

Université de Montréal

Développements théoriques et empiriques des tests  
lisses d'ajustement des modèles ARMA vectoriels

par

**Gabriel Desrosiers**

Département de mathématiques et de statistique  
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales  
en vue de l'obtention du grade de  
Maître ès sciences (M.Sc.)  
en Statistique

18 décembre 2020



# Sommaire

---

Lors de la validation des modèles de séries chronologiques, une hypothèse qui peut s'avérer importante porte sur la loi des données. L'approche préconisée dans ce mémoire utilise les tests lisses d'ajustement. Ce mémoire apporte des développements théoriques et empiriques des tests lisses pour les modèles autorégressifs moyennes mobiles (ARMA) vectoriels. Dans des travaux précédents, Ducharme et Lafaye de Micheaux (2004) ont développé des tests lisses d'ajustement reposant sur les résidus des modèles ARMA univariés. Tagne Tatsinkou (2016) a généralisé les travaux dans le cadre des modèles ARMA vectoriels (VARMA), qui s'avèrent potentiellement utiles dans les applications avec données réelles. Des considérations particulières au cas multivarié, telles que les paramétrisations structurées dans les modèles VARMA sont abordées.

Les travaux de Tagne Tatsinkou (2016) sont complétés selon les angles théoriques et des études de simulations additionnelles sont considérées. Les nouveaux tests lisses reposent sur des familles de polynômes orthogonaux. Dans cette étude, une attention particulière est accordée aux familles de Legendre et d'Hermite. La contribution théorique majeure est une preuve complète que la statistique de test est invariante aux transformations linéaires affines lorsque la famille d'Hermite est adoptée. Les résultats de Tagne Tatsinkou (2016) représentent une première étape importante, mais ils sont incomplets quant à l'utilisation des résidus du modèle.

Les tests proposés reposent sur une famille de densités sous les hypothèses alternatives d'ordre  $k$ . La sélection automatique de l'ordre maximal, basée sur les résultats de Ledwina (1994), est discutée. La sélection automatique est également implantée dans nos études de simulations.

Nos études de simulations incluent des modèles bivariés et un modèle trivarié. Dans une étude de niveaux, on constate la bonne performance des tests lisses. Dans une étude de puissance, plusieurs compétiteurs ont été considérés. Il est trouvé que les tests lisses affichent des propriétés intéressantes de puissance lorsque les données proviennent de modèles VARMA avec des innovations dans la classe de lois normales contaminées.

*Titre court:* Tests lisses d'ajustement des modèles ARMA vectoriels.

*2010 Mathematics Subject Classification:* Primaire: 62M10; Secondaire: 62M15

*Mots clés:* Test lisse d'ajustement; Modèle VARMA; Test de normalité; Polynômes d'Hermite; Polynômes de Legendre; Procédure de sélection automatique; Invariance.

## Summary

---

When validating time series models, the distribution of the observations represents a potentially important assumption. In this Master's Thesis, the advocated approach uses smooth goodness-of-fit test statistics. This research provides theoretical and empirical developments of the smooth goodness of fit tests for vector autoregressive moving average models (VARMA). In previous work, Ducharme and Lafaye de Micheaux (2004) developed smooth goodness-of-fit tests designed for the residuals of univariate ARMA models. Later, Tagne Tatsinkou (2016) generalized the work within the framework of vector ARMA (VARMA) models, which prove to be potentially useful in real applications. Structured parameterizations, which are considerations specific to the multivariate case, are discussed.

The works of Tagne Tatsinkou (2016) are completed, according to theoretical angles, and additional simulation studies are also considered. The new smooth tests are based on families of orthogonal polynomials. In this study, special attention is given to Legendre's family and Hermite's family. The major theoretical contribution in this work is a complete proof that the test statistic is invariant to linear affine transformations when the Hermite family is adopted. The results of Tagne Tatsinkou (2016) represent an important first step, but they were incomplete with respect to the use of the model residuals.

The proposed tests are based on a family of densities under alternative hypotheses of order  $k$ . A data driven method to choose the maximal order, based on the results of Ledwina (1994), is discussed. In our simulation studies, the automatic selection is also implemented.

Our simulation studies include bivariate models and a trivariate model. In the level study, we can appreciate the good performance of the smooth tests. In the power study, several competitors were considered. We found that the smooth tests displayed interesting

power properties when the data came from VARMA models with innovations in the class of contaminated normal distributions.

*Short title:* Smooth goodness-of-fit tests in vector ARMA models

*2010 Mathematics Subject Classification:* Primary: 62M10; Secondary: 62M15

*Keywords:* Smooth goodness-of-fit tests; VARMA models; Non-normality test; Hermite polynomials; Legendre polynomials; Automatic selection; Invariance

# Table des matières

---

<b>Sommaire</b> .....	iii
<b>Summary</b> .....	v
<b>Liste des tableaux</b> .....	xi
<b>Liste des figures</b> .....	xiii
<b>Remerciements</b> .....	xv
<b>Chapitre 1. Introduction</b> .....	1
<b>Chapitre 2. Préliminaires</b> .....	7
2.1. Processus stationnaire .....	7
2.1.1. Cas univarié .....	7
2.1.2. Cas vectoriel .....	11
2.2. Formulation, propriétés et modèle ARMA .....	12
2.3. Modèle VARMA .....	15
<b>Chapitre 3. Statistiques de test</b> .....	17
3.1. Le principe du test de Neyman .....	18
3.1.1. Construction du test lisse de Neyman .....	18
3.1.2. Dérivation de la statistique de test .....	19
3.1.3. Le choix de l'ordre $k$ .....	21
3.2. Tests de normalité .....	22

3.2.1.	Coefficient d'asymétrie .....	23
3.2.2.	Coefficient d'aplatissement.....	24
3.2.3.	Test de Jarque-Bera.....	25
<b>Chapitre 4.</b>	<b>Nouveaux tests .....</b>	<b>27</b>
4.1.	Préliminaires .....	27
4.2.	Tests lisses d'ajustement pour le modèle VARMA.....	29
4.3.	Polynômes d'Hermite.....	34
4.4.	Polynômes de Legendre.....	36
4.5.	Sélection du $R$ automatique .....	38
<b>Chapitre 5.</b>	<b>Sur l'invariance du test lisse avec les polynômes d'Hermite.</b>	<b>41</b>
5.1.	Notation .....	41
5.1.1.	Polynômes orthogonaux multivariés généralisés .....	44
5.1.2.	Polynômes orthogonaux à une densité quelconque.....	44
5.1.3.	Polynômes orthogonaux d'Hermite multivariés .....	46
5.2.	Preuve de l'invariance de $U_k$ .....	49
5.2.1.	Preuve du théorème 4.3.1.....	51
<b>Chapitre 6.</b>	<b>Simulations.....</b>	<b>55</b>
6.1.	Modèles de simulations .....	55
6.2.	Tests pour les simulations .....	57
6.3.	Niveaux empiriques.....	58
6.4.	Puissances empiriques.....	59
<b>Conclusion</b>	<b>.....</b>	<b>75</b>



<b>Annexe A</b> .....	77
Développement des matrices $A_{m,R}$ et $B_{m,R}$ .....	77
Développement de la matrice $A_{m,R}$ de l'équation (4.2.12) .....	77
<b>Annexe B</b> .....	87
Preuve du théorème 5.1.2 .....	87
<b>Bibliographie</b> .....	91



## Liste des tableaux

---

4.1	Indices multiples ordonnés pour $m = 4$ et $k = 3$ .....	30
6.1	Matrices des paramètres des modèles de séries chronologiques pour les modèles $DGP_1, DGP_2, DGP_3$ et $DGP_4$ .....	56
6.2	Matrices des variances et des covariances des bruits blancs pour les modèles $DGP_1, DGP_2, DGP_3$ et $DGP_4$ .....	56
6.3	Niveaux empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 %.....	58
6.4	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_1$ .....	62
6.5	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_1$ avec des normales contaminées. 63	
6.6	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_1$ avec des ARCH.....	64
6.7	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_2$ .....	65
6.8	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_2$ avec des normales contaminées. 66	
6.9	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_2$ avec des ARCH.....	67

6.10	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_3$ .....	68
6.11	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_3$ avec des normales contaminées. 69	
6.12	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_3$ avec des ARCH. ....	70
6.13	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_4$ .....	71
6.14	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_4$ avec des normales contaminées. 72	
6.15	Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle $DGP_4$ avec des ARCH. ....	73

# Liste des figures

---

2.1 Bénéfices trimestriels par action de Johnson & Johnson, 84 trimestres, de 1960-I à 1980-IV..... 8



# Remerciements

---

Je veux remercier mon directeur de recherche, Monsieur Pierre Duchesne, pour sa confiance, son aide et son support.

Je veux aussi remercier mes parents et les fonds de recherche de mon directeur pour le soutien financier.

Je remercie Roxanne et Christiane d'avoir apporté des commentaires constructifs à ce projet.





# Chapitre 1

---

## Introduction

Un aspect central en statistique porte sur la modélisation. Parmi les objectifs importants se trouve l'inférence statistique (intervalles de confiance et tests d'hypothèses), menant à l'estimation de paramètres, aux choix des variables importantes et au calcul de prévisions. Un autre aspect qui peut s'avérer important repose sur les hypothèses de la distribution des données. Il est possible de tester si les données disponibles semblent compatibles avec une certaine loi. Choisir correctement la loi des données peut faciliter l'inférence, par exemple celle reposant sur la méthode du maximum de vraisemblance. Il est alors possible d'estimer efficacement les paramètres du modèle.

En statistique classique, supposons que l'on dispose d'un échantillon aléatoire de taille  $n$ , que l'on peut noter  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . L'hypothèse nulle pourrait prendre la forme suivante :

$$H_0 : F(\mathbf{x}) = F_0(\mathbf{x}, \boldsymbol{\theta}),$$

où  $\boldsymbol{\theta} \in \Theta$  avec  $\Theta$ , l'espace des paramètres sous  $H_0$  et  $F$  est une fonction de répartition. Dans l'hypothèse nulle citée précédemment,  $F_0$  est une fonction de répartition connue, mais reposant sur  $\boldsymbol{\theta}$ , un vecteur de paramètres. De façon générale, l'hypothèse alternative prendra la forme suivante :

$$H_1 : F(\mathbf{x}) = G_0(\mathbf{x}, \boldsymbol{\omega}),$$

avec  $G_0$  la fonction de répartition sous l'alternative, où  $\boldsymbol{\omega} \in \Omega$  avec  $\Omega$ , l'espace des paramètres sous  $H_1$ . Ici aussi,  $\boldsymbol{\omega}$  peut être un paramètre vectoriel. Cette façon générale de

poser le problème de test se trouve dans Rayner, Thas et Best (2009). Ainsi, sous l'hypothèse alternative  $H_1$ , n'importe quelle autre distribution pourrait être spécifiée, différente de celle sous l'hypothèse nulle.

Une hypothèse nulle particulièrement importante porte sur la normalité. Plus spécifiquement, si  $F_0$  est la fonction de répartition d'une loi normale. Cette hypothèse nulle s'avère en fait fondamentale, car beaucoup de procédures statistiques permettent de faire de l'inférence exacte lorsque les données suivent une loi normale. Ceci ne veut pas dire que le rejet de l'hypothèse invalide forcément les inférences, compte tenu du théorème limite central par exemple, ou encore puisque certaines méthodes sont plus ou moins robustes à l'hypothèse de normalité. Cependant, si la normalité est correctement présumée, plusieurs procédures jouissent de propriétés d'optimalité.

Historiquement, il semblerait que le premier test d'ajustement soit celui de Pearson (1900). Le  $\chi^2$  de Pearson est un des tests statistiques les plus utilisés. Plackett (1983) qualifie l'article de Pearson comme un des plus grands monuments statistiques du 20<sup>ième</sup> siècle. Supposons qu'il y a  $m$  classes : ces tests consistent à détecter la différence entre le nombre dans les classes prévues avec celles observées. Ce test est détaillé dans le chapitre 2 de Rayner, Thas et Best (2009).

Historiquement, les tests lisses d'ajustement ont été proposés par Neyman (1937). Le principe de cette méthodologie est de projeter la densité sous  $H_0$  dans un espace orthogonal à la densité testée. Donc, si les données sont orthogonales à la densité testée, la statistique de test sera proche de zéro et l'hypothèse nulle aura tendance à ne pas être rejetée. Un élément important de cette approche est que les variables aléatoires continues peuvent être transformées en variables aléatoires possédant une loi uniforme en utilisant la transformation reposant sur la fonction de répartition. Le test original de Neyman (1937) utilise une famille orthogonale à la loi uniforme, soit les polynômes de Legendre. Le test possède des propriétés intéressantes pour des petits échantillons. Pour compléter ces propriétés, ce test fut obtenu avec la fonction dite du score afin de dégager les propriétés asymptotiques du test du score. Ces éléments sont discutés plus profondément dans la section 4.2 de Rayner, Thas et Best (2009).

L'approche préconisée dans ce mémoire repose sur l'étude des tests lisses. Cependant, la littérature sur les tests de normalité est vaste et plusieurs autres approches ont été considérées. Mardia (1970) a introduit un test basé sur le coefficient d'asymétrie et un autre basé sur le coefficient d'aplatissement, soient respectivement le troisième et le quatrième moment. Ces quantités sont estimées et comparées à ce que l'on devrait obtenir si les données provenaient de la loi normale. Il présente la version univariée et la version multivariée. Jarque et Bera (1980, 1987) ont proposé un test basé sur les coefficients d'asymétrie et d'aplatissement. Ce test est souvent utilisé en vertu de la facilité à le programmer, puisqu'il requiert seulement d'estimer les troisième et les quatrième moments. Ces méthodes possèdent de bonnes propriétés de puissance, sauf naturellement dans les cas où le troisième et le quatrième moments sont similaires à ce qui serait obtenu si les données provenaient d'une loi normale. Ces tests seront détaillés dans le chapitre 3. D'autres tests tels que ceux de Kolmogorov-Smirnov et Anderson-Darling sont basés sur les fonctions de répartition empiriques. Henze (2002) a également étudié ces tests.

Terui et Imano (2003) ont développé un test de normalité dans le contexte des séries chronologiques multivariées. Ils ont adapté le test de Kariya, Tsay, Terui et Li (1999) en version multivariée. Ce test est similaire à celui présenté dans ce mémoire puisqu'il utilise les polynômes d'Hermite. Cependant, le test de Terui et Imano (2003) se limite au produit de deux polynômes d'Hermite (bivarié) et se fait directement sur les données. Le test présenté dans ce mémoire sera plus complexe dans la mesure où il est valide quand il repose sur des résidus de modèles.

Ducharme et Lafaye de Micheaux (2004) ont proposé un test de normalité pour l'innovation dans un modèle ARMA avec une moyenne connue. Ce test est basé sur l'approche du test lisse de Neyman (1937) et l'approche de la sélection automatique de Ledwina (1994). L'estimation des paramètres du modèle se fait avec la méthode du maximum de vraisemblance. Ils ont montré que le test avait plusieurs avantages dans le cas indépendant et identiquement distribué, autant théoriques qu'empiriques. Aussi, leurs simulations montrent que le test est performant en terme de puissance.

Le nouveau test effectué dans ce mémoire repose sur les résidus dans le contexte de séries chronologiques vectorielles. Une question naturelle porte sur la validité des procédures calculées sur les résidus, lorsqu'elles sont justifiées et étudiées sur des données brutes (c'est-à-dire sur des réalisations finies de procédures stationnaires). Il est possible de se demander s'il est juste de travailler avec des résidus qui reposent sur l'étape d'estimation et qui typiquement sont des variables aléatoires dépendantes. Pierce (1985) montre que tous les tests de normalité basés sur les résidus d'un modèle autorégressif ont la même distribution que les tests basés sur des observations indépendantes et identiquement distribuées. Ceci suggère que les tests de normalité peuvent être appliqués sous certaines conditions sur des résidus.

Le but de ce mémoire est d'adapter le test de Ducharme et Lafaye de Micheaux (2004) dans le contexte de modèles ARMA vectoriels ou VARMA. Les motivations sont aussi de poursuivre les travaux du chapitre 4 de la thèse de Tagne Tatsinkou (2016). Le test adapte le principe du test lisse de Neyman (1937) sur les résidus d'un modèle VARMA. Les familles des polynômes d'Hermite et Legendre sont utilisées. Il est important d'adapter le test en version multivariée car la loi normale multivariée implique que les lois marginales sont normales, mais savoir uniquement que les marginales sont normales n'implique pas la multinormalité. Aussi, une approche consistant à appliquer des tests de normalité justifiés dans le cas univarié, composante par composante d'une série chronologique vectorielle, ne pourra détecter des alternatives à la multinormalité ne survenant pas dans les lois marginales normales. Plus précisément, exprimé avec des fonctions de copules, il est possible en général que les lois marginales soient normales, mais que la distribution conjointe ne soit tout de même pas multinormale. Dans de telles situations, des tests de normalité univariés à la normalité n'auront pas beaucoup de puissance.

Ce mémoire est divisé de la façon suivante : le chapitre 2 présente la théorie des séries chronologiques requises pour comprendre ce mémoire. Le chapitre 3 décortique le test lisse de Neyman (1937) et les tests de normalité de Mardia (1970) et Jarque et Bera (1980, 1987). Le chapitre 4 est une partie importante du mémoire. Il présente le test lisse d'ajustement sur les résidus dans le contexte d'un modèle VARMA. L'utilisation de la famille des polynômes d'Hermite et Legendre est présentée ainsi qu'une sélection du

choix de l'ordre maximal selon l'approche de Ledwina (1994). Le chapitre 5 contient une contribution théorique originale, complétant la démarche de Tagne Tatsinkou (2016), et présentant la preuve de l'invariance du test lisse utilisant les polynômes d'Hermite. Le chapitre 6 compare les différents tests de normalité à l'aide de simulations de type Monte Carlo. Quatre modèles sont utilisés afin d'évaluer la performance des tests sur leurs niveaux et puissances empiriques. Une conclusion sur les forces et faiblesses des expériences empiriques sera ensuite effectuée. Des pistes pour des travaux de recherche futurs sont aussi abordées.



# Chapitre 2

---

## Préliminaires

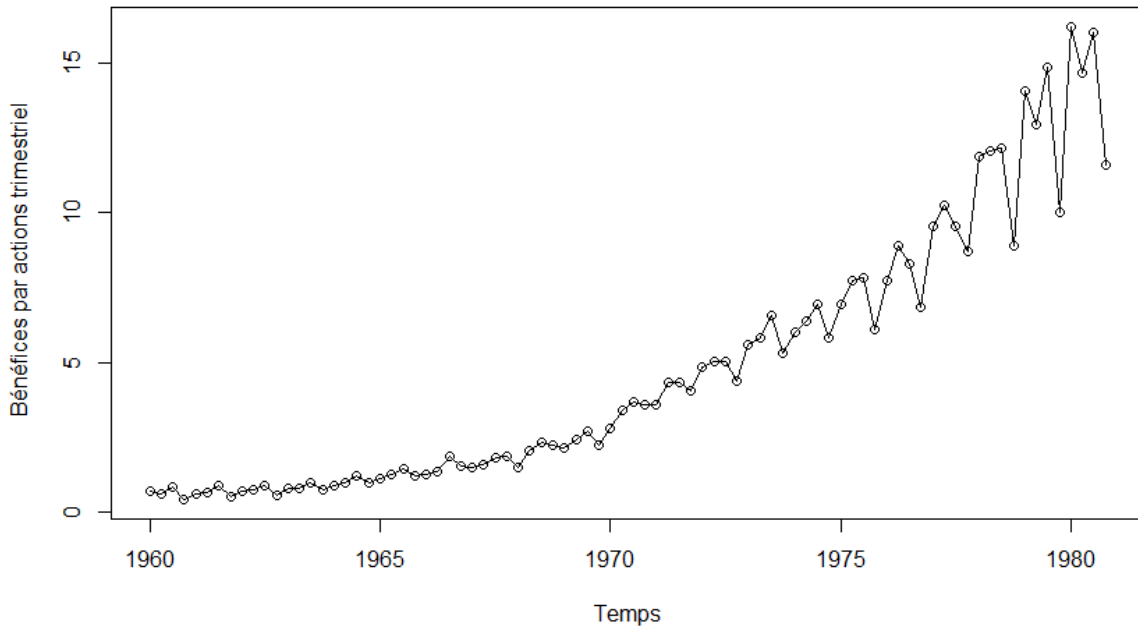
Une série chronologique est une séquence d'observations mesurée dans le temps. Dans les modèles que l'on va considérer, ces mesures sont effectuées à des moments équidistants (par exemple des jours, des mois ou des années). Ces séries trouvent naturellement des applications dans plusieurs domaines dont l'économie, les sciences sociales, l'épidémiologie et la finance, pour ne nommer que ces applications. L'objectif de ce chapitre est de définir ce qu'est un processus stationnaire, et de fournir les mesures statistiques les plus utiles pour les décrire. Les processus stationnaires dans les cas univarié et multivarié sont d'abord abordés. Ensuite, les classes des modèles ARMA et ARMA vectoriels sont définies.

### 2.1. Processus stationnaire

#### 2.1.1. Cas univarié

Une série chronologique est une réalisation finie d'un processus stochastique. Les concepts probabilistes dans l'étude des processus stochastiques demeurent utiles. Un concept important est celui de la stationnarité. Comme il sera démontré ultérieurement, certaines quantités possèdent des propriétés d'invariance, comme la fonction de moyenne et la fonction d'autocovariance. D'un point de vue pratique, on pourra alors considérer l'inférence statistique.

Pour bien illustrer une série chronologique, un exemple est présenté dans la figure 2.1. La série chronologique représente la série de Johnson & Johnson des bénéfices par action à chaque trimestre de 1960 à 1980 et cet exemple se trouve dans Shumway et Stoffer (2017, p.



**Figure 2.1.** Bénéfices trimestriels par action de Johnson & Johnson, 84 trimestres, de 1960-I à 1980-IV.

2). Les données sont exprimées en dollars sur 84 trimestres (correspondant ainsi à 21 ans). L'objectif est de savoir si la série possède une composante de tendance. Dans la figure 2.1, on peut apprécier une composante de tendance quadratique. Un modèle de séries chronologiques peut être plus approprié qu'un modèle de régression linéaire classique car les observations sont potentiellement corrélées. La définition 2.1.1 présente une première définition du concept de la stationnarité.

**Définition 2.1.1.** *Le processus  $\{X_t, t \in \mathbb{Z}\}$  est dit strictement stationnaire si les distributions conjointes de  $(X_{t_1}, \dots, X_{t_k})^T$  et  $(X_{t_1+h}, \dots, X_{t_k+h})^T$  sont identiques pour tout entier positif  $k$  et pour tout  $t_1, \dots, t_k, h \in \mathbb{Z}$ .*

Cette définition de stationnarité est peu conviviale, car il est difficile de vérifier cette condition avec des données. Le prochain type de stationnarité, la stationnarité au sens large, est plus facile à vérifier car elle est basée sur les deux premiers moments théoriques du processus.

**Définition 2.1.2.** *Le processus  $\{X_t, t \in \mathbb{Z}\}$  est stationnaire au sens large si  $\forall t, s :$*



- (i)  $\mathbb{V}(X_t) = \gamma_X(0, 0) < \infty$ , *indépendant de  $t$ ,*
  - (ii)  $\mathbb{E}(X_t) = \mu_X$ , *est constant en  $t$ ,*
  - (iii)  $\gamma_X(t, s) = \mathbb{Cov}(X_t, X_s)$ , *est fonction du délai  $t - s$ ,*
- où  $\gamma_X$  est la fonction d'autocovariance du processus  $\{X_t\}$ .

Lorsqu'un processus stochastique est stationnaire au sens strict avec les deux premiers moments finis, il est aussi stationnaire au sens large. L'application inverse n'est pas vraie, sauf dans le cas d'un processus gaussien où les deux définitions sont équivalentes.

Il est possible de définir le délai  $h$  tel que  $h = |t - s|$ , ce qui permet de travailler dans une distance de  $h$  et non seulement entre deux instants spécifiques. Il est alors possible de réécrire  $\mathbb{Cov}(X_t, X_{t+h}) = \gamma_X(t + h, t) = \gamma_X(h)$ . Notons que cette fonction est symétrique telle que  $\gamma_X(h) = \gamma_X(-h)$ . Si la série est stationnaire, il est possible de simplifier l'expression  $\gamma_X(t+h, t) = \mathbb{Cov}(X_h, X_0) = \gamma_X(h, 0)$ . La matrice d'autocorrélation  $\mathbf{\Gamma}_n$  peut être définie :

$$\mathbf{\Gamma}_n = \begin{pmatrix} \gamma_X(0) & \gamma_X(1) & \cdots & \gamma_X(n-1) \\ \gamma_X(1) & \gamma_X(0) & \cdots & \gamma_X(n-2) \\ \vdots & \vdots & & \vdots \\ \gamma_X(n-1) & \gamma_X(n-2) & \cdots & \gamma_X(0) \end{pmatrix}. \quad (2.1.1)$$

Tout comme la fonction d'autocovariance, la fonction d'autocorrélation d'un processus stationnaire est aussi définie pour le délai  $h$ , dans la définition 2.1.3.

**Définition 2.1.3.** *Soit  $\{X_t, t \in \mathbb{Z}\}$ , un processus stochastique stationnaire. La fonction d'autocorrélation de délai  $h$  est définie comme :*

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}, \quad h \in \mathbb{Z}. \quad (2.1.2)$$

Les propriétés théoriques sont intéressantes pour décrire un processus stochastique. Cependant, il est fréquent de disposer que d'un seul échantillon. L'hypothèse de stationnarité facilite grandement l'inférence. On utilise la moyenne et la fonction d'autocovariance échantillonnale de délai  $h$  afin d'estimer leurs équivalents théoriques :

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t. \quad (2.1.3)$$

Il est possible d'estimer l'autocovariance de délais  $h$  :

$$\hat{\gamma}_X(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}), \quad (2.1.4)$$

avec  $\hat{\gamma}_X(h) = \hat{\gamma}_X(-h)$  pour  $h = 0, 1, \dots, n-1$ . Le domaine d'estimation est limité en raison du terme  $x_{t+h}$  puisqu'il n'est pas disponible pour  $t+h > n$ . La fonction d'autocorrélation échantillonnage estimée de délai  $h$  s'écrit :

$$\hat{\rho}_X(h) = \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}, \quad |h| \leq n-1. \quad (2.1.5)$$

Dans les modèles de séries chronologiques, le terme d'erreur est appelé un bruit blanc. Le bruit blanc faible est constitué de variables aléatoires non corrélées. La définition 2.1.4 présente le bruit blanc fort où les variables aléatoires sont non corrélées et indépendantes.

**Définition 2.1.4.** Soit  $\{\epsilon_t, t \in \mathbb{Z}\}$ , un processus stochastique. Il sera dit *bruit blanc fort* si les trois conditions suivantes sont satisfaites :

- (i)  $\mathbb{E}(\epsilon_t) = 0, \quad \forall t \in \mathbb{Z},$
- (ii)  $\mathbb{E}(\epsilon_t^2) = \sigma^2, \quad \forall t \in \mathbb{Z},$
- (iii) Le processus  $\epsilon = \{\epsilon_t, t \in \mathbb{Z}\}$  est constitué de variables aléatoires indépendantes et identiquement distribuées (i.i.d).

La définition 2.1.5 présente le bruit blanc faible, où les composantes du processus sont seulement non-corrélées.

**Définition 2.1.5.** Un *bruit blanc* sera dit *faible* lorsque la condition (iii) de la définition 2.1.4 est remplacée par la condition (iv) :

- (iv)  $\text{Cov}(\epsilon_t, \epsilon_s) = 0, \quad \forall t \neq s.$

Le bruit blanc fort est noté  $\{\epsilon_t\} \sim IID(0, \sigma^2)$ , tandis que le bruit blanc faible est noté  $\{\epsilon_t\} \sim BB(0, \sigma^2)$ . Un processus linéaire peut s'écrire comme une combinaison linéaire d'un bruit blanc.

**Définition 2.1.6.** Le processus  $\{X_t\}$  est un *processus linéaire* s'il admet une *représentation* :

$$X_t = \mu + \sum_{j=-\infty}^{+\infty} \psi_j w_{t-j},$$

$\forall t \in \mathbb{Z}$ , avec  $\{w_t\} \sim BB(0, \sigma_w^2)$  et  $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$ .

La fonction d'autocovariance du processus  $\{X_t\}$  comme défini dans la définition 2.1.6 satisfait :

$$\gamma_X(h) = \sigma_w^2 \sum_{j=-\infty}^{+\infty} \psi_j \psi_{j+h}. \quad (2.1.6)$$

Le processus linéaire dépend du passé, présent ou futur. D'un point de vue de prévision, un processus s'exprimant en fonction du futur n'est pas très naturel, donc des modèles qui ne dépendront pas du futur sont considérés. Par exemple, un modèle causal avec  $\psi_j = 0$  lorsque  $j < 0$  est présenté.

La théorie univariée s'adapte naturellement au contexte multivarié à l'aide de matrices et de vecteurs. La théorie dans le cas vectoriel est présentée dans la prochaine section.

### 2.1.2. Cas vectoriel

Un processus de dimension  $m$  s'écrit  $\{\mathbf{X}_t\}$  où  $\mathbf{X}_t = (X_{t1}, X_{t2}, \dots, X_{tm})^T$ , soit  $m$  observations au temps  $t$ . Les définitions de l'espérance, de l'autocovariance et de la stationnarité sont généralisées au cas vectoriel.

**Définition 2.1.7.** *L'espérance du vecteur  $\mathbf{X}_t$  est définie par :*

$$\mathbb{E}(\mathbf{X}_t) = \boldsymbol{\mu}_t = (\mu_{t1}, \mu_{t2}, \dots, \mu_{tm})^T.$$

La matrice des variances et des covariances du processus fait l'objet de la prochaine définition.

**Définition 2.1.8.** *La fonction d'autocovariance matricielle entre  $\mathbf{X}_{t+h}$  et  $\mathbf{X}_t$  est définie par :*

$$\boldsymbol{\Gamma}(t+h, t) = \mathbb{E} [(\mathbf{X}_{t+h} - \boldsymbol{\mu}_{t+h})(\mathbf{X}_t - \boldsymbol{\mu}_t)^T] = (\gamma_{ij}(t+h, t))_{i,j=1}^m.$$

Tout comme dans le cas univarié, la stationnarité est un concept important. La définition 2.1.9 présente la stationnarité au sens large, aussi appelée stationnarité faible.

**Définition 2.1.9.** *Le processus  $\{\mathbf{X}_t, t \in \mathbb{Z}\}$  est stationnaire au sens large si :*

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}, \quad \forall t,$$

$$\Gamma(t+h, t) = \Gamma(h), \quad \forall t, h.$$

La matrice des variances et des covariances de  $\mathbf{X}_t$  est  $\Gamma(t, t) = \Sigma_X$ . Il est utile de définir le bruit blanc de façon analogue au cas univarié.

**Définition 2.1.10.** *Le vecteur  $\{\mathbf{Z}_t, t \in \mathbb{Z}\}$  est un bruit blanc de moyenne  $\mathbf{0}$  et de matrice de covariance  $\Sigma_Z$  et on écrit  $\{\mathbf{Z}_t\} \sim BB(\mathbf{0}, \Sigma_Z)$ , s'il est de moyenne  $\mathbf{0}$ , si  $\mathbf{Z}_t$  et  $\mathbf{Z}_s$  sont non corrélés,  $t \neq s$  tel que :*

$$\Gamma(h) = \begin{cases} \Sigma_Z, & \text{si } h = 0, \\ \mathbf{0}, & \text{sinon.} \end{cases}$$

La matrice  $\Sigma_Z$  est assumée non singulière. Le bruit blanc multivarié est souvent utilisé pour construire une variété de séries chronologiques multivariées, tel que le processus linéaire présenté dans la définition 2.1.11.

**Définition 2.1.11.** *Le processus  $\{\mathbf{X}_t, t \in \mathbb{Z}\}$  est un processus linéaire s'il admet la représentation suivante :*

$$\mathbf{X}_t = \sum_{j=-\infty}^{+\infty} \Psi_j \mathbf{Z}_{t-j},$$

$\forall t \in \mathbb{Z}$  où  $\{\mathbf{Z}_t\} \sim BB(\mathbf{0}, \Sigma_Z)$  et  $\Psi_j$  est une suite de matrices de taille  $m \times m$  telle que  $\sum_{j=-\infty}^{\infty} \|\Psi_j\| < \infty$ , avec  $\|\Psi_j\| = \sqrt{\text{tr}(\Psi_j^T \Psi_j)}$  étant la norme euclidienne de la matrice  $\Psi_j$  et  $\text{tr}(\cdot)$  la somme des éléments sur la diagonale de la matrice.

Sous certaines conditions, le modèle général autorégressif moyenne mobile peut s'écrire comme un processus linéaire. Ce modèle est introduit dans la prochaine section.

## 2.2. Formulation, propriétés et modèle ARMA

Avant de définir la classe des modèles ARMA, il est nécessaire d'introduire l'opérateur retard. Il est très utilisé dans l'étude des séries chronologiques pour alléger la notation et il peut être utile comme outil de modélisation afin d'enlever une tendance dans une série. Aussi, dans la méthodologie de Box et Jenkins, cet opérateur est utilisé afin d'éliminer les tendances non-désirées, grâce à la différenciation de la série. Cette méthode est étudiée plus profondément et avec plus de détails dans Box, Jenkins et Reinsel (2008). L'opérateur retard est défini dans la définition 2.2.1.

**Définition 2.2.1.** Posons  $B^0 = 1$ . Pour  $k \geq 1$ , l'opérateur retard est défini par :

$$\begin{aligned} BX_t &= X_{t-1}, \\ B^2X_t &= B(X_{t-1}) = X_{t-2}, \\ &\vdots \\ B^kX_t &= X_{t-k}. \end{aligned}$$

Un polynôme exprimé en fonction de l'opérateur retard  $B$  de degré  $k$  s'écrit comme suit :

$$P(B) = \sum_{i=0}^k a_i B^i.$$

La classe des modèles ARMA est présentée dans la définition 2.2.2.

**Définition 2.2.2.** Une série chronologique  $\{X_t, t \in \mathbb{Z}\}$  est un processus ARMA( $p, q$ ) si la série est stationnaire et si elle satisfait l'équation aux différences stochastique suivante :

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}, \quad (2.2.1)$$

avec  $\phi_p \neq 0$  et  $\theta_q \neq 0$ . Les innovations sont  $\{\epsilon_t\} \sim BB(0, \sigma_\epsilon^2)$ . Les paramètres  $\phi_1, \dots, \phi_p; \theta_1, \dots, \theta_q$  sont les coefficients du modèle. Un processus  $\{X_t, t \in \mathbb{Z}\}$  est dit ARMA( $p, q$ ) de moyenne  $\mu$  si  $\{X_t - \mu, t \in \mathbb{Z}\}$  est un processus ARMA( $p, q$ ).

En utilisant l'opérateur retard, l'équation précédente peut être réécrite de façon plus compacte :

$$\Phi(B)X_t = \Theta(B)\epsilon_t, \quad (2.2.2)$$

avec :

$$\begin{aligned} \Theta(B) &= 1 + \theta_1 B + \dots + \theta_q B^q, \\ \Phi(B) &= 1 - \phi_1 B - \dots - \phi_p B^p. \end{aligned}$$

Si  $p = 0$ , le processus est une moyenne mobile d'ordre  $q$ , notée MA( $q$ ). Si  $q = 0$ , le processus est autorégressif d'ordre  $p$ , noté AR( $p$ ). Dans l'équation (2.2.1), il est possible d'avoir des facteurs communs dans  $\Theta(B)$  et  $\Phi(B)$ , ce qui permet de réduire l'ordre du modèle. Il faut aussi avoir des facteurs différents car si tous les facteurs sont identiques, le modèle se réduira à  $X_t = \epsilon_t$ . Ce modèle est ainsi un bruit blanc, et aucune inférence

n'est possible. Les concepts de causalité et d'inversibilité sont définis pour le modèle ARMA( $p, q$ ).

**Définition 2.2.3.** *Le processus  $\{X_t, t \in \mathbb{Z}\}$ , ARMA( $p, q$ ) est causal s'il existe une suite de constantes  $\{\psi_j, j = 0, 1, \dots\}$  telle que  $\sum_{j=0}^{\infty} |\psi_j| < \infty$  et  $\forall t \in \mathbb{Z}$ :*

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}.$$

La causalité équivaut à la condition présentée dans le théorème 2.2.1, qui stipule que les racines de  $\Phi(z)$  doivent être à l'extérieur du cercle unitaire.

**Théorème 2.2.1.** *Le processus  $\{X_t, t \in \mathbb{Z}\}$ , ARMA( $p, q$ ) est causal si et seulement si :*

$$\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0, \quad \forall |z| \leq 1.$$

La définition 2.2.4 présente l'inversibilité, où le processus est représenté sous la forme autorégressive d'ordre infini.

**Définition 2.2.4.** *Le processus  $\{X_t, t \in \mathbb{Z}\}$ , ARMA( $p, q$ ) est inversible s'il existe une suite de constante  $\{\pi_j, j = 0, 1, \dots\}$  telle que  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  et  $\forall t \in \mathbb{Z}$  :*

$$\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

Tout comme dans le cas de la causalité, l'inversibilité correspond à la condition que les racines doivent être à l'extérieur du cercle unitaire, tel que présentée dans le théorème 2.2.2.

**Théorème 2.2.2.** *Le processus  $\{X_t, t \in \mathbb{Z}\}$ , ARMA( $p, q$ ) est inversible si et seulement si :*

$$\Theta(z) = 1 - \Theta_1 z - \dots - \Theta_p z^p \neq 0, \quad \forall |z| \leq 1.$$

Ces définitions et ces théorèmes sont disponibles dans le chapitre 3 de Brockwell et Davis (1991) et Shumway et Stoffer (2017). La causalité n'est pas une propriété du processus  $\{X_t\}$ , mais plutôt la relation entre les processus  $\{X_t\}$  et  $\{\epsilon_t\}$  de l'équation (2.2.1) dans la définition d'un processus ARMA. La même analogie peut se faire pour l'inversibilité. L'inversibilité est utile lors de l'étude des prévisions. Dans la prochaine section, la classe des modèles VARMA est présentée.

### 2.3. Modèle VARMA

Les séries chronologiques multivariées permettent non seulement d'établir la dépendance entre une composante de la série, mais aussi entre les composantes. Dans ce modèle, à un temps  $t$  donné,  $m$  valeurs sont observées et notées  $\mathbf{X}_t = (X_{t1}, X_{t2}, \dots, X_{tm})^T$ . La classe des modèles VARMA est présentée dans la définition 2.3.1.

**Définition 2.3.1.** *Le processus  $\{\mathbf{X}_t, t \in \mathbb{Z}\}$  est un processus VARMA( $p, q$ ) s'il est stationnaire et satisfait l'équation aux différences stochastique suivante :*

$$\mathbf{X}_t - \Phi_1 \mathbf{X}_{t-1} - \dots - \Phi_p \mathbf{X}_{t-p} = \mathbf{Z}_t + \Theta_1 \mathbf{Z}_{t-1} + \dots + \Theta_q \mathbf{Z}_{t-q},$$

avec  $\Phi_p \neq \mathbf{0}, \Theta_p \neq \mathbf{0}$  et  $\{\mathbf{Z}_t\} \sim BB(\mathbf{0}, \Sigma_Z)$  qui sont les innovations, les paramètres autorégressifs et de moyennes mobiles sont  $\Phi_1, \dots, \Phi_p; \Theta_1, \dots, \Theta_q$ , qui sont des matrices de dimensions  $m \times m$ .

Le processus  $\{\mathbf{X}_t, t \in \mathbb{Z}\}$  est un processus VARMA( $p, q$ ) de moyenne  $\boldsymbol{\mu}$  si  $\{\mathbf{X}_t - \boldsymbol{\mu}, t \in \mathbb{Z}\}$  est un processus VARMA( $p, q$ ). Comme dans le cas univarié, la côté de gauche est le côté autorégressif, tandis que le droit est le côté moyenne mobile.

En utilisant l'opérateur retard, l'équation précédente peut être réduite à :

$$\Phi(B)\mathbf{X}_t = \Theta(B)\mathbf{Z}_t,$$

avec :

$$\Theta(B) = \mathbf{I}_m + \Theta_1 B + \dots + \Theta_q B^q, \quad (2.3.1)$$

$$\Phi(B) = \mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p. \quad (2.3.2)$$

Si  $p = 0$ , le processus est une moyenne mobile d'ordre  $q$ , notée VMA( $q$ ). Si  $q = 0$ , le processus est autorégressif d'ordre  $p$ , noté VAR( $p$ ).

Dans le modèle univarié, la causalité et l'inversibilité sont définis dans les définitions 2.2.3 et 2.2.4. Les différences principales avec le cas multivarié sont que les coefficients  $\psi$  et  $\pi$  sont remplacés par des matrices  $\Psi$  et  $\Pi$  qui nécessitent que leurs sommes en valeur absolue dans un sens matriciel soient finies.

**Théorème 2.3.1.** *Le processus  $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ , VARMA( $p, q$ ) est causal si  $\det\{\Phi(z)\} \neq 0$  pour tout  $z \in \mathbb{C}$ ,  $|z| \leq 1$  et  $\sum_{j=0}^{\infty} \|\Psi_j\| < \infty$  où  $\|\cdot\|$  représente la norme euclidienne :*

$$\mathbf{X}_t - \boldsymbol{\mu} = \sum_{j=0}^{\infty} \Psi_j \mathbf{Z}_{t-j},$$

où  $\det\{\cdot\}$  est le déterminant d'une matrice. Le concept d'inversibilité pour la classe des modèles VARMA fait l'objet du théorème 2.3.2.

**Théorème 2.3.2.** *Le processus  $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ , VARMA( $p, q$ ) est inversible si  $\det\{\Theta(z)\} \neq 0$  pour tout  $z \in \mathbb{C}$ ,  $|z| \leq 1$  et  $\sum_{j=0}^{\infty} \|\Pi_j\| < \infty$  :*

$$\mathbf{Z}_t = \sum_{j=0}^{\infty} \Pi_j (\mathbf{X}_{t-j} - \boldsymbol{\mu}).$$

Ces théorèmes et leurs preuves se trouvent dans la section 11.3 de Brockwell et Davis (1991). Dans le prochain chapitre, le test lisse d'ajustement de Neyman (1937) ainsi que trois tests de normalité sont présentés.



# Chapitre 3

---

## Statistiques de test

Il existe deux grandes classes de tests d'ajustement. Les tests *omnibus* possèdent une bonne puissance contre toutes les alternatives, mais typiquement ils ne sont pas les tests les plus puissants. Ils ne donnent habituellement aucune indication sur les écarts de distribution entre l'hypothèse nulle et l'hypothèse alternative. Ces tests sont souvent basés sur la fonction de répartition empirique ou la fonction caractéristique empirique, tel que les tests de Kolmogorov-Smirnov, d'Anderson-Darling ou de Cramér-von Mises. Ainsi, la fonction de répartition empirique estime la fonction de répartition théorique dont les données sont issues, mais les formes ou les classes particulières de distributions ne sont pas prises en compte. Les tests *directionnels* sont très puissants contre des alternatives spécifiques mais ils ne sont pas forcément convergents pour un aussi grand nombre d'alternatives que les tests omnibus. Les tests directionnels donnent aussi une indication plus précise sur l'alternative. Par exemple, le rejet de l'hypothèse nulle d'un test de normalité basé sur le coefficient d'aplatissement indique que le coefficient d'aplatissement de la loi alternative s'éloigne de celui de la loi normale. Les tests lisses d'ajustement sont des tests partiellement omnibus, dans le sens que ces tests possèdent habituellement une bonne puissance pour de nombreuses distributions respectant certaines caractéristiques particulières (ou encore possédant une puissance non triviale dans certaines directions spécifiques), et ces procédures font que la puissance est concentrée vers une alternative donnée.

D'abord, le test lisse de Neyman est abordé. Les tests directionnels de normalité de Jarque-Bera et de Mardia, aussi utilisés dans les simulations de ce mémoire, sont ensuite présentés.

## 3.1. Le principe du test de Neyman

### 3.1.1. Construction du test lisse de Neyman

Le test de Neyman (1937) cherche à déterminer si les observations sont compatibles avec des données provenant d'une loi uniforme. Une des bases de ce test est que toute distribution peut être transformée en distribution uniforme en utilisant la transformation reposant sur la fonction de répartition. Avec cette transformation, il est possible de faire le test sur n'importe quelle base de données. Le test de Neyman repose sur une famille d'alternatives à la loi uniforme. La famille prend une forme particulière et est composée de polynômes orthogonaux à la loi uniforme. De plus, on doit choisir un entier  $k$ . Cette famille est parfois appelée famille d'imbrication et se formule ainsi :

$$g_k(y, \boldsymbol{\theta}) = C(\boldsymbol{\theta}) \exp \left[ \sum_{i=1}^k \theta_i h_i(y) \right], \quad 0 < y < 1, \quad (3.1.1)$$

où  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)^T$ ,  $\{h_i(y)\}$  est un ensemble contenant des polynômes orthogonaux par rapport à la loi uniforme et  $C(\boldsymbol{\theta})$ , qui est une constante dépendante de  $\boldsymbol{\theta}$  nécessaire pour que la densité intègre à un. Aussi,  $g_k$  est une densité paramétrée par  $\boldsymbol{\theta}$  avec l'argument  $y$  qui sont les observations. Les polynômes orthogonaux à la loi uniforme sont les polynômes de Legendre. Ceux-ci seront présentés dans le chapitre 4. Le test d'hypothèse à la loi uniforme se ramène au test paramétrique suivant :

$$H_0 : \boldsymbol{\theta} = \mathbf{0},$$

$$H_1 : \boldsymbol{\theta} \neq \mathbf{0}.$$

Remarquons que lorsque  $\boldsymbol{\theta} = \mathbf{0}$ , la densité  $g_k(y, \boldsymbol{\theta})$  en (3.1.1) sera un, soit la densité de la loi  $\mathbb{U}[0, 1]$ , où  $\mathbb{U}$  désigne la loi uniforme. Lorsque le test fut construit en 1937, il ne respectait pas les critères modernes des tests d'hypothèses, c'est-à-dire sans biais, puissant et de niveau  $\alpha$  pour toutes les hypothèses alternatives à la loi uniforme (voir la section 4.1 de Rayner, Thas et Best (2009)). Une solution décrite dans Rayner, Thas et Best (2009) est de transformer les  $\theta_i$  en  $\delta_i = \theta_i \sqrt{n}$ . Ce test s'appelle *Le principe généralisé du lemme de Neyman et Pearson* et sera présenté dans la prochaine section. Rayner, Thas et Best

(2009) ont montré que le test optimal, qui rejette pour de grandes valeurs, est :

$$\Psi_k^2 = \sum_{i=1}^k U_i^2, \quad (3.1.2)$$

avec :

$$U_i = \frac{1}{\sqrt{n}} \sum_{j=1}^n h_i(Y_j). \quad (3.1.3)$$

Il a été montré que ce test est le plus puissant sous certaines conditions et il rejette pour de grandes valeurs de  $\Psi_k^2$ . Cependant, la distribution conjointe des  $U_i$  est la normale multivariée et les conditions sont satisfaites asymptotiquement. Par conséquent, il peut être montré, sous certaines conditions, que la distribution asymptotique sous l'hypothèse nulle de  $\Psi_k^2$  est de loi  $\chi^2$ .

### 3.1.2. Dérivation de la statistique de test

Plutôt que d'utiliser l'approche originale de Neyman, il est possible d'obtenir la statistique de test de Neyman (1937) comme un test du score. Ceci a l'avantage de faire des liens avec les propriétés générales du test du rapport de vraisemblance. De plus, ceci facilite l'étude des propriétés asymptotiques. Supposons un échantillon  $x_1, \dots, x_n$ , d'une distribution continue de densité  $f(x)$ . Cette densité doit être connue pour construire la statistique de test. Ce développement se retrouve dans la section 4.2 de Rayner, Thas et Best (2009). La densité alternative d'ordre  $k$  à  $f(x)$  est définie par :

$$g_k(x, \boldsymbol{\theta}) = C(\boldsymbol{\theta})f(x) \exp \left[ \sum_{i=1}^k \theta_i h_i(x) \right], \quad (3.1.4)$$

$\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)^T$  et  $C(\boldsymbol{\theta})$ , une constante dépendante de  $\boldsymbol{\theta}$  nécessaire pour que la densité intègre à un. L'ensemble des  $\{h_i(x)\}_{i=0}^{\infty}$  est un ensemble de fonctions orthogonales à  $f(x)$  satisfaisant :

$$\int_{-\infty}^{+\infty} h_r(x)h_s(x)f(x)dx = \delta_{rs}, \quad r, s = 0, 1, 2, \dots \quad (3.1.5)$$

Il est convenu de poser  $h_0(x) = 1, \forall x$  de telle sorte que :

$$\int_{-\infty}^{+\infty} h_r(x)f(x)dx = \delta_{r0}, \quad r = 0, 1, 2, \dots$$

Le choix de  $h_0(x) = 1$  assure que  $\mathbb{E}[h_r(x)] = \delta_{r0} = 0$  pour tout  $r \geq 1$ . La fonction  $\delta_{rs}$  est une fonction indicatrice telle que :

$$\delta_{rs} = \begin{cases} 1, & \text{si } r = s, \\ 0, & \text{sinon.} \end{cases} \quad (3.1.6)$$

Pour tester l'hypothèse que les données proviennent de la densité  $f(x)$ , cela revient à tester que le vecteur  $\boldsymbol{\theta}$  dans l'expression (3.1.4) est nul :

$$H_0 : \boldsymbol{\theta} = \mathbf{0},$$

$$H_1 : \boldsymbol{\theta} \neq \mathbf{0}.$$

Tel que mentionné précédemment, si  $\boldsymbol{\theta} = \mathbf{0}$ , alors la densité  $g_k(x, \boldsymbol{\theta})$  devient égale à  $f(x)$ . La fonction de vraisemblance  $L$  correspondant à la densité (3.1.4) est :

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{j=1}^n \left\{ C(\boldsymbol{\theta}) \exp \left( \sum_{i=1}^k \theta_i h_i(x_j) \right) f(x_j) \right\}, \\ &= C(\boldsymbol{\theta})^n \exp \left( \sum_{i=1}^k \sum_{j=1}^n \theta_i h_i(x_j) \right) \prod_{j=1}^n f(x_j). \end{aligned} \quad (3.1.7)$$

Puis, le logarithme de la vraisemblance est obtenue :

$$\log L(\boldsymbol{\theta}) = n \log C(\boldsymbol{\theta}) + \sum_{i=1}^k \sum_{j=1}^n \theta_i h_i(x_j) + \sum_{j=1}^n \log f(x_j). \quad (3.1.8)$$

Pour être en mesure de dériver la log-vraisemblance, il faut dériver  $\log C(\boldsymbol{\theta})$ . Pour ce faire, le lemme 4.2.1 de Rayner, Thas et Best (2009, p. 57) est nécessaire et il est présenté dans le lemme 3.1.1. Pour plus de détails sur les notations, consulter Rayner, Thas et Best (2009).

**Lemme 3.1.1.** *La dérivée de  $\log C(\boldsymbol{\theta})$  satisfait :*

$$-\frac{\partial \log C(\boldsymbol{\theta})}{\partial \theta_i} = \mathbb{E}_k \{ h_i(x) \},$$

et

$$-\frac{\partial^2 \log C(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = \text{Cov}_k \{ h_i(x), h_j(x) \}.$$

Il est maintenant possible de procéder au calcul de la dérivée et d'obtenir le test du score basé sur la statistique de Neyman (voir le Théorème 4.2.2 dans Rayner, Thas et Best (2009)). Le résultat est énoncé dans le théorème 3.1.1.

**Théorème 3.1.1.** *La statistique du score pour tester  $H_0 : \boldsymbol{\theta} = \mathbf{0}$  contre  $H_1 : \boldsymbol{\theta} \neq \mathbf{0}$  reposant sur un échantillon aléatoire  $x_1, \dots, x_n$  provenant d'une distribution  $g_k(x, \boldsymbol{\theta})$  est :*

$$S_k = \sum_{i=1}^k U_i^2,$$

avec :

$$U_i = \frac{1}{\sqrt{n}} \sum_{j=1}^n h_i(Y_j).$$

Dans le théorème 4.2.2 de Rayner, Thas et Best (2009), les  $U_i$  ne sont pas exactement le  $i^{\text{ème}}$  élément du score, ils diffèrent du facteur  $\sqrt{n}$ . La différence entre la statistique  $S_k$  et  $\Psi_k$  de Neyman est que la première statistique est basée sur des polynômes orthogonaux à une densité quelconque  $f(x)$  alors que la deuxième statistique est basée sur les polynômes de Legendre. Donc,  $\Psi_k$  est un cas particulier de  $S_k$ . La statistique de test  $S_k$  suit une  $\chi_k^2$  car la matrice d'information est de rang  $k$ . Le choix de l'ordre  $k$  peut s'avérer complexe. Dans la prochaine section, une procédure de sélection automatique de l'ordre  $k$  est discutée.

### 3.1.3. Le choix de l'ordre $k$

Le test lisse de Neyman nécessite que le nombre de composantes  $k$  soit spécifié. Un aspect délicat associé à la performance de ce test est justement le choix de l'ordre  $k$ . Ce choix est critique, car s'il est trop petit ou trop grand, il peut mener à une perte de puissance. Ledwina (1994) propose une méthode de sélection automatique fonction des données afin d'obtenir le meilleur choix  $k$  parmi un ensemble  $\{1, \dots, d\}$ , où  $d$  est une borne supérieure qui doit être spécifiée. Le critère de sélection bayésien est utilisé afin de déterminer le meilleur choix de  $k$ . Pour ce faire, le logarithme de la vraisemblance défini en (3.1.8) d'un échantillon aléatoire doit être calculé pour une alternative d'ordre  $k$  pour la loi uniforme avec les polynômes de Legendre. Le critère d'information bayésien est :

$$BIC_k = 2 \sup_{\boldsymbol{\theta} \in R^k} \{\log(L_k) - k \log n\}.$$

Le choix optimal, noté par  $K$ , est le plus petit  $k$  qui maximise le critère  $BIC_k$  et qui s'écrit comme suit :

$$K = \min \{k \in \{1, 2, \dots, d\} \text{ et } BIC_k \geq BIC_r, \quad r = 1, 2, \dots, d\}.$$

Avec ce critère de sélection, la statistique de test devient :

$$S_K = U_1^2 + \dots + U_K^2.$$

Les degrés de liberté de la khi-deux sont différents de ceux du test  $\Psi_k^2$ . Les degrés de liberté du test ne sont plus  $K$  étant donné que sous l'hypothèse nulle,  $K$  est maintenant aléatoire. Ledwina (1994) a montré que, sous certaines conditions et dans son contexte, que  $k$  tend en probabilité vers un lorsque la taille de l'échantillon augmente. Ce résultat est mentionné dans le théorème 3.1.2.

**Théorème 3.1.2.** *Sous l'hypothèse nulle, lorsque  $n \rightarrow \infty$ ,  $K \xrightarrow{p} 1$  et  $S_K \xrightarrow{d} \chi_1^2$ .*

Lorsque la taille de l'échantillon augmente, le  $K$  choisi converge vers la plus petite valeur possible de l'ensemble  $\{1, 2, \dots, d\}$ . Cependant, dans un échantillon fini, des études de simulation faites par Ledwina (1994) démontrent que la densité n'est pas toute concentrée à  $K = 1$  sous l'hypothèse nulle, mais qu'il peut aussi y avoir de la masse ailleurs, soit à  $K > 1$ . Donc, pour un niveau  $\alpha$ , le point critique de la distribution  $\chi_1^2$  est plus petit que le  $\alpha$  d'une  $\chi_2^2$ . En utilisant le test, il faut s'attendre à ce que le niveau empirique soit plus élevé que le niveau  $\alpha$  en utilisant les points critiques d'une  $\chi_1^2$ . Le niveau empirique va tendre vers le niveau  $\alpha$  lorsque  $n \rightarrow \infty$ . Sous l'hypothèse alternative, il est faux que  $K \xrightarrow{p} 1$  et  $K$  devrait être plus grand. Utiliser la région critique d'une  $\chi_1^2$  aura ainsi pour effet d'augmenter la puissance. Ledwina (1994) a aussi montré que le test reposant sur la sélection automatique est efficace pour les alternatives d'ordre  $k \leq d$ . Dans le cas où l'hypothèse nulle n'est pas vraie, ces tests rejettent l'hypothèse nulle avec une probabilité qui tend vers un. Des tests de normalité sont présentés dans la prochaine section.

### 3.2. Tests de normalité

Les moments de la loi normale caractérisent cette distribution. Plus particulièrement, le troisième moment (coefficient d'asymétrie) et le quatrième moment (coefficient d'aplatissement) influencent la forme de la distribution de la loi normale. L'estimation de ces

moments permet la construction de certains tests de normalité. Dans cette section, trois tests invariants de normalité seront présentés. D'abord, les tests de Mardia (1970) sont introduits, suivi du test de Jarque et Bera (1980, 1987). Posons  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)^T$ , un vecteur aléatoire de dimension  $m$ . Une composante importante des tests invariants par rapport à la loi normale multivariée est la mesure de Mahalanobis, telle que mentionnée dans Henze (2002).

**Proposition 3.2.1.** *Toute statistique de test invariante pour tester la loi normale multivariée, est une fonction de la distance de Mahalanobis  $\{D_{jk} : 1 \leq j, k \leq n\}$ , où*

$$D_{jk} = (\mathbf{X}_j - \bar{\mathbf{X}})^T \mathbf{S}^{-1} (\mathbf{X}_k - \bar{\mathbf{X}}) = \mathbf{Y}_j^T \mathbf{Y}_k, \quad (3.2.1)$$

avec  $\bar{\mathbf{X}} = n^{-1} \sum_{j=1}^n \mathbf{X}_j$  et  $\mathbf{S} = n^{-1} \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T$ . Les variables  $\bar{\mathbf{X}}$  et  $\mathbf{S}$  sont respectivement la moyenne du vecteur et l'estimateur de la variance-covariance. Les résidus centrés et réduits peuvent s'écrire de la façon suivante :  $\mathbf{Y}_j = \mathbf{S}^{-1/2}(\mathbf{X}_j - \bar{\mathbf{X}})$  où  $\mathbf{S}^{-1/2}$  est la matrice symétrique et définie positive de la racine carrée de  $\mathbf{S}^{-1}$ . Ils seront utilisés lors de la présentation du test de Jarque-Bera.

### 3.2.1. Coefficient d'asymétrie

Avant de présenter la mesure du coefficient d'asymétrie de Mardia pour des données vectorielles, le test pour des données scalaires est introduit :

$$\sqrt{b_{n,1}^{(1)}} = \frac{1}{n} \sum_{j=1}^n \left( \frac{X_j - \bar{X}}{\sigma} \right)^3, \quad (3.2.2)$$

avec  $\sigma^2 = \sum_{j=1}^n (X_j - \bar{X})^2 / n$ . Mardia (1970) a généralisé ce test pour des vecteurs aléatoires :

$$b_{n,m}^{(1)} = \frac{1}{n^2} \sum_{j,k=1}^n D_{jk}^3, \quad (3.2.3)$$

$$= \frac{1}{n^2} \sum_{j,k=1}^n (\mathbf{Y}_j^T \mathbf{Y}_k)^3. \quad (3.2.4)$$

Mardia (1975) montre que  $b_{n,m}^{(1)}$  a un lien naturel avec les études de robustesse sur le test- $T^2$  de Hotelling sous l'hypothèse de non-normalité. La statistique de test  $b_{n,m}^{(1)}$  est invariante pour des transformations orthogonales et aussi invariante pour des transformations linéaires affines. Une transformation linéaire affine peut s'écrire comme la fonction

suivante :  $T(\mathbf{X}) = \mathbf{a}\mathbf{X} + \mathbf{b}$ , où  $\mathbf{a}$  est une matrice et  $\mathbf{b}$  est un vecteur. Si  $\mathbf{X}$  est sphérique symétrique et  $\mathbb{E}\|\mathbf{X}\|^6 < \infty$ , la distribution sera :

$$nb_{n,m}^{(1)} \xrightarrow{d} \alpha_1 \chi_m^2 + \alpha_2 \chi_{m(m-1)(m+4)/6}^2, \quad (3.2.5)$$

avec :

$$\alpha_1 = \frac{3}{m} \left[ \frac{\mathbb{E}\|\mathbf{X}\|^6}{m+2} - 2\mathbb{E}\|\mathbf{X}\|^4 + m(m+2) \right], \quad (3.2.6)$$

$$\alpha_2 = \frac{6\mathbb{E}\|\mathbf{X}\|^6}{m(m+2)(m+4)}, \quad (3.2.7)$$

et  $\chi_m^2$  et  $\chi_{m(m-1)(m+4)/6}^2$  sont indépendantes. Si  $\mathbf{X} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{I}_m)$ , les paramètres deviennent à  $\alpha_1 = \alpha_2 = 6$  :

$$nb_{n,m}^{(1)} \xrightarrow{d} 6\chi_{m(m+1)(m+2)/6}^2. \quad (3.2.8)$$

Pour tester la normalité,  $b_{n,m}^{(1)}$  est calculé et  $H_0$  est rejetée pour des grandes valeurs de  $nb_{n,m}^{(1)}/6$ . Mardia (1970) montre que faire un test de loi normale multivariée reposant sur  $b_{n,m}^{(1)}$  n'est pas convergent contre certaines alternatives satisfaisant  $\mathbb{E}\|\mathbf{X}\|^6 < \infty$ . Ainsi, la puissance empirique est grande contre des alternatives à ailes lourdes et elle est petite contre les distributions sphériques symétriques à ailes moins lourdes.

### 3.2.2. Coefficient d'aplatissement

Dans le cas univarié, le test basé sur le coefficient d'aplatissement est :

$$b_{n,1}^{(2)} = \frac{1}{n} \sum_{j=1}^n \left( \frac{X_j - \bar{X}}{\sigma} \right)^4, \quad (3.2.9)$$

avec  $\sigma^2 = \sum_{j=1}^n (X_j - \bar{X})^2/n$ . Mardia (1970) a aussi adapté cette mesure pour des vecteurs aléatoires :

$$b_{n,m}^{(2)} = \frac{1}{n} \sum_{j=1}^n D_{jj}^2, \quad (3.2.10)$$

$$= \frac{1}{n} \sum_{j=1}^n \|\mathbf{Y}_j\|^4. \quad (3.2.11)$$

La statistique  $b_{n,m}^{(2)}$  a aussi des connections avec les études de robustesse sur la taille d'un test de normalité de la matrice de covariance. Pour plus de détails, consulter Mardia



(1974). Ce test, tout comme celui reposant sur le coefficient d'asymétrie, est invariant aux transformations orthogonales et invariant aux transformations linéaires affines. Sous l'hypothèse de normalité, Mardia (1970) a montré que la distribution de la statistique est asymptotiquement :

$$\sqrt{n}(b_{n,m}^{(2)} - m(m+2)) \xrightarrow{d} \mathcal{N}(0, 8m(m+2)).$$

Le test de normalité basé sur  $b_{n,m}^{(2)} - m(m+2)$  est valide si et seulement si  $\mathbb{E}\|\mathbf{X}\|^4 \neq m(m+2)$ . Le test est rejeté pour des petites et des grandes valeurs de la statistique de test. Les tests de Mardia utilisent les coefficients d'asymétrie et d'aplatissement séparément. Jarque et Bera (1980, 1987) les utilisent dans la même statistique de test, qui est présentée dans la prochaine section.

### 3.2.3. Test de Jarque-Bera

Le test de Jarque-Bera est obtenu du test du multiplicateur de Lagrange sur la famille des distributions Pearson. En effectuant cette procédure, la statistique de test (3.2.12) est obtenue :

$$LM = n [b_1^2/6 + (b_2 - 3)^2/24], \quad (3.2.12)$$

avec  $b_1 = \hat{\mu}_3/\hat{\mu}_2^{3/2}$ ,  $b_2 = \hat{\mu}_4/\hat{\mu}_2^2$ , où  $\hat{\mu}_j = \sum_{i=1}^n (X_i - \bar{X})^j/n$  et  $\bar{X} = \sum_{i=1}^n X_i/n$ . La quantité  $\hat{\mu}_j$  est l'estimateur centré du  $j^{\text{ième}}$  moment de la loi normale. Les variables  $b_1$  et  $b_2$  sont respectivement les coefficients d'asymétrie et d'aplatissement de l'échantillon. L'hypothèse nulle est rejetée si la statistique est plus grande que le niveau de signification d'une  $\chi_2^2$ .

Le test  $LM$  (3.2.12) est un test du score, donc il est asymptotiquement le plus puissant. Jarque et Bera (1987) montrent que  $LM$  a une bonne puissance, même pour des familles qui ne sont pas membres de celle de Pearson. Ils ont aussi démontré à l'aide de simulations que la statistique  $LM$  est invariante au paramètre d'échelle, où le paramètre d'échelle est un facteur multiplicatif. Au niveau technique, l'avantage de ce test est la facilité à le programmer.

Lütkepohl (2005) présente une version multivariée de ce test dans les modèles autorégressifs vectoriels. La version présentée est basée sur les résidus car elle peut s'avérer plus puissante que celle basée sur les observations. Ce point a été démontré pour un échantillon

composé de données univariées par Lütkepohl et Schneider (1989). D'abord, introduisons les composantes :

$$\mathbf{b}_1 = (b_{11}, \dots, b_{m1})^T,$$

$$\mathbf{b}_2 = (b_{12}, \dots, b_{m2})^T,$$

où  $b_{k1} = n^{-1} \sum_{t=1}^n \mathbf{Y}_{tk}^3$ ,  $b_{k2} = n^{-1} \sum_{t=1}^n \mathbf{Y}_{tk}^4$ . La statistique de Jarque-Bera est définie dans l'équation (3.2.13) :

$$L_{JB} = \frac{n[\hat{\mathbf{b}}_1^T \hat{\mathbf{b}}_1 / 6 + (\hat{\mathbf{b}}_2 - \mathbf{3}_m)^T (\hat{\mathbf{b}}_2 - \mathbf{3}_m)]}{24}, \quad (3.2.13)$$

avec  $\mathbf{3}_m = (3, \dots, 3)^T$ . Lütkepohl (2005) a montré que la distribution de la statistique sur les résidus d'un VAR( $p$ ) est :

$$L_{JB} \xrightarrow{d} \chi_{2m}^2.$$

Les trois tests seront utilisés lors des simulations afin de les comparer avec les tests lisses d'ajustement proposés dans la section suivante.

# Chapitre 4

---

## Nouveaux tests

Ce chapitre présente un test lisse d'ajustement basé sur Neyman (1937) et il repose également sur le chapitre 4 de la thèse de Tagne Tatsinkou (2016). Ce test d'ajustement s'effectue dans le contexte de séries chronologiques vectorielles avec une paramétrisation structurée. Plus particulièrement, le test est fait pour des modèles autorégressifs moyennes mobiles vectoriels (VARMA). La théorie est présentée de façon générale en utilisant une famille quelconque de polynômes orthogonaux par rapport à la loi normale multivariée. Par la suite, les familles de polynômes d'Hermite et Legendre sont présentées ainsi qu'une procédure de sélection automatique visant à déterminer l'ordre de la famille d'emboîtement.

### 4.1. Préliminaires

Soit  $\mathbf{Y} = \{\mathbf{Y}_t, t \in \mathbb{Z}\}$ , un processus stochastique de dimension  $m$  où  $\mathbf{Y}_t = (Y_{t1}, \dots, Y_{tm})^T$ . Le processus  $\mathbf{Y}$  peut s'écrire sous la forme d'un processus VARMA avec une paramétrisation structurée s'il satisfait :

$$\Phi(B)(\mathbf{Y}_t - \boldsymbol{\mu}) = \Theta(B)\mathbf{Z}_t, \quad (4.1.1)$$

tel qu'énoncé dans la définition 2.3.1. Les polynômes AR et MA des équations (2.3.1) et (2.3.2), sont définis par :

$$\Theta(B) = \mathbf{I}_m + \Theta_1 B + \dots + \Theta_q B^q,$$

$$\Phi(B) = \mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p,$$

où  $\mathbf{I}_m$  représente la matrice identité d'ordre  $m$ . Les matrices  $\Phi_k$  et  $\Theta_k$  sont de dimension  $m \times m$  et leurs éléments sont supposés être des fonctions deux fois différentiables avec des dérivées continues du vecteur de paramètres  $\beta$ , de dimension  $b \times 1$ , avec  $b \leq m^2(p + q)$ . Lorsque  $\beta = \text{vec}(\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q)$ , où  $\text{vec}(\cdot)$  représente l'opérateur qui empile les colonnes d'une matrice en un vecteur, nous nous trouvons dans la situation où il y a  $m^2(p + q)$  paramètres libres. Il est supposé que  $\det\{\Phi(z; \beta)\}$  et  $\det\{\Theta(z; \beta)\}$  n'ont pas de racines communes. Aussi, le processus  $\{\mathbf{Y}_t\}$  est assumé stationnaire, de manière à ce que  $\det\{\Phi(z; \beta)\} \neq 0$ , pour tout  $|z| \leq 1$ . En plus d'assumer que les fonctions matricielles dans le modèle (4.1.1) doivent être différentiables, il est supposé que la matrice hessienne :

$$\mathbf{H}^T = \left( \frac{\partial \phi_1^T}{\partial \beta}, \dots, \frac{\partial \phi_p^T}{\partial \beta}, \frac{\partial \theta_1^T}{\partial \beta}, \dots, \frac{\partial \theta_q^T}{\partial \beta} \right), \quad (4.1.2)$$

est de rang plein  $b$ , où  $\phi_j = \text{vec}\{\Phi_j(\beta)\}$  et  $\theta_j = \text{vec}\{\Theta_j(\beta)\}$ . Aussi, le modèle est paramétrisé de manière unique de sorte que si  $\Phi_j(\beta_1) = \Phi_j(\beta_2)$  pour  $j = 1, \dots, p$  et  $\Theta_j(\beta_1) = \Theta_j(\beta_2)$  pour  $j = 1, \dots, q$ , ceci implique que  $\beta_1 = \beta_2$ . Pour avoir plus de détails sur l'utilisation de la paramétrisation structurée ou pour obtenir des exemples, on peut consulter Ahn (1988).

Soit  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ , un échantillon de taille  $n$  de dimension  $m$  provenant du modèle (4.1.1). L'estimation des paramètres se fait avec la méthode du maximum de la vraisemblance conditionnelle. La fonction (4.1.3) du log-vraisemblance est utilisée puisqu'elle facilite les calculs. Cette formule prend la forme suivante :

$$\tilde{l}_c(\nu) = -\frac{n}{2} \log\{\det(\Sigma_Z)\} - \frac{1}{2} \sum_{t=1}^n \mathbf{Z}_t^T(\beta, \mu) \Sigma_Z^{-1} \mathbf{Z}_t(\beta, \mu), \quad (4.1.3)$$

où  $\nu = (\beta^T, \mu^T, \sigma^T)^T$ , un vecteur de dimension  $l \times 1$  et  $l = b + m + (m(m + 1)/2)$ . Le vecteur  $\sigma$  est obtenu en mettant les colonnes de  $\Sigma_Z$  dans un vecteur, en gardant que les éléments sur et en dessous de la diagonale. Les résidus du modèle peuvent s'écrire de la façon suivante :

$$\left\{ \begin{array}{l} \mathbf{Z}_1(\boldsymbol{\beta}, \boldsymbol{\mu}) = \mathbf{Y}_1 - \boldsymbol{\mu}, \\ \mathbf{Z}_2(\boldsymbol{\beta}, \boldsymbol{\mu}) = (\mathbf{Y}_2 - \boldsymbol{\mu}) - \boldsymbol{\Phi}_1(\mathbf{Y}_1 - \boldsymbol{\mu}) - \boldsymbol{\Theta}_1 \mathbf{Z}_1(\boldsymbol{\beta}, \boldsymbol{\mu}), \\ \quad \quad \quad \vdots \\ \mathbf{Z}_n(\boldsymbol{\beta}, \boldsymbol{\mu}) = (\mathbf{Y}_n - \boldsymbol{\mu}) - \sum_{i=1}^p \boldsymbol{\Phi}_i(\mathbf{Y}_{n-i} - \boldsymbol{\mu}) - \sum_{j=1}^q \boldsymbol{\Theta}_j \mathbf{Z}_{n-j}(\boldsymbol{\beta}, \boldsymbol{\mu}). \end{array} \right.$$

Les résidus peuvent aussi s'écrire sous la forme compacte :

$$\mathbf{Z}_t(\boldsymbol{\beta}, \boldsymbol{\mu}) = \boldsymbol{\Theta}^{-1}(B, \boldsymbol{\beta}) \boldsymbol{\Phi}(B, \boldsymbol{\beta}) \mathbf{Y}_t^*,$$

avec  $\mathbf{Y}_t^* = (\mathbf{Y}_t - \boldsymbol{\mu})$ . Or, il existe une méthode d'estimation alternative à maximiser (4.1.3). En effet, Lütkepohl (2005) montre qu'au lieu de maximiser  $\tilde{l}_c(\boldsymbol{\nu})$ , il est possible de minimiser  $\det\{\tilde{\Sigma}_{\mathbf{Z}}(\boldsymbol{\beta}, \boldsymbol{\mu})\}$  (voir la section 12.3 de Lütkepohl (2005)), où

$$\det\{\tilde{\Sigma}_{\mathbf{Z}}(\boldsymbol{\beta}, \boldsymbol{\mu})\} = \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_t(\boldsymbol{\beta}, \boldsymbol{\mu}) \mathbf{Z}_t(\boldsymbol{\beta}, \boldsymbol{\mu})^T.$$

En minimisant  $\det\{\tilde{\Sigma}_{\mathbf{Z}}(\boldsymbol{\beta}, \boldsymbol{\mu})\}$ , l'estimateur par le maximum de vraisemblance est obtenu et il est noté  $\hat{\boldsymbol{\nu}}_{CML}$ . Dans la prochaine section, un test lisse d'ajustement par rapport à la loi normale est proposé.

## 4.2. Tests lisses d'ajustement pour le modèle VARMA

Dans ce mémoire, le test lisse qui sera introduit cherche à déterminer si le terme d'erreur du modèle VARMA provient ou non de la loi normale. Le test s'effectue sur les résidus d'un modèle VARMA. Notons que ce développement peut se faire sur les données brutes. L'hypothèse nulle de ce test est que le bruit blanc  $\{\mathbf{Z}_t, t \in \mathbb{Z}\}$  du modèle (4.1.1) suit une loi normale multivariée et se traduit par l'hypothèse suivante :

$$H_0 : \{\mathbf{Z}_t, t \in \mathbb{Z}\} \text{ est i.i.d., } \mathbf{Z}_t \sim \mathcal{N}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{Z}}).$$

Sous  $H_0$ , la densité de  $\mathbf{Z}_t$  est ainsi :

$$\phi_{\mathbf{Z}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{Z}}) = (2\pi)^{-\frac{m}{2}} \{\det(\boldsymbol{\Sigma}_{\mathbf{Z}})\}^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right). \quad (4.2.1)$$

La construction d'un test lisse d'ajustement requiert le choix d'une famille de polynômes orthogonaux par rapport à la loi normale multivariée. Plus particulièrement, la

famille orthogonale est choisie par rapport à la loi normale standard  $\mathcal{N}_m(\mathbf{0}, \mathbf{I}_m)$  et elle est notée  $\{H_i(y), i = 0, 1, \dots\}$ . Donc, les données doivent être centrées et réduites afin d'obtenir la loi normale multivariée standardisée. Pour ce faire, posons  $\mathbf{y} = \mathbf{C}(\mathbf{x} - \boldsymbol{\mu})$  avec  $\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^T = \mathbf{I}_m$ . La distribution de  $\mathbf{y}$  est maintenant  $\mathcal{N}_m(\mathbf{0}, \mathbf{I}_m)$ . Pour un  $\boldsymbol{\Sigma}$  quelconque, il existe plusieurs candidats de matrice  $\mathbf{C}$ . Cette transformation des données est souvent utilisée dans les tests lisses par rapport à la loi normale multivariée, particulièrement avec  $\mathbf{C} = \boldsymbol{\Sigma}^{-1/2}$  qui désigne l'inverse de la racine carrée d'une matrice. Par exemple, voir Gnanadesikan (1977). Pour le vecteur  $\mathbf{y} = (y_1, \dots, y_m)^T$ , nous définissons :

$$\begin{aligned} L_{\mathbf{r}:k}(\mathbf{y}) &= H_{r_1}(y_1) \cdot \dots \cdot H_{r_m}(y_m), \\ &= \prod_{i=1}^m H_{r_i}(y_i), \end{aligned} \tag{4.2.2}$$

avec  $\sum_{i=1}^m r_i = k$  et  $H_{r_i}(\cdot)$  est un polynôme de degré  $r_i$ ,  $i = 1, \dots, m$ . Le produit de polynômes  $L_{\mathbf{r}:k}(\mathbf{y})$  est un polynôme de degré  $k$ . Les fonctions  $L_{\mathbf{r}:k}$  sont ordonnées selon le degré pour que les fonctions  $L_{\mathbf{r}:k}$  de degré  $k$  soient avant les fonctions  $L_{\mathbf{r}:k+1}$  de degré  $k + 1$ . Ce système de fonctions est discuté dans Koziol (1986) et dans la section 6.4 de Rayner, Thas et Best (2009). Un argument d'analyse combinatoire révèle qu'il y a  $d(m, k) = \binom{m+k-1}{k}$  polynômes de degré  $k$  lorsque la dimension du vecteur  $\mathbf{r}$  est  $m$ . Voici un exemple de tous les polynômes de degré 3 de dimension 4 :

**Tableau 4.1.** Indices multiples ordonnés pour  $m = 4$  et  $k = 3$ .

$s$	$r$	$s$	$r$	$s$	$r$	$s$	$r$
1	(0, 0, 0, 3)	6	(0, 1, 1, 1)	11	(1, 0, 0, 2)	16	(1, 2, 0, 0)
2	(0, 0, 1, 2)	7	(0, 1, 2, 0)	12	(1, 0, 1, 1)	17	(2, 0, 0, 1)
3	(0, 0, 2, 1)	8	(0, 2, 0, 1)	13	(1, 0, 2, 0)	18	(2, 0, 1, 0)
4	(0, 0, 3, 0)	9	(0, 2, 1, 0)	14	(1, 1, 0, 1)	19	(2, 1, 0, 0)
5	(0, 1, 0, 2)	10	(0, 3, 0, 0)	15	(1, 1, 1, 0)	20	(3, 0, 0, 0)

Il est utile de définir une famille de fonctions des polynômes qui tient compte de l'ordre de la génération des polynômes de degré  $k$  :

$$\mathcal{F} = \bigcup_{k=1}^{\infty} \bigcup_{s=1}^{d(m,k)} \{L_{s:k}, 1 \leq s \leq d(m,k)\},$$

avec l'indice  $s$  qui est un des  $d(m,k)$  polynômes de degré  $k$ . Il est important de tenir compte de l'ordre utilisé lors du calcul de la distribution de la statistique de test. La première étape du test est de définir la densité de la famille alternative à la loi normale :

$$g_R(\mathbf{x}, \boldsymbol{\eta}, \boldsymbol{\Sigma}) = c(\boldsymbol{\eta}, \boldsymbol{\Sigma}) \exp \left[ \sum_{k=1}^R \sum_{s=1}^{d(m,k)} \eta_{s:k} L_{s:k} \{ \mathbf{C}(\mathbf{x} - \boldsymbol{\mu}) \} \right] \phi_{\mathbf{Z}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{Z}}), \quad (4.2.3)$$

avec  $L_{s:k}$  sont les polynômes orthogonaux par rapport à  $\phi_{\mathbf{Z}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{Z}})$ , tel que défini par la relation (3.1.5). Le vecteur  $\boldsymbol{\eta}_R = (\eta_{1:1}, \dots, \eta_{d(m,1):1}, \dots, \eta_{R:1}, \dots, \eta_{D(m,R):R})^T$  est le vecteur des coefficients de  $L_{s:k}$  dans l'équation (4.2.3). Ce vecteur est de dimension  $D(m, R) \times 1$ , où  $D(m, R) = \sum_{k=1}^R d(m, k)$ . Le test d'hypothèse d'une alternative d'ordre  $R$  à la loi normale multivariée peut se traduire par :

$$H_0 : \boldsymbol{\eta}_R = \mathbf{0},$$

$$H_1 : \boldsymbol{\eta}_R \neq \mathbf{0}.$$

Il faut obtenir la statistique du score de Rao de (4.2.3) pour être en mesure de tester cette hypothèse (voir le chapitre 6 de Rayner, Thas et Best (2009)). La statistique de test est :

$$V_{s:k}(\hat{\boldsymbol{\nu}}_{CML}) = n^{-1/2} \sum_{t=1}^n L_{s:k} \left\{ \hat{\mathbf{C}} \mathbf{Z}_t(\hat{\boldsymbol{\nu}}_{CML}) \right\}, \quad (4.2.4)$$

avec  $\hat{\boldsymbol{\nu}}_{CML}$  qui sont les estimateurs obtenus avec la méthode du maximum de la vraisemblance. La matrice  $\hat{\mathbf{C}}$  satisfait  $\hat{\mathbf{C}} \hat{\boldsymbol{\Sigma}}_{\mathbf{Z}} \hat{\mathbf{C}}^T = \mathbf{I}_m$ . Le vecteur (4.2.5) représente tous les polynômes de degré  $k$  de la statistique (4.2.4) et le vecteur (4.2.6) est le vecteur de la statistique avec tous les polynômes de degré 1 à  $R$  :

$$\mathbf{V}_k(\hat{\boldsymbol{\nu}}_{CML}) = (V_{1:k}(\hat{\boldsymbol{\nu}}_{CML}), \dots, V_{d(m,k):k}(\hat{\boldsymbol{\nu}}_{CML}))^T, \quad (4.2.5)$$

$$\mathbf{W}_R(\hat{\boldsymbol{\nu}}_{CML}) = (\mathbf{V}_1^T(\hat{\boldsymbol{\nu}}_{CML}), \dots, \mathbf{V}_R^T(\hat{\boldsymbol{\nu}}_{CML}))^T. \quad (4.2.6)$$

De façon similaire, les expressions vectorielles (4.2.7) et (4.2.8), basées sur l'équation (4.2.2) sont introduites :

$$\mathbf{L}_k(\mathbf{x}) = (L_{1:k}(\mathbf{x}), \dots, L_{d(m,k):k}(\mathbf{x}))^T, \quad (4.2.7)$$

$$\mathcal{L}_R(\mathbf{x}) = (\mathbf{L}_1^T(\mathbf{x}), \dots, \mathbf{L}_R^T(\mathbf{x}))^T. \quad (4.2.8)$$

Ces vecteurs sont respectivement de dimension  $d(m, k) \times 1$  et  $D(m, R) \times 1$ . Il est intéressant de réécrire l'équation (4.2.4) avec un vecteur de paramètres arbitraires  $\boldsymbol{\nu}$  tel que :

$$V_{s:k}(\boldsymbol{\nu}) = n^{-1/2} \sum_{t=1}^n L_{s:k} \{ \mathbf{CZ}_t(\boldsymbol{\beta}, \boldsymbol{\mu}) \}. \quad (4.2.9)$$

Sous  $H_0$ ,  $\{\mathbf{Z}_t\}$  est un bruit blanc gaussien et  $\mathbf{L}_k(\mathbf{CZ}_t)$  est de moyenne  $\mathbf{0}$  et de matrice des variances et des covariances  $\mathbf{I}_{d(m,k)}$ . Cependant, sous  $H_1$ , le paramètre est différent tel que  $\boldsymbol{\eta}_R = \boldsymbol{\eta}_1$ , avec  $\boldsymbol{\eta}_1 \neq \mathbf{0}$ . La puissance de  $\mathbf{L}_k(\mathbf{CZ}_t)$  devrait s'enligner vers  $g_k(\cdot, \boldsymbol{\nu}_1)$ . Or, le vecteur  $\boldsymbol{\nu}_0$  est inconnu et la distribution de  $\mathbf{W}_R$  est maintenant étudiée dans le théorème 4.2.1.

**Théorème 4.2.1.** *Soit  $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ , un processus VARMA( $p, q$ ) stationnaire et inversible généré par le modèle (4.1.1). Soit  $\hat{\boldsymbol{\nu}}_{CML}$ , l'estimateur par le maximum de vraisemblance conditionnelle de  $\boldsymbol{\nu}_0$ . Alors, la distribution de la statistique  $\mathbf{W}_R(\hat{\boldsymbol{\nu}}_{CML})$  est :*

$$\mathbf{W}_R(\hat{\boldsymbol{\nu}}_{CML}) \xrightarrow{d} \mathcal{N}_{D(m,R)}(\mathbf{0}_{D(m,R)}, \boldsymbol{\Omega}_{m,R}), \quad (4.2.10)$$

avec :

$$\boldsymbol{\Omega}_{m,R} = \mathbf{I}_{D(m,R)} - \frac{1}{4} \mathbf{B}_{m,R} (\mathbf{I}_{m^2} + \mathbf{K}_{mm}) \mathbf{B}_{m,R}^T - \mathbf{A}_{m,R} \mathbf{A}_{m,R}^T, \quad (4.2.11)$$

où  $\mathbf{K}_{mm}$  est la matrice de commutation et  $D(m, R) = \sum_{k=1}^R d(m, k)$ . Les matrices  $\mathbf{A}_{m,R}$  et  $\mathbf{B}_{m,R}$  sont non stochastiques et calculées sous l'hypothèse nulle de normalité :

$$\mathbf{A}_{m,R} = \mathbb{E}_0 \{ \mathcal{L}_R(\mathbf{CZ}_t) (\mathbf{CZ}_t)^T \}, \quad (4.2.12)$$

$$\mathbf{B}_{m,R} = \mathbb{E}_0 \{ \mathcal{L}_R(\mathbf{CZ}_t) \text{vec}^T(\mathbf{CZ}_t \mathbf{Z}_t^T \mathbf{C}^T) \}, \quad (4.2.13)$$

où  $\mathbb{E}_0$  désigne l'espérance mathématique sous l'hypothèse nulle. Les matrices (4.2.12) et (4.2.13) sont respectivement de dimensions  $D(m, R) \times m$  et  $D(m, R) \times m^2$ .



Les propriétés de base de la matrice  $\mathbf{K}_{mm}$  sont exposées dans la section 16.3 de Harville (2002). Le calcul des matrices  $\mathbf{A}_{m,R}$  et  $\mathbf{B}_{m,R}$  va dépendre de la famille de polynômes choisie. La preuve du théorème 4.2.1 est présentée dans la thèse de Tagne Tatsinkou (2016). De ce théorème et de la preuve, l'estimation du paramètre  $\boldsymbol{\beta}$  n'a pas d'impact sur la distribution asymptotique. Par contre, l'estimation des paramètres  $\boldsymbol{\mu}$  et  $\boldsymbol{\Sigma}$  ont un impact asymptotique. En effet, Pierce (1985) montre que tous les tests de normalité basés sur les résidus d'un modèle autorégressif ont la même distribution que les tests basés sur des observations indépendantes et identiquement distribuées avec la moyenne et la variance estimées.

L'application du théorème 4.2.1 consiste à construire une forme quadratique qui permet de tester l'hypothèse nulle de normalité. Si  $\boldsymbol{\Omega}_{m,R}$  est inversible, la statistique de test suivante converge vers une distribution khi-deux :

$$S_{m,R} = \mathbf{V}_R^T(\hat{\boldsymbol{\nu}}_{CML})\boldsymbol{\Omega}_{m,R}^{-1}\mathbf{V}_R(\hat{\boldsymbol{\nu}}_{CML}) \xrightarrow{d} \chi_{D(m,R)}^2. \quad (4.2.14)$$

Si  $\boldsymbol{\Omega}_{m,R}$  n'est pas inversible, il est possible d'introduire l'inverse généralisé  $\boldsymbol{\Omega}_{m,R}^-$  :

$$S_{m,R} = \mathbf{V}_R^T(\hat{\boldsymbol{\nu}}_{CML})\boldsymbol{\Omega}_{m,R}^-\mathbf{V}_R(\hat{\boldsymbol{\nu}}_{CML}) \xrightarrow{d} \chi_{\delta}^2. \quad (4.2.15)$$

avec  $\delta = \text{rang}(\boldsymbol{\Omega}_{m,R})$ . Les quantités  $a_u$  et  $b_u$  sont les composantes des matrices  $\mathbf{A}_{m,R}$  et  $\mathbf{B}_{m,R}$  :

$$a_u = \int_{-\infty}^{\infty} H_u(x)x\phi(x)dx, \quad (4.2.16)$$

$$b_u = \int_{-\infty}^{\infty} H_u(x)x^2\phi(x)dx. \quad (4.2.17)$$

Le résultat 4.2.1 fournit une expression générale des matrices  $\mathbf{A}_{m,R}$  et  $\mathbf{B}_{m,R}$ .

**Résultat 4.2.1.** Soit  $L_{r:k}(\mathbf{y})$  défini par l'équation (4.2.2) pour un ordre fixé de  $k$ . Alors :

$$\mathbb{E}_0\{L_{r:k}(\mathbf{F})F_j\} = \begin{cases} a_k, & \text{si } r_j = k \text{ et } r_i = 0, i \neq j, \\ 0, & \text{sinon,} \end{cases} \quad (4.2.18)$$

et

$$\mathbb{E}_0\{L_{r:k}(\mathbf{F})F_jF_l\} = \begin{cases} b_k, & \text{si } r_j = k \text{ et } r_i = 0, i \neq j, \\ b_k, & \text{si } r_l = k \text{ et } r_i = 0, i \neq l, \\ a_{r_j}a_{r_l}, & \text{si } r_j + r_l = k \text{ et } r_i = 0, i \neq j, i \neq l, \\ 0, & \text{sinon,} \end{cases} \quad (4.2.19)$$

où  $\mathbf{r} = (r_1, \dots, r_m)$  avec  $\sum_{i=1}^m r_i = k$  et  $\mathbf{F} = (F_1, \dots, F_m)^T = \mathbf{C}\mathbf{Z}$ , avec  $\mathbf{Z} \sim \mathcal{N}_m(\mathbf{0}, \Sigma_{\mathbf{Z}})$ . Le développement des matrices  $\mathbf{A}_{m,R}$  et  $\mathbf{B}_{m,R}$  ainsi que des exemples sont présentés en annexes. Deux familles importantes de polynômes orthogonaux sont les polynômes d’Hermite et de Legendre. La construction du test lisse d’ajustement utilisant ces familles est discutée dans les prochaines sections.

### 4.3. Polynômes d’Hermite

Les polynômes d’Hermite orthogonaux à la loi normale sont présentés dans la définition 4.3.1.

**Définition 4.3.1.** *Le polynôme d’Hermite de degré  $n$  est défini tel que :*

$$\mathcal{H}_n(y) = (-1)^n \exp(y^2/2) \frac{d^n}{dx^n} [\exp(-y^2/2)], \quad n \geq 1,$$

et  $H_0(y) = 1$ . Plus de détails sont fournis dans Sansone (1991). Les polynômes d’Hermite peuvent être aussi définis par la condition (4.3.1) énoncée dans Szegő (1975).

$$\int_{-\infty}^{\infty} \exp(-y^2) \mathcal{H}_n(y) \mathcal{H}_m(y) dy = \pi^{\frac{1}{2}} 2^n n! \delta_{nm}, \quad n, m = 0, 1, 2, \dots \quad (4.3.1)$$

où  $\delta_{rs}$  est la fonction de Kronecker tel que présentée en (3.1.6). Dans l’expression (4.3.1), il est possible de voir que pour être orthogonaux à la loi normale, les polynômes devront être divisés par  $\sqrt{n!}$ . Donc, les polynômes d’Hermite normalisés peuvent être écrits explicitement, c’est-à-dire les polynômes d’Hermite orthogonaux à la loi  $\mathcal{N}(0, 1)$  :

$$\begin{aligned} \mathcal{H}_0(x) &= 1, \\ \mathcal{H}_1(x) &= x, \\ \mathcal{H}_2(x) &= \frac{x^2 - 1}{\sqrt{2!}}, \\ \mathcal{H}_3(x) &= \frac{x^3 - 3x}{\sqrt{3!}}, \\ \mathcal{H}_4(x) &= \frac{x^4 - 6x^2 + 3}{\sqrt{4!}}, \\ \mathcal{H}_5(x) &= \frac{x^5 - 10x^3 + 15x}{\sqrt{5!}}, \\ \mathcal{H}_6(x) &= \frac{x^6 - 15x^4 + 45x^2 - 15}{\sqrt{6!}}. \end{aligned}$$

Ces polynômes sont reliés par deux relations de récurrence, soit :

$$\mathcal{H}_{n+1}(x) = x \frac{\mathcal{H}_n(x)}{\sqrt{n+1}} - \sqrt{\frac{n}{n+1}} \mathcal{H}_{n-1}(x), \quad (4.3.2)$$

et

$$\mathcal{H}'_n(x) = \sqrt{n} \mathcal{H}_{n-1}(x). \quad (4.3.3)$$

Le fait que les polynômes d'Hermite sont orthonaux à la loi normale se montre facilement : il suffit de montrer que  $\mathcal{H}_1$  et  $\mathcal{H}_2$  sont orthogonaux à  $\mathcal{N}(0, 1)$  et on utilise l'équation (4.3.2) pour établir la récurrence. Dans le cas des polynômes d'Hermite, les  $a_u$  et  $b_u$  des expressions (4.2.16) et (4.2.17) peuvent être écrites explicitement. En effet,  $a_u = \delta_{u1}$  et  $b_u = \sqrt{\delta_{u2}}$ . Il est possible de simplifier la statistique générale du théorème 4.2.1 à l'équation (4.3.4).

$$S_{m,r} = \sum_{k=3}^R U_k(\hat{\boldsymbol{\nu}}_{CML}), \quad (4.3.4)$$

avec :

$$U_k(\hat{\boldsymbol{\nu}}_{CML}) = \|\mathbf{V}_k(\hat{\boldsymbol{\nu}}_{CML})\|^2 = \sum_{s=1}^{d(m,k)} V_{s:k}^2(\hat{\boldsymbol{\nu}}_{CML}). \quad (4.3.5)$$

Rayner, Thas et Best (2009) ont montré que  $\mathbf{V}_1(\hat{\boldsymbol{\nu}}_{CML})$  et  $\mathbf{V}_2(\hat{\boldsymbol{\nu}}_{CML})$  sont égaux à zéro lorsque la normalité est testée. Ainsi, la matrice des variances et des covariances  $\boldsymbol{\Omega}_{m,R}$  de  $\mathbf{W}_R(\hat{\boldsymbol{\nu}}_{CML})$  converge vers une matrice singulière de rang  $D(m, R) - d(m, 1) - d(m, 2)$ . La distribution de la statistique (4.3.4) est :

$$S_{m,R} \xrightarrow{d} \chi_{D(m,R)-d(m,1)-d(m,2)}^2.$$

Si  $R = 4$ , la distribution de la statistique  $S_{m,4}$  est la suivante :

$$S_{m,4} = U_3(\hat{\boldsymbol{\nu}}_{CML}) + U_4(\hat{\boldsymbol{\nu}}_{CML}) \xrightarrow{d} \chi_{m(m+1)(m+2)(m+7)/24}^2.$$

Koziol (1987) montre que  $S_{m,3}$  et  $S_{m,4}$  sont invariants aux transformations linéaires affines. Pour montrer l'invariance de la statistique (4.3.4), la démonstration doit être faite pour  $R \geq 3$ . Le lemme 4.3.1 montre l'invariance de la composante  $U_k(\boldsymbol{\nu})$ .

**Lemme 4.3.1.** Soit  $U_k(\boldsymbol{\nu})$  définie par :

$$U_k(\boldsymbol{\nu}) = \sum_{s=1}^{d(m,k)} V_{s:k}^2(\boldsymbol{\nu}),$$

pour un  $\boldsymbol{\nu}$  quelconque, où  $V_{s:k}(\boldsymbol{\nu})$  est définie par (4.3.5). Alors  $U_k(\boldsymbol{\nu})$  est invariant aux transformations orthogonales pour tout  $k \geq 3$ .

La preuve du lemme 4.3.1 sera présentée dans le chapitre 5. Ce lemme est important afin de montrer l'invariance de la statistique (4.3.5). Le théorème 4.3.1 énonce l'invariance aux transformations linéaires affines de la statistique  $S_{m,R}$ .

**Théorème 4.3.1.** Soit la statistique de test définie par (4.3.5). La statistique  $U_k(\hat{\boldsymbol{\nu}}_{CML})$  est invariante aux transformations linéaires affines pour  $k \geq 3$ . Par conséquent, la statistique de test  $S_{m,R}$  définie par (4.2.4) est invariante aux transformations linéaires affines pour  $R \geq 3$ .

Pour généraliser la preuve, il est suffisant de montrer que chaque composante  $U_k(\hat{\boldsymbol{\nu}}_{CML})$  est invariante aux transformations linéaires. La preuve dépend du produit des fonctions  $\mathcal{H}_n(y)$ ,  $n \geq 1$ . Ce produit est un polynôme d'Hermite multivarié et une base de  $\mathcal{V}_k^m$ , un espace de fonctions orthogonales. Plus de détails sont donnés dans le chapitre 5. La preuve est présentée dans le prochain chapitre et est basée sur l'approche de Mardia (1970). Il existe un lien intéressant entre les statistiques de Mardia et les statistiques  $U_3$  et  $U_4$ .

**Remarque 4.3.1.** Notons que  $U_3$  est l'équivalent algébrique du coefficient d'asymétrie de Mardia. Il est plus facile de programmer la statistique de Mardia plutôt que la formule  $U_3$ . De plus,  $U_4$  n'est pas exactement équivalent au coefficient d'aplatissement de Mardia. Koziol (1989) montre qu'il existe la relation  $24nU_4 = n^2b_{2,m} - 6n^2b_{2,mp} + 3n^2m(m+2)$ .

Dans la prochaine section, l'autre famille des polynômes orthogonaux utilisée est présentée, soit les polynômes de Legendre.

## 4.4. Polynômes de Legendre

Les polynômes de Legendre orthogonaux à la loi uniforme sont présentés dans la définition 4.4.1.

**Définition 4.4.1.** *La famille des polynômes de Legendre peut être définie comme une solution d'une certaine équation différentielle (Sansone (1991)). Plus précisément, le polynôme de Legendre  $P_n$  de degré  $n$  est la solution de l'équation différentielle suivante :*

$$(1 - x^2) \frac{d^2 P_n}{dx^2} - 2x \frac{dP_n}{dx} + n(n + 1)P_n = 0. \quad (4.4.1)$$

Soit  $\mathcal{U}$ , la densité de la loi  $\mathbb{U}[-1, 1]$ , telle que :

$$\mathcal{U}(x) = \begin{cases} -\frac{1}{2}, & \text{si } -1 \leq x \leq 1, \\ 0, & \text{sinon.} \end{cases}$$

Les polynômes de Legendre vérifient la relation :

$$\langle P_m, P_n \rangle_{\mathcal{U}} = \frac{1}{2n + 1} \delta_{mn}.$$

Contrairement aux polynômes d'Hermite, les polynômes de Legendre sont orthogonaux à la loi uniforme, il faut donc trouver une façon de les utiliser pour tester la normalité. Pour ce faire, les données sont transformées avec la fonction de répartition  $F(x)$ . La fonction de répartition de la loi normale centrée et réduite sera utilisée puisque la normalité des données doit être testée. La transformation (4.4.2) utilisée est :

$$2\Phi(x) - 1, \quad (4.4.2)$$

où  $\Phi(x)$  est la fonction de répartition de la loi  $\mathcal{N}(0, 1)$ . On rappelle que si  $x \sim \mathcal{N}(0, 1)$  alors  $2\Phi(x) - 1 \sim \mathbb{U}(-1, 1)$ .

**Remarque 4.4.1.** *En utilisant la densité de la loi  $\mathbb{U}[-1, 1]$ , il faut multiplier les polynômes de Legendre par  $\sqrt{2}$  pour maintenir l'orthogonalité.*

Seulement les sept premiers polynômes de Legendre sont utilisés lors des simulations. Ils peuvent être écrits ainsi :

$$\mathcal{L}_0(x) = 1,$$

$$\mathcal{L}_1(x) = \sqrt{2}(1,224745x),$$

$$\mathcal{L}_2(x) = \sqrt{2}(2,371708x^2 - 0,7905694),$$

$$\mathcal{L}_3(x) = \sqrt{2}(4,677072x^3 - 2,806243x),$$

$$\mathcal{L}_4(x) = \sqrt{2}(9,280777x^4 - 7,954951x^2 + 0,7954951),$$

$$\mathcal{L}_5(x) = \sqrt{2}(18,46851x^5 - 20,52057x^3 + 4,397265x),$$

$$\mathcal{L}_6(x) = \sqrt{2}(-0,7967218 + 16,73116x^2 - 50,19347x^4 + 36,80855x^6).$$

Il est facile de vérifier si ces polynômes sont orthogonaux à la loi  $\mathbb{U}[-1, 1]$  en utilisant la définition 3.1.5. Il est alors possible d'utiliser la statistique  $S_{m,R}$  présentée dans le théorème 4.2.1 afin de tester la normalité en utilisant les polynôme de Legendre après avoir transformé les données avec (4.4.2).

Dans le cas des polynômes de Legendre, les constantes  $a_u$  sont nulles lorsque  $u$  est pair et les  $b_u$  sont nulles lorsque  $u$  est impair. Il n'est pas possible d'avoir des formules explicites pour les  $a_u$  et  $b_u$ . Or, il est possible d'avoir des approximations précises par intégration numérique. La statistique  $S_{m,R}$  utilisant les polynômes de Legendre n'est pas invariante aux transformations linéaires affines puisque les données sont transformées de façon non linéaire. L'invariance n'est pas toujours nécessaire, spécifiquement dans le cas des modèles VARMA. La prochaine section présente une procédure de sélection automatique de l'ordre  $k$ . Celle-ci est appliquée sur la famille des polynômes orthogonaux d'Hermite et la famille des polynômes orthogonaux de Legendre.

## 4.5. Sélection du $R$ automatique

Dans cette section, la sélection automatique du chapitre 3 sera adaptée et elle est utilisée dans les simulations. La sélection du  $R$  se fait sur la statistique  $S_{m,r}$  définie par (4.2.14). L'intervalle du  $R$  considéré est de  $d = 3$  à  $D = 6$  dans les simulations du chapitre 6. Les valeurs de  $R$  sont choisies dans l'ensemble  $\{d, d + 1, \dots, D\}$ . Le critère de sélection peut donc se traduire ainsi :

$$\hat{R} = \min_{3 \leq s \leq 6} \left\{ S_{m,s} - \frac{1}{2} D(m, s) \log n \right\}.$$

Si plusieurs critères sont égaux, le plus petit  $R$  associé est utilisé. Comme l'indique le théorème 3.1.2,  $\hat{R}$  converge vers la plus petite valeur, soit 3 dans ce cas, lorsque la taille augmente. Ce résultat est vrai sous  $H_0$ , comme l'indique le théorème 4.5.1.

**Théorème 4.5.1.** *Sous l'hypothèse nulle de normalité, le choix du  $R$  va converger vers  $d$  tel que :*

$$\hat{R} \xrightarrow{p} d. \quad (4.5.1)$$

On note que  $\hat{R}$  est aléatoire, et il est possible d'établir le résultat asymptotique suivant :

$$S_{m,\hat{R}} \xrightarrow{d} \chi_{rang(\Omega_m,d)}^2.$$

Par exemple, si la dimension est  $m = 2$ , les degrés de liberté seront 9 et si  $m = 3$ , les degrés de liberté seront plutôt 19.

**Remarque 4.5.1.** *Notons que dans le cas de la famille d'Hermite, ce sera la statistique (4.3.4) de la section 4.2.1 qui sera utilisée. Les degrés de liberté changent et sont respectivement 4 et 10 pour  $m = 2$  et  $m = 3$ .*

Dans la prochaine section, la preuve d'invariance de la statistique de test  $S_{m,r}$  reposant sur l'utilisation des polynômes d'Hermite est présentée.





# Chapitre 5

---

## Sur l'invariance du test lisse avec les polynômes d'Hermite

Cette section est consacrée à la preuve du théorème 4.3.1. Cette preuve nécessite la compréhension de la statistique  $U_k$  définie par l'équation (4.3.5), que nous rappelons être définie comme suit :

$$U_k(\hat{\boldsymbol{\nu}}_{CML}) = \|\mathbf{V}_k(\hat{\boldsymbol{\nu}}_{CML})\|^2 = \sum_{s=1}^{d(m,k)} V_{s:k}^2(\hat{\boldsymbol{\nu}}_{CML}),$$

avec :

$$V_{s:k}(\hat{\boldsymbol{\nu}}_{CML}) = n^{-1/2} \sum_{t=1}^n L_{s:k} \left\{ \hat{\mathbf{C}} \mathbf{Z}_t(\hat{\boldsymbol{\nu}}_{CML}) \right\}.$$

Les statistiques de  $U_k$  sont fonctions de polynômes homogènes, plus particulièrement des polynômes d'Hermite multivariés. Il sera montré que les composantes de la statistique, faisant intervenir  $\mathbf{L}_k(\mathbf{x})$ , sont orthogonales à la loi normale multivariée. Aussi, elles engendrent une base orthogonale à la densité gaussienne multivariée. Il est prouvé que le vecteur  $\mathbf{L}_k(\mathbf{x})$ , défini par (4.2.7) permet de définir une base orthonormale de  $\mathcal{V}^m$ , définie par (5.1.3).

### 5.1. Notation

Il faut d'abord rappeler des notations d'algèbre linéaire. Celles-ci se trouvent dans le chapitre 4 de Meyer (2000) et dans les chapitres 3 et 4 de Rahman (2017). Un espace

vectorel implique deux ensembles  $\mathcal{V}$  et  $\mathcal{F}$ , ainsi que l'addition de vecteurs et la multiplication par un scalaire. La définition 5.1.1 d'un espace vectoriel met en relation ces quatre éléments.

**Définition 5.1.1.**  $\mathcal{V}$  est un espace vectoriel sur le domaine  $\mathcal{F}$  quand les opérations d'additions et multiplications vectorielles satisfont :

- (A1)  $\mathbf{x} + \mathbf{y} \in \mathcal{V}$  pour tout  $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ . Cette propriété est appelée « propriété de fermeture pour l'addition de vecteur ».
- (A2)  $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$  pour tout  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$ .
- (A3)  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$  pour tout  $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ .
- (A4) Il existe un élément  $\mathbf{0} \in \mathcal{V}$  tel que  $\mathbf{x} + \mathbf{0} = \mathbf{x}$  pour tout  $\mathbf{x} \in \mathcal{V}$ .
- (A5) Pour tout  $\mathbf{x} \in \mathcal{V}$ , il existe un élément  $(-\mathbf{x}) \in \mathcal{V}$  tel que  $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ .
- (M1)  $\alpha\mathbf{x} \in \mathcal{V}$  pour tout  $\alpha \in \mathcal{F}$  et  $\mathbf{x} \in \mathcal{V}$ . Cette propriété est appelée « propriété de fermeture pour la multiplication de vecteur ».
- (M2)  $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$  pour tout  $\alpha, \beta \in \mathcal{F}$  et  $\mathbf{x} \in \mathcal{V}$ .
- (M3)  $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$  pour tout  $\alpha \in \mathcal{F}$  et  $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ .
- (M4)  $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$  pour tout  $\alpha, \beta \in \mathcal{F}$  et  $\mathbf{x} \in \mathcal{V}$ .
- (M5)  $1\mathbf{x} = \mathbf{x}$  pour tout  $\mathbf{x} \in \mathcal{V}$ .

Si un sous-ensemble  $\mathcal{S}$  d'un espace vectoriel satisfait aussi les conditions d'un espace vectoriel, il sera appelé un sous-espace vectoriel. La définition 5.1.2 présente un sous-espace vectoriel.

**Définition 5.1.2.** Soit  $\mathcal{S}$ , un sous-ensemble non vide d'un espace vectoriel  $\mathcal{V}$  dans  $\mathcal{F}$ . Si  $\mathcal{S}$  est aussi un espace vectoriel dans  $\mathcal{F}$  utilisant les mêmes opérations d'additions et multiplications qu'un scalaire, donc  $\mathcal{S}$  est dit un sous-espace de  $\mathcal{V}$ . Ce n'est pas nécessaire de vérifier les dix conditions pour déterminer si le sous-ensemble est un sous-espace. Il faut seulement vérifier les conditions de fermeture (A1) et (M1). Donc, un sous-espace non vide  $\mathcal{S}$  d'un espace vectoriel  $\mathcal{V}$  est un sous-espace de  $\mathcal{V}$  si et seulement si :

- (A1)  $\mathbf{x}, \mathbf{y} \in \mathcal{S} \implies \mathbf{x} + \mathbf{y} \in \mathcal{S}$ .
- (M1)  $\mathbf{x} \in \mathcal{S} \implies \alpha\mathbf{x} \in \mathcal{S}$  pour tout  $\alpha \in \mathcal{F}$ .

Un exemple d'un sous-espace vectoriel dans un espace vectoriel  $\mathbb{R}$  est  $\mathcal{L} = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{y} = \alpha\mathbf{x}\}$ . Ainsi, le sous-espace  $\mathcal{L}$  satisfait la définition 5.1.1 d'un espace vectoriel. Un élément

important de la preuve est de montrer que les composantes de  $\mathbf{L}_k(\mathbf{x})$  définies par (4.2.7), aussi voir (4.2.2), forment une base dans un certain espace tel que défini en (5.1.3).

**Définition 5.1.3.** *Un ensemble de vecteurs  $\mathcal{W}$  est appelé une base de  $\mathcal{V}$  si  $\mathcal{W}$  est un système linéairement indépendant et si chaque vecteur de  $\mathcal{V}$  est une combinaison linéaire d'éléments dans  $\mathcal{W}$ .*

Plus particulièrement, il sera montré que  $\mathbf{L}_k(\mathbf{x})$  est un ensemble orthonormé dans un espace vectoriel, ce qui mènera à une base orthonormale, tel que décrit dans la définition 5.1.4.

**Définition 5.1.4.** *Posons  $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ . On dira que  $\mathcal{B}$  est un ensemble orthonormé si  $\|\mathbf{u}_i\| = 1$  pour tout  $i$  et  $\mathbf{u}_i \perp \mathbf{u}_j$  pour tout  $i \neq j$ , c'est-à-dire :*

$$\mathbf{u}_i^T \mathbf{u}_j = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j, \end{cases}$$

où  $\perp$  signifie perpendiculaire. Tout ensemble orthonormé est linéairement indépendant. Également, tout ensemble orthonormé de  $n$  vecteurs d'un espace de dimension  $n$  est une base orthonormée de  $\mathcal{V}$ . Aussi, il est possible de passer d'une base à une autre en utilisant une transformation linéaire qui est nommée une matrice de passage, tel qu'énoncé dans le théorème 5.1.1.

**Théorème 5.1.1.** *Pour une base orthonormale  $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ , un vecteur  $\mathbf{v}$  quelconque peut s'écrire de la façon suivante :*

$$\mathbf{v} = \sum_{j=1}^n (\mathbf{v}^T \mathbf{u}_j) \mathbf{u}_j = \sum_{j=1}^n \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j. \quad (5.1.1)$$

Un élément important de la preuve est de montrer que la matrice de passage entre deux bases orthonormales est orthogonale. Dans l'Annexe B, on introduit la matrice de passage entre deux bases. Ce résultat est énoncé dans le théorème 5.1.2.

**Théorème 5.1.2.** *Une matrice de passage  $\mathbf{P}$  entre deux bases orthonormales est une matrice orthogonale.*

La preuve du théorème 5.1.2 est fournie en annexe. Les composantes du vecteur  $\mathbf{L}_k(\mathbf{x})$  sont des polynômes orthogonaux multivariés généralisés. Ceux-ci font l'objet de la prochaine section.

### 5.1.1. Polynômes orthogonaux multivariés généralisés

Il est important d'adapter la théorie d'algèbre dans le contexte de la statistique  $U_k$ . Soit  $\mathbf{x} = (x_1, \dots, x_m)^T$ , un vecteur de dimension  $m$ . Le produit de ces variables est un monôme et il est noté  $\mathbf{x}^{\mathbf{j}} = x_1^{j_1} \dots x_m^{j_m}$  avec un degré total de :

$$|\mathbf{j}| = j_1 + \dots + j_m, \quad (5.1.2)$$

avec  $j_i \geq 0$ . Par exemple, si  $\mathbf{x} = (x_1, x_2, x_3)^T$  et  $\mathbf{j} = (2, 4, 1)$ , nous avons alors que  $\mathbf{x}^{\mathbf{j}} = x_1^2 x_2^4 x_3$ , un monôme de degré 7. Une combinaison linéaire de  $\mathbf{x}^{\mathbf{j}}$ , avec  $|\mathbf{j}| = k \in \mathbb{N}$  est l'espace des polynômes homogènes de degré  $k$  noté :

$$\mathcal{P}_k^m : \{\mathbf{x}^{\mathbf{j}} : |\mathbf{j}| = k, \mathbf{j} \in \mathbb{N}^m\}. \quad (5.1.3)$$

L'espace des polynômes homogènes de degré maximal  $k$  est noté :

$$\Pi_R^m : \{\mathbf{x}^{\mathbf{j}} : 0 \leq |\mathbf{j}| \leq R, \mathbf{j} \in \mathbb{N}^m\}, \quad (5.1.4)$$

et l'espace de tous les polynômes de  $m$  variables est noté par :

$$\Pi^m = \mathbb{R}[x_1, \dots, x_m]. \quad (5.1.5)$$

La dimension de  $\mathcal{P}_k^m$  est  $d(m, k) = \binom{m+k-1}{k}$  et la dimension de  $\Pi_R^m$  est  $D(m, R) = \sum_{k=1}^R d(m, k)$ . Ces affirmations se retrouvent dans la section 3 de Rahman (2017).

### 5.1.2. Polynômes orthogonaux à une densité quelconque

Soit  $\mathbf{X} = (X_1, \dots, X_m)^T$ ,  $m \in \mathbb{N}$ , un vecteur aléatoire avec une fonction de répartition continue  $F_{\mathbf{X}}(\mathbf{x})$  et une fonction de densité  $f_{\mathbf{X}}(\mathbf{x}) = \partial^m F_{\mathbf{X}}(\mathbf{x}) / \partial x_1 \dots \partial x_m$ . En utilisant l'équation (5.1.7), les moments de tous les ordres de la fonction de densité sont supposés exister, pour tout  $\mathbf{j} \in \mathbb{N}^m$  :

$$\mu_{\mathbf{j}} = \mathbb{E} \{ |\mathbf{X}^{\mathbf{j}}| \} = \int |\mathbf{x}^{\mathbf{j}}| f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} < \infty. \quad (5.1.6)$$

Pour n'importe quelle paire de polynômes  $P, Q \in \Pi^m$ , le produit scalaire par rapport à la densité  $f_{\mathbf{X}}(\mathbf{x})$  est :

$$\langle P, Q \rangle_{f_{\mathbf{X}}(\mathbf{x})} = \int P(\mathbf{x}) Q(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \mathbb{E} \{ P(\mathbf{X}) Q(\mathbf{X}) \}. \quad (5.1.7)$$

Il s'ensuit que la norme induite par ce produit est :

$$\|P\|_{f_{\mathbf{X}}(\mathbf{x})} = \sqrt{\langle P, P \rangle_{f_{\mathbf{X}}(\mathbf{x})}} = \left( \int P^2(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \right)^{1/2} = \sqrt{\mathbb{E}\{P^2(\mathbf{X})\}}.$$

Les polynômes  $P \in \Pi^m$  et  $Q \in \Pi^m$  sont orthogonaux par rapport à  $f_{\mathbf{X}}(\mathbf{x})$  si  $\langle P, Q \rangle_{f_{\mathbf{X}}(\mathbf{x})} = 0$ . La définition 5.1.5 présente la définition d'un polynôme orthogonal provenant de Dunkl et Xu (2014).

**Définition 5.1.5.** *Un polynôme  $P \in \Pi_k^m \subset \Pi^m$  est un polynôme orthogonal de degré  $k \in \mathbb{N}$  par rapport à un produit scalaire  $\langle \cdot, \cdot \rangle_{f_{\mathbf{X}}(\mathbf{x})}$  ou alternativement par rapport à la densité  $f_{\mathbf{X}}(\mathbf{x})$ , si le polynôme est orthogonal à tous les polynômes de degré inférieur tel que :*

$$\langle P, Q \rangle_{f_{\mathbf{X}}(\mathbf{x})} = 0, \quad \forall Q \in \Pi^m \text{ avec } \deg Q < \deg P,$$

où  $\deg$  signifie degré. Il existe une infinité d'ensembles de polynômes orthogonaux multivariés, noté  $\{P_{\mathbf{j}}(x) : \mathbf{j} \in \mathbb{N}\}$ , avec  $P_{\mathbf{0}} = 1$  et  $P_{\mathbf{j}} \neq 0$ , qui sont orthogonaux par rapport au produit scalaire  $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$  satisfaisant :

$$\langle P_{\mathbf{j}}, P_{\mathbf{i}} \rangle_{f_{\mathbf{X}}(\mathbf{x})} = 0, \quad \text{quand } |\mathbf{j}| \neq |\mathbf{i}|,$$

pour  $\mathbf{i} \in \mathbb{N}^m$  et  $|\mathbf{j}|$  qui est le degré total du polynôme multivarié tel que défini en (5.1.2). De la définition 5.1.5, chaque polynôme  $P_{\mathbf{j}} \in \Pi^m$  est un polynôme orthogonal. Ceci implique qu'il est orthogonal à tous les polynômes de degré inférieur, mais pas forcément orthogonal à un polynôme du même degré.

Soit  $\mathcal{V}_0^m$ , l'espace des fonctions constantes. Pour chaque  $1 \leq k \leq \infty$  noté par  $\mathcal{V}_k^m \subset \Pi_k^m$ , l'espace de tous les polynômes orthogonaux de degré  $k$  qui sont orthogonaux à tous les polynômes dans  $\Pi_{k-1}^m$  est écrit de la façon suivante :

$$\mathcal{V}_k^m = \{P_k^m : \langle P, Q \rangle_{f_{\mathbf{X}}(\mathbf{x})} = 0 \quad \forall Q \in \Pi_{k-1}^m\}, \quad 1 \leq k \leq \infty. \quad (5.1.8)$$

La dimension de  $\mathcal{V}_k^m$  est  $d(m, k)$  à condition que  $f_{\mathbf{X}}(\mathbf{x})$  soit non vide. Il existe plusieurs choix de bases pour  $\mathcal{V}_k^m$  ; les bases de  $\mathcal{V}_k^m$  ne sont pas forcément mutuellement orthogonales. Les bases ne sont pas uniques dans le cas multivarié. Soit  $\{P_{\mathbf{j}} : |\mathbf{j}| = k, \mathbf{j} \in \mathbb{N}^m\} \subset \mathcal{V}_k^m$ , une base de  $\mathcal{V}_k^m$  comprenant  $d(m, k)$  fonctions de bases. Chaque fonction  $P_{\mathbf{j}}(\mathbf{x})$  est un polynôme orthogonal multivarié de degré  $|\mathbf{j}|$ , tel que :

$$\mathcal{V}_k^m = \{P_{\mathbf{j}} : |\mathbf{j}| = k, \mathbf{j} \in \mathbb{N}\}, \quad 1 \leq k \leq \infty. \quad (5.1.9)$$

Selon la définition de 5.1.5,  $P_j$  est orthogonal à  $P_i$  quand  $|\mathbf{j}| \neq |\mathbf{i}|$ . Donc, pour deux sous-espaces de polynômes  $\mathcal{V}_k^m$  et  $\mathcal{V}_r^m$ , avec  $0 \leq k, r \leq \infty$ , ceux-ci sont orthogonaux quand  $k \neq r$ . L'absence d'orthogonalité entre deux polynômes du même degré peut caractériser l'orthogonalité. En effet, il existe deux type d'orthogonalité, soit l'orthogonalité faible et l'orthogonalité forte.

**Définition 5.1.6.** Soit  $\mathbf{X} = (X_1, \dots, X_m)^T$ ,  $m \in \mathbb{N}$ , un vecteur aléatoire de fonction de densité  $f_{\mathbf{X}}(\mathbf{x})$  avec tous les moments finis en valeur absolue. Une famille de polynômes orthogonaux multivariés  $\{P_{\mathbf{k}}(\mathbf{x}) : \mathbf{k} \in \mathbb{N}\}$  est dite faiblement orthogonale par rapport à la densité  $f_{\mathbf{X}}(\mathbf{x})$  si pour tout  $\mathbf{k}, \mathbf{j} \in \mathbb{N}$  :

$$\langle P_{\mathbf{k}}, P_{\mathbf{j}} \rangle_{f_{\mathbf{X}}(\mathbf{x})} = 0, \quad \text{quand } |\mathbf{k}| \neq |\mathbf{j}|,$$

et est dite fortement orthogonale si pour tout  $\mathbf{k}, \mathbf{j} \in \mathbb{N}$  :

$$\langle P_{\mathbf{k}}, P_{\mathbf{j}} \rangle_{f_{\mathbf{X}}(\mathbf{x})} = 0, \quad \text{quand } \mathbf{k} \neq \mathbf{j}.$$

Si une famille de polynômes est fortement orthogonale, alors elle est aussi faiblement orthogonale. En effet, il est possible de vérifier en prenant les polynômes d'Hermite présentés dans le chapitre 4. Par exemple, si nous posons  $P_{\mathbf{k}}(\mathbf{x}) = H_1(x_1) \cdot H_1(x_2)$  et  $P_{\mathbf{j}}(\mathbf{x}) = H_0(x_1) \cdot H_2(x_2)$ , il est possible de vérifier que le produit scalaire de ces polynômes de degré deux sont orthogonaux à la loi normale multivariés. Ainsi, il est possible de prouver que la famille des polynômes d'Hermite est fortement et faiblement orthogonale, tel qu'énoncé dans la section suivante. Cependant, la réciproque n'est généralement pas vraie, par exemple quand les variables sont statistiquement dépendantes. Lorsque  $\mathbf{X}$  est formé de variables indépendantes, le système de polynômes multivariés est fortement et faiblement orthogonal seulement si celui-ci est obtenu par combinaison linéaire d'espaces polynomiaux univariés. Dans la prochaine section, les polynômes orthogonaux d'Hermite multivariés sont présentés, tel que décrit à la page 308 de Rahman (2017).

### 5.1.3. Polynômes orthogonaux d'Hermite multivariés

Plutôt qu'utiliser des polynômes orthogonaux à une densité quelconque, le développement des polynômes orthogonaux est fait sur la densité normale multivariée dans  $\mathbb{R}^m$  avec tous ses moments finis et connus. Les polynômes orthogonaux de  $\mathbf{x}$  par rapport au

produit scalaire satisfont :

$$\langle P, Q \rangle = \int_{\mathbb{R}^m} P(\mathbf{x})Q(\mathbf{x})\phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}})d\mathbf{x} = \mathbb{E}\{P(\mathbf{X})Q(\mathbf{X})\}, \quad (5.1.10)$$

où  $\phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}})$  est la densité de la loi normale multivariée décrite en (4.2.1). Une famille très connue de polynômes orthogonaux à la loi normale est la famille des polynômes d’Hermite. Ils ont été présentés en version univariée dans le chapitre 4 et sont maintenant présentés dans leur version multivariée. Une façon populaire d’obtenir la définition est en dérivant la densité de la loi normale multivariée provenant de Dunkl et Xu (2014), Holmquist (1996) et Takemura et Takeuchi (1988).

**Définition 5.1.7.** Soit  $\mathbf{X} = (X_1, \dots, X_m)^T$ ,  $m \in \mathbb{N}$ , un vecteur aléatoire de  $\mathbb{R}^m$  avec une moyenne zéro, une matrice des variances et des covariances et définie positive  $\boldsymbol{\Sigma}_{\mathbf{X}}$  et soit  $\phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}})$  la densité de la loi normale multivariée. Les polynômes orthogonaux d’Hermite multivariés  $H_{\mathbf{r}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}})$ ,  $\mathbf{r} = (r_1, \dots, r_m)$ , de degré  $|\mathbf{r}| = r_1 + \dots + r_m$  sont définis par :

$$H_{\mathbf{r}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}}) = \frac{(-1)^{|\mathbf{r}|}}{\phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}})} \left( \frac{\partial}{\partial \mathbf{x}} \right)^{\mathbf{r}} \phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{X}}), \quad (5.1.11)$$

où  $(\partial/\partial \mathbf{x})^{\mathbf{r}} = \partial^{r_1+\dots+r_m}/\partial x_1^{r_1} \dots \partial x_m^{r_m}$ . Si les variables aléatoires gaussiennes sont indépendantes, la matrice des variances et des covariances est une matrice diagonale telle que  $\boldsymbol{\Sigma}_{\mathbf{X}} = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$  avec  $0 < \sigma_i^2 < \infty$ ,  $i = 1, \dots, m$  étant la variance de la  $i^{\text{ème}}$  variable. Si  $\sigma_i^2 = 1$ , pour chaque  $i$ , la matrice est  $\boldsymbol{\Sigma}_{\mathbf{X}} = \mathbf{I}_m$ . Donc, la loi normale multivariée de  $\mathbf{X}$  devient le produit des lois marginales normales des  $x_i$  et la définition 5.1.7 d’un polynôme d’Hermite multivarié de degré  $|\mathbf{k}|$  se simplifie aux produits des polynômes d’Hermite univariés :

$$H_{\mathbf{k}}(\mathbf{x}, \boldsymbol{\mu}, \mathbf{I}_m) = H_{r_1}(x_1) \cdot \dots \cdot H_{r_m}(x_m),$$

avec  $\sum_{i=1}^m r_i = k$ . Selon la définition 5.1.7, la famille des polynômes d’Hermite multivariés  $\{H_{\mathbf{k}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}), \mathbf{k} \in \mathbb{N}^m\}$  pour des variables gaussiennes est faiblement orthogonale par rapport au produit scalaire  $\langle \cdot, \cdot \rangle_{\phi_{\mathbf{X}}}$  tel que :

$$\langle H_{\mathbf{k}}, H_{\mathbf{j}} \rangle_{\phi_{\mathbf{X}}} = \mathbb{E} \{H_{\mathbf{k}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})H_{\mathbf{j}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})\} = 0, \quad |\mathbf{k}| \neq |\mathbf{j}|.$$

Ceci implique que  $H_{\mathbf{k}}$  est orthogonal à tous les polynômes de degré différent, mais pas forcément orthogonal à des polynômes du même degré. Cependant, si les variables gaussiennes sont indépendantes,  $H_{\mathbf{k}}$  est un polynôme d’Hermite multivarié fortement orthogonal. Ceci est dû à la structure du produit des polynômes où deux polynômes univariés de degré différent sont orthogonaux. Le théorème 5.1.3, provenant du corollaire 9 de Rahman (2017), présente les premiers moments des polynômes d’Hermite multivariés.

**Théorème 5.1.3.** *Soit  $\mathbf{X} = (X_1, \dots, X_m)^T$  provenant de la loi gaussienne indépendante avec une moyenne zéro, une matrice des variances et des covariances  $\Sigma_{\mathbf{X}} = \mathbf{I}_m$  résultant à des polynômes d’Hermite orthonormaux multivariés  $\Psi_{\mathbf{j}}(\mathbf{X}, \boldsymbol{\mu}, \mathbf{I}_m)$ . Les premiers moments sont :*

$$\mathbb{E} \{ \Psi_{\mathbf{j}}(\mathbf{X}, \boldsymbol{\mu}, \mathbf{I}_m) \} = \begin{cases} 1, & \text{si } \mathbf{j} = \mathbf{0}, \\ 0, & \text{si } \mathbf{j} \neq \mathbf{0}. \end{cases}$$

et

$$\mathbb{E} \{ \Psi_{\mathbf{j}}(\mathbf{X}, \boldsymbol{\mu}, \mathbf{I}_m) \Psi_{\mathbf{k}}(\mathbf{X}, \boldsymbol{\mu}, \mathbf{I}_m) \} = \begin{cases} 1, & \text{si } \mathbf{j} = \mathbf{k}, \\ 0, & \text{si } \mathbf{j} \neq \mathbf{k}. \end{cases}$$

La proposition 5.1.1 montre que les polynômes d’Hermite constituent une base  $\mathcal{V}_k^m$ . Soit  $L^2(\mathbb{R}^m, \mathcal{B}^m, \phi_{\mathbf{X}})$ , un espace d’Hilbert de fonction carré intégrable par rapport à la densité gaussienne  $\phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \Sigma_{\mathbf{X}})$  de  $\mathbb{R}^m$ .

**Proposition 5.1.1.** *Soit  $\mathbf{X} = (X_1, \dots, X_m)^T : (\omega, \mathcal{F}) \rightarrow (\mathbb{R}^m, \mathcal{B}^m)$ ,  $m \in \mathbb{N}$ , un vecteur aléatoire de  $\mathbb{R}^m$  avec une moyenne zéro, une matrice des variances et des covariances symétrique et définie positive  $\Sigma_{\mathbf{X}}$  et soit  $\phi_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \Sigma_{\mathbf{X}})$  la densité de la loi normale multivariée. Soit  $\{H_{\mathbf{r}} : |\mathbf{r}| = k, \mathbf{r} \in \mathbb{N}^m\}$ , un ensemble de polynômes orthogonaux multivariés d’Hermite de degré  $k$  avec la densité de la loi gaussienne  $\phi_{\mathbf{X}}$ , alors  $H_{\mathbf{r}}(\mathbf{x})$  est une base de  $\mathcal{V}_k^m$ .*

La proposition 5.1.1 et sa preuve sont présentées dans la proposition 11 de la page 313 de Rahman (2017). Aussi, la preuve de l’invariance de la statistique  $U_k$  énoncée en (4.3.5), est présentée dans le prochaine section.



## 5.2. Preuve de l'invariance de $U_k$

D'abord, le lemme 4.3.1 est prouvé et il sera utilisé pour montrer l'invariance de la statistique de test  $U_k$  aux transformations linéaires affines.

Décortiquons la statistique (4.3.5) :

$$U_k(\hat{\boldsymbol{\nu}}_{CML}) = \|\mathbf{V}_k(\hat{\boldsymbol{\nu}}_{CML})\|^2 = \sum_{s=1}^{d(m,k)} V_{s:k}^2(\hat{\boldsymbol{\nu}}_{CML}),$$

avec

$$V_{s:k}(\hat{\boldsymbol{\nu}}_{CML}) = n^{-1/2} \sum_{t=1}^n L_{s:k} \left\{ \hat{\mathbf{C}} \mathbf{Z}_t(\hat{\boldsymbol{\nu}}_{CML}) \right\}.$$

Le vecteur  $\mathbf{L}_k(\mathbf{x}) = (L_{1:k}(x), \dots, L_{d(m,k):k}(x))^T$  est un vecteur composé de produits de polynômes d'Hermite univariés dont le degré total est  $k$ , tel que défini en (4.2.2) :

$$L_{r:k}(\mathbf{x}) = H_{r_1}(x_1) \cdot \dots \cdot H_{r_m}(x_m) = \prod_{i=1}^m H_{r_i}(x_i),$$

avec  $\sum_{i=1}^m r_i = k$  et  $H_{r_i}(x_i)$  est un polynôme de degré  $r_i$ . Le produit de polynômes  $L_{r:k}(\mathbf{x})$  est une combinaison linéaire de variables dont le degré total est  $k$  formant ainsi un polynôme homogène de degré  $k$ . Grâce à la section 5.1, il est possible d'affirmer que chaque composante du vecteur  $\mathbf{L}_k(\mathbf{x})$  est un polynôme orthogonal d'Hermite multivarié de degré  $k$ . De plus, comme les lois normales sont indépendantes et que les données sont centrées et réduites, les polynômes vont être fortement orthogonaux. Ce qui implique que chaque élément de  $\mathbf{L}_k(\mathbf{x})$  est orthogonal un à un. La proposition 5.1.1 permet de conclure que  $\mathbf{L}_k(\mathbf{x})$  est une base de  $\mathcal{V}_k^m$ .

### *Preuve du lemme 4.3.1.*

Un élément important de la preuve du théorème 4.3.1 est l'utilisation du lemme 5.2.1.

**Lemme 5.2.1.** *Lorsque la famille des polynômes d'Hermite est utilisée, le résultat suivant est obtenu. Pour toute matrice orthogonale  $\mathbf{Q}$  de taille  $m \times m$ , il existe une matrice orthogonale  $\mathbf{P}(m, k)$  de taille  $d(m, k) \times d(m, k)$  pour tout vecteur  $\mathbf{Z} \in \mathbb{R}^m$ , tel que :*

$$\mathbf{L}_k(\mathbf{Q}\mathbf{Z}) = \mathbf{P}(m, k)\mathbf{L}_k(\mathbf{Z}).$$

*Démonstration.* Soit  $\mathbf{L}_k(\mathbf{Q}\mathbf{Z})$  et  $\mathbf{L}_k(\mathbf{Z})$ , deux bases de  $\mathcal{V}_k^m$ . L'idée de la preuve est de montrer que la matrice de passage reliant les deux bases est orthogonale. D'abord, montrons que les bases sont orthonormées. Selon la définition 5.1.4, il faut que le produit scalaire entre chaque élément distinct de  $\mathbf{L}_k(\mathbf{Z})$  soit zéro et que la norme d'un élément soit un. Le produit scalaire, tel que défini en (5.1.10) par rapport à la loi normale multivariée, est :

$$\langle L_{i:k}(\mathbf{Z}), L_{j:k}(\mathbf{Z}) \rangle_{\phi_{\mathbf{X}}} = \mathbb{E}\{L_{i:k}(\mathbf{Z})L_{j:k}(\mathbf{Z})\}.$$

En utilisant le deuxième moment des polynômes d'Hermite multivariés qui se retrouve dans le théorème 5.1.3, on trouve que :

$$\mathbb{E}\{L_{i:k}(\mathbf{Z})L_{j:k}(\mathbf{Z})\} = \begin{cases} 1, & \text{si } j = i, \\ 0, & \text{si } j \neq i, \end{cases}$$

donc  $\mathbf{L}_k$  est une base orthonormée. Il est possible de conclure que  $\mathbf{P}(m, k)$  est orthogonale en utilisant le théorème 5.1.2.

*Application du lemme 5.2.1.* Soit une matrice orthogonale  $\mathbf{Q}$  :

$$V_{s:k}^{\mathbf{Q}}(\boldsymbol{\nu}) = \frac{1}{n^{1/2}} \sum_{t=1}^n L_{s:k} \left\{ \mathbf{Q}\tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\}, \quad U_k^{\mathbf{Q}}(\boldsymbol{\nu}) = \sum_{s=1}^{d(m,k)} \left\{ V_{s:k}^{\mathbf{Q}}(\boldsymbol{\nu}) \right\}^2 = \|\mathbf{V}_k^{\mathbf{Q}}\|^2,$$

où  $\tilde{\mathbf{Z}}(\boldsymbol{\nu}) = \mathbf{C}\mathbf{Z}_t(\boldsymbol{\beta}, \boldsymbol{\mu})$ . Le lemme 4.2.1 stipule que  $U_k^{\mathbf{Q}}(\boldsymbol{\nu}) = U_k(\boldsymbol{\nu})$ . Développons  $U_k^{\mathbf{Q}}(\boldsymbol{\nu})$  et  $U_k(\boldsymbol{\nu})$ .

$$\begin{aligned} U_k(\boldsymbol{\nu}) &= \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\} \right\}^T \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\} \right\} = \|\mathbf{V}_k\|^2, \\ U_k^{\mathbf{Q}}(\boldsymbol{\nu}) &= \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \mathbf{Q}\tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\} \right\}^T \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \mathbf{Q}\tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\} \right\}, \\ &= \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\} \right\}^T \mathbf{P}(m, k)^T \mathbf{P}(m, k) \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \tilde{\mathbf{Z}}_t(\boldsymbol{\nu}) \right\} \right\}, \\ &= U_k(\boldsymbol{\nu}). \end{aligned}$$

### 5.2.1. Preuve du théorème 4.3.1

La preuve du théorème 4.3.1 utilise les mêmes étapes que celle du lemme 1 de Mardia (1970). La principale différence avec la démarche démontrée ici est qu'un échantillon est utilisé, alors que dans le cas de l'approche de Mardia (1970) un seul vecteur aléatoire est utilisé. De plus, la preuve de Mardia montre que la statistique est invariante avec un  $\mathbf{U}$  satisfaisant  $\mathbf{U}\Sigma\mathbf{U}^T = \mathbf{I}_m$ . Comme décrit dans la section 4.2, les résidus sont centrés avec  $\mathbf{C}$  satisfaisant  $\mathbf{C}\Sigma\mathbf{C}^T = \mathbf{I}_m$ . La statistique (4.3.5), avec l'estimation des paramètres est :

$$U_k(\tilde{\mathbf{Z}}_t) = \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \hat{\mathbf{C}} \hat{\mathbf{Z}}_t(\boldsymbol{\beta}, \boldsymbol{\mu}) \right\} \right\}^T \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k \left\{ \hat{\mathbf{C}} \hat{\mathbf{Z}}_t(\boldsymbol{\beta}, \boldsymbol{\mu}) \right\} \right\},$$

avec  $\tilde{\mathbf{Z}}_t$ , les résidus du modèle qui sont décrits dans la section 4.1. Considérons l'échantillon de vecteurs  $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n)$  et posons

$$\mathbf{X} = \mathbf{A}\mathbf{Y}, \quad (5.2.1)$$

avec  $\mathbf{A}$  une matrice inversible. Soit  $\tilde{\mathbf{Z}}_Y$  et  $\tilde{\mathbf{Z}}_X$ , les résidus du modèle pour les échantillons  $\mathbf{Y}$  et  $\mathbf{X}$ . En raison de la méthode d'estimation par la méthode de la vraisemblance, la relation

$$\tilde{\mathbf{Z}}_X = \mathbf{A}\tilde{\mathbf{Z}}_Y, \quad (5.2.2)$$

est maintenue. Les moyennes de  $\tilde{\mathbf{Z}}_Y$  et  $\tilde{\mathbf{Z}}_X$  sont  $\mathbf{0}$  et leurs matrices des variances et des covariances sont respectivement  $\Sigma_{\tilde{\mathbf{Z}}_Y}$  et  $\Sigma_{\tilde{\mathbf{Z}}_X}$ , des matrices définies positives. Il existe des matrices inversibles  $\mathbf{G}_{\tilde{\mathbf{Z}}_Y}$  et  $\mathbf{G}_{\tilde{\mathbf{Z}}_X}$  satisfaisant :

$$\mathbf{G}_{\tilde{\mathbf{Z}}_Y} \Sigma_{\tilde{\mathbf{Z}}_Y} \mathbf{G}_{\tilde{\mathbf{Z}}_Y}^T = \mathbf{I}_m, \quad (5.2.3)$$

$$\mathbf{G}_{\tilde{\mathbf{Z}}_X} \Sigma_{\tilde{\mathbf{Z}}_X} \mathbf{G}_{\tilde{\mathbf{Z}}_X}^T = \mathbf{I}_m, \quad (5.2.4)$$

avec  $\mathbb{V}(\tilde{\mathbf{Z}}_X) = \Sigma_{\tilde{\mathbf{Z}}_X} = \mathbf{A}\Sigma_{\tilde{\mathbf{Z}}_Y}\mathbf{A}^T$ .

L'extension de  $U_k$  à cet échantillon de vecteurs est donnée par :

$$U_k^*(\tilde{\mathbf{Z}}_Y) = \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k(\mathbf{G}_{\tilde{\mathbf{Z}}_Y} \tilde{\mathbf{Z}}_Y) \right\}^T \left\{ \frac{1}{n^{1/2}} \sum_{t=1}^n \mathbf{L}_k(\mathbf{G}_{\tilde{\mathbf{Z}}_Y} \tilde{\mathbf{Z}}_Y) \right\} = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_Y} \tilde{\mathbf{Z}}_Y). \quad (5.2.5)$$

Comme la décomposition de (5.2.3) n'est pas unique ( $\mathbf{G}_{\tilde{\mathbf{Z}}_Y}$  n'est pas unique), ce que l'on doit montrer pour conclure à l'invariance est de s'assurer que  $U_k^*$  étendra  $U_k$  de façon unique. En d'autres termes, cela signifie que  $U_k(\mathbf{G}\tilde{\mathbf{Z}}_Y) = U_k(\mathbf{M}\tilde{\mathbf{Z}}_Y)$ , si  $\mathbf{G}$  et  $\mathbf{M}$  satisfont l'équation (5.2.3). Or, il est possible de trouver un deuxième candidat pour  $\tilde{\mathbf{Z}}_Y$ .

$$U_k^*(\tilde{\mathbf{Z}}_X) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_X}\tilde{\mathbf{Z}}_X) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_X}\mathbf{A}\tilde{\mathbf{Z}}_Y) = U_k(\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{Z}}_Y),$$

avec  $\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y} = \mathbf{G}_{\tilde{\mathbf{Z}}_X}\mathbf{A}$  qui satisfait :

$$\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y}\Sigma_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y}^T = \mathbf{G}_{\tilde{\mathbf{Z}}_X}\mathbf{A}\Sigma_{\tilde{\mathbf{Z}}_Y}\mathbf{A}^T\mathbf{G}_{\tilde{\mathbf{Z}}_X}^T = \mathbf{G}_{\tilde{\mathbf{Z}}_X}\Sigma_{\tilde{\mathbf{Z}}_X}\mathbf{G}_{\tilde{\mathbf{Z}}_X}^T = \mathbf{I}_m.$$

Nous montrons que  $U_k^*$  étend  $U_k$  de façon unique. Pour ce faire, il est utile d'introduire le lemme 5.2.2 qui montre une relation orthogonale entre deux matrices qui satisfont l'équation (5.2.3).

**Lemme 5.2.2.** *Soit  $\Sigma$  une matrice symétrique et définie positive de dimension  $n \times n$ . Considérons une matrice  $\mathbf{U}$  telle que :*

$$\mathbf{U}\Sigma\mathbf{U}^T = \mathbf{I}.$$

Alors, pour une seconde matrice  $\mathbf{V}$  satisfaisant :

$$\mathbf{V}\Sigma\mathbf{V}^T = \mathbf{I},$$

alors forcément :

$$\mathbf{V} = \mathbf{C}\mathbf{U},$$

où  $\mathbf{C}$  est une matrice orthogonale.

*Démonstration.* Pour montrer le lemme, notons que la matrice  $\mathbf{U}$  existe, puisque  $\Sigma$  étant symétrique et définie positive, alors elle est orthogonalement diagonalisable et :

$$\begin{aligned}\Sigma &= \mathbf{P}\mathbf{D}_\lambda\mathbf{P}^T, \\ &= \sum_{i=1}^n \lambda_i \mathbf{P}_i \mathbf{P}_i^T,\end{aligned}$$

où  $\mathbf{D}_\lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  et  $\lambda_i > 0$ ,  $i = 1, \dots, n$ . Ici, on a que  $\mathbf{P} = (\mathbf{P}_1, \dots, \mathbf{P}_n)$ , une matrice orthogonale. Posant  $\mathbf{U} = \mathbf{D}_1/\sqrt{\lambda}\mathbf{P}^t$ , où  $\mathbf{D}_1/\sqrt{\lambda} = \text{diag}(1/\sqrt{\lambda_1}, \dots, 1/\sqrt{\lambda_n})$ , on montre bien l'existence de  $\mathbf{U}$ , puisque :

$$\begin{aligned} \mathbf{U}\Sigma\mathbf{U}^T &= \mathbf{D}_1/\sqrt{\lambda}\mathbf{P}^T\Sigma\mathbf{P}\mathbf{D}_1/\sqrt{\lambda}, \\ &= \mathbf{D}_1/\sqrt{\lambda}\mathbf{D}_\lambda\mathbf{D}_1/\sqrt{\lambda}, \\ &= \mathbf{I}. \end{aligned}$$

Posons maintenant  $\mathbf{C} = \mathbf{V}\mathbf{U}^{-1}$ . Afin de montrer que  $\mathbf{C}$  est orthogonale, on calcule directement  $\mathbf{C}\mathbf{C}^T$  :

$$\begin{aligned} \mathbf{C}\mathbf{C}^T &= \mathbf{V}\mathbf{U}^{-1}(\mathbf{U}^{-1})^T\mathbf{V}^T, \\ &= \mathbf{V}\mathbf{U}^{-1}(\mathbf{U}^T)^{-1}\mathbf{V}^T, \\ &= \mathbf{V}\Sigma\mathbf{V}^T, \\ &= \mathbf{I}, \end{aligned}$$

ce qui montre le lemme.

Il est alors possible de montrer que  $U_k^*$  s'étend  $U_k$  de façon unique. Soit  $\mathbf{G}_{\tilde{\mathbf{Z}}_Y}$  et  $\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y}$ , deux candidats satisfaisant (5.2.3). En utilisant le lemme 5.2.2, il existe une matrice orthogonale  $\mathbf{Q}$  telle que  $\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y} = \mathbf{Q}\mathbf{G}_{\tilde{\mathbf{Z}}_Y}$ . Alors, il est possible de réécrire  $U_k(\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{Z}}_Y)$  de la façon suivante :

$$U_k(\tilde{\mathbf{G}}_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{Z}}_Y) = U_k(\mathbf{Q}\mathbf{G}_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{Z}}_Y) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{Z}}_Y).$$

Cette égalité est obtenue grâce au lemme 5.2.1. Il est alors possible de conclure, s'il existe deux candidats pour  $\tilde{\mathbf{Z}}_Y$ , que  $U_k^*$  s'étendra de façon unique tel que :

$$U_k^*(\tilde{\mathbf{Z}}_X) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_X}\mathbf{A}\tilde{\mathbf{Z}}_Y),$$

et donc, en utilisant encore (5.2.5) et (5.2.2), il est possible de conclure sur l'invariance.

$$U_k^*(\tilde{\mathbf{Z}}_X) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_X}\tilde{\mathbf{Z}}_X) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_Y}\mathbf{A}^{-1}\tilde{\mathbf{Z}}_X) = U_k(\mathbf{G}_{\tilde{\mathbf{Z}}_Y}\tilde{\mathbf{Z}}_Y) = U_k^*(\tilde{\mathbf{Z}}_Y).$$

Dans le prochain chapitre, les résultats des simulations de type Monte Carlo sont présentés.

# Chapitre 6

---

## Simulations

Dans les chapitres 3 et 4, plusieurs tests de normalité ont été introduits. Pour comparer la performance de ces tests, des simulations ont été faites et seront montrées dans ce chapitre. Plus précisément, le niveau de signification et la puissance empirique seront évalués. Pour ce faire, quatre modèles de séries chronologiques multivariées de taille  $n$  ont été choisis. Ces tailles sont 50, 100, 200 et 400. Les tests ont été comparés selon les modèles et leurs tailles d'échantillons avec une simulation Monte Carlo. La fréquence de rejet sous l'hypothèse nulle quand celle-ci est vraie a été regardée sous un seuil nominal de 5 %. La fréquence de rejet sous différentes distributions a été évaluée afin de déterminer la puissance empirique des tests.

### 6.1. Modèles de simulations

Les processus générateurs des données (DGP) suivants ont été utilisés :

$$\text{i) DGP}_1 : \mathbf{Y}_t = \boldsymbol{\mu}_1 + \boldsymbol{\Phi}_1 \mathbf{Y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{Y}_{t-2} + \mathbf{Z}_{1t},$$

$$\text{ii) DGP}_2 : \mathbf{Y}_t = \boldsymbol{\mu}_2 + \boldsymbol{\Phi}_3 \mathbf{Y}_{t-1} + \mathbf{Z}_{2t},$$

$$\text{iii) DGP}_3 : \mathbf{Y}_t = \boldsymbol{\Phi}_4 \mathbf{Y}_{t-1} + \mathbf{Z}_{3t} + \boldsymbol{\Theta}_1 \mathbf{Z}_{3,t-1},$$

$$\text{iii) DGP}_4 : \mathbf{Y}_t = \boldsymbol{\Phi}_5 \mathbf{Y}_{t-1} + \mathbf{Z}_{4t} + \boldsymbol{\Theta}_2 \mathbf{Z}_{4,t-1}.$$

Le premier modèle est un VAR(2) bivarié, le deuxième est un VAR(1) trivarié. Les modèles trois et quatre sont des VARMA(1,1) de dimension deux. Il est important de

savoir que les valeurs zéros des paramètres sont présumées connues et donc, il ne faut pas les estimer. Ceci mène à des modèles avec une paramétrisation structurée. Ce choix est justifié car le nombre de paramètres libres peut atteindre  $m^2(p + q) + m$  dans un modèle ARMA vectoriel avec une moyenne inconnue. Ce nombre croît rapidement en raison du terme  $m^2$ . Plus le nombre de paramètres à estimer est élevé, plus nos estimateurs peuvent être imprécis, surtout lorsque la taille  $n$  est petite. Notons que  $m$  désigne la dimension des modèles. Les paramètres sont détaillés dans les tableaux 6.1 et 6.2 :

**Tableau 6.1.** Matrices des paramètres des modèles de séries chronologiques pour les modèles DGP<sub>1</sub>, DGP<sub>2</sub>, DGP<sub>3</sub> et DGP<sub>4</sub>.

$$\begin{aligned} \boldsymbol{\mu}_1 &= \begin{pmatrix} 0,02 \\ 0,03 \end{pmatrix}, & \boldsymbol{\Phi}_1 &= \begin{pmatrix} 0,5 & 0,1 \\ 0,4 & 0,5 \end{pmatrix}, & \boldsymbol{\Phi}_2 &= \begin{pmatrix} 0,0 & 0,0 \\ 0,25 & 0,0 \end{pmatrix}, \\ \boldsymbol{\mu}_2 &= \begin{pmatrix} 0,01 \\ 0,02 \\ 0,00 \end{pmatrix}, & \boldsymbol{\Phi}_3 &= \begin{pmatrix} 0,5 & 0,0 & 0,0 \\ 0,1 & 0,1 & 0,3 \\ 0,0 & 0,2 & 0,3 \end{pmatrix}, & \boldsymbol{\Phi}_4 &= \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,5 \end{pmatrix}, \\ \boldsymbol{\Theta}_1 &= \begin{pmatrix} 0,1 & 0,0 \\ 0,0 & 0,1 \end{pmatrix}, & \boldsymbol{\Phi}_5 &= \begin{pmatrix} 0,7 & -0,1 \\ 0,0 & -0,6 \end{pmatrix}, & \boldsymbol{\Theta}_2 &= \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,8 \end{pmatrix}. \end{aligned}$$

Les modèles VARMA(1, 1) considérés sont identifiables, voir la section 2.3.3 de Reinsel (1997) pour les conditions d'identifiabilité dans les VARMA(1, 1).

**Tableau 6.2.** Matrices des variances et des covariances des bruits blancs pour les modèles DGP<sub>1</sub>, DGP<sub>2</sub>, DGP<sub>3</sub> et DGP<sub>4</sub>.

$$\begin{aligned} \boldsymbol{\Sigma}_1 &= \begin{pmatrix} 0,09 & 0,00 \\ 0,00 & 0,04 \end{pmatrix}, & \boldsymbol{\Sigma}_2 &= \begin{pmatrix} 2,25 & 0,00 & 0,00 \\ 0,00 & 1,00 & 0,50 \\ 0,00 & 0,50 & 0,74 \end{pmatrix}, \\ \boldsymbol{\Sigma}_3 &= \begin{pmatrix} 1,00 & 0,50 \\ 0,50 & 1,25 \end{pmatrix}, & \boldsymbol{\Sigma}_4 &= \begin{pmatrix} 1,00 & 0,00 \\ 0,00 & 1,00 \end{pmatrix}. \end{aligned}$$



L'espérance et la variance des bruits blancs  $\{\mathbf{Z}_{it}\}$  sont  $\mathbb{E}(\mathbf{Z}_{it}) = 0$  et  $\mathbb{V}(\mathbf{Z}_{it}) = \boldsymbol{\Sigma}_i$ ,  $i = 1, 2, 3, 4$ . Dans ce cas précis, les valeurs des zéros des matrices des variances et des covariances sont inconnues, il faut donc les estimer. La prochaine section présente les tests utilisés dans les simulations de type Monte Carlo.

## 6.2. Tests pour les simulations

Soit  $\hat{\boldsymbol{\nu}}_{i,CML}$ , l'estimateur du maximum de vraisemblance (*CML*) conditionnel d'un des modèles  $DGP_i$  tel que décrit dans le chapitre 4. Soit  $\hat{\mathbf{Z}}_{it}$ ,  $t = 1, \dots, n$  les résidus du modèle associé. L'estimateur de  $\boldsymbol{\Sigma}$ , noté  $\hat{\boldsymbol{\Sigma}}_{CML}$  satisfait :

$$\hat{\boldsymbol{\Sigma}}_{CML} = n^{-1} \sum_{t=1}^n \hat{\mathbf{Z}}_{it} \hat{\mathbf{Z}}_{it}^T.$$

Soit  $\hat{\mathbf{C}}$ , la racine carré de  $\hat{\boldsymbol{\Sigma}}_{CML}^{-1}$ , satisfaisant  $\hat{\mathbf{C}}\hat{\mathbf{C}}^T = \hat{\boldsymbol{\Sigma}}_{CML}^{-1}$ . Les résidus normalisés sont définis par :

$$\hat{\boldsymbol{w}}_{it} = \hat{\mathbf{C}}(\hat{\mathbf{Z}}_{it} - \bar{\mathbf{Z}}_{it}) = (\hat{w}_{it1}, \dots, \hat{w}_{itm})^T, \quad (6.2.1)$$

avec  $\bar{\mathbf{Z}}_{it}$  qui est la moyenne des  $\hat{\mathbf{Z}}_{it}$ , pour  $t = 1, \dots, n$ . Il est important de rappeler que l'objectif est de déterminer si le terme d'erreur provient d'une loi normale. C'est pourquoi tous les tests sont basés sur les résidus normalisés, soit des résidus centrés et réduits.

**Remarque 6.2.1.** *Il est très important de centrer les résidus. Des résultats très différents peuvent être obtenus dans le cas de données non centrées.*

Les statistiques de Jarque-Bera et de Mardia du chapitre 3 font parties des simulations. Le test de Jarque-Bera sera noté *JB* et les coefficients d'aplatissement et d'asymétrie seront respectivement notés  $b_1$  et  $b_2$ . Les arguments des tests sont les résidus normalisés  $\hat{\boldsymbol{w}}_{it}$  tel que présenté en (6.2.1).

Les statistiques du test lisse du chapitre 4 avec les polynômes d'Hermite et Legendre ont été utilisées avec les résidus normalisés. Par exemple, pour  $R = 4$ , la notation des tests est  $L_{H4}$  pour Hermite et  $L_{L4}$  pour Legendre. Les statistiques utilisant le critère de sélection automatique sont aussi utilisées. Elles sont notées respectivement  $H_A$  pour Hermite et  $L_A$  pour Legendre. Notons que  $L_{AM}$  et  $L_{ASD}$  sont la moyenne et l'écart-type

du  $R$  choisi pour Legendre. Les équivalents pour Hermite sont  $H_{AM}$  et  $H_{ASD}$ . Dans la prochaine section, les résultats des simulations pour les niveaux empiriques sont présentés.

### 6.3. Niveaux empiriques

Le tableau des niveaux empiriques est présenté dans le tableau 6.3. Il est possible de distinguer les quatre modèles utilisés ainsi que leurs quatre tailles d'échantillons.

**Tableau 6.3.** Niveaux empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 %.

	$n$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
DGP <sub>1</sub>	50	4,8	5,0	5,3	5,7	10,3	3,10	0,39	4,7	6,5	6,5	5,5	8,5	3,12	0,48	5,1	4,7	2,1
	100	5,8	4,4	4,5	4,4	9,7	3,05	0,23	5,2	6,8	7,7	6,6	9,4	3,12	0,55	5,2	5,2	3,2
	200	3,8	2,8	2,9	3,3	6,4	3,03	0,19	4,6	6,1	8,3	7,5	8,4	3,13	0,58	4,5	4,6	4,4
	400	5,1	4,9	4,0	4,4	5,6	3,02	0,14	5,2	5,9	8,0	8,3	8,0	3,09	0,48	4,5	5,2	4,6
DGP <sub>2</sub>	50	6,0	3,8	3,5	3,2	6,9	3,010	0,11	3,8	6,2	6,3	5,1	5,7	3,07	0,41	4,8	3,8	2,6
	100	5,0	4,2	5,0	4,0	5,3	3,003	0,05	5,6	8,6	10,7	9,7	7,8	3,08	0,47	4,6	5,6	3,3
	200	5,4	4,4	4,0	4,6	5,5	3,001	0,03	3,8	9,5	10,8	11,3	5,9	3,07	0,43	4,6	3,8	2,9
	400	5,0	5,2	4,5	5,1	5,0	3,000	0,00	4,8	6,7	11,6	12,7	6,6	3,06	0,42	5,1	4,8	3,8
DGP <sub>3</sub>	50	4,1	4,0	4,8	4,4	11,2	3,11	0,40	4,2	6,0	5,3	4,6	8,5	3,13	0,52	4,3	4,2	2,3
	100	5,1	5,4	3,9	4,6	8,4	3,05	0,22	6,8	7,3	7,8	7,3	10,0	3,12	0,54	5,2	6,8	2,2
	200	4,5	5,1	5,0	5,2	6,8	3,03	0,17	4,3	6,4	7,1	7,6	8,1	3,09	0,45	4,4	4,3	4,4
	400	6,3	5,6	4,6	4,7	7,5	3,01	0,11	5,2	7,1	9,4	10,5	8,8	3,11	0,53	6,6	5,2	5,0
DGP <sub>4</sub>	50	4,8	4,4	4,3	3,8	10,8	3,11	0,40	5,2	6,9	5,9	5,7	9,3	3,14	0,55	5,2	5,2	2,8
	100	4,7	3,4	4,5	4,0	7,5	3,04	0,23	4,5	6,2	6,8	6,3	8,5	3,13	0,56	4,2	4,5	2,7
	200	4,9	5,5	5,4	5,8	7,0	3,03	0,17	5,1	6,7	9,0	8,6	9,3	3,12	0,55	5,3	5,1	2,9
	400	5,0	5,5	4,7	5,4	6,1	3,01	0,10	4,6	6,9	10,0	9,7	8,4	3,11	0,54	4,7	4,6	5,1

Les tests avec les polynômes de Legendre performent bien. Le niveau tourne autour de 5 % pour  $R = 3, 4, 5, 6$  et ce, peu importe le modèle. Remarquons aussi la convergence de la sélection automatique vers  $R = 3$  lorsque la taille de l'échantillon augmente. Ce résultat illustre le résultat du théorème 4.5.1. Notons que celle-ci se fait plus vite dans le modèle trivarié, soit le modèle DGP<sub>2</sub>.

Les tests avec les polynômes d’Hermite semblent rejeter plus souvent que le seuil de 5 %. En effet, le taux de rejet est plus élevé lorsque  $n$  augmente pour  $H_4$ ,  $H_5$  et  $H_6$ . La convergence de la sélection paraît plus lente dans cette famille. Comme les résultats de  $H_3$  sont identiques à  $b_1$ , l’équivalence algébrique entre  $H_3$  et  $b_1$  est montrée empiriquement.

Les performances des polynômes de Legendre sont supérieurs aux polynômes d’Hermite et  $b_2$ . Elles sont plutôt comparables à  $JB$  et  $b_1$ . L’utilisation des polynômes d’Hermite,  $JB$  et  $b_1$  semble équivalente. La prochaine section présente les résultats des simulations pour les puissances empiriques.

## 6.4. Puissances empiriques

Dans cette section, plusieurs distributions du terme d’erreur ont été simulées afin de comparer les puissances empiriques des tests. Deux méthodes ont été utilisées. La première est basée sur la valeur- $p$  et la deuxième est basée sur les quantiles du test sous  $H_0$ .

Les erreurs suivantes ont été utilisées dans leurs versions centrées. D’abord, la loi normale ( $N$ ), des lois Student indépendantes de degré  $d$  ( $t_d$ ), des lois uniformes standards ( $U$ ), des lois exponentielles de paramètres un ( $E$ ), des lois logistiques ( $L$ ), des lois log-normales ( $LN$ ), des lois du khi-deux avec  $d$  degrés de liberté ( $KH_d$ ) et des lois Weibull ( $WB$ ) ont été simulées. Quelques copules méta-gaussiennes ont été explorées. Une copule méta-gaussienne est une copule gaussienne qui génère une distribution non gaussienne. La première ( $MG_1$ ) possède des lois marginales formées de lois logistiques et de Student avec quatre degrés de liberté. La seconde ( $MG_2$ ) a toutes ses lois marginales provenant de lois logistiques alors que pour la troisième ( $MG_3$ ), elles proviennent toutes des lois de Student avec quatre degrés de liberté. De plus, une copule de type Clayton (*Clayton*) a été incluse. La loi normale contaminée ( $CN$ ) est définie de la façon suivante :

$$\mathbf{Z}_t \stackrel{\text{i.i.d.}}{\sim} p \cdot \mathcal{N}_m(\mathbf{0}, \Sigma) + (1 - p)\mathcal{N}_m(\mathbf{0}, \sigma^2 \Sigma),$$

où l’on a spécifié  $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$  avec  $\rho = 0,1$ . La matrice  $\Sigma$  peut s’écrire :

$$\Sigma = (1 - \rho)\mathbf{I}_m + \rho \mathbf{1}_m \mathbf{1}_m^T,$$

pour un certain  $m$ , avec  $\mathbf{1}_m = (1, \dots, 1)^T$ .

La génération de l'erreur  $\mathbf{Z}_t$  se fait de la façon suivante :  $\mathbf{Z}_t$  a une probabilité  $p$  d'être  $\mathcal{N}_m(\mathbf{0}, \mathbf{\Sigma})$  et une probabilité  $(1 - p)$  d'être  $\mathcal{N}_m(\mathbf{0}, \sigma^2 \mathbf{\Sigma})$ . Ainsi, cette loi est un mélange de lois normales. Rappelons que la loi normale multivariée implique que les lois marginales sont normales et que la réciproque est fautive. En utilisant cette distribution, il sera possible d'observer si les tests sont en mesure de détecter la normale lorsque la loi générée est un mélange de normale. Les choix  $p$  sont 0,9 et 0,95 et ceux de  $\sigma^2$  explorés sont 2, 5, 10, 15 et 20.

Les erreurs ARCH(1) sont définies de façon récursive. Soit  $\mathbf{Z}_0 = \mathbf{0}$  et

$$\mathbf{V}_0 = \begin{pmatrix} \frac{3}{7} & 0,1 \\ 0,1 & \frac{3}{9} \end{pmatrix}.$$

Si  $\mathbf{Z}_{t-1} = (Z_{t-1,1}, Z_{t-1,2})^T$ , calculons :

$$\mathbf{V}_t = \begin{pmatrix} 0,3 + \alpha_1 Z_{t-1,1}^2 & 0,1 \\ 0,1 & 0,3 + \alpha_2 Z_{t-1,2}^2 \end{pmatrix},$$

et posons  $\mathbf{Z}_t = \mathbf{V}_t^{\frac{1}{2}} \boldsymbol{\nu}_t$ , avec  $\boldsymbol{\nu}_t$  qui est la loi normale bivariée standard. Notons que  $\mathbf{V}_0 = \mathbb{E}(\mathbf{V}_t)$ . Une procédure similaire a été employée lorsque  $m = 3$  avec :

$$\mathbf{V}_0 = \begin{pmatrix} \frac{3}{7} & 0,1 & 0 \\ 0,1 & \frac{3}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

et

$$\mathbf{V}_t = \begin{pmatrix} 0,3 + \alpha_1 Z_{t-1,1}^2 & 0,1 & 0 \\ 0,1 & 0,3 + \alpha_2 Z_{t-1,2}^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les valeurs du paramètre  $\alpha_1$  choisies sont 0,8, 0,7 et 0,6 et celles de  $\alpha_2$  sont 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8 et 0,9.

Dans les tableaux 6.4 à 6.15, les tableaux des puissances sont présentés par modèle. Pour chaque modèle, trois tableaux ont été faits. Les tableaux 6.4, 6.7, 6.10 et 6.13 contiennent les erreurs normales, Student, uniformes, exponentielles, khi-deux, Weibull, logistiques, lognormales, les copules et Clayton. Les tableaux 6.5, 6.8, 6.11 et 6.14 examinent en détails les normales contaminées et tandis que les tableaux 6.6, 6.9, 6.12 et

6.15 analyse les erreurs ARCH. Pour chaque distribution, la puissance avec la valeur- $p$  et la puissance basée sur les quantiles sous  $H_0$  sont présentées. Cette dernière est présentée entre parenthèses. La taille de toutes les simulations des puissances est  $n = 50$ .

Dans les tableaux 6.4, 6.7, 6.10 et 6.13, les nouveaux tests sont puissants contre les lois exponentielles, Weibull et lognormales. Ils semblent moins puissants lorsqu'on augmente les degrés de liberté de la Student et du khi-deux. Les tests ne semblent pas bien performer contre la loi logistique et les copules, spécifiquement lorsque les marginales sont logistiques.

Dans les tableaux 6.5, 6.8, 6.11 et 6.14, la puissance augmente lorsque  $\sigma^2$  augmente. Plus  $\sigma^2$  augmente, plus la loi devient différente à une loi normale multivariée. Les tests sont puissants lorsque les marginales sont normales. Rappelons que  $p = 0,9$  signifie qu'en moyenne seulement cinq observations ont une plus grande variance.

Dans les tableaux 6.6, 6.9, 6.12 et 6.15, les tests sont plus puissants lorsque  $\alpha_1 = 8$  et lorsque  $\alpha_2$  augmente.

Globalement, les tests avec les polynômes d'Hermite semblent plus puissants qu'avec les polynômes de Legendre.

**Tableau 6.4.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>1</sub>.

BB	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
$N$	5,2 (5,7)	4,8 (4,8)	5,1 (4,7)	4,8 (4,7)	11,9 (4,9)	3,11	0,40	6,1 (6,3)	6,9 (4,0)	6,5 (4,6)	5,0 (4,5)	9,8 (4,2)	3,11	0,48	5,3 (5,0)	6,1 (6,3)	1,6 (4,5)
$t_5$	45,9 (46,6)	46,7 (46,8)	47,3 (46,3)	47,4 (46,6)	55,0 (43,2)	3,79	1,15	45,0 (45,1)	56,3 (51,8)	55,6 (52,3)	54,4 (53,5)	59,7 (51,7)	4,42	1,38	48,4 (48,1)	45,0 (45,1)	49,0 (55,5)
$t_8$	24,6 (25,1)	26,6 (26,6)	25,6 (24,9)	27,6 (27,0)	35,9 (23,3)	3,41	0,85	26,3 (26,9)	37,6 (30,6)	35,3 (31,5)	33,2 (32,2)	41,2 (30,7)	3,84	1,21	26,4 (26,0)	26,3 (26,9)	25,9 (33,0)
$U$	37,3 (39,2)	21,6 (21,7)	15,0 (13,8)	15,1 (14,6)	39,0 (10,5)	3,05	0,28	0,1 (0,1)	0,1 (0,0)	0,0 (0,0)	0,0 (0,0)	0,1 (0,0)	3,00	0,00	0,0 (0,0)	0,1 (0,1)	35,5 (72,0)
$E$	99,2 (99,4)	99,0 (99,0)	98,5 (98,3)	98,3 (98,2)	99,6 (97,2)	4,57	1,08	99,6 (99,6)	97,1 (93,8)	97,7 (96,1)	93,7 (92,6)	99,6 (86,6)	5,16	1,16	99,3 (99,2)	99,6 (99,6)	73,0 (78,1)
$KH_4$	87,3 (87,9)	87,1 (87,1)	81,1 (79,7)	78,9 (78,6)	89,9 (77,8)	3,90	1,02	89,4 (89,6)	80 (72,7)	77,6 (70,7)	70,3 (67,1)	89,7 (60,2)	4,43	1,35	88,0 (87,6)	89,4 (89,6)	46,6 (52,0)
$KH_{10}$	45,6 (46,4)	45,3 (45,3)	40,9 (39,2)	39,3 (38,8)	54,2 (36,8)	3,43	0,83	55,0 (55,7)	45,5 (38,6)	41,4 (35,8)	36,0 (34,8)	56,5 (31,8)	3,75	1,19	52,5 (52,1)	55,0 (55,7)	19,4 (25,3)
$WB$	99,2 (99,3)	99,4 (99,4)	98,7 (98,6)	98,4 (98,2)	99,4 (98,5)	4,66	1,04	99,3 (99,3)	97,1 (95,6)	97,5 (95,9)	95,2 (94,7)	99,3 (87,4)	5,19	1,14	99,2 (99,2)	99,3 (99,3)	74,9 (79,4)
$L$	22,1 (23,5)	21,7 (21,7)	20,5 (19,7)	20,2 (20,1)	30,6 (16,8)	3,24	0,63	15,9 (16,5)	28,9 (22,5)	26,0 (22,8)	24,1 (23,0)	33,4 (22,5)	3,60	1,04	28,3 (28,1)	15,9 (16,5)	18,1 (24,1)
$LN$	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	5,66	0,69	100 (100)	99,9 (99,9)	99,9 (99,9)	99,7 (99,7)	100 (99,2)	5,88	0,48	100 (100)	100 (100)	97,1 (98,2)
$MG_1$	42,6 (43,6)	41,5 (41,5)	41,8 (39,5)	42,5 (42,3)	50,8 (37,2)	3,63	1,07	37,1 (37,5)	51,5 (44,9)	49,5 (46,4)	47,4 (46,3)	55,3 (45,7)	4,20	1,30	50,2 (50,0)	37,1 (37,5)	39,9 (45,2)
$MG_2$	22,8 (23,4)	23,5 (23,5)	21,9 (21,0)	22,2 (21,9)	31,9 (19,6)	3,27	0,70	21,7 (22,2)	29,9 (24,6)	28,4 (24,9)	26,8 (25,1)	35,8 (24,2)	3,67	1,11	30,3 (29,7)	21,7 (22,2)	19,2 (25,3)
$MG_3$	60,2 (61,3)	60,8 (60,8)	59,8 (59,3)	60,1 (59,3)	66,0 (55,8)	4,02	1,28	52,7 (53,7)	67,7 (64,2)	66,6 (63,4)	63,9 (63,2)	70,9 (62,6)	4,71	1,34	68,6 (68,5)	52,7 (53,7)	58,2 (63,6)
$CT$	35,3 (37,0)	30,9 (31,0)	26,3 (24,9)	25,7 (24,8)	41,3 (21,8)	3,21	0,59	42,2 (43,3)	33,4 (25,1)	25,7 (21,5)	21,6 (20,0)	45,0 (17,5)	3,40	0,91	19,9 (19,7)	42,2 (43,3)	9,3 (13,7)

**Tableau 6.5.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>1</sub> avec des normales contaminées.

p	$\sigma^2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
0,9	2	8,1 (8,8)	9,4 (9,4)	8,7 (8,5)	9,1 (8,9)	15,4 (8,2)	3,16	0,50	8,4 (8,7)	11,4 (9,0)	10,7 (9,0)	9,8 (9,2)	14,7 (8,8)	3,25	0,73	8,5 (8,3)	8,4 (8,7)	6,1 (8,6)
	5	39,1 (39,7)	43,1 (43,3)	43,3 (42,9)	43,7 (43,0)	50,6 (40,3)	3,75	1,11	40,9 (41,2)	53,4 (49,0)	53,9 (51,6)	52,6 (51,8)	57,6 (50,4)	4,37	1,37	46,0 (45,7)	40,9 (41,2)	43,3 (47,5)
	10	73,7 (73,9)	76,8 (76,8)	76,3 (75,5)	77,7 (77,3)	81,5 (73,2)	4,63	1,32	74,4 (74,8)	84,9 (81,7)	84,5 (82,8)	83,7 (82,7)	86,5 (81,8)	5,34	1,13	77,1 (77,1)	74,4 (74,8)	79,1 (82,3)
	15	86,6 (86,9)	87,9 (87,9)	87,7 (87,1)	88,0 (87,8)	90,0 (86,7)	5,15	1,20	85,1 (85,1)	91,2 (90,0)	91,0 (89,9)	90,6 (89,8)	92,2 (89,7)	5,63	0,91	87,8 (87,6)	85,1 (85,1)	88,5 (89,9)
	20	92,3 (92,4)	93,1 (93,1)	92,8 (92,8)	93,4 (93,3)	94,0 (91,5)	5,42	1,05	90,7 (90,7)	95,8 (95,3)	95,7 (95,0)	95,4 (95,2)	96,3 (95,0)	5,81	0,67	93,0 (93,0)	90,7 (90,7)	93,8 (94,6)
0,95	2	4,7 (5,3)	4,9 (5,0)	5,4 (5,4)	6,7 (6,3)	11,1 (5,0)	3,13	0,44	5,8 (6,0)	9,2 (6,1)	8,9 (6,8)	7,5 (6,9)	12,4 (6,1)	3,18	0,62	6,8 (6,8)	5,8 (6,0)	3,2 (5,6)
	5	23,7 (24,0)	27,3 (27,3)	26,6 (25,5)	28,9 (28,5)	35 (25,9)	3,51	0,99	27,9 (28,7)	37,4 (32,7)	37,5 (34,5)	35,5 (34,6)	41,4 (33,9)	3,94	1,29	29,3 (29,1)	27,9 (28,7)	26,2 (30,8)
	10	50,0 (50,9)	53,3 (53,3)	53,0 (52,2)	54,1 (53,9)	58,5 (51,0)	4,17	1,34	54,2 (54,5)	62,7 (60,5)	62,4 (60,9)	61,4 (60,6)	65,2 (59,9)	4,71	1,41	53,8 (53,7)	54,2 (54,5)	54,2 (58,0)
	15	65,2 (65,8)	67,5 (67,5)	67,1 (66,8)	67,2 (67,0)	72,2 (66,4)	4,65	1,37	65,9 (66,1)	72,6 (70,8)	72,5 (70,7)	71,0 (70,8)	74,7 (70,8)	5,06	1,33	66,9 (66,7)	65,9 (66,1)	66,4 (69,6)
	20	75,2 (75,8)	76,5 (76,6)	76,4 (76,2)	77,0 (76,9)	78,9 (74,7)	5,01	1,32	76,2 (76,5)	80,4 (79,6)	80,6 (79,7)	80,2 (79,7)	81,6 (79,3)	5,34	1,20	77,1 (77,0)	76,2 (76,5)	76,7 (78,2)

**Tableau 6.6.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>1</sub> avec des ARCH.

$\alpha_1$	$\alpha_2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
8	0,1	29,6 (29,7)	28,4 (28,4)	29,4 (28,8)	28,7 (28,6)	36,9 (26,7)	3,45	0,93	21,2 (21,5)	32,7 (29,0)	30,9 (29,2)	30,5 (29,4)	37,5 (28,3)	3,77	1,17	33,5 (33,2)	21,2 (21,5)	22,7 (27,7)
	0,2	28,7 (29,2)	27,9 (27,9)	26,5 (25,7)	27,0 (26,3)	36,5 (24,0)	3,39	0,86	22,3 (23,1)	33,9 (29,0)	32,1 (28,1)	30,2 (29,1)	37,9 (27,6)	3,74	1,14	33,8 (33,7)	22,3 (23,1)	20,9 (25,7)
	0,3	32,2 (32,9)	32,1 (32,2)	31,1 (30,2)	30,1 (29,5)	40,5 (28,4)	3,42	0,89	26,6 (26,8)	38,9 (33,6)	36,1 (31,8)	33,1 (31,8)	43,5 (31,4)	3,83	1,19	37,4 (36,9)	26,6 (26,8)	24,6 (29,6)
	0,4	33,2 (33,9)	32,2 (32,2)	32,6 (31,7)	31,9 (31,4)	41,8 (29,7)	3,47	0,92	25,2 (25,3)	40,3 (34,5)	37,5 (34,1)	35,6 (34,1)	43,9 (33,0)	3,89	1,22	38,8 (38,7)	25,2 (25,3)	27,5 (32,4)
	0,5	37,2 (38,0)	37,9 (37,9)	36,4 (35,4)	35,8 (35,1)	45,0 (32,8)	3,48	0,94	30,4 (30,7)	44,0 (38,8)	42,6 (38,8)	40,2 (38,4)	48,2 (37,9)	3,95	1,22	44,2 (43,8)	30,4 (30,7)	31,3 (37,5)
	0,6	40,7 (41,8)	40,3 (40,4)	39,4 (39,0)	39,9 (39,4)	47,4 (35,8)	3,54	1,01	31,3 (31,7)	46,1 (41,0)	43,4 (40,0)	41,4 (40,0)	50,6 (40,3)	4,06	1,27	45,5 (45,2)	31,3 (31,7)	35,2 (41,5)
	0,7	43,9 (44,7)	44,1 (44,1)	43,5 (42,5)	43,0 (42,4)	52,2 (38,2)	3,56	1,00	33,4 (33,8)	49,2 (43,4)	47,4 (42,9)	45,4 (43,6)	55,1 (42,8)	4,14	1,28	49,9 (49,6)	33,4 (33,8)	38,7 (44,0)
	0,8	47,3 (48,1)	46,4 (46,4)	44,9 (43,3)	44,0 (43,0)	53,5 (40,7)	3,60	1,05	35,7 (36,3)	51,9 (45,5)	49,5 (44,8)	46,9 (45,0)	56,5 (43,7)	4,18	1,29	51,6 (51,2)	35,7 (36,3)	38,9 (45,4)
	0,9	53,5 (54,6)	53,0 (53,0)	51,6 (50,4)	50,0 (49,5)	58,6 (46,7)	3,70	1,13	41,0 (41,4)	56,9 (51,1)	54,5 (50,7)	52,7 (50,9)	62,5 (49,4)	4,28	1,28	58,1 (57,8)	41,0 (41,4)	45,1 (51,3)
7	0,1	22,5 (23,0)	22,1 (22,1)	20,8 (20,4)	21,6 (20,9)	29,4 (19,4)	3,27	0,69	16,8 (17,2)	26,3 (22,5)	25,0 (22,6)	23,8 (22,8)	30,8 (21,7)	3,58	1,05	26,3 (26,1)	16,8 (17,2)	15,4 (20,3)
	0,2	25,4 (25,9)	25,0 (25,0)	23,6 (22,7)	23,0 (22,6)	31,9 (21,6)	3,30	0,74	20,0 (20,2)	28,6 (24,5)	26,5 (24,4)	25,1 (23,4)	32,1 (24,1)	3,62	1,08	28,1 (28,0)	20,0 (20,2)	17,1 (21,2)
	0,3	25,5 (26,4)	27,1 (27,1)	25,9 (24,6)	26,8 (26,1)	33,4 (23,7)	3,33	0,76	22,2 (22,4)	31,7 (27,1)	29,1 (26,2)	27,6 (26,3)	35,5 (26,1)	3,67	1,10	31,8 (31,6)	22,2 (22,4)	19,7 (24,7)
	0,4	28,7 (29,4)	26,3 (26,3)	26,8 (25,8)	26,2 (25,8)	35,2 (22,2)	3,36	0,82	22,2 (22,3)	31,6 (26,8)	31,1 (27,6)	28,1 (27,2)	36,0 (27,2)	3,73	1,15	30,2 (29,8)	22,2 (22,3)	21,8 (27,1)
	0,5	28,0 (28,8)	28,2 (28,2)	27,8 (26,9)	26,6 (26,2)	36,1 (26,2)	3,38	0,83	22,2 (22,4)	33,3 (29,2)	32,4 (29,2)	30,4 (29,5)	38,7 (28,4)	3,74	1,14	33,1 (32,6)	22,2 (22,4)	23,6 (27,5)
	0,6	34,2 (35,9)	34,3 (34,3)	32,9 (32,0)	32,4 (31,9)	43,0 (30,4)	3,47	0,91	27,5 (27,7)	40,1 (35,6)	38,5 (34,0)	35,5 (34,5)	44,9 (34,1)	3,90	1,22	38,7 (38,3)	27,5 (27,7)	28,2 (33,8)
	0,7	38,3 (39,0)	39,0 (39,0)	38,2 (37,1)	37,8 (37,4)	46,0 (34,3)	3,49	0,97	30,2 (30,6)	45,6 (40,0)	43,7 (39,8)	41,2 (39,8)	51,0 (39,6)	4,04	1,26	45,7 (45,1)	30,2 (30,6)	33,0 (38,6)
	0,8	44,9 (45,5)	44,4 (44,4)	42,8 (41,8)	43,6 (42,9)	51,9 (40,2)	3,59	1,04	35,2 (35,6)	50,1 (44,3)	48,1 (44,8)	47,1 (45,1)	54,2 (44,2)	4,16	1,30	51,6 (51,0)	35,2 (35,6)	40,2 (44,6)
	0,9	46,8 (47,2)	46,0 (46,0)	45,0 (44,2)	43,7 (43,1)	54,0 (40,7)	3,64	1,09	36,7 (37,0)	51,8 (46,5)	50,7 (46,2)	48,5 (46,0)	57,0 (44,7)	4,18	1,28	53,2 (53,0)	36,7 (37,0)	41,0 (45,8)
6	0,1	16,2 (16,6)	16,5 (16,5)	15,7 (15,3)	16,0 (15,7)	23,3 (14,1)	3,23	0,66	13,7 (13,8)	21,2 (17,3)	20,4 (17,3)	19,3 (18,4)	24,9 (17,0)	3,45	0,93	20,6 (20,5)	13,7 (13,8)	11,6 (15,6)
	0,2	18,9 (19,4)	17,8 (17,8)	16,7 (16,0)	16,9 (16,3)	25,9 (15,2)	3,21	0,61	16,3 (16,3)	23,6 (19,2)	22,0 (18,5)	20,2 (19,3)	28,2 (18,7)	3,49	0,96	23,2 (22,7)	16,3 (16,3)	11,9 (16,0)
	0,3	22,4 (23,3)	22,5 (22,5)	21,0 (20,4)	21,3 (21,2)	31,1 (19,7)	3,30	0,72	16,9 (17,0)	27,5 (22,0)	25,6 (21,8)	23,5 (22,1)	32,2 (21,0)	3,59	1,05	25,7 (25,3)	16,9 (17,0)	15,9 (21,5)
	0,4	21,9 (22,7)	21,2 (21,4)	21,1 (20,5)	21,5 (21,1)	30,8 (19,3)	3,29	0,71	19,6 (20,0)	26,8 (22,4)	25,5 (22,2)	23,3 (22,7)	30,8 (22,2)	3,57	1,04	27,0 (26,8)	19,6 (20,0)	15,9 (20,8)
	0,5	26,0 (26,6)	26,7 (26,7)	24,9 (23,9)	24,6 (24,1)	34,7 (23,4)	3,34	0,78	20,1 (20,3)	31,3 (26,4)	28,9 (26,5)	26,8 (25,8)	36,3 (25,9)	3,70	1,11	31,9 (31,6)	20,1 (20,3)	20,8 (25,8)
	0,6	30,1 (31,0)	29,6 (29,6)	28,1 (27,4)	28,2 (28,1)	37,3 (24,6)	3,38	0,84	24,6 (25,4)	35,1 (29,8)	32,9 (29,2)	30,5 (29,3)	40,1 (28,6)	3,79	1,17	35,6 (35,3)	24,6 (25,4)	24,1 (27,9)
	0,7	32,8 (33,6)	32,1 (32,1)	32,7 (31,8)	31,3 (30,4)	40,2 (29,1)	3,42	0,87	25,1 (25,7)	37,5 (32,4)	35,5 (31,3)	33,0 (31,8)	43,0 (32,2)	3,82	1,17	36,7 (36,3)	25,1 (25,7)	25,0 (30,7)
	0,8	40,8 (41,6)	40,7 (40,7)	39,4 (38,5)	39,3 (38,6)	49,1 (36,3)	3,55	1,01	32,2 (32,7)	45,8 (41,2)	43,6 (40,4)	41,7 (40,0)	49,9 (39,5)	4,02	1,26	45,6 (45,3)	32,2 (32,7)	34,3 (39,6)
	0,9	45,1 (46,3)	45,6 (45,6)	44,5 (43,6)	44,3 (43,8)	52,9 (40,3)	3,62	1,07	34,1 (34,4)	50,6 (45,4)	48,2 (45,0)	47,3 (46,1)	54,6 (44,2)	4,16	1,29	51,0 (50,3)	34,1 (34,4)	39,8 (44,7)



**Tableau 6.7.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>2</sub>.

BB	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
$N$	5,0 (5,0)	4,9 (4,9)	4,7 (5,4)	3,9 (3,8)	6,1 (6,1)	3,01	0,10	3,7 (4,4)	7,5 (5,8)	7,1 (3,8)	5,6 (2,2)	5,9 (5,3)	3,07	0,40	5,6 (5,5)	3,7 (98,8)	2,1 (53,3)
$t_5$	57,1 (57,3)	61,0 (60,8)	61,4 (63,2)	64,7 (64,6)	60,7 (60,9)	3,53	0,95	64,1 (65,0)	77,3 (75,2)	76,4 (69,5)	74,4 (63,2)	75,1 (73,9)	4,88	1,35	59,2 (59,2)	64,1 (99,9)	66,3 (89,7)
$t_8$	31,4 (31,4)	36,6 (36,5)	34,4 (36,5)	36,8 (36,5)	35,9 (35,9)	3,21	0,60	37,2 (38,1)	52,1 (50,5)	51,5 (42,8)	48,1 (37,4)	48,7 (47,0)	4,10	1,37	34,9 (34,8)	37,2 (99,7)	36,6 (70,7)
$U$	45,5 (46,0)	24,6 (24,2)	11,4 (13,5)	8,5 (8,4)	45,8 (46,3)	3,00	0,05	0,0 (0,0)	0,1 (0,1)	0,1 (0,0)	0,0 (0,0)	0,1 (0,1)	3,00	0,03	0,0 (0,0)	0,0 (91,6)	60,6 (99,7)
$E$	100 (100)	100 (100)	99,3 (99,3)	98,4 (98,4)	100 (100)	4,05	0,85	100 (100)	98,6 (98,3)	97,7 (93,9)	94,0 (86,7)	100 (99,9)	5,36	1,11	100 (100)	100 (100)	81,1 (94,4)
$KH_4$	93,8 (93,8)	91,7 (91,7)	83,1 (84,4)	79,2 (79,2)	94,0 (94,0)	3,41	0,61	94,5 (94,7)	85,4 (83,5)	81,5 (71,3)	73,2 (59,2)	94,6 (93,4)	4,59	1,39	96,5 (96,4)	94,5 (100)	49,7 (80,8)
$KH_{10}$	54,8 (54,9)	51,0 (50,8)	38,4 (40,2)	35,9 (35,5)	56,0 (56,1)	3,11	0,36	60,6 (62,2)	51,5 (48,1)	46,2 (35,5)	40,3 (27,6)	61,0 (55,5)	3,78	1,25	68,7 (68,4)	60,6 (99,9)	19,9 (60,4)
$WB$	99,8 (99,8)	99,8 (99,8)	99,3 (99,3)	98,8 (98,8)	99,8 (99,8)	4,05	0,87	99,8 (99,8)	98,7 (98,4)	98,0 (94,9)	95,8 (87,0)	99,8 (99,8)	5,38	1,10	100 (100)	99,8 (100)	82,7 (95,1)
$L$	24,6 (24,8)	24,3 (24,3)	22,6 (24,3)	20,3 (20,0)	26,8 (27,0)	3,05	0,24	24,8 (26,6)	38,9 (35,8)	37,9 (28,9)	34,6 (21,1)	36,9 (34,6)	3,69	1,16	41,0 (41,0)	24,8 (99,8)	19,9 (62,6)
$LN$	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	5,45	0,81	100 (100)	100 (100)	100 (99,9)	99,8 (99,5)	100 (100)	5,95	0,32	100 (100)	100 (100)	99,3 (99,9)
$MG_1$	44,2 (44,5)	41,6 (41,3)	37,8 (39,0)	36,6 (36,6)	45,5 (45,7)	3,20	0,62	41,7 (43,0)	56,0 (53,4)	53,8 (46,8)	50,8 (39,2)	53,8 (52,0)	4,19	1,37	57,8 (57,7)	41,7 (100)	37,0 (71,8)
$MG_2$	25,2 (25,4)	24,1 (24,1)	21,2 (22,8)	20,9 (20,7)	26,0 (26,2)	3,05	0,28	24,2 (25,3)	37,4 (35,1)	36,8 (28,4)	33,9 (21,6)	33,9 (32,2)	3,67	1,16	38,4 (38,4)	24,2 (99,6)	19,7 (61,4)
$MG_3$	70,8 (71,1)	67,4 (67,2)	64,4 (66,0)	64,0 (63,8)	71,8 (72,0)	3,52	1,02	64,5 (65,0)	77,9 (76,7)	75,5 (69,8)	74,4 (65,0)	75,3 (74,4)	4,94	1,34	79,0 (79,0)	64,5 (100)	67,9 (87,9)
$CT$	66,4 (66,7)	57,5 (57,3)	45,8 (47,3)	42,7 (42,5)	67,2 (67,5)	3,08	0,32	69,5 (71,4)	53,5 (50,1)	42,8 (30,9)	36,2 (22,4)	70,1 (66,0)	3,60	1,11	50,1 (50,1)	69,5 (100)	16,0 (58,2)

**Tableau 6.8.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>2</sub> avec des normales contaminées.

p	$\sigma^2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
0,9	2	8,8 (9,0)	10,8 (10,4)	10,4 (11,1)	11,8 (11,5)	10,5 (10,7)	3,04	0,25	11,1 (11,7)	19,0 (17,3)	18,4 (14,6)	17,4 (9,5)	17,1 (16,1)	3,31	0,85	10,8 (10,7)	11,1 (99,1)	6,3 (52,5)
	5	46,8 (47,0)	54,3 (54,1)	52,4 (54,0)	56,3 (56,3)	52,7 (52,8)	3,38	0,76	58,3 (59,3)	70,6 (69,3)	71,1 (64,9)	69,2 (60,0)	68,1 (67,5)	4,79	1,40	49,2 (49,2)	58,3 (99,9)	54,8 (82,0)
	10	85,5 (85,6)	87,6 (87,6)	87,3 (87,7)	88,1 (88,1)	86,9 (87,0)	4,34	1,26	87,1 (87,6)	93,7 (93,1)	93,5 (92,1)	93,1 (89,8)	92,9 (92,6)	5,68	0,87	84,8 (84,7)	87,1 (100,0)	88,9 (95,6)
	15	94,1 (94,1)	95,6 (95,6)	95,2 (95,3)	95,5 (95,5)	95,2 (95,2)	5,05	1,18	94,9 (94,9)	97,4 (97,1)	97,3 (96,5)	97,1 (95,7)	97,3 (96,9)	5,87	0,57	94,0 (94,0)	94,9 (100,0)	95,5 (98,7)
	20	97,0 (97,0)	97,8 (97,7)	97,8 (97,8)	97,8 (97,8)	97,5 (97,5)	5,44	1,00	97,2 (97,2)	98,6 (98,6)	98,7 (98,5)	98,6 (98,4)	98,6 (98,5)	5,95	0,37	96,9 (96,9)	97,2 (100,0)	97,9 (99,0)
0,95	2	6,8 (6,8)	6,3 (6,3)	4,8 (5,2)	6,3 (6,1)	8,3 (8,3)	3,02	0,15	7,0 (8,0)	11,8 (10,9)	12,6 (8,0)	11,8 (6,6)	11,0 (9,7)	3,20	0,71	8,9 (8,9)	7,0 (99,0)	2,7 (48,7)
	5	30,6 (30,6)	35,7 (35,7)	35,1 (36,4)	38,2 (38,1)	35,4 (35,4)	3,25	0,65	37,9 (38,8)	51,6 (49,9)	52,1 (46,1)	49,8 (40,5)	48,9 (48,1)	4,23	1,41	34,4 (34,3)	37,9 (99,7)	34,6 (67,4)
	10	63,5 (63,6)	66,9 (66,8)	66,1 (67,0)	67,1 (67,0)	65,6 (65,7)	3,98	1,23	69,0 (69,6)	75,2 (74,6)	75,4 (73,0)	74,9 (71,1)	74,9 (74,3)	5,12	1,33	63,6 (63,6)	69,0 (100)	66,4 (85,0)
	15	75,8 (75,8)	79,1 (79,0)	78,3 (78,8)	79,5 (79,5)	78,0 (78,0)	4,58	1,35	79,3 (79,5)	83,9 (83,4)	84,4 (82,0)	83,7 (81,1)	83,2 (82,8)	5,42	1,16	76,4 (76,4)	79,3 (99,9)	76,5 (90,1)
	20	82,2 (82,2)	83,2 (83,2)	83,2 (83,8)	84,0 (84,0)	83,0 (83,0)	4,93	1,33	83,5 (83,9)	87,0 (86,4)	87,2 (85,8)	86,4 (85,1)	86,1 (85,8)	5,55	1,06	81,8 (81,8)	83,5 (99,9)	82,5 (92,4)

**Tableau 6.9.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>2</sub> avec des ARCH.

$\alpha_1$	$\alpha_2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
8	0,1	26,1 (26,1)	24,4 (24,4)	21,6 (22,3)	21,3 (21,1)	27,1 (27,1)	3,06	0,33	18,8 (19,5)	33,1 (30,0)	30,3 (24,3)	26,6 (19,1)	30,1 (28,9)	3,58	1,09	33,5 (33,4)	18,8 (99,4)	13,6 (56,8)
	0,2	25,9 (26,1)	22,8 (22,7)	20,2 (21,9)	19,4 (19,2)	27,2 (27,4)	3,06	0,34	19,0 (19,7)	29,7 (28,4)	28,7 (22,0)	26,3 (17,4)	28,6 (27,2)	3,54	1,05	30,7 (30,6)	19,0 (99,7)	13,9 (54,9)
	0,3	28,5 (28,8)	26,0 (25,7)	23,0 (24,5)	22,3 (22,0)	30,3 (30,6)	3,09	0,42	22,3 (23,3)	33,9 (31,9)	33,0 (27,3)	30,2 (21,7)	33,8 (31,7)	3,67	1,15	34,3 (34,2)	22,3 (99,8)	17,4 (56,2)
	0,4	30,0 (30,1)	27,2 (27,1)	24,8 (26,5)	24,0 (23,8)	31,4 (31,5)	3,07	0,36	23,6 (24,7)	38,3 (36,8)	36,0 (29,1)	32,4 (23,7)	36,3 (35,1)	3,73	1,18	38,1 (37,9)	23,6 (99,4)	19,1 (60,1)
	0,5	30,8 (31,0)	29,5 (29,5)	27,0 (28,7)	25,6 (25,5)	32,1 (32,3)	3,08	0,38	23,3 (24,5)	38,8 (36,5)	37,5 (30,0)	35,0 (25,1)	34,9 (34,2)	3,75	1,21	39,5 (39,5)	23,3 (99,7)	21,2 (63,0)
	0,6	35,1 (35,5)	34,2 (34,2)	32,1 (33,3)	30,7 (30,4)	36,7 (37,1)	3,07	0,36	27,6 (28,8)	44,6 (42,8)	41,7 (34,6)	38,5 (28,8)	41,5 (40,2)	3,86	1,25	41,7 (41,4)	27,6 (99,1)	25,9 (63,3)
	0,7	39,6 (39,8)	37,4 (37,3)	33,2 (34,5)	31,6 (31,4)	40,6 (40,8)	3,09	0,43	31,3 (32,0)	45,8 (44,0)	43,9 (37,3)	40,6 (30,2)	44,9 (43,5)	3,93	1,27	47,9 (47,9)	31,3 (99,9)	27,9 (62,2)
	0,8	42,2 (42,3)	38,6 (38,5)	36,3 (37,2)	35,2 (35,0)	43,2 (43,3)	3,12	0,48	31,8 (33,0)	48,0 (45,7)	46,0 (38,8)	43,2 (32,6)	45,2 (43,7)	3,97	1,29	51,2 (51,2)	31,8 (99,5)	29,0 (64,8)
	0,9	48,5 (48,6)	46,4 (46,2)	41,5 (43,0)	39,0 (38,8)	49,9 (50,0)	3,13	0,51	37,3 (39,1)	56,9 (54,7)	53,6 (46,8)	51,0 (39,9)	54,2 (53,2)	4,19	1,34	58,7 (58,7)	37,3 (99,6)	35,7 (70,1)
7	0,1	22,2 (22,3)	21,2 (21,2)	19,5 (20,5)	16,8 (16,7)	23,1 (23,2)	3,05	0,32	17,6 (18,3)	27,1 (25,6)	25,5 (20,9)	23,8 (16,2)	25,9 (24,1)	3,51	1,05	28,1 (28,1)	17,6 (99,3)	12,0 (55,4)
	0,2	22,2 (22,4)	21,1 (21,1)	18,6 (19,8)	17,5 (17,4)	23,5 (23,7)	3,04	0,26	19,0 (19,9)	27,5 (25,9)	26,6 (22,3)	24,9 (17,5)	27,2 (25,7)	3,53	1,06	27,3 (27,3)	19,0 (99,9)	13,5 (55,4)
	0,3	23,2 (23,2)	21,1 (20,7)	18,4 (19,5)	18,7 (18,6)	24,5 (24,5)	3,06	0,32	18,1 (18,4)	29,7 (28,3)	28,3 (22,5)	25,9 (18,2)	28,4 (27,3)	3,56	1,07	31,2 (31,2)	18,1 (98,6)	13,5 (56,4)
	0,4	23,5 (23,7)	22,5 (22,3)	18,7 (19,9)	18,5 (18,4)	24,8 (25,0)	3,05	0,26	19,1 (20,2)	30,6 (29,4)	29,7 (23,0)	27,3 (17,6)	29,8 (28,0)	3,56	1,07	31,3 (31,3)	19,1 (99,5)	14,5 (57,3)
	0,5	23,7 (23,8)	22,7 (22,5)	21,6 (21,9)	20,5 (20,3)	24,1 (24,2)	3,05	0,29	19,1 (19,3)	32,1 (29,8)	31,4 (25,0)	29,7 (19,4)	30,0 (28,8)	3,60	1,10	33,0 (33,0)	19,1 (99,3)	16,8 (58,1)
	0,6	29,4 (29,5)	29,3 (29,2)	25,1 (26,5)	23,7 (23,5)	30,4 (30,5)	3,10	0,41	23,9 (24,5)	37,6 (35,4)	36,2 (28,6)	33,0 (23,1)	34,6 (33,4)	3,71	1,17	36,7 (36,7)	23,9 (99,2)	19,4 (61,0)
	0,7	31,0 (31,0)	28,7 (28,7)	26,3 (28,0)	25,6 (25,3)	32,1 (32,1)	3,07	0,33	24,0 (24,8)	39,2 (36,8)	37,1 (30,4)	33,9 (24,5)	36,4 (34,9)	3,75	1,19	40,9 (40,8)	24,0 (99,9)	20,2 (60,5)
	0,8	36,4 (36,5)	33,9 (33,8)	30,7 (32,1)	28,0 (27,9)	38,0 (38,1)	3,09	0,43	25,4 (26,4)	42,4 (40,4)	40,9 (33,3)	37,3 (26,5)	39,7 (38,2)	3,82	1,22	46,0 (45,9)	25,4 (99,5)	24,8 (63,4)
	0,9	44,5 (44,6)	40,5 (40,4)	37,3 (39,1)	35,5 (35,2)	46,1 (46,2)	3,11	0,46	33,1 (34,2)	51,9 (50,1)	50,4 (42,3)	47,0 (34,6)	49,2 (47,7)	4,07	1,31	54,0 (54,0)	33,1 (99,4)	30,9 (69,0)
6	0,1	15,9 (15,9)	14,9 (14,9)	11,0 (12,0)	10,7 (10,7)	16,9 (16,9)	3,03	0,22	12,4 (13,1)	20,6 (18,5)	20,7 (14,4)	18,4 (11,3)	18,9 (17,8)	3,34	0,89	20,6 (20,6)	12,4 (99,4)	7,2 (50,8)
	0,2	15,6 (16,0)	14,8 (14,7)	13,7 (14,5)	12,5 (12,5)	17,3 (17,7)	3,04	0,22	13,3 (14,5)	20,5 (19,3)	19,6 (14,5)	17,2 (12,0)	19,1 (17,8)	3,36	0,91	20,2 (20,1)	13,3 (99,6)	8,2 (52,6)
	0,3	17,0 (17,2)	17,1 (17,0)	15,0 (15,9)	13,3 (13,1)	18,5 (18,7)	3,04	0,20	14,4 (15,0)	24,0 (22,5)	22,3 (17,2)	20,6 (13,5)	22,9 (21,6)	3,42	0,95	22,5 (22,4)	14,4 (99,7)	8,8 (51,6)
	0,4	18,9 (19,2)	17,9 (17,8)	15,4 (16,4)	15,4 (15,2)	20,3 (20,6)	3,03	0,22	14,7 (15,8)	27,6 (25,0)	27,5 (19,6)	23,9 (14,7)	24,1 (22,5)	3,46	1,01	27,4 (27,4)	14,7 (98,7)	10,1 (56,9)
	0,5	20,4 (20,4)	20,3 (20,2)	17,8 (18,8)	17,5 (17,4)	22,3 (22,3)	3,05	0,26	16,5 (18,1)	29,6 (27,5)	28,5 (21,8)	25,6 (17,0)	27,2 (26,2)	3,53	1,05	28,8 (28,7)	16,5 (99,7)	12,4 (54,6)
	0,6	25,1 (25,2)	23,5 (23,3)	22,1 (23,0)	21,4 (21,0)	27,2 (27,3)	3,06	0,34	19,7 (20,6)	32,9 (31,5)	31,9 (25,1)	29,0 (20,7)	31,5 (30,3)	3,63	1,12	33,5 (33,4)	19,7 (99,2)	16,3 (55,6)
	0,7	30,0 (30,3)	29,3 (29,2)	26,7 (27,9)	26,1 (25,9)	31,5 (31,7)	3,07	0,32	24,8 (26,6)	36,7 (34,5)	35,2 (29,9)	32,6 (23,8)	35,2 (33,1)	3,71	1,17	37,4 (37,2)	24,8 (99,8)	19,5 (59,4)
	0,8	33,1 (33,1)	31,0 (31,0)	28,1 (30,0)	27,3 (27,2)	34,8 (34,8)	3,09	0,40	24,4 (25,6)	41,2 (38,9)	39,2 (31,8)	34,9 (26,3)	38,4 (36,9)	3,80	1,22	41,3 (41,2)	24,4 (99,2)	21,7 (61,1)
	0,9	39,7 (39,8)	37,5 (37,4)	32,9 (34,5)	33,3 (33,1)	40,8 (40,9)	3,09	0,42	29,4 (30,7)	46,7 (44,6)	44,4 (36,9)	40,5 (29,8)	44,3 (43,5)	3,93	1,27	48,4 (48,3)	29,4 (99,7)	26,1 (64,5)

**Tableau 6.10.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>3</sub>.

BB	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
$N$	4,2 (6,3)	4,4 (4,6)	4,5 (5,0)	3,9 (4,6)	9,7 (4,0)	3,09	0,35	4,5 (5,8)	5,4 (4,4)	5,3 (5,0)	4,6 (4,9)	8,6 (4,7)	3,12	0,49	3,6 (4,3)	4,5 (5,8)	1,4 (4,7)
$t_5$	48,7 (51,9)	51,2 (52,9)	51,1 (51,6)	51,4 (52,8)	59,4 (47,4)	3,87	1,18	48,2 (49,7)	60,2 (57,1)	58,5 (58,1)	55,6 (57,1)	63,9 (54,2)	4,50	1,37	52,8 (54,0)	48,2 (49,7)	52,7 (56,4)
$t_8$	26,6 (28,7)	29,3 (30,3)	27,8 (28,4)	28,0 (28,6)	36,5 (25,1)	3,45	0,90	28,1 (29,2)	38,2 (35,6)	38,5 (37,5)	35,9 (36,9)	43,0 (34,3)	3,88	1,22	30,6 (32,0)	28,1 (29,2)	28,4 (32,5)
$U$	43,4 (50,9)	29,0 (31,5)	19,6 (20,9)	20,9 (22,9)	44,8 (11,3)	3,05	0,30	0,1 (0,1)	0,2 (0,1)	0,1 (0,1)	0,0 (0,0)	0,2 (0,0)	3,00	0,03	0,0 (0,1)	0,1 (0,1)	40,6 (78,6)
$E$	99,0 (99,5)	99,3 (99,4)	99,1 (99,1)	98,8 (98,9)	99,3 (97,8)	4,68	1,05	99,6 (99,6)	97,2 (96,2)	97,6 (97,4)	94,8 (95,3)	99,6 (88,7)	5,19	1,17	99,6 (99,6)	99,6 (99,6)	74,3 (76,6)
$KH_4$	89,6 (90,6)	90,6 (91,3)	85,0 (85,6)	82,6 (83,7)	92,7 (81,2)	3,98	1,05	93,1 (93,8)	83,5 (79,1)	82,4 (81,3)	73,4 (74,6)	93,1 (64,6)	4,48	1,35	90,5 (91,6)	93,1 (93,8)	48,9 (51,9)
$KH_{10}$	49,5 (52,8)	47,4 (49,4)	40,8 (41,7)	39,5 (40,6)	55,9 (35,8)	3,41	0,79	56,5 (58,7)	45,9 (41,4)	42,3 (41,3)	35,8 (36,3)	58,1 (30,8)	3,73	1,18	55,2 (56,5)	56,5 (58,7)	20,6 (23,1)
$WB$	97,2 (97,6)	96,3 (96,5)	94,7 (94,9)	93,9 (94,2)	97,8 (93,6)	4,49	1,14	97,6 (97,8)	93,1 (91,8)	93,2 (92,9)	90,1 (90,7)	97,6 (84,5)	5,06	1,20	96,7 (97,0)	97,6 (97,8)	72,9 (76,2)
$L$	22,9 (25,5)	23,3 (24,2)	21,6 (22,4)	21,6 (22,8)	31,8 (20,2)	3,31	0,74	20,4 (22,1)	29,9 (25,7)	28,3 (27,5)	26,2 (27,2)	35,7 (24,8)	3,65	1,09	28,5 (30,3)	20,4 (22,1)	19,1 (22,6)
$LN$	99,9 (99,9)	99,8 (99,9)	99,8 (99,8)	99,8 (99,8)	99,9 (99,8)	5,67	0,68	99,9 (99,9)	99,7 (99,7)	99,8 (99,8)	99,4 (99,5)	99,9 (98,8)	5,86	0,53	99,9 (99,9)	99,9 (99,9)	96,6 (97,1)
$MG_1$	45,6 (48,2)	44,2 (46,6)	44,2 (44,9)	43,2 (44,8)	52,9 (39,6)	3,67	1,10	36,9 (38,8)	52,2 (49,3)	51,0 (49,6)	48,5 (48,9)	57,5 (48,0)	4,28	1,33	52,4 (53,4)	36,9 (38,8)	43,1 (47,0)
$MG_2$	22,3 (26,4)	23,4 (24,3)	22,2 (22,9)	22,0 (23,2)	31,3 (19,1)	3,30	0,73	21,8 (23,0)	29,9 (28,0)	29,5 (28,9)	27,8 (28,4)	35,3 (26,3)	3,68	1,12	29,5 (30,9)	21,8 (23,0)	19,4 (23,3)
$MG_3$	61,2 (63,7)	60,3 (61,9)	58,5 (59,7)	59,4 (60,8)	66,8 (54,5)	3,96	1,26	51,3 (52,3)	66,9 (64,8)	65,6 (64,9)	64,0 (64,6)	70,5 (61,2)	4,67	1,34	68,2 (69,1)	51,3 (52,3)	57,5 (61,0)
$CT$	39,9 (45,3)	35,1 (37,4)	29,2 (30,0)	27,7 (29,0)	46,9 (23,8)	3,23	0,61	48,0 (51,4)	34,9 (29,9)	27,0 (25,8)	23,1 (23,7)	50,0 (18,9)	3,37	0,87	21,3 (23,0)	48,0 (51,4)	10,2 (13,4)

**Tableau 6.11.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>3</sub> avec des normales contaminées.

p	$\sigma^2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
0,9	2	7,1 (8,3)	8,2 (8,8)	8,8 (9,1)	8,7 (9,1)	15,6 (8,3)	3,17	0,53	8,1 (8,9)	12,2 (10,3)	10,5 (10,4)	9,6 (10,0)	14,8 (9,3)	3,25	0,73	7,5 (8,3)	8,1 (8,9)	4,8 (9,0)
	5	39,6 (41,9)	42,8 (43,9)	42,8 (43,2)	43,4 (45,1)	50,5 (40,0)	3,73	1,09	40,5 (42,3)	53,8 (50,9)	54,2 (52,5)	51,5 (51,8)	57,4 (50,3)	4,37	1,37	46,0 (47,2)	40,5 (42,3)	43,3 (46,8)
	10	75,2 (77,4)	79,0 (79,5)	79,8 (80,0)	81,0 (81,5)	82,0 (75,8)	4,69	1,32	74,9 (76,2)	85,4 (84,5)	85,4 (84,9)	84,5 (84,7)	87,9 (83,7)	5,39	1,10	77,9 (78,3)	74,9 (76,2)	80,8 (82,5)
	15	88,4 (88,9)	89,8 (90,1)	89,5 (89,5)	89,9 (90,5)	91,6 (88,8)	5,19	1,19	86,4 (87,2)	93,8 (92,8)	93,2 (93,1)	93,0 (93,1)	94,2 (92,5)	5,70	0,83	90,2 (90,5)	86,4 (87,2)	90,9 (91,3)
	20	93,0 (93,3)	94,0 (94,1)	94,6 (94,7)	94,7 (94,8)	95,1 (93,3)	5,50	0,99	92,0 (92,2)	95,7 (95,5)	95,9 (95,8)	95,5 (95,5)	96,1 (95,5)	5,84	0,64	93,7 (94,1)	92,0 (92,2)	94,5 (95,0)
0,95	2	6,5 (7,8)	7,0 (7,9)	7,2 (7,5)	6,7 (8,4)	13,3 (6,1)	3,13	0,46	6,9 (7,6)	8,8 (7,5)	8,0 (7,4)	7,2 (7,5)	11,7 (7,2)	3,20	0,67	6,3 (7,1)	6,9 (7,6)	4,1 (6,5)
	5	24,1 (26,9)	27,6 (28,6)	28,7 (29,5)	29,5 (30,4)	34,1 (25,3)	3,54	1,00	28,2 (29,2)	36,8 (35,4)	36,8 (36,2)	34,6 (35,5)	40,3 (34,1)	3,97	1,31	31,1 (31,7)	28,2 (29,2)	27,1 (31,0)
	10	54,4 (56,2)	58,0 (58,9)	57,3 (57,7)	57,3 (58,2)	63,1 (55,9)	4,29	1,35	56,8 (57,2)	64,6 (63,5)	64,6 (64,1)	63,2 (63,6)	66,4 (63,0)	4,81	1,41	57,9 (57,9)	56,8 (57,2)	57,2 (60,2)
	15	70,0 (72,0)	72,1 (72,5)	72,2 (72,3)	71,6 (72,3)	76,0 (70,3)	4,80	1,36	70,8 (71,5)	77,0 (75,8)	77,7 (77,4)	76,5 (76,6)	78,3 (75,7)	5,21	1,28	72,4 (72,6)	70,8 (71,5)	72,2 (73,8)
	20	76,8 (77,9)	77,8 (78,2)	78,0 (78,1)	78,9 (79,1)	79,6 (76,5)	5,05	1,31	76,4 (77,1)	81,1 (80,4)	81,3 (81,2)	80,8 (80,9)	82,0 (80,3)	5,38	1,18	78,3 (78,5)	76,4 (77,1)	78,3 (79,2)

**Tableau 6.12.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>3</sub> avec des ARCH.

$\alpha_1$	$\alpha_2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
8	0,1	31,1 (33,5)	30,8 (31,6)	29,6 (30,3)	29,7 (30,3)	39,1 (26,8)	3,42	0,90	23,8 (25,1)	33,9 (31,7)	31,4 (30,8)	30,1 (31,1)	37,7 (29,2)	3,75	1,16	33,5 (34,6)	23,8 (25,1)	22,0 (26,1)
	0,2	30,6 (33,5)	31,0 (31,7)	29,6 (30,7)	29,2 (30,6)	37,8 (25,7)	3,37	0,84	23,1 (23,8)	35,4 (33,1)	34,6 (34,1)	32,9 (33,5)	39,9 (32,4)	3,82	1,19	35,9 (36,4)	23,1 (23,8)	24,6 (27,6)
	0,3	32,5 (34,5)	31,2 (32,6)	30,8 (31,6)	30,8 (31,9)	40,9 (27,0)	3,42	0,90	23,6 (24,9)	36,9 (33,6)	35,2 (34,4)	33,7 (34,6)	40,9 (31,9)	3,83	1,20	35,8 (36,9)	23,6 (24,9)	25,6 (29,6)
	0,4	34,6 (36,6)	33,4 (35,2)	32,4 (32,9)	30,7 (32,3)	43,1 (27,9)	3,44	0,92	24,7 (26,1)	37,4 (35,1)	37,1 (36,2)	34,6 (35,5)	42,2 (34,9)	3,87	1,20	36,7 (38,4)	24,7 (26,1)	27,5 (31,1)
	0,5	34,4 (37,7)	35,5 (36,8)	34,9 (35,1)	34,7 (36,0)	43,8 (31,8)	3,52	0,99	28,6 (29,9)	42,8 (39,3)	41,2 (40,3)	38,6 (39,4)	46,6 (37,9)	3,96	1,24	41,2 (42,1)	28,6 (29,9)	29,8 (34,0)
	0,6	36,4 (38,8)	34,5 (36,1)	33,9 (34,6)	34,4 (35,1)	44,5 (30,4)	3,50	0,97	28,6 (29,8)	41,8 (39,4)	40,6 (39,8)	39,1 (39,5)	46,9 (37,8)	3,97	1,25	40,2 (41,1)	28,6 (29,8)	33,0 (36,5)
	0,7	43,8 (46,5)	43,3 (45,6)	42,4 (43,1)	41,1 (42,6)	51,5 (37,5)	3,62	1,08	35,9 (37,3)	50,6 (48,2)	48,2 (47,8)	45,5 (46,3)	54,4 (44,3)	4,14	1,30	48,1 (48,9)	35,9 (37,3)	38,5 (42,0)
	0,8	46,9 (49,3)	47,3 (48,8)	45,9 (46,9)	44,0 (45,7)	54,3 (40,9)	3,67	1,12	35,3 (37,3)	51,7 (49,3)	50,1 (49,5)	47,8 (48,8)	56,7 (46,8)	4,21	1,30	51,8 (53,8)	35,3 (37,3)	42,7 (46,5)
	0,9	49,2 (51,7)	49,4 (50,8)	48,2 (48,8)	46,8 (47,8)	56,6 (43,3)	3,69	1,12	36,9 (38,8)	54,6 (52,1)	52,8 (52,5)	51,6 (52,2)	59,0 (50,9)	4,30	1,31	54,0 (55,4)	36,9 (38,8)	44,5 (48,7)
7	0,1	23,2 (24,8)	22,8 (24,5)	21,6 (22,5)	21,8 (23,0)	31,2 (20,0)	3,30	0,75	18,0 (19,5)	27,2 (25,3)	27,0 (26,3)	26,1 (26,3)	31,4 (24,5)	3,63	1,09	26,8 (27,9)	18,0 (19,5)	17,8 (21,3)
	0,2	25,4 (27,6)	24,9 (26,0)	23,5 (24,0)	21,4 (22,5)	32,8 (19,9)	3,29	0,73	17,5 (19,1)	28,8 (26,2)	27,5 (26,8)	25,9 (26,8)	32,7 (24,6)	3,62	1,07	26,7 (27,8)	17,5 (19,1)	18,2 (21,7)
	0,3	23,4 (25,4)	23,6 (25,1)	22,7 (23,1)	21,8 (22,6)	32,4 (20,1)	3,29	0,71	18,4 (19,5)	28,8 (27,1)	27,4 (26,8)	25,5 (26,2)	33,8 (24,3)	3,60	1,05	27,6 (28,5)	18,4 (19,5)	17,7 (21,9)
	0,4	24,8 (27,5)	25,5 (26,9)	24,6 (24,9)	24,0 (24,9)	32,3 (20,9)	3,31	0,74	20,9 (22,1)	30,5 (27,5)	29,3 (28,6)	27,2 (27,6)	35,3 (26,5)	3,68	1,12	29,2 (30,1)	20,9 (22,1)	20,3 (23,7)
	0,5	33,0 (35,1)	31,9 (33,3)	31,7 (32,4)	30,9 (32,0)	41,3 (27,8)	3,40	0,87	26,3 (27,3)	38,0 (35,3)	35,4 (34,9)	33,7 (34,6)	42,5 (32,5)	3,84	1,20	38,6 (39,3)	26,3 (27,3)	26,6 (30,4)
	0,6	37,2 (38,7)	37,0 (38,8)	35,4 (36,2)	35,2 (36,4)	45,3 (31,6)	3,50	0,96	29,5 (30,6)	40,9 (38,5)	38,9 (38,5)	36,9 (37,6)	45,2 (36,0)	3,91	1,22	41,1 (41,9)	29,5 (30,6)	29,7 (33,8)
	0,7	42,2 (43,9)	41,6 (43,0)	40,1 (40,7)	39,6 (41,0)	48,8 (35,6)	3,54	1,00	33,3 (34,3)	47,7 (44,8)	46,1 (45,5)	44,2 (45,3)	51,8 (43,1)	4,07	1,26	46,9 (48,5)	33,3 (34,3)	36,8 (41,1)
	0,8	45,4 (48,1)	45,8 (47,8)	43,5 (44,1)	44,2 (44,9)	53,4 (39,7)	3,61	1,04	35,0 (36,5)	52,0 (48,7)	47,9 (47,3)	46,7 (47,2)	56,4 (44,8)	4,17	1,29	51,7 (52,5)	35,0 (36,5)	38,1 (42,1)
	0,9	49,8 (53,1)	50,5 (51,8)	49,9 (50,3)	48,7 (50,2)	56,2 (43,0)	3,67	1,11	38,2 (39,8)	56,0 (54,1)	54,5 (53,7)	52,8 (53,6)	60,9 (51,6)	4,34	1,31	56,2 (57,4)	38,2 (39,8)	47,0 (50,9)
6	0,1	17,3 (19,2)	16,5 (17,7)	16,8 (17,3)	16,8 (17,8)	25,1 (14,2)	3,24	0,65	12,6 (14,1)	21,1 (19,0)	20,8 (20,2)	19,3 (19,7)	24,6 (17,9)	3,48	0,98	21,0 (21,8)	12,6 (14,1)	12,0 (15,2)
	0,2	17,6 (19,9)	17,2 (18,5)	17,5 (17,8)	16,5 (17,8)	24,1 (14,4)	3,24	0,64	14,0 (14,8)	22,0 (19,8)	21,3 (20,7)	19,6 (20,1)	25,5 (17,9)	3,47	0,96	21,0 (22,0)	14,0 (14,8)	12,9 (15,4)
	0,3	19,4 (21,4)	20,2 (22,0)	19,2 (20,0)	18,5 (19,6)	29,9 (16,7)	3,28	0,69	17,8 (19,8)	25,5 (22,4)	23,7 (22,3)	21,4 (22,1)	29,9 (21,5)	3,55	1,02	24,4 (25,6)	17,8 (19,8)	14,4 (18,3)
	0,4	24,2 (26,9)	25,1 (26,8)	24,6 (25,0)	22,3 (23,6)	33,8 (20,5)	3,32	0,73	19,2 (21,0)	30,4 (27,7)	29,9 (29,1)	27,3 (28,4)	35,9 (26,1)	3,68	1,11	28,5 (28,8)	19,2 (21,0)	19,7 (23,8)
	0,5	24,6 (26,5)	24,2 (25,6)	23,0 (23,5)	23,4 (24,3)	31,9 (18,8)	3,31	0,75	20,6 (22,0)	30,7 (28,4)	28,8 (28,3)	27,1 (27,7)	34,8 (25,6)	3,66	1,10	29,5 (30,3)	20,6 (22,0)	18,6 (23,0)
	0,6	25,5 (28,2)	24,8 (26,6)	25,6 (26,0)	25,4 (26,9)	33,7 (20,8)	3,32	0,75	19,7 (21,0)	32,5 (29,5)	31,4 (30,4)	29,1 (30,1)	37,1 (28,7)	3,72	1,12	32,1 (33,2)	19,7 (21,0)	22,7 (26,9)
	0,7	34,4 (37,2)	36,4 (37,4)	34,7 (35,2)	33,8 (34,7)	42,9 (31,2)	3,44	0,89	27,2 (29,1)	41,0 (38,7)	39,8 (39,4)	38,3 (39,1)	45,9 (35,8)	3,90	1,20	40,8 (41,7)	27,2 (29,1)	30,2 (33,3)
	0,8	39,9 (42,4)	40,2 (41,7)	38,4 (38,8)	38,7 (40,2)	47,4 (34,4)	3,51	0,96	31,4 (33,1)	46,3 (43,3)	44,3 (43,6)	43,1 (43,3)	52,7 (40,7)	4,05	1,24	45,8 (46,7)	31,4 (33,1)	34,1 (38,9)
	0,9	45,3 (47,6)	45,7 (46,8)	45,7 (46,3)	44,8 (45,9)	54,0 (40,1)	3,64	1,09	35,2 (36,8)	50,7 (49,1)	48,9 (48,2)	47,1 (48,1)	56,0 (46,8)	4,20	1,31	50,3 (51,2)	35,2 (36,8)	39,7 (44,7)

**Tableau 6.13.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>4</sub>.

BB	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
$N$	3,9 (4,4)	4,8 (5,7)	5,0 (5,6)	5,0 (6,1)	11,4 (5,2)	3,11	0,39	3,8 (3,7)	5,7 (4,0)	5,0 (3,2)	4,5 (3,6)	7,8 (3,7)	3,10	0,46	4,6 (4,5)	3,8 (3,7)	1,7 (3,8)
$t_5$	44,8 (45,6)	47,1 (47,9)	47,0 (47,8)	48,2 (50,1)	54,1 (43,0)	3,87	1,20	45,0 (44,9)	57,9 (53,8)	57,3 (51,9)	54,9 (52,5)	61,5 (53,0)	4,43	1,36	49,2 (48,6)	45,0 (44,9)	49,4 (51,1)
$t_8$	27,0 (27,5)	29,4 (30,2)	30,3 (31,5)	30,9 (33,1)	37,8 (26,3)	3,46	0,93	29,8 (29,8)	38,6 (34,8)	38,4 (33,2)	36,2 (33,6)	43,7 (35,1)	3,93	1,26	32,2 (31,5)	29,8 (29,8)	29,5 (31,6)
$U$	36,1 (37,7)	21,3 (22,7)	15,2 (16,3)	18,5 (21,3)	39,5 (12,6)	3,06	0,30	0,0 (0,0)	0,1 (0,0)	0,0 (0,0)	0,0 (0,0)	0,2 (0,0)	3,00	0,04	0,0 (0,0)	0,0 (0,0)	32,0 (64,9)
$E$	99,3 (99,4)	99,5 (99,7)	98,7 (99,0)	98,2 (98,6)	99,6 (98,7)	4,61	1,07	99,7 (99,7)	97,5 (95,7)	97,9 (95,3)	94,9 (93,1)	99,7 (90,1)	5,19	1,13	99,6 (99,5)	99,7 (99,7)	73,7 (74,9)
$KH_4$	88,2 (88,6)	87,9 (88,5)	83,4 (85,0)	80,4 (83,5)	90,3 (79,8)	3,92	1,06	90,7 (90,7)	81,5 (76,0)	80,4 (71,6)	71,9 (68,0)	90,8 (63,2)	4,48	1,35	89,4 (89,2)	90,7 (90,7)	47,8 (49,5)
$KH_{10}$	45,9 (47,3)	43,7 (45,1)	38,5 (40,4)	38,0 (40,4)	52,9 (34,7)	3,35	0,74	53,7 (53,4)	44,4 (37,8)	39,9 (32,9)	35,0 (31,8)	55,0 (30,3)	3,70	1,15	51,9 (51,2)	53,7 (53,4)	18,3 (20,0)
$WB$	97,3 (97,4)	97,5 (97,6)	95,7 (95,7)	94,3 (95,4)	98,3 (94,4)	4,53	1,15	98,1 (98,1)	94,9 (93,1)	94,9 (90,8)	91,1 (88,8)	98,1 (86,2)	5,13	1,18	97,5 (97,3)	98,1 (98,1)	73,0 (73,7)
$L$	22,1 (22,9)	22,7 (23,4)	21,5 (22,5)	21,9 (23,5)	30,3 (19,6)	3,29	0,72	20,7 (20,5)	30,3 (25,8)	28,6 (24,4)	26,8 (24,8)	34,3 (25,6)	3,67	1,12	31,1 (30,1)	20,7 (20,5)	21,4 (22,4)
$LN$	100 (100)	100 (100)	99,9 (99,9)	99,9 (99,9)	100 (99,9)	5,64	0,73	99,9 (99,9)	99,8 (99,7)	99,9 (99,8)	99,7 (99,6)	99,9 (99,5)	5,86	0,51	100 (100)	99,9 (99,9)	96,5 (96,8)
$MG_1$	40,8 (41,4)	42,0 (43,0)	41,6 (43,4)	40,9 (43,5)	49,6 (38,5)	3,65	1,08	37,0 (36,9)	50,1 (45,6)	48,2 (43,4)	46,1 (43,9)	54,6 (45,2)	4,20	1,32	49,2 (48,7)	37,0 (36,9)	39,7 (41,0)
$MG_2$	24,8 (25,1)	23,8 (24,4)	22,7 (23,7)	22,7 (25,1)	32,3 (19,7)	3,30	0,74	22,8 (22,7)	31,5 (27,0)	28,9 (24,3)	26,8 (24,5)	37,5 (26,0)	3,65	1,08	31,9 (31,2)	22,8 (22,7)	21,4 (22,2)
$MG_3$	61,2 (62,5)	60,7 (61,8)	59,8 (60,9)	59,7 (61,6)	66,4 (55,9)	4,07	1,32	53,2 (53,1)	66,9 (63,2)	66,2 (61,4)	62,9 (61,2)	69,7 (63,1)	4,71	1,35	67,5 (66,9)	53,2 (53,1)	58,9 (60,3)
$CT$	35,1 (36,4)	29,6 (30,8)	26,0 (27,3)	23,5 (26,5)	41,0 (20,7)	3,21	0,60	42,7 (42,4)	30,5 (25,5)	24,6 (16,8)	19,6 (16,9)	45,3 (16,8)	3,36	0,87	18,2 (17,7)	42,7 (42,4)	9,5 (9,8)

**Tableau 6.14.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle  $DGP_4$  avec des normales contaminées.

p	$\sigma^2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
0,9	2	8,5 (9,1)	8,7 (9,2)	8,8 (9,2)	8,2 (10,1)	17,5 (8,8)	3,18	0,52	6,8 (6,8)	12,0 (9,6)	12,0 (9,2)	10,0 (9,2)	15,1 (10,0)	3,26	0,75	8,2 (7,7)	6,8 (6,8)	5,9 (7,5)
	5	36,4 (36,9)	40,8 (41,9)	40,6 (42,0)	40,7 (43,1)	49,1 (39,6)	3,71	1,08	43,1 (42,8)	54,9 (50,1)	53,7 (47,9)	51,7 (49,2)	58,5 (49,8)	4,33	1,37	43,2 (42,6)	43,1 (42,8)	41,4 (43,3)
	10	75,2 (75,4)	78,4 (79,0)	77,9 (78,6)	79,8 (81,5)	82,8 (76,7)	4,65	1,30	76,1 (75,9)	86,0 (84,7)	86,3 (84,3)	85,1 (84,3)	88,1 (85,4)	5,42	1,07	78,1 (77,6)	76,1 (75,9)	80,8 (81,1)
	15	88,4 (89,0)	90,0 (90,3)	90,0 (90,2)	89,5 (90,1)	91,5 (88,9)	5,22	1,16	85,7 (85,7)	92,4 (91,8)	92,3 (91,5)	92,0 (91,6)	93,0 (91,9)	5,68	0,85	89,6 (89,4)	85,7 (85,7)	90,3 (90,5)
	20	92,4 (92,5)	93,4 (93,4)	93,3 (93,5)	93,2 (93,6)	94,1 (92,6)	5,43	1,04	89,7 (89,5)	95,0 (94,3)	94,5 (94,2)	94,4 (94,3)	95,4 (94,3)	5,8	0,72	93,1 (93,1)	89,7 (89,5)	93,1 (93,2)
0,95	2	5,5 (5,6)	5,5 (6,1)	5,2 (5,8)	6,1 (7,4)	12,4 (5,2)	3,11	0,40	4,5 (4,5)	8,2 (5,5)	7,8 (5,8)	6,9 (6,1)	10,4 (5,9)	3,16	0,58	5,3 (5,0)	4,5 (4,5)	3,6 (5,4)
	5	22,6 (23,2)	27,5 (28,1)	25,7 (26,8)	27,0 (28,6)	34,7 (25,3)	3,51	0,97	26,0 (26,0)	36,0 (32,7)	36,1 (32,6)	34,6 (32,6)	39,6 (33,5)	3,91	1,28	27,7 (27,1)	26,0 (26,0)	25,3 (26,3)
	10	52,6 (53,1)	56,6 (57,2)	56,6 (57,2)	57,9 (59,3)	61,2 (55,3)	4,29	1,35	55,8 (55,7)	64,5 (61,9)	64,9 (61,6)	62,9 (61,6)	66,5 (62,3)	4,78	1,39	56,9 (56,4)	55,8 (55,7)	56,4 (57,3)
	15	66,3 (66,4)	69,0 (69,6)	68,4 (69,2)	69,4 (70,6)	73,8 (68,2)	4,74	1,37	67,7 (67,6)	74,1 (72,9)	74,0 (72,9)	73,7 (72,7)	76,1 (73,6)	5,14	1,30	68,2 (67,9)	67,7 (67,6)	68,6 (68,9)
	20	75,4 (75,8)	77,4 (77,7)	76,9 (77,2)	77,6 (78,6)	80,1 (76,9)	5,03	1,31	75,7 (75,7)	80,5 (79,4)	80,5 (79,3)	80,2 (79,5)	81,2 (79,8)	5,36	1,20	75,8 (75,7)	75,7 (75,7)	77,3 (77,5)



**Tableau 6.15.** Puissances empiriques (en pourcentage) des différents tests basés sur 1000 itérations au niveau nominal 5 % du modèle DGP<sub>4</sub> avec des ARCH.

$\alpha_1$	$\alpha_2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_A$	$L_{AM}$	$L_{ASD}$	$H_3$	$H_4$	$H_5$	$H_6$	$H_A$	$H_{AM}$	$H_{ASD}$	$JB$	$b_1$	$b_2$
8	0,1	26,6 (27,0)	26,2 (27,4)	26,3 (27,5)	27,5 (29,0)	34,5 (24,5)	3,37	0,81	19,9 (19,7)	30,2 (27,2)	29,0 (25,8)	27,5 (25,6)	34,2 (26,6)	3,68	1,13	30,7 (30,1)	19,9 (19,7)	20,2 (21,7)
	0,2	30,7 (31,4)	29,2 (30,5)	28,0 (28,9)	27,6 (30,3)	38,1 (26,1)	3,38	0,85	21,5 (21,3)	35,7 (30,2)	33,1 (28,7)	31,5 (29,6)	40,2 (30,3)	3,79	1,16	34,9 (34,0)	21,5 (21,3)	22,4 (24,5)
	0,3	31,1 (31,6)	30,0 (30,9)	28,9 (30,2)	28,6 (31,1)	38,8 (27,4)	3,41	0,88	22,5 (22,4)	34,0 (30,2)	32,0 (27,8)	30,3 (28,1)	39,7 (29,6)	3,77	1,16	35,9 (35,4)	22,5 (22,4)	23,4 (24,8)
	0,4	31,4 (31,8)	29,9 (30,7)	29,2 (29,8)	27,7 (29,5)	38,3 (26,8)	3,41	0,86	23,7 (23,7)	36,2 (32,2)	33,5 (29,7)	31,5 (30,0)	40,3 (30,9)	3,81	1,18	36,6 (36,2)	23,7 (23,7)	24,6 (26,1)
	0,5	35,0 (35,7)	35,1 (35,8)	35,1 (35,8)	34,2 (36,3)	42,9 (32,3)	3,48	0,93	27,3 (26,8)	40,2 (36,3)	38,6 (34,8)	35,9 (34,3)	44,4 (35,5)	3,90	1,22	39,6 (39,5)	27,3 (26,8)	29,0 (30,6)
	0,6	39,8 (40,3)	39,6 (39,9)	38,4 (39,6)	39,9 (41,9)	46,8 (35,3)	3,52	1,01	30,4 (30,3)	45,6 (42,0)	44,8 (40,0)	42,9 (40,2)	51,7 (42,5)	4,07	1,26	45,7 (45,0)	30,4 (30,3)	34,8 (36,2)
	0,7	44,2 (44,6)	44,5 (45,9)	42,4 (43,2)	42,2 (44,4)	52,0 (39,5)	3,54	0,99	35,0 (35,0)	50,3 (45,9)	48,9 (43,7)	47,4 (45,0)	56,3 (45,9)	4,16	1,27	49,6 (48,6)	35,0 (35,0)	39,4 (40,8)
	0,8	49,4 (49,8)	50,6 (51,6)	47,7 (48,8)	48,5 (50,7)	57,9 (44,9)	3,68	1,10	38,9 (38,7)	56,5 (52,9)	55,4 (50,2)	52,5 (50,2)	61,8 (52,6)	4,33	1,30	55,2 (54,5)	38,9 (38,7)	44,9 (46,0)
	0,9	56,1 (56,7)	55,2 (55,8)	53,9 (54,9)	54,0 (55,8)	62,3 (52,1)	3,80	1,18	43,7 (43,6)	61,1 (57,5)	58,9 (54,9)	57,4 (55,0)	64,4 (56,5)	4,42	1,31	60,7 (60,1)	43,7 (43,6)	50,7 (51,7)
8	0,1	25,6 (26,1)	24,5 (25,5)	22,6 (24,4)	23,6 (25,8)	31,5 (20,9)	3,29	0,77	18,2 (17,9)	28,8 (25,9)	27,1 (23,5)	26,1 (24,3)	31,9 (25,2)	3,63	1,10	28,7 (27,7)	18,2 (17,9)	19,1 (20,7)
	0,2	21,9 (22,8)	22,2 (23,4)	20,8 (21,6)	21,2 (22,1)	29,8 (18,4)	3,31	0,77	19,3 (19,1)	28,9 (25,1)	26,3 (21,6)	23,7 (21,6)	32,6 (24,4)	3,6	1,05	27,4 (26,8)	19,3 (19,1)	16,7 (18,2)
	0,3	22,5 (22,8)	22,1 (23,3)	21,1 (22,0)	21,3 (23,0)	30,1 (19,7)	3,26	0,69	18,1 (18,0)	26,7 (23,4)	26,1 (22,8)	24,0 (23,2)	31,5 (24,1)	3,61	1,08	27,4 (26,8)	18,1 (18,0)	17,5 (18,7)
	0,4	26,2 (26,8)	27,4 (28,1)	26,0 (27,1)	26,2 (28,1)	35,3 (24,9)	3,37	0,83	21,4 (21,2)	33,9 (29,1)	31,7 (27,4)	29,9 (27,3)	37,6 (29,1)	3,74	1,14	32,0 (31,4)	21,4 (21,2)	22,6 (23,7)
	0,5	26,7 (27,4)	26,6 (27,8)	26,3 (27,2)	26,3 (28,4)	34,9 (23,0)	3,35	0,80	22,2 (22,0)	34,1 (29,8)	30,9 (26,2)	28,7 (27,0)	38,4 (28,2)	3,73	1,15	32,8 (32,1)	22,2 (22,0)	22,7 (24,2)
	0,6	37,2 (37,7)	38,0 (39,2)	36,7 (37,8)	36,0 (38,6)	44,7 (33,1)	3,44	0,89	27,7 (27,7)	42,9 (39,9)	42,1 (35,3)	39,6 (36,6)	48,0 (38,5)	3,96	1,23	41,6 (40,7)	27,7 (27,7)	31,0 (32,6)
	0,7	40,6 (40,8)	41,6 (42,7)	39,7 (41,0)	39,4 (42,5)	49,4 (38,0)	3,55	1,00	31,4 (31,4)	44,9 (41,1)	43,8 (38,8)	41,4 (39,1)	50,1 (41,1)	4,04	1,26	45,2 (44,5)	31,4 (31,4)	33,2 (34,3)
	0,8	44,3 (44,7)	43,1 (44,0)	41,8 (43,6)	41,6 (43,9)	52,0 (38,9)	3,55	1,00	32,1 (31,8)	47,6 (43,6)	45,3 (40,6)	43,5 (41,8)	52,0 (42,5)	4,11	1,28	48,7 (47,9)	32,1 (31,8)	37,9 (38,7)
	0,9	50,0 (50,6)	50,1 (51,2)	49,4 (50,1)	49,1 (52,3)	57,8 (47,2)	3,7	1,12	37,7 (37,5)	54,6 (50,6)	53,6 (48,1)	51,1 (49,2)	60,5 (50,5)	4,28	1,29	53,1 (52,4)	37,7 (37,5)	43,9 (45,4)
8	0,1	18,6 (18,8)	18,2 (18,8)	17,6 (18,4)	18,0 (20,3)	26,1 (16,6)	3,25	0,69	14,5 (14,2)	21,9 (19,0)	21,4 (18,1)	20,3 (18,9)	25,9 (19,1)	3,48	0,97	22,1 (21,4)	14,5 (14,2)	13,0 (14,5)
	0,2	20,2 (20,9)	19,9 (21,2)	20,3 (21,2)	18,5 (21,2)	28,4 (19,0)	3,25	0,66	14,9 (14,9)	24,1 (21,0)	23,0 (20,5)	21,4 (20,0)	28,3 (21,3)	3,57	1,05	23,3 (22,8)	14,9 (14,9)	14,6 (16,2)
	0,3	18,9 (19,5)	19,3 (20,1)	18,9 (20,0)	19,7 (20,9)	26,7 (17,4)	3,27	0,69	17,4 (17,2)	25,6 (21,8)	24,3 (20,2)	21,9 (20,2)	29,5 (21,7)	3,54	1,01	24,4 (23,6)	17,4 (17,2)	13,9 (15,3)
	0,4	21,7 (22,2)	21,4 (22,7)	20,9 (22,2)	22,0 (24,3)	29,7 (20,0)	3,31	0,74	18,9 (18,8)	27,6 (23,5)	26,3 (23,1)	25,2 (23,6)	32,5 (24,7)	3,63	1,10	27,5 (27,0)	18,9 (18,8)	18,0 (19,4)
	0,5	23,9 (24,3)	24,4 (25,0)	22,6 (23,9)	23,6 (25,0)	33,4 (23,5)	3,32	0,75	18,7 (18,7)	30,3 (25,7)	30,0 (24,9)	27,2 (24,7)	34,8 (25,9)	3,66	1,10	29,4 (28,7)	18,7 (18,7)	19,8 (21,1)
	0,6	27,8 (28,6)	27,6 (28,8)	28,1 (29,3)	26,9 (28,7)	36,7 (24,6)	3,35	0,78	22,7 (22,6)	33,9 (29,3)	32,7 (28,4)	30,4 (28,7)	39,0 (29,9)	3,76	1,16	33,8 (33,3)	22,7 (22,6)	24,9 (26,1)
	0,7	34,9 (35,8)	35,4 (36,5)	34,1 (34,6)	34,6 (36,5)	42,6 (32,6)	3,48	0,94	26,9 (26,5)	40,7 (37,1)	39,3 (35,0)	37,5 (35,0)	45,0 (36,5)	3,93	1,23	41,6 (41,1)	26,9 (26,5)	30,3 (31,7)
	0,8	37,9 (38,3)	37,3 (38,5)	36,6 (38,2)	36,7 (38,8)	46,3 (34,3)	3,49	0,95	28,6 (28,4)	44,8 (41,1)	42,5 (37,6)	40,5 (37,8)	48,5 (39,4)	3,98	1,24	44,8 (44,3)	28,6 (28,4)	31,9 (33,1)
	0,9	45,9 (46,5)	45,2 (45,9)	42,8 (44,1)	42,8 (45,2)	53,7 (40,7)	3,60	1,06	34,2 (33,9)	49,9 (46,1)	47,5 (43,2)	45,4 (42,6)	54,9 (45,0)	4,15	1,29	50,6 (50,1)	34,2 (33,9)	38,9 (40,5)



## Conclusion

---

Dans ce mémoire, des nouveaux tests de normalité ont été développés sur les résidus de modèles de séries ARMA vectoriels, basés sur le principe de Neyman (1937).

Les tests construits reposent sur le principe de Neyman (1937) qui projette la densité sous  $H_0$  dans un espace orthogonal. Le choix de deux familles orthogonales a été présenté, soit les polynômes d’Hermite et de Legendre. L’utilisation des polynômes d’Hermite permet d’avoir un test invariant aux transformations linéaires affines. De plus, la preuve de l’invariance de la statistique de test utilisant les polynômes d’Hermite a été démontrée. Dans le cas de Legendre, une transformation s’est imposée puisque la famille est orthogonale dans la loi uniforme. Aussi, une procédure de sélection automatique basée sur Ledwina (1994) a été présentée.

Pour évaluer la performance des tests, des simulations de Monte Carlo ont été faites avec quatre modèles et quatre tailles d’échantillon. Pour le niveau empirique, les nouveaux tests de normalité performant mieux que le test traditionnel du coefficient d’asymétrie et ils sont comparables aux tests de Jarque-Bera et Mardia. Ils sont assez puissants contre des distributions avec des ailes différentes à la loi normale. Les nouveaux tests sont aussi performants dans le cas des normales contaminées. Cependant, ils ne sont pas aussi performants dans les modèles ARCH.

Pour conclure, ce test pourrait être essayé dans d’autres contextes afin de voir sa performance. Aussi, il pourrait être utilisé pour tester d’autres distributions que la loi normale. En effet, le test pourrait être construit avec une famille orthogonale à une autre distribution où l’utilisation des polynômes de Legendre permet de tester n’importe quelles distributions grâce à la fonction de répartition. Tel que décrit par Tsay (2013), dans les

séries financières, d'autres distributions peuvent être utilisées, soit lognormal, des normales contaminées, etc. Nous laissons ces considérations pour des travaux futurs.

## Annexe A

---

### Développement des matrices $A_{m,R}$ et $B_{m,R}$

Les matrices  $A_{m,R}$  et  $B_{m,R}$  sont composées des quantités  $a_u$  et  $b_u$  tels que définis par les équations (4.2.16) et (4.2.17). Nous rappelons que ces quantités sont exprimé de la façon suivante:

$$a_u = \int_{-\infty}^{\infty} H_u(x)x\phi(x)dx,$$
$$b_u = \int_{-\infty}^{\infty} H_u(x)x^2\phi(x)dx.$$

### Développement de la matrice $A_{m,R}$ de l'équation (4.2.12)

Commençons le développement de la matrice  $A_{m,R}$ .

$$\begin{aligned} A_{m,R} &= \mathbb{E}_0\{\mathcal{L}_R(\mathbf{C}\mathbf{Z}_t)(\mathbf{C}\mathbf{Z}_t)^T\}, \quad \mathbf{Z}_t \sim \mathcal{N}_m(\mathbf{0}, \Sigma_{\mathbf{Z}}), \\ &= \mathbb{E}_0\{\mathcal{L}_R(\mathbf{F})(\mathbf{F})^T\}, \quad \mathbf{F} = (F_1, \dots, F_m)^T \text{ et } \mathbf{F} = \mathbf{C}\mathbf{Z}_t \sim \mathcal{N}_m(\mathbf{0}, \mathbf{I}_m), \\ &= [\mathbb{E}_0\{\mathcal{L}_R(\mathbf{F})F_1\}, \dots, \mathbb{E}_0\{\mathcal{L}_R(\mathbf{F})F_m\}], \\ &= [\mathbb{E}_0\{(\mathbf{L}_1(\mathbf{F}), \dots, \mathbf{L}_R(\mathbf{F}))F_1\}, \dots, \mathbb{E}_0\{((\mathbf{L}_1(\mathbf{F}), \dots, \mathbf{L}_R(\mathbf{F}))F_m)\}], \\ &= [\mathbb{E}_0\{(L_{1:1}(\mathbf{F}), \dots, L_{d(m,k):k}(\mathbf{F}))F_1\}, \dots, \mathbb{E}_0\{(L_{1:1}(\mathbf{F}), \dots, L_{d(m,k):k}(\mathbf{F}))F_m\}]. \end{aligned}$$

Il est possible de réécrire la matrice  $A_{m,R}$  sous la forme d'une matrice de dimension  $D(m, R) \times m$ .

$$A_{m,r} = \begin{pmatrix} \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_1\} & \cdots & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_m\} \\ \vdots & & \vdots \\ \mathbb{E}_0\{\mathbf{L}_R(\mathbf{F})F_1\} & \cdots & \mathbb{E}_0\{\mathbf{L}_R(\mathbf{F})F_m\} \end{pmatrix},$$

$$\begin{aligned}
&= \begin{pmatrix} \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1\} & \cdots & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_m\} \\ \vdots & & \vdots \\ \mathbb{E}_0\{L_{R:d(m,R)}(\mathbf{F})F_1\} & \cdots & \mathbb{E}_0\{L_{R:d(m,R)}(\mathbf{F})F_m\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{\prod_{i=1}^m H_{1_i}(F_i)F_1\} & \cdots & \mathbb{E}_0\{\prod_{i=1}^m H_{1_i}(F_i)F_m\} \\ \vdots & & \vdots \\ \mathbb{E}_0\{\prod_{i=1}^m H_{R_i}(F_i)F_1\} & \cdots & \mathbb{E}_0\{\prod_{i=1}^m H_{R_i}(F_i)F_m\} \end{pmatrix}.
\end{aligned}$$

Pour mieux comprendre la matrice  $A_{m,R}$ , quelques exemples sont donnés. Posons  $m = 2$  et  $R = 1$ , il existe alors deux polynômes de degré 1, soit  $(1, 0)$  et  $(0, 1)$ .

$$\begin{aligned}
A_{2,1} &= \begin{pmatrix} \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_1\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_1\} & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_1\} & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_2\} \end{pmatrix}, \quad H_0(x) = 1, \\
&= \begin{pmatrix} \mathbb{E}_0\{H_1(F_1)F_1\} & \mathbb{E}_0\{H_1(F_1)F_2\} \\ \mathbb{E}_0\{H_1(F_2)F_1\} & \mathbb{E}_0\{H_1(F_2)F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \int_{-\infty}^{\infty} H_1(F_1)F_1\Phi(F_1)dF_1 & \int_{-\infty}^{\infty} \Phi(F_2)dF_2 & \int_{-\infty}^{\infty} F_2\Phi(F_2)dF_2 & \int_{-\infty}^{\infty} H_1(F_1)\Phi(F_1)dF_1 \\ \int_{-\infty}^{\infty} F_1\Phi(F_1)dF_1 & \int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2 & \int_{-\infty}^{\infty} H_1(F_2)F_2\Phi(F_2)dF_2 & \int_{-\infty}^{\infty} \Phi(F_1)dF_1 \end{pmatrix}, \\
&= \begin{pmatrix} \int_{-\infty}^{\infty} H_1(F_1)F_1\Phi(F_1)dF_1 & 0 \int_{-\infty}^{\infty} H_1(F_1)\Phi(F_1)dF_1 \\ 0 \int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2 & \int_{-\infty}^{\infty} H_1(F_2)F_2\Phi(F_2)dF_2 \end{pmatrix}, \\
&= \begin{pmatrix} a_1 & 0 \\ 0 & a_1 \end{pmatrix}.
\end{aligned}$$

Posons  $m = 2$  et  $R = 3$ , le développement de la matrice  $A_{2,3}$  se fait de la façon suivante :

$$A_{2,3} = \begin{pmatrix} \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_1\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_2\} \\ \mathbb{E}_0\{\mathbf{L}_2(\mathbf{F})F_1\} & \mathbb{E}_0\{\mathbf{L}_2(\mathbf{F})F_2\} \\ \mathbb{E}_0\{\mathbf{L}_3(\mathbf{F})F_1\} & \mathbb{E}_0\{\mathbf{L}_3(\mathbf{F})F_2\} \end{pmatrix},$$

$$\begin{aligned}
&= \begin{pmatrix} \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{2:1}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{2:1}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{2:2}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{2:2}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{2:3}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{2:3}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{3:1}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{3:1}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{3:2}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{3:2}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{3:3}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{3:3}(\mathbf{F})F_2\} \\ \mathbb{E}_0\{L_{3:4}(\mathbf{F})F_1\} & \mathbb{E}_0\{L_{3:4}(\mathbf{F})F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_1\} & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_1\} & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_2\} \\ \mathbb{E}_0\{H_2(F_1)H_0(F_2)F_1\} & \mathbb{E}_0\{H_2(F_1)H_0(F_2)F_2\} \\ \mathbb{E}_0\{H_1(F_1)H_1(F_2)F_1\} & \mathbb{E}_0\{H_1(F_1)H_1(F_2)F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_2(F_2)F_1\} & \mathbb{E}_0\{H_0(F_1)H_2(F_2)F_2\} \\ \mathbb{E}_0\{H_3(F_1)H_0(F_2)F_1\} & \mathbb{E}_0\{H_3(F_1)H_0(F_2)F_2\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_1\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_2\} \\ \mathbb{E}_0\{H_1(F_1)H_2(F_2)F_1\} & \mathbb{E}_0\{H_1(F_1)H_2(F_2)F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_3(F_2)F_1\} & \mathbb{E}_0\{H_0(F_1)H_3(F_2)F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \int_{-\infty}^{\infty} H_1(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} \Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} H_2(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} \Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} H_1(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_2(F_2)\Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} H_3(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} \Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} H_2(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} H_1(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_2(F_2)\Phi(F_2)dF_2 \\ \int_{-\infty}^{\infty} F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_3(F_2)\Phi(F_2)dF_2 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
& \left. \begin{aligned}
& \int_{-\infty}^{\infty} H_1(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} \Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_1(F_2)F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} H_2(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} H_1(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_1(F_2)F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} \Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_2(F_2)F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} H_3(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} H_1(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_2(F_2)F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} H_2(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_1(F_2)F_2\Phi(F_2)dF_2 \\
& \int_{-\infty}^{\infty} \Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_3(F_2)F_2\Phi(F_2)dF_2
\end{aligned} \right\} , \\
& = \begin{pmatrix} a_1 & 0 \\ 0 & a_1 \\ a_2 & 0 \\ 0 & 0 \\ 0 & a_2 \\ a_3 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & a_3 \end{pmatrix} .
\end{aligned}$$

La seule ligne qui n'a pas été justifiée est  $\int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2$ , elle peut se réécrire comme :

$$\int_{-\infty}^{\infty} H_1(F_2)\Phi(F_2)dF_2 = \int_{-\infty}^{\infty} H_0(F_2)H_1(F_2)\Phi(F_2)dF_2 = 0$$

car  $H_0(F_i) = 1$ . Alors l'équation est égale à 0 parce que les polynômes  $H_i(\cdot)$  sont des polynômes orthogonaux. Étudions un dernier exemple avec  $m = 3$ . Lorsque  $m = 3$ , et  $R = 3$ , il y a  $d(m, R) = d(3, 3) = \binom{5}{3} = 10$  polynômes de degré 3. Or, il y a trois cas qui sont intéressants à étudier, soit  $(3, 0, 0)$ ,  $(2, 1, 0)$  et  $(1, 1, 1)$ . Les autres cas sont des permutations de ceux-ci. Si les trois cas sont présentés dans une matrice, nous avons alors





$$\begin{aligned}
&= \begin{pmatrix} \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1F_1\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1F_2\} & \cdots & \mathbb{E}_0\{L_{1:1}(\mathbf{X})F_mF_m\} \\ \vdots & & & \vdots \\ \mathbb{E}_0\{L_{R:d(m,R)}(\mathbf{F})F_1F_1\} & \mathbb{E}_0\{L_{R:d(m,R)}(\mathbf{F})F_1F_2\} & \cdots & \mathbb{E}_0\{L_{R:d(m,R)}(\mathbf{F})F_mF_m\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{\prod_{i=1}^m H_{1_i}(F_i)F_1F_1\} & \mathbb{E}_0\{\prod_{i=1}^m H_{1_i}(F_i)F_1F_2\} & \cdots & \mathbb{E}_0\{\prod_{i=1}^m H_{1_i}(F_i)F_mF_m\} \\ \vdots & & & \vdots \\ \mathbb{E}_0\{\prod_{i=1}^m H_{R_i}(F_i)F_1F_1\} & \mathbb{E}_0\{\prod_{i=1}^m H_{R_i}(F_i)F_1F_2\} & \cdots & \mathbb{E}_0\{\prod_{i=1}^m H_{R_i}(F_i)F_mF_m\} \end{pmatrix}.
\end{aligned}$$

Pour mieux comprendre, quelques exemples suivront. Posons  $m = 2$ ,  $R = 1$ , il y a 2 polynômes de degrés 1, soit  $(1, 0)$  et  $(0, 1)$ .

$$\begin{aligned}
B_{2,1} &= \begin{pmatrix} \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_1F_1\} & \mathbb{E}_0\{L_1(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_2F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1F_1\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_2F_2\} \\ \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_1F_1\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_2F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_1F_1\} & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_1F_2\} & & \\ \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_1F_1\} & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_1F_2\} & & \\ & & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_2F_1\} & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_2F_2\} \\ & & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_2F_1\} & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_2F_2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{H_1(F_1)F_1^2\} & \mathbb{E}_0\{H_1(F_1)F_1F_2\} & \mathbb{E}_0\{H_1(F_1)F_2F_1\} & \mathbb{E}_0\{H_1(F_1)F_2^2\} \\ \mathbb{E}_0\{H_1(F_2)F_1^2\} & \mathbb{E}_0\{H_1(F_2)F_1F_2\} & \mathbb{E}_0\{H_1(F_2)F_2F_1\} & \mathbb{E}_0\{H_1(F_2)F_2^2\} \end{pmatrix}, \\
&= \begin{pmatrix} b_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_1 \end{pmatrix}.
\end{aligned}$$

car

$$\mathbb{E}_0\{H_1(F_1)F_1^2\} = \int_{-\infty}^{\infty} H_1(F_1)F_1^2\Phi(F_1)dF_1 = b_1,$$

$$\mathbb{E}_0\{H_1(F_1)F_1F_2\} = \mathbb{E}_0\{H_1(F_1)F_1\}\mathbb{E}_0\{F_2\} = 0, \quad F_i \sim N(0, 1),$$

et

$$\begin{aligned}
\mathbb{E}_0\{H_1(F_1)F_2^2\} &= \int_{-\infty}^{\infty} H_1(F_1)\Phi(F_1)dF_1 \int_{-\infty}^{\infty} \Phi(F_2)F_2^2dF_2, \\
&= 1 \int_{-\infty}^{\infty} H_1(F_1)H_0(F_1)\Phi(F_1)dF_1 \quad H_i \text{ sont orthogonaux à } \Phi(F),
\end{aligned}$$

$$= 0.$$

Si nous posons  $m = 2$  et  $R = 3$ , nous avons alors que le développement de la matrice  $B_{2,3}$  est le suivant :

$$\begin{aligned}
B_{2,3} &= \begin{pmatrix} \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_1\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{\mathbf{L}_1(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{\mathbf{L}_2(\mathbf{F})F_1^2\} & \mathbb{E}_0\{\mathbf{L}_2(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{\mathbf{L}_2(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{\mathbf{L}_2(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{\mathbf{L}_3(\mathbf{F})F_1^2\} & \mathbb{E}_0\{\mathbf{L}_3(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{\mathbf{L}_3(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{\mathbf{L}_3(\mathbf{F})F_2^2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{1:1}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{1:2}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{2:1}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{2:1}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{2:1}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{2:1}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{2:2}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{2:2}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{2:2}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{2:2}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{2:3}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{2:3}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{2:3}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{2:3}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{3:1}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{3:1}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{3:1}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{3:1}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{3:2}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{3:2}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{3:2}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{3:2}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{3:3}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{3:3}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{3:3}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{3:3}(\mathbf{F})F_2^2\} \\ \mathbb{E}_0\{L_{3:4}(\mathbf{F})F_1^2\} & \mathbb{E}_0\{L_{3:4}(\mathbf{F})F_1F_2\} & \mathbb{E}_0\{L_{3:4}(\mathbf{F})F_2F_1\} & \mathbb{E}_0\{L_{3:4}(\mathbf{F})F_2^2\} \end{pmatrix}, \\
&= \begin{pmatrix} \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_1^2\} & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_1^2\} & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_2(F_1)H_0(F_2)F_1^2\} & \mathbb{E}_0\{H_2(F_1)H_0(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_1(F_1)H_1(F_2)F_1^2\} & \mathbb{E}_0\{H_1(F_1)H_1(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_2(F_2)F_1^2\} & \mathbb{E}_0\{H_0(F_1)H_2(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_3(F_1)H_0(F_2)F_1^2\} & \mathbb{E}_0\{H_3(F_1)H_0(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_1^2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_1(F_1)H_2(F_2)F_1^2\} & \mathbb{E}_0\{H_1(F_1)H_2(F_2)F_1F_2\} \\ \mathbb{E}_0\{H_0(F_1)H_3(F_2)F_1^2\} & \mathbb{E}_0\{H_0(F_1)H_3(F_2)F_1F_2\} \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
& \left( \begin{array}{cc}
\mathbb{E}_0\{H_1(F_1)H_0(F_2)F_2F_1\} & \mathbb{E}_0\{H_1(F_1)H_0(F_2)F_2^2\} \\
\mathbb{E}_0\{H_0(F_1)H_1(F_2)F_2F_1\} & \mathbb{E}_0\{H_0(F_1)H_1(F_2)F_2^2\} \\
\mathbb{E}_0\{H_2(F_1)H_0(F_2)F_2F_1\} & \mathbb{E}_0\{H_2(F_1)H_0(F_2)F_2^2\} \\
\mathbb{E}_0\{H_1(F_1)H_1(F_2)F_2F_1\} & \mathbb{E}_0\{H_1(F_1)H_1(F_2)F_2^2\} \\
\mathbb{E}_0\{H_0(F_1)H_2(F_2)F_2F_1\} & \mathbb{E}_0\{H_0(F_1)H_2(F_2)F_2^2\} \\
\mathbb{E}_0\{H_3(F_1)H_0(F_2)F_2F_1\} & \mathbb{E}_0\{H_3(F_1)H_0(F_2)F_2^2\} \\
\mathbb{E}_0\{H_2(F_1)H_1(F_2)F_2F_1\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_2^2\} \\
\mathbb{E}_0\{H_1(F_1)H_2(F_2)F_2F_1\} & \mathbb{E}_0\{H_1(F_1)H_2(F_2)F_2^2\} \\
\mathbb{E}_0\{H_0(F_1)H_3(F_2)F_2F_1\} & \mathbb{E}_0\{H_0(F_1)H_3(F_2)F_2^2\}
\end{array} \right), \\
= & \left( \begin{array}{cccc}
b_1 & 0 & 0 & 0 \\
0 & 0 & 0 & b_1 \\
b_2 & 0 & 0 & 0 \\
0 & a_1a_1 & a_1a_1 & 0 \\
0 & 0 & 0 & b_2 \\
b_3 & 0 & 0 & 0 \\
0 & a_2a_1 & a_2a_1 & 0 \\
0 & a_1a_2 & a_1a_2 & 0 \\
0 & 0 & 0 & b_2
\end{array} \right),
\end{aligned}$$

car

$$\begin{aligned}
\mathbb{E}_0\{H_1(F_1)H_2(F_2)F_1F_2\} &= \int_{-\infty}^{\infty} H_1(F_1)F_1\Phi(F_1)dF_1 \int_{-\infty}^{\infty} H_2(F_2)\Phi(F_2)dF_2, \\
&= a_1a_2.
\end{aligned}$$

Remarquons que la deuxième et la troisième colonne sont identiques. Lorsque  $d = 3$ , et  $R = 3$ , il y a  $d(m, R) = d(3, 3) = \binom{5}{3} = 10$  polynômes de degré 3. Or, il y a trois cas intéressants, soit  $(3, 0, 0)$ ,  $(2, 1, 0)$  et  $(1, 1, 1)$ . Les autres possibilités sont des permutations de ceux-ci. En considérant les exemples précédents, pour qu'une ligne soit non nulle, il faut que la somme  $r_i + r_j = k$  pour  $x_i$  et  $x_j$ . Si  $i = j$ , le résultat est  $b_k$  et 0 ailleurs, sinon nous allons être en présence d'un produit de  $a_ja_i$ . Le résultat attendu à la ligne associée à  $H_1(x_1), H_1(x_2), H_1(x_3) = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)$  en raison des polynômes orthogonaux. Donc, le seul développement intéressant est celui de

$H_2(x_1), H_1(x_2), H_0(x_3)$  qui est présenté dans la matrice suivante :

$$\begin{aligned}
& \begin{pmatrix} \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_1F_1\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_2F_1\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_3F_1\} \end{pmatrix} \\
& \begin{pmatrix} \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_1F_2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_1F_3\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_2F_2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_2F_3\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_3F_2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)H_0(F_3)F_3F_3\} \end{pmatrix}, \\
& = \begin{pmatrix} \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_1^2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_1F_2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_1F_3\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_2F_1\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_2F_2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_2F_3\} \\ \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_3F_1\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_3F_2\} & \mathbb{E}_0\{H_2(F_1)H_1(F_2)F_3F_3\} \end{pmatrix}, \\
& = \begin{pmatrix} 0 & a_1a_2 & 0 \\ a_1a_2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
& = \begin{pmatrix} 0 & a_1a_2 & 0 & a_1a_2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned}$$

Dans la matrice  $B_{m,r}$ , il y a toujours deux colonnes pareilles, soit pour  $F_iF_j$  lorsque  $i \neq j$ .



## Annexe B

---

### Preuve du théorème 5.1.2

Supposons que  $\mathcal{T} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$  et  $\mathcal{R} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$  sont deux bases orthonormales. Ainsi, chaque vecteur  $\mathbf{w}_i$ ,  $i = 1, \dots, n$ , peut s'exprimer de manière unique comme suit :

$$\mathbf{w}_i = a_{i1}\mathbf{u}_1 + \dots + a_{in}\mathbf{u}_n, \quad i = 1, \dots, n,$$

ou encore :

$$\mathbf{w}^T = \sum_{k=1}^n a_{ik} \mathbf{u}_k^T.$$

En multipliant par  $\mathbf{u}_j$ , pour un certain  $j \in \{1, \dots, n\}$ , on trouve que :

$$\begin{aligned} \mathbf{w}_i^T \mathbf{u}_j &= \sum_{k=1}^n a_{ik} \mathbf{u}_k^T \mathbf{u}_j, \\ &= a_{ij}, \end{aligned}$$

puisque  $\mathbf{u}_i^T \mathbf{u}_j = 0$ ,  $i \neq j$  et  $\mathbf{u}_i^T \mathbf{u}_j = 1$ . On a donc :

$$\begin{aligned} \mathbf{w}_1 &= \langle \mathbf{w}_1, \mathbf{u}_1 \rangle \mathbf{u}_1 + \dots + \langle \mathbf{w}_1, \mathbf{u}_n \rangle \mathbf{u}_n, \\ \mathbf{w}_2 &= \langle \mathbf{w}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 + \dots + \langle \mathbf{w}_2, \mathbf{u}_n \rangle \mathbf{u}_n, \\ &\vdots \\ \mathbf{w}_n &= \langle \mathbf{w}_n, \mathbf{u}_1 \rangle \mathbf{u}_1 + \dots + \langle \mathbf{w}_n, \mathbf{u}_n \rangle \mathbf{u}_n, \end{aligned}$$

ou encore :

$$\mathbf{w}_i = \sum_{j=1}^n \langle \mathbf{u}_i, \mathbf{w}_j \rangle \mathbf{u}_j.$$

Ainsi, on peut écrire :

$$\mathbf{U} = \mathbf{W}\mathbf{P},$$

où  $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ ,  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$  et  $\mathbf{P}$  est la matrice  $n \times n$  :

$$\mathbf{P} = (\langle \mathbf{u}_i, \mathbf{w}_j \rangle)_{i,j=1,\dots,n},$$

ou encore :

$$\mathbf{P} = (\mathbf{P}_{ij}),$$

avec  $\mathbf{P}_{ij} = \langle \mathbf{u}_i, \mathbf{w}_j \rangle$ . On montre que la matrice  $\mathbf{P}$  est orthogonale en montrant directement que :

$$\mathbf{P}\mathbf{P}^T = \mathbf{I}_n,$$

En effet, posons  $\mathbf{Q} = (\mathbf{Q}_{ij}) = \mathbf{P}\mathbf{P}^T$ . On aura que :

$$\begin{aligned} \mathbf{Q}_{ij} &= \sum_{k=1}^n \langle \mathbf{u}_i, \mathbf{w}_k \rangle \langle \mathbf{w}_k, \mathbf{u}_j \rangle, \\ &= \sum_{k=1}^n (\mathbf{u}_i^T \mathbf{w}_k) (\mathbf{w}_k^T \mathbf{u}_j), \\ &= \mathbf{u}_i^T \left( \sum_{k=1}^n (\mathbf{w}_k \mathbf{w}_k^T) \right) \mathbf{u}_j, \end{aligned}$$

or  $\sum_{k=1}^n (\mathbf{w}_k \mathbf{w}_k^T) = \mathbf{I}_n$ . En effet, les  $n$  colonnes de  $\mathbf{W}$  sont linéairement indépendantes, car  $\mathbf{R}$  est une base, et  $\mathbf{W}$  est inversible. On note que :

$$\mathbf{W}^T \mathbf{W} = \mathbf{I}.$$

Or  $\mathbf{W}^T = \mathbf{W}^{-1}$  et donc :

$$\mathbf{W}\mathbf{W}^T = \mathbf{I},$$

ce qui revient à dire que :

$$\sum_{k=1}^n \mathbf{w}_k \mathbf{w}_k^T = \mathbf{I}_n.$$



Alternativement, quelque soit  $\mathbf{v}$ , on montre que  $(\sum_{k=1}^n \mathbf{w}_k \mathbf{w}_k^T) \mathbf{v} = \mathbf{v}$ . En effet,  $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{w}_i$ .

$$\begin{aligned} \sum_{k=1}^n \mathbf{w}_k \mathbf{w}_k^T \mathbf{v} &= \left( \sum_{k=1}^n \mathbf{w}_k \mathbf{w}_k^T \right) \left( \sum_{i=1}^n \alpha_i \mathbf{w}_i \right), \\ &= \sum_{k=1}^n \mathbf{w}_k \left( \sum_{i=1}^n \alpha_i \langle \mathbf{w}_k, \mathbf{w}_i \rangle \right), \\ &= \sum_{k=1}^n \mathbf{w}_k \alpha_k, \\ &= \mathbf{v}. \end{aligned}$$

Puisque ceci est vrai quelque soit  $\mathbf{v}$ , ceci entraîne que  $\sum_{k=1}^n \mathbf{w}_k \mathbf{w}_k^T = \mathbf{I}_n$ . Ainsi, il découle que :

$$Q_{ij} = \mathbf{u}_i^T \mathbf{u}_j,$$

et  $Q_{ij} = 0$ ,  $i \neq j$ , ou  $Q_{ij} = 1$ , si  $i = j$ . Ceci montre que  $\mathbf{P}$  est une matrice orthogonale et ceci complète la preuve du théorème 5.1.2.



# Bibliographie

---

- Ahn, S. K. (1988), ‘Distributions for residual autocovariance in multivariate autoregressive models with structured parameterization’, *Biometrika* **75**, 590–593.
- Box, G. E. P., Jenkins, G. M. et Reinsel, G. C. (2008), *Time Series Analysis Forecasting and Control*, 4<sup>ième</sup> Édition. John Wiley & Sons Inc., Hoboken, New Jersey.
- Brockwell, P. J. et Davis, R. A. (1991), *Time Series : Theory and Methods*, 2<sup>ième</sup> Édition. Springer-Verlag, New York.
- Cox, D., Little, J., et O’Shea, D. (2005), *Using Algebraic Geometry*, 2<sup>ième</sup> Édition. Springer-Verlag, New York.
- Ducharme, G. R. et Lafaye de Micheaux, P. (2004), ‘Goodness-of-fit tests of normality for the innovations in ARMA models’, *Journal of Time Series Analysis* **25**, 373–395.
- Dunkl, C. F. et Xu, Y. (2014), *Orthogonal Polynomials of Several Variables*, 2<sup>ième</sup> Édition. Encyclopedia of mathematics and its application, vol. 155, Cambridge University Press, Cambridge.
- Gnanadesikan, R. (1977), *Methods for Statistical Data Analysis of Multivariate Observations*, Wiley, New York.
- Harville, D. A. (1997), *Matrix Algebra from a Statistician’s Perspective*, Springer-Verlag, New York.
- Henze, N. (2002), ‘Invariant tests for multivariate normality : A critical review’, *Statistical Papers* **43**, 467–506.
- Holmquist, B. (1995), ‘The  $d$ -variate vector Hermite polynomial of order  $k$ ’, *Linear Algebra and Its Applications* **237/238**, 155–190.
- Jarque, C. M. et Bera, A. K. (1980), ‘Efficient tests for normality, homoscedasticity and serial independence of regression residuals’, *Economics Letters* **6**, 255–259.

- Jarque, C. M. et Bera, A. K. (1987), 'A test for normality of observations and regression residuals', *International Statistical Review* **55**, 255–259.
- Johnson, L. W., Riess, R. D. et Arnold, J. T. (2002), *Introduction to Linear Algebra*, 5<sup>ième</sup> Édition. Addison-Wesley, Boston.
- Kariya, T., Tsay, R. S., Terui, N. et Li, H. (1999), 'Tests for multinormality with applications to time series', *Communications in Statistics - Theory and Methods* **28**, 519–536.
- Kilian, L. et Demiroglu, U. (2000), 'Residuals-based tests for normality in autoregressions : Asymptotic theory and simulation evidence', *Journal of Business & Economic Statistics* **18**, 40–50.
- Koziol, J. A. (1986), 'Assessing multivariate normality : A compendium', *Communications in Statistics - Theory and Methods* **15**, 2763–2783.
- Koziol, J. A. (1987), 'An alternative formulation of Neyman's smooth goodness of fit tests under composite alternatives', *Metrika* **34**, 17–24.
- Koziol, J. A. (1989), 'A note on measures of multivariate kurtosis', *Biometrical Journal* **5**, 619–624.
- Ledwina, T. (1994), 'Data-driven version of Neyman's smooth test of fit', *Journal of the American Statistical Association* **89**, 1000–1005.
- Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, New York.
- Lütkepohl, H. et Schneider, W. (1989), 'Testing for nonnormality of autoregressive time series', *Computational Statistics Quarterly* **5**, 151–168.
- Mardia, K. V. (1970), 'Measures of multivariate skewness and kurtosis with applications', *Biometrika* **57**, 519–530.
- Mardia, K. V. (1974), 'Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies', *Sankhyā : The Indian Journal of Statistics* **36**, 115–128.
- Mardia, K.V. (1975), 'Assessment of multinormality and the robustness of Hotelling's T<sup>2</sup>-test', *Applied Statistics* **24**, 163–171.
- Meyer, C. (2000), *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphie.

- Neyman, J. (1937), ‘Smooth test for goodness of fit’, *Skandinavisk Aktuarietidskrift* **20**, 149–199.
- Pearson, K. (1900), ‘On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling’, *Philosophical Magazine* **50**, 157–175.
- Pierce, D. (1985), ‘Testing normality in autoregressive models’, *Biometrika* **72**, 293–297.
- Plackett, R. L. (1983), ‘Karl Pearson and the chi-squared test’, *International Statistical Review* **51**, 59–72.
- Rahman, S. (2017), ‘Wiener–Hermite polynomial expansion for multivariate Gaussian probability measures’, *Journal of Mathematical Analysis and Applications* **454**, 303–334.
- Rayner, J. C. W., Thas, O. et Best, D. J. (2009), *Smooth Tests of Goodness of Fit : Using R*, 2<sup>ième</sup> Édition. John Wiley & Sons, Singapour.
- Reinsel, G. C. (1993), *Elements of Multivariate Time Series Analysis*, Springer Series in Statistics, Springer-Verlag, New York.
- Reinsel, G. C. (1997), *Elements of Multivariate Time Series Analysis*, 2<sup>ième</sup> Édition. Springer Series in Statistics, Springer-Verlag, New York.
- Sansone, G. (1991), *Orthogonal Functions*, Dover Publications Inc., New York.
- Shumway, R. H. et Stoffer, D. S. (2017), *Time Series Analysis and Its Applications : With R Examples*, 4<sup>ième</sup> Édition. Springer, New York.
- Szegö, G. (1975), *Orthogonal Polynomials*, 4<sup>ième</sup> Édition. American Mathematical Society, New York.
- Tagne Tatsinkou, J. F. (2016), ‘Sur les tests lisses d’ajustement dans le contexte des séries chronologiques’, [thèse de doctorat], Université de Montréal, Montréal.
- Takemura, A. et Takeuchi, K. (1988), ‘Some results on univariate and multivariate Cornish–Fisher expansion : Algebraic properties and validity’, *Sankhyā : The Indian Journal of Statistics* **50**, 111–136.
- Terui, N. et Imano, Y. (2003), ‘Testing multivariate gaussianity by Hermitian polynomials transformation’, Technical Report 66, Graduate School of Economics and Management, Tohoku University, Sendai.

Tsay, R. S. (2013), *An Introduction to Analysis of Financial Data with R*, John Wiley & Sons Inc., Hoboken, New Jersey.

Wong, W. (1997), 'Frequency domain tests of multivariate gaussianity and linearity', *Journal of Time Series Analysis* **18**, 181–194.