = P(\chi^2(k, 0) > \chi_\alpha^2)$$

### I.3.3 Comparaison avec le test de Neyman-Pearson

On se donne une suite de probabilités  $\nu_n$  de la forme :

$$\nu_n = (1 + g_n \frac{c_n}{\sqrt{n}}) \cdot \mu, \quad n \geq 1$$

où  $g_n \in L^2(\mu)$  et  $c_n \in \mathbb{R}^+$ .

Prenons

$$T_n = \sum_{i=1}^n \log(1 + g_n(X_{in}) \frac{c_n}{\sqrt{n}})$$

alors le test de Neyman-Pearson est de la forme

$$T_n > \log(w_n)$$

ou encore

$$\frac{1}{\gamma} [c_n^{-1} T_n + \frac{c_n}{2} \|g_n\|^2] > \frac{1}{\gamma} [c_n^{-1} \log(w_n) + \|g_n\|^2]$$

où

$$\|g_n\|^2 \rightarrow \gamma^2$$

Le test est de niveau asymptotique  $\alpha \in ]0, 1[$  si

$$Z_n = \frac{1}{\gamma} [c_n^{-1} \log(w_n) + \frac{c_n}{2} \|g_n\|^2] \rightarrow N_\alpha \quad (1)$$

avec

$$P(N(0, 1) > N_\alpha) = \alpha$$

Si  $c_n \rightarrow c$ , alors

$$w_n \rightarrow \exp[N_\alpha \gamma - \frac{1}{2} \gamma^2 c^2]$$

Si la relation (1) est vérifiée, on peut chercher la puissance asymptotique du test.

Sous la suite d'alternatives, on a

$$\frac{1}{\gamma} [c_n^{-1} \log(w_n) - \frac{c_n}{2} \|g_n\|^2] \rightarrow N(0, 1)$$

alors la puissance asymptotique est la limite de

$$P\{N(0, 1) > \frac{1}{\gamma} [c_n^{-1} \log(w_n) - \frac{c_n}{\gamma} \|g_n\|^2]\}$$

ou encore la limite de

$$P\{N(0, 1) > g_n - \frac{c_n}{\gamma} \|g_n\|^2\}$$

La limite appartient à  $]0, 1[$ .

Si  $c_n \rightarrow c$ , elle vaut alors  $P(N(0, 1) > N_\alpha - c\gamma)$ .

Si  $c_n \rightarrow +\infty$ , alors la puissance du test asymptotique est égale à 1.

**Remarque :**

Le test hilbertien de noyau fixe et de niveau asymptotique  $\alpha \in ]0, 1[$  a une puissance asymptotique qui vaut au plus  $P(\chi^2(k, \gamma^2) > \chi_\alpha^2)$  (ceci pour  $c_n = 1$ , et  $\|g_n\|^2 \rightarrow \gamma^2$ ). L'efficacité asymptotique d'un tel test vaut donc

$$e = \frac{P(\chi^2(k, \gamma^2) > \chi_\alpha^2)}{P(N(0, 1) > N_\alpha(0, 1) - \gamma)}$$

Cas particulier pour  $k = 1$ . Comme  $\chi_\alpha = N_{\alpha/2}$ , il vient facilement

$$e = \frac{P\{|N(0, 1) - \gamma| > N_{\alpha/2}(0, 1)\}}{P(N(0, 1) > N_\alpha(0, 1) - \gamma)}$$

### I.3.4 Cas d'un noyau variable de dimension finie

On se donne ici une suite  $(K_n, n \geq 1)$  de noyaux définis positifs de rangs respectifs  $k_n$ , c'est à dire :

$$K_n(x, y) = \sum_{j=1}^{k_n} e_{jn}(x) e_{jn}(y), \quad (x, y) \in E^2$$

et  $e_{0n} \equiv 1$ .

### 1.3.4.1 Loi limite de $\|S_n(\cdot)\|^2$

**Proposition 7.** *Sous les conditions suivantes*

1.

$$k_n \rightarrow +\infty, \text{ et } k_n^{11}/n \rightarrow 0$$

2.

$$n \sum_{j=1}^{k_n} \left( \int e_{jn}(x) d\nu_n(x) \right)^2 \rightarrow l$$

3.

$$\limsup_{n \rightarrow +\infty} |C_{\nu_n} - I_{k_n}|_{k_n \times k_n} < 1$$

4.

$$M = \sup_{x \in E} \sup_{1 \leq j < k_n, n > 1} |e_{jn}(x)| < +\infty$$

alors on a

$$\frac{\|S_n(\cdot)\|^2 - k_n}{\sqrt{2k_n}} \xrightarrow{\mathcal{L}} N(0, 1)$$

*Démonstration.* voir [18] p 30-32 ■

**Remarque :**

1. Si l'hypothèse 2) de la proposition ci dessus est remplacée par :

$$\frac{n}{\sqrt{k_n}} \sum_{j=1}^{k_n} \left( \int e_{jn} d\nu_n \right)^2 \rightarrow l'$$

alors on obtient

$$\frac{\|S_n(\cdot)\|^2 - k_n}{\sqrt{2k_n}} \xrightarrow{\mathcal{L}} N(l'/\sqrt{2}, 1)$$

2. Nadaraya, dans [62], a obtenu la loi limite de la proposition 7 sous des conditions plus fortes.

### I.3.4.2 Etude du test

Soit  $\alpha \in ]0, 1[$ .

Si

$$w_n = N_\alpha \sqrt{2k_n} + k_n$$

avec  $P(N > N_\alpha) = \alpha$  ( $N$  suit une  $N(0,1)$ ). Alors

$$P_\mu[\|S_n(\cdot)\|^2 > w_n] = [P_\mu[\frac{\|S_n(\cdot)\|^2 - k_n}{\sqrt{2k_n}} > N_\alpha]$$

$$\longrightarrow P_\mu[N > N_\alpha] = \alpha$$

Le test est donc de niveau asymptotique  $\alpha$ , et sa puissance asymptotique s'écrit

$$P[N + t/\sqrt{2} > N_\alpha]$$

**Remarque :**

Si

$$c_n = k_n^{1/4} \text{ et } \|g_n\|^2 \rightarrow \gamma^2$$

où l'on a posé

$$\frac{d\nu_n}{d\mu} = 1 + g_n \frac{k_n^{1/4}}{\sqrt{n}}$$

alors la puissance asymptotique du test de Neyman-Pearson de niveau asymptotique  $\alpha$  vaut 1, et donc l'efficacité asymptotique du test hilbertien optimal sera égale à sa puissance asymptotique, c'est à dire :

$$e = P[N + \frac{\|g\|}{\sqrt{2}} > N_\alpha]$$

où  $g_n$  tend fortement vers  $g$ .

### I.3.5 Cas d'un noyau de dimension infinie

Dans cette partie, nous donnons les résultats sans démonstration, elles sont très techniques. Pour les détails, voir [17], p 8-20.

### 1.3.5.1 Hypothèses et notations

Soit  $(\{X_{1n}, \dots, X_{nn}\}, n \geq 1)$  une suite de séquences de variables aléatoires indépendantes et de même loi. On s'intéresse ici au comportement asymptotique des statistiques :

$$S_n(\nu_n) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_{in}, \cdot) - \int K(t, \cdot) d\nu_n(t) \right]$$

et

$$S_n(\cdot) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_{in}, \cdot) - 1 \right]$$

$$S_n(\cdot) = S_n(\nu_n + \int K'(t, \cdot) d\nu_n(t))$$

où  $K' = K - 1$ .

Lorsque la loi de  $(X_{1n}, \dots, X_{nn})$  est  $\nu_n^{\otimes n}$ , on pose

$$\hat{c}_{jn} = \frac{1}{n} \sum_{i=1}^n c_j(X_{in})$$

et

$$c_{jn} = \int c_j d\nu_n$$

$$K(x, t) = \sum_{j=0}^{+\infty} \lambda_j c_j(x) c_j(t), \quad (x, t) \in E^2$$

où  $(c_j), j \geq 0$  est un système orthonormal de  $L^2(\mu)$  tel que  $c_0 = 1$ , et où

$$\lambda_0 = 1; \quad \forall j \geq 0, |\lambda_j| \geq |\lambda_{j+1}| > 0; \quad \sum \lambda_j^2 < +\infty$$

Posons également

$$\begin{aligned} h_n &= \int K'(t, s) h(s) d\mu(s) \nu_n(t) \\ &= \sum_{j=1}^{+\infty} \lambda_j \int c_j h d\mu \int c_j d\nu_n \end{aligned}$$

où  $h \in L^2(\mu)$  tel que  $\int c_j h d\mu \neq 0$ .

### 1.3.5.2 Proposition 8

**Proposition 8.** *Sous les hypothèses suivantes :*

1.

$$\lim_{n \rightarrow +\infty} \sum_{j=1}^{+\infty} \lambda_j^2 (\sqrt{n} \int e_j \nu_n - \alpha_j)^2 = 0$$

avec

$$\sum_{j=1}^{+\infty} \lambda_j^2 \alpha_j^2 < +\infty$$

2.

$$\lim_{n \rightarrow +\infty} \int e_i e_j d\nu_n = \delta_{ij}, \quad i, j \geq 1$$

$$\sum |\lambda_j| < \infty$$

Alors on a

a)

$$S_n \xrightarrow{\mathcal{L}} Z$$

où  $Z$  est une variable aléatoire gaussienne à valeurs dans  $L^2(\mu)$ , de moyenne

$$\sum_{j \geq 1} \lambda_j \alpha_j e_j(\cdot)$$

et d'opérateur de covariance  $K' \circ K'$

b)

$$n \sum_{j \geq 1} \lambda_j^2 \hat{c}_{jn}^2 \xrightarrow{\mathcal{L}} \sum_{j \geq 1} \lambda_j^2 (U_j + \alpha_j)^2$$

Démonstration. Voir [17] p 6-7. ■

### I.3.6 vitesse de convergence et efficacité du test

#### I.3.6.1 vitesse de convergence

**Proposition 9.** *Sous les hypothèses de la proposition 8, on a :*

$$\sup_{\alpha \in \mathbb{R}^+} |P\{\|S_n(\cdot)\|^2 \leq \alpha\} - P\{\|Z\|^2 \leq \alpha\}|$$

$$\leq \gamma_n + (d_1 + d_2) \Gamma_{r(n)}^{2/3} + d_3 \frac{r_n^4}{\sqrt{n}}$$

$$\Gamma_{r(n)} = \sum_{j=r(n)+1}^{+\infty} \lambda_j^2 (1 + \alpha_j)^2$$

$$\gamma_n = C_0 M^{3/2} \left( \sum_{j=1}^{r_n} \lambda_j \right)^3 / 2 \left( \sum_{l=1}^{r_n} \alpha_{ln}^{-3/2} \right) \frac{r_n^4}{\sqrt{n}}$$

et les  $\alpha_n$  sont les valeurs propres de la matrice de variance-covariance  $D_{v_n}$  de terme général.

$$(D_{v_n})_{i,j} = \lambda_i \lambda_j \left[ \int c_i c_j d\nu_n - \int c_i d\nu_n \int c_j d\nu_n \right]$$

Démonstration. Voir [17] p 10-18. ■

#### I.3.6.2 Efficacité et optimalité asymptotique du test

On obtient les résultats suivants pour l'étude asymptotique.

1. Si

$$\|S_n(\cdot)\|^2 \xrightarrow{\mathcal{L}} \|Z\|^2$$

alors le test de région critique

$$\{\|S_n(\cdot)\|^2 \geq w_n\}$$

et de niveau asymptotique  $\alpha$  a pour puissance asymptotique :

$$P_K(\alpha) = P\{\|Z\|^2 \geq w_\alpha\}$$

où  $w_\alpha$  est défini par

$$\alpha = P\left\{ \sum_{j \geq 1} \lambda_j^2 U_j^2 \geq w_\alpha \right\}$$



2. Si

$$\nu_n = \left(1 + \frac{g_n}{\sqrt{n}}\right) \cdot \mu$$

où  $g_n$  converge faiblement avec

$$\|g_n\|^2 \rightarrow \gamma^2$$

et

$$\frac{\sup_{x \in E} |g_n(x)|^3}{\sqrt{n}} \rightarrow 0$$

alors le test de Neyman-Pearson de  $\mu$  contre  $\nu_n$  de niveau asymptotique  $\alpha$  a pour puissance  $P[U_1 \geq N_\alpha - \gamma]$ , où  $N_\alpha$  vérifie  $\alpha = P[U_1 > N_\alpha]$ .  
On en déduit alors que le test

$$\|S_n(\cdot)\|^2 > w_n$$

a pour efficacité asymptotique :

$$e(\alpha) = \frac{P\left[\sum_{j \geq 1} \lambda_j^2 (U_j + \alpha_j)^2 \geq w_n\right]}{P[U_1 \geq N_\alpha - \gamma]}$$

Par ailleurs, les résultats obtenus par Gregory dans [47] permettent d'avoir des tests asymptotiquement optimaux. Ces tests optimaux correspondent aux noyaux de la forme

$$K = 1 + \frac{g}{\|g\|} \otimes \frac{g}{\|g\|} + \sum_{j \geq 2} \lambda_j e_j \otimes e_j$$

où  $\lambda_j < 1$  pour tout  $j \geq 2$ .

### 1.3.7 Efficacité asymptotique de BAHADUR

#### 1.3.7.1 Définition

**Définition 1.** Soient  $(T_n)$  et  $(V_n)$  deux séquences de statistiques basées sur  $n$  observations, choisies pour tester l'hypothèse nulle de base  $H$  contre l'hypothèse alternative.

On suppose que l'alternative est caractérisée par un paramètre  $\theta$  tel que  $\theta = \theta_0$  soit dans  $H$ .

Notons par  $N_T(\alpha, \beta, \theta)$  la taille de l'échantillon nécessaire à la séquence  $(T_n)$  pour obtenir la puissance  $\beta$  et le niveau  $\alpha$  pour la valeur alternative du paramètre  $\theta$ .

$N_V(\alpha, \beta, \theta)$  est défini de la même façon pour la séquence  $(V_n)$ .

L'efficacité relative de la séquence  $(T_n)$  par rapport à la séquence  $(V_n)$  est définie par la quantité :

$$e_{T,V}(\alpha, \beta, \theta) = \frac{N_V(\alpha, \beta, \theta)}{N_T(\alpha, \beta, \theta)}$$

La quantité

$$\lim_{n \rightarrow \infty} e_{T,V}(\alpha, \beta, \theta)$$

si elle existe est appelée : Bahadur asymptotic relative efficiency "ARE" (efficacité asymptotique de BAHADUR). Pour les détails, voir [65] p10-11.

En fait, l'efficacité asymptotique de Bahadur a été introduite par Bahadur depuis 1960 dans son article "Stochastic comparison of tests" dans *Ann. Math. Statist.* 31, p276-295.

### 1.3.7.2 Hypothèses et notations

Les hypothèses sont celles de 1.3.5.1, où on a posé

$$K = \sum_{j>0} \lambda_j e_j \otimes e_j$$

$$K' = K - 1 = \sum_{j>1} \lambda_j e_j \otimes e_j \quad (1)$$

$$\sup_{j>1} \sup_{x \in E} |e_j(x)| < +\infty, \quad |\lambda_j| \rightarrow 0, \quad \text{et} \quad \sum_{j=1}^{+\infty} \lambda_j^2 < \infty$$

Posons

$$\Delta(\nu) = \left\| \int K'(x, \cdot) d\nu(x) \right\|, \quad \nu \in H_1$$

### 1.3.7.3 Efficacité relative de Bahadur

#### n) Le test hilbertien

**Proposition 10.** *La pente de Bahadur du test défini par*

$$\|S_n(\cdot)\|^2 \geq u_n$$

*est donnée par*

$$C_S^{K'}(\nu) = \frac{\Delta(\nu)}{\lambda_1^2} (1 + O(1))$$

*dès que  $\Delta(\nu) > 0$ .*

*Démonstration.* Voir Nikitin [65] p 6-7 et 23. ■

**Remarques :**

Si

$$\nu = (1 + h) \cdot \mu$$

avec

$$|h| < 1, \int h d\mu = 0, \text{ et } \int h^2 d\mu > 0$$

alors on obtient

$$C_S^{K'}(\nu) = \sum_{j \geq 1} \left( \frac{\lambda_j}{\lambda_1} \right)^2 \left( \int h e_j d\mu \right)^2$$

par conséquent, le meilleur noyau (dans le sens de l'optimalité du test) de la forme ci-dessus  $K' = K - 1$  s'écrit :

$$K'_h = \frac{h}{\|h\|} \otimes \frac{h}{\|h\|}$$

(voir Gregory section 1.3.6.1).

Par suite :

$$C_S^{K'}(\nu) = \|h\|^2$$

b) Le test de Neyman-Pearson

La pente de Bahadur pour le test de Neyman-Pearson est donnée par

$$C_f^{(N,P)} = 2 \int (1+h) \log(1+h) d\mu$$

(voir Van der Vaart [75] p.208).

alors l'efficacité relative de Bahadur du test fonctionnel basé sur  $K'_h$  est :

$$E_n(h) = \frac{\|h\|^2}{2 \int (1+h) \log(1+h) d\mu}$$

Notons que

$$\lim_{\|h\|_\infty \rightarrow 0} E_n(h) = 1$$

Ce qui prouve l'efficacité relative du test fonctionnel par rapport au test de Neyman-Pearson.

## I.4 Tests du $\chi^2$ généralisés

Dans cette section, il s'agira de comparer un test hilbertien au test du  $\chi^2$ . La base orthonormale choisie est celle des polynômes de Legendre en raison de leur simplicité de mise en oeuvre.

Ce test Hilbertien sera simplement appelé "test de Legendre".

### I.4.1 Tests hilbertiens

Nous supposons ici que les variables aléatoires sont à valeurs dans un intervalle  $D$  de longueur finie ou infinie. Soit  $m$  une probabilité sur  $D$ , on se propose de tester l'hypothèse  $H$  "la loi des  $X_i$  est  $m$ ". Comme au I.2.2, on a :

$$a_j = \int_D e_j dm; j = 0, 1, \dots, k.$$

$$\hat{a}_{jn} = \frac{1}{n} \sum_{i=1}^n e_j(x_i).$$

D'où la statistique de test

$$T_n = n \sum_{j=0}^k |\hat{a}_{jn} - a_j|^2$$

et un test de région critique  $T_n > c_n$ , où  $c_n$  est choisi de manière convenable.

Remarque :

Le test du  $\chi^2$  correspond au choix d'un système de la forme

$$|m(B_j)|^{-1/2} 1_{B_j}, \quad j = 0, 1, \dots, k$$

où les  $B_j$  forment une partition de  $D$  telle que  $m(B_j) > 0$  pour tout  $j$ .

Pour effectuer le test de Legendre, on commence par amener les observations sur  $[-1, 1]$  en posant

$$Y_i = 2F(X_i) - 1, \quad i = 1, \dots, n$$

Où  $F$  désigne la fonction de répartition de  $m$  supposée continue. Sous  $H$ , les  $Y_i$  suivent la loi uniforme  $U$  sur  $[-1, 1]$ .

Le système  $(c_j)$  est obtenu en orthonormalisant par rapport à  $U$  les polynômes  $1, x, x^2, \dots$ .

On obtient pour tout  $j$  :

$$c_j(x) = (1 + 2j)^{1/2} P_j(x), \quad x \in [-1, 1]$$

où

$$P_j(x) = \frac{1}{2^j j!} \frac{d^j}{dx^j} (x^2 - 1)^j, \quad x \in [-1, 1]$$

Voici les premières fonctions  $c_j$  :

$$c_0(x) = 1$$

$$c_1(x) = \sqrt{3}x$$

$$c_2(x) = \frac{1}{2}\sqrt{5}(3x^2 - 1)$$

$$c_3(x) = \frac{1}{2}\sqrt{7}(5x^3 - 3x)$$

$$c_4(x) = \frac{3}{8}(35x^4 - 30x^2 - 3)$$

$$c_5(x) = \frac{1}{8}\sqrt{11}(63x^5 - 70x^3 + 15x)$$

$$c_6(x) = \frac{1}{16}\sqrt{13}(231x^6 - 315x^4 + 105x^2 - 5)$$

Il reste à examiner la question pratique.

## 1.4.2 Choix des cellules dans les tests du $\chi^2$

Des études faites ont modifié les critères de choix des cellules dans le test du  $\chi^2$ .

Si le choix des cellules équiprobables sous  $H$  est considéré comme acquis, c'est le nombre de cellules à choisir qui a été critiqué par les auteurs tels que Best-Rayner (1981,1982,1985), Kallenberg (1985).

Ils se sont rendus compte que le nombre  $k'=k+1$  de cellules assurant une bonne puissance du test était beaucoup plus petit que ne l'indiquaient les règles usuelles de Mann-Wald (1942) et Kendal (1973) qui conseillent de choisir  $k'$  de l'ordre de  $n^{2/5}$ .

Les alternatives étant contaminantes et contiguës, le niveau à 5% et  $n$  (gal à 100). BEST-RAYNER proposent une règle du type suivant :  $2 \leq k' \leq 6$ . Si la queue de l'alternative est plus faible que celle de  $m$ ,  $k'$  voisin de 10 si les queues sont analogues,  $15 \leq k' \leq 20$  si la queue de l'alternative est nettement plus lourde. Dans des conditions du même genre la formule de Mann-Wald donne  $k' = 23$ . La différence est donc sensible. Les simulations dont nous rendons compte ici confirment la nécessité de choisir  $k'$  petit pour avoir un bon test.

## 1.4.3 Choix du nombre de polynômes dans le test de Legendre

Le choix de ce nombre est lié au choix de la dimension dans la construction de l'estimateur de la densité par projection.

Dans [15], Hart donne des méthodes pour le choix de l'ordre des coefficients de Fourier intervenant dans l'estimation de la densité.

En pratique, on fait une première simulation pour déterminer un ordre de grandeur raisonnable pour le nombre de polynômes de Legendre entrant dans la statistique  $T_n$ .

Pour des raisons de commodité, le problème a été transposé sur  $D = [0, 1]$  en considérant les  $F(\chi_i)$  au lieu des  $Y_i$  définis au 1.4.1, ce qui permet de prendre en compte comme alternatives un certain nombre de lois Bêta. Dans chaque cas, il a été tiré 100 échantillons et la puissance empirique a été calculée en pourcentage de rejets. Voici les résultats obtenus pour un niveau à 5% et des échantillons de taille 50. A la lecture de ce tableau, on constate que des valeurs de  $k'$  relativement faibles permettent d'obtenir une bonne puissance.

| Paramètres de B(r, s) |     | Nombre de Polynômes retenus |    |     |    |    |    |     |     |
|-----------------------|-----|-----------------------------|----|-----|----|----|----|-----|-----|
| r                     | s   | 1                           | 2  | 3   | 4  | 5  | 6  | 7   | 8   |
| 0,5                   | 0,5 | 13                          | 88 | 87  | 91 | 92 | 93 | 97  | 100 |
| 0,5                   | 0,8 | 79                          | 91 | 92  | 95 | 95 | 95 | 99  | 98  |
| 0,8                   | 0,8 | 6                           | 19 | 22  | 26 | 24 | 25 | 52  | 43  |
| 0,5                   | 1   | 98                          | 98 | 100 | 98 | 98 | 98 | 100 | 100 |
| 0,8                   | 1   | 26                          | 23 | 24  | 22 | 21 | 19 | 58  | 81  |
| 1                     | 1   | 4                           | 5  | 6   | 8  | 8  | 7  | 30  | 31  |
| 0,9                   | 0,9 | 9                           | 7  | 10  | 10 | 11 | 11 | 38  | 37  |
| 1                     | 0,8 | 29                          | 29 | 28  | 27 | 28 | 29 | 75  | 73  |
| 11                    | 11  | 3                           | 5  | 6   | 4  | 3  | 2  | 29  | 32  |

#### I.4.4 Comparaison empirique des tests de Legendre et du $\chi^2$

La comparaison des deux tests s'effectue dans les mêmes conditions que précédemment : calcul des puissances empiriques au niveau 5% et avec des échantillons de taille 50 ou 100.

Compte tenu des résultats préliminaires le nombre de cellules varie de 2 à 10 pour le test du  $\chi^2$  et le nombre de polynômes de LEGENDRE de 2 à 6. Nous donnons en Appendice 2 des courbes de puissance pour différentes hypothèses alternatives  $H_a$ . Le nombre de cellules (respectivement de polynômes) est en abscisse et les puissances en pourcentage en ordonnée. Les courbes de puissance du test de LEGENDRE sont en traits pleins, celles du test du  $\chi^2$  sont en pointillés.

Il apparaît clairement que le test de LEGENDRE est performant et que sa mise en œuvre est simple. Le choix du nombre de polynômes n'est pas un problème majeur car une valeur telle que  $k' = 5$  assure une bonne puissance dans des cas très variés.

La comparaison des deux tests montre la supériorité du test de LEGENDRE lorsque la variance est plus grande sous l'alternative que sous  $H_0$ . Dans le cas contraire on peut considérer que les deux tests sont au pire équivalents.

On trouvera au paragraphe L5 ci-dessous en appendices, le théorème de Sazonov et les courbes de puissances pour différentes hypothèses alternatives  $H_a$  (cf [19] p.49-52).

Après avoir étudié les tests hilbertiens pour les variables aléatoires indépendantes, nous nous intéresserons au prochain chapitre aux tests hilbertiens pour les processus stationnaires et mélangeants. En effet, les variables aléatoires étudiées sont le plus souvent corrélées, il est donc intéressant d'examiner comment les résultats obtenus ci-dessus peuvent être généralisés au cas des variables dépendantes.

## I.5 Appendices

### I.5.1 Appendice 1 : le théorème de Sazonov

Soit  $(U_n)_{n \geq 1}$  une suite de v.a à valeurs dans  $\mathbb{R}^k$ , indépendantes et de même loi  $\nu$ , centrées et admettant un moment d'ordre 2. Soit  $C$  la classe de convexes de  $\mathbb{R}^k$ . Enfin, soit  $t = (t_1, \dots, t_n)$  une partie finie de  $\mathbb{R}^k$  telle que les produits scalaires

$$(U_i, t_i); i = 1, \dots, k$$

soient des v.a.r non corrélées. Alors

$$\sup_{B \in C} |P_n(B) - N(B)| \leq c_0 k^3 \left| \sum_{i=1}^k \rho_i^{(t)} \right| n^{-1/2}, \quad n \geq 1$$

où  $P_n$  désigne la loi de

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n U_i$$

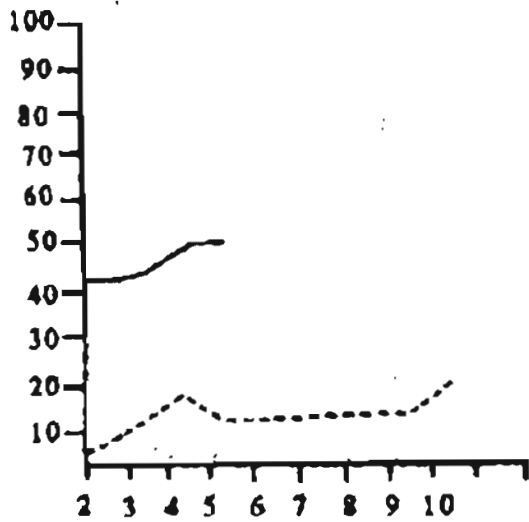
$N$  est la loi normale possédant les mêmes moments d'ordre 1 et 2 que  $\nu$ , et

où on a posé  $\rho_i^{(t)} = \frac{E(|(U_i, t_i)|^3)}{(E(|(U_i, t_i)|^2)^{3/2}}$ ,  $c_0$  désignant une constante universelle

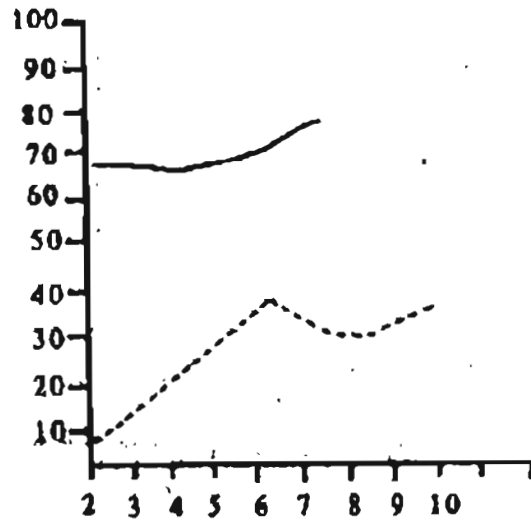
Preuve du théorème de Sazonov Voir [71] p 208

### Appendice 2 : Courbes de puissances pour des hypothèses alternatives $H_a$

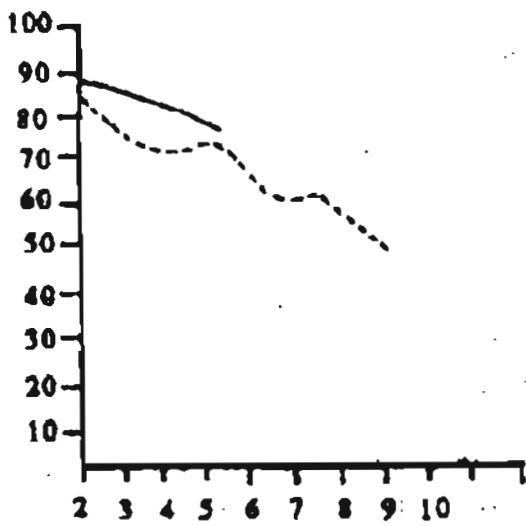




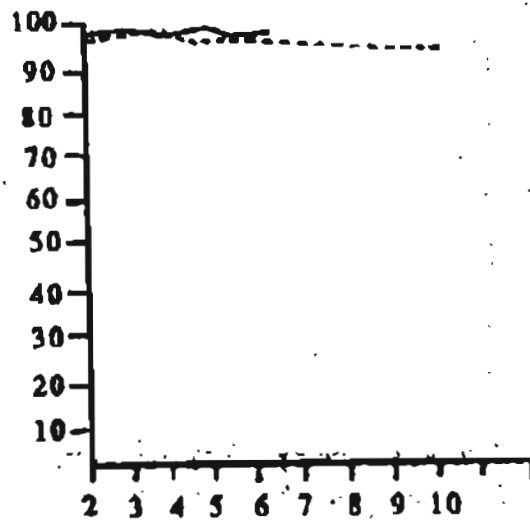
H : N (0,1)  
 $H_a : N (0.25/16)$   
 $n = 50$



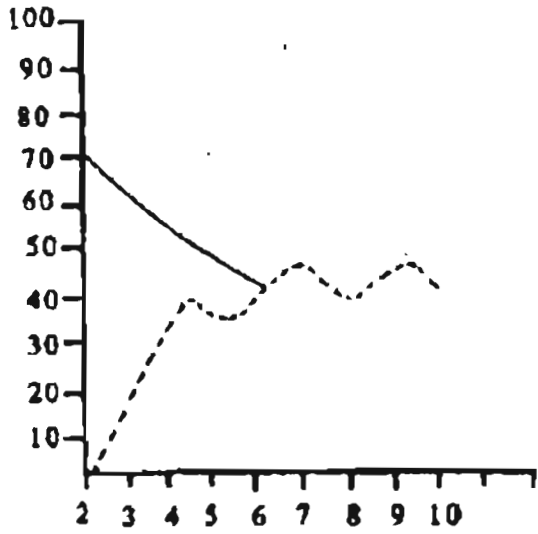
H : N (0,1)  
 $H_a : N (0.25/16)$   
 $n = 100$



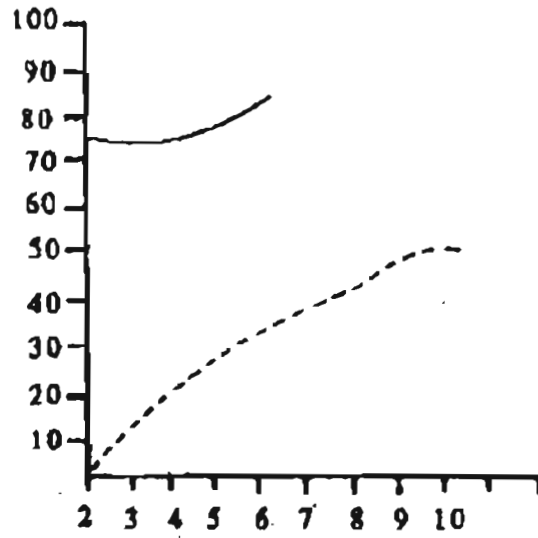
H : N (0,1)  
 $H_a : N (0.5,1)$   
 $n = 50$



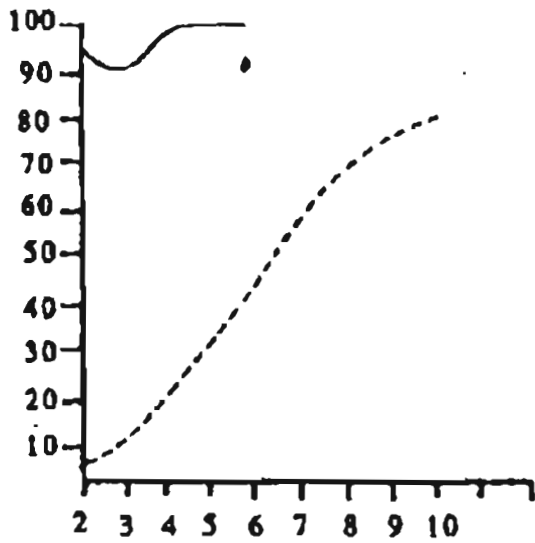
H : N (0,1)  
 $H_a : N (0.5,1)$   
 $n = 100$



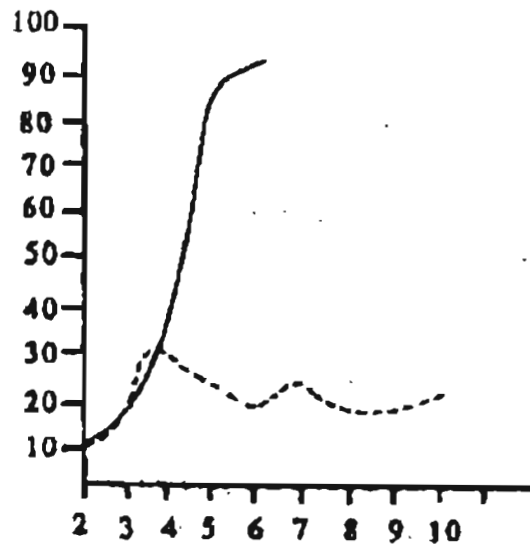
H : N (0,1)  
 $H_a$  : N (0,0.5)  
 $n = 50$



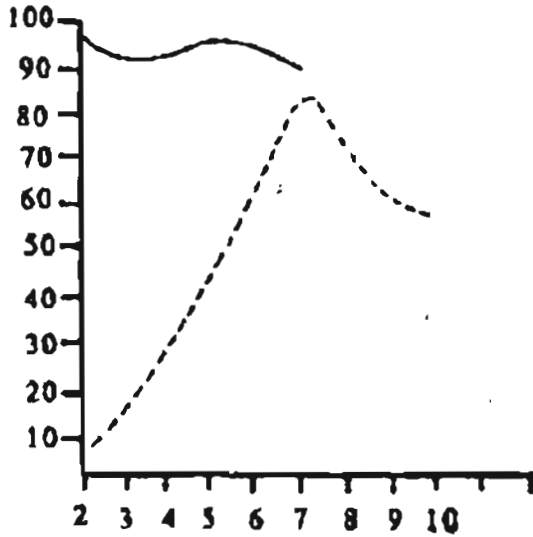
H : N (0,1)  
 $H_a$  : N (0,2)  
 $n = 50$



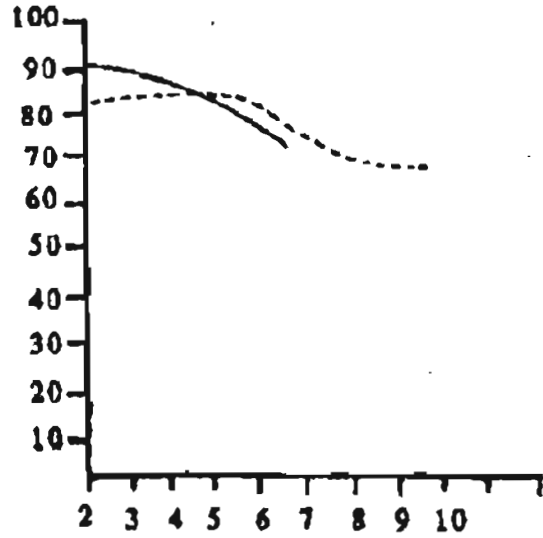
H : N (0,1)  
 $H_a$  : STUDENT (1)  
 $n = 50$



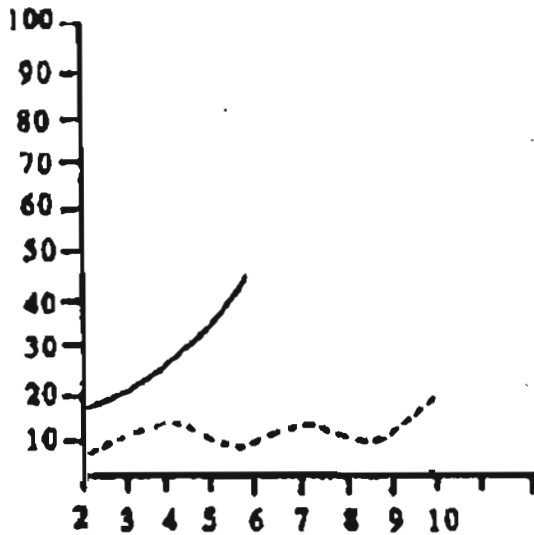
H : N (0,1)  
 $H_a$  : CAUCHY (0,0.5)  
 $n = 50$



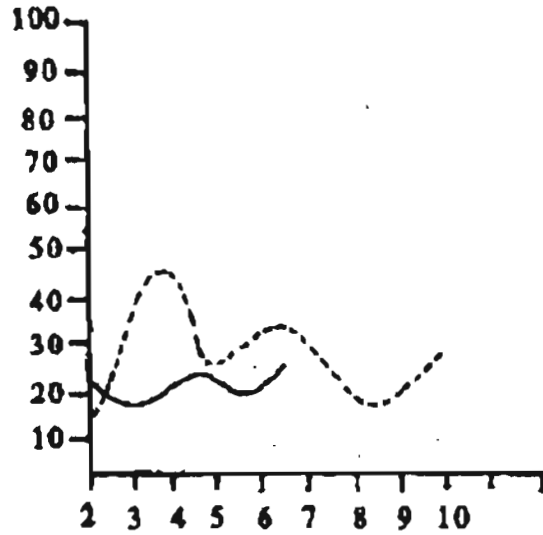
H : CAUCHY (0,1)  
 $H_0$  : STUDENT (25)  
 $n = 50$



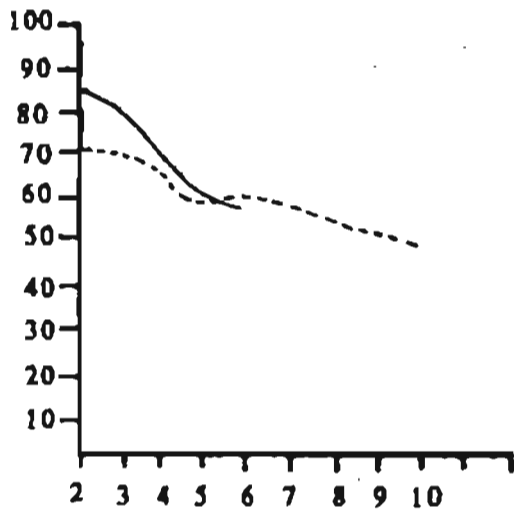
H : U (0,1)  
 $H_0$  : BETA (3/2,1)



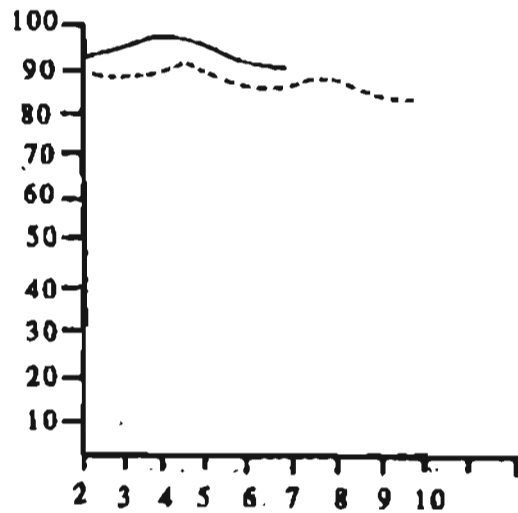
H : N (0,1)  
 $H_0$  : 0,9 N (0,1) + 0,1 N (0,25)  
 $n = 50$



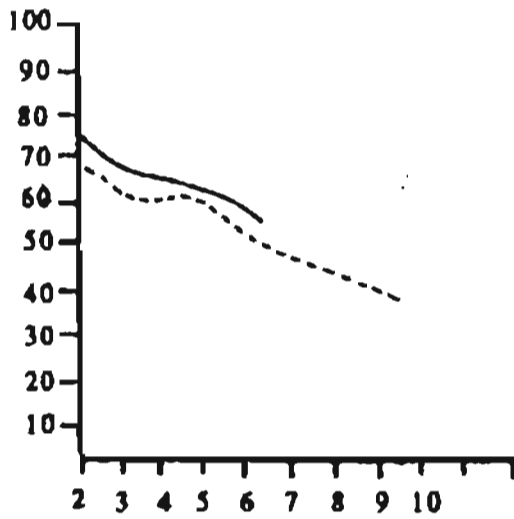
H : N (0,1)  
 $H_0$  : 0,8 N (0,1) + 0,2 N (0,0,04)



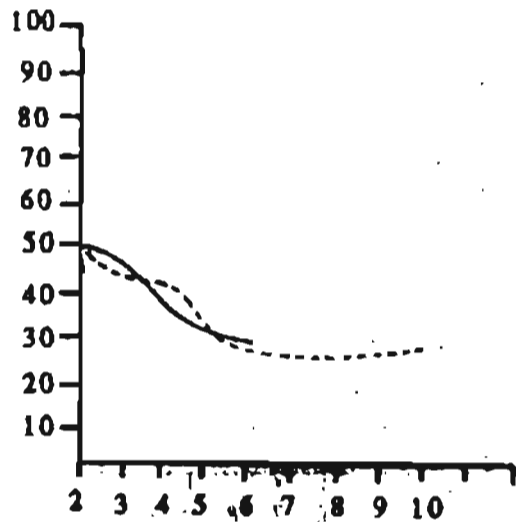
H : U (0,1)  
H<sub>0</sub> : BETA (1,1,4)  
n = 100



H : exp (1)  
H<sub>0</sub> : exp (0,5)  
n = 50



H : exp (0,5)  
H<sub>0</sub> : Γ (0,8,0,6)  
n = 50



H : exp (1)  
H<sub>0</sub> : Γ (0,9,1,2)  
n = 50

# Chapitre II

## ests hilbertiens pour un processus stationnaire et mélangeant

### II.1. Introduction

Dans ce chapitre, on étudie une classe de tests basés sur des normes hilbertiennes pour un processus strictement stationnaire et fortement mélangeant. Des lois limites sont obtenues ainsi que des vitesses de convergence qui précisent le comportement asymptotique des statistiques utilisées et on établit une condition suffisante pour qu'un test de la classe soit convergent.

Les résultats obtenus constituent une généralisation de ceux obtenus par D. Bosq dans [14], [15], [16], [17] (cf. chapitre I). En effet, à partir des statistiques fonctionnelles utilisant des variables aléatoires indépendantes, il a été établi quelques résultats généraux relatifs à ces tests. Nous généralisons certains de ces résultats au cas de processus strictement stationnaires et fortement mélangeant dans le cas discret. Dans la section 2 de ce chapitre, nous donnons des notations et hypothèses générales. Dans la section 3, nous étudions la loi limite de la statistique hilbertienne considérée dans le cas d'un noyau de dimension  $k$ . La vitesse de convergence est abordée à la section 4. Enfin, à la section 5, la construction et les propriétés asymptotiques du test sont étudiées.

### II.2. Notations et hypothèses générales.

Soit  $(X_n, n \geq 1)$  un processus strictement stationnaire et fortement mélangeant défini sur l'espace probabilisé  $(\Omega, \mathcal{A}, P)$  à valeurs dans un espace mesurable

# Chapitre II

## Tests hilbertiens pour un processus stationnaire et mélangeant

### II.1. Introduction

Dans ce chapitre, on étudie une classe de tests basés sur des normes hilbertiennes pour un processus strictement stationnaire et fortement mélangeant. Des lois limites sont obtenues ainsi que des vitesses de convergence qui précisent le comportement asymptotique des statistiques utilisées et on établit une condition suffisante pour qu'un test de la classe soit convergent.

Les résultats obtenus constituent une généralisation de ceux obtenus par D. Bosq dans [14], [15], [16], [17] (cf. chapitre I). En effet, à partir des statistiques fonctionnelles utilisant des variables aléatoires indépendantes, il a été établi quelques résultats généraux relatifs à ces tests. Nous généralisons certains de ces résultats au cas de processus strictement stationnaires et fortement mélangeant dans le cas discret. Dans la section 2 de ce chapitre, nous donnons des notations et hypothèses générales. Dans la section 3, nous étudions la loi limite de la statistique hilbertienne considérée dans le cas d'un noyau de dimension  $k$ . La vitesse de convergence est abordée à la section 4. Enfin, à la section 5, la construction et les propriétés asymptotiques du test sont étudiées.

### II.2. Notations et hypothèses générales.

Soit  $(X_n, n \geq 1)$  un processus strictement stationnaire et fortement mélangeant défini sur l'espace probabilisé  $(\Omega, \mathcal{A}, P)$  à valeurs dans un espace mesurable

$(E, B)$ .

Soit  $\mu$  une probabilité sur  $(E, B)$ , on cherche à éprouver l'hypothèse  $H_0$  " la loi des  $X_i$  est  $\mu$ ".

Pour cela, on commence par effectuer un lissage de la loi empirique à l'aide d'un opérateur défini par un noyau  $K$  (cf. les hypothèses et notations du chapitre I portant sur le noyau  $K$  où on prend successivement  $K$  de dimension fixe  $k$ ,  $K$  de dimension  $k_n$  et enfin  $K$  de dimension infinie).

On considère la statistique à valeurs dans  $L^2(\mu)$  définie par :

$$S_n(\cdot) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right]$$

Sous  $H_0$ , la statistique  $S_n(\cdot)$  définie ci-dessus est égale à :

$$S_n(\cdot) = \sqrt{n} [T_n(\cdot) - E[T_n(\cdot)]]$$

En effet,

$$\forall t \in E, E[T_n(t)] = E \left[ \sum_{i=1}^n K(X_i, t) \right]$$

Or

$$\begin{aligned} E[T_n(t)] &= \int_E \sum_{i=1}^n K(x, t) d\mu(x), \text{ sous } H_0 \\ &= n \int_E K(x, t) d\mu(x) \quad (\text{car les } X_i \text{ sont de même loi}) \end{aligned}$$

$$\int_E K(x, t) d\mu(x) = \langle K(\cdot, t), 1(\cdot) \rangle = 1$$

Par suite  $E[T_n(t)] = n$  et donc,

$$\begin{aligned} S_n(\cdot) &= \frac{1}{\sqrt{n}} [T_n(\cdot) - E[T_n(\cdot)]] \\ &= \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right] \end{aligned}$$

L'objet de l'étude consiste à tester l'hypothèse

$H_0$  " la loi commune des  $X_i$  est  $\mu$  avec  $\forall j \geq 2, \int c_j d\mu = 0$ " contre l'hypothèse

$H_1$  : " la loi commune des  $X_i$  est  $\nu$ , tel qu'il existe au moins un  $j_0 \geq 2$  tel que  $\int c_{j_0} d\nu \neq 0$ ".

Rappelons maintenant la définition d'un processus strictement stationnaire et  $\alpha$ -fortement mélangeant.

**Définitions :**

**Définition 2.**  $(X_n)_{n \geq 1}$  est dit strictement stationnaire, si la distribution des vecteurs aléatoires  $X_m, X_{m+1}, \dots, X_{m+n}$  est indépendante de  $m$ , pour tout  $n \in \mathbb{N}^*$  (voir [35] p.94).

**Définition 3.**  $(X_n)_{n \geq 1}$  est un processus fortement mélangeant, de coefficient de mélangeance  $\alpha$  (on dit encore  $\alpha$ -fortement mélangeant) s'il vérifie :

$\forall n \in \mathbb{N}^*, \forall A \in \mathcal{M}_1^n, \forall B \in \mathcal{M}_{l+n}^\infty, |P(A \cap B) - P(A)P(B)| \leq \alpha(n)$ , où  $\mathcal{M}_1^n$  désigne la  $\sigma$ -algèbre engendrée par  $X_1, X_2, \dots, X_l$  et  $\mathcal{M}_{l+n}^\infty$  celle qui l'est par les variables aléatoires  $X_{l+n}, \dots$ ; et  $\lim_{n \rightarrow \infty} \alpha(n) = 0$  où  $\alpha(n)$  est une suite décroissante de nombres réels positifs ( voir [78] p.156).

## II.3. Loi limite de $\|S_n(\cdot)\|^2$ dans le cas d'un noyau de dimension finie $k$

### II.3.1. Loi limite de $S_n(\cdot)$

$$\text{A partir de } S_n(\cdot) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right]$$

on montre facilement que :

$$S_n(\cdot) = \sqrt{n} \sum_{j=2}^k \hat{a}_{j,n} e_j(\cdot) \text{ où } \hat{a}_{j,n} = \frac{1}{n} \sum_{i=1}^n c_j(X_i)$$

$$\text{C'est-à-dire } S_n(\cdot) = \sum_{j=2}^k \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \right] e_j(\cdot)$$

La recherche de la loi limite se fera en plusieurs étapes.

1<sup>ère</sup> étape : on cherche la loi limite de  $\frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i)$

$(X_n)_{n \geq 1}$  étant strictement stationnaire et  $\alpha$ -fortement mélangeant alors les  $(c_j(X_n))$  le sont également car les  $c_j$  sont mesurables.



**Théorème 1.** *Sous les hypothèses :*

$$1) \sup_{x \in E} \sup_{j \in \{2, \dots, k\}} |c_j(x)| = M < +\infty$$

$$2) \sum_{i=1}^{\infty} \alpha(i) < +\infty$$

3) *La loi des  $X_i$  est  $\mu$  alors*

$$a) \forall j = 2, \dots, k; 1 + 2 \sum_{\nu=1}^{\infty} E|c_j(X_1) c_j(X_{\nu+1})| < +\infty$$

$$b) \forall j = 2, \dots, k; \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z_j \sim N(0, \sigma_j^2)$$

$$\text{où } \sigma_j^2 = 1 + 2 \sum_{\nu=1}^{\infty} E|c_j(X_1) c_j(X_{\nu+1})|$$

**Démonstration.** (voir [3] p.172 et [50] p.440). ■

**2<sup>ème</sup> étape :** On cherche la loi limite du vecteur

$$S'_n = \begin{pmatrix} S_n^2 \\ \vdots \\ S_n^k \end{pmatrix} \text{ où } S_n^j = \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i)$$

Le théorème de Cramer Wold, permet de trouver la loi limite de  $S'_n$ .

Théorème de Cramer Wold : [10] p.49

**Théorème 2.** *Dans  $\mathbb{R}^k$ , la variable aléatoire  $X_n$  converge en loi vers la variable aléatoire  $X$  si et seulement si toute combinaison linéaire des composantes de  $X_n$  converge en loi vers la combinaison linéaire correspondante des composantes de  $X$ .*

En l'appliquant à notre situation, il suffit de montrer que :

$$\left[ \sum_{j=2}^k t_j \left( \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \right) \right] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \sum_{j=2}^k t_j Z_j, \quad t_j \in \mathbb{R} \text{ où } Z_j \text{ suit la loi normale } N(0, \sigma_j^2).$$

Donnons nous la variable aléatoire normale.

$$Z = \begin{pmatrix} Z_2 \\ \vdots \\ Z_j \\ \vdots \\ Z_k \end{pmatrix} \text{ de matrice de covariance } (\sigma_{jj'}) \begin{matrix} j = 2, \dots, k \\ j' = 2, \dots, k \end{matrix}$$

$$\text{on } \left\{ \begin{array}{l} \text{pour } j=j', \sigma_j^2 = 1 + 2 \sum_{\nu=1}^{\infty} E|c_j(X_1)c_j(X_{\nu+1})| \\ \text{pour } j \neq j', \sigma_{jj'} = \sum_{\nu=1}^{\infty} E\{c_j(X_1)c_j(X_{\nu+1}) + c_{j'}(X_1)c_{j'}(X_{\nu+1})\} \end{array} \right.$$

posons :  $\sum_{j'=2, \dots, k} (\sigma_{jj'})$   $j = 2, \dots, k$  et montrons que :

$$\sum_{j=2}^k t_j \left( \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \right) \xrightarrow{L} \sum_{j=2}^k t_j Z_j, \quad t_j \in \mathbb{R}$$

$$\sum_{j=2}^k t_j \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \right] = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[ \sum_{j=2}^k t_j c_j(X_i) \right]$$

posons  $Y_i = \sum_{j=2}^k t_j c_j(X_i)$ , d'où

$$\sum_{j=2}^k t_j \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \right] = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i.$$

Il est immédiat que :  $(Y_n)_{n \geq 1}$  est un processus strictement stationnaire, fortement mélangeant (car la tribu engendrée par  $(X_n)$  contient celle engendrée par  $(c_j(X_n))$  les  $c_j$  étant mesurables).

Sous les deux hypothèses suivantes :

1)  $(Y_n)_{n > 1}$  uniformément borné.

2)  $\sum_{i=1}^{\infty} \alpha(i) < +\infty$ ,

on peut appliquer le théorème 1 au processus  $(Y_n)_{n > 1}$ .

•  $Y_n$  est uniformément borné si  $c_j(X_n)$  l'est, en effet

$$|Y_n| = \left| \sum_{j=2}^k t_j c_j(X_n) \right| \leq \sum_{j=2}^k |t_j| |c_j(X_n)| \leq M \sum_{j=2}^k |t_j| < +\infty$$

$$\left( M = \sup_{x \in E} \sup_{j=2, \dots, k} |c_j(x)| \right).$$

•  $\sum_{i=1}^{\infty} \alpha(i) < +\infty$  (hypothèse 2 du Théorème 1 car  $(c_j(X_n))$  est  $\alpha$ -fortement mélangeant).

On a donc :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \xrightarrow[n \rightarrow +\infty]{L} \mathcal{N}(0, c^2) \text{ avec } c^2 = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \right]^2$$

$$\mathbb{E} \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \right]^2 = \frac{1}{n} \mathbb{E} \left[ \sum_{i=1}^n \sum_{j=2}^k t_j c_j(X_i) \right]^2 = \frac{1}{n} \mathbb{E} \left[ \sum_{j=2}^k t_j \left( \sum_{i=1}^n c_j(X_i) \right) \right]^2$$

$$= \frac{1}{n} \mathbb{E} \left[ \sum_{j=2}^k t_j^2 \left( \sum_{i=1}^n c_j(X_i) \right)^2 \right] + \frac{1}{n} \mathbb{E} \left[ \sum_{j \neq j'} t_j t_{j'} \left( \sum_{i=1}^n c_j(X_i) \right) \left( \sum_{i=1}^n c_{j'}(X_i) \right) \right].$$

Posons

$$A = \frac{1}{n} \mathbb{E} \left[ \sum_{j=2}^k t_j^2 \left( \sum_{i=1}^n c_j(X_i) \right)^2 \right]$$

$$B = \frac{1}{n} \mathbb{E} \left[ \sum_{j \neq j'} t_j t_{j'} \left( \sum_{i=1}^n c_j(X_i) \right) \left( \sum_{i=1}^n c_{j'}(X_i) \right) \right].$$

$$\text{Alors } A = \frac{1}{n} \mathbb{E} \left[ \sum_{j=2}^k t_j^2 \mathbb{E} \left( \sum_{i=1}^n c_j(X_i) \right)^2 \right]$$

$$= \frac{1}{n} \sum_{j=2}^k t_j^2 \left[ \sum_{i=1}^n (\mathbb{E} c_j^2(X_i)) \right] + \sum_{i \neq i'} \mathbb{E} [c_j(X_i) c_j(X_{i'})]$$

Soit par stationnarité du processus  $(c_j(X_n))_{n \geq 1}$ .

$$A = \sum_{j=2}^k t_j^2 \left[ 1 + 2 \mathbb{E} \sum_{\nu=1}^{n-1} \frac{n-\nu}{n} \mathbb{E} (c_j(X_1) c_j(X_{\nu+1})) \right]$$

$$B = \frac{1}{n} \mathbb{E} \left[ \sum_{j \neq j'} t_j t_{j'} \left( \sum_{i=1}^n c_j(X_i) \right) \left( \sum_{i=1}^n c_{j'}(X_i) \right) \right]. \text{ En utilisant la stationnarité}$$

du processus  $(X_n)_{n \geq 1}$  de variables aléatoires, on a :

$$B = \sum_{j \neq j'} t_j t_{j'} \left[ \sum_{\nu=1}^{n-1} \frac{n-\nu}{n} [\mathbb{E} (c_j(X_1) c_j(X_{\nu+1})) + \mathbb{E} (c_{j'}(X_1) c_{j'}(X_{\nu+1}))] \right]$$

D'où le théorème 1 permet d'écrire

$$\mathbb{E}[S_n^j]^2 = 1 + 2 \sum_{\nu=1}^{n-1} \frac{n-\nu}{n} [\mathbb{E} (c_j(X_1) c_j(X_{\nu+1}))] \xrightarrow[n \rightarrow +\infty]{} \sigma_j^2$$

$$\text{et } \sigma_j^2 = 1 + 2 \sum_{\nu=1}^{\infty} \mathbb{E} (c_j(X_1) c_j(X_{\nu+1}))$$

Par conséquent  $A_{n \rightarrow +\infty} \sum_{j=2}^k t_j^2 \sigma_j^2$ .

D'autre part le processus  $(X_n)_{n \geq 1}$  de variables aléatoires réels est strictement stationnaire,  $\alpha$ -fortement mélangeant et  $\mu$ -centré. Les  $c_j(X_n)$  étant uniformément bornés alors, on a :  $\forall j = 2, \dots, k$ ,

$$\sum_{\nu=1}^{\infty} \alpha(\nu) < +\infty \implies \begin{cases} \sum_{\nu=1}^{\infty} |\mathbb{E}(c_j(X_1) c_j(X_{\nu+1}))| < +\infty \\ \sum_{\nu=1}^{\infty} |\mathbb{E}(c_{j'}(X_1) c_j(X_{\nu+1}))| < +\infty \end{cases}$$

Par suite  $\sum_{\nu=1}^{n-1} \frac{n-\nu}{n} \mathbb{E}[(c_j(X_1) c_j(X_{\nu+1})) + c_{j'}(X_1) c_j(X_{\nu+1})] \xrightarrow{n \rightarrow +\infty} \sigma_{jj'}$

avec  $\sigma_{jj'} = \sum_{\nu=1}^{\infty} \mathbb{E}(c_j(X_1) c_j(X_{\nu+1})) + \sum_{\nu=1}^{\infty} \mathbb{E}(c_{j'}(X_1) c_j(X_{\nu+1}))$

d'où

$$B = \sum_{j/j'} t_j t_{j'} \left[ \sum_{\nu=1}^{n-1} \frac{n-\nu}{n} [\mathbb{E}(c_j(X_1) c_j(X_{\nu+1})) + \mathbb{E}(c_{j'}(X_1) c_j(X_{\nu+1}))] \right] \xrightarrow{n \rightarrow +\infty} \sum_{j/j'} t_j t_{j'} \sigma_{jj'}$$

On a donc montré que :

$$A_{n \rightarrow +\infty} \sum_{j=2}^k t_j^2 \sigma_j^2 \text{ et } B_{n \rightarrow +\infty} \sum_{j/j'} t_j t_{j'} \sigma_{jj'}$$

On obtient alors

$$\mathbb{E} \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \right]^2 \xrightarrow{n \rightarrow +\infty} c^2 = \sum_{j=2}^k t_j^2 \sigma_j^2 + \sum_{j/j'} t_j t_{j'} \sigma_{jj'} < +\infty$$

Car si  $\sum_{\nu=1}^{\infty} \alpha(\nu) < +\infty$  alors  $\sum_{\nu=1}^{\infty} |\mathbb{E}(c_j(X_1) c_j(X_{\nu+1}))| < +\infty$

et  $\sum_{i=1}^{\infty} |\mathbb{E}(c_{j'}(X_1) c_j(X_{\nu+1}))|^2 < +\infty$

Après avoir montré que  $\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \xrightarrow{\mathcal{L}} N(0, c^2)$ , nous allons chercher la loi

de la v.a.  $\sum_{j=2}^k t_j Z_j$ .

La v.a.  $Z = \begin{pmatrix} Z_2 \\ \vdots \\ Z_j \\ \vdots \\ Z_k \end{pmatrix}$  étant normale  $\sum_{j=2}^k t_j Z_j$  l'est aussi

(cf. Théorème 7.4. p.67 de[39])

d'où  $\sum_{j=2}^k t_j Z_j$  suit une loi normale  $N\left(0, \text{Var} \left[ \sum_{j=2}^k t_j Z_j \right]\right)$

(puisque les  $Z_j$  sont des v.a. centrées).

$$\text{Var} \left[ \sum_{j=2}^k t_j Z_j \right] = \mathbb{E} \left[ \sum_{j=2}^k t_j Z_j \right]^2 = \sum_{j=2}^k t_j^2 \mathbb{E}[Z_j^2] + \sum_{j \neq j'}^k t_j t_{j'} \mathbb{E}[Z_j Z_{j'}]$$

$$\mathbb{E}[Z_j^2] = \sigma_j^2, \quad j=2, \dots, k.$$

$$\mathbb{E}[Z_j Z_{j'}] = \sigma_{jj'}, \quad j, j'=2, \dots, k.$$

$$\text{donc } \text{Var} \left[ \sum_{j=2}^k t_j Z_j \right] = \sum_{j=2}^k t_j^2 \sigma_j^2 + \sum_{j \neq j'}^k t_j t_{j'} \sigma_{jj'} = c^2.$$

$$\text{D'où, on a bien, } \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \sum_{j=2}^k t_j Z_j$$

$$\text{c'est-à-dire : } \sum_{j=2}^k t_j S_n^j \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \sum_{j=2}^k t_j Z_j.$$

Le théorème de Cramer Wold nous permet alors d'affirmer que

$$S_n \xrightarrow{\mathcal{L}} Z = \begin{pmatrix} Z_2 \\ \vdots \\ Z_j \\ \vdots \\ Z_k \end{pmatrix}$$

où  $Z$  est une v.a. normale dans  $\mathbb{R}^{k-1}$  de moyenne nulle et de matrice de covariance  $\Sigma$  définie par :  $\Sigma = (\sigma_{jj'})$   $\begin{matrix} j=2, \dots, k \\ j'=2, \dots, k \end{matrix}$  où  $\begin{cases} \sigma_j^2 & \text{si } j=j' \\ \sigma_{jj'} & \text{si } j \neq j' \end{cases}$

3ème étape : on cherche la loi limite de  $S_n(\cdot)$

**Théorème 3.** *Sous les hypothèses suivantes :*

1) la loi  $X_i$  est  $\mu$

$$2) \sup_{x \in E} \sup_{j \in \{1, \dots, k\}} |c_j(x)| = M < +\infty$$

$$3) \sum_{i=1}^{\infty} \alpha(i) < +\infty$$

alors la statistique  $S_n(\cdot)$  tend en loi vers la variable aléatoire

$$U = \sum_{j=2}^k Z_j c_j(\cdot).$$

**Démonstration.**

$$\begin{aligned} S_n(\cdot) &= \frac{1}{\sqrt{n}} \sum_{j=2}^k \sum_{i=1}^n c_j(X_i) c_j(\cdot) = \sum_{j=2}^k \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n c_j(X_i) \right] c_j(\cdot) \\ &= \sum_{j=2}^k S_n^j c_j(\cdot) = g(S_n^*) \end{aligned}$$

L'application  $g$  est linéaire continue  $\|S_n(\cdot)\|_{1,2(\mu)}^2 = \|S_n^*(\cdot)\|_{\mathbb{R}^{k-1}}^2$

La convergence en loi se conservant par continuité alors

$$S_n(\cdot) \xrightarrow{L} \sum_{j=2}^k Z_j c_j(\cdot) \quad \blacksquare$$

## II.3.2. Comportement asymptotique de $\|S_n(\cdot)\|^2$

### II.3.2.1. La loi des $X_i$ est $\mu$

**Proposition 11.** Sous les conditions suivantes :

$$1) \sup_{x \in E} \sup_{j \in \{1, \dots, k\}} |c_j(x)| = M < +\infty$$

$$2) \sum \alpha(i) < +\infty$$

alors  $\|S_n(\cdot)\|^2 \xrightarrow[n \rightarrow \infty]{L} Q = \sum_{j=2}^k Z_j^2$  où  $Q$  est une variable aléatoire dont la fonction

caractéristique est égale à  $[\det(I_{k-1} - 2it\Sigma)]^{-1/2}$ ,  $I_{k-1}$  désignant la matrice unitaire  $\mathbb{R}^{k-1}$ .

**Démonstration. Démonstration :**

L'application  $S_n(\cdot) \rightarrow \|S_n(\cdot)\|^2$  étant continue

Alors  $\|S_n(\cdot)\|^2 \xrightarrow{L} Q = \sum_{j=2}^k Z_j^2$  et on conclut en utilisant le théorème 1 p.88

de [12].  $\blacksquare$

### II.3.2.2. la loi des $X_i$ est $\nu$ .

**Proposition 12.** 1) si la loi de probabilité des  $X_n$  est  $\nu$  telle qu'il existe  $j \in \{2, \dots, k\}$  pour lequel  $c_j$  soit  $\nu$ -intégrable et d'intégrale non nulle.

2) Si  $\sup_{x \in E} \sup_{j \in \{2, \dots, k\}} |c_j(x)| = M < +\infty$

3)  $\sum_{i=1}^{\infty} \sqrt{\alpha(i)} < \infty$

alors  $\frac{1}{n} \sum_{i=1}^n c_j(X_i) \xrightarrow[n \rightarrow \infty]{p.s.} E[c_j(X_1)] \neq 0$

de plus  $\|S_n(\cdot)\|^2 \xrightarrow[n \rightarrow \infty]{p.s.} +\infty$

Démonstration. Elle consiste à utiliser la loi forte des grands nombres dans le cas des processus stationnaires (voir [48] p.18). ■

## II.4. Vitesse de convergence

L'étude de la vitesse de convergence se fera en considérant les 3 cas suivants :

- 1) le cas où le noyau  $K$  est de dimension finie  $k$
- 2) le cas où le noyau  $K$  est de dimension finie et varie en fonction de la taille de l'échantillon
- 3) le cas où le noyau  $K$  est de dimension infinie.

L'étude de la vitesse de convergence consiste à obtenir des majorations de la vitesse de convergence de la statistique  $\|S_n(\cdot)\|^2$  vers sa loi limite.

### II.4.1. Cas du noyau de dimension $k$

. Le résultat principal est donné par le

#### Théorème 4.

- 1) Si la loi des  $X_i$  est  $\mu$
- 2) si  $\limsup \sup_{x \in E, j \in \{2, \dots, k\}} |c_j(x)| = M_1 < +\infty$
- 3) si  $\sum_{i \geq 1}^{\infty} \sqrt{\alpha(i)} < \infty$

pour  $n$  assez grand, on a :

$$\sup_{a \in \mathbb{R}^1} |P(\|S_n(\cdot)\|^2 \leq a^2) - P[Q \leq a^2]| = O\left[\frac{1}{n^{(k-1)/k}}\right]^{1/4} + O\left[\frac{1}{\sqrt[4]{n}} a^{(k-2)/2}\right] + O\left[n^{9/16} \alpha(n^{1/4}) a^{(k-2)/2}\right] \quad (2)$$

où  $Q$  est une variable aléatoire dont la fonction caractéristique est donnée par  $[\det(I_{k-1} - 2it \sum)]^{-1/2}$

Démonstration. Elle est assez technique voir [41] p.41 à 68 pour les détails. ■

**Corollaire 1.** *Sous les hypothèses du théorème 4 ; il existe un ensemble de suites  $\{a_n\}$  tel que quelle que soit la suite  $(b_n)$  ( $\lim b_n = +\infty$ ) appartenant à  $(a_n)$  tel que :*

$$\sup_{a < b_n} |P(\|S_n(\cdot)\|^2 \leq a^2) - P[Q \leq a^2]| \xrightarrow{n \rightarrow \infty} 0 ;$$

plus précisément on a :

$$\sup_{a < b_n} |P(\|S_n(\cdot)\|^2 \leq a^2) - P[Q \leq a^2]| = O\left[\frac{1}{n^{1/(k-1)}}\right]^{1/4} + O\left[\frac{1}{\sqrt[4]{n}} b_n^{(k-2)/2}\right] + O\left[n^{9/16} \alpha(n^{1/4}) b_n^{(k-2)/2}\right] \quad (3)$$

Remarque :

La suite  $(b_n)$  existe , il suffit de prendre  $b_n = \log(2+n)$  et le remplacer dans (2) pour obtenir le résultat.

**Corollaire 2.** *Sous les hypothèses du théorème 4 :*

1) si  $b_n = n^{1/(k-1)}$

2) si  $\alpha(n) = a n^\rho$  ,  $a > 0$  ,  $0 \leq \rho < 1$  et  $n \geq 1$

Alors

$$\sup_{a < b_n} |P(\|S_n(\cdot)\|^2 \leq a^2) - P[Q \leq a^2]| = O\left[\frac{1}{n^\gamma}\right]^{1/4} \quad (4)$$

où  $\gamma = \inf\left(\frac{k-1}{k}, \frac{k}{2(k-1)}\right)$  ,  $k \geq 2$ .

## II.4.2. Cas d'un noyau de dimension finie $k_n$ qui varie avec la taille de l'échantillon

. Le résultat principal est le suivant :

**Théorème 5 .**



1) si la loi des  $X_i$  est  $\mu$

2) sous les hypothèses

$$a) n \in \mathbb{N}^*, \sup_{x \in E} \sup_{j \in \{2, \dots, k_n\}} |c_j(x)| = M_2 < +\infty$$

$$b) \lim_{n \rightarrow +\infty} k_n = +\infty \text{ et } \lim_{n \rightarrow +\infty} \frac{k_n}{n} = 0$$

Pour  $n$  assez grand, on a :

$$\begin{aligned} \sup_{a > 0} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U\|^2 \leq a^2]| = & O\left[\frac{k_n^2}{n}\right]^{1/6} + O\left[a \frac{k_n^6}{n}\right]^{1/24} \\ & + O\left[\frac{1}{k_n} a^{k_n}\right] + O\left[(na)^{k_n} \alpha(n^{1/4})\right]^{1/6} \end{aligned} \quad (5)$$

$$\text{où } U_{k_n} = \sum_{j=2}^{k_n} Z_j c_j(\cdot) \text{ avec } Z_j \approx N(0, \sigma_j^2)$$

**Démonstration.** Il suffit de réadapter la démonstration du cas précédent au cas où le noyau est de dimension finie qui varie avec la taille de l'échantillon (pour plus de détails voir [41] pp.73-80) ■

**Remarque.**

On montre de même que, comme précédemment, il existe une suite  $(b_n)$  telle que  $\lim b_n = +\infty$  et  $\sup_{n < b_n} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U_{k_n}\|^2 \leq a^2]| \xrightarrow[n \rightarrow +\infty]{} 0$

Plus précisément, on montre le corollaire suivant :

**Corollaire 3.** Sous les hypothèses du théorème 5, si en plus

$$1) b_n = \log(2+n), n \geq 1$$

$$2) k_n = \lfloor n^\beta \rfloor, 0 < \beta < 1/6$$

$$3) \alpha(n) = a \rho^n, a > 0, 0 \leq \rho < 1/n, n \geq 1$$

Alors

$$\sup_{n < b_n} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U_{k_n}\|^2 \leq a^2]| = O[n^{\delta/4} \log(2+n)] \quad (6)$$

$$\text{où } -\frac{1}{6} < \delta < 0$$

**Remarque :**

D'autres choix de  $k_n$  et  $b_n$  peuvent être faits. Cependant le problème de l'optimalité de la vitesse de convergence reste posé, puisque la convergence de la vitesse de convergence dépend de la suite  $(b_n)$ .

### II.4.3. Cas où le noyau est de dimension infinie

$$\text{posons : } S_n(\cdot) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right]$$

$$K(X_i, \cdot) = \sum_{j=1}^{\infty} \lambda_j e_j(X_i) e_j(\cdot)$$

$$U = \sum_{j=2}^{\infty} \lambda_j Z_j e_j(\cdot) \text{ où } Z_j \sim N(0, \sigma_j^2)$$

Le théorème ci-dessous donne la vitesse de convergence de  $\|S_n(\cdot)\|^2$  vers sa loi limite.

#### **Théorème 6**

Les  $X_i$  étant de loi  $\mu$ , si en plus :

1)  $\sup_{x \in F} \sup_{j \geq 1} |e_j(x)| = M < +\infty$

2)  $\sum_{i=1}^{\infty} \sqrt{a(i)} < +\infty$

3)  $\lim_{n \rightarrow +\infty} k_n = +\infty$  avec  $\lim_{n \rightarrow +\infty} \frac{k_n^4}{n} = 0$

Alors, il existe une suite  $(b_n)$  telle que  $\lim b_n = +\infty$  et

$$\sup_{0 < a \leq b_n} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U\|^2 \leq a^2]| = O\left[\frac{k_n^6}{n}\right]^{1/6} + O\left[\left(\frac{k_n^6}{n}\right)^{1/24} b_n\right]$$

$$+ O\left[\frac{b_n}{k_n}\right]^{k_n} + O\left[\Lambda_{k_n}^{2/3}\right] + O\left[\Lambda_{k_n}'^{2/3}\right] + O[(nb_n)^{k_n} \alpha(n^{1/4})]$$

(7)

où  $\Lambda_{k_n} = \sum_{j=k_n+1}^{\infty} \lambda_j^2 \sigma_j^2$ ,  $\Lambda_{k_n}' = \sum_{j=k_n+1}^{\infty} \lambda_j^2$ .

**Démonstration.** Elle se fait en trois étapes, pour les détails voir [41] p.p.39-87.

On considère successivement le cas où le noyau  $K$  est de dimension finie  $k$ , puis le cas où la dimension varie avec la taille de l'échantillon  $k_n$ , et enfin le cas où la dimension est infinie.

**1<sup>ère</sup> étape :**  $K$  de dimension fixe  $k$ .

On étudie  $\sup_{a \in \mathbb{R}^+} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U_{k_n}\|^2 \leq a^2]|$

a) on adopte la démarche utilisée par D. Bosq dans [13] et on pose alors que :

$$S_n = Y_n + Z_n$$

on fait l'hypothèse  $(C_1)$  pour toute la suite de cette partie.

$$C_1 : \begin{cases} \sup_{x \in E} \sup_{j \in \{2, \dots, k\}} |c_j(x)| = M_1 < +\infty \\ \sum_{i=1}^{\infty} \sqrt{\alpha(i)} < +\infty \end{cases}$$

et on pose

$$W_i = \begin{pmatrix} \frac{1}{\sqrt{n}} c_2(X_i) \\ \vdots \\ \frac{1}{\sqrt{n}} c_k(X_i) \end{pmatrix} \quad \begin{cases} E[||W_i||^2] = \Delta_{n,k}^2 \\ E[||S'_n||^2] = A_{n,k}^2 \end{cases}$$

$\Delta_{n,k}^2$  et  $A_{n,k}^2$  existent du fait de  $(C_1)$ .

On fait alors le regroupement suivant :

Soit  $h$  le plus grand entier tel que :

$$E \left[ \left\| \sum_{i=1}^h W_i \right\|^2 \right] \leq \Delta_{n,k} A_{n,k}$$

Soit  $m$  la partie entière de  $\left[ \frac{A_{n,k}}{\Delta_{n,k}} \right]^{1/2}$

On pose :

$$Y_1 = W_1 + \dots + W_h$$

$$Y_2 = W_{h+m+1} + \dots + W_{2h+m}$$

$$\vdots$$

$$Y_i = W_{(i-1)(h+m)+1} + \dots + W_{ih+(i-1)m}$$

$$\vdots$$

$$Y_l = \dots$$

$$Z_1 = W_{h+1} + \dots + W_{h+m}$$

$$Z_2 = W_{2h+m+1} + \dots + W_{2h+2m}$$

$$\vdots$$

$$Z_i = W_{ih+(i-1)m+1} + \dots + W_{i(h+m)}$$

$$\vdots$$

$$Z_l = \dots$$

$$Z_{l+1} = \begin{cases} 0 & \text{si } lm+lh=n \\ W_{l(m+h)+1} + \dots + W_n & \text{sinon} \end{cases}$$

$l$  étant déterminé par l'impossibilité de construire  $Y_{l+1}$  et  $Z_{l+1}$  à la fois selon le schéma  $Y_1, Z_1, \dots, Y_l, Z_l, \dots$  tel que

$l(m+h) \leq n < (l+1)(m+h)$ , pour les détails des calculs voir [15] p.p.42-47.

or  $\|S'_n\|_{\mathcal{L}^2(\mu)}^2 = \|S_n(\cdot)\|_{L^2(\mu)}^2$

donc ce qui permettra de remplacer  $\|S_n(\cdot)\|^2$  par  $\|Y_n + Z_n\|^2$

b) on utilise ensuite les résultats d'Essen [38] p.p. 1-125 les calculs sont assez techniques.

**2<sup>me</sup> étape** :  $K$  de dimension  $k_n$ .

On reprend la même démarche que précédemment en posant comme hypothèses : en plus des conditions  $C_1$ , l'hypothèse  $C_2$  suivante :

$$C_2 : \begin{cases} \sup_{x \in E} \sup_{j \in \{2, \dots, k_n\}} |e_j(x)| = M_2 < +\infty \\ \lim_{n \rightarrow +\infty} k_n = +\infty, \quad \lim_{n \rightarrow +\infty} \frac{k_n^4}{n} = 0 \end{cases}$$

et on cherche

$$\sup_{n \in \mathbb{N}^*} |\mathbb{P}[\|S'_{k_n}(\cdot)\|^2 \leq n^2] - \mathbb{P}[\|U_{k_n}\|^2 \leq n^2]|$$

$$\text{où } S'_{k_n}(\cdot) = \sqrt{n} \sum_{j=2}^{k_n} \hat{\alpha}_{j,n} e_j(\cdot)$$

$$U_{k_n} = \sum_{j=2}^{k_n} Z_j e_j(\cdot)$$

$\|U_{k_n}\|^2$  est une variable aléatoire dont la fonction caractéristique est donnée par

$$[\det(I_{k_n} - 2it \sum_{j=2}^{k_n} \sigma_{jj}^{-1})]^{-1/2}$$

$$\sum_{j=2}^{k_n} = (\sigma_{jj}^{-1}) \quad 2 \leq j, j' \leq k_n$$

**3<sup>me</sup> étape** :  $K$  de dimension infinie.

a) introduction et notation.

On pose ici :

$$S_{k_n}(\cdot) = \sqrt{n} \sum_{j=2}^{k_n} \lambda_j \hat{\alpha}_{j,n} e_j(\cdot); \quad \lim_{n \rightarrow +\infty} k_n = +\infty$$

$$U_{k_n} = \sum_{j=2}^{k_n} \lambda_j Z_j e_j(\cdot)$$

$$U = \sum_{j=2}^{\infty} \lambda_j Z_j e_j(\cdot)$$

$$\text{et } S_n(\cdot) = \sqrt{n} \sum_{j=2}^{\infty} \lambda_j \hat{\alpha}_{j,n} e_j(\cdot)$$

Et on étudie

$$\sup_{n \in \mathbb{N}^*} |\mathbb{P}[\|S_n(\cdot)\|^2 \leq n^2] - \mathbb{P}[\|U\|^2 \leq n^2]|$$

En plus des hypothèses  $C_1$  et  $C_2$  on suppose  $C_3$  réaliste

$$C_3 : \begin{cases} \sup_{x \in E} \sup_{j \geq 1} |c_j(x)| = M_3 < +\infty \\ \sum_{j \geq 1} \lambda_j^2 < +\infty \end{cases}$$

b) Étude de la vitesse de convergence.

Il faut majorer la quantité suivante :

$$\sup_{a \in \mathbb{R}^1} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U\|^2 \leq a^2]|$$

On montre que :

$$\sup_{a \in \mathbb{R}^1} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U\|^2 \leq a^2]| \leq$$

$$\sup_{a \in \mathbb{R}^1} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|S_{k_n}\|^2 \leq a^2]| + \sup_{a \in \mathbb{R}^1} |P[\|S_{k_n}(\cdot)\|^2 \leq a^2] - P[\|U_{k_n}\|^2 \leq a^2]| + \sup_{a \in \mathbb{R}^1} |P[\|U_{k_n}(\cdot)\|^2 \leq a^2] - P[\|U\|^2 \leq a^2]|$$

La recherche de la vitesse de convergence consistera à majorer les 3 quantités ci-dessous, à savoir

$$\sup_{a \in \mathbb{R}^1} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|S_{k_n}\|^2 \leq a^2]|,$$

$$\sup_{a \in \mathbb{R}^1} |P[\|S_{k_n}(\cdot)\|^2 \leq a^2] - P[\|U_{k_n}\|^2 \leq a^2]|, \text{ et}$$

$$\sup_{a \in \mathbb{R}^1} |P[\|U_{k_n}(\cdot)\|^2 \leq a^2] - P[\|U\|^2 \leq a^2]|$$

Pour les détails des calculs, ils sont similaires à [41] p.p. 82-89. ■

**Remarques :**

1) La suite  $(b_n)$  a été choisie de telle manière qu'il y ait convergence du test, c'est-à-dire soit  $\{(a_n)\}$  l'ensemble des suites telles que quelle que soit la suite  $(b_n)$  appartenant à cet ensemble, on ait :  $\sup_{a \leq b_n} |P[\|S_n(\cdot)\|^2 \leq a^2] - P[\|U_{k_n}\|^2 \leq a^2]| \rightarrow 0$

2) l'ensemble des suites  $(a_n)$  n'est pas vide, il suffit de prendre par exemple  $b_n = \log(2+n)$ ,  $n \geq 1$  pour assurer le résultat, en supposant bien entendu que :

$$\sum_{j=1}^{\infty} \lambda_j^2 \sigma_j^2 < +\infty \quad k_n \text{ et } \alpha(n) \text{ étant convenablement choisis.}$$

3) Exemples :

**Corollaire 4.** 1) Si la loi des  $X_i$  est  $\mu$

2) sous les hypothèses suivantes :

$$\sup_{x \in E} \sup_{j \geq 1} |c_j(x)| = M < +\infty$$

- $\sum_{j=1}^{\infty} \lambda_j^2 < +\infty ; \sum_{j=1}^{\infty} \lambda_j^2 \sigma_j^2 < +\infty$
  - $k_n = [nr^\beta], 0 < \beta < 1/6$  où  $[.]$  désigne la partie entière
  - $\alpha(n) = a\rho^n ; a > 0$  et  $\rho < \frac{1}{n}, n \geq 1$
- Alors

$$\sup_{n < k_n(2+n)} |P(\|S(\cdot)\|^2 \leq a^2) - P(\|U\|^2 \leq a^2)| = O[n^{\delta/4} \log(2+n)] + O\left[\Lambda_{k_n}^{2/3}\right] + O\left[\Lambda_{k_n}^{2/3}\right] \quad ((8))$$

où  $\frac{1}{6} < \delta < \theta$ .

**Corollaire 5.** Sous les hypothèses du corollaire 4 et si en plus  $\Delta_{k_n} = O\left[\frac{1}{k_n}\right]$

et  $\Lambda_{k_n} = O\left[\frac{1}{k_n}\right]$

alors

$$\sup_{n < k_n(2+n)} |P(\|S(\cdot)\|^2 \leq a^2) - P(\|U\|^2 \leq a^2)| = O[n^{\delta/4} \log(2+n)] \quad ((9))$$

où  $\frac{1}{6} < \delta < \theta$ .

**Remarque :**

Les vitesses de convergence obtenues vont nous permettre d'établir quelques résultats généraux relatifs à toute une classe de tests dont le test du  $\chi^2$  est un cas particulier.

En effet, dans le cas où le noyau est de dimension finie et si on pose

$$K(x,t) = \sum_{j=1}^k p_j^{-1} \mathbf{1}_{A_j}(x) \mathbf{1}_{A_j}(t), (x,t) \in E \times E \text{ où les } A_j \text{ forment une } k\text{-partition}$$

de  $E(p_j = \mu(A_j))$ , on retrouve le test du  $\chi^2$  usuel, puisque :

$$\|S_n(\cdot)\|^2 = \sum_{j=1}^k \frac{1}{np_j} \left[ \sum_{i=1}^n \mathbf{1}_{A_j}(x_i) - np_j \right]^2. \text{ ( cf 1 2.2)}$$

## II.5. Construction et propriétés asymptotiques du test

### II.5.1. Construction du test.

#### II.5.1.1. Cas du noyau de dimension finie $k$

Nous allons tester l'hypothèse  $H_0$  "La loi des  $X_i$  est  $\mu$ " contre l'hypothèse  $H_1(k)$  "La loi des  $X_i$  est une probabilité  $\nu$  telle que  $\int e_j d\nu$  existe et est non nulle pour au moins un  $j$  dans  $\{2, \dots, k\}$ "  
 sous  $H_0 : \|S_n(\cdot)\|^2 \xrightarrow{L} Q$  (Proposition 11) sous  $H_1(k) : \|k_n(\cdot)\|^2 \xrightarrow{P_n} +\infty$  (proposition 12) Alors le test est de la forme :

$$n \sum_{j=2}^k \hat{\eta}_{jn}^2 > w_n.$$

Soit  $\alpha \in ]0,1[$  et  $w$  tels que

$P[Q > w] = \alpha$ . On en déduit que le test  $n \sum_{j=2}^k \hat{\eta}_{jn}^2$  est de niveau asymptotique  $\alpha$ .

#### II.5.1.2. Cas où le noyau est de dimension finie qui varie avec la taille de l'échantillon $k_n$

Pour tester l'hypothèse  $H_0$  " la loi commune des  $X_i$  est  $\mu$ " on peut utiliser un test de la forme

$$n \sum_{j=2}^{k_n} \hat{\eta}_{jn}^2 > w_n.$$

L'alternative  $H_1(k_n)$  sera l'ensemble des lois  $\nu$  tel que, pour au moins un  $j \in \{2, \dots, k_n\}$   
 $\int e_j d\nu$  existe et soit non nulle.  $H_1(k_n)$  contient alors toutes les lois (différentes de  $\mu$ ) et admettant une densité appartenant à  $L^2(\mu)$ .

### II.5.1.3. Cas du noyau de dimension infinie

Pour tester l'hypothèse  $H_0$  : " La loi des  $X_i$  est  $\mu$ " contre l'hypothèse  $H_1(\lambda_n)$  (où  $1 \geq |\lambda_n| > |\lambda_{n+1}| > 0$  pour  $n \geq 1$  et  $\sum_{n \geq 2} \lambda_n^2 < +\infty$ ). " La loi des  $X_i$  est  $\nu$  telle qu'il existe au moins un  $j \geq 2$  t.q.  $\int e_j d\nu$  existe et soit non nulle", on peut utiliser le test de région critique  $\|S_n(\cdot)\|^2 > w_n$ .

### II.5.2. Propriétés asymptotiques du test : convergence du test

Rappelons qu'un test est convergent si son niveau  $\alpha_n$  tend vers 0 et sa puissance  $p_n(\nu)$  tend vers 1 pour toute loi de l'alternative.

#### II.5.2.1. Cas du noyau de dimension finie k

Si  $\alpha_n$  est le niveau du test  $n \sum_{j=2}^k \hat{a}_{jn}^2 > w_n$  alors  $\alpha_n = P \left[ n \sum_{j=2}^k \hat{a}_{jn}^2 \geq w_n \right]$ .

**Proposition 13.** 1) Sous les hypothèses :

a)  $\sup_{x \in E} \sup_{j \in \{2, \dots, k\}} |e_j(x)| = M < +\infty$

b)  $\left(\frac{1}{n}\right)^{1/4} b_n^{(k-1)/4} \xrightarrow[n \rightarrow \infty]{} 0$

c)  $n^{9/16} \alpha(n^{1/4}) b_n^{(k-1)/4} \xrightarrow[n \rightarrow \infty]{} 0$

d)  $\begin{cases} w_n \leq b_n \\ w_n \xrightarrow[n \rightarrow \infty]{} +\infty \\ \frac{w_n}{n} \xrightarrow[n \rightarrow \infty]{} 0 \end{cases}$

Alors le test  $n \sum_{j=2}^k \hat{a}_{jn}^2 > w_n$  est convergent.

2) La convergence de la puissance est uniforme sur tout  $H_2 \subset H_1(k)$  telle que :

$$\delta(H_2) = \inf_{\nu \in H_2} \max_{j \in \{2, \dots, k\}} \left[ \int e_j d\nu \right]^2 > 0$$



3) Soit  $(\nu_n) \subset H_1(k)$  telle que :

$$\frac{n}{4(k-1)} \sum_{j=2}^k \left[ \int e_j d\nu \right]^2 \geq w_n ; n \geq 1$$

Alors si a) et b) sont vérifiées, on a simultanément  $\alpha_n \rightarrow 0$  et  $p_n(\nu_n) \rightarrow 1$ .

**Démonstration.** 1) utilisant le corollaire 1 on a :

$$\sup_{w_n \leq b_n} |P(Q > W_n) - \alpha_n| = O\left[\frac{1}{n^{(k-2)/4}}\right]^{1/4} + O\left[\left(\frac{1}{n}\right)^{1/4} (b_n)^{(k-2)/4}\right] + O\left[n^{9/16} \alpha(n^{1/4}) b_n^{(k-2)/4}\right]$$

D'où sous les hypothèses a, b et c on a :

$$\sup_{w_n \leq b_n} |\alpha_n - P[Q \leq W_n]| \xrightarrow{n \rightarrow +\infty} 0$$

$$\text{donc si } w_n \leq b_n \quad |\alpha_n - P[Q \leq W_n]| \xrightarrow{n \rightarrow +\infty} 0$$

par suite  $\alpha_n \rightarrow 0 \iff w_n \rightarrow +\infty$ .

Supposons maintenant  $\frac{w_n}{n} \rightarrow 0$ .

On va alors considérer  $\nu \in H_1(k)$  et  $j = 2, \dots, k$  tel que  $a_j = \int e_j d\nu \neq 0$  on va se donner un entier  $m > 1$ .

Soit  $N(a_j, m)$  le plus petit entier tel que  $\frac{w_n}{n} < \frac{a_j^2}{m^2}$  pour  $n \geq N(a_j, m)$ .

Cet entier existe puisque  $\frac{w_n}{n} \rightarrow 0$ .

$$P\left[|\hat{a}_{j,n} - a_j| \geq \frac{m-1}{m} |a_j|\right] \leq \frac{E[\hat{a}_{j,n} - a_j]^4}{a_j^4} \left(\frac{m}{m-1}\right)^4.$$

$$P\left[|\hat{a}_{j,n} - a_j| \geq \frac{m-1}{m} |a_j|\right] = 1 - P\left[|\hat{a}_{j,n} - a_j| < \frac{m-1}{m} |a_j|\right] = 1 - p_n(\nu)$$

$p_n(\nu)$  étant la puissance du test.

On déduit alors :

$$1 - p_n(\nu) \leq \frac{E[\hat{a}_{j,n} - a_j]^4}{a_j^4} \left(\frac{m}{m-1}\right)^4$$

$$\frac{E[\hat{a}_{j,n} - a_j]^4}{a_j^4} \leq 3072 \times 2^4 M^4 \sum_{i=1}^{\infty} \sqrt{\alpha(i)} \times \frac{1}{n^2} \quad (\text{cf [29] p.99}),$$

$$\text{donc } 1 - p_n(\nu) \leq \frac{3072 \times 2^4 \times M^4}{a_j^4} \left(\frac{m}{m-1}\right)^4 \sum_{i=1}^{\infty} \sqrt{\alpha(i)} \times \frac{1}{n^2}$$

donc  $1 - p_n(\nu) \xrightarrow{n \rightarrow +\infty} 0$ , d'où le résultat annoncé.

2) Pour  $n$  assez grand on a :

$$\frac{w_n}{n} < \frac{\delta(H_2)}{m^2}, m > 1.$$

En se servant de la démonstration précédente et en remplaçant  $a_j^2$  par  $\delta(H_2)$

alors :

$$\sup_{\nu \in H_2} [1 - p_n(\nu)] \leq \frac{3072 \times 2^4 \times M^4}{\delta^2(H_2)} \left(\frac{m}{m-1}\right)^4 \sum_{i=1}^{\infty} \sqrt{\alpha(i)} \times \frac{1}{n^2} \xrightarrow{n \rightarrow \infty} 0$$

donc :  $\sup_{\nu \in H_2} [1 - p_n(\nu)] \rightarrow 0$

d'où la convergence uniforme annoncée.

3) Soit  $j_0 \in \{2, \dots, k\}$  tel que  $|\int c_{j_0} d\nu_n|^2 = \max_{2 \leq j \leq k} |\int c_j d\nu_n|^2$

$$\text{pour } n > 1 \quad \frac{n}{4(k-1)} \sum_{j=2}^k \left[ \int c_j d\nu_n \right]^2 \geq w_n,$$

d'où

$$\frac{n}{4(k-1)} \sum_{j=2}^k \left[ \int c_j d\nu_n \right]^2 \geq \frac{w_n}{n} \implies \frac{w_n}{n} \leq \frac{\left[ \int c_{j_0} d\nu_n \right]^2}{4}$$

$$\sqrt{\frac{w_n}{n}} \leq \frac{\left| \int c_{j_0} d\nu_n \right|}{2} \quad (\text{ici on a pris } m=2)$$

En remplaçant alors  $a_j$  par  $\int c_{j_0} d\nu_n$  et  $m$  par 2 dans la démonstration de 1)

on a :

$$|1 - p_n(\nu_n)| \leq \frac{3072 \times 2^4 \times M^4 \sum_{i=2}^{\infty} \alpha(i)}{w_n^2(H_2)} \xrightarrow{n \rightarrow \infty} 0$$

Sous a), b), c),  $w_n \rightarrow +\infty$ , alors  $\alpha_n \rightarrow 0$  ■

### Remarques

1) Les conditions :  $w_n \rightarrow +\infty$  et  $\frac{w_n}{n} \rightarrow 0$  sont nécessaires pour assurer la convergence du test mais ne sont pas suffisantes.

$$2) |1 - p(\nu_n)| \leq \frac{3072 \times 2^4 \times \sum_{i \geq 1} \sqrt{\alpha(i)}}{w_n^2}$$

$$\text{donc } |1 - p(\nu_n)| = O\left[\frac{1}{w_n^2}\right]$$

Ce qui nous permet d'avoir la vitesse de convergence du test.

### II.5.2.2. Cas où le noyau est de dimension finie qui varie avec la taille de l'échantillon.

$$\text{Si } \alpha_n \text{ est le niveau du test } n \sum_{j=2}^k \hat{\alpha}_{j,n}^2 > w_n \text{ alors } \alpha_n = P\left[ n \sum_{j=2}^k \hat{\alpha}_{j,n}^2 > w_n \right].$$

**Proposition 14.** *Sous les conditions*

$$1) \ a) \ n \in \mathbb{N}^*, \sup_{x \in E} \sup_{j \in \{2, \dots, k_n\}} |c_j(x)| = M_1 < +\infty$$

$$b) \ \sum_{i=1}^{\infty} \sqrt{\alpha(i)} < +\infty$$

$$2) \ a) \ \left(\frac{k_n}{n}\right)^{1/24} b_n^{1/2} \rightarrow 0$$

$$b) \ \left(\frac{1}{k_n}\right)^{k_n} b_n^{k_n} \rightarrow 0$$

$$c) \ \alpha(n^{1/4}) (nb_n)^{k_n} \rightarrow 0$$

$$3) \ \begin{cases} w_n \leq b_n \\ w_n \xrightarrow{n \rightarrow +\infty} +\infty \\ \frac{w_n}{n} \rightarrow 0 \end{cases}$$

Alors le test  $n \sum_{j=2}^k \hat{\alpha}_{j_n}^2 > w_n$  est convergent.

**Démonstration.** Elle utilise le théorème 6 et l'existence de la suite  $(b_n)$  annoncée plus haut. ■

**Proposition 15.** *Sous l'hypothèse 1 de la proposition 14*

1) Soit  $H_3 \subset H_1(k_n)$  telle que, pour tout entier  $n$  assez grand

$$\inf_{\nu \in H_3} \sup_{j=2, \dots, k_n} \left| \int c_j d\nu \right| \geq m \sqrt{\frac{w_n}{n}}$$

où  $m$  est un entier donné supérieur à 1, alors la convergence de la puissance est uniforme sur  $H_3$  dès que  $w_n \rightarrow +\infty$ , plus précisément, on a :

$$\sup_{\nu \in H_3} [1 - p_n(\nu)] = o\left[\frac{1}{w_n^2}\right]$$

2) Soit  $(\nu_n) \subset H_1(k_n)$  et  $j_n \in \mathbb{N}$  tel que  $\left| \int c_{j_n} d\nu \right| \geq m \sqrt{\frac{w_n}{n}}$

alors  $p_n(\nu_n) \rightarrow 1$  dès que  $w_n \rightarrow +\infty$

3)  $\alpha_n$  étant le niveau du test  $\alpha_n = () \left[ \frac{k_n^2}{w_n} \right]$

**Démonstration.** Pour 1) et 2) la démonstration est analogue à celle de la Proposition 14.

$$3) \ \alpha_n = P \left[ n \sum_{j=2}^{k_n} \hat{\alpha}_{j_n}^2 \geq w_n \right] = P \left[ n \sum_{j=2}^{k_n} \hat{\alpha}_{j_n}^2 > \frac{w_n}{n} \right] \leq \sum_{j=2}^{k_n} P \left[ \hat{\alpha}_{j_n}^2 > \frac{w_n}{n(k_n - 1)} \right]$$

$$\begin{aligned}
&< \sum_{j=2}^{k_n} P \left[ \left| \hat{a}_{j,n} \right| > \sqrt{\frac{w_n}{n(k_n - 1)}} \right] \\
\text{or } P \left[ \left| \hat{a}_{j,n} \right| > \sqrt{\frac{w_n}{n(k_n - 1)}} \right] &\leq \frac{E \left[ \hat{a}_{j,n}^2 \right]}{w_n} nk_n \\
E \left[ \hat{a}_{j,n}^2 \right] &\leq \frac{1}{4^2} \left( 1 + 8M_1^2 \sum_{i=1}^{\infty} \alpha(i) \right) nk_n \\
\text{Par suite} \\
\alpha_n \leq P \left[ \left| \hat{a}_{j,n} \right| > \sqrt{\frac{w_n}{n(k_n - 1)}} \right] &\leq \left( 1 + 8M_1^2 \sum_{i \geq 1} \alpha(i) \right) \frac{k_n^2}{w_n} \\
\text{D'où le résultat.}
\end{aligned}$$

### 11.5.2.3. Cas du noyau de dimension infinie

$$\alpha_n \text{ étant le niveau du test } n \sum_{j=2}^{\infty} \hat{a}_{j,n}^2 > w_n \text{ alors } \alpha_n = P \left[ n \sum_{j=2}^{\infty} \hat{a}_{j,n}^2 > w_n \right]$$

**Proposition 16.** *Sous l'hypothèse :*

$$\begin{cases} \sup_{x \in E} \sup_{j \geq \ell} |e_j(x)| = M < +\infty \\ \sum \sqrt{\alpha(i)} < +\infty \end{cases}$$

*Si en plus :*

- 1)  $\left( \frac{k_n}{n} \right)^{1/2\ell} b_n^{1/4} \xrightarrow{n \rightarrow +\infty} 0$
- 2)  $\left( \frac{1}{k_n} \right)^{k_n} b_n^{k_n} \xrightarrow{n \rightarrow +\infty} 0$
- 3)  $(nb_n)^{k_n} \alpha(n^{1/4}) \xrightarrow{n \rightarrow +\infty} 0$
- 4)  $\sum_{j \geq \ell} \lambda_j^2 < +\infty$ ; et  $\sum_{j \geq \ell} \lambda_j^2 \sigma_j^2 < +\infty$
- 5)  $\begin{cases} w_n \leq b_n \\ w_n \xrightarrow{n \rightarrow +\infty} +\infty \\ \frac{w_n}{n} \xrightarrow{n \rightarrow +\infty} 0 \end{cases}$

alors le test  $n \sum_{j=2}^{\infty} \hat{a}_{j,n}^2 = \|S_n(\cdot)\|^2 > w_n$  est convergent

**Démonstration.** On utilise le Théorème 6,

Sous les conditions 1) 2) 3) 4) 5) on a alors :

$$\sup_{\alpha_n \leq \alpha_n} |\alpha_n - P[\|U\|^2 > w_n]| \xrightarrow{n \rightarrow +\infty} 0$$

par suite  $\alpha_n \rightarrow 0$

car

$w_n \rightarrow +\infty$

Soit  $\nu \in H_1(\lambda_n)$  et soit  $j_0 \geq 2$  tel que  $\int c_{j_0} d\nu = a_{j_0} \neq 0$ .

Comme  $\|S_n(\cdot)\|^2 \leq w_n$  alors  $n\lambda_{j_0}^2 \hat{a}_{j_0 n} \leq w_n$

d'où  $P[\|S_n(\cdot)\|^2 \leq w_n] \leq P\left[|\hat{a}_{j_0 n}| \leq \sqrt{\frac{w_n}{n}} |\lambda_{j_0}^{-1}|\right]$ .

Comme  $\frac{w_n}{n} \rightarrow 0$ , on aura pour  $n$  assez grand

$$\frac{w_n}{n} \leq \frac{\lambda_{j_0}^2 \lambda_{j_0}^2}{m^2} (m > 1),$$

on obtient alors

$$P[\|S_n(\cdot)\|^2 \leq w_n] \leq P\left[|\hat{a}_{j_0 n} - a_{j_0}| > \frac{m-1}{m} |a_{j_0}|\right]$$

$$P\left[|\hat{a}_{j_0 n} - a_{j_0}| \geq \frac{m-1}{m} |a_{j_0}|\right] \leq \frac{E[\hat{a}_{j_0 n} - a_{j_0}]^4}{a_{j_0}^4} \left(\frac{m}{m-1}\right)^4.$$

$$\text{Or } 1 - p_n(\nu) \leq P\left[|\hat{a}_{j_0 n} - a_{j_0}| \geq \frac{m-1}{m} |a_{j_0}|\right]$$

car  $P[\|S_n(\cdot)\|^2 \leq w_n] = 1 - p_n(\nu)$

On en déduit alors :

$$1 - p_n(\nu) \leq \frac{E[\hat{a}_{j_0 n} - a_{j_0}]^4}{a_{j_0}^4} \left(\frac{m}{m-1}\right)^4$$

$$\text{or } E[\hat{a}_{j_0 n} - a_{j_0}]^4 \leq 3072 \times 2^4 \times M^4 \sum_{i=1}^{\infty} \sqrt{\alpha(i)} \times \frac{1}{n^2} \text{ (cf [30] p.99).}$$

$$\text{donc } 1 - p_n(\nu) \leq \frac{3072 \times 2^4 \times M^4}{a_{j_0}^4} \left(\frac{m}{m-1}\right)^4 \times \sum_{i=1}^{\infty} \sqrt{\alpha(i)} \times \frac{1}{n^2}$$

par conséquent  $1 - p_n(\nu) \xrightarrow{n \rightarrow +\infty} 0$ , d'où le résultat annoncé. ■

### Remarques

1) La convergence de la puissance du test est uniforme sur tout  $H_2 \subset H_1(\lambda_n)$  tel que

$$\delta(H_2) = \inf_{\nu \in H_2} \max_{j \geq 2} \left| \int c_j d\nu \right|^2 > 0$$

$$2) \sup_{\nu \in H_2} (1 - p_n(\nu)) = O\left(\frac{1}{n^2}\right)$$

La vitesse de convergence de la puissance du test est  $O(\frac{1}{n^2})$  mais cette vitesse dépend des  $a_j$  qui sont donnés par  $a_j = \int e_j d\nu$ .

3) dans les conditions du Corollaire 5, on obtient

$$\alpha_n = O\left[\frac{1}{\log(2+n)}\right]^2.$$

4) Enfin, d'une manière générale il se pose le problème de l'optimalité asymptotique du test qui vient d'être construit.

Après avoir étudié la convergence du test construit le chapitre suivant sera consacré à l'application sur un autoregressif d'ordre un.

# Chapitre III

## Tests du $\chi^2$ Généralisé

### Application au processus autoregressif

#### III.1. Introduction

Une manière naturelle de construire un test d'adéquation consiste à utiliser un estimateur de la loi dont on teste l'ajustement. Le calcul de la distance de la loi présumée à son estimateur permet d'apprécier la validité de l'hypothèse de départ. Ainsi, le célèbre test du  $\chi^2$  est fondé sur un estimateur élémentaire de la densité : l'histogramme des fréquences. Or il est bien connu que les estimateurs lisses de la densité approchent la loi des observations beaucoup mieux que l'histogramme. Au chapitre précédent, sous certaines conditions

de régularité nous avons montré (II.3.1) que  $S_n \xrightarrow{L} Z = \begin{pmatrix} Z_1 \\ \vdots \\ Z_k \end{pmatrix}$  où  $Z$  est

une variable aléatoire normale dans  $\mathbb{R}^{k-1}$  de moyenne nulle et de matrice de covariance  $\Sigma$ . Toute une classe de tests a été construite, ainsi que la convergence de ces tests étudiée au chapitre précédent. Ici, il s'agira d'estimer les éléments de la matrice de variance covariance  $\Sigma$  et de faire des simulations sur un autoregressif d'ordre un afin de faire un test statistique de l'hypothèse : les observations suivent une loi fixée notée par  $\mu$ . Ainsi, à la section 2, les éléments de la matrice  $\Sigma$  sont estimés. L'approximation de la loi limite de la statistique  $\|S_n(\cdot)\|^2$  est étudiée à la section 3. À la section 4, les simulations sont indiquées.

## III.2. Estimation de $\Sigma = (\sigma_{jj'})_{2 \leq j, j' \leq k}$

### III.2.1 Hypothèses et notations

Nous supposons pour toute la suite que la matrice  $\Sigma$  est définie positive. La matrice  $\Sigma = (\sigma_{jj'})_{2 \leq j, j' \leq k}$  sera désormais notée par  $\Sigma = (\sigma_{jl})_{2 \leq j, l \leq k}$ . L'estimation de  $\Sigma$  revient à estimer les éléments  $(\sigma_{jl})_{2 \leq j, l \leq k}$ . Nous allons nous inspirer des résultats de Carbon[30] pp 25-30 et Ignaccolo[51] pp 5-15 pour

l'estimation des paramètres.

Posons  $\sigma_{jl}(v) = E(Y_{ij}Y_{i+v,l})$  où  $Y_{ij} = e_j(X_i)$ .

Par stationnarité du vecteur  $Y_i = (Y_{i1}, \dots, Y_{ik})$ , il vient que :

$\text{cov}(Y_1, Y_v) = \text{cov}(Y_t, Y_{t+v})$  et pour chaque élément on a :

$E(Y_{ij}Y_{i+v,l}) = E(Y_{ij}Y_{v+t,l})$  d'où  $\sigma_{jl}(v) = E(Y_{ij}Y_{v+t,l})$ .

De plus l'égalité  $\sigma_{jl}(-v) = E(Y_{ij}Y_{t+v,l}) = E(Y_{t-v,j}Y_{t,l}) = \sigma_{lj}(v)$ .

Posant  $v = i-1$ ,

on obtient  $\sigma_{jl}(i-1) = E(Y_{ij}Y_{i,l})$  et  $\sigma_{jl}[-(i-1)] = \sigma_{lj}(i-1)$ , aussi il est possible d'écrire  $\sigma_{jl} = \sum_{v=-2}^{\infty} \sigma_{jl}(v)$ .

Par la suite supposons  $\sum_{v=-2}^{\infty} |\sigma_{jl}(v)| < +\infty$  pour  $j, l = 2, k$  (ce qui est réalisé dès que  $\sum_{i>2} \alpha(i) < +\infty$  (cf Chapitre V).

Soit  $l_n$  une suite de nombres réels pour  $n \in \mathbb{N}^*$  telle que  $l_n < n$  et  $\lim_{n \rightarrow +\infty} l_n = \lim_{n \rightarrow +\infty} \frac{n}{l_n} = +\infty$ .

Considérons l'estimation classique de la covariance croisée du processus stationnaire bivarié  $(Y_{ij}Y_{it})_{i \in \mathbb{Z}}$  défini par :

$$\begin{cases} \hat{\sigma}_{jl}(v) = \frac{1}{n-v} \sum_{t=1}^{n-v} Y_{tj}Y_{t+v,l} \text{ pour } 0 \leq v \leq l_n \\ \hat{\sigma}_{jl}(v) = \hat{\sigma}_{lj}(-v) \text{ pour } -l_n \leq v \leq 0 \end{cases}$$

Cet estimateur est sans biais, si  $E[\hat{\sigma}_{jl}(v)] = \sigma_{jl}(v)$ .

Nous pouvons définir un estimateur de  $\sigma_{jl}(v)$  par :  $\hat{\sigma}_{jl} = \sum_{v=-l_n}^{l_n} w_n(v) \hat{\sigma}_{jl}(v)$  où la fonction poids  $w_n$  est donnée par :

$$w_n(v) = \left(1 - \frac{|v|}{n}\right) \eta\left(\frac{v}{l_n}\right) \mathbf{1}_{\{|v| \leq l_n\}}(v)$$



avec  $\eta$  satisfaisant les conditions suivantes :

$$\begin{cases} \eta(0) = 1 \\ |\eta(x)| \leq B \text{ pour } |x| \leq 1 \text{ (} B \geq 1 \text{)} \\ \eta(x) = \eta(-x) \end{cases}$$

de plus  $\eta$  est continue en 0.

### III.2.2 Estimation

Avec les hypothèses et notations ci-dessus, on a les résultats suivants (cf [5] pp 12-14) :

**Proposition 17 :**

L'estimateur  $\hat{\sigma}_{jl}$  est asymptotiquement sans biais.

**Proposition 18 :**

Pour  $j$  et  $l$  fixés, posons :

$$Z_{lv} = Y_{tj} Y_{v+t,l} - \sigma_{jl}(v) \text{ pour } 0 \leq v \leq l_n \text{ et}$$

$$Z_{lv} = Y_{t-v,j} Y_{t,l} - \sigma_{jl}(-v) \text{ pour } -l_n \leq v \leq 0.$$

Si pour tout  $t \in [1, n]$  et tout  $v \in [-l_n, l_n]$  il existe deux constantes  $C_1$  et  $C_2$  telles que :

$$|Z_{lv}| \leq C_1 \text{ et } |Z_{lv}^2| \leq C_2$$

alors pour tout entier  $q_n \in [1, \frac{n-l_n}{2}]$  et tout  $\varepsilon > 0$  on a :

$$\begin{aligned} P \left[ \sup_{|v| \leq l_n} |\hat{\sigma}_{jl}(v) - \sigma_{jl}(v)| > \varepsilon \right] &\leq 4(1 + 2l_n) \exp \left[ -\frac{\varepsilon^2}{8c^2} q_n \right] \\ &+ 22(1 + 2l_n) \left( \sqrt{1 + \frac{4c}{\varepsilon}} \right) q_n \alpha \left( \left[ \frac{n-l_n}{2q_n} \right] \right) \end{aligned} \quad (10)$$

pour tout  $v$ , où  $c = \max(C_1, C_2)$ .

De plus, si

$$\lim_{n \rightarrow +\infty} l_n \exp(-\alpha q_n) = \lim_{n \rightarrow +\infty} l_n q_n \alpha \left( \left[ \frac{n-l_n}{2q_n} \right] \right) = 0 \quad (11)$$

où  $\alpha$  est une constante dépendant de  $\varepsilon$  et  $c$ ,

alors  $\hat{\sigma}_{jl}(v) \xrightarrow[n \rightarrow +\infty]{P} \sigma_{jl}(v)$ .

Si, pour tout  $\varepsilon > 0$ ;

$$\sum_n \left\{ 4(1 + 2l_n) \exp \left[ -\frac{\varepsilon^2}{8c^2} l_n \right] + 22(1 + 2l_n) \left( \sqrt{1 + \frac{4c}{\varepsilon}} \right) l_n \alpha \left( \left[ \frac{n - l_n}{2l_n} \right] \right) \right\} < +\infty \quad (12)$$

alors  $\hat{\sigma}_{jl}(v) \xrightarrow[n \rightarrow +\infty]{p.s.} \sigma_{jl}(v)$ .

**Proposition 19 :**

Sous les hypothèses de la proposition 18, pour  $n$  suffisamment grand, pour tout entier  $l_n \in \left[ 1, \frac{n - l_n}{2} \right]$  et pour tout  $\varepsilon > 0$  on a :

$$P \left[ |\hat{\sigma}_{jl} - \sigma_{jl}| > \varepsilon \right] \leq 4(1 + 2l_n) \exp \left[ -\frac{\varepsilon^2}{32c^2 B^2 (1 + 2l_n)^2} l_n \right] + 22(1 + 2l_n) \left( \sqrt{1 + \frac{8cB(1 + 2l_n)}{\varepsilon}} \right) l_n \alpha \left( \left[ \frac{n - l_n}{2l_n} \right] \right) \quad (13)$$

où  $c = \max(C_1, C_2)$ .  
Si, de plus

$$\lim_{n \rightarrow +\infty} l_n \exp \left( -\Lambda \frac{l_n}{l_n^2} \right) = \lim_{n \rightarrow +\infty} (l_n)^{\frac{3}{2}} l_n \alpha \left( \left[ \frac{n - l_n}{2l_n} \right] \right) = 0 \quad (14)$$

où  $\Lambda$  est une constante dépendant de  $\varepsilon$ ,  $c$  et  $B$ ;  
alors  $\hat{\sigma}_{jl} \xrightarrow[n \rightarrow +\infty]{p.s.} \sigma_{jl}$ . Si,

$$\sum \left\{ 4(1 + 2l_n) \exp \left[ -\frac{\varepsilon^2}{32c^2 B^2 (1 + 2l_n)^2} l_n \right] + 22(1 + 2l_n) \left( \sqrt{1 + \frac{8cB(1 + 2l_n)}{\varepsilon}} \right) l_n \alpha \left( \left[ \frac{n - l_n}{2l_n} \right] \right) \right\} \leq +\infty \quad (15)$$

alors  $\hat{\sigma}_{jl} \xrightarrow[n \rightarrow +\infty]{p.s.} \sigma_{jl}$ .

Si, pour tout  $\varepsilon > 0$ ;

$$\sum_n \left\{ 4(1+2l_n) \exp \left[ -\frac{\varepsilon^2}{8c^2 l_n} \right] + 22(1+2l_n) \left( \sqrt{1 + \frac{4c}{\varepsilon}} \right) l_n \alpha \left( \left[ \frac{n-l_n}{2l_n} \right] \right) \right\} < +\infty \quad (12)$$

alors  $\hat{\sigma}_{ji}(v) \xrightarrow[n \rightarrow +\infty]{P.S.} \sigma_{ji}(v)$ .

**Proposition 19 :**

Sous les hypothèses de la proposition 18, pour  $n$  suffisamment grand, pour tout entier  $q_n \in \left[ 1, \frac{n-l_n}{2} \right]$  et pour tout  $\varepsilon > 0$  on a :

$$P \left[ |\hat{\sigma}_{ji} - \sigma_{ji}| > \varepsilon \right] \leq 4(1+2l_n) \exp \left[ -\frac{\varepsilon^2}{32c^2 B^2 (1+2l_n)^2} \right] + 22(1+2l_n) \left( \sqrt{1 + \frac{8cB(1+2l_n)}{\varepsilon}} \right) l_n \alpha \left( \left[ \frac{n-l_n}{2l_n} \right] \right) \quad (13)$$

où  $c = \max(C_1, C_2)$ .

Si, de plus

$$\lim_{n \rightarrow +\infty} l_n \exp \left( -A \frac{l_n}{l_n^2} \right) = \lim_{n \rightarrow +\infty} (l_n)^{\frac{3}{2}} l_n \alpha \left( \left[ \frac{n-l_n}{2l_n} \right] \right) = 0 \quad (14)$$

où  $A$  est une constante dépendant de  $\varepsilon$ ,  $c$  et  $B$ ;

alors  $\hat{\sigma}_{ji} \xrightarrow[n \rightarrow +\infty]{P.} \sigma_{ji}$ . Si,

$$\sum_n \left\{ 4(1+2l_n) \exp \left[ -\frac{\varepsilon^2}{32c^2 B^2 (1+2l_n)^2} \right] + 22(1+2l_n) \left( \sqrt{1 + \frac{8cB(1+2l_n)}{\varepsilon}} \right) l_n \alpha \left( \left[ \frac{n-l_n}{2l_n} \right] \right) \right\} \leq +\infty \quad (15)$$

alors  $\hat{\sigma}_{ji} \xrightarrow[n \rightarrow +\infty]{P.S.} \sigma_{ji}$ .

Démonstration : Voir [51] pp 12-14.

**Corollaire 6 :** Si  $\begin{cases} \alpha(n) \leq \rho^n \text{ où } 0 < \rho < 1 \\ l_n n^\gamma \text{ où } 0 < \gamma < \frac{1}{2} \\ q_n n^\beta \text{ où } 2\gamma < \beta < 1 \end{cases}$

alors les conditions (14) sont vérifiées et la convergence en probabilité de  $\hat{\sigma}_{ji}$  vers  $\sigma_{ji}$  en découle.

**Corollaire 7 :** Si  $\begin{cases} \alpha(n) \leq n^{-\delta} \text{ avec } \delta > 0 \\ l_n n^\gamma \text{ où } 0 < \gamma < \frac{2\delta}{4\delta+7} \\ q_n n^\beta \text{ où } 2\gamma < \beta < \frac{\delta+3\gamma}{1+\delta} \end{cases}$

alors les conditions (14) sont vérifiées et la convergence en probabilité s'en suit.

**REMARQUE :** Nous obtenons donc un estimateur consistant de  $\sigma_{ji}$  pour tout  $j$  et  $l$ . On pose  $\hat{\Sigma}_n = (\hat{\sigma}_{jl})_{2 \leq j, l \leq k}$ .

### III.3. Approximation de la loi limite de $\|S(\cdot)\|^2$

#### III.3.1. Loi limite de $\|S_n(\cdot)\|^2$ en fonction des valeurs propres de $\Sigma$

La matrice  $\Sigma$  étant définie positive alors ses valeurs propres que nous noterons par  $\lambda_j$  sont telles que : pour tout  $j \in \{2, \dots, k\}$ ,  $\lambda_j > 0$ . Nous obtenons le résultat suivant :

**Proposition 20 :**

Sous les conditions de la proposition 11,  $\Sigma$  étant définie positive : Alors

$$\|S_n(\cdot)\|^2 \underset{n \rightarrow \infty}{\overset{L.P.}{\rightarrow}} \sum_{j=2}^k \lambda_j U_j^2$$

où les v.a.  $V_j \sim N(0,1)$  et sont indépendantes.

**Démonstration.** Elle est analogue à celle donnée dans [45] p.11-13. Les valeurs propres  $\lambda_j$  étant inconnues pour la réalisation de la simulation nous les estimons à partir de la matrice  $\hat{\Sigma}_n$ .

### III.3.2. Estimation des valeurs propres $\lambda_j ; j=2, \dots, k$

Nous utiliserons les résultats de [22] p.102-104 pour l'estimation des  $\lambda_j, j=2, \dots, k$ . On y établit que si  $C$  est un opérateur de covariance et  $C_n$  son estimateur empirique;  $\lambda_j, j \geq 1$  les valeurs propres de  $C$  c'est-à-dire  $C(v_j) = \lambda_j v_j$ , où  $\{v_j, j \geq 1\}$  est un système orthonormal constitué par les vecteurs propres de  $C$ ; alors  $\lambda_{jn}, j \geq 1$  qui sont des estimateurs empiriques de  $\lambda_j, j \geq 1$  sont donnés par :  $C(v_{jn}) = \lambda_{jn} v_{jn}, j \geq 1$ . Les  $\lambda_{jn}, j \geq 1$  sont donc les valeurs propres de  $C_n$  associées aux vecteurs propres  $v_{jn}, j \geq 1$ , qui constituent un système orthonormal. On montre en particulier que si :  $\lambda_{1n} > \lambda_{2n} > \dots > \lambda_{kn} \geq 0 = \lambda_{n+1,n} = \lambda_{n+2,n} = \dots$  alors, pour les opérateurs compacts, on a :  $\sup_{j \geq 1} |\lambda_{jn} - \lambda_j| < \|C_n - C\|$

Nous appliquons ces résultats à notre situation. En effet nous travaillons dans  $\mathbb{R}^{k+1}$  et les matrices  $\widehat{\Sigma}_n$  et  $\Sigma$  sont bornées (pour tous  $j, j' = 2, \dots, k ; \sigma_j^2 < 1 \infty$  et  $|\sigma_{jt}| < 1 \infty ; \widehat{\sigma}_{jt}$  et  $\widehat{\sigma}_j^2$  sont des estimateurs convergents de  $\sigma_{jt}$  et  $\sigma_j^2$ ), on peut donc utiliser les résultats de [22].

Désignons par  $\lambda_j ; j=2, \dots, k$  les valeurs propres de la matrice  $\Sigma$ , comme  $\Sigma$  est supposée être définie positive alors ses valeurs propres sont toutes positives (pour tous  $j, j' = 2, \dots, k ; \lambda_j > 0$ ). Soient  $\lambda_{jn} ; j=2, \dots, k$  les valeurs propres de  $\widehat{\Sigma}_n$  ( $\widehat{\Sigma}_n$  étant symétrique ses valeurs propres sont réelles). Par ailleurs, on a montré que  $\widehat{\Sigma}_n \xrightarrow[n \rightarrow \infty]{p.s.} \Sigma$  donc, par continuité  $\lambda_{jn} \xrightarrow[n \rightarrow \infty]{p.s.} \lambda_j ; j=2, \dots, k$  et  $\lambda_{jn}$  est positive à partir d'un certain rang aléatoire (p.s).

### III.3.3. Approximation de la loi limite de $\|S_n(\cdot)\|^2$

$\widehat{\Sigma}_n$  étant symétrique, elle est diagonalisable et est semblable à la matrice  $\widehat{D}_n$  définie par :

$$\widehat{D}_n = \begin{pmatrix} \lambda_{2n} & 0 & \dots & 0 \\ 0 & \lambda_{3n} & \dots & \vdots \\ \vdots & \dots & \dots & 0 \\ 0 & \dots & 0 & \lambda_{kn} \end{pmatrix} \xrightarrow[n \rightarrow \infty]{p.s.} D = \begin{pmatrix} \lambda_2 & 0 & \dots & 0 \\ 0 & \lambda_3 & \dots & \vdots \\ \vdots & \dots & \dots & 0 \\ 0 & \dots & 0 & \lambda_k \end{pmatrix}$$

Remarquons que  $\Sigma$  est semblable à  $D$ .

Prenons maintenant : pour tout  $j = 2, \dots, k$ ; pour tout  $n > 1 ; \lambda_{jn}^* = \max(0, \lambda_{jn})$

On a  $\lambda_{jn}^* \xrightarrow[n \rightarrow \infty]{p.n.} \lambda_j$  et pour tout  $n \geq 1$   $\lambda_{jn}^* > 0$  donc :

$$\hat{D}_n^* = \begin{pmatrix} \lambda_{2n}^* & 0 & \dots & 0 \\ 0 & \lambda_{3n}^* & \dots & \vdots \\ \vdots & \dots & \dots & 0 \\ 0 & \dots & 0 & \lambda_{kn}^* \end{pmatrix} \xrightarrow[n \rightarrow \infty]{p.n.} D = \begin{pmatrix} \lambda_2 & 0 & \dots & 0 \\ 0 & \lambda_3 & \dots & \vdots \\ \vdots & \dots & \dots & 0 \\ 0 & \dots & 0 & \lambda_k \end{pmatrix}$$

Nous utiliserons les estimateurs  $\lambda_{jn}^*$  pour approximer la loi limite de  $\|S_n(\cdot)\|^2$ .

**Proposition 21.**

Sous les conditions des propositions 11 et 19 et si  $\Sigma$  est définie positive, alors  $\sum_{j=2}^k \lambda_{jn}^* U_j^2 \xrightarrow[n \rightarrow \infty]{p.n.} \sum_{j=2}^k \lambda_j U_j^2$  où les  $U_j$ ,  $j=2, \dots, k$  suivent des lois normales centrées réduites et indépendantes.

*Démonstration.* Puisque  $\lambda_{jn}^* \xrightarrow[n \rightarrow \infty]{p.n.} \lambda_j$  pour tout  $j=2, \dots, k$  alors  $\sum_{j=2}^k \lambda_{jn}^* U_j^2 \xrightarrow[n \rightarrow \infty]{p.n.}$

$$\sum_{j=2}^k \lambda_j U_j^2. \blacksquare$$

Remarque :

De la proposition ci-dessus on déduit donc que pour  $n$  assez grand la variable aléatoire  $\hat{U}_j^2 = \sum_{j=2}^k \lambda_{jn}^* U_j^2$  approxime presque sûrement la variable aléatoire  $\sum_{j=2}^k \lambda_j U_j^2$  qui est la loi limite de  $\|S_n(\cdot)\|^2$ .

$\hat{U}_j^2$  sera utilisée dans les simulations pour tester l'hypothèse nulle  $H_0$ .

### III.4. Simulation

#### III.4.1. Introduction

On va réaliser une simulation sur un autorégressif d'ordre 1. On pose donc  $X_t = \rho X_{t-1} + \varepsilon_t$   $|\rho| < 1$ ,  $(\varepsilon_t)_{t \geq 1}$  une suite de variables aléatoires indépendantes et équadistribuées de loi normale  $N(0,1)$ . On construit un échantillon de taille  $n$  du processus en appelant pour chaque  $t$  une variable aléatoire  $\varepsilon_t$  générée par un sous-programme et en utilisant l'équation ci-dessus pour  $X_0$  donné. Nous

avons choisi des valeurs de  $\rho$  pour générer le processus  $(X_t)_{t \geq 1}$ . La matrice  $\widehat{\Sigma}_n$  étant fonction du système orthogonal considéré, nous avons choisi pour les calculs le système orthonormal constitué par les polynômes de Legendre. A partir de  $\widehat{\Sigma}_n$ , on obtient les valeurs propres  $\lambda_{jn}$  et par suite  $\lambda_{jn}^*$ ,  $j=2, \dots, k$ , ce qui permettra d'avoir  $\widehat{U}_n^2$  en fonction de  $k$ . Pour effectuer le test de Legendre on commence par amener les observations sur  $[-1,1]$  en posant :  $Y_i = 2F(X_i) - 1$ ,  $i=1, 2, \dots, n$ . Où  $F$  désigne la fonction de répartition de la loi des  $X_i$ . Sous  $H_0$  les  $Y_i$  suivent la loi uniforme  $U$  sur  $[-1,1]$  (cf. chapitre 1, paragraphe 4).

### III.4.2. Test d'hypothèse

Notre test d'adéquation sera : tester  $H_0$  : " $\rho = 0$ " c'est-à-dire que le processus  $(X_t)$  est gaussien  $N(0,1)$  et les variables aléatoires sont indépendantes contre  $H_1$  : " $\rho \neq 0$ ", le processus  $(X_t)$  est constitué de variables aléatoires dépendantes. On va considérer des cas en choisissant des valeurs de  $\rho$  comme suit :

$$\begin{array}{lll} \rho = 0, 20; & \rho = 0, 50; & \rho = 0, 70 \\ \rho = -0, 20; & \rho = -0, 50; & \rho = -0, 70 \end{array}$$

On se donne  $\alpha$  le risque de première espèce. Cette donnée permet de déterminer la région de rejet de l'hypothèse nulle  $H_0$  : " $\rho = 0$ ". Cette région critique est donnée par, sous  $H_0$  :  $\{(x_1, \dots, x_n) : \|S_n(\cdot)\|^2 > w_n\}$  où  $w_n$  est déterminé par  $\alpha = P \left[ \sum_{j=2}^k \lambda_{jn}^* U_j^2 \geq w_n \right]$ , puisque pour  $n$  assez grand,  $\sum_{j=2}^k \lambda_{jn}^* U_j^2$

approxime presque sûrement la variable aléatoire  $\sum_{j=2}^k \lambda_j U_j^2$  qui est la loi limite de  $\|S_n(\cdot)\|^2$ . Si  $\beta$  est la puissance du test, on a :  $\beta = P_\nu \left[ \|S_n(\cdot)\|^2 \geq w_n \right]$  où  $\nu \in H_1$  : " $\rho \neq 0$ ".

Nous présentons ici, les puissances empiriques du test obtenues comme pourcentage de rejet sur  $m=1000$  simulations. Les tableaux ci-dessous donnent les puissances empiriques pour 1000 simulations, en faisant varier  $n$  et  $\rho$  et pour  $\alpha = 5\%$ .

Les colonnes représentent :

-  $k$  la dimension du noyau  $K$

-  $\chi^2$  : test de Pearson pour es observations fortement mélangants.

-  $\chi^2_{ind}$  : test de Pearson pour des observations indépendantes.

-leg : test avec les polynômes de Legendre pour des observations fortement mélangées.

legind : test avec les polynômes de Legendre pour des observations indépendantes.

Les tableaux 3.1 à 3.8 sont présentés en appendice.

Dans tous les tableaux, les valeurs en gras indiquent la puissance la plus grande lorsque  $k$  change, pour avoir une modification sur le choix de  $k$ .

On constate de manière générale, pour  $k$  fixé le test de Legendre a une puissance empirique plus grande que le test de Pearson.

Pour des valeurs positives de  $\rho$ , on observe que la puissance maximale du test de Legendre est obtenue pour des  $k$  plus petites que celle du test de Pearson.

### **III.4.3. Appendice : Tableaux des puissances empiriques**



**TABLEAU 3.1** : Puissances empiriques pour  $\rho = 0,2$  et  $n = 100$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 10.1     | 9.5            | 9.7        | 11.8          |
| 3        | 8.9      | 9.5            | 8.3        | 10.6          |
| 4        | 8.1      | 8.2            | 7.2        | 9.9           |
| 5        | 8.4      | 8.5            | 6.5        | 10.1          |
| 6        | 7.8      | 6.9            | 6.5        | 10.3          |
| 7        | 7.4      | 8.2            | 6.2        | 10.4          |
| 8        | 6.2      | 6.7            | 6.2        | 9.6           |
| 9        | 6.2      | 7.3            | 5.9        | 10.4          |
| 10       | 6        | 7.6            | 5.3        | 10.2          |

**TABLEAU 3.2** : Puissances empiriques pour  $\rho = -0,2$  et  $n = 100$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 9.0      | 3.0            | 9.9        | 3.3           |
| 3        | 8.9      | 4.1            | 8.0        | 4.2           |
| 4        | 8.8      | 3.8            | 7.1        | 3.5           |
| 5        | 7.4      | 4.2            | 7.3        | 4.3           |
| 6        | 7.7      | 4.7            | 7.3        | 5.9           |
| 7        | 6.7      | 3.9            | 6.8        | 5.7           |
| 8        | 5.6      | 3.7            | 6.8        | 6.1           |
| 9        | 6.0      | 4.0            | 5.4        | 5.5           |
| 10       | 6.1      | 4.4            | 4.2        | 5.2           |

**TABLEAU 3.3** : Puissances empiriques pour  $\rho = 0,2$  et  $n = 50$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 12.1     | 7.2            | 13.8       | 9.9           |
| 3        | 12.3     | 8.4            | 14.2       | 9.9           |
| 4        | 12.5     | 7.7            | 12.2       | 9.8           |
| 5        | 12.5     | 8              | 11.8       | 9.7           |
| 6        | 10.9     | 7.4            | 10.7       | 9             |
| 7        | 10.2     | 6.4            | 9.9        | 8.6           |
| 8        | 10.2     | 8.2            | 9.1        | 8.6           |
| 9        | 8        | 6.9            | 8.1        | 8.4           |
| 10       | 7.9      | 6.7            | 7.5        | 8.1           |

**TABLEAU 3.4** : Puissances empiriques pour  $\rho = 0,2$  et  $n = 250$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 7.4      | 8.9            | 7.8        | 11.6          |
| 3        | 7.3      | 8              | 6.7        | 10.1          |
| 4        | 6.2      | 8.2            | 6.3        | 9.8           |
| 5        | 6.2      | 7.3            | 6.3        | 9.6           |
| 6        | 6.9      | 8.1            | 5.6        | 9             |
| 7        | 5.2      | 6.9            | 5.7        | 8.6           |
| 8        | 6.1      | 7.4            | 4.9        | 9.7           |
| 9        | 5.2      | 7.4            | 4.9        | 9.3           |
| 10       | 5.3      | 7.7            | 4.8        | 8.9           |

**TABLEAU 3.5** : Puissances empiriques pour  $\rho = 0,5$  et  $n = 100$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 10.8     | 29             | 21.3       | 51.2          |
| 3        | 13.4     | 29.7           | 17.5       | 49.1          |
| 4        | 15.6     | 30.5           | 16.5       | 49.5          |
| 5        | 14.6     | 31.8           | 16         | 48.2          |
| 6        | 13.5     | 30.6           | 14.9       | 48.8          |
| 7        | 14.4     | 33.2           | 14.6       | 48.1          |
| 8        | 13.1     | 31.2           | 14         | 47.3          |
| 9        | 13.3     | 31.4           | 12.3       | 47.4          |
| 10       | 13.7     | 32             | 11.1       | 47.7          |

**TABLEAU 3.6** : Puissances empiriques pour  $\rho = -0,5$  et  $n = 100$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 18.00    | 7.7            | 34.1       | 33.6          |
| 3        | 21.0     | 11.2           | 30.1       | 30.8          |
| 4        | 20.3     | 11.7           | 29.9       | 35.3          |
| 5        | 19.7     | 12.7           | 27.1       | 33.8          |
| 6        | 20.4     | 12.9           | 24.9       | 35.0          |
| 7        | 19.2     | 15.0           | 22.7       | 34.3          |
| 8        | 17.8     | 14.6           | 21.7       | 35.0          |
| 9        | 18.1     | 16.5           | 19.3       | 34.2          |
| 10       | 18.2     | 17.8           | 18.3       | 35.6          |

**TABLEAU 3.7** : Puissances empiriques pour  $\rho = 0,7$  et  $n = 100$ 

| <b>k</b> | $\chi^2$ | $\chi^2_{ind}$ | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 20.9     | 62.1           | 47.9       | 91.2          |
| 3        | 25.5     | 70.7           | 33.8       | 91.2          |
| 4        | 29.9     | 74.7           | 43.6       | 92.5          |
| 5        | 29.5     | 76.8           | 35.1       | 92.3          |
| 6        | 31.9     | 76.9           | 39.8       | 92.7          |
| 7        | 32.9     | 78.8           | 34.3       | 92.1          |
| 8        | 32.7     | 79.6           | 35.5       | 91.9          |
| 9        | 33.2     | 80.7           | 30.7       | 91.8          |
| 10       | 33.3     | 79.3           | 31.9       | 91.9          |

**TABLEAU 3.8** : Puissances empiriques pour  $\rho = -0,7$  et  $n = 100$ 

| <b>k</b> | $\chi^2$ | <b>chi2ind</b> | <b>leg</b> | <b>legind</b> |
|----------|----------|----------------|------------|---------------|
| 2        | 47.8     | 32.1           | 72.6       | 82.7          |
| 3        | 54.4     | 47.6           | 69.8       | 80.1          |
| 4        | 59.3     | 54.7           | 71.2       | 84.8          |
| 5        | 60.9     | 57.4           | 69.6       | 83.5          |
| 6        | 58.4     | 59.9           | 68.0       | 85.4          |
| 7        | 60.8     | 62.0           | 66.2       | 83.9          |
| 8        | 60.4     | 63.0           | 63.1       | 84.3          |
| 9        | 59.6     | 63.2           | 60.9       | 83.7          |
| 10       | 59.1     | 63.7           | 57.8       | 84.1          |

# CHAPITRE IV

## Tests Banachiques : le cas de $C[0,1]$

### IV.1. Introduction

Soit  $(X_n, n \geq 1)$  une suite de variables aléatoires indépendantes et de même loi  $P_X$ , définies sur l'espace probabilisé  $(\Omega, \mathcal{A}, P)$  et à valeurs dans  $[0,1]$  muni de sa tribu borélienne  $\mathcal{B}$ .

Soit  $\mu$  une probabilité sur  $([0,1], \mathcal{B})$ . On se propose de tester l'hypothèse " $P_X = \mu$ ".

Bosq a étudié une classe de tests basés sur des normes hilbertiennes. Nous étudierons ici un test banachique associé à la norme uniforme de  $C[0,1]$ .

Dans la section 2 de ce chapitre nous donnons les notations. La loi limite de la statistique  $S_n(\cdot)$  est étudiée à la section 3. À la section 4 nous étudions le niveau du test.

### IV.2. Notations et définitions

Considérons  $e_1, e_2, \dots, e_k; k \geq 2$  une base orthonormale de  $E_k$ , sous espace de dimension finie de  $L^2(\mu)$ . On suppose que les  $e_j(\cdot)$  sont continues sur  $[0,1]$  et que  $e_1 \approx \mathbb{1}_{[0,1]}$ . Le noyau reproduisant de l'espace  $E_k$  engendré par  $e_1, e_2, \dots, e_k$  noté  $K(x, t)$  est donné par (cf. Chapitre I)

$$K(x, t) = \sum_{j=1}^k e_j(x)e_j(t) \text{ où } (x, t) \in [0,1] \times [0,1].$$

Définissons la statistique  $S_n(t)$  par :

$$S_n(t) = \sqrt{n} \sum_{j=2}^k \hat{a}_{j,n} c_j(t) \text{ où } \hat{a}_{j,n} = \frac{1}{n} \sum_{i=1}^n c_j(X_i) \text{ et } t \in [0, 1].$$

Nous allons chercher la loi limite de  $S_n(\cdot)$ .

### IV.3. Loi limite de $S_n(\cdot)$

Le résultat principal va nous être fourni par le Théorème de Jain et Marcus [54] p.220, que nous appliquerons à nos hypothèses

**Théorème 7** Théorème de Jain et Marcus

Soit  $(Y_n(\cdot))_{n \in \mathbb{N}}$  une suite de variables aléatoires indépendantes et de même loi définies sur  $(\Omega, \mathcal{A}, P)$  à valeur dans  $C(T)$  où  $T$  est un espace compact muni de sa distance  $d$ .

On fait les hypothèses suivantes :

1)  $E[Y_1(t)] = 0, E[Y_1^2(t)] < +\infty, t \in T$

2) Il existe une variable aléatoire réelle  $M$  telle que  $E[M^2] < +\infty$  et  $|Y_1(\mathbf{w}, s) - Y_1(\mathbf{w}, t)| \leq M(\mathbf{w})d(s, t), \mathbf{w} \in \Omega; s, t \in T$

alors si  $(T, d)$  vérifie la condition

$$[3]) \limsup_{h \rightarrow 0} \sup_{t \in T} \int_0^h \left( \log \frac{1}{m(B(t, \epsilon))} \right)^{1/2} d\epsilon = 0$$

où  $m$  est une probabilité sur  $T$  et  $B(t, \epsilon) = \{s : d(s, t) < \epsilon\}$

on a :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \xrightarrow{\mathcal{L}} Z$$

où  $Z$  est une variable aléatoire gaussienne centrée à valeurs dans  $C(T)$ .

**Démonstration.** Voir [54]p.220.221. Dans les chapitres précédents, nous avons vu que

$$\begin{aligned} S(\cdot) &= \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_{i,\cdot}) - 1 \right] \\ &= \frac{1}{\sqrt{n}} \left[ \sum_{i=1}^n K(X_{i,\cdot}) - n \right] \end{aligned}$$

nous allons alors montrer que le théorème ci-dessus s'applique aux variables aléatoires

$(K(X_i, \cdot) - 1, i \geq 1)$  dès que les  $e_j$  sont lipschitziennes.

En effet, on a :

$$\begin{aligned} |K(X_1(\omega), s) - 1 - K(X_1(\omega), t) + 1| &= |K(X_1(\omega), s) - K(X_1(\omega), t)| \\ &< M(\omega) |s - t| \quad (T = [0, 1]) \end{aligned}$$

car si les  $e_j$  sont lipschitziennes, on obtient

$$|K(X_1(\omega), s) - K(X_1(\omega), t)| = \left| \sum_{j=1}^k e_j(X_1)(e_j(s) - e_j(t)) \right|$$

or  $e_j \approx \mathbf{1}_{0,1}$ , donc

$$\begin{aligned} |K(X_1(\omega), s) - K(X_1(\omega), t)| &= \left| \sum_{j=2}^k e_j(X_1)(e_j(s) - e_j(t)) \right| \\ &< \sum_{j=2}^k M_j d_j |s - t| \end{aligned}$$

$M_j = \sup |e_j| < \infty$  car les  $e_j$  sont lipschitziennes donc continues et  $|e_j(s) - e_j(t)| \leq d_j |s - t|$

Définissons la constante  $M$  par :  $M = \sum_{j=2}^k M_j d_j$ .

1) et 2) sont par conséquent vérifiés.

Pour vérifier la condition 3), on choisit  $m$  égale à la mesure de Lebesgue

sur  $[0, 1]$  et on considère  $\sup_{t \in [0, 1]} \int_0^h \left( \log \frac{1}{m|B(t, \varepsilon)|} \right)^{1/2} d\varepsilon$

on a alors :

$$\begin{cases} m|B(t, \varepsilon)| = \varepsilon & \text{si } \varepsilon \leq t \leq 1 - \varepsilon \\ m|B(t, \varepsilon)| > \varepsilon/2 & \text{sinon} \end{cases}$$

d'où

$$\sup_{t \in [0, 1]} \int_0^h \left( \log \frac{1}{m|B(t, \varepsilon)|} \right)^{1/2} d\varepsilon \leq \int_0^h \left( \log \frac{2}{\varepsilon} \right)^{1/2} d\varepsilon$$

pour  $h$  assez petit,  $0 < \varepsilon < h$   $\log \frac{2}{\varepsilon} > 1$ .

on a :

$$\int_0^h \left( \log \frac{2}{\varepsilon} \right)^{1/2} d\varepsilon \leq \int_0^h \left( \log \frac{2}{\varepsilon} \right) d\varepsilon$$

et  $\int_0^h \left( \log \frac{2}{\varepsilon} \right) d\varepsilon = h \log 2 - h \log h + h$

et par suite

$$h \log 2 - h \log h + h \xrightarrow[h \rightarrow 0]{} 0$$

d'où la vérification de la condition 3).

Le théorème 7 s'applique donc à notre situation, d'où le

### Théorème 8

Sous les conditions suivantes : 1) les variables aléatoires  $(X_i)_{i \geq 1}$  sont de même loi, indépendantes et centrées.

2) Les  $e_j$  sont lipschitziennes.

$$\text{alors : } S_n(\cdot) = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n K(X_i, \cdot) - 1 \right] \xrightarrow{\mathcal{L}} Z(\cdot)$$

où  $Z(\cdot)$  est une variable aléatoire à valeurs dans  $C[0,1]$  définie par le processus gaussien centré de même opérateur de covariance que  $K(X_i, \cdot) - 1$ .

### Remarque

L'opérateur de covariance de la variable  $Z$  est définie par :

$$E[(K(X_1, s) - 1)(K(X_1, t) - 1)]$$

Ce qui donne après un simple calcul  $K(s, t) - 1$  : pour

$$(s, t) \in [0, 1] \times [0, 1].$$

$$\text{Par ailleurs } S_n(\cdot) = \left[ \sum_{j=2}^k \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n e_j(X_i) \right] e_j(\cdot) \right]$$

et  $\frac{1}{\sqrt{n}} \sum_{i=1}^n e_j(X_i) \xrightarrow{\mathcal{L}} Y_j, j = 2, \dots, k$  (Théorème central limite).

Les  $Y_j$  sont normales, centrées, réduites et indépendantes. Par conséquent :

$$S_n(\cdot) \xrightarrow{\mathcal{L}} \sum_{j=2}^k Y_j e_j(\cdot) \quad (\text{on utilise la continuité du produit scalaire}).$$

A cause de l'unicité de la limite on a :

$$Z(s) = \sum_{j=2}^k Y_j e_j(s) \quad \text{où } s \in [0, 1]$$

En considérant la norme uniforme, on a :

$$\|Z\| = \sup_{s \in [0,1]} |Z(s)| = \sup_{s \in [0,1]} \left| \sum_{j=2}^k Y_j e_j(s) \right|$$

## IV.4. Etude du niveau du test

### IV.4.1. Introduction

Soit  $\alpha \in ]0, 1[$  tel que  $P[\|Z\| \geq C_\alpha] = \alpha$

On peut alors construire un test de région critique  $\{(x_1, \dots, x_n) : \|S_n\|^2 > C_\alpha\}$  de niveau asymptotique  $\alpha$ .

La recherche de  $\alpha$  conduit à la détermination de la loi de  $\|Z\| = \sup_{s \in [0,1]} |Z(s)|$ .

Cette loi est très difficile à expliciter, nos recherches bibliographiques ne nous ont pas permis de résoudre la question. Certains auteurs comme M.A. Lifshits dans [60] proposent des minoration de  $\alpha$ , nous allons donc adapter ses résultats à notre processus.

Mais examinons d'abord un exemple, en considérant que les  $e_j$  sont uniformément bornés.

### IV.4.2. Majoration de $\alpha$ avec $\sup_{s \in \mathbb{E}} \sup_{j=1, \dots, k} |e_j(x)| =: D < +\infty$

$$\|Z\| = \sup \left| \sum_{j=2}^k Y_j e_j(s) \right| \leq \sum_{j=2}^k |Y_j| \sup |e_j(x)| \leq D \sum_{j=2}^k |Y_j|.$$

Posons alors  $X = D \sum_{j=2}^k |Y_j|$ , par suite  $\|Z\| \leq X$  et donc

$$\alpha = P[\|Z\| \geq C_\alpha] \leq P[X \geq C_\alpha].$$

Un calcul technique qui utilise les fonctions caractéristiques donne, en désignant par  $f_X$  la fonction densité de  $X$ ,



$$f_X(x) = \frac{2}{D\sqrt{2\pi(k-1)\left(1 - \frac{4}{2\pi}\right)}} \exp\left[-\frac{1}{2}\left(\frac{x - \frac{2D(k-1)}{\sqrt{2\pi}}}{D\sqrt{(k-1)\left(1 - \frac{4}{2\pi}\right)}}\right)^2\right] \mathbf{1}_{\mathbb{R}^+}(x)$$

On déduit alors :

$$P[X \geq C_n] = \int_{C_n}^{\infty} f_X(x) dx = 1 - \int_{-\infty}^{C_n} f_X(x) dx.$$

Si  $F_X$  désigne la fonction de répartition de  $X$ , on obtient  $P[X > C_n] = 1 - F_X(C_n)$  et comme  $\alpha \leq P[X \geq C_n]$ , alors

$$\alpha < 1 - F_X(C_n)$$

#### IV.4.3. Cas où le système orthonormal est la base trigonométrique

Nous supposons que la base orthonormale  $e_1, e_2, \dots, e_k$  de l'espace  $E_k \subset L_2(\mu)$  est constituée de fonctions trigonométriques

$$V(Z(s)) = E\left[\sum_{j=2}^k Y_j e_j(s)\right]^2 = E\left[\sum_{j=2}^k Y_j^2 e_j^2(s) + \sum_{j \neq j'} Y_j Y_{j'} e_j(s) e_{j'}(s)\right]$$

or  $E(Y_j Y_{j'}) = 0$  si  $j \neq j'$  (indépendance des  $Y_j$  qui sont centrés), donc

$$V(Z(s)) = \sum_{j=2}^k E[Y_j^2 e_j^2(s)] = k-1 \text{ D'où } V(Z(s)) = \text{constante.}$$

On utilise alors les résultats de M.A. Lifshits dans [59] en les adaptant à notre processus pour chercher le comportement de :

$$P\{\|Z\| \geq C_n\} = P\left[\sup_{s \in [0,1]} |Z(s)| \geq C_n\right]$$

##### IV.4.3.1. hypothèses et notations

Nous reprenons ici les données et hypothèses générales de [60] p.p. 177-178.

Soit  $(T, \rho)$  un espace métrique.

##### Définition 5 :

Un sous-ensemble  $A$  de  $T$  est appelé  $\varepsilon$ -distinguable si  $\rho(s,t) \geq \varepsilon$  pour tous points  $s$  et  $t$  distincts de  $A$ .

**Notations :**

· On note par  $M(t, \epsilon)$  le plus grand nombre possible d'éléments dans un sous-ensemble de  $T$   $\epsilon$ -distinguable.

Exemple :

si  $T_\epsilon = \{t_1, t_2, \dots, t_n\}$  est un sous-ensemble maximal de  $T$   $\epsilon$ -distinguable, alors  $M(T, \epsilon) = n$  et pour tout  $i \neq j$ , on a :  $\rho(t_i, t_j) > \epsilon$ .

· la quantité  $E(T, \epsilon) = \log M(T, \epsilon)$  est appelée la capacité métrique de l'espace  $T$ .

**IV.4.3.2. Quelques résultats préliminaires**

Nous donnons maintenant quelques résultats techniques que nous utiliserons par la suite :

**Théorème 9** [60] p. 193 : borne inférieure de Sudakov.

Soit  $(X_t, t \in T)$  un processus gaussien centré et  $\epsilon > 0$ , alors

$$E \left[ \sup_{t \in T} X_t \right] \geq 2^{-1/2} v [M(t, \epsilon)] |E(T, \epsilon)|^{1/2} \times \epsilon \tag{16}$$

où

$$v(n) = \begin{cases} 0,64 & \text{si } 1 < n < 23 \\ 2^{1/2} - (\log n)^{-1/2} & \text{si } n > 24 \end{cases} \tag{17}$$

Démonstration.

Voir [60] p.193-194.

**Théorème 10** [60] p.195

Soit  $(X_t, t \in T)$  un processus gaussien centré de variance constante  $\sigma^2 > 0$ . Pour tout  $t \in T$ ,  $V(X_t) = \sigma^2$

alors, pour  $r \in \mathbb{R}$ , on a :

$$P \left[ \sup_{t \in T} X_t > r \right] \geq \frac{1 - \Phi(r/\sigma - \varphi(r))}{2} \tag{18}$$

où  $\Phi(t) = \int_{-\infty}^t \exp\left(-\frac{x^2}{2}\right) dx$

$$\varphi(r) = \sup_{r < \sigma} \left\{ - \left( \frac{\varepsilon^2}{2\sigma^3} + \frac{3\varepsilon^4}{8\sigma^5} \right) r + \varepsilon^2 [M(t, \varepsilon)] \sigma^{-1} \left[ \frac{E(t, \varepsilon)}{2} \right]^{1/2} \right\} \quad (19)$$

La fonction  $\varphi$  étant définie au Théorème 9.

Démonstration : voir [60] p.195-196

Application :

Les résultats précédents sont appliqués au cas d'un processus gaussien qui est notre cas d'étude.

Si  $(X_t, t \in [0, 1])$  est un processus gaussien tel que les deux conditions suivantes soient remplies :

1)

$$V(X_t) = 1 \quad \forall t \in [0, 1] \quad (20)$$

2) pour tous  $s$  et  $t$  vérifiant la condition  $|s - t| < \delta_0$  ( donné ), on a :

$$a^2 |s - t|^{2\beta} \leq E |X_s - X_t|^2, \quad 0 < \beta \leq 1 \quad (21)$$

alors M-A Lifshits montre dans [59] p.197 que :  $P \left[ \sup_{t \in [0, 1]} X_t > r \right] > \alpha_r$  avec

$$\alpha_r \approx r^{-1+1/2\beta} \exp\left(-\frac{r^2}{2}\right) \quad (22)$$

Nous appliquerons ce dernier résultat à notre processus  $Z(s)$ .

On montrera tout d'abord que le processus  $Z(s)$  que nous définirons ci-dessous vérifie les conditions (20) et (21).

#### IV.4.3.3. Minoration de $P \left[ \sup_{s \in [0, 1]} |Z(s)| \geq C_\alpha \right]$

On a montré que, avec la base trigonométrique  $V(Z(s)) = k - 1$ , d'où :

$$V \left[ \frac{1}{\sqrt{k-1}} Z(s) \right] = 1$$

Posons  $Z'(s) = \frac{1}{\sqrt{k-1}}Z(s)$ .

$V(Z'(s)) = 1$  (d'où hypothèse (20) vérifiée par  $Z'(s)$ ).

Par ailleurs, calculons  $E[Z'(s)-Z'(t)]^2$  :

$$E[Z'(s)-Z'(t)]^2 = \frac{1}{k-1} \sum_{j=2}^k E[Y_j^2 (e_j(s) - e_j(t))^2] + \frac{1}{k-1} \sum_{j \neq j'} E[Y_j Y_{j'} (e_j(s) - e_j(t)) (e_{j'}(s) - e_{j'}(t))].$$

Comme les  $Y_j$  sont indépendantes, centrées et réduites, alors :

$$E[Z'(s)-Z'(t)]^2 = \frac{1}{k-1} \sum_{j=2}^k (e_j(s) - e_j(t))^2. \quad (23)$$

Pour  $\delta_0$  fixé, montrons qu'on peut trouver  $\alpha \in \mathbb{R}$  tel que :

$$\alpha^2 |s-t|^{2\beta} \leq \frac{1}{k-1} \sum_{j=2}^k (e_j(s) - e_j(t))^2 \text{ avec } 0 < \beta \leq 1.$$

Les fonctions  $e_j$  étant continues, on sait que :

$$\inf_{s,t} \sum_{j=2}^k (e_j(s) - e_j(t))^2 \text{ existe pour } (s,t) \in [0,1] \times [0,1].$$

$|s-t| < \delta_0$

Soit  $\Lambda$  cette quantité, alors il suffit de prendre  $\alpha$  tel que :

$$\alpha^2 < \frac{\Lambda}{(k-1)\delta_0^{2\beta}};$$

donc (21) est vérifiée pour  $Z'(s)$ .

En conséquence, nous pouvons appliquer (22) au processus  $Z'(s)$

$P[\sup Z'(s) > \gamma] < P[\sup |Z'(s)| > \gamma]$  alors

$P[\sup Z'(s) > \gamma] \geq \alpha_\gamma \Rightarrow P[\sup |Z'(s)| > \gamma] \geq \alpha_\gamma$  avec

$$\alpha_\gamma \approx \gamma^{-1+1/2\beta} \exp\left(-\frac{\gamma^2}{2}\right).$$

Or

$$P[\sup |Z(s)| > r] = P\left[\sup \frac{1}{\sqrt{k-1}} |Z(s)| > \frac{r}{\sqrt{k-1}}\right] = P\left[\sup |Z'(s)| > \frac{r}{\sqrt{k-1}}\right].$$

Ce qui donne, en notant par  $\alpha$  le niveau du test :

$$\alpha = P[\sup |Z(s)| > C_\alpha] > \eta_\alpha \quad (24)$$

$$\eta_n \approx \left[ \frac{C_n}{\sqrt{k-1}} \right]^{1+1/2\beta} \exp \left( - \frac{(C_n/\sqrt{k-1})^2}{2} \right)$$

#### IV.4.4. Cas où les $c_j$ sont lipschitziennes

Nous faisons l'hypothèse lipschitzienne pour utiliser un résultat annoncé par Fernique en 1964 et démontré par Marcus en 1970 ([57] p.219 - 220) ci-dessous.

**Lemme 1** [57] p.219.

Soit  $(X_t, t \in [0, 1])$  un processus normal avec

$$V(X_0) = \sigma^2 > 0 \quad (25)$$

telle que :

$$E[X_s - X_t]^2 \leq C |s - t|^\beta, \quad 0 < \beta < 2 \quad (26)$$

Alors, il existe une constante  $C_\beta > 0$  dépendant uniquement de  $\beta$ , tel que pour tout  $x$ , on a :

$$P \left[ \sup_{t \in [0,1]} X_t > x \right] \leq 1 \exp \left( - \frac{C_\beta x^2}{C} \right) + \frac{1}{2} \exp \left( - \frac{x^2}{8\sigma^2} \right) \quad (27)$$

Si nous supposons que les  $c_j$  sont lipschitziennes, alors

$\forall j = 2, \dots, k;$

$\exists d_j > 0; |c_j(x) - c_j(t)| \leq d_j |x - t|$  donc  $|c_j(x) - c_j(t)|^2 \leq d_j^2 |x - t|^2$

Vérifions alors les hypothèses du lemme précédent à notre processus  $Z(s)$ .

$$V[Z(0)] = V \left[ \sum_{j=2}^k Y_j c_j(0) \right] = \sum_{j=2}^k c_j^2(0)$$

Comme  $c_j \neq 0 \forall j = 2, \dots, k$ , alors :

$$V[Z(0)] = \sum_{j=2}^k c_j^2(0) = \sigma^2 \quad \text{d'où (25).}$$

De même;

$$E[Z(s) - Z(t)]^2 = \sum_{j=2}^k [c_j(s) - c_j(t)]^2$$

par suite  $E[Z(s)-Z(t)]^2 < \sum_{j=2}^k d_j^2 |s-t|^2$

prenons  $C = \sum_{j=2}^k d_j^2 = C_0 \sigma^2$

on obtient alors  $E[Z(s)-Z(t)]^2 \leq C |s-t|^2$

Il suffit de prendre  $\beta = 2$  et (26) est vérifiée.

On peut donc appliquer (27) en remarquant que :

$$P \left[ \sup_{t \in [0,1]} Z(s) > \epsilon \right] \leq P[\sup |Z(s)| > \epsilon] \leq 2P[\sup Z(s) > \epsilon]$$

on a alors :  $\alpha = P[\|Z\| \geq D_\alpha] = P[\sup |Z(s)| > D_\alpha]$

c'est à dire

$$\alpha \leq 8 \exp\left(-\frac{C_0 D_\alpha^2}{C}\right) + \exp\left(-\frac{D_\alpha^2}{8 \sum_{j=2}^k r_j^2(t)}\right) \quad (28)$$

où  $\{\|Z\| > D_\alpha\}$  désigne la région critique du test de niveau  $\alpha$ .

## Conclusion

Pour la recherche du niveau de test, nous avons examiné trois cas qui sont en fait des cas assez particuliers.

On n'a donc pas pu obtenir la valeur exacte du niveau du test compte tenu du fait que la détermination de la loi de  $Z$  n'a pas été possible. C'est un problème très difficile et dans la littérature également ce problème reste ouvert. Le problème a été abordé de manière générale par des majorations ou minoration.

Ultérieurement nous nous intéresserons à l'utilisation du Théorème de Slepian-Schiffli suivant :

### Théorème 11 [60] p.189

Soit  $X = (X_1, X_2, \dots, X_n)$  et  $Y = (Y_1, Y_2, \dots, Y_n)$  deux processus gaussiens centrés, tels que :

1) Pour tout  $i = 1, 2, \dots, n$ ;  $V(X_i) = V(Y_i)$

2) Pour tout  $i \neq j$ ;  $E[X_i X_j] \leq E[Y_i Y_j]$

alors, pour tout  $r \in \mathbb{R}$ , on a :

$$P[\sup X_i > r] \geq Z[\sup Y_i > r]$$

Dans notre cas :

$$\text{Cov}(Z(s), Z(t)) = E[Z(s)Z(t)] = \sum_{j=2}^k c_j(s)c_j(t)$$

Nous espérons ultérieurement faire d'autres approches pour l'étude des tests Banachiques : le cas de  $C[0, 1]$  par l'utilisation du théorème de SLEPIAN-SCHLAFLI ci-dessus. Il s'agira alors de trouver un processus gaussien dont la covariance  $r(s, t)$  ne dépend que de  $|s - t|$  c'est-à-dire pour obtenir un processus stationnaire au second ordre de telle sorte que :  $\sum_{j=2}^k c_j(s)c_j(t) < r(s, t)$ ,

dans ce cas on pourrait obtenir une meilleure approximation du niveau du test.

Un autre problème important à étudier serait le cas où le noyau  $K$  est de dimension infinie.

# CHAPITRE V Tests de non-influence

## V.1 Introduction et hypothèses générales

### V.1.1 Introduction

L'estimation et les tests non paramétriques pour la fonction de régression tant pour un processus à temps continu que discret ont été étudiés par de nombreux auteurs. Plusieurs méthodes ont été utilisées dont celle du noyau, celle des séries orthogonales et celle des moindres carrés.

Box et Watson ont étudié dans [25, 26], la robustesse du test de régression par rapport à la non normalité. L'estimation du paramètre de décentrage dans les tests de régression a été examinée par Bertraud [5]; Yatchew [78] a étudié les tests non paramétriques de régression basés sur la méthode des moindres carrés.

J. D. Hart a étudié dans son livre "Nonparametric smoothing and lack of fit tests" [19] comment une variable  $Y$  se comporte si elle est fonction d'une autre variable déterministe  $x$ . La relation entre  $x$  et  $Y$  s'exprimant par  $Y = \tau(x) + \varepsilon$ ,  $\tau$  désignant la fonction de régression,  $\varepsilon$  le terme d'erreur et  $x$  étant la variable déterministe. Il considère une série  $(x_1, y_1), \dots, (x_n, y_n)$  qui contient une information sur la fonction  $\tau$  et à partir de ces données, diverses conjectures ou estimations de  $\tau$  sont effectuées.

Par la suite Hart et Hart et al. ont développé des tests sous plusieurs alternatives portant sur la régression, on peut citer parmi d'autres, des tests pour la régression simple et la régression multiple [1], comparant les puissances de la combinaison de plusieurs modèles à partir des simulations. Dans [59] Lee et Hart utilisent les fonctions trigonométriques pour estimer  $\tau$  et donnent des méthodes pour choisir l'ordre des coefficients de Fourier.

Dans ce travail nous nous intéressons à une régression de la forme  $Y =$



$r(X) + \varepsilon$  où  $X$  est une variable aléatoire, pour tester l'hypothèse de non-influence  $H_0 : r(x) = C$  où  $C$  est une constante.

Après avoir défini l'estimateur de  $r(x)$  et la statistique du test, nous étudions dans la deuxième Section le test pour les observations indépendantes. Le cas des observations corrélées est traité dans la Section 3. Les simulations apparaissent à la Section 4.

Considérons les couples de v.a.  $(X_i, Y_i)$  avec  $1 \leq i < n$ , définies sur le même espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  et à valeurs dans l'espace mesuré  $(E \times \mathbb{R}, \mathcal{B} \times \mathcal{B}_2, \nu \times \lambda)$  où  $\nu$  est une mesure  $\sigma$ -finie sur  $E$  et  $\lambda$  la mesure de Lebesgue. On suppose que les  $X_i$  sont de même loi connue  $\mu$ . On suppose également que les  $(X_i, Y_i)$  sont de même loi pour laquelle existe la densité  $f$ .

Soit  $f(x, y)$  la densité commune des  $(X_i, Y_i)$ . On suppose de plus l'existence de  $g(x) = \int y f(x, y) dy$  et d'une fonction de régression  $r$  de  $Y$  par rapport à  $X$ , définie par

$$r(x) = E(Y|X = x) = \begin{cases} \frac{g(x)}{f(x)} & \text{si } f \neq 0 \\ E(Y) & \text{sinon} \end{cases}$$

On va considérer  $r(x) = E(Y|X = x)$  dans le modèle  $Y = r(X) + \varepsilon$  (*modèle de régression aléatoire*) et on se propose de tester l'hypothèse  $H_0 : r(x) = C$  où  $C$  est une constante.

Sans perte de généralité, nous prenons  $C = 0$  (si  $C$  est connue) puisque c'est possible d'appliquer le test aux observations  $Y_i - C$  en considérant le modèle  $Y - C = r(X) + \varepsilon$ .

De plus si  $r_0(x)$  est une fonction spécifiée on peut utiliser le test aussi pour l'hypothèse  $H_0 : r(x) = C + r_0(x)$  en appliquant le test aux observations  $Y_i - C - r_0(X_i)$ .

On suppose que  $r \in L^2(\mu)$ , donc  $r(x)$  s'écrit  $r(x) = \sum_{j=0}^{\infty} b_j e_j(x)$  avec les coefficients de Fourier  $b_j = \langle r, e_j \rangle = \int r(x) e_j(x) d\mu(x)$ .

### V.1.2 Estimation de $r(x)$

Pour  $k$  entier positif fixé, considérons un système orthonormal  $\{c_0, c_1, \dots, c_k\}$  avec  $c_0 = 1$ , dans  $E_k$ , qui engendre le sous-espace, de dimension  $k + 1$ . L'estimateur de la fonction de régression par projection sur  $E_k$  est défini par :

$$r_n(\mathbf{X}, \mathbf{Y}; x) = r_n(x) = \sum_{j=0}^k \hat{b}_{jn} c_j(x) \quad (29)$$

où  $\hat{b}_{jn} = \frac{1}{n} \sum_{i=1}^n Y_i c_j(X_i)$  est l'estimateur empirique sans biais du coefficient de Fourier  $b_j$ .

En effet

$$\begin{aligned} b_j &= \int r(x) c_j(x) d\mu(x) = \int E(Y|X=x) c_j(x) d\mu(x) \\ &= E(E(Y|X=x) c_j(x)) = E(Y c_j(X)) \end{aligned}$$

car on a  $E(YX) = E(E(YX|X)) = E(XE(Y|X))$ . On observe que pour  $j = 0$  on a  $\hat{b}_{0n} = \frac{1}{n} \sum_{i=1}^n Y_i = \bar{Y}$ .

### V.1.3 Définition de la statistique du test

Pour tester l'hypothèse  $H_0 : "r(x) = 0"$ , considérons la distance  $d(r_n, 0) = \|r_n - 0\|$  dans  $L^2(\mu)$  entre la fonction de régression estimée et celle qui est vraie sous  $H_0$ , où  $\|\cdot\|$  est la norme dans  $L^2(\mu)$ .

En considérant la statistique hilbertienne

$$R_n = \sqrt{n} r_n$$

et sa norme dans  $L^2(\mu)$ ,  $\|R_n\| = \sqrt{n} \|r_n\|$ , on a

$$\|R_n\|^2 = n \sum_{j=0}^k \hat{b}_{jn}^2.$$

Alors le test associé à  $R_n$  rejette  $H_0$  pour de grandes valeurs de  $\|R_n\|^2$ .

On veut tester l'hypothèse  $H_0 : "r(x) = 0"$  pour tout  $x$  contre  $H_1 : "r(x) \neq 0"$  ce qui est équivalent, en considérant les coefficients de Fourier, à

$$\begin{cases} H_0 : b_j = 0 \forall j \geq 0 \\ H_1 : \exists j \geq 0 : b_j \neq 0 \end{cases}$$

En effet nous considérons l'hypothèse alternative  $H_1(k) : "r(x) \neq 0"$  avec  $r$  telle qu'il existe un  $j \in \{0, 1, \dots, k\}$  pour lequel  $b_j \neq 0$ .

Dans toute la suite nous adopterons les notations suivantes :  
pour chaque  $j$  définissons les v.a. réelles centrées

$$D_{ij} = Y_i c_j(X_i) - E(Y_i c_j(X_i)), \quad 1 \leq i \leq n$$

et posons

$$n^{-1/2} S_{nj} = n^{-1/2} \sum_{i=1}^n D_{ij} = \sqrt{n} (\hat{b}_{jn} - b_j).$$

considérons également les vecteurs

$$n^{-1/2} \mathbf{S}_n = \begin{pmatrix} n^{-1/2} S_{n0} \\ \vdots \\ n^{-1/2} S_{nk} \end{pmatrix} = \sqrt{n} \left[ \begin{pmatrix} \hat{b}_{0n} \\ \vdots \\ \hat{b}_{kn} \end{pmatrix} - \begin{pmatrix} b_0 \\ \vdots \\ b_k \end{pmatrix} \right] = \sqrt{n} (\mathbf{B}_n - \mathbf{b})$$

et pour  $1 \leq i \leq n$

$$\mathbf{D}_i = \begin{pmatrix} D_{i0} \\ \vdots \\ D_{ik} \end{pmatrix}$$

et la combinaison linéaire  $V_i = \sum_{j=0}^k c_j D_{ij} = \mathbf{c}^T \mathbf{D}_i$  avec

$$\mathbf{c} = (c_0, \dots, c_k)^T \in \mathbb{R}^{k+1}.$$

## V.2 Cas des observations indépendantes

### V.2.1 Comportement asymptotique de $\|R_n\|^2$

Les théorèmes suivants nous donnent les lois limites de la statistique

$$\|R_n\|^2 = n \sum_{j=0}^k (\hat{b}_{jn} - b_j)^2$$

sous l'hypothèse  $H_1 : "r \neq 0"$  (théorème 12), et de la statistique  $\|R_n\|^2$  sous l'hypothèse nulle (théorème 13).

### **Théorème 12**

Sous  $H_1$  on a

$$\|R_n\|^2 \xrightarrow{\mathcal{L}} Q'$$

où la v.n.  $Q'$  a pour fonction caractéristique  $\varphi(t) = |\det(I_{k+1} - 2it\Sigma)|^{-1/2}$ ,  $\Sigma$  désignant la matrice avec les éléments définis par

$$\sigma_{jl} = E(|Y_i c_j(X_i) - b_j| |Y_i c_l(X_i) - b_l|) \quad (30)$$

et  $I_{k+1}$  désigne la matrice identique d'ordre  $k+1$ .

*Démonstration.* L'application du Théorème Central Limite dans le cas multivarié à la suite  $(D_i)_{1 \leq i \leq n}$ , où les  $D_i$  sont de même loi et indépendantes (iid), nous permet d'écrire

$$\sqrt{n}(\mathbf{B}_n - \mathbf{b}) = n^{-1/2} \sum_{i=1}^n \xrightarrow{\mathcal{L}} \mathbf{Z} \sim \mathcal{N}_{k+1}(0, \Sigma)$$

où  $\Sigma$  a l'élément  $\sigma_{jl} = E(D_{ij} D_{il}) = E(|Y_i c_j(X_i) - b_j| |Y_i c_l(X_i) - b_l|)$

En utilisant la conservation de la convergence en loi par continuité, on a

$$n(\mathbf{B}_n - \mathbf{b})^T I_{k+1} (\mathbf{B}_n - \mathbf{b}) \xrightarrow{\mathcal{L}} \mathbf{Z}^T I_{k+1} \mathbf{Z}$$

pour les formes quadratiques.

Pour conclure il suffit de remarquer que

$$n(\mathbf{B}_n - \mathbf{b})^T I_{k+1} (\mathbf{B}_n - \mathbf{b}) = n \sum_{j=0}^k (\hat{b}_{jn} - b_j)^2 - \|R_n\|^2,$$

$\mathbf{Z}^T I_{k+1} \mathbf{Z} = Q'$  et d'utiliser le théorème 1 de Barra ([4] p. 88).

■

### **Théorème 13**

Sous  $H_0$  : " $r = 0$ " et si  $E(Y_i^2 | X_i) = \gamma$ , où  $\gamma$  est une constante non nulle, on a

$$\|R_n\|^2 \xrightarrow{\mathcal{L}} \gamma Q$$

où  $Q \sim \chi_{k+1}^2$ .

*Démonstration.* Sous  $H_0$  les coefficients de Fourier sont égaux à zéro, c'est-à-dire  $\mathbf{b} = \mathbf{0}$ . Donc à partir de la démonstration du théorème précédent on a

$$\|R_n\|^2 = n \mathbf{B}_n^T I_{k+1} \mathbf{B}_n \stackrel{L}{\rightsquigarrow} \mathbf{Z}^T I_{k+1} \mathbf{Z}.$$

De plus la formule (30) devient

$$\begin{aligned} \sigma_{jl} &= E(Y_i^2 e_j(X_i) e_l(X_i)) \\ &= E(E(Y_i^2 | X_i) e_j(X_i) e_l(X_i)) \end{aligned}$$

et avec la condition  $E(Y_i^2 | X_i) = \gamma$  on trouve

$$\sigma_{jl} = \begin{cases} 0 & \text{si } j \neq l \\ CE(e_j^2(X_i)) = \gamma & \text{si } j = l \end{cases}$$

c'est-à-dire  $\Sigma = \gamma I_{k+1}$ .

Alors  $\mathbf{Z}^T I_{k+1} \mathbf{Z} = \sum_{j=0}^k \lambda_j (U_j)^2 = \gamma Q$ , puisque  $\lambda_j = \gamma$  sont les valeurs propres de  $\Sigma$  et les v.a.  $U_j \sim \mathcal{N}(0, 1)$  sont indépendantes. ■

Les théorèmes suivants donnent les vitesses de convergence vers les lois limites.

### **Théorème 14**

Sous les conditions du théorème 12, posons

$$\Delta_n = \sup_{t \in \mathcal{R}} |P(\|R_n\|^2 \leq t) - P(Q \leq t)|.$$

Si  $E(\|\mathbf{D}_i\|^3) = M < \infty$  et  $\Sigma$  est définie positive alors

$$\Delta_n \leq c_0 M \frac{(k+1)^4}{\min_{0 \leq j \leq k} \lambda_j^{3/2}} n^{-1/2}$$

où  $c_0$  est une constante universelle et les  $\lambda_j$  sont les valeurs propres de  $\Sigma$ .

*Démonstration.* Observons que

$$\Delta_n = \sup_{u \in \mathcal{R}^{k+1}} |P(\sqrt{n}(\mathbf{B}_n - \mathbf{b}) \leq u) - P(\mathbf{Z} \leq u)|$$

et rappelons que  $\sqrt{n}(\mathbf{B}_n - \mathbf{b}) = n^{-1/2} \sum_{i=1}^n \mathbf{D}_i$ . Alors on peut appliquer le théorème de Sazonov (voir Appendice I chapitre I) aux v.a.  $\mathbf{D}_i$ , à valeurs dans  $\mathbb{R}^{k+1}$ , iid, centrées, avec la même matrice de variance-covariance que  $\mathbf{Z}$ . Si nous prenons comme  $(\mathbf{t}_0, \mathbf{t}_1, \dots, \mathbf{t}_k)$  la base des vecteurs propres unitaires de  $\Sigma$ , les produits scalaires  $\langle \mathbf{D}_i, \mathbf{t}_j \rangle$  sont des v.a. réelles non corrélées puisque

$$E(\langle \mathbf{D}_i, \mathbf{t}_j \rangle \langle \mathbf{D}_i, \mathbf{t}_l \rangle) = \mathbf{t}_j^T \Sigma \mathbf{t}_l = \mathbf{t}_j^T (\lambda_l \mathbf{t}_l) = \lambda_l \delta_{jl}$$

où  $\lambda_l$  est la valeur propre associée à  $\mathbf{t}_l$ . Par le théorème de Sazonov, on trouve

$$\Delta_n \leq c_0 (k+1)^3 \sum_{j=0}^k \rho_j^{(0)} n^{-1/2}$$

avec

$$\rho_j^{(0)} = \frac{E(\langle \mathbf{D}_1, \mathbf{t}_j \rangle^3)}{[E(|\langle \mathbf{D}_1, \mathbf{t}_j \rangle|^2)]^{3/2}} = \frac{E(\|\mathbf{D}_1\|^3 \|\mathbf{t}_j\|^3)}{\lambda_j^{3/2}} = \frac{M}{\lambda_j^{3/2}}$$

Pour conclure il suffit de remarquer que

$$\sum_{j=0}^k \rho_j^{(0)} < \frac{M(k+1)}{\left(\min_{0 \leq j < k} \lambda_j\right)^{3/2}}$$

■

### Théorème 15

Sous les conditions du théorème 13, posons

$$\Delta_n = \sup_{t \in \mathbb{R}} |P(\|\hat{R}_n\|^2 < t) - P(Q \leq t)|.$$

Si  $E(|\mathbf{D}_i|^3) < \infty$  pour chaque  $j$  alors

$$\Delta_n \leq c_0 \frac{(k+1)^3}{\gamma^{3/2}} \sum_{j=0}^k E(|\mathbf{D}_i|^3) n^{-1/2}$$

où  $c_0$  est une constante universelle.

Démonstration. Calculant la démonstration du théorème 14 on voit que la base des vecteurs propres  $(t_0, t_1, \dots, t_k)$  de  $\Sigma$  coïncide avec la base canonique de  $\mathbb{R}^{k+1}$ , les produits scalaires  $\langle D_i, t_j \rangle$  sont les v.a. réelles non corrélées  $D_{ij}$  et les valeurs propres  $\lambda_j$  qui sont toutes égales à  $\gamma$ .

■

### **Théorème 16**

Sous  $H_1 : "r(x) \neq 0"$  avec  $r$  tel qu'il existe un  $j \in \{0, 1, \dots, k\}$  pour lequel  $b_j \neq 0$ , on a

$$\|R_n\|^2 \xrightarrow{P^*} \infty.$$

Démonstration. D'après la loi forte des grands nombres on a

$$\bar{b}_{jn} = \frac{1}{n} \sum_{i=1}^n Y_i e_j(X_i) \xrightarrow{P^*} E(Y_i e_j(X_i)) = b_j$$

avec  $b_j \neq 0$  (en considérant le  $j$  tel que  $b_j \neq 0$ ).

Donc  $\hat{b}_{jn}^2 \xrightarrow{P^*} b_j^2$  et par conséquent  $n \sum_{j=0}^k \hat{b}_{jn}^2 \xrightarrow{P^*} \infty$ . ■

### **V.2.2 Construction et consistance du test**

Ce qui précède permet de dire que si  $\alpha \in ]0, 1[$  tel que

$$P(\chi_{k+1}^2 > m) = \alpha$$

on peut alors construire un test de région critique  $\{(x_1, \dots, x_n) : \|R_n\|^2 > m\}$  de niveau asymptotique  $\alpha$  et, sous l'hypothèse nulle, on a

$$\left| P\left(\sum_{j=0}^k \hat{b}_{jn}^2 > \frac{m}{n}\right) - \alpha \right| = O(n^{-1/2})$$

Comme la constante  $\gamma$  est inconnue, pour l'utilisation du test on l'estimera par la moyenne empirique

$$\hat{\gamma} = \frac{1}{k+1} \sum_{j=0}^k \hat{\gamma}_j = \frac{1}{n(k+1)} \sum_{j=0}^k \sum_{i=1}^n Y_i^2 e_j^2(X_i).$$

Rappelons que le test de région critique  $\|R_n\|^2 > w_n$  est *consistent* si on a simultanément :

$$n) \alpha_n = P(\|R_n\|^2 > w_n) \xrightarrow{n \rightarrow \infty} 0.$$

$$1. \beta_n(r_1) = P(\|R_n\|^2 > w_n) \xrightarrow{n \rightarrow \infty} 1, r_1 \in H_1.$$

Le lemme suivant est utilisé dans la preuve du théorème sur la consistance.

### Lemme 2

Pour tout  $\eta > 0$  il existe une fonction de régression  $r_1 \in H_1(k)$  telle que

$$0 < \sum_{j=0}^k b_j^2(r) < \eta.$$

*Démonstration.* Pour  $\eta > 0$  donné, prenons  $r_1(x) = b_1 e_1(x)$  dans  $H_1(k)$  choisissant  $0 < b_1 < \sqrt{\eta}$ , alors  $\sum_{j=0}^k b_j^2 = b_1^2 < \eta$ . ■

Le théorème suivant donne les conditions nécessaires et suffisantes pour la consistance du test.

### Théorème 17

Le test de région critique  $\|R_n\|^2 > w_n$  est consistant par rapport à  $H_1(k)$  si et seulement si

$$w_n \xrightarrow{n \rightarrow \infty} \infty \quad \text{et} \quad \frac{w_n}{n} \xrightarrow{n \rightarrow \infty} 0.$$

*Démonstration.* L'affirmation du théorème est équivalente aux trois points suivants :

1.  $\alpha_n \rightarrow 0 \iff w_n \rightarrow \infty$ ,
2.  $\frac{w_n}{n} \rightarrow 0 \iff \beta_n(r_1) \rightarrow 1$ ,
3.  $\beta_n(r_1) \rightarrow 1 \iff \frac{w_n}{n} \rightarrow 0$ .

Alors prouvons les trois implications.

1. D'après le théorème 15 on a  $|\alpha_n - P(cQ > w_n)| = O(n^{-1/2})$ , donc  $\alpha_n \rightarrow 0$  si et seulement si  $w_n \rightarrow \infty$ .
2. Selon  $H_1(k)$  il existe un  $j \in \{0, 1, \dots, k\}$ , soit  $j_0$ , tel que  $b_{j_0} = E(Y e_{j_0}(X)) \neq 0$ . Soient un entier  $m > 1$  et  $N(b_{j_0}, m)$  le plus petit entier tel que



$\frac{w_n}{n} \leq \frac{b_{j_0}^2}{m^2}$  pour  $n \geq N(b_{j_0}, m)$ .  
Comme

$$\|R_n\|^2 \leq w_n \quad \longrightarrow \quad \sum_{j=0}^k \widehat{b}_{jn}^2 \leq \frac{w_n}{n} \quad \longrightarrow \quad |\widehat{b}_{j_0 n}| < \sqrt{\frac{w_n}{n}},$$

pour  $n > N(b_{j_0}, m)$  on a  $|\widehat{b}_{j_0 n}| \leq \frac{|b_{j_0}|}{m}$ , d'où

$$|\widehat{b}_{j_0 n} - b_{j_0}| \geq |b_{j_0}| - |\widehat{b}_{j_0 n}| \geq |b_{j_0}| - \frac{|b_{j_0}|}{m} = |b_{j_0}| \frac{m-1}{m}.$$

Alors pour  $n > N(b_{j_0}, m)$

$$\begin{aligned} 1 - \beta_n(\tau_1) &= P(\|R_n\|^2 \leq w_n) \leq P\left(|\widehat{b}_{j_0 n} - b_{j_0}| > |b_{j_0}| \frac{m-1}{m}\right) \\ &\leq \frac{E\left(|\widehat{b}_{j_0 n} - b_{j_0}|^2\right)}{|b_{j_0}|^2 \left(\frac{m-1}{m}\right)^2} = \frac{\mathbf{Var}\left(\widehat{b}_{j_0 n}\right)}{b_{j_0}^2 (m-1)^2} \\ &\leq \frac{1}{n} \frac{E\left(D_{j_0}^2\right)}{b_{j_0}^2 (m-1)^2}. \end{aligned}$$

Donc  $1 - \beta_n(\tau_1) \rightarrow 0$  pour  $n \rightarrow \infty$ .

3. On va supposer que  $\frac{w_n}{n} \rightarrow 0$ . Alors ils existent  $\eta > 0$  et une partie infinie  $\mathbb{N}_1 \subseteq \mathbb{N}$  tels que  $\frac{w_n}{n} > \eta$  pour  $n \in \mathbb{N}_1$ .

Considérons  $\nu \in H_1(k)$  tel que  $0 < \sum_{j=0}^k b_j^2 < \eta$ , dont l'existence est assurée d'après le Lemme 2.

La condition  $\beta_n(\tau_1) \rightarrow 1$  donne pour  $n \in \mathbb{N}_1$

$$P\left(\sum_{j=0}^k \widehat{b}_{jn}^2 < \eta\right) \leq P\left(\sum_{j=0}^k \widehat{b}_{jn}^2 \leq \frac{w_n}{n}\right) \underset{n \rightarrow \infty}{\rightarrow} 0 \quad (31)$$

D'autre part (voir la preuve du théorème 16)

$$\sum_{j=0}^k \widehat{b}_{jn}^2 \xrightarrow{P^*} \sum_{j=0}^k \widehat{b}_j^2 < \eta$$

d'où

$$P\left(\sum_{j=0}^k \widehat{b}_{jn}^2 \leq \eta\right) \leq P\left(\sum_{j=0}^k (\widehat{b}_{jn}^2 - b_j^2) \leq \eta - \sum_{j=0}^k b_j^2\right) \underset{n \rightarrow \infty}{\rightarrow} 1$$

ce qui est contradictoire.

■

Remarque Les résultats obtenus ici sont similaires à ceux de Bosq dans [14] page 25.

### V.3 Cas des observations corrélées

Pour toute la suite nous supposons avoir des observations engendrées par le processus  $(X_t, Y_t)_{t \in \mathbb{Z}}$  qui est  $\alpha$ -mélangeant avec pour coefficient de mélangeance  $\alpha_{XY}$ .

#### Lemme 3

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus  $\alpha$ -mélangeant,  $f : (E, \mathcal{B}_E) \rightarrow (F, \mathcal{B}_F)$  une application mesurable. Si  $Y_t = f(X_t)$ , le processus  $(Y_t)_{t \in \mathbb{Z}}$  est  $\alpha$ -mélangeant avec  $\alpha_Y(n) \leq \alpha_X(n) \forall n = 1, 2, \dots$ .

#### Corollaire 8

- Les v.a.  $D_{ij}$  sont  $\alpha$ -mélangeant pour chaque  $j$  fixé, avec  $\alpha_{D_j}(n) \leq \alpha_{XY}(n)$ , puisque c'est possible d'écrire  $D_{ij} = f_j(X_i, Y_i)$  avec  $f_j$  mesurable.
- Les v.a.  $V_i$  sont  $\alpha$ -mélangeant avec  $\alpha_V(n) \leq \alpha_{XY}(n)$ , puisque c'est possible d'écrire  $V_i = \sum_{j=0}^k c_j f_j(X_i, Y_i) = g(X_i, Y_i)$  avec  $g$  mesurable.

#### V.3.1 Comportement asymptotique de $\|R_n\|^2$

Nous donnons ici un théorème général de type Tikhomirov [75] (voir Appendice, théorème 24) que nous utiliserons dans les démonstrations suivantes.

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus centré, stationnaire,  $\alpha$ -mélangeant et à valeurs dans  $\mathbb{R}^k$ , donc  $X_t = (X_{t1}, \dots, X_{tk})^T$ . Considérons la somme  $S_n = \sum_{i=1}^n X_i$ , la combinaison linéaire  $V_i = c^T X_i$  avec  $c = (c_1, \dots, c_k)^T \in \mathbb{R}^k$  et l'ensemble des conditions suivantes :

- (A)  $E(|X_{i1}|^{2^{1+\delta_X}}) < \infty$  pour  $\delta_X > 0$ ;  
 (B)  $E(|V_i|^{2^{1+\delta_V}}) < \infty$  pour  $\delta_V > 0$  (on sait que B  $\rightarrow$  A);  
 (C)  $\sigma^2 > 0$  où  $\sigma^2 = E(|V_1|^2) + 2 \sum_{i=2}^{\infty} E(V_1 V_i) = \sum_{i=-\infty}^{\infty} E(V_1 V_i)$ .

**Théorème 18**

Sous les conditions (B)–(C), posons

$$\Delta_n = \sup_{u \in \mathbb{R}} \left| P \left( \|n^{-1/2} \mathbf{S}_n\|^2 \leq u \right) - P \left( \|Z\|^2 \leq u \right) \right|$$

avec  $Z \sim \mathcal{N}_k(0, \Sigma)$ , où la matrice  $\Sigma$  est définie par

$$\sigma_{ij} = E(X_{1j} X_{1i}) + \sum_{i=2}^{\infty} E(X_{1j} X_{1i}) + \sum_{i=2}^{\infty} E(X_{1j} X_{1i}) . \quad (32)$$

Alors on a :

- i) si  $\alpha_X(n) = O(n^{-\beta(2^{1+\delta})(1+\delta)/\delta^2})$  pour  $\beta > 1$ , alors il existe une constante  $\gamma_1$  telle que  $\Delta_n \leq \gamma_1 n^{-\frac{\beta(2^{1+\delta})}{2(1+\delta)}}$ ,
  - ii) si  $\alpha_X(n) = O(e^{-\beta n})$  pour  $\beta > 1$ , alors il existe une constante  $\gamma_2$  telle que  $\Delta_n \leq \gamma_2 n^{-\frac{\delta}{2}} \log^{1+\delta}$ ,
- avec  $\delta = \max(\delta_X, \delta_V)$ .

De plus,

$$\|n^{-1/2} \mathbf{S}_n\|^2 \xrightarrow{\mathcal{L}} \|Z\|^2 = \sum_{j=1}^k \lambda_j^2 U_j^2 \quad (33)$$

où les v.v.  $U_j \sim \mathcal{N}(0, 1)$  sont indépendantes et  $\lambda_j^2$  les valeurs propres de la matrice  $\Sigma$ .

*Démonstration.* Examinons les détails seulement pour la i), comme la ii) se démontre de manière similaire.

Avec les notations et les conditions ci-dessus, on peut appliquer le théorème de Tikhomirov à la suite  $(V_i)$ , laquelle est centrée et à valeurs réelles par définition,  $\alpha$ -mélangeant par le corollaire 6b et stationnaire.

Alors il existe une constante  $\gamma_1$  telle que

$$\sup_{t \in \mathbb{R}} \left| P \left( n^{-1/2} \sum_{i=1}^n V_i \leq t \right) - P(N_\sigma \leq t) \right| \leq \gamma_1 n^{-\frac{\beta(2^{1+\delta})}{2(1+\delta)}} \quad (34)$$

où  $N_n \sim \mathcal{N}(0, \sigma^2)$ , puisque  $\sigma^2 < \infty$  (voir remarque en Appendice A ce chapitre).

Mais

$$\begin{aligned} \sum_{i=1}^n V_i &= c_1 X_{i1} + \dots + c_k X_{ik} + \dots + c_1 X_{ni} + \dots + c_k X_{nk} \\ &= c_1 \sum_{i=1}^n X_{i1} + \dots + c_k \sum_{i=1}^n X_{ik} \\ &= \sum_{j=1}^k c_j S_{nj} = c^T S_n. \end{aligned}$$

Donc  $\forall c \in \mathbb{R}^k$

$$c^T S_n \stackrel{\mathcal{L}}{\rightarrow} \mathcal{N}(0, \sigma^2)$$

et si  $\sigma^2 = c^T \Sigma c$ , où  $\Sigma$  est la matrice de variance-covariance du vecteur  $Z$ , définie strictement positive parce que  $\sigma^2 > 0$  par la condition (C), on a aussi

$$\begin{aligned} c^T S_n &\stackrel{\mathcal{L}}{\rightarrow} \mathcal{N}(0, c^T \Sigma c) \\ c^T S_n &\stackrel{\mathcal{L}}{\rightarrow} c^T Z. \end{aligned} \tag{35}$$

Il reste à voir que  $\sigma^2 = c^T \Sigma c$  avec  $\Sigma$  symétrique.

Comme nous avons appliqué le théorème de Tikhomirov aux  $V_i$ , on a  $\sigma^2 = E(|V_1|^2) + 2 \sum_{i=2}^{\infty} E(V_1 V_i)$ .

Mais

$$\begin{aligned} E(V_1 V_i) &= E\left([c^T X_1] [c^T X_i]^T\right) \\ &= E(c^T X_1 X_i^T c) = c^T E(X_1 X_i^T) c \\ &= c^T \Sigma_{1i} c \end{aligned}$$

et de la même façon

$$E(V_i V_1) = c^T \Sigma_{i1} c$$

où on a posé  $\Sigma_{1i} = E(X_1 X_i^T)$  et  $\Sigma_{i1} = E(X_i X_1^T)$ .

On observe que  $E(V_1 V_i) = E(V_i V_1)$ , alors que  $\Sigma_{1i} = \Sigma_{i1}^T$ .

Donc

$$\begin{aligned} \sigma^2 &= c^T \Sigma_{11} c + \sum_{i=2}^{\infty} c^T \Sigma_{1i} c + \sum_{i=2}^{\infty} c^T \Sigma_{i1} c \\ &= c^T \left( \Sigma_{11} + \sum_{i=2}^{\infty} \Sigma_{1i} + \sum_{i=2}^{\infty} \Sigma_{i1} \right) c \\ &= c^T \Sigma c. \end{aligned}$$

Alors l'élément général de  $\Sigma$  est

$$\sigma_{jt} = E(X_{1j} X_{1t}) + \sum_{i=2}^{\infty} E(X_{1j} X_{it}) + \sum_{i=2}^{\infty} E(X_{it} X_{1j}) \quad (36)$$

et  $\sigma_{jt} = \sigma_{tj}$ .

Par (35) en appliquant le critère de Cramér-Wold on a

$$n^{-1/2} S_n \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}_k [0, \Sigma]. \quad (37)$$

La convergence en loi se conserve par continuité et donc pour les formes quadratiques on a

$$n^{-1} S_n^T I_k S_n \xrightarrow{\mathcal{L}} Z^T I_k Z.$$

De plus

$$n^{-1} S_n^T I_k S_n = \|n^{-1/2} S_n\|^2$$

et

$$Z^T I_k Z = \sum_{j=1}^k \lambda_j^2 (U_j)^2 = \sum_{j=1}^k \lambda_j^2 \chi_j^2$$

où  $\lambda_j^2 > 0$  sont les valeurs propres de  $\Sigma$  et les v.a.  $U_j \sim \mathcal{N}(0, 1)$  sont indépendantes.

Par rapport à la vitesse de convergence, nous avons

$$\begin{aligned}
& \sup_{v \in \mathbb{R}^3} \left| P(n^{-1/2} c^T S_n < t) - P(N_n \leq t) \right| \\
&= \sup_{v \in \mathbb{R}^3} \left| P(c^T [n^{-1/2} S_n] < v) - P(c^T Z < v) \right| \\
&= \sup_{v \in \mathbb{R}^3} \left| P\left(\|c^T [n^{-1/2} S_n]\|^2 \leq v^2\right) - P\left(\|c^T Z\|^2 \leq v^2\right) \right| \\
&= \sup_{v \in \mathbb{R}^3} \left| P\left(\|n^{-1/2} S_n\|^2 \leq \frac{v^2}{\|c\|^2}\right) - P\left(\|Z\|^2 \leq \frac{v^2}{\|c\|^2}\right) \right| \\
&= \sup_{u \in \mathbb{R}^3} \left| P\left(\|n^{-1/2} S_n\|^2 \leq u^2\right) - P\left(\|Z\|^2 \leq u^2\right) \right| \\
&= \Delta_n
\end{aligned}$$

où on a posé  $u^2 = \frac{v^2}{\|c\|^2}$ .

D'après (34) on a l'existence d'une constante  $\gamma_1$  pour laquelle l'inégalité est vraie

$$\Delta_n \leq \gamma_1 n^{-\frac{k_V (k_V - 1)}{2(k_V + 1)}}. \quad (38)$$

#### Remarque

Le critère de Cràmer-Wold a donné la relation (37) qui établit la convergence dans  $\mathbb{R}^k$ .

#### V.3.1.1 Loi limite de $\|R_n\|^2$

Les théorèmes suivants nous donnent les lois limites de la statistique  $\|R_n'\|^2 = n \sum_{j=0}^k (\hat{b}_{j,n} - b_j)^2$  sous l'hypothèse  $H_1 : r(x) \neq 0$  (théorème 19) et de la statistique du test  $\|R_n\|^2$  sous l'hypothèse nulle (théorème 20).

Dans les deux cas donnons des inégalités qui caractérisent la convergence de la statistique considérée à sa distribution limite.

Considérons l'ensemble des conditions suivantes :

- (A)  $E\left[|D_{1j}|^{2+\delta_D}\right] < \infty$  pour  $\delta_D > 0$ ;
- (B)  $E(|V_i|^{2+\delta_V}) < \infty$  pour  $\delta_V > 0$ ;
- (C)  $\sigma^2 > 0$  où  $\sigma^2 = E(|V_1|^2) + 2 \sum_{i=2}^{\infty} E(V_1 V_i) = \sum_{i=-\infty}^{+\infty} E(V_1 V_i)$ .

### Théorème 19

Sous les conditions (B) (C), posons

$$\Delta_n = \sup_{u \in \mathbb{R}} \left| P \left( \|R'_n\|^2 \leq u \right) - P \left( \|U\|^2 \leq u \right) \right|$$

avec  $U = \sum_{j=0}^k \lambda_j U_j e_j$ , où  $\lambda_j^2$  sont les valeurs propres de la matrice  $\Sigma$  définie par

$$\sigma_{ji} = E(D_{1j} D_{1i}) + \sum_{i=2}^{\infty} E(D_{1j} D_{ii}) + \sum_{i=2}^{\infty} E(D_{ij} D_{ii}) \quad (39)$$

et les v.a.  $U_j \sim \mathcal{N}(0, 1)$  sont indépendantes, on a :

- i) si  $\alpha_{XY}(n) = O(n^{-\beta(2+\delta)(1+\delta)/\delta^2})$  pour  $\beta > 1$ , alors il existe une constante  $\gamma_1$  telle que  $\Delta_n \leq \gamma_1 n^{-\frac{\beta(\delta-1)}{2(\delta+1)}}$ ,
- ii) si  $\alpha_{XY}(n) = O(e^{-\beta n})$  pour  $\beta > 1$ , alors il existe une constante  $\gamma_2$  telle que  $\Delta_n \leq \gamma_2 n^{-\frac{\delta}{2}} \log^{1+\delta}$ ,

avec  $\delta = \max(\delta_Y, \delta_X)$ .

De plus,

$$\|R'_n\|^2 \stackrel{\mathcal{L}}{\sim} \sum_{j=0}^k \lambda_j^2 U_j^2. \quad (40)$$

Démonstration. Il suffit d'appliquer le théorème 18 aux v.a.  $D_i$  à valeurs dans  $\mathbb{R}^{k+1}$ . ■

#### Remarque

Comme les v.a.  $U_j^2 \sim \chi_1^2$  sont indépendantes, la v.a.  $\|U\|^2$  a pour fonction caractéristique  $\varphi(t) = [\det(I_{k+1} - 2it\Sigma)]^{-1/2}$ , c'est-à-dire la loi  $\chi^2$  généralisée avec un degré de liberté (voir [4]).

En particulierisant le théorème précédent nous obtenons le comportement asymptotique de  $\|R'_n\|^2$  sous l'hypothèse nulle.

### Théorème 20

### V.3.1.2 Estimation de la matrice de variance-covariance

La distribution asymptotique dans le théorème 19 (ainsi que dans le théorème 20) dépend des valeurs propres  $\lambda_j^2$  de la matrice inconnue  $\Sigma$  d'éléments (39). Par conséquent il est nécessaire d'estimer les éléments de cette matrice pour obtenir l'estimation de ses valeurs propres. Introduisons d'abord

$$\sigma_{jl}(v) = E(D_{0j}D_{vt}).$$

Par la stationnarité des  $\mathbf{D}_i$  on a  $\text{Cov}(\mathbf{D}_0, \mathbf{D}_n) = \text{Cov}(\mathbf{D}_t, \mathbf{D}_{t+v})$  et pour chaque élément  $E(D_{0j}D_{vt}) = E(D_{tj}D_{t+v,l})$  donc

$$\sigma_{jl}(v) = E(D_{tj}D_{t+v,l});$$

de plus l'égalité

$$\sigma_{jl}(-v) = E(D_{tj}D_{t-v,l}) = E(D_{t-v,l}D_{tj}) = \sigma_{lj}(v)$$

est satisfaite.

En posant  $v = i - 1$  on obtient  $\sigma_{jl}(i - 1) = E(D_{1j}D_{il})$  et  $\sigma_{jl}(-(i - 1)) = \sigma_{lj}(i - 1)$ ; aussi, il est possible d'écrire

$$\sigma_{jl} = \sum_{v=-\infty}^{+\infty} \sigma_{jl}(v).$$

Par la suite nous supposons que

$$\sum_{v=-\infty}^{+\infty} |\sigma_{jl}(v)| < \infty \quad j, l = 1, \dots, k. \quad (43)$$

Soit  $\ell_n$  une suite d'entiers positifs vérifiant  $\ell_n < n$  et  $\lim_{n \rightarrow \infty} \ell_n = \lim_{n \rightarrow \infty} \frac{n}{\ell_n} = \infty$ . Nous considérons l'estimateur classique de la covariance croisée du processus stationnaire bivarié  $(D_{tj}, D_{tl})_{t \in \mathbb{Z}}$

$$\hat{\sigma}_{jl}(v) = \frac{1}{n-v} \sum_{t=1}^{n-v} D_{tj}D_{t+v,l} \quad \text{pour } 0 < v \leq \ell_n$$



et

$$\hat{\sigma}_{jl}(v) = \hat{\sigma}_{lj}(-v) \quad \text{pour } -\ell_n \leq v < 0.$$

Cet estimateur est sans biais, puisque  $\mathbb{E}(\hat{\sigma}_{jl}(v)) = \sigma_{jl}(v)$ .

Nous pouvons définir un estimateur pour  $\sigma_{jl}$  par

$$\hat{\sigma}_{jl} = \sum_{v=-\ell_n}^{\ell_n} w_n(v) \hat{\sigma}_{jl}(v)$$

où les poids sont

$$w_n(v) = \left[ 1 - \frac{|v|}{n} \right] \eta \left[ \frac{v}{\ell_n} \right] \mathbb{1}_{\{|v| < \ell_n\}}(v)$$

avec  $\eta$  vérifiant les conditions suivantes

$$\begin{cases} \eta(0) = 1 \\ |\eta(x)| \leq M & \text{pour } |x| < 1 \quad (M > 1) \\ \eta(x) = \eta(-x) \end{cases}$$

De plus il faut que  $\eta(x)$  soit continue en  $x = 0$ .

Nous donnons des résultats analogues à ceux obtenus par Iguaccolo [51] pages 5-8 où les  $Y_{ij} = e_j(X_i) - E(e_j(X_i))$  sont remplacées par les  $D_{ij} = Y_i e_j(X_i) - E(Y_i e_j(X_i))$ .

### Proposition 22

L'estimateur  $\hat{\sigma}_{jl}$  est asymptotiquement sans biais.

### Proposition 23

Pour  $j$  et  $l$  fixés, soient

$$Z_{i,v} = D_{ij} D_{li,v} - \sigma_{jl}(v) \quad \text{pour } 0 \leq v \leq \ell_n$$

et

$$Z_{i,v} = D_{li,v} D_{ij} - \sigma_{jl}(-v) \quad \text{pour } -\ell_n \leq v < 0.$$

Si pour tout  $t$  ( $1 \leq t \leq n$ ) et pour tout  $v$  ( $|v| \leq \ell_n$ ) ils existent  $b_1$  et  $b_2$  tels que

$$|Z_{t,v}| \leq b_1 \quad \text{et} \quad |Z_{t,v}^-| \leq b_2$$

alors pour tout entier  $q_n \in \left[1, \frac{n - l_n}{2}\right]$  et pour tout  $\varepsilon > 0$

$$P \left( \sup_{|v| \leq \ell_n} |\hat{\sigma}_{jt}(v) - \sigma_{jt}(v)| > \varepsilon \right) \leq 4(2\ell_n + 1) \exp \left( -\frac{\varepsilon^2}{8b^2} q_n \right) \quad (44)$$

$$+ 22(2\ell_n + 1) \left(1 + \frac{4b}{\varepsilon}\right)^{1/2} q_n \alpha \left( \left[ \frac{n - l_n}{2q_n} \right] \right)$$

pour tout  $v$ , où  $b = \max(b_1, b_2)$ .

De plus si

$$\lim_{n \rightarrow \infty} \ell_n \exp(-aq_n) = 0 \quad \text{et} \quad \lim_{n \rightarrow \infty} \ell_n q_n \alpha \left( \left[ \frac{n - l_n}{2q_n} \right] \right) = 0, \quad (45)$$

où  $a$  est une constante dépendant de  $\varepsilon$  et  $b$ , alors

$$\hat{\sigma}_{jt}(v) \xrightarrow{P} \sigma_{jt}(v),$$

pendant que si pour tout  $\varepsilon > 0$

$$\sum_n 4(2\ell_n + 1) \exp \left( -\frac{\varepsilon^2}{8b^2} q_n \right) + 22(2\ell_n + 1) \left(1 + \frac{4b}{\varepsilon}\right)^{1/2} q_n \alpha \left( \left[ \frac{n - l_n}{2q_n} \right] \right) < \infty \quad (46)$$

alors

$$\hat{\sigma}_{jt}(v) \xrightarrow{P} \sigma_{jt}(v).$$

### Proposition 24

Sous les suppositions de la proposition 21, pour  $n$  assez grand, pour tout entier  $q_n \in \left[1, \frac{n - l_n}{2}\right]$  et pour tout  $\varepsilon > 0$

$$P(|\hat{\sigma}_{jt} - \sigma_{jt}| > \varepsilon) \leq 4(2\ell_n + 1) \exp \left( -\frac{\varepsilon^2}{32b^2 M^2 (2\ell_n + 1)^2} q_n \right) \quad (47)$$

$$+ 22(2\ell_n + 1) \left[ 1 + \frac{8bM}{\varepsilon} (2\ell_n + 1) \right]^{1/2} q_n \alpha \left( \left[ \frac{n - l_n}{2q_n} \right] \right)$$

où  $b = \max(b_1, b_2)$ .

De plus si

$$\lim_{n \rightarrow \infty} \ell_n \exp\left(-A \frac{q_n}{\ell_n^2}\right) = 0 \quad \text{et} \quad \lim_{n \rightarrow \infty} \ell_n^{3/2} q_n \alpha\left(\left[\frac{n - l_n}{2q_n}\right]\right) = 0, \quad (48)$$

où  $A$  est une constante dépendant de  $\varepsilon$ ,  $M$  et  $b$ , alors

$$\hat{\sigma}_{jl} \xrightarrow{P} \sigma_{jl},$$

pendant que si pour tout  $\varepsilon > 0$

$$\sum_n \left\{ 4(2\ell_n + 1) \exp\left(-\frac{\varepsilon^2}{32b^2 M^2 (2\ell_n + 1)^2}\right) + 22(2\ell_n + 1) \left|1 + \frac{8bM}{\varepsilon}(2\ell_n + 1)\right|^{1/2} q_n \alpha\left(\left[\frac{n - l_n}{2q_n}\right]\right) \right\} < \infty \quad (49)$$

alors

$$\hat{\sigma}_{jl} \xrightarrow{P^*} \sigma_{jl}.$$

### Corollaire 9

Si  $\alpha(n) \leq \rho^n$  avec  $0 < \rho < 1$ ,  $\ell_n \sim n^\gamma$  avec  $0 < \gamma < 1/2$  et  $q_n \sim n^\beta$  avec  $2\gamma < \beta < 1$  alors les conditions (48) sont vérifiées et  $\hat{\sigma}_{jl}$  converge en probabilité vers  $\sigma_{jl}$ .

### Corollaire 10

Si  $\alpha(n) \leq n^{-\delta}$  avec  $\delta > 0$ ,  $\ell_n \sim n^\gamma$  avec  $0 < \gamma < 2\delta/(4\delta + 7)$  et  $q_n \sim n^\beta$  avec  $2\gamma < \beta < (\delta - \frac{3}{2}\gamma)/(1 + \delta)$  alors les conditions (48) sont vérifiées et  $\hat{\sigma}_{jl}$  converge en probabilité vers  $\sigma_{jl}$ .

### Corollaire 11

Si  $\alpha(n) \leq \exp(-\tau n)$  avec  $\tau > 0$ ,  $\ell_n \sim n^\gamma$  avec  $0 < \gamma < 1/2$  et  $q_n \sim n^\beta$  avec  $2\gamma < \beta < 1$  alors les conditions (48) sont vérifiées et  $\hat{\sigma}_{jl}$  converge en probabilité vers  $\sigma_{jl}$ .

Nous obtenons ainsi un estimateur consistant de  $\sigma_{jl}$  pour chaque  $j$  et  $l$ . Donc, on a  $\hat{\Sigma}_n \xrightarrow{P^*} \Sigma$ , où  $\hat{\Sigma}_n = (\hat{\sigma}_{jl})_{1 \leq j, l \leq k+1}$ , par rapport à la convergence pour chaque composante. Cependant, alors que  $\Sigma$  est définie positive, il est possible

que  $\hat{\Sigma}_n$  ne le soit pas. Pour obtenir une matrice d'estimateurs définie positive, nous posons

$$\zeta_n = \left| \inf_{\|\mathbf{c}\| \leq 1} \mathbf{c}^T \hat{\Sigma}_n \mathbf{c} \right| + \frac{\tau}{n}$$

avec  $\mathbf{c} \in \mathbb{R}^{k+1}$  et  $\tau$  une constante positive convenablement choisie. Définissons la nouvelle matrice

$$\hat{\Sigma}_n^1 = \hat{\Sigma}_n + \zeta_n I_{k+1}.$$

**Proposition 25**

La matrice  $\hat{\Sigma}_n^1$  est définie positive et  $\hat{\Sigma}_n^1 \xrightarrow{p.s.} \Sigma$ .

*Démonstration.* Pour prouver que  $\hat{\Sigma}_n^1$  est définie positive il suffit de démontrer que la forme quadratique  $\mathbf{c}^T \hat{\Sigma}_n^1 \mathbf{c}$  est strictement positive pour un vecteur de norme un.

Soit  $\mathbf{c} \in \mathbb{R}^k$  avec  $\|\mathbf{c}\| = 1$  on a :

$$\begin{aligned} \mathbf{c}^T \hat{\Sigma}_n^1 \mathbf{c} &= \mathbf{c}^T \left( \hat{\Sigma}_n + \zeta_n I_{k+1} \right) \mathbf{c} \\ &= \mathbf{c}^T \hat{\Sigma}_n \mathbf{c} + \zeta_n \|\mathbf{c}\|^2 \\ &= \mathbf{c}^T \hat{\Sigma}_n \mathbf{c} + \zeta_n \\ &> 0 \end{aligned}$$

parce que dans le cas où  $\mathbf{c}^T \hat{\Sigma}_n^1 \mathbf{c} < 0$ , la définition de  $\zeta_n$  entraîne  $-\mathbf{c}^T \hat{\Sigma}_n^1 \mathbf{c} < \zeta_n$ .

si  $\mathbf{c} \in \mathbb{R}^k$  tel que  $\|\mathbf{c}\| \leq 1$ , on a :

$$\left| \mathbf{c}^T \hat{\Sigma}_n \mathbf{c} - \mathbf{c}^T \Sigma \mathbf{c} \right| = \left| \mathbf{c}^T \left( \hat{\Sigma}_n - \Sigma \right) \mathbf{c} \right| \leq \left\| \hat{\Sigma}_n - \Sigma \right\| \|\mathbf{c}\|^2 \leq \left\| \hat{\Sigma}_n - \Sigma \right\|. \quad (50)$$

Comme  $k$  est fini,  $\hat{\Sigma}_n \xrightarrow{p.s.} \Sigma$  entraîne que

$$\left\| \hat{\Sigma}_n - \Sigma \right\| = \sum_{j=1}^k \sum_{i=1}^k \left[ \hat{\sigma}_{ji}(n) - \sigma_{ji}(n) \right]^2 \xrightarrow{p.s.} 0$$

et par (50) on a

$$\mathbf{c}^T \hat{\Sigma}_n \mathbf{c} \xrightarrow{p.s.} \mathbf{c}^T \Sigma \mathbf{c} \quad \text{uniformément par rapport à } \mathbf{c},$$

posons  $f_n(\mathbf{c}) = \mathbf{c}^T \hat{\Sigma}_n \mathbf{c}$  et  $f(\mathbf{c}) = \mathbf{c}^T \Sigma \mathbf{c}$ ; donc nous avons  $f_n(\mathbf{c}) \xrightarrow{p.s.} f(\mathbf{c})$  uniformément par rapport à  $\mathbf{c}$ . Or, pour tout  $n$  il existe  $\mathbf{c}_n$  tel que  $f_n(\mathbf{c}_n) = \inf f_n(\mathbf{c})$  et aussi il existe  $\mathbf{c}_0$  tel que  $f(\mathbf{c}_0) = \inf f(\mathbf{c})$ .

Il faut prouver que  $f_n(\mathbf{c}_n) \rightarrow f(\mathbf{c}_0)$ .

Comme  $f(\mathbf{c}_0) \leq f(\mathbf{c}_n)$  et pour  $n$  assez grand  $f_n(\mathbf{c}_n) \leq f(\mathbf{c}_n) + \varepsilon$ , on a  $f(\mathbf{c}_0) - f_n(\mathbf{c}_n) \leq \varepsilon$  pour  $n$  assez grand.

De la même façon  $f_n(\mathbf{c}_n) \leq f(\mathbf{c}_0)$  et pour  $n$  assez grand  $f_n(\mathbf{c}_n) \leq f(\mathbf{c}_0) + \varepsilon$ , donnent  $f(\mathbf{c}_0) - f_n(\mathbf{c}_n) \leq \varepsilon$  pour  $n$  assez grand.

Par suite pour tout  $\varepsilon > 0$  et pour  $n$  assez grand on a  $|f_n(\mathbf{c}_n) - f(\mathbf{c}_0)| \leq \varepsilon$ .

De plus  $f(\mathbf{c}_0) = 0$  parce que la matrice  $\Sigma$  est définie positive et donc

$$\inf_{\|\mathbf{c}\| \leq 1} \mathbf{c}^T \hat{\Sigma}_n \mathbf{c} \xrightarrow{p.s.} \inf_{\|\mathbf{c}\| \leq 1} \mathbf{c}^T \Sigma \mathbf{c} = 0$$

Par conséquent  $\zeta_n \xrightarrow{p.s.} 0$  et alors  $\hat{\Sigma}_n \xrightarrow{p.s.} \Sigma$ . ■

### Remarque

Par continuité, pour les valeurs propres  $\hat{\lambda}_{jn}^2$  de  $\hat{\Sigma}_n$  on obtient

$$\hat{\lambda}_{jn}^2 \xrightarrow{p.s.} \lambda_j^2$$

et il vient :

$$\sum_{j=0}^k \hat{\lambda}_{jn}^2 U_j^2 \xrightarrow{p.s.} \sum_{j=0}^k \lambda_j^2 U_j^2.$$

Alors on peut utiliser la distribution de  $\hat{U}_n^2 = \sum_{j=0}^k \hat{\lambda}_{jn}^2 U_j^2$  comme approximation de la distribution de  $\|R_n\|^2$  comme cas particulier) puisque

$$\begin{aligned} \left| P \left( \|R_n\|^2 \leq u \right) - P \left( \hat{U}_n^2 < u \right) \right| &\leq \\ &\leq \left\{ \left| P \left( \|R_n\|^2 \leq u \right) - P \left( \|U\|^2 < u \right) \right| + \right. \\ &\quad \left. + \left| P \left( \|U\|^2 \leq u \right) - P \left( \hat{U}_n^2 < u \right) \right| \right\} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

$$\begin{aligned} \left| P \left( \|R_n\|^2 \leq u \right) - P \left( \widehat{U}^2 \leq u \right) \right| &\leq \\ &\leq \left| P \left( \|R_n\|^2 \leq u \right) - P \left( \|U\|^2 \leq u \right) \right| \\ &\quad + \left| P \left( \|U\|^2 \leq u \right) - P \left( \widehat{U}^2 \leq u \right) \right| \end{aligned}$$

### V.3.2 Construction et consistance du test

Ce qui précède permet de dire que si  $\alpha \in ]0, 1[$  tel que

$$P \left( \sum_{j=0}^k \lambda_j^2 U_j^2 > w \right) = \alpha$$

on peut alors construire un test de région critique  $\{(x_1, \dots, x_n) : \|R_n\|^2 > w\}$  de niveau asymptotique  $\alpha$ .

Comme nous avons vu au paragraphe V.3.1.2 pour appliquer le test on peut utiliser la distribution de la v.a.  $\widehat{U}_n^2$ , obtenue avec les valeurs propres estimées.

Pour obtenir la consistance du test faisons une hypothèse supplémentaire à l'ensemble (A) (C) du paragraphe V.3.1.1, c'est-à-dire :

$$(D) \quad \sup_{1 \leq i \leq n} \sup_{x \in E} \max_{0 \leq j \leq k} |Y_i c_j(x)| = M < \infty.$$

Le théorème suivant donne les conditions nécessaires et suffisantes pour la consistance du test.

#### **Théorème 22**

Sous les conditions (B) (D) et si

$$\sum_{i=1}^{\infty} \alpha^*(i) < \infty, \quad (51)$$

où  $\alpha^*(i)$  sont les coefficients de mélangeance pour les v.a.  $Y_i c_j(X_i)$ , le test de région critique  $\{\|R_n\|^2 > w\}$  est consistant par rapport à  $H_1(k)$  si et seulement si

$$w_n \xrightarrow[n \rightarrow \infty]{} \infty \quad \text{et} \quad \frac{w_n}{n} \xrightarrow[n \rightarrow \infty]{} 0.$$

Démonstration. L'affirmation du théorème est équivalente aux trois points suivants :

1.  $\alpha_n \rightarrow 0 \Leftrightarrow w_n \rightarrow \infty$ ,
2.  $\sum_{i=1}^{\infty} \alpha^*(i) < \infty$  et  $\frac{w_n}{n} \rightarrow 0 \Rightarrow \beta_n(r_1) \rightarrow 1$ ,
3.  $\beta_n(r_1) \rightarrow 1 \Rightarrow \frac{w_n}{n} \rightarrow 0$ .

Alors prouvons les trois implications.

1) Pour la v.a.  $\|U^*\|^2 = \sum_{j=0}^k \lambda_j^{*2} \chi_j^2$  on a  $P(\|U^*\|^2 > w_n) \leq \frac{1}{w_n} \sum_{j=0}^k \lambda_j^{*2}$  qui tend vers zéro si et seulement si  $w_n \rightarrow \infty$ .

D'après i) du Théorème 20 on a  $|\alpha_n - P(\|U^*\|^2 > w_n)| \leq \gamma_1 n^{-\frac{s(\theta-1)}{2(\theta+1)}}$  donc  $\alpha_n \rightarrow 0$  tend vers zéro si et seulement si  $w_n \rightarrow \infty$ .

Ce que nous avons également si on considère ii) du Théorème 20.

2) Selon  $H_1(k)$  il existe un  $j \in \{0, 1, \dots, k\}$ , soit  $j_0$ , tel que  $b_{j_0} = \mathbb{E}(Y_{e_{j_0}}(X)) \neq 0$ . Soient un entier  $m > 1$  et  $N(b_{j_0}, m)$  le plus petit entier tel que

$$\frac{w_n}{n} < \frac{b_{j_0}^2}{m^2} \text{ pour } n \geq N(b_{j_0}, m).$$

Comme

$$\|R_n\|^2 < w_n \rightarrow \sum_{j=0}^k \widehat{b}_{jn}^2 < \frac{w_n}{n} \rightarrow |\widehat{b}_{j_0 n}| < \sqrt{\frac{w_n}{n}},$$

pour  $n \geq N(b_{j_0}, m)$  on a  $|\widehat{b}_{j_0 n}| \leq \frac{|b_{j_0}|}{m}$ , d'où

$$|\widehat{b}_{j_0 n} - b_{j_0}| \geq |b_{j_0}| - |\widehat{b}_{j_0 n}| > |b_{j_0}| - \frac{|b_{j_0}|}{m} = |b_{j_0}| \frac{m-1}{m}.$$

Alors pour  $n \geq N(b_{j_0}, m)$

$$\begin{aligned} 1 - \beta_n(r_1) &= P(\|R_n\|^2 \leq w_n) \leq P\left(|\widehat{b}_{j_0 n} - b_{j_0}| \geq |b_{j_0}| \frac{m-1}{m}\right) \\ &\leq \frac{E\left(|\widehat{b}_{j_0 n} - b_{j_0}|^2\right)}{|b_{j_0}|^2 \left(\frac{m-1}{m}\right)^2} = \frac{\text{Var}\left(\widehat{b}_{j_0 n}\right)}{b_{j_0}^2} \frac{m^2}{(m-1)^2}. \end{aligned}$$

Majorons  $\text{Var}(\widehat{b}_{jn})$ , en posant  $R(i) = \text{Cov}(Y_1 e_{j_0}(X_1), Y_{i+1} e_{j_0}(X_{i+1}))$ .

$$\begin{aligned} \text{Var}(\widehat{b}_{jn}) &= \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n Y_i e_{j_0}(X_i)\right) = \frac{1}{n} \sum_{|i| \leq n-1} \left(1 - \frac{|i|}{n}\right) R(i) \\ &\leq \frac{1}{n} \left[ |R(0)| + 2 \sum_{i=1}^{n-1} |R(i)| \right] \\ &\leq \frac{1}{n} \left[ |R(0)| + 8M^2 \sum_{i=1}^{n-1} \alpha^*(i) \right] \end{aligned}$$

où dans la dernière majoration nous avons appliqué l'inégalité (voir [38])  $|\text{Cov}(\mathbf{Y}_s, \mathbf{Y}_{s+t})| \leq 4 \|\cdot\|_\infty \|\cdot\|_\infty \alpha(t)$  aux v.a.  $Y_i e_j(X_i)$  bornées par hypothèse.

Donc  $1 - \beta_n(\tau_1) \rightarrow 0$  pour  $n \rightarrow \infty$ .

- 3) On va supposer que  $\frac{m_n}{n} \rightarrow 0$ . Alors existent  $\eta > 0$  et une partie infinie  $\mathbb{N}_1 \subseteq \mathbb{N}$  tels que  $\frac{m_n}{n} > \eta$  pour  $n \in \mathbb{N}_1$ .

Considérons  $\nu \in H_1(k)$  tel que  $0 < \sum_{j=0}^k b_j^2 < \eta$ , dont l'existence est assurée d'après le Lemme 2.

La condition  $\beta_n(\tau_1) \rightarrow 1$  donne pour  $n \in \mathbb{N}_1$

$$P\left(\sum_{j=0}^k \widehat{b}_{jn}^2 < \eta\right) \leq P\left(\sum_{j=0}^k \widehat{b}_{jn}^2 < \frac{m_n}{n}\right) \xrightarrow{n \rightarrow \infty} 0. \quad (52)$$

D'autre part (voir la preuve du théorème 21)

$$\sum_{j=0}^k \widehat{b}_{jn}^2 \xrightarrow{p.s.} \sum_{j=0}^k b_j^2 < \eta$$

d'où

$$P\left(\sum_{j=0}^k \widehat{b}_{jn}^2 \leq \eta\right) \leq P\left(\sum_{j=0}^k (\widehat{b}_{jn}^2 - b_j^2) \leq \eta - \sum_{j=0}^k b_j^2\right) \xrightarrow{n \rightarrow \infty} 1$$

ce qui est contradictoire.



## V.4 SIMULATIONS

### V.4.1 INTRODUCTION

Des résultats de quelques simulations sont présentés pour l'hypothèse de normalité  $H_0 : X_t \sim N(0, 1)$  avec une taille d'échantillon  $n=100$  et un niveau asymptotique  $\alpha$ . Nous avons simulé à partir d'un processus autoregressif d'ordre 1,  $AR(1)$  par la relation :

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

où  $\epsilon_t$  est bruit gaussien avec :

$$V(\epsilon_t) = (1 - \rho^2)V(X_t)$$

et le paramètre d'autocorrélation  $\rho$  prend une valeur fixe précisée dans chaque cas.

### V.4.2 NIVEAU EMPIRIQUE

Dans ce paragraphe, nous présentons les résultats relatifs au niveau empirique, obtenu comme pourcentage de rejets sur le nombre de simulations. Nous avons simulé des données à partir d'une loi normale standard ou d'une exponentielle et nous avons testé  $H_0$  en utilisant soit la loi limite obtenue dans le cadre i.i.d. (voir  $T_C^{iid}$  et  $T_L^{iid}$ , pour la base du  $\chi^2$  et la base de Legendre), soit la loi limite dans le cadre  $\alpha$ -mélangeant ( $T_C$  et  $T_L$ ) pour les deux systèmes considérés et pour  $k$  variables ;  $k=2, \dots, 10$ .

#### V.4.2.1 $H_0 : X_t \sim N(0, 1)$

Dans ce cas, nous avons effectué  $m=1000$  simulations et les tableaux 5.1-5.4 donnent les pourcentages de rejet pour différentes valeurs du niveau nominal  $\alpha$  ( $\alpha=1\%$ ,  $3\%$ ,  $5\%$ ,  $10\%$ ). Pour chaque tableau la valeur de  $\rho$  est fixée.

Dans le but de voir ce qui se passe en utilisant les tests pour le cadre i.i.d. sur des données corrélées, nous avons effectué le test aussi avec la loi limite du cas i.i.d..

On peut observer que, pour une corrélation positive  $\rho = 0.5$  (voir le tableau 5.1),

L'hypothèse de normalité est rejetée très souvent par les tests  $T_L^{iid}$  et  $T_C^{iid}$ , alors que  $T_L$  et  $T_C$  se comportent mieux. Pour une corrélation faible (voir tableau 5.2), le test  $T_L$  semble être le meilleur en général, même si l'écart n'est pas très important. La corrélation négative montre un comportement différent pour les diverses valeurs de  $\alpha$  et, pour  $k > 7$ , les tests  $T_L$ ,  $T_C$ ,  $T_C^{iid}$  semblent presque équivalents par rapport au niveau empirique.

#### V.4.2.2 $H_0 : X_t \sim Exp(1)$

. Pour ce cas, nous avons effectué  $m=200$  simulations dans le cas d'une corrélation positive  $\rho=0.5$  (voir tableau 5.5) et  $\rho=0.2$  (voir tableau 5.6). On peut voir, ici aussi, qu'une corrélation positive de type AR(1) avec  $\rho=0.5$  amène les tests pour des observations indépendantes à rejeter l'hypothèse nulle très souvent, pour tout niveau  $\alpha$ . Lorsque la corrélation est modérée, la différence de comportement entre les tests de type "i.i.d." et les tests pour des observations corrélées diminue. On peut aussi dire que, en général,  $T_L$  présente un niveau empirique plus proche que  $T_C$  à un niveau  $\alpha$  fixé.

### V.4.3 PUISSANCE EMPIRIQUE

#### V.4.3.1 $H_0 : X_t \sim N(0, 1)$

. les tableaux suivants (tableau 5.7-5.12) donnent les puissances empiriques du test pour  $H_0 : X_t \sim N(0,1)$  contre  $H_1 : X_t \sim N(\mu, \sigma^2)$ , obtenues comme pourcentage de rejet sur  $m=1000$  simulations. Des échantillons de taille  $n=100$  sont tirés à partir de la loi sous  $H_1$ , qui est précisée dans chaque tableau. De plus, différentes valeurs de  $\rho$  et de  $\alpha$  sont utilisées.

Dans tous les tableaux, les valeurs en gras indiquent la puissance la plus grande lorsque  $k$  change, pour avoir tout de suite une indication sur le choix de  $k$ .

La première alternative considérée est  $H_1 : N(0, \sigma^2)$  avec  $\sigma^2 = 25/16$ , dont les résultats des simulations sont dans le tableau 5.7. On peut voir que, pour  $k$  fixé,  $T_L$  a une puissance empirique plus grande que  $T_C$  pratiquement toujours. La puissance de  $T_L$  est plus grande pour  $k=2$  et elle décroît lorsque  $k$  augmente.

Pour des corrélations positives, le tableau 5.8 montre les puissances empiriques pour l'alternative  $H_1 : N(0, \sigma^2)$  avec  $\sigma^2 = 0.5 = \sigma_0^2/2$  et  $\sigma^2 = 2 = 2\sigma_0^2/2$ ,  $\sigma_0^2 = 1$  indiquant la variance sous  $H_0$ .

Ici aussi  $T_L$  a une puissance plus grande que  $T_C$ , soit dans le cas d'une variance plus petite que 1, soit dans le cas d'une variance plus grande que 1. Relativement à  $k$ , nous remarquons que la puissance maximale de  $T_L$  est obtenue pour des  $k$  plus petites que celles de  $T_C$ .

Pour se faire une idée de la fonction puissance du test lorsque  $\sigma^2$  change (et donc l'alternative change), on peut voir les graphiques dans le tableau 5.9 et 5.10, où nous avons considéré des situations intermédiaires avec  $k=4$  et  $k=5$ . Le tableau 5.11 donne les résultats des simulations pour l'alternative  $H_1 : N(\mu,1)$  avec  $\mu=2$  et  $\sigma^2=0.5$ , pour différentes valeurs de  $\rho$  et  $\alpha$ . La valeur  $k=2$  donne, presque toujours, la puissance empirique maximale soit pour  $T_L$ , soit pour  $T_C$ .  $T_L$  se comporte mieux que  $T_C$ . Mais on peut observer que la différence entre les deux puissances n'est pas très importante.

Enfin, nous avons combiné les deux alternatives précédentes, en considérant  $H_1 : N(\mu,\sigma^2)$  avec  $\mu=0.5$  et  $\sigma^2=25/16$ . Le tableau 5.12 montre des puissances de  $T_L$  presque toujours supérieures à celles de  $T_C$ , mais ici aussi les deux tests sont presque équivalents.

On peut dire que pour une alternative avec le paramètre de location de la loi normale différent de zéro, le test  $T_C$  a un bon comportement; alors que, si l'alternative change pour le paramètre d'échelle,  $T_L$  se comporte mieux.

Dans le tableau 5.13 sont présentées les puissances empiriques obtenues avec 200 simulations des échantillons pour  $n=100$  à partir d'un processus  $AR(1)$  avec  $\rho=0.2$  où le bruit est contaminé. Plus précisément  $\epsilon_t \sim (1 - \delta)N(0, 1) + \delta N(0; 0.04)$ . Pour tout degré de contamination, il n'y a pas beaucoup de différence entre les puissances empiriques de  $T_L$  et  $T_C$ . On peut observer que, lorsque la contamination augmente, la puissance s'accroît aussi.

#### V.4.3.2 $H_0 : X_t \sim Exp(1)$

Les lois alternatives considérées dans ce cas sont obtenues en variant le paramètre  $\lambda$ . Pour chaque cas nous avons tiré  $m=200$  échantillons de taille  $n=100$ . Dans les tableaux 5.14-5.15, les puissances empiriques sont présentées aussi pour  $\lambda=1$  ( $H_0$ ) pour  $\rho$  fixé et donc on peut lire les résultats sur chaque colonne lorsque  $\lambda$  change. On ne peut pas dire dans ce cas, qu'il y a une supériorité nette de  $T_L$  sur  $T_C$  soit pour  $\rho=0.2$  soit pour  $\rho=0.5$ .

## V.5 APPENDICE

### V.5.1 APPENDICE1 : Théorèmes limites pour un processus $\alpha$ -mélangeant

#### Théorème 23

[Davydov, 1968] Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus centré, stationnaire,  $\alpha$ -mélangeant et à valeurs réelles et considérons la somme  $S_n = \sum_{i=1}^n X_i$ .

Si  $E(|X_1|^{2+\delta}) < \infty$  pour  $\delta > 0$  et

$$\sum_{n=1}^{\infty} \alpha(n)^{\frac{\delta}{2+\delta}} < \infty \quad (1)$$

alors la série  $\sum_{i=-\infty}^{+\infty} E(X_0 X_i)$  converge absolument avec somme  $\sigma^2 \geq 0$  et

$$n \text{Var} \left( \frac{S_n}{n} \right) \xrightarrow[n \rightarrow \infty]{} \sigma^2 \quad (2)$$

Démonstration. Voir le théorème 1.5 p.34 dans Bosq [21]. ■

Nous donnons ici dans un seul théorème les résultats pour un processus  $\alpha$ -mélangeant présentés dans l'article de Tikhomirov (1980) [74], en une version adaptée pour nous (similaire à celle présentée dans [37] p.48).

#### Théorème 24

[Tikhomirov, 1980] Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus centré, stationnaire,  $\alpha$ -mélangeant et à valeurs réelles. Posons

$$\Delta_n = \sup_{t \in \mathbb{R}} |P(n^{-1/2} S_n \leq t) - P(N_n \leq t)|$$

où  $N_n \sim \mathcal{N}(0, \sigma^2)$ .

Si  $E|X_1|^{2+\delta} < \infty$  pour  $\delta > 0$  et si  $\sigma^2 > 0$ , on a

- i) si  $\alpha(n) = O(n^{-\beta(2+\delta)(1+\delta)/\delta^2})$  pour  $\beta > 1$  alors il existe une constante  $\kappa_1$  telle que  $\Delta_n \leq \kappa_1 n^{-\frac{\beta(2+\delta)}{2(\beta+1)}}$  ;

ii) si  $\alpha(n) = O(e^{-\beta n})$  pour  $\beta > 1$ , alors il existe une constante  $\kappa_2$  telle que  $\Delta_n \leq \kappa_2 n^{-\frac{1}{2}} \log^{1+\delta}$ .

Démonstration. Voir les théorèmes 1 et 2 de l'article original [75]. ■

Remarque

Les vitesses de décroissance des coefficients de mélangeance dans le théorème ci dessus impliquent la condition (1) et donc on a

$$\sigma^2 = E(|X_1|^2) + 2 \sum_{i=2}^{\infty} E(X_1 X_i) < \infty.$$

Remarque

La somme  $\sigma^2$  peut être zéro, mais pour un processus stationnaire, centré,  $\alpha$  mélangeant avec  $E(X_1^2) < \infty$ , Bradley [28] donne une condition nécessaire et suffisante pour que  $\sigma^2 > 0$  : sous l'hypothèse  $\sum_{n=1}^{\infty} \frac{1}{n} |E(X_0 X_n)| < \infty$  la somme  $\sigma^2$  est finie et  $\sigma^2 > 0$  si et seulement si  $\text{Var}(S_n) \sim \frac{1}{n} \rightarrow \infty$ .

## V.5.2 APPENDICE2 :TABLEAUX :Niveau empirique- Puissance empirique

**TABLEAU 5.1** : Niveau empirique ; données simulées par un AR (1) gaussien avec  $\rho=0.5$

| k                      | $T_c$ | $T_c^{hd}$ | $T_L$ | $T_L^{hd}$ |
|------------------------|-------|------------|-------|------------|
| Niveau $\alpha = 1\%$  |       |            |       |            |
| 2                      | 4.6   | 12.8       | 7.3   | 32.7       |
| 3                      | 5.4   | 14.9       | 5.8   | 31.4       |
| 4                      | 4.6   | 16.0       | 4.9   | 33.8       |
| 5                      | 4.9   | 16.7       | 4.4   | 33.1       |
| 6                      | 4.3   | 16.7       | 4.1   | 34.1       |
| 7                      | 4.1   | 16.1       | 3.4   | 33.7       |
| 8                      | 3.3   | 16.3       | 2.7   | 35.0       |
| 9                      | 3.5   | 16.5       | 2.0   | 34.5       |
| 10                     | 3.4   | 15.6       | 2.7   | 33.6       |
| Niveau $\alpha = 3\%$  |       |            |       |            |
| 2                      | 10.1  | 21.3       | 15.3  | 43.1       |
| 3                      | 10.0  | 25.3       | 13.0  | 41.7       |
| 4                      | 11.0  | 24.5       | 12.9  | 44.9       |
| 5                      | 11.0  | 25.3       | 11.3  | 43.9       |
| 6                      | 10.4  | 24.9       | 9.8   | 44.1       |
| 7                      | 9.8   | 25.7       | 9.2   | 43.4       |
| 8                      | 9.9   | 24.9       | 9.3   | 43.3       |
| 9                      | 9.5   | 24.1       | 8.5   | 43.8       |
| 10                     | 9.0   | 23.9       | 7.9   | 43.1       |
| Niveau $\alpha = 5\%$  |       |            |       |            |
| 2                      | 14.0  | 27.4       | 21.9  | 49.0       |
| 3                      | 16.8  | 30.2       | 19.0  | 48.0       |
| 4                      | 15.2  | 30.1       | 19.3  | 50.9       |
| 5                      | 16.3  | 32.7       | 17.9  | 50.0       |
| 6                      | 15.9  | 30.6       | 17.5  | 49.3       |
| 7                      | 15.3  | 31.4       | 15.9  | 49.4       |
| 8                      | 15.6  | 30.6       | 14.6  | 50.0       |
| 9                      | 15.0  | 30.5       | 13.5  | 49.6       |
| 10                     | 14.6  | 31.9       | 13.3  | 48.3       |
| Niveau $\alpha = 10\%$ |       |            |       |            |
| 2                      | 24.2  | 35.2       | 33.0  | 58.3       |
| 3                      | 26.2  | 39.2       | 30.9  | 58.1       |
| 4                      | 25.2  | 38.9       | 33.0  | 59.1       |
| 5                      | 27.8  | 41.6       | 31.5  | 59.2       |
| 6                      | 27.1  | 41.4       | 31.0  | 58.9       |
| 7                      | 26.2  | 40.6       | 28.6  | 57.9       |
| 8                      | 26.7  | 40.8       | 28.0  | 57.6       |
| 9                      | 26.1  | 39.2       | 26.5  | 57.1       |
| 10                     | 26.9  | 40.1       | 25.6  | 58.0       |

**TABLEAU 5.2** : Niveau empirique : données simulées par un AR (1) gaussien avec  $\rho=0.2$

| k                      | $T_c$ | $T_c^{nd}$ | $T_L$ | $T_L^{nd}$ |
|------------------------|-------|------------|-------|------------|
| Niveau $\alpha = 1\%$  |       |            |       |            |
| 2                      | 2.7   | 3.4        | 3.5   | 4.2        |
| 3                      | 3.1   | 2.9        | 2.5   | 2.9        |
| 4                      | 2.3   | 2.5        | 1.9   | 3.3        |
| 5                      | 1.9   | 1.9        | 2.0   | 3.0        |
| 6                      | 1.5   | 2.0        | 1.5   | 2.9        |
| 7                      | 1.6   | 2.3        | 1.3   | 2.9        |
| 8                      | 1.5   | 1.5        | 0.9   | 3.0        |
| 9                      | 0.9   | 1.4        | 0.8   | 2.7        |
| 10                     | 1.1   | 1.5        | 0.8   | 3.2        |
| Niveau $\alpha = 3\%$  |       |            |       |            |
| 2                      | 8.3   | 7.4        | 6.2   | 7.5        |
| 3                      | 6.9   | 6.0        | 6.0   | 7.5        |
| 4                      | 5.7   | 6.3        | 5.3   | 7.7        |
| 5                      | 5.1   | 5.4        | 5.0   | 6.9        |
| 6                      | 4.8   | 5.4        | 4.1   | 6.2        |
| 7                      | 4.5   | 5.0        | 3.7   | 6.1        |
| 8                      | 4.6   | 4.4        | 3.5   | 5.9        |
| 9                      | 3.5   | 3.8        | 2.6   | 5.6        |
| 10                     | 3.1   | 4.8        | 2.1   | 6.0        |
| Niveau $\alpha = 5\%$  |       |            |       |            |
| 2                      | 11.5  | 9.9        | 9.0   | 10.5       |
| 3                      | 9.8   | 8.8        | 8.6   | 10.1       |
| 4                      | 8.8   | 9.2        | 8.5   | 11.2       |
| 5                      | 9.2   | 9.1        | 7.9   | 10.9       |
| 6                      | 7.8   | 9.2        | 7.8   | 10.1       |
| 7                      | 7.1   | 7.7        | 7.2   | 9.4        |
| 8                      | 6.9   | 7.9        | 6.8   | 9.3        |
| 9                      | 5.1   | 6.6        | 5.4   | 8.8        |
| 10                     | 5.9   | 7.4        | 5.5   | 8.3        |
| Niveau $\alpha = 10\%$ |       |            |       |            |
| 2                      | 17.6  | 16.5       | 15.9  | 16.0       |
| 3                      | 17.4  | 14.0       | 15.9  | 17.4       |
| 4                      | 16.7  | 14.8       | 17.2  | 16.4       |
| 5                      | 17.2  | 15.1       | 15.7  | 17.4       |
| 6                      | 16.5  | 15.7       | 16.2  | 18.2       |
| 7                      | 14.4  | 13.3       | 15.0  | 17.3       |
| 8                      | 15.4  | 14.2       | 14.3  | 16.6       |
| 9                      | 12.9  | 12.6       | 14.3  | 15.4       |
| 10                     | 13.5  | 12.6       | 13.1  | 16.0       |

**TABLEAU 5.3** : Niveau empirique ; données simulées par un AR (1) gaussien  
avec  $\rho = -0.2$

| k                      | $T_c$ | $T_c^{nd}$ | $T_l$ | $T_l^{nd}$ |
|------------------------|-------|------------|-------|------------|
| Niveau $\alpha = 1\%$  |       |            |       |            |
| 2                      | 3.5   | 0.4        | 3.6   | 1.0        |
| 3                      | 2.5   | 0.8        | 2.1   | 0.8        |
| 4                      | 1.6   | 0.6        | 1.9   | 0.8        |
| 5                      | 1.6   | 0.4        | 1.2   | 0.9        |
| 6                      | 1.6   | 1.0        | 1.1   | 1.0        |
| 7                      | 1.2   | 0.5        | 0.3   | 1.2        |
| 8                      | 1.0   | 0.7        | 0.2   | 1.1        |
| 9                      | 1.1   | 0.6        | 0.1   | 1.0        |
| 10                     | 0.5   | 0.6        | 0.2   | 0.8        |
| Niveau $\alpha = 3\%$  |       |            |       |            |
| 2                      | 7.1   | 2.0        | 6.8   | 2.7        |
| 3                      | 6.1   | 2.1        | 5.8   | 2.3        |
| 4                      | 6.2   | 1.6        | 5.0   | 2.5        |
| 5                      | 5.3   | 2.5        | 4.6   | 3.3        |
| 6                      | 5.9   | 2.8        | 3.2   | 3.5        |
| 7                      | 4.3   | 2.6        | 2.9   | 3.7        |
| 8                      | 2.9   | 2.5        | 2.5   | 3.6        |
| 9                      | 2.7   | 2.4        | 2.2   | 3.2        |
| 10                     | 3.1   | 2.2        | 1.3   | 3.3        |
| Niveau $\alpha = 5\%$  |       |            |       |            |
| 2                      | 10.4  | 3.5        | 9.8   | 4.0        |
| 3                      | 11.1  | 3.6        | 9.2   | 3.9        |
| 4                      | 9.1   | 4.1        | 8.8   | 3.7        |
| 5                      | 8.9   | 4.8        | 7.9   | 4.0        |
| 6                      | 9.2   | 4.4        | 7.1   | 5.2        |
| 7                      | 7.0   | 4.2        | 6.3   | 5.7        |
| 8                      | 5.9   | 4.1        | 5.5   | 6.0        |
| 9                      | 5.6   | 3.8        | 4.7   | 6.0        |
| 10                     | 4.9   | 3.8        | 4.4   | 5.2        |
| Niveau $\alpha = 10\%$ |       |            |       |            |
| 2                      | 18.0  | 7.5        | 16.9  | 7.4        |
| 3                      | 17.4  | 9.2        | 16.8  | 8.1        |
| 4                      | 16.9  | 8.4        | 16.8  | 8.8        |
| 5                      | 16.4  | 9.1        | 15.5  | 8.7        |
| 6                      | 16.1  | 9.5        | 15.4  | 9.0        |
| 7                      | 17.4  | 9.4        | 15.0  | 9.8        |
| 8                      | 14.3  | 9.0        | 14.3  | 10.3       |
| 9                      | 13.0  | 7.3        | 13.6  | 11.2       |
| 10                     | 13.0  | 7.4        | 12.3  | 10.7       |



**TABLEAU 5.4** : Niveau empirique : données simulées par un AR (1) gaussien avec  $\rho = -0.5$

| k                      | $T_c$ | $T_c^{ind}$ | $T_l$ | $T_l^{ind}$ |
|------------------------|-------|-------------|-------|-------------|
| Niveau $\alpha = 1\%$  |       |             |       |             |
| 2                      | 8.8   | 1.9         | 17.6  | 19.5        |
| 3                      | 8.4   | 3.9         | 14.6  | 16.8        |
| 4                      | 9.1   | 5.1         | 12.1  | 20.6        |
| 5                      | 7.5   | 4.6         | 9.9   | 19.9        |
| 6                      | 6.6   | 6.1         | 8.2   | 21.7        |
| 7                      | 6.6   | 5.9         | 7.0   | 20.8        |
| 8                      | 5.6   | 7.8         | 5.4   | 22.5        |
| 9                      | 5.9   | 7.2         | 5.0   | 21.6        |
| 10                     | 4.8   | 7.7         | 3.4   | 23.5        |
| Niveau $\alpha = 3\%$  |       |             |       |             |
| 2                      | 16.5  | 4.8         | 26.9  | 27.5        |
| 3                      | 17.8  | 7.9         | 24.1  | 25.6        |
| 4                      | 17.1  | 9.3         | 23.3  | 29.9        |
| 5                      | 15.6  | 10.3        | 20.5  | 28.2        |
| 6                      | 16.4  | 13.1        | 17.9  | 29.7        |
| 7                      | 14.4  | 13.7        | 15.9  | 29.5        |
| 8                      | 15.5  | 14.6        | 15.3  | 30.3        |
| 9                      | 14.4  | 13.3        | 13.4  | 30.7        |
| 10                     | 12.0  | 14.1        | 12.7  | 31.3        |
| Niveau $\alpha = 5\%$  |       |             |       |             |
| 2                      | 21.6  | 7.8         | 32.5  | 33.2        |
| 3                      | 22.1  | 11.6        | 31.6  | 30.0        |
| 4                      | 22.8  | 14.9        | 29.1  | 35.3        |
| 5                      | 21.7  | 15.9        | 28.2  | 33.3        |
| 6                      | 22.2  | 16.3        | 26.3  | 35.3        |
| 7                      | 21.1  | 19.3        | 22.9  | 34.1        |
| 8                      | 21.1  | 19.6        | 22.0  | 35.5        |
| 9                      | 19.0  | 18.2        | 20.9  | 35.5        |
| 10                     | 18.2  | 17.8        | 19.3  | 35.6        |
| Niveau $\alpha = 10\%$ |       |             |       |             |
| 2                      | 28.9  | 15.5        | 44.3  | 41.7        |
| 3                      | 32.0  | 20.0        | 42.9  | 39.3        |
| 4                      | 34.7  | 23.0        | 40.9  | 43.5        |
| 5                      | 32.6  | 24.7        | 39.3  | 42.4        |
| 6                      | 34.5  | 25.6        | 38.1  | 43.3        |
| 7                      | 34.3  | 25.2        | 36.8  | 43.5        |
| 8                      | 34.2  | 26.4        | 34.6  | 44.4        |
| 9                      | 33.3  | 27.2        | 34.1  | 43.0        |
| 10                     | 31.8  | 24.9        | 34.1  | 44.4        |

**TABLEAU 5.5 :** Niveau empirique ; données simulées par un AR (1) exponentielle  
avec  $\rho = 0.5$

| k                      | $T_C$ | $T_C^{iid}$ | $T_L$ | $T_L^{iid}$ |
|------------------------|-------|-------------|-------|-------------|
| Niveau $\alpha = 1\%$  |       |             |       |             |
| 2                      | 7.5   | 14.5        | 7.0   | 21.5        |
| 3                      | 4.5   | 15.0        | 6.0   | 21.5        |
| 4                      | 3.5   | 16.5        | 5.5   | 22.0        |
| 5                      | 4.0   | 17.5        | 4.5   | 22.5        |
| 6                      | 4.0   | 15.0        | 3.0   | 20.0        |
| 7                      | 3.5   | 15.0        | 2.5   | 21.0        |
| 8                      | 2.0   | 13.0        | 2.5   | 20.0        |
| 9                      | 1.5   | 12.0        | 2.5   | 17.0        |
| 10                     | 1.5   | 11.5        | 2.5   | 16.5        |
| Niveau $\alpha = 3\%$  |       |             |       |             |
| 2                      | 12.5  | 24.5        | 13.0  | 29.0        |
| 3                      | 9.5   | 29.0        | 10.5  | 32.0        |
| 4                      | 9.5   | 29.5        | 9.0   | 33.5        |
| 5                      | 6.0   | 26.5        | 8.0   | 33.0        |
| 6                      | 6.5   | 25.5        | 5.0   | 33.5        |
| 7                      | 6.5   | 25.0        | 4.5   | 32.0        |
| 8                      | 5.5   | 23.0        | 5.0   | 30.5        |
| 9                      | 4.5   | 22.0        | 5.5   | 30.0        |
| 10                     | 5.5   | 22.5        | 5.0   | 28.5        |
| Niveau $\alpha = 5\%$  |       |             |       |             |
| 2                      | 14.5  | 32.0        | 16.0  | 33.0        |
| 3                      | 12.0  | 33.5        | 13.5  | 36.5        |
| 4                      | 11.0  | 34.0        | 12.5  | 38.0        |
| 5                      | 11.5  | 36.0        | 11.0  | 39.5        |
| 6                      | 9.0   | 36.0        | 10.0  | 41.0        |
| 7                      | 8.0   | 34.5        | 9.0   | 39.0        |
| 8                      | 7.5   | 31.0        | 7.5   | 38.5        |
| 9                      | 7.0   | 30.5        | 7.5   | 36.5        |
| 10                     | 6.0   | 27.5        | 8.5   | 36.5        |
| Niveau $\alpha = 10\%$ |       |             |       |             |
| 2                      | 21.5  | 38.5        | 23.0  | 41.5        |
| 3                      | 22.5  | 44.5        | 22.0  | 46.5        |
| 4                      | 19.5  | 43.0        | 22.5  | 47.0        |
| 5                      | 22.0  | 45.5        | 18.5  | 52.5        |
| 6                      | 15.0  | 46.0        | 17.5  | 49.0        |
| 7                      | 18.0  | 44.0        | 16.5  | 49.0        |
| 8                      | 15.0  | 47.5        | 16.0  | 48.0        |
| 9                      | 14.0  | 40.0        | 17.5  | 47.5        |
| 10                     | 13.5  | 38.5        | 16.5  | 45.0        |

**TABLEAU 5.6** : Niveau empirique ; données simulées par un AR (1) exponentielle avec  $\rho = 0.2$

| k                      | $T_c$ | $T_c^{hd}$ | $T_L$ | $T_L^{hd}$ |
|------------------------|-------|------------|-------|------------|
| Niveau $\alpha = 1\%$  |       |            |       |            |
| 2                      | 3.5   | 4.0        | 2.0   | 4.5        |
| 3                      | 2.0   | 2.5        | 2.0   | 3.5        |
| 4                      | 1.5   | 2.0        | 3.5   | 3.0        |
| 5                      | 1.5   | 2.5        | 2.5   | 4.0        |
| 6                      | 0.5   | 2.0        | 2.0   | 4.5        |
| 7                      | 3.0   | 2.5        | 2.0   | 4.5        |
| 8                      | 1.5   | 2.0        | 2.0   | 3.5        |
| 9                      | 0.5   | 2.5        | 1.0   | 3.0        |
| 10                     | 1.5   | 2.0        | 1.5   | 2.5        |
| Niveau $\alpha = 3\%$  |       |            |       |            |
| 2                      | 8.0   | 9.5        | 7.0   | 10.0       |
| 3                      | 5.0   | 6.5        | 5.5   | 8.5        |
| 4                      | 4.5   | 5.5        | 5.5   | 8.0        |
| 5                      | 5.5   | 6.5        | 5.0   | 7.5        |
| 6                      | 3.5   | 6.0        | 4.5   | 7.0        |
| 7                      | 4.0   | 5.5        | 3.5   | 6.5        |
| 8                      | 3.0   | 4.5        | 3.5   | 6.0        |
| 9                      | 3.5   | 5.5        | 3.0   | 6.0        |
| 10                     | 3.5   | 5.0        | 3.5   | 5.5        |
| Niveau $\alpha = 5\%$  |       |            |       |            |
| 2                      | 9.5   | 11.0       | 9.5   | 12.5       |
| 3                      | 8.0   | 11.0       | 9.0   | 12.0       |
| 4                      | 6.0   | 7.5        | 9.0   | 12.0       |
| 5                      | 8.0   | 7.5        | 8.0   | 9.5        |
| 6                      | 7.0   | 10.0       | 8.0   | 10.0       |
| 7                      | 6.5   | 6.5        | 6.5   | 8.5        |
| 8                      | 6.5   | 8.5        | 5.5   | 7.5        |
| 9                      | 6.0   | 8.5        | 4.0   | 6.5        |
| 10                     | 5.5   | 7.0        | 5.0   | 6.5        |
| Niveau $\alpha = 10\%$ |       |            |       |            |
| 2                      | 15.0  | 14.0       | 19.0  | 24.0       |
| 3                      | 18.0  | 18.5       | 15.0  | 17.5       |
| 4                      | 13.5  | 13.0       | 15.0  | 17.5       |
| 5                      | 14.0  | 12.5       | 14.0  | 16.0       |
| 6                      | 16.0  | 15.5       | 13.5  | 15.0       |
| 7                      | 14.5  | 16.0       | 14.0  | 15.0       |
| 8                      | 13.0  | 16.0       | 14.0  | 13.5       |
| 9                      | 14.5  | 17.5       | 11.5  | 13.5       |
| 10                     | 15.0  | 13.0       | 13.5  | 13.0       |

**TABLEAU 5.7** : Puissances empiriques pour  $H_0 : N(0,1)$  contre  $H_1 : N\left(0, \frac{25}{16}\right)$

| $\rho$ | k    | $\alpha = 1\%$ |             | $\alpha = 3\%$ |             | $\alpha = 5\%$ |             | $\alpha = 10\%$ |             |
|--------|------|----------------|-------------|----------------|-------------|----------------|-------------|-----------------|-------------|
|        |      | $T_c$          | $T_L$       | $T_c$          | $T_L$       | $T_c$          | $T_L$       | $T_c$           | $T_L$       |
| -0.5   | 2    | 19.8           | <b>41.6</b> | 27.5           | <b>54.7</b> | 34.8           | <b>61.4</b> | 45.8            | <b>73.1</b> |
|        | 3    | 21.4           | 34.8        | 32.9           | 49.6        | 39.2           | 56.5        | 49.8            | 68.3        |
|        | 4    | <b>22.6</b>    | 31.4        | 34.5           | 48.0        | 42.7           | 56.4        | 54.6            | 68.6        |
|        | 5    | 21.1           | 26.8        | <b>36.1</b>    | 42.8        | <b>43.3</b>    | 53.2        | 56.4            | 66.3        |
|        | 6    | 19.5           | 22.3        | 32.9           | 40.0        | 42.2           | 49.7        | <b>56.7</b>     | 66.2        |
|        | 7    | <b>18.1</b>    | 20.4        | 33.9           | 36.9        | 40.7           | 46.9        | 55.8            | 63.7        |
|        | 8    | 17.8           | 17.0        | 32.7           | 34.4        | 41.3           | 44.9        | 56.1            | 60.5        |
|        | 9    | 17.3           | 14.6        | 30.8           | 31.9        | 40.1           | 42.6        | 56.5            | 58.2        |
|        | 10   | 15.1           | 13.5        | 29.4           | 38.3        | 39.3           | 39.2        | 54.6            | 58.8        |
|        | -0.2 | 2              | 16.9        | <b>47.8</b>    | 27.6        | <b>62.3</b>    | 33.4        | <b>69.5</b>     | 44.5        |
| 3      |      | 20.6           | 37.3        | 32.9           | 53.9        | 39.2           | 61.7        | 52.4            | 72.7        |
| 4      |      | <b>20.8</b>    | 38.5        | 34.4           | 54.1        | 42.8           | 63.6        | 56.4            | 77.2        |
| 5      |      | 20.1           | 32.1        | 34.6           | 47.9        | 43.6           | 58.6        | 57.8            | 73.1        |
| 6      |      | 20.1           | 29.0        | <b>36.1</b>    | 46.1        | 43.4           | 56.5        | 57.9            | 72.6        |
| 7      |      | 27.5           | 23.5        | 33.0           | 41.1        | 42.6           | 51.6        | 57.9            | 69.9        |
| 8      |      | 16.5           | 21.6        | 33.8           | 40.1        | <b>44.0</b>    | 50.6        | <b>59.3</b>     | 68.5        |
| 9      |      | 16.5           | 17.1        | 31.1           | 35.0        | 41.0           | 57.7        | 56.7            | 65.5        |
| 10     |      | 15.3           | 15.4        | 30.2           | 32.8        | 40.7           | 46.0        | 56.8            | 64.3        |
| 0.2    |      | 2              | 9.5         | <b>32.5</b>    | 18.2        | <b>50.4</b>    | 25.7        | <b>59.7</b>     | 37.3        |
|        | 3    | 12.8           | 31.9        | 21.7           | 38.2        | 31.8           | 48.0        | 45.0            | 64.9        |
|        | 4    | 13.1           | 37.7        | 26.1           | 45.0        | 33.6           | 56.0        | 46.5            | 72.2        |
|        | 5    | <b>13.7</b>    | 22.5        | 24.6           | 38.5        | 32.8           | 50.8        | 48.8            | 66.5        |
|        | 6    | 12.9           | 22.0        | 24.0           | 42.0        | 32.3           | 52.7        | 52.7            | 68.8        |
|        | 7    | 12.6           | 17.7        | 25.5           | 36.9        | 34.9           | 46.7        | 51.1            | 64.7        |
|        | 8    | 12.9           | 17.2        | <b>26.4</b>    | 35.1        | 35.3           | 46.0        | 50.7            | 65.1        |
|        | 9    | 10.9           | 14.6        | 25.8           | 31.4        | <b>35.7</b>    | 42.8        | <b>53.4</b>     | 62.3        |
|        | 10   | 10.5           | 13.6        | 22.3           | 30.0        | 35.1           | 41.6        | 51.8            | 61.0        |
|        | 0.5  | 2              | 7.0         | <b>17.2</b>    | 13.0        | <b>31.4</b>    | 17.9        | <b>40.5</b>     | 29.9        |
| 3      |      | 7.7            | 11.5        | 15.2           | 22.9        | 22.9           | 31.6        | 35.9            | 48.7        |
| 4      |      | 8.4            | 13.5        | 17.9           | 27.6        | 24.6           | 37.3        | 38.2            | 55.6        |
| 5      |      | 8.6            | 14.2        | <b>18.9</b>    | 24.1        | 25.3           | 32.6        | 41.0            | 51.4        |
| 6      |      | 8.5            | 11.1        | 18.4           | 24.7        | 25.4           | 33.9        | 43.5            | 53.7        |
| 7      |      | <b>9.0</b>     | 9.0         | <b>18.9</b>    | 21.9        | 26.7           | 31.2        | 43.1            | 49.5        |
| 8      |      | 8.1            | <b>8.5</b>  | <b>18.6</b>    | 21.2        | <b>26.8</b>    | 30.8        | 43.7            | 49.4        |
| 9      |      | 7.3            | 7.6         | 17.4           | 18.3        | 25.0           | 29.0        | 43.4            | 46.4        |
| 10     |      | 7.3            | 7.3         | 17.0           | 17.9        | 26.1           | 28.9        | <b>44.2</b>     | 45.4        |

**TABIEAU 5.8** : Puissances empiriques pour  $H_0 : N(0,1)$  contre  $H_1 : N(0,0.5)$   
et  $H_1 : N(0,2)$

| $\rho$           | k   | $\alpha = 1\%$ |             | $\alpha = 3\%$ |             | $\alpha = 5\%$ |             | $\alpha = 10\%$ |             |
|------------------|-----|----------------|-------------|----------------|-------------|----------------|-------------|-----------------|-------------|
|                  |     | $T_c$          | $T_L$       | $T_c$          | $T_L$       | $T_c$          | $T_L$       | $T_c$           | $T_L$       |
| $\sigma^2 = 0.5$ |     |                |             |                |             |                |             |                 |             |
| 0.2              | 2   | 43.9           | <b>92.5</b> | 58.0           | <b>96.6</b> | 65.5           | <b>97.5</b> | 76.3            | <b>99.3</b> |
|                  | 3   | 55.4           | 87.8        | 73.1           | 94.4        | 61.0           | 96.8        | 87.9            | 98.6        |
|                  | 4   | <b>56.5</b>    | 83.5        | <b>75.4</b>    | 93.4        | 81.7           | 96.0        | 89.9            | 98.6        |
|                  | 5   | 52.0           | 79.5        | 74.5           | 91.1        | <b>82.8</b>    | 94.8        | 91.3            | 98.1        |
|                  | 6   | 49.4           | 74.2        | 72.9           | 89.7        | 82.2           | 93.2        | <b>91.6</b>     | 98.0        |
|                  | 7   | 44.1           | 68.3        | 68.6           | 86.6        | 79.2           | 91.8        | 90.9            | 97.3        |
|                  | 8   | 41.0           | 62.5        | 66.0           | 82.9        | 78.3           | 90.3        | 91.3            | 96.7        |
|                  | 9   | 36.4           | 56.9        | 62.3           | 79.4        | 76.1           | 87.7        | 88.5            | 95.5        |
|                  | 10  | 31.3           | 52.4        | 58.8           | 75.3        | 70.9           | 85.4        | 87.4            | 95.3        |
|                  | 0.5 | 2              | 32.2        | <b>74.4</b>    | 45.9        | 66.8           | 54.8        | <b>88.5</b>     | 67.1        |
| 3                |     | 40.1           | 62.6        | 56.1           | 78.4        | 66.3           | 85.3        | 79.2            | 91.4        |
| 4                |     | 41.1           | 58.3        | <b>59.6</b>    | 80.1        | 70.0           | 84.8        | 81.4            | 91.5        |
| 5                |     | <b>41.4</b>    | 55.7        | 59.1           | <b>82.8</b> | <b>70.4</b>    | 82.8        | <b>83.6</b>     | 90.5        |
| 6                |     | 38.8           | 51.4        | 58.1           | <b>82.8</b> | 69.3           | 81.3        | 83.5            | 89.5        |
| 7                |     | 35.9           | 49.1        | 56.8           | 81.0        | 67.4           | 79.5        | 81.7            | 89.3        |
| 8                |     | 33.2           | 45.6        | 55.1           | 80.3        | 68.2           | 76.9        | <b>83.7</b>     | 88.1        |
| 9                |     | 29.1           | 42.7        | 53.2           | 77.9        | 67.2           | 75.3        | 82.2            | 86.7        |
| 10               |     | 25.4           | 39.5        | 50.4           | 76.7        | 63.2           | 73.9        | 79.6            | 86.2        |
| $\sigma^2 = 2$   |     |                |             |                |             |                |             |                 |             |
| 0.2              | 2   | 20.2           | <b>73.3</b> | 32.9           | <b>87.6</b> | 43.6           | <b>91.8</b> | 59.7            | 96.4        |
|                  | 3   | 29.9           | 57.5        | 48.8           | 75.7        | 58.7           | 84.4        | 74.5            | 93.5        |
|                  | 4   | 35.9           | 70.8        | 55.0           | 87.1        | 65.4           | <b>91.9</b> | 78.8            | <b>96.8</b> |
|                  | 5   | 41.0           | 59.2        | 60.4           | 79.5        | 70.8           | 87.6        | 83.1            | 95.7        |
|                  | 6   | 41.7           | <b>64.6</b> | 62.4           | 82.1        | 72.2           | 90.3        | 82.4            | 96.7        |
|                  | 7   | 42.2           | 55.8        | 61.8           | 75.8        | 73.4           | 85.2        | 85.3            | 95.2        |
|                  | 8   | <b>42.9</b>    | 55.6        | <b>63.9</b>    | 77.7        | <b>73.7</b>    | 87.1        | 85.6            | 95.9        |
|                  | 9   | 40.8           | 47.3        | 62.2           | 74.3        | 72.9           | 82.5        | 86.4            | 93.6        |
|                  | 10  | 41.4           | <b>48.8</b> | 63.2           | 74.1        | 73.3           | 83.6        | <b>86.9</b>     | 93.7        |
|                  | 0.5 | 2              | 13.0        | <b>47.1</b>    | 23.7        | <b>67.4</b>    | 31.1        | <b>77.5</b>     | 45.7        |
| 3                |     | 18.6           | 28.3        | 31.6           | 50.4        | 42.2           | 63.4        | 60.3            | 81.2        |
| 4                |     | 21.3           | 40.7        | 37.5           | 66.8        | 49.5           | 77.2        | 65.5            | <b>89.5</b> |
| 5                |     | 23.0           | 30.5        | 40.7           | 55.3        | 51.7           | 69.7        | 69.4            | 85.3        |
| 6                |     | 22.9           | 33.7        | 41.8           | 61.0        | 54.8           | 73.7        | 72.4            | 87.4        |
| 7                |     | 24.1           | 27.2        | 44.9           | 52.7        | 57.1           | 66.4        | 74.3            | 83.7        |
| 8                |     | <b>24.5</b>    | 27.9        | 44.5           | 53.9        | 58.2           | 66.1        | 76.2            | 84.0        |
| 9                |     | 23.8           | 23.0        | <b>45.5</b>    | 48.7        | <b>58.8</b>    | 63.4        | 76.2            | 82.6        |
| 10               |     | 22.6           | 23.7        | 45.3           | 50.0        | 58.5           | 64.0        | <b>77.8</b>     | 82.1        |

**TABLEAU 5.9** : Puissances empiriques ( $\rho = 0.2, \alpha = 5\%$ ) pour  $H_0 : N(0,1)$  contre  $H_1 : N(0, \sigma^2)$

| $\sigma^2$ | (k = 4) |       | (k = 5) |       |
|------------|---------|-------|---------|-------|
|            | $T_c$   | $T_L$ | $T_c$   | $T_L$ |
| 0.5        | 81.7    | 96    | 82.8    | 94.8  |
| 1          | 8.8     | 8.5   | 9.2     | 7.9   |
| 25/16      | 33.6    | 56    | 32.8    | 50.8  |
| 2          | 65.4    | 91.9  | 70.8    | 87.6  |

**TABLEAU 5.10** : Puissances empiriques ( $\rho = 0.5, \alpha = 5\%$ ) pour  $H_0 : N(0,1)$  contre  $H_1 : N(0, \sigma^2)$

| $\sigma^2$ | (k = 4) |       | (k = 5) |       |
|------------|---------|-------|---------|-------|
|            | $T_c$   | $T_L$ | $T_c$   | $T_L$ |
| 0.5        | 70      | 84.8  | 70.4    | 82.8  |
| 1          | 15.2    | 19.3  | 16.3    | 17.9  |
| 25/16      | 24.6    | 37.3  | 25.3    | 32.6  |
| 2          | 49.5    | 77.2  | 51.7    | 69.7  |

**TABLEAU 5.11** : Puissances empiriques pour  $H_0: N(0,1)$  contre  $H_1: N(0.5,1)$

| $\rho$ | k    | $\alpha = 1\%$ |       | $\alpha = 3\%$ |       | $\alpha = 5\%$ |       | $\alpha = 10\%$ |       |
|--------|------|----------------|-------|----------------|-------|----------------|-------|-----------------|-------|
|        |      | $T_c$          | $T_j$ | $T_c$          | $T_L$ | $T_c$          | $T_L$ | $T_c$           | $T_L$ |
| -0.5   | 2    | 99.3           | 99.8  | 99.9           | 99.9  | 100.0          | 100.0 | 100.0           | 100.0 |
|        | 3    | 98.1           | 98.1  | 99.7           | 99.7  | 100.0          | 99.9  | 100.0           | 100.0 |
|        | 4    | 96.8           | 94.4  | 99.5           | 99.4  | 99.9           | 99.8  | 100.0           | 100.0 |
|        | 5    | 95.1           | 89.6  | 99.4           | 98.3  | 99.9           | 99.6  | 100.0           | 100.0 |
|        | 6    | 92.4           | 84.7  | 98.9           | 97.0  | 99.7           | 99.4  | 99.9            | 100.0 |
|        | 7    | 88.2           | 80.4  | 97.9           | 95.7  | 99.1           | 99.2  | 100.0           | 100.0 |
|        | 8    | 84.0           | 73.3  | 97.2           | 92.9  | 98.7           | 98.0  | 99.8            | 99.8  |
|        | 9    | 80.6           | 67.9  | 93.8           | 90.3  | 97.9           | 96.7  | 99.7            | 99.6  |
|        | 10   | 77.0           | 62.4  | 93.7           | 88.0  | 97.8           | 95.5  | 99.6            | 99.4  |
|        | -0.2 | 2              | 97.1  | 98.6           | 98.9  | 99.6           | 99.3  | 99.8            | 99.8  |
| 3      |      | 95.2           | 97.3  | 99.2           | 99.4  | 99.6           | 99.7  | 99.2            | 99.9  |
| 4      |      | 95.3           | 95.5  | 99.1           | 98.9  | 99.7           | 99.5  | 100.0           | 99.9  |
| 5      |      | 92.6           | 91.7  | 98.1           | 98.3  | 99.0           | 99.3  | 99.7            | 99.9  |
| 6      |      | 89.4           | 87.3  | 97.4           | 97.2  | 98.8           | 98.9  | 99.9            | 99.8  |
| 7      |      | 85.6           | 82.5  | 96.0           | 95.9  | 98.3           | 98.4  | 99.6            | 99.6  |
| 8      |      | 82.2           | 78.4  | 93.3           | 93.3  | 96.9           | 97.4  | 99.4            | 99.1  |
| 9      |      | 80.8           | 73.5  | 93.9           | 91.2  | 97.0           | 95.6  | 99.6            | 99.1  |
| 10     |      | 77.7           | 67.3  | 92.8           | 88.5  | 95.9           | 93.2  | 99.1            | 98.9  |
| 0.2    |      | 2              | 85.4  | 89.3           | 93.2  | 95.6           | 95.9  | 97.2            | 97.9  |
|        | 3    | 85.4           | 84.8  | 93.2           | 94.3  | 95.7           | 96.5  | 98.0            | 98.3  |
|        | 4    | 82.6           | 80.4  | 91.2           | 91.5  | 94.8           | 95.0  | 98.1            | 97.8  |
|        | 5    | 80.1           | 76.3  | 90.2           | 89.3  | 94.7           | 93.3  | 97.2            | 97.3  |
|        | 6    | 75.7           | 71.2  | 88.7           | 86.5  | 93.1           | 92.3  | 96.5            | 96.7  |
|        | 7    | 73.4           | 65.9  | 84.8           | 83.2  | 92.1           | 90.0  | 96.8            | 96.3  |
|        | 8    | 69.2           | 71.6  | 86.3           | 81.3  | 91.1           | 88.3  | 95.8            | 95.8  |
|        | 9    | 65.9           | 57.1  | 83.4           | 77.0  | 89.1           | 86.7  | 95.5            | 95.5  |
|        | 10   | 64.8           | 52.8  | 81.2           | 75.5  | 87.1           | 84.6  | 94.4            | 94.1  |
|        | 0.5  | 2              | 65.0  | 63.4           | 76.1  | 77.1           | 81.8  | 82.4            | 88.1  |
| 3      |      | 61.6           | 59.0  | 74.7           | 74.4  | 77.7           | 80.4  | 88.7            | 88.6  |
| 4      |      | 60.0           | 53.1  | 73.2           | 70.1  | 79.8           | 77.5  | 86.8            | 87.2  |
| 5      |      | 55.7           | 49.0  | 72.0           | 67.5  | 78.3           | 74.5  | 87.4            | 86.8  |
| 6      |      | 52.3           | 44.9  | 70.0           | 64.2  | 77.2           | 73.6  | 86.4            | 85.6  |
| 7      |      | 50.4           | 41.4  | 67.0           | 61.5  | 75.9           | 71.9  | 85.7            | 84.1  |
| 8      |      | 48.3           | 37.5  | 66.8           | 58.6  | 75.7           | 68.9  | 84.5            | 83.1  |
| 9      |      | 46.2           | 35.1  | 65.2           | 56.4  | 73.0           | 67.2  | 83.4            | 81.5  |
| 10     |      | 44.2           | 31.4  | 64.0           | 54.4  | 71.1           | 64.2  | 82.2            | 80.9  |

**TABLEAU 5.12** : Puissances empiriques pour  $H_0 : N(0,1)$  contre  $H_1 : N(0.5, \frac{25}{16})$

| $\rho$ | k  | $\alpha = 1\%$ |             | $\alpha = 3\%$ |             | $\alpha = 5\%$ |             | $\alpha = 10\%$ |             |
|--------|----|----------------|-------------|----------------|-------------|----------------|-------------|-----------------|-------------|
|        |    | $T_C$          | $T_L$       | $T_C$          | $T_L$       | $T_C$          | $T_L$       | $T_C$           | $T_L$       |
| 0.2    | 2  | 76.1           | <b>91.8</b> | 86.6           | <b>96.2</b> | 90.8           | <b>98.0</b> | 96.0            | <b>99.2</b> |
|        | 3  | 78.8           | <b>88.0</b> | <b>90.0</b>    | 95.3        | 93.8           | 96.9        | 96.4            | 99.0        |
|        | 4  | 79.5           | <b>88.3</b> | <b>90.0</b>    | 95.0        | 94.1           | 96.7        | <b>97.6</b>     | 99.0        |
|        | 5  | 79.2           | <b>83.9</b> | 89.7           | 93.9        | <b>94.4</b>    | 95.9        | 97.3            | 98.6        |
|        | 6  | 77.1           | <b>83.0</b> | 89.7           | 93.2        | <b>94.4</b>    | 96.1        | 97.4            | 98.2        |
|        | 7  | 77.8           | 79.0        | 89.6           | 92.9        | 93.1           | 95.4        | 97.1            | 97.8        |
|        | 8  | 76.6           | 77.1        | <b>88.4</b>    | 91.2        | 92.9           | 95.4        | 96.9            | 97.6        |
|        | 9  | 75.8           | 72.2        | 88.5           | 89.7        | 92.7           | 93.9        | 96.9            | 97.4        |
|        | 10 | 75.0           | 69.0        | 97.7           | 87.7        | 91.9           | 92.6        | 97.2            | 97.0        |



**TABLEAU 5.13** : Puissances empiriques pour  $H_0 : N(0,1)$  contre  $H_1$  tel que  $\varepsilon_t \sim (1 - \delta)N(0,1) + \delta N(0,0.04)$  avec  $\rho = 0.2$

| $\delta$ | k   | $\alpha = 1\%$ |       | $\alpha = 3\%$ |       | $\alpha = 5\%$ |       | $\alpha = 10\%$ |       |
|----------|-----|----------------|-------|----------------|-------|----------------|-------|-----------------|-------|
|          |     | $T_c$          | $T_L$ | $T_c$          | $T_L$ | $T_c$          | $T_L$ | $T_c$           | $T_L$ |
| 0.1      | 2   | 1.5            | 3.5   | 6.0            | 5.5   | 11.5           | 11.0  | 20.0            | 21.5  |
|          | 3   | 1.0            | 1.5   | 4.0            | 5.0   | 9.5            | 9.5   | 20.0            | 19.0  |
|          | 4   | 1.0            | 0.5   | 4.5            | 6.5   | 9.5            | 12.0  | 21.0            | 18.5  |
|          | 5   | 0.0            | 0.5   | 5.5            | 6.0   | 9.0            | 9.5   | 19.5            | 19.5  |
|          | 6   | 1.5            | 0.5   | 3.0            | 4.0   | 7.5            | 7.5   | 13.5            | 16.0  |
|          | 7   | 1.0            | 0.0   | 4.0            | 2.0   | 6.5            | 6.5   | 14.0            | 14.0  |
|          | 8   | 1.0            | 0.5   | 3.5            | 3.5   | 5.5            | 5.5   | 13.5            | 14.0  |
|          | 9   | 0.0            | 0.5   | 2.5            | 3.3   | 5.5            | 6.0   | 14.5            | 13.0  |
|          | 10  | 0.0            | 0.5   | 2.0            | 2.5   | 4.5            | 6.0   | 13.5            | 12.0  |
|          | 0.2 | 2              | 1.0   | 4.5            | 5.5   | 8.5            | 11.0  | 13.0            | 28.5  |
| 3        |     | 3.0            | 4.5   | 8.5            | 7.5   | 14.0           | 11.5  | 25.0            | 25.0  |
| 4        |     | 2.0            | 1.5   | 7.5            | 7.5   | 13.0           | 14.0  | 25.5            | 30.5  |
| 5        |     | 1.0            | 2.5   | 5.0            | 6.0   | 10.0           | 13.5  | 21.5            | 30.0  |
| 6        |     | 1.0            | 1.0   | 5.5            | 7.0   | 12.5           | 9.0   | 24.5            | 26.5  |
| 7        |     | 1.0            | 1.0   | 6.0            | 5.0   | 9.5            | 9.0   | 24.0            | 27.0  |
| 8        |     | 0.5            | 0.5   | 3.5            | 4.5   | 8.5            | 8.5   | 23.0            | 24.0  |
| 9        |     | 0.5            | 0.0   | 4.5            | 3.5   | 8.0            | 7.5   | 22.0            | 21.0  |
| 10       |     | 0.0            | 0.0   | 2.5            | 3.0   | 6.0            | 6.5   | 21.0            | 19.5  |
| 0.3      |     | 2              | 4.0   | 6.0            | 19.0  | 18.5           | 30.5  | 27.5            | 53.5  |
|          | 3   | 4.5            | 6.0   | 18.5           | 16.5  | 27.5           | 26.0  | 45.5            | 43.0  |
|          | 4   | 3.5            | 4.5   | 14.0           | 19.5  | 26.0           | 29.0  | 46.0            | 54.5  |
|          | 5   | 2.0            | 2.5   | 15.5           | 18.5  | 25.5           | 27.5  | 49.0            | 51.5  |
|          | 6   | 2.5            | 3.5   | 12.0           | 17.0  | 24.0           | 29.0  | 48.5            | 49.5  |
|          | 7   | 2.5            | 3.5   | 14.0           | 14.5  | 29.5           | 27.5  | 46.0            | 48.0  |
|          | 8   | 3.0            | 2.5   | 10.0           | 11.0  | 21.0           | 24.5  | 46.5            | 46.5  |
|          | 9   | 2.0            | 1.0   | 14.5           | 11.0  | 22.5           | 25.0  | 42.5            | 44.5  |
|          | 10  | 2.0            | 1.0   | 8.0            | 10.0  | 21.0           | 26.0  | 45.0            | 44.0  |
|          | 0.4 | 2              | 19.5  | 24.5           | 48.5  | 42.0           | 63.0  | 52.5            | 88.0  |
| 3        |     | 20.0           | 24.0  | 37.0           | 40.0  | 47.0           | 51.5  | 69.5            | 69.5  |
| 4        |     | 15.0           | 21.5  | 41.0           | 45.5  | 59.0           | 60.0  | 87.5            | 81.0  |
| 5        |     | 14.5           | 21.5  | 43.0           | 44.5  | 61.5           | 57.0  | 82.0            | 81.0  |
| 6        |     | 15.0           | 18.5  | 39.0           | 41.5  | 57.0           | 61.0  | 84.5            | 84.5  |
| 7        |     | 14.0           | 16.5  | 35.0           | 41.0  | 53.0           | 60.0  | 82.5            | 85.0  |
| 8        |     | 15.0           | 13.5  | 38.0           | 38.5  | 59.0           | 59.5  | 83.5            | 86.0  |
| 9        |     | 12.5           | 13.5  | 32.5           | 36.0  | 53.5           | 57.5  | 83.5            | 86.0  |
| 10       |     | 12.5           | 13.5  | 33.0           | 37.5  | 55.0           | 56.5  | 82.0            | 87.0  |

**TABLEAU 5.14** : Puissances empiriques pour  $H_0 : \text{Exp}(1)$  contre  $H_1 : \text{Exp}(\lambda)$  avec  $\rho = 0.2$

| $\lambda$ | k    | $\alpha = 1\%$ |       | $\alpha = 3\%$ |       | $\alpha = 5\%$ |       | $\alpha = 10\%$ |       |
|-----------|------|----------------|-------|----------------|-------|----------------|-------|-----------------|-------|
|           |      | $T_c$          | $T_L$ | $T_c$          | $T_L$ | $T_c$          | $T_L$ | $T_c$           | $T_L$ |
| 0.5       | 2    | 94.5           | 98.0  | 97.0           | 99.5  | 99.0           | 100.0 | 99.0            | 100.0 |
|           | 3    | 94.0           | 99.0  | 98.5           | 100.0 | 99.0           | 100.0 | 99.5            | 100.0 |
|           | 4    | 95.5           | 98.5  | 98.5           | 100.0 | 99.5           | 100.0 | 100.0           | 100.0 |
|           | 5    | 96.5           | 97.5  | 99.0           | 100.0 | 99.5           | 100.0 | 100.0           | 100.0 |
|           | 6    | 97.5           | 96.5  | 99.0           | 100.0 | 99.0           | 100.0 | 100.0           | 100.0 |
|           | 7    | 95.0           | 95.5  | 98.0           | 99.0  | 99.5           | 100.0 | 100.0           | 100.0 |
|           | 8    | 94.5           | 94.5  | 99.5           | 98.5  | 99.5           | 100.0 | 100.0           | 100.0 |
|           | 9    | 93.0           | 92.5  | 99.5           | 99.0  | 100.0          | 99.5  | 100.0           | 100.0 |
|           | 10   | 94.5           | 91.0  | 99.0           | 97.5  | 100.0          | 99.5  | 100.0           | 100.0 |
|           | 0.75 | 2              | 34.5  | 36.0           | 43.5  | 49.0           | 50.0  | 57.5            | 63.0  |
| 3         |      | 30.0           | 33.0  | 40.0           | 48.5  | 47.5           | 53.5  | 63.0            | 66.5  |
| 4         |      | 30.5           | 32.5  | 41.5           | 48.0  | 48.5           | 57.5  | 61.5            | 66.0  |
| 5         |      | 29.5           | 30.5  | 40.5           | 44.0  | 46.0           | 54.5  | 57.5            | 65.5  |
| 6         |      | 26.5           | 26.5  | 38.5           | 42.5  | 47.0           | 50.0  | 60.0            | 65.0  |
| 7         |      | 24.5           | 23.0  | 40.0           | 41.5  | 45.0           | 47.0  | 56.0            | 63.5  |
| 8         |      | 20.5           | 21.0  | 33.0           | 40.5  | 43.5           | 48.0  | 55.5            | 60.0  |
| 9         |      | 26.0           | 17.0  | 34.0           | 36.5  | 45.0           | 45.0  | 55.5            | 59.5  |
| 10        |      | 18.0           | 13.5  | 31.5           | 31.5  | 40.0           | 45.5  | 55.5            | 57.5  |
| 1         |      | 2              | 3.5   | 2.0            | 8.0   | 7.0            | 9.5   | 9.5             | 15.0  |
|           | 3    | 2.0            | 2.0   | 5.0            | 5.5   | 8.0            | 9.0   | 18.0            | 15.0  |
|           | 4    | 1.5            | 3.5   | 4.5            | 5.5   | 6.0            | 9.0   | 13.5            | 15.0  |
|           | 5    | 1.5            | 2.5   | 5.5            | 5.0   | 8.0            | 8.0   | 14.0            | 14.0  |
|           | 6    | 0.5            | 2.0   | 3.5            | 4.5   | 7.0            | 8.0   | 16.0            | 13.5  |
|           | 7    | 3.0            | 2.0   | 4.0            | 3.5   | 6.5            | 6.5   | 14.5            | 14.0  |
|           | 8    | 1.5            | 2.0   | 3.0            | 3.5   | 6.5            | 5.5   | 13.5            | 14.0  |
|           | 9    | 0.5            | 1.0   | 3.5            | 3.0   | 6.0            | 4.0   | 14.5            | 11.5  |
|           | 10   | 1.5            | 1.5   | 3.5            | 3.5   | 5.5            | 5.0   | 15.0            | 13.0  |
|           | 1.25 | 2              | 18.0  | 23.0           | 29.5  | 35.0           | 37.0  | 42.5            | 45.0  |
| 3         |      | 14.0           | 16.0  | 28.0           | 29.0  | 34.5           | 37.5  | 46.5            | 51.5  |
| 4         |      | 10.5           | 12.5  | 22.5           | 22.0  | 32.0           | 31.0  | 43.0            | 48.0  |
| 5         |      | 6.5            | 9.5   | 19.5           | 20.5  | 27.5           | 29.0  | 41.5            | 45.5  |
| 6         |      | 9.0            | 6.5   | 18.5           | 17.5  | 26.0           | 25.0  | 38.5            | 44.0  |
| 7         |      | 7.5            | 5.0   | 16.0           | 16.0  | 26.0           | 23.0  | 36.5            | 41.0  |
| 8         |      | 5.5            | 5.5   | 14.5           | 13.0  | 18.5           | 23.5  | 35.5            | 48.5  |
| 9         |      | 5.0            | 5.0   | 15.0           | 14.5  | 22.0           | 21.5  | 34.0            | 37.0  |
| 10        |      | 3.0            | 4.5   | 7.5            | 11.0  | 19.5           | 21.0  | 32.5            | 35.5  |
| 1.5       |      | 2              | 65.5  | 70.5           | 78.0  | 85.0           | 81.5  | 89.0            | 89.0  |
|           | 3    | 52.5           | 61.0  | 76.0           | 78.0  | 80.5           | 84.0  | 89.5            | 92.5  |
|           | 4    | 43.5           | 51.0  | 71.0           | 70.5  | 79.5           | 80.0  | 88.0            | 90.0  |
|           | 5    | 47.0           | 43.0  | 67.5           | 69.5  | 78.5           | 78.0  | 85.0            | 88.5  |
|           | 6    | 39.0           | 38.0  | 64.5           | 62.5  | 74.5           | 73.0  | 86.5            | 90.0  |
|           | 7    | 39.0           | 32.5  | 61.0           | 58.5  | 72.0           | 71.0  | 87.0            | 86.5  |
|           | 8    | 28.0           | 29.5  | 58.0           | 50.5  | 71.5           | 65.5  | 86.5            | 85.0  |
|           | 9    | 28.5           | 26.0  | 52.0           | 49.5  | 65.0           | 61.5  | 83.5            | 83.5  |
|           | 10   | 25.5           | 25.0  | 48.5           | 47.0  | 52.5           | 60.0  | 75.0            | 96.5  |

**TABLEAU 5.15** : Puissances empiriques pour  $H_0 : \text{Exp}(1)$  contre  $H_1 : \text{Exp}(\lambda)$  avec  $\rho = 0.5$

| $\lambda$ | k    | $\alpha = 1\%$ |       | $\alpha = 3\%$ |       | $\alpha = 5\%$ |       | $\alpha = 10\%$ |       |
|-----------|------|----------------|-------|----------------|-------|----------------|-------|-----------------|-------|
|           |      | $T_c$          | $T_l$ | $T_c$          | $T_l$ | $T_c$          | $T_l$ | $T_c$           | $T_l$ |
| 0.5       | 2    | 74.0           | 76.5  | 87.5           | 91.0  | 92.5           | 94.5  | 96.0            | 98.5  |
|           | 3    | 74.0           | 73.0  | 86.5           | 88.5  | 90.5           | 91.0  | 95.0            | 96.5  |
|           | 4    | 72.5           | 68.0  | 84.0           | 86.5  | 89.0           | 92.0  | 95.5            | 95.5  |
|           | 5    | 73.0           | 67.0  | 83.0           | 83.0  | 90.0           | 93.0  | 96.5            | 96.0  |
|           | 6    | 69.5           | 62.5  | 83.0           | 81.0  | 87.5           | 90.0  | 93.5            | 97.0  |
|           | 7    | 67.5           | 57.0  | 84.0           | 79.0  | 88.5           | 87.0  | 94.5            | 96.0  |
|           | 8    | 65.0           | 56.0  | 82.5           | 74.0  | 89.0           | 85.5  | 94.0            | 95.5  |
|           | 9    | 65.5           | 54.5  | 83.5           | 72.0  | 88.5           | 82.5  | 95.0            | 94.0  |
|           | 10   | 64.5           | 52.0  | 83.5           | 71.0  | 88.0           | 78.5  | 94.5            | 93.0  |
|           | 0.75 | 2              | 19.0  | 19.5           | 29.5  | 33.0           | 36.5  | 38.0            | 48.0  |
| 3         |      | 17.5           | 13.5  | 27.0           | 27.0  | 32.5           | 34.5  | 44.5            | 51.5  |
| 4         |      | 17.0           | 11.0  | 29.0           | 23.5  | 36.0           | 33.5  | 45.5            | 48.5  |
| 5         |      | 13.0           | 10.5  | 24.0           | 21.0  | 35.5           | 29.0  | 47.0            | 46.0  |
| 6         |      | 12.0           | 9.5   | 21.0           | 18.5  | 33.5           | 25.5  | 47.5            | 45.0  |
| 7         |      | 11.0           | 7.5   | 21.5           | 18.0  | 31.5           | 25.5  | 46.5            | 41.0  |
| 8         |      | 10.5           | 8.5   | 20.5           | 16.5  | 29.0           | 23.0  | 49.5            | 39.5  |
| 9         |      | 9.0            | 7.5   | 21.0           | 15.5  | 30.5           | 23.0  | 44.0            | 36.5  |
| 10        |      | 8.0            | 6.5   | 19.0           | 15.5  | 26.5           | 23.0  | 42.0            | 36.0  |
| 1         |      | 2              | 7.5   | 7.5            | 12.5  | 13.0           | 14.5  | 16.0            | 21.5  |
|           | 3    | 4.5            | 6.0   | 9.5            | 10.5  | 12.0           | 13.5  | 22.5            | 22.0  |
|           | 4    | 3.5            | 5.5   | 9.5            | 9.5   | 11.0           | 12.5  | 19.5            | 22.5  |
|           | 5    | 4.0            | 4.5   | 6.0            | 8.0   | 11.5           | 11.0  | 22.0            | 18.5  |
|           | 6    | 4.0            | 3.0   | 6.5            | 5.0   | 9.0            | 10.0  | 15.0            | 17.5  |
|           | 7    | 3.5            | 2.5   | 6.5            | 4.5   | 8.0            | 9.0   | 18.0            | 16.5  |
|           | 8    | 2.0            | 2.0   | 5.5            | 5.0   | 7.5            | 7.5   | 15.0            | 16.0  |
|           | 9    | 1.5            | 2.5   | 4.5            | 5.5   | 7.0            | 7.5   | 14.0            | 17.5  |
|           | 10   | 1.5            | 2.5   | 5.5            | 5.0   | 6.0            | 8.5   | 13.5            | 16.5  |
|           | 1.25 | 2              | 17.5  | 23.5           | 27.0  | 33.0           | 34.0  | 41.5            | 47.5  |
| 3         |      | 15.0           | 15.5  | 24.5           | 28.0  | 28.5           | 36.5  | 42.0            | 49.5  |
| 4         |      | 10.0           | 14.0  | 20.5           | 24.5  | 27.0           | 32.0  | 40.5            | 47.0  |
| 5         |      | 13.5           | 13.5  | 21.0           | 22.0  | 27.5           | 28.5  | 41.5            | 42.5  |
| 6         |      | 10.0           | 12.5  | 19.0           | 21.0  | 25.5           | 28.5  | 38.0            | 43.5  |
| 7         |      | 8.5            | 12.5  | 17.5           | 20.5  | 25.5           | 26.5  | 37.5            | 43.0  |
| 8         |      | 9.0            | 11.5  | 18.0           | 20.0  | 24.0           | 26.5  | 35.0            | 40.5  |
| 9         |      | 6.5            | 10.0  | 17.0           | 17.5  | 23.0           | 25.5  | 38.0            | 39.5  |
| 10        |      | 6.5            | 8.5   | 17.0           | 16.0  | 22.0           | 22.5  | 38.0            | 37.0  |
| 1.5       |      | 2              | 46.5  | 58.5           | 59.0  | 68.0           | 65.5  | 74.0            | 76.5  |
|           | 3    | 42.5           | 46.5  | 56.0           | 62.5  | 66.5           | 68.5  | 75.5            | 78.0  |
|           | 4    | 35.0           | 40.0  | 53.5           | 62.0  | 62.0           | 64.5  | 72.5            | 75.5  |
|           | 5    | 38.0           | 35.5  | 52.0           | 54.5  | 61.0           | 64.0  | 70.5            | 73.5  |
|           | 6    | 33.0           | 33.0  | 47.0           | 52.5  | 57.0           | 60.5  | 71.5            | 72.5  |
|           | 7    | 25.5           | 30.0  | 49.5           | 48.0  | 61.0           | 59.0  | 72.0            | 71.0  |
|           | 8    | 27.5           | 28.5  | 44.0           | 45.0  | 55.5           | 57.0  | 69.0            | 69.5  |
|           | 9    | 25.0           | 26.0  | 40.0           | 40.5  | 54.5           | 54.0  | 68.0            | 70.5  |
|           | 10   | 25.0           | 22.5  | 37.5           | 37.5  | 52.5           | 49.0  | 69.5            | 68.0  |

# Chapitre VI

## Test non paramétrique de la densité marginale d'un processus stationnaire à temps continu

### VI.1. Introduction

L'estimateur non paramétrique de la densité a été étudié par de nombreux auteurs ([31], [74], [80], [4], [34] entre autres). Plusieurs méthodes ont été utilisées dont celle du noyau, celle des ondelettes, et l'estimation par séries orthogonales. Cette dernière a d'abord été discutée par Cencov(1962) et plus tard par Schwartz(1967), Watson(1969), Bosq(1970) pour des variables aléatoires indépendantes, puis Ahmad(1979) pour des variables aléatoires dépendantes.

Frenay (2001) a étudié l'estimation par des fonctions orthogonales d'un processus stationnaire fortement mélangeant à temps continu. Ce cas a été étudié par Delcroix(1980) dans le cadre plus général des  $\delta$ -séquences.

Dans le présent travail, nous rappelons tout d'abord quelques résultats sur l'estimateur de la densité marginale d'un processus à temps continu par projection orthogonale obtenus dans [40], ensuite nous étudierons des tests d'hypothèses portant sur la densité marginale de ce processus. Le plan de la suite du chapitre sera le suivant : les données du problème sont abordées dans la section 2. A la section 3, la densité  $f$  est estimée. Les tests d'hypothèses sont envisagés à la section 4. A la section 5, on étudie la vitesse de convergence en montrant que nous pouvons appliquer les résultats obtenus au chapitreII(section 4) à l'estimateur de la densité et, enfin on aborde la con-

vergence du test à la section 6.

## VI.2. Les données du problème

Soit  $(X_t, t \in [0, T])$ ,  $T > 1$  un processus strictement stationnaire à temps continu et  $\alpha$ -mélangeant défini sur un espace probabilisé  $(\Omega, \mathcal{A}, P)$  à valeurs dans un espace mesurable  $(E, \mathcal{B})$ , où  $E$  est l'espace des observations,  $\mathcal{B}$  une tribu sur  $E$ , et soit  $\mu$  une probabilité sur  $(E, \mathcal{B})$ . Rappelons que la  $\alpha$ -mélangeance a été définie au chapitre II.

Soit  $L^2(\mu)$  l'espace de Hilbert séparable dont une base orthonormale est donnée par  $e_0, \dots, e_k, \dots$ , avec  $e_0 \equiv 1$ , et les  $e_j$ ,  $j \geq 0$  des applications mesurables. On suppose également que la densité marginale commune des  $X_t$  est  $f$ .

Soit  $\mathcal{P}$  une famille de mesures de probabilités sur  $(E, \mathcal{B})$  dominée par  $\mu$ , et  $\nu$  un élément de  $\mathcal{P}$ . On suppose que  $f$  est la densité de  $\nu$  par rapport à  $\mu$ , i.e :

$$f = \frac{d\nu}{d\mu} \quad \mu - p.s$$

et  $f \in L^2(\mu)$ , espace vectoriel muni du produit scalaire suivant :

$$\langle h, g \rangle = \int hgd\mu, \quad h, g \in L^2(\mu)$$

Soit  $H = L^2(E, \mathcal{B}, \nu)$ , et  $H_{k_T}$  le sous espace de  $H$  engendré par les fonctions  $e_j$ ,  $j = 0, \dots, k$ .

Comme précédemment (chap II), le problème est de tester l'hypothèse  $H_0$  : " La loi des  $X_t$  est  $\mu$ " contre l'hypothèse  $H_1$  : " La loi des  $X_t$  est  $\nu$ ,  $\nu \neq \mu$ ". Du point de vue pratique, on est amené à tester l'hypothèse  $H_0$  : " $f=1$ " contre  $H_1$  : " $f \neq 1$ ".

Rappelons d'abord quelques résultats contenus dans [40] sur l'estimation de  $f$ .

## VI.3. Estimation de $f$

$f$  étant de carré intégrable, peut être développée en série de la forme :

$$f = \sum_{j=1}^{+\infty} u_j e_j$$

où

$$a_j = \int f e_j d\mu$$

On désire estimer par projection la fonction inconnue  $f$  sur un sous espace  $E_{k_T}$  de  $L^2(\mu)$  de dimension finie  $k_T$  qui varie avec  $T$ .  
L'estimateur de  $f$  est donné par

$$\hat{f}_T(x) = \sum_{j=0}^{k_T} \hat{a}_{j,T} e_j(x)$$

avec

$$\lim_{T \rightarrow +\infty} k_T = +\infty$$

et

$$\frac{1}{T} \int_0^T e_j(X_t) dt = \hat{a}_{j,T}$$

Plusieurs types de convergence seront examinés, l'estimateur obtenu sera utilisé pour les tests d'hypothèses.

### VI.3.1. Convergence en moyenne quadratique

#### **Théorème 25**

Sous les conditions suivantes :

- $f$  est continue et  $f \in H$
- $\alpha(u) \in L^1(\lambda)$  et

$$\frac{k_T^2}{T} \xrightarrow{T \rightarrow \infty} 0$$

$$\sup_{j \geq 0, x \in E} |e_j(x)| = C < +\infty$$

Alors :

1.

$$E[\hat{f}_T(x) - f(x)]^2 \xrightarrow{T \rightarrow \infty} 0 \quad (53)$$

c.à.d,  $\hat{f}_T(x)$  converge en moyenne quadratique vers  $f(x)$

2. De plus, si il existe  $r > 2$  tel que

$$\left| \sum_{j=k_T+1}^{+\infty} a_j \right|^2 \sim k_T^{2-r}$$

et si

$$k_T = T^{\frac{1}{r}}$$

alors

$$E[\hat{f}_T(x) - f(x)]^2 = o\{T^{\frac{2-r}{r}}\} \quad (54)$$

### Démonstration

1)

$$E[\hat{f}_T(x) - f(x)]^2 = E[\hat{f}_T(x) - E[\hat{f}_T(x)]]^2 + [E[\hat{f}_T(x)] - f(x)]^2$$

Après un calcul rapide, on obtient que :

$$E[\hat{f}_T(x)] = \sum_{j=0}^{k_T} a_j c_j(x)$$

d'où  $\hat{f}_T(x)$  est un estimateur sans biais de la projection de la densité  $f(x)$  sur le sous espace  $H_k$ , et donc  $\hat{f}_T(x)$  est un estimateur asymptotiquement sans biais de  $f(x)$ .

Par suite,

$$[E[\hat{f}_T(x)] - f(x)]^2 \rightarrow_{T \rightarrow +\infty} 0 \quad (55)$$

$$\left\{ \begin{aligned} E[\hat{f}_T(x) - E[\hat{f}_T(x)]]^2 &= \text{var } \hat{f}_T(x) = E\left[\sum_{j=0}^{k_T} (\hat{a}_{jT} - E[\hat{a}_{jT}])e_j(x)\right]^2 \\ &= \sum_{j=0}^{k_T} (\hat{a}_{jT} - E[\hat{a}_{jT}])^2 e_j(x)^2 + 2 \sum_{0 \leq j < l \leq k_T} e_j e_l \text{cov}(\hat{a}_{jT}, \hat{a}_{lT}) \quad (56) \\ &= \sum_{j=0}^{k_T} e_j^2 \text{var}(\hat{a}_{jT}) + 2 \sum_{0 \leq j < l \leq k_T} e_j e_l \text{cov}(\hat{a}_{jT}, \hat{a}_{lT}) \end{aligned} \right.$$

Nous allons d'abord majorer  $\text{var}(\hat{a}_{jT})$  puis  $\text{cov}(\hat{a}_{jT}, \hat{a}_{lT})$ .

$$\text{var}(\hat{a}_{jT}) = \text{var}\left[\frac{1}{T} \int_0^T e_j(x_t) dt\right]$$

Le théorème de Fubini permet d'écrire

$$\text{var}(\hat{a}_{jT}) = \frac{1}{T^2} \int_{[0,T]^2} \text{cov}(e_j(x_t), e_j(x_s)) ds dt$$

Nous allons appliquer l'inégalité de Billingsley et la majoration de la fonction  $e_j$ , nous obtenons alors :

$$\text{var}(\hat{a}_{jT}) \leq \frac{1}{T^2} 4c^2 \int_{[0,T]^2} \alpha(|t-s|) ds dt \quad (57)$$

$$\leq \frac{1}{T} 4c \int_{[0,T]} \alpha(u) du \quad (58)$$



$$\text{cov}(\hat{a}_{jT}, \hat{a}_{lT}) \leq [\text{var}(\hat{a}_{jT})\text{var}(\hat{a}_{lT})]^{\frac{1}{2}}$$

On a alors

$$\text{cov}(\hat{a}_{jT}, \hat{a}_{lT}) \leq \frac{1}{T} 4c \int_0^T \alpha(u) du \quad (59)$$

En insérant (57) et (59) dans (55), on a :

$$\left\{ \begin{array}{l} \text{var}[\hat{f}_T(x)] \\ \leq \\ \sum_{j=0}^{k_T} e_j^2(x) \frac{1}{T} 4c \int_0^T \alpha(u) du + 2 \sum_{0 \leq j < l \leq k_T} e_j e_l \frac{1}{T} 4c \int_0^T \alpha(u) du \\ \leq \\ \text{var}[\hat{f}_T(x)] \leq \sum_{j=0}^{k_T} 4c^3 \frac{k_T}{T} \int_0^T \alpha(u) du + 4c^3 \frac{k_T^2}{T} \int_0^T \alpha(u) du \end{array} \right.$$

En utilisant le fait que  $k_T \geq 1$  et que  $\alpha(u)$  est intégrable, on a finalement :

$$\text{var}[\hat{f}_T(x)] = O\left[\frac{k_T^2}{T}\right] \quad (60)$$

d'où (55) et (60) entraînent (53).

2) Si maintenant nous supposons

$$\left[ \sum_{j=1+\lfloor k_T \rfloor}^{+\infty} a_j \right]^2 \simeq k_T^{2-r}$$

et si nous prenons  $k_T = T^{\frac{1}{r}}$ , alors

$$|E[\hat{f}_T(x)] - f(x)|^2 \leq c \left[ \sum_{j=1}^{+\infty} a_j \right]^2 \leq c (k_T)^{2-r}$$

Par conséquent

$$E[\hat{f}_T(x)] - f(x) = 0 \left[ \frac{k_T^2}{T} \right] + 0 [k_T^{2-r}] = 0 \left[ T^{-\frac{2-r}{r}} \right]$$

Ce qui donne (54)

### **Proposition 26**

Si  $f \in H$  et si

$$|e_j(x)| \leq c(1+j)^{-\gamma}$$

avec  $0 \leq \gamma \leq 1/2$ , où  $c$  est une constante. Alors il existe une constante  $A$  telle que pour tout  $T$  et tout  $\tau \leq \frac{T}{2}$ , on ait

$$V[\hat{f}_T(x)] \leq A \left\{ \frac{k_T^{2-2\gamma}}{T\tau} \sigma_\tau^2 + k_T^{2-4\gamma} \alpha(\tau) \right\} \quad (61)$$

où

$$\sigma_\tau^2 = \sup_{0 \leq j \leq k_T, 0 \leq t \leq \tau} E \left[ \int_t^{t+\tau} (e_j(X_s) - E(e_j(X_s))) ds \right]^2$$

**Démonstration**

$$\text{var}[\hat{f}_T(x)] = E \left[ \sum_{j=0}^{k_T} [\hat{a}_{jT} - E(\hat{a}_{jT})] e_j(x) \right]^2$$

Nous utiliserons l'inégalité de Holder appliquée aux sommes.

$$\left\{ \begin{array}{l} \text{var}[\hat{f}_T(x)] \leq E\left[\left(\sum_{j=0}^{k_T} e_j^2(x)\right) \left(\left[\sum_{j=0}^{k_T} [a_{jT} - E(\hat{a}_{jT})]^2\right]\right)\right] \\ \leq \left(\sum_{j=0}^{k_T} e_j^2(x)\right) \left(\sum_{j=0}^{k_T} E[\hat{a}_{jT} - E(\hat{a}_{jT})]^2\right) \end{array} \right. \quad (62)$$

Posons  $\varphi_s^j = e_j(x_s) - E[e_j(x_s)]$

On a :

$$E[\varphi_s^j] = 0, \quad \sup_j \sup_s |\varphi_s^j| \leq 2c(1+j)^{-\gamma} \quad \text{et} \quad E[\varphi_s^j] < +\infty$$

d'où

$$\left\{ \begin{array}{l} \text{var}[\hat{f}_T(x)] \leq c^2 \left(\sum_{j=0}^{k_T} (1+j)^{-2\gamma}\right) \left(\sum_{j=0}^{k_T} E\left[\frac{1}{T^2} \int_0^T \varphi_s^j ds\right]^2\right) \\ \leq \frac{c^2 k_T^{1-2\gamma}}{T^2} \sum_{j=0}^{k_T} E\left(\int_0^T \varphi_s^j ds\right)^2 \end{array} \right.$$

En utilisant le résultat de Leblanc[59] page 35 à savoir :

**Lemme 4 :** Soit  $(X_s)$  un processus mesurable à valeurs réelles, localement intégrable, à temps continu, tel que pour tout  $s$ ,  $\|X\|_\infty \leq B$ ,  $E(x_s) = 0$  et  $E(X_s^2) < +\infty$ .

Pour  $p \in [2, +\infty[$ , il existe une constante  $D$  telle que pour tout  $T$  positif,  $B$  et  $\tau \leq \frac{T}{2}$  :

$$E\left(\left|\int_0^T X_s ds\right|^p\right) \leq D \left\{ \left(\frac{T}{\tau}\right)^{\frac{p}{2}} \sigma_\tau^p + \frac{T}{\tau} \sigma_\tau^2 (\tau B)^{p-2} + (\tau B)^p \left(\frac{T}{\tau}\right)^p \alpha(\tau) \right\} \quad (63)$$

où

$$\sigma_\tau^2 = \sup_{0 \leq t \leq T} E \left( \int_t^{t+\tau} X_s ds \right)^2$$

On obtient en tenant compte de (62)

$$\left\{ \begin{aligned} \text{var}[\hat{f}_T(x)] &\leq c^2 D \frac{k_T^{1-2\gamma}}{T^2} \sum_{j=0}^{k_T} \left\{ 2 \frac{T}{\tau} \sigma_\tau^2 + (1+j)^{-2\gamma} T^2 \alpha(\tau) \right\} \\ &\leq c^2 D \left( \frac{k_T^{1-2\gamma}}{T^2} \sigma_\tau^2 + k_T^{2-4\gamma} \alpha(\tau) \right) \end{aligned} \right. \quad (64)$$

En posant  $C^2 D = A$  dans (64), on obtient (61).

**Remarque :**

**La condition**

$$|e_j(x)| \leq c(1+j)^{-\gamma}$$

n'est pas irréalisable.

En effet, si la base orthonormale est constituée des fonctions d'Hermites, on a

$$|e_j(x)| \leq c_1(1+j)^{-1/12}$$

sur  $\mathbb{R}$ , pour tout  $j$ .

Pour les fonctions trigonométriques, on a

$$|e_j(x)| \leq c_2$$

sur  $[-\Pi, \Pi]$ , où  $c_2$  est une constante

### VI.3.2. Convergence presque sûre

#### Théorème 26

Soit  $f \in H$ .

- Si

$$k_T = \left(\frac{T}{\log T}\right)^{\frac{1}{r}}$$

avec  $r > 2$ , et si  $\sigma_7^2 \leq c_1 \tau$ , où  $c_1$  est une constante.

- Si

$$\alpha(u) \leq c_2 u^{-u}, \quad u > 0$$

( respectivement  $\alpha(u) \leq c_3 \alpha^{-u}$ ) avec

$$1 < \alpha < \frac{7r + 6 - 6\gamma}{2r - 4 + 4\gamma}$$

(respectivement  $\alpha > 1$ ),  $c_2, c_3$  étant des constantes, alors

1.

$$\frac{1}{\log_m(T)} \left(\frac{T}{\log T}\right)^{\frac{r-2+2\gamma}{2r}} [\hat{f}_T(x) - E(f_T(x))] \xrightarrow{T \rightarrow \infty} 0 \text{ p.s.} (65)$$

où  $m > 1$

2. Si de plus

$$\left[ \sum_{j=k_T+1}^{\infty} a_j \right]^2 \sim k_T^{2-r}$$

alors

$$\frac{1}{\log_m(T)} \left(\frac{T}{\log T}\right)^{\frac{r-2+2\gamma}{2r}} [f_T(x) - f(x)] \xrightarrow{T \rightarrow \infty} 0 \text{ p.s.} (66)$$

Démonstration . Voir [40] p 59-62. ■

**Remarque :** le théorème précédent permet ainsi d'établir la convergence presque sûre de l'estimateur  $\hat{f}_T$ .

Après avoir donné ces quelques éléments, abordons maintenant le test d'hypothèse.

## VI.4. Tests d'hypothèses

Rappelons qu'on veut tester l'hypothèse  $H_0$  : " La loi des  $X_t$  est  $\mu$ " (ou encore "  $f$ , densité commune des  $X_t$  est 1") contre l'hypothèse  $H_1$  : " La loi des  $X_t$  n'est pas  $\mu$ " (ou encore " $f \neq 1$ "). Dans cette partie, nous supposons  $T$  entier et nous posons  $T = n$ ,  $n > 1$  ( $k_T$  devient  $k_n$  et  $\hat{f}_T(x)$  devient  $\hat{f}_n(x)$ )

### VI.4.1. La loi des $X_t$ est $\mu$

#### VI.4.1.1. Préliminaires

Comme précédemment, on pose

$$f = \sum_{j \geq 0} a_j e_j$$

où

$$a_j = \int f e_j d\mu$$

avec nos nouvelles notations, on a :

$$\hat{f}_n(x) = \sum_{j=0}^{k_n} \hat{a}_{jn} e_j(x)$$

où

$$\hat{a}_{jn} = \frac{1}{n} \int_0^n e_j(X_t) dt.$$

C'est à dire encore

$$\hat{a}_{jn} = \frac{1}{n} \sum_{i=1}^n \int_{i-1}^i e_j(X_t) dt;$$

posons

$$Y_i^j = \int_{i-1}^i e_j(X_t) dt$$

donc

$$\hat{a}_{jn} = \frac{1}{n} \sum_{i=1}^n Y_i^j$$

$(X_t)$  étant un processus strictement stationnaire et les  $e_j$  mesurables, alors les  $(e_j(X_t))$  le sont également. Par conséquent :

$$(Y_i^j)_{i \geq 1} = \left( \int_{i-1}^i e_j(X_t) dt \right)_{i \geq 1}$$

est strictement stationnaire et mélangeant.

Sous l'hypothèse nulle,  $\hat{a}_{jn}$  est centré pour  $j \geq 1$ .

En effet

$$E[Y_i^j] = E \left[ \int_0^1 e_j(X_t) dt \right]$$

En appliquant Fubini, on a :

$$E[Y_i^j] = \int_0^1 E[e_j(X_t)] dt$$

et sous  $H_0$  on obtient :

$$E[e_j(X_t)] = \int 1 \cdot e_j d\mu = 0, j \geq 1$$

par suite,

$$E[Y_i^j] = 0, j \geq 1$$

VI.4.1.2. Loi limite de  $\sqrt{n}\hat{a}_{jn} = \sqrt{n}\bar{Y}_n^j, j \geq 1$

$$\bar{Y}_n^j = \frac{1}{n} \sum_{i=1}^n Y_i^j$$

$(X_t)_{t \in \{0,1,n\}}$  étant un processus strictement stationnaire et  $\alpha$ -mélangeant alors le processus stationnaire  $(Y_i^j)_{i \geq 1}$ , qui est une fonction mesurable de  $(X_t)$  est également  $\alpha$ -mélangeant, et on a

$$\alpha_{Y^j} \leq \alpha_X(i), \forall i \geq 1$$

où  $\alpha$  désigne le coefficient de mélangeance.

**Proposition 27**

$(Y_i^j)_{j \geq 1}$  étant un processus réel, centré stationnaire tel que

- si Pour tout  $\gamma > 0$ , on a

$$- \sup E[|Y_i^j|^{2+\gamma}] < +\infty$$

$$- \sum_{i \geq 1} \alpha_{Y^j}(i)^{\frac{\gamma}{\gamma+2}} < +\infty$$

Alors

1. La série  $\sum_{i \in \mathbb{N}^*} cov(Y_1^j, Y_i^j)$  converge absolument et sa limite vaut  $\sigma_j^2$ , avec

$$\sigma_j^2 = \sum_{i \in \mathbb{N}^*} cov(Y_1^j, Y_i^j) \quad (67)$$

de plus

$$nV[\bar{Y}_n^j] \longrightarrow \sigma_j^2 \quad (68)$$

2. Si en plus on a

$$\alpha_Y(m) \leq am^{-\beta}$$

où  $a$  est une constante positive et  $\beta > \frac{2+\gamma}{\gamma}$ , alors

$$\lim_{n \rightarrow +\infty} \sqrt{n} \bar{Y}_n^j \xrightarrow{\mathcal{L}} N_j \sim N(0, \sigma_j^2) \quad (69)$$



Démonstration . Elle est similaire à celles dans [20] p 34-36 et [69] p.7 et 64.

En considérant notre processus  $(Y_i^j)_{i \geq 1}$ , il suffit de réécrire la démonstration en remplaçant  $(X_t)_{t \in \mathbb{Z}}$  par  $(Y_i^j)_{i \geq 1}$ , avec:

$$Y_i^j = \int_{i-1}^i e_j(X_t) dt$$

Remarque :

$$\sigma_j^2 = \sum_{i=1}^{\infty} \text{cov}(Y_1^j, Y_i^j) = \text{cov}(Y_1^j, Y_1^j) + \sum_{i=2}^{\infty} \text{cov}(Y_1^j, Y_i^j)$$

$$\begin{aligned} \text{cov}(Y_1^j, Y_1^j) &= V(Y_1^j) = V\left[\int_0^1 e_j(X_t) dt\right] \\ &= E[(Y_1^j)^2] \end{aligned}$$

On a classiquement

$$V(Y_1^j) = \int_0^1 \int_0^1 \text{cov}(e_j(X_t), e_j(X_s)) dt ds$$

Alors

$$\begin{aligned} \text{cov}(Y_1^j, Y_i^j) &= E\left[\int_0^1 e_j(X_t) dt \int_{i-1}^i e_j(X_s) ds\right] \\ &= \int_0^1 \int_{i-1}^i \text{cov}(e_j(X_t), e_j(X_s)) dt ds \\ &= E[Y_1^j Y_i^j] \end{aligned}$$

### VI.4.1.3. Loi limite de $n \| 1 - \hat{f}_n \|^2$

La loi des  $X_i$  étant  $\mu$  (hypothèse  $H_0$ ), alors

$$a_j = \int 1 \cdot e_j d\mu = \delta_{j,0}.$$

Tester l'hypothèse "la loi des  $X_i$  est  $\mu$ " revient à tester l'hypothèse, comme nous l'avons indiqué plus haut : " $f=1$ ".

D'où la recherche du comportement de  $n \| 1 - \hat{f}_n \|^2$  vers sa loi limite en terme de recherche de la vitesse de convergence.

$$1 - \hat{f}_n(x) = 1 - \sum_{j=0}^{k_n} \hat{a}_{jn} e_j(x)$$

mais  $e_0 = 1$ , et  $a_{0n} = \frac{1}{n} \int_0^n 1 dt = 1$ , par suite

$$1 - \hat{f}_n(x) = \sum_{j=1}^{k_n} \hat{a}_{jn} e_j(x)$$

donc

$$n \| 1 - \hat{f}_n \|^2 = \sum_{j=1}^{k_n} (\sqrt{n} \hat{a}_{jn})^2.$$

Nous supposons maintenant que  $k_n = k$  est fixé.

Loi limite de  $S_n$  Posons

$$S_n = \sqrt{n} \begin{pmatrix} \hat{a}_{1n} \\ \hat{a}_{2n} \\ \dots \\ \hat{a}_{kn} \end{pmatrix}$$

et

$$Y_i = \begin{pmatrix} Y_i^1 \\ Y_i^2 \\ \dots \\ Y_i^k \end{pmatrix}$$

Posons aussi

$$A_n = \frac{1}{\sqrt{n}} S_n.$$

$$V_i = \sum_{j=1}^k C_j Y_i^j = C^t Y_i$$

où

$$C^t = (C_1, \dots, C_k) \in \mathbb{R}^k$$

### Proposition 28

Sous les hypothèses de la proposition 27, on a

$$S_n \xrightarrow{\mathcal{L}} N = (N_1, \dots, N_k)^t$$

où

$$\forall j = 1, \dots, k, N_j \sim N(0, \sigma_j^2)$$

et de matrice de variance covariance  $\Sigma$  donnée par

$$\Sigma = (\sigma_{jj'})_{1 \leq j, j' \leq k}$$

$$\Sigma = \left( \sum_{\nu=1}^{+\infty} \text{cov}(Y_1^j, Y_{\nu+1}^{j'}) \right)_{1 \leq j, j' \leq k}$$

c'est à dire

$$\sigma_{jj'} = \sum_{\nu=1}^{+\infty} \text{cov}(Y_1^j, Y_{\nu+1}^{j'})$$

Démonstration . Elle est similaire à celle dans [45] p 11-13 où  $e_j(X_i)$  est remplacé par  $Y_i^j = \int_{i-1}^i e_j(X_t)dt$  ■

Loi limite de  $\sum_{j=1}^k (\sqrt{n}\hat{a}_{jn})^2$ , k fixé

**Proposition 29**

Sous les hypothèses de la proposition 27, alors

$$\| S_n(\cdot) \|^2 = \sum_{j=1}^k (\sqrt{n}\hat{a}_{jn})^2 \longrightarrow Q \quad (70)$$

dont la fonction caractéristique est donnée par

$$\phi_Q(t) = \det(I_k - 2it \sum)^{-1/2}$$

où  $I_k$  désigne l'identité de  $R^k$  et  $\sum$  la matrice de covariance définie ci dessus.

Démonstration . On utilise le fait que la convergence en loi se conserve par continuité et le théorème 1 p88 de [4]. ■

**Remarque :**

Cas où k est fonction de la taille de l'échantillon ( $k = k_n$ )

$$n \| 1 - \hat{f}_n \|^2 = \sum_{j=1}^{k_n} (\sqrt{n}\hat{a}_{jn})^2, \quad \lim k_n = +\infty$$

La dimension dépend de n, nous ne chercherons donc pas la loi limite de  $n \| 1 - \hat{f}_n \|^2$ , mais nous chercherons plutôt à majorer la quantité

$$\sup_{a \in R^+} | P[n \| 1 - \hat{f}_n \|^2 \leq a^2] - P[Q_n \leq a^2] | \quad (71)$$

où  $Q_n$  est une variable aléatoire dont la fonction caractéristique est donnée par

$$\phi_{Q_n}(t) = [\det(I_{k_n} - 2it\Sigma_n)]^{-1/2}$$

où  $I_{k_n}$  est la matrice identité dans  $R^{k_n}$  et

$$\Sigma_n = (\sigma_{jj'})_{1 \leq j, j' \leq k_n}$$

On utilise le fait que si  $k$  est fixe la v.a  $\sum_{j=1}^k (\sqrt{n}\hat{a}_{jn})^2$  converge en loi vers la v.a  $Q$ , ce qui permet d'avoir la forme de la loi de  $Q_n$ .

#### VI.4.2. La loi des $X_t$ est $\nu \neq \mu$

Nous sommes donc dans l'hypothèse  $H_1$

$$n \|1 - \hat{f}_n\|^2 := \sum_{j=1}^k (\sqrt{n}\hat{a}_{jn})^2$$

$$\hat{a}_{jn} - a_j = \frac{1}{n} \sum_{i=1}^n (Y_i^j - a_j)$$

#### Proposition 30

Si la loi des  $X_t$  est  $\nu$  telle qu'il existe un  $j \in 1, 2, \dots, k_n$  pour lequel  $e_j$  soit  $\nu$ -intégrable et  $\int e_j d\nu \neq 0$ , alors

$$n \|1 - \hat{f}_n\|^2 \xrightarrow{p.s.} +\infty \quad (72)$$

Démonstration. Le processus  $(e_j(X_t))$  étant stationnaire et  $\alpha$ -mélangeant, il est donc ergodique (Bradley p 172).

On peut alors appliquer la loi forte des grands nombres pour les processus, il vient :

$$\hat{a}_{jn} \xrightarrow{p.s.} a_j$$

et par continuité

$$n \sum_{i=1}^{k_n} \hat{a}_{jn}^2 \xrightarrow{p.s.} +\infty$$

Remarque : construction du test.

Pour tester l'hypothèse  $H_0$  : "La loi des  $X_t$  est  $\mu$ " contre l'hypothèse  $H_1$  : "La loi des  $X_t$  est  $\nu \neq \mu$ ", la proposition 30 permet d'affirmer qu'on peut utiliser un test de région critique de la forme

$$n \sum_{i=1}^{k_n} \hat{a}_{jn}^2 > w_n$$

Si  $\alpha_n$  est le niveau du test on aura :

$$\alpha_n = P\left[n \sum_{i=1}^{k_n} \hat{a}_{jn}^2 > w_n\right]$$

## VI.5. Etude de la vitesse de convergence

Les résultats que nous avons obtenus dans [41] p 76-79 peuvent s'appliquer au processus  $(Y_i^j)_{i \geq 1}$ .

En effet les hypothèses portant sur le processus  $e_j(X_t)$  dans la proposition 6 de [41] p 79 à vérifier par le processus  $(Y_i^j)_{i \geq 1}$  sont les suivantes :

1. La loi des  $X_t$  est  $\mu$
2. Les  $(Y_i^j)_{i \geq 1}$  sont uniformément bornés
3. Une condition sur la mélangeance et la dimension  $k_n$ .

Seul le 2<sup>e</sup> point est à vérifier, puisque les points 1 et 3 sont des hypothèses que nous pouvons nous donner.

Si le processus  $e_j(X_t)$  est uniformément borné, alors  $(Y_i^j)_{i \geq 1}$  l'est également.

En effet

$$\sup_{x \in E, j \in \{1, \dots, k_n\}} |Y_i^j| \leq M$$

(où  $M = \sup_j \sup_x |e_j(x)|$ ).

On a donc

**Proposition 31**

1. Si la loi des  $X_t$  est  $\mu$
2. Si le processus  $e_j(X_t)$  est uniformément borné
3. si

$$\sum_{i=1}^{\infty} \sqrt{\alpha(i)} < \infty$$

4. Si

$$\lim_{n \rightarrow +\infty} k_n = +\infty$$

et

$$\lim_{n \rightarrow +\infty} \frac{k_n^4}{n} = 0$$

alors, pour  $n$  assez grand, on a

$$\begin{aligned} & \sup_{a \in \mathbb{R}^+} |P[n \|1 - \hat{f}_n\|^2 \leq a^2] - P[\|U_{k_n}(\cdot)\|^2 \leq a^2]| \\ &= O\left[\left(\frac{k_n^6}{n}\right)^{1/6}\right] + O\left[\left(a \frac{k_n^6}{n}\right)^{1/24}\right] + O\left[\left(\frac{1}{k_n}\right)^{k_n} a^{k_n}\right] + O\left[(na)^{k_n} \alpha n^{1/4}\right] \end{aligned}$$

où

$$U_{k_n}(\cdot) = \sum_{i=1}^{k_n} N_j e_j(\cdot), \quad N_j \sim N(0, \sigma_j^2)$$

*Démonstration* . Voir [41] p 76-79. Il suffit de remplacer les  $e_j(X_n)$  par  $Y_i^j$ , avec  $(Y_i^j)_{i \in \mathbb{N}}$  constituant un processus stationnaire et mélangant, et on obtient le résultat ■

**Corollaire 12 :**

**Sous les hypothèses de la proposition 31, si en plus**

- 1.

$$k_n = [n^\beta], \quad 0 < \beta < 1/6$$

2.

$$\alpha(n) = a\rho^n, \quad a > 0, \quad \rho < 1/n, \quad n \geq 1$$

alors

$$\sup_{a \leq \log(2+n)} |P[n \mid \|1 - \hat{f}_n\|^2 \leq a^2] - P[\|U_{k_n}(\cdot)\|^2 \leq a^2]| \quad (73)$$

$$= O[n^{\delta/4} \log(2+n)]$$

où  $-1/6 < \delta < 0$ .

Remarque :

On a choisit  $a < \log(2+n)$  pour assurer la convergence vers 0 de

$$\sup_{a \in \mathbb{R}^+} |P[n \mid \|1 - \hat{f}_n\|^2 \leq a^2] - P[\|U_{k_n}(\cdot)\|^2 \leq a^2]|$$

## VI.6. Convergence du test

Dans la construction du test nous avons posé

$$\alpha_n = P[n \mid \|1 - \hat{f}_n\|^2 > w_n]$$

où  $\alpha_n$  est le niveau du test

### Proposition 32

1. Si  $(e_j(X_l))$  est uniformément borné.
2. Si  $\sum \alpha(i) < +\infty$

Alors

$$\alpha_n = O\left[\frac{k_n^2}{w_n}\right] \quad (74)$$



Démonstration .

$$\alpha_n = P\{n \| 1 - \hat{f}_n \|^2 > w_n\}$$

$$= P\left\{\sum_{j=1}^{k_n} \hat{a}_{jn}^2 > w_n/n\right\}$$

$$\leq \sum_{j=1}^{k_n} P\left[\hat{a}_{jn}^2 > \frac{w_n}{nk_n}\right]$$

$$\leq \sum_{j=1}^{k_n} P\left\{|\hat{a}_{jn}| > \sqrt{\frac{w_n}{nk_n}}\right\}$$

or

$$P\left\{|\hat{a}_{jn}| > \sqrt{\frac{w_n}{nk_n}}\right\} \leq \frac{E[\hat{a}_{jn}^2]^2}{w_n} nk_n$$

Comme

$$\hat{a}_{jn} = \frac{1}{n} \sum_{r=1}^n \int_{i-1}^i c_j(X_t) dt$$

On applique l'inégalité de Billingsley p 157 et on obtient

$$\alpha_n \leq (1 + 8M^2) \sum_{i=1}^{\infty} \alpha(i) \frac{k_n^2}{w_n}$$

donc

$$\alpha_n = O\left[\frac{k_n^2}{w_n}\right]$$

Remarque :

Dans cette partie, on retrouve les résultats pour un processus à temps discret obtenus au chapitre II. Corollaire 13

Sous les hypothèses du corollaire 12

$$| \alpha_n - P[ \| U_{k_n}(\cdot) \|^2 > w_n ] | = O[n^{\delta/4} \log(2+n)] \quad (75)$$

où  $-1/6 < \delta < 0$ .

Après ces quelques résultats obtenus pour un processus à temps continu ( $t \in [0, T]$ ) avec  $[T] = n$ ,  $n \in \mathbb{N}^*$ ; on pourrait généraliser au cas où  $T$  n'est plus entier.

On pose  $T = n\beta$  avec  $[T] = n$  ( $T \geq 1$ ).

De  $[T] \leq T < [T] + 1$ , on tire que :  $1 \leq \beta < 2$ .

Et

$$\left\{ \begin{aligned} \hat{a}_{jr} &= \frac{1}{T} \int_0^r e_j(X_t) dt \\ &= \frac{1}{n\beta} \int_0^{n\beta} e_j(X_t) dt \\ &= \frac{1}{n\beta} \sum_{i=1}^n \int_{(i-1)\beta}^{i\beta} e_j(X_t) dt \end{aligned} \right.$$

$Y_i^j$  introduit précédemment devient

$$Y_i^j = \frac{1}{\beta} \int_{(i-1)\beta}^{i\beta} e_j(X_t) dt$$

et donc,

$$\hat{a}'_{jT} = \frac{1}{n} \sum_{i=1}^n Y_i^{(j)}.$$

L'étude réalisée ci-dessus peut donc facilement s'appliquer à  $\hat{a}'_{jT}$ .

## Conclusion Générale

Il ressort de la synthèse faite sur les tests hilbertiens dans le cas des variables aléatoires indépendantes que beaucoup de résultats ont été obtenus qui permettent de montrer l'efficacité asymptotique et l'optimalité des tests construits.

Ainsi la loi limite de la statistique de test pour les hypothèses adjacentes à l'hypothèse nulle a été étudiée de même que les propriétés asymptotiques locales au voisinage de l'hypothèse nulle. On déduit des résultats relatifs à l'efficacité asymptotique de ces tests ainsi que des critères pour qu'un test soit asymptotiquement optimal.

L'introduction de la pente de Bahadur dans les tests hilbertiens a permis de prouver l'efficacité relative du test fonctionnel par rapport au test de Neyman-Pearson.

Des considérations théoriques montrent que l'utilisation des systèmes orthogonaux lisses permettent de construire des tests plus efficaces que le test usuel. Les simulations réalisées (p.p 32-35) confirment cette idée. L'étude des tests hilbertiens a donc permis de montrer leur supériorité par rapport au test du  $\gamma^2$  usuel. Le praticien aura donc avantage à utiliser ces tests.

Notre tentative de généralisation de certains de ces résultats au cas des processus stationnaires tout en donnant des résultats nouveaux pose le problème de l'optimalité des tests que nous avons construits. Une piste de recherche serait donc l'examen des conditions nécessaire et suffisante de convergence du test, car pour certains cas de noyau, nous n'avons obtenu que des conditions suffisantes de convergence, ceci à cause de l'introduction des suites  $(b_n)$  qui réalisent la convergence du test.

En ce qui concerne les tests fonctionnels dans un espace de Banach, les résultats obtenus appellent des généralisations au cas de

noyaux de dimensions infinies, de même que le problème de la loi de  $\sup_{t \in [0,1]} \left| \sum_{j=1}^k Z_j e_j(t) \right|$  est posé. Cela ne nous a pas permis de trouver la valeur du niveau du test. Nous espérons y revenir ultérieurement.

Pour le test de non-influence, les deux cas examinés, c'est-à-dire 1<sup>er</sup> cas les variables aléatoires  $(X_i, Y_i)_{1 \leq i \leq n}$  sont dépendantes; 2<sup>ème</sup> cas les v.a.  $(X_i, Y_i)_{1 \leq i \leq n}$  sont indépendantes nous ont permis de montrer que le test construit est consistant. Les simulations réalisées donnent les régions critiques du test «  $r(x) = 0$  » contre «  $r(x) \neq 0$  » en fonction de la dimension  $k$  du noyau. Il serait intéressant d'envisager de tester à la fois la loi de  $X$  et  $r(x) = 0$ , car dans notre étude nous avons supposé que la loi de  $X$  qui est  $\mu$  est connue et l'hypothèse  $H_0$  a porté sur «  $r(x) = 0$  ». La double hypothèse sera ainsi présentée par :

$$H'_0 = \begin{cases} P_x = \mu \\ r(x) = 0 \end{cases}$$

Ce qui nécessitera la recherche de la loi du couple  $(\|S_n\|^2, \|R_n\|^2)$ . Nous espérons y revenir.

Pour l'étude des tests non paramétriques de la densité marginale après avoir obtenu ces quelques résultats, on peut envisager considérer un processus  $(X_t, t \in [0, T])$  qui est observé aux temps  $\delta_n, 2\delta_n, \dots, n\delta_n$  où  $\delta_n \rightarrow 0$  quand  $n \rightarrow +\infty$  et  $T = T_n = n\delta_n \rightarrow +\infty$  quand  $n \rightarrow +\infty$ .

On réalise en fait une discrétisation des observations et on obtient  $X_{\delta_n}, X_{2\delta_n}, \dots, X_{n\delta_n}$ .

$\hat{a}_{jT}$  et  $Y_j^j$  deviennent donc respectivement  $\hat{a}_{jT}''$  et  $Y_j''$  donnés par

$$\left\{ \begin{aligned} \hat{a}_{jT}'' &= \frac{1}{n\delta_n} \int_0^{n\delta_n} e_j(X_t) dt \\ &= \frac{1}{n\delta_n} \sum_{i=1}^n \int_{(i-1)\delta_n}^{i\delta_n} e_j(X_t) dt \end{aligned} \right.$$

et

$$Y_i^{nj} = \frac{1}{\delta_n} \int_{(i-1)\delta_n}^{i\delta_n} e_j(X_t) dt$$

C'est à dire encore

$$\hat{\alpha}_{jT}'' = \frac{1}{n} \sum_{i=1}^n Y_i^{nj}.$$

L'étude des lois limites des statistiques introduites ci-dessus ainsi que la vitesse de convergence sont assez délicates du fait de l'introduction de  $\delta_n$ . Nous espérons y revenir.

Enfin, des simulations sur un autoregressif pourront être envisagées ultérieurement à partir des résultats obtenus à la proposition 29; ce qui nécessitera l'estimation de la matrice  $\Sigma$ .

# Bibliographie

- [1] Aerts M., G. Claseskens, J.D. Hart. Testing lack of fit in multiple regression. *Biometrika* 87 (2000), n° 2 p.405-424.
- [2] Ahmad I.A. Strong consistency of density estimation by orthogonal series methods for dependent variables with applications. *Ann. Inst. Stat. Math.*, 31 p.279-288 (1979).
- [3] Anderson T.W. The statistical Analysis of time series J.Wiley, New-York (1971).
- [4] Barra J.R. Notions fondamentales de statistique mathématique, Dunod, Paris (1971).
- [5] Bertaud M. Estimation du paramètre de décentrage dans les tests de regression, *Metron* 42 (1984) n° 1-2 ; p.27-34.
- [6] Best D.J. and J.C.W. Rayner. Are two classes enough for the  $\chi^2$  goodness of fit test ? *Statistica Neerlandica*, 35 p.157-163 (1981).
- [7] Best D.J. and J.C.W. Rayner. The choice of class probabilities and number classes for simple  $\chi^2$  goodness of fit test. *Sankhya* (1981).
- [8] Best D.J. and J.C.W. Rayner. Uniformity testing when alternatives have low order. *Sankhya* (1982)
- [9] Bickel P.J., M.Rosenblatt. On some global measure of the deviations of density function estimates. *Ann. Statist.* p 1071-1095 (1973)
- [10] Billingsley P. Convergence of Probability measures . J.Wiley, New-York (1968)
- [11] Billingsley P. Convergence of Probability measures, seconde Edition. J. Wiley, New-York (1999)

- [12] Blacher R. A new-Chi-squared independence test. Rapp. Rech. N° 512, IMAG (Grenoble) (1985)
- [13] Bosq D. Inégalités de Bernstein pour les processus stationnaires et mélangeants. Applications C.R.A.S., A t.281, p.1095-1098. (1975).
- [14] Bosq D. Tests d'ajustement hilbertiens. Publications UER Mathématiques Pures et Appliquées. Lille I n° 125 p.1-39 (1978).
- [15] Bosq D. Tests hilbertiens et tests du  $\chi^2$ ; C.R.A.S. , A. t.286 p.946-948. (1978).
- [16] Bosq D. Sur une classe de tests qui contient le test du  $\chi^2$ . Publication de l'I.S.U.P. Volume 25 Fasc.1-2. (1980)
- [17] Bosq D. Comportement asymptotique de tests hilbertiens de dimension infinie sous des hypothèses adjacentes. Publ.IRMA – Lille I, Vol 3, fasc.6, p.1-22 (1981).
- [18] Bosq D. Lois limites et efficacité asymptotique des tests hilbertiens de dimension finie sous des hypothèses adjacentes. Statistiques et Analyse des données; Vol.8, n° 1 p.1-40 (1983).
- [19] Bosq D. Tests du  $\chi^2$  Généralisés. Comparaison avec le test du  $\chi^2$  classique. Rev. Stat. Appli. Vol 27, n° 1 p.43-52 (1989)
- [20] Bosq D. Nonparametric statistics for stochastic processes- Lecture notes in statistics. Springer Verlag Berlin/New-York (1996)
- [21] Bosq D. Nonparametric statistic for stochastic processus. Second ed. Springer Verlag, New-York - Estimation and Prediction (1998).
- [22] Bosq D. Linear processes in function spaces-Theory and applications. Lecture note in statistics n° 1149.Springer-Verlag New-York (2000)
- [23] Bosq D. Functional tests of fit. Goodness of fit tests and model validity. Statistics of Industry and Technology. Birkhäuser Boston (2002)
- [24] Bosq D. et J.P. Lecoutre. Théorie de l'estimation fonctionnelle. Economie et Statistique avancée. Economica (1987)



- [25] Box D.E.P. and Watson G.S., Robustness to non-normality of regression tests *Biometrika* 49 p.93-106 (1962)
- [26] Box D.E.P. and Watson G.S. Corrigenda : “ Robustness to non-normality of regression tests” *Biomtrica* 52 (1965)
- [27] Bradley R.C. Basic properties of strong mixing conditions. Dependence in probability and statistics (Oberwolfach, 1985), Birkhäuser Boston, Boston, M.A. 1986, p.165-192.
- [28] Bradley R.C. On quantiles and the central limit question for strongly mixing sequences, *J. theor. Proba.* 10 (1977) n° 2 p.507-555, dedicated to Murray Rosenblatt.
- [29] Carbon M. Sur l'estimation asymptotique d'une classe de paramètre fonctionnels pour un processus stationnaire. Thèse 3ème cycle Lille I (1982)
- [30] Carbon M. Inégalités de grandes déviations dans les processus – Application à l'estimation fonctionnelle. Doctorat – Univ. Paris 6. (1988)
- [31] Centsov. Estimation of an unknown density function from observations. *Soviet. Math.* 3 p.1559-1562 (1962)
- [32] Csàki E. A relation between chungs and strassen's law of iterated logarithm. *Z. Wahr. Verw. Geb.* Springer Verlag (1980)
- [33] Davydov. Y. A Convergence of distributions generated by stationary stochastic processes. *Theor. Proba. Appl.* 13, 691-696 (1968)
- [34] Delecroix M. Sur l'estimation des densités d'un processus stationnaire à temps continu. *Publ. Inst. Stat. Univ. Paris 6* (1-2) p.17-40 (1980)
- [35] Diebolt J., Posse P. On the density of the maximum of smooth Gaussian processes. *Ann. Prob.* Vol. 24 n° 3 p.1104-1129 (1996).
- [36] Doob J.L. *Stochastic processes.* Wiley, New-York. (1967)
- [37] Doukhan P. *Mixing properties and examples.* Springer-Verlag, New-York (1994)
- [38] Essen C.G. Fourier analysis of distribution functions. *Acta. Math.* 77, pp.1-125 (1945)
- [39] Fourgeaud C. and Fuchs A. *Statistique.* Dunod (1967)

- [40] Frenay A. Sur l'estimation de la densité marginale d'un processus à temps continu par projection orthogonale. Pub. Inst. Stat. Univ. Paris – Vol 45 n°1 p.53-92 (2001)
- [41] Gadiaga D. Sur une classe de tests qui contient le test du  $\chi^2$  le cas d'un processus stationnaire. Thèse 3ème cycle Lille I. (1982)
- [42] Gadiaga D. Tests hilbertiens et test du  $\chi^2$  pour un processus stationnaire et mélangeant . CRAS t.296 p.171-174. (1983)
- [43] Gadiaga D. Tests hilbertiens pour un processus stationnaire et mélangeant. Preprint, soumis à la revue de l'I.S.U.P. (2002)
- [44] Gadiaga D. Tests du  $\chi^2$  Généralisés. Application au processus autoregressif. Preprint, soumis à la revue de Statistique Appliquée. (2002)
- [45] Gadiaga D. et Ignaccolo R. Tests de non influence. Preprint - L.S.T.A . Université Paris 6. (2002)
- [46] Gregory G. Large sample theory for U-statistics and test of fit. Ann. Stat., Vol. 5 n° 1 p.110-123 (1977)
- [47] Gregory G. On efficiency and optimality of quadratic tests. Ann. Stat. Vol.8 n° 1 p.116-131 (1980)
- [48] Grenander U. and Rosenblatt R. Statistical analysis of stationary time series Wiley, New-York (1957)
- [49] Hart J.D. Nonparametric smoothing and lack-of-fit tests. Springer Verlag, New-York (1997)
- [50] Ibragimov I.A. and Linnik I.U. Independent and stationary sequences of random variables. Nauka, Moscow (1965).
- [51] Ignaccolo R. Goodness of fit test for dependent observations Prépublication, (2002)
- [52] Ignaccolo R. Tests d'ajustement fonctionnels par des observations corrélées. Thèse de Doctorat présentée à l'univ Paris6 et univ Milan. Nov 2002.
- [53] Izraelewitch , Laffite J., Larault Z., Roubert B. Le test du  $\chi^2$  et le test de Legendre Projet ISUP – Paris 6 (1988)
- [54] N.C. Jain, M.B.Marcus : Central limit theorem for  $C(S)$ -valued random variables. J funct.anal.19, 216-231 (1975).

- [55] Kallenberg W.C.C.N., Oosterhoof J., Schiever B.F.S. The number of classes in chi-squared goodness of fit test. *JASA* 80, n° 392 p.959-968 (1985).
- [56] Kendal M.G. The advanced theory of statistics. Vol.2 Charles Griffin, Londres (1973)
- [57] Leadbetter M.R., Georg Lindgren, Holger Rootzen. Extremes and related properties of random sequences and processes. Springer Series in Statistics. Springer Verlag (1983)
- [58] Leblanc F. Estimation par ondelettes de la densité marginale d'un processus stochastique. Publ. Paris – (1995)
- [59] Lee G. and Hart J.D. Model selection criteria with data dependent penalty, with applications to data driven. Neyman smooth tests. *J. Nonparametric. Stat.*, 12 n° 5 p.683-707 (2000)
- [60] Lifshits M.A. Gaussian random functions mathematics and its applications. Kluwer Academic Publishers (1995)
- [61] Mann H.B. and Wald A. On the choice of class intervals in the application of the chi-squared test. *Ann. Math. Stat.*, 13 p.306-317 (1942).
- [62] Nadaraja E.A. A quadratic measure of the deviation of a density estimation. *Theo. Prob. Appli.* 21 p.843-850 (1976)
- [63] Neveu J. Bases mathématiques du calcul des probabilités. Masson et Cie (1970)
- [64] Neveu J. Processus aléatoires gaussiens. Presses de l'Université de Montréal Canada (1968)
- [65] Neyman J. Smooth test for goodness of fit. *Skandd, Aktman*, 20 p.119-128 (1937)
- [66] Y.Nikitin. Asymptotic efficiency of non parametric tests. Cambridge Univ. Press (1995)
- [67] Pearson K. On the criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical magazine, series 5*, 50, p.157-175 (1900).
- [68] Riesz et Nagy Leçons d'analyse fonctionnelle. Gauthier Villars (1955)

- [69] Rio E. Théorie asymptotique des processus faiblement dépendants. Springer Verlag – Berlin (2000)
- [70] Rosenblatt M. A central limit theorem and a strong mixing condition. Proceeding Vol 42 p.42-47 (1956)
- [71] Roussas G. Contiguity of Probability measures. Cambridge Univ. Press (1978)
- [72] Sazonov V.V. On the multidimensional Central Limit Theorem. Shankhya A., 30 p.192-204 (1968)
- [73] Sazonov V.V. On W2 criterion. Shankhya A., 30 p.205-210 (1968)
- [74] Schwartz S.C. Estimation of probability density by a orthogonal series. Ann. Math. Stat. 38 p.1261-1265 (1967)
- [75] Shiryaev A.N. Probability. Springer-Verlag New-York (1996)
- [76] Tikhomirov A.N. Convergence rate in the central limit theorem for weakly dependent randomvariables. Theeor. Veroyatnost Primenen, 25 n° 4 p.800-818 (1980)
- [77] Van der Vaart A.W. Asymptotic statistics. Cambridge Univ. Press (1998)
- [78] Yakchew A.J. Nonparametric regression tests based on least squares. Econometric theor. 8 n° 4 p.435-451 (1992)
- [79] Walter Ph. The central limit problem for the mixing sequences of random variables Z. Wahr. Verw. Geb. P.155-171 (1969)
- [80] Watson G.S. Density estimation by orthogonal series. Ann. Math. Stat. 40 p.1496-1498 (1969).

**Annexe :**  
**Curriculum Vitae**

## CURRICULUM VITAE

*Dembo GADIAGA*

*Statisticien/Probabiliste*

- Docteur 3<sup>ème</sup> cycle en mathématiques appliquées  
Option probabilités statistiques
- Maître assistant de mathématiques appliquées :  
UFR/SEG et UFR/SEA(Département de mathématique  
et Informatique) Université de Ouagadougou

Date de naissance : 1950 à Thiou, province du Yatenga

Nationalité : Burkinabe.

### Etudes secondaires

-1964 à 1972

Diplôme : Baccalauréat de l'Enseignement du 2<sup>ème</sup> degré (BAC C) mention assez bien, Juin 1972

### Etudes universitaires

-1972 à 1974 : Université de Niamey, République du Niger

Diplôme : DUESS Mathématiques et Physiques (M-P) mention assez bien.

-1974 à 1982 : Université de Lille I – France

Diplôme

1982 Thèse de Doctorat 3<sup>ème</sup> cycle  
option probabilités statistiques  
mention très honorable.

### Associations scientifiques

- Membre du CEDRES.UFR/SEG – Université de Ouagadougou.
- Directeur Adjoint du CEDRES (Centre d'Etude, de Documentation, de Recherches Economiques et Sociales) 1983 à 1986.

- Secrétaire Général de la CIERRA : Conférence des Institutions d'Enseignement et de Recherche Economique en Afrique de 1987 à 1989.
- Membre de la Société Française de Statistique.

Travaux scientifiques

1) sur le plan mathématiques

a) Publications

- Lois et estimations du spectre en Juin 1978

Mémoire de DEA Université Lille I.

- Sur une classe de tests qui contient le test du  $\chi^2$  : le cas d'un processus stationnaire/ Octobre 1982.

Thèse de 3<sup>ème</sup> cycle – Université Lille I.

- Tests hilbertiens et test du  $\chi^2$  pour un processus stationnaire et mélangeant.

CRAS t.296 p.171-174.

b) Publications en cours.

- Tests hilbertiens pour un processus stationnaire et mélangeant, preprint, soumis à la revue de L'I.S.U.P. –Juin 2002.

-Tests du  $\chi^2$  Généralisés : applications au processus autorégressif, preprint, soumis à la revue de Statistique Appliquée. Institut Henri Poincaré (I.H.P.) Juillet 2002.

- Tests de non influence, en collaboration avec R.Ignaccolo, preprint L.S.T.A.

c) Ouvrage didactique.

Statistiques appliquées : Inférence statistique. Introduction aux modèles linéaires à une équation.

Cours de statistiques intégré à la formation des économistes. Ouvrage destiné aux étudiants du 2<sup>ème</sup> cycle.

Edité par le CEDRES – Université de Ouagadougou – Aout 1987.

2) monographie – Recherche appliquée en Economie.

Travaux de recherche réalisés sur le terrain par des enquêtes.

- Les artisans ruraux formés de la boucle de la Volta Noire : impact socio-économique en collaboration avec Thiombiano. Edité par le CEDRES - Université de Ouagadougou -Septembre 1983.
- Problématique de Développement des Industries rurales : quelques études de cas au Burkina Faso. En collaboration avec T.H. Kaboré. Edité par le CEDRES -Université de Ouagadougou- Juin 1990

#### Missions d'enseignement invitées

- Mission d'enseignement AUPELF (Association des Universités Partiellement ou Entièrement de Langue française) à l'Université du Benin, Cotonou (I.N.E.). Cours aux élèves Ingénieurs Statisticiens Economistes portant sur le processus stochastique en temps discret. Juillet 1990
- Mission d'enseignement de l'Association des Universités partiellement ou entièrement de langue française. (AUPELF) Université du Bénin. Cotonou. Cours aux élèves Ingénieurs Statisticiens Economistes portant sur le processus stochastique en temps discret. Février 1991.

#### Expérience professionnelle – Responsabilités administratives.

- à partir de 1979, assistant de Mathématiques appliquées à l'ESSEC. Université de Ouagadougou :
- Chef du département : analyse économique et technique quantitative, de 1982 à 1983.
- Responsable de la Cellule Technique quantitative : statistique, mathématique, économétrie, à partir d'octobre 1983.
- Directeur Adjoint du CEDRES de 1983 à 1986. (Centre d'Etude de Documentation et de Recherches Economiques et sociale)
- Directeur de l'ESSEC de 1986 à 1988. (Ecole Supérieur des Sciences Economiques)
- Secrétaire Générale de la CIEREA. (Conférence des Institutions d'Enseignement et de Recherche Economique en Afrique)
- Vice Doyen aux affaires académiques (VDA) FASEG (Faculté des Sciences Economiques et de Gestion), 1997-2000.
- Participation à la réalisation du projet – conservation des eaux, des sols, et agroforesterie CES/AGF – Etude de base – 1989, projet exécuté par la SEPIA (Société d'Etudes de Projets d'Investissement en Afrique).



- Participation à la réalisation d'une étude portant sur les projections des populations au Burkina de 1990 à 2005, étude réalisée par la SEPJA.
- Membre d'équipes de recherches dans le cadre du CEDRES : projets portant sur :
  - artisanat rural dans la boucle de la Volta Noire, 18 mois (1981-1982)
  - technologies villageoises de 1985 à 1991.
  - CEDRES/AGRISK sur le problème de sécurité alimentaire au Burkina de 1985 à 1990.
  - Gestion des usages conflictuels des ressources en eau au Burkina. 1999-2000

Séminaires et conférences Internationaux invités.

- Novembre 1982 , participation à la conférence organisée par le PNUD et l'ONED sur la promotion de l'Industrie en Afrique, à Brazzaville.
- Juillet 1984, séminaire International en Analyse fonctionnelle (Institut de Recherche en mathématique appliquée) organisé par l'IRMA – Université Nationale de Côte d'Ivoire, Abidjan.
- Février 1986, 1<sup>ère</sup> conférence des institutions d'enseignement et de recherche en Sciences Economiques d'Afrique – 10<sup>ème</sup> anniversaire du CEDRES à Ouagadougou suivi d'un colloque des Economistes.
- Avril 1987, 2<sup>ème</sup> conférence des Institutions d'enseignement et de recherche en Sciences Economiques d'Afrique-10<sup>ème</sup> anniversaire du CEDRES à Ouagadougou suivi d'un colloque des Economistes
- Février 1988, Niamey, Séminaire organisé par la CIEREA sur l'intégration africaine, cas de la CEDEAO.
- Avril 1988, séminaire sur la promotion de l'expertise nationale et le rôle des cadres à Berlin Ouest : séminaire organisé par la Banque Mondiale. Allemagne de l'Ouest.
- Septembre 1994, atelier pour la mise en œuvre des programmes d'enseignement du PTCI (Programme de troisième cycle inter-universitaire à Bengerville (côte d'Ivoire).
- 9-12 mars 1996 – 9<sup>ème</sup> conférence de la CIEREA à Dakar.
- 11 au 13 mars 1998 Atelier sur l'harmonisation des programmes des 1<sup>er</sup> et 2<sup>ème</sup> cycle universitaires par le PTCI à Abidjan.