

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES ET INFORMATIQUE
APPLIQUÉES

PAR
YAN GIRARD

SÉRIES CHRONOLOGIQUES À UNE ET PLUSIEURS VARIABLES :
SYNTHÈSE DES MÉTHODES CLASSIQUES ET MODÈLES À BASE DE
COPULES

MAI 2011

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

Table des matières

Liste des tableaux	vii
Liste des figures	viii
Chapitre 1. Introduction	1
Chapitre 2. Quelques modèles de séries chronologiques	5
2.1 Notions de base en séries chronologiques	5
2.1.1 Fonctions d'autocovariance et d'autocorrélation	5
2.1.2 Autocorrélation partielle	7
2.1.3 Stationnarité	8
2.1.4 Processus <i>bruits blancs</i>	9
2.1.5 Théorème de Wold	9
2.1.6 Opérateur délai	11
2.2 Processus aléatoires stationnaires	13
2.2.1 Modèle à moyenne mobile : processus MA	13
2.2.2 Modèle autorégressif : processus AR	15
2.2.3 Processus ARMA	20
2.2.4 Modèle ARMA à composantes saisonnières	22
2.3 Processus aléatoires non stationnaires	23
2.3.1 Généralités	23

2.3.2	Modèle à tendance stationnaire: processus TS	25
2.3.3	Modèle à différence stationnaire: processus DS	26
2.4	Processus stochastiques non linéaires	28
2.4.1	Processus ARCH	28
2.4.2	Modèles ARCH généralisés : processus GARCH	29
Chapitre 3. Méthode de Box & Jenkins		30
3.1	Validation des données	30
3.2	Identification du modèle	31
3.2.1	Estimation du paramètre d'intégration	31
3.2.2	Estimation des ordres p et q	32
3.2.3	Tests d'hétéroscédasticité	34
3.2.4	Tests de non stationnarité	35
3.2.5	Transformations non linéaires	36
3.3	Estimation des paramètres	37
3.3.1	Maximum de vraisemblance	37
3.3.2	Application au processus AR(1)	38
3.4	Adéquation et validation du modèle	40
3.4.1	Test sur la moyenne des résidus	40
3.4.2	Tests sur les autocorrélations	41
3.4.3	Test de normalité des résidus	43
3.5	Prévision	44
Chapitre 4. Théorie des copules		48
4.1	Introduction	48
4.2	Présentation de quelques copules	51
4.2.1	Copule d'Ali-Mikhail-Haq	51

4.2.2	Copule de Farlie–Gumbel–Morgenstern	52
4.2.3	Copule de Clayton	53
4.2.4	Copule de Gumbel–Hougaard	54
4.2.5	Copule de Frank	54
4.2.6	Copule Normale	55
4.3	Mesures de dépendance	57
4.4	Estimation des paramètres	59
Chapitre 5.	Un modèle multivarié à base de copules	61
5.1	Inconvénients des modèles classiques	62
5.2	Description d'un nouveau modèle	64
5.3	Estimation des paramètres	68
5.4	Procédure d'adéquation pour la copule	70
5.5	Prévision	76
5.6	Extension multivariée	77
5.7	Illustration sur des données simulées	80
Chapitre 6.	Études de simulation	82
6.1	Performance des estimateurs	82
6.1.1	Estimateurs MV du processus AR(1)	82
6.1.2	Estimation du paramètre de dépendance de la copule	84
6.2	Performance de la méthode d'adéquation	89
6.2.1	Plan de simulation	89
6.2.2	Résultats	90
6.2.3	Interprétation des résultats	94
Chapitre 7.	Analyse des indices boursiers <i>Standard & Poor 500</i> et <i>Nasdaq composite</i>	95

7.1	Présentation des données et analyses univariées	96
7.1.1	Série SP500: indice <i>Standard & Poor 500</i>	96
7.1.2	Série NASDAQ: indice <i>Nasdaq composite</i>	100
7.2	Analyse multivariée	103
7.2.1	Sélection d'une copule	103
7.2.2	Prévision	105
	Conclusion	109
	Bibliographie	111
	Annexe A. Programmes en Matlab	116
A.1	Simulation de nombre aléatoires	116
A.1.1	Paires indépendantes à partir d'une copule	116
A.1.2	Deux processus AR(1) liés par une copule	119
A.2	Estimation de paramètres	120
A.2.1	Processus AR(1)	120
A.2.2	Processus AR(1) liés par une copule	121
A.2.3	EQM du maximum de vraisemblance	122
A.2.4	EQM de l'inversion du tau de Kendall	123
A.3	Procédure d'adéquation	124
A.3.1	Calcul de la statistique W_T	124
A.3.2	Estimation des valeurs critiques	125
A.3.3	Puissance	126
A.4	Procédures utilitaires	128
A.4.1	Vraisemblance d'un processus AR(1)	128
A.4.2	Vraisemblance de processus AR(1) liés par une copule	129
A.4.3	Tau de Kendall	131

A.4.4 Calcul de rangs	132
A.4.5 Inversion du tau de Kendall	133

LISTE DES TABLEAUX

3.1	Quelques exemples de la transformation de Box–Cox	36
5.1	Résultats des tests d’adéquation pour les données simulées . . .	81
6.1	EQM des EMC des paramètres d’un processus AR(1)	83
6.2	EQM de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule AMH	85
6.3	EQM de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule FGM	86
6.4	EQM de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Clayton	86
6.5	EQM de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de GH	87
6.6	EQM de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Frank	87
6.7	EQM de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule Normale	88
6.8	Puissance de la procédure d’adéquation quand $T = 50$	91
6.9	Puissance de la procédure d’adéquation quand $T = 100$	92
6.10	Puissance de la procédure d’adéquation quand $T = 250$	93
7.1	Résultats des tests d’adéquation de copules pour les séries <i>Standard & poor 500</i> et <i>Nasdaq composite</i>	106

LISTE DES FIGURES

2.1	Réalisations des processus $X_t = \varepsilon_t + 2\varepsilon_{t-3}$ et $X_t = 4 + \varepsilon_t + 0.8\varepsilon_{t-1}$	14
2.2	Réalisations des processus $X_t = -0.5X_{t-1} + 0.2X_{t-3} + \varepsilon_t$ et $X_t = 0.8 + 0.7X_{t-1} + 0.25X_{t-2} + \varepsilon_t$	17
2.3	Modèle avec rupture abrupte de moyenne	24
2.4	Modèle à tendance linéaire déterministe	24
2.5	Marche aléatoire sans dérive	25
4.1	Densité de la copule de Ali-Mikhail-Haq	51
4.2	Densité de la copule de Farlie–Gumbel–Morgenstern	52
4.3	Densité de la copule de Clayton	53
4.4	Densité de la copule de Gumbel–Hougaard	54
4.5	Densité de la copule de Frank	55
4.6	Densité de la copule Normale	56
5.1	Deux séries AR(1) de dépendance Clayton	81
7.1	Valeurs journalières de fermeture de l'indice <i>Standard & poor</i> <i>500</i> pour la période du 1er juin 2000 au 1er juin 2007	96
7.2	Fonctions d'autocorrélation et d'autocorrélation partielle em- piriques pour l'indice <i>Standard & poor 500</i>	97
7.3	Analyse de la série <i>Standard & poor 500</i> sous l'hypothèse d'un modèle autorégressif d'ordre un	98
7.4	Résidus du modèle AR(1) pour la série SP500	99
7.5	Fonctions d'autocorrélation et d'autocorrélation partielle em- piriques pour les résidus de l'indice <i>Standard & poor 500</i>	99

7.6	Valeurs journalières de fermeture de l'indice <i>Nasdaq composite</i> pour la période du 1er juin 2000 au 1er juin 2007	100
7.7	Fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour l'indice <i>Nasdaq composite</i>	101
7.8	Analyse de la série <i>Nasdaq composite</i> sous l'hypothèse d'un modèle GARCH(1,1)	102
7.9	Fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour les résidus de l'indice <i>Nasdaq composite</i>	102
7.10	Valeurs journalières de fermeture des indices <i>Standard & poor 500</i> et <i>Nasdaq composite</i> pour la période du 1er juin 2000 au 1er juin 2007	103
7.11	Nuage de points des résidus de la série bivariée SP500–Nasdaq	104
7.12	Nuage de points des résidus normalisés de la série bivariée SP500–Nasdaq	105
7.13	Nuages des résidus et d'observations indépendantes tirées de la copule Normale	107
7.14	Espérances conditionnelles $E(\tilde{X}_{T+1,1} \tilde{X}_{T+1,2} = x)$ et $E(\tilde{X}_{T+1,2} \tilde{X}_{T+1,1} = x)$	108

CHAPITRE 1

INTRODUCTION

Depuis toujours, l'Homme a voulu prédire l'avenir, que ce soit pour prendre de meilleures décisions ou simplement pour satisfaire sa curiosité. Les premières tentatives étaient basées sur l'astrologie ou autres superstitions. Toutefois, ces méthodes de prédiction ne peuvent pas être considérées comme rigoureuses aux yeux de la science. Heureusement, la science a énormément évolué au cours des dernières décennies; les méthodes de prévision qui en découlent ne font pas exception.

Une méthode de prévision très populaire est basée sur l'étude rigoureuse de *séries chronologiques*. Cette approche permet de prédire, par exemple, de nombreux phénomènes naturels et financiers. Une série chronologique est constituée de valeurs observées à des intervalles de temps réguliers. Par exemple, les débits annuels sur un cours d'eau ou encore les valeurs mensuelles de titres boursiers sont des séries chronologiques.

À la base, l'étude formelle des séries chronologiques consiste à trouver un modèle mathématique qui explique le mieux possible les données observées. À partir de ce modèle, il est possible de faire de la prévision. Cependant, la

justesse des prévisions dépend fortement de la qualité du modèle choisi. Il est donc primordial de trouver des modèles qui reflètent le mieux possible la réalité afin de minimiser les erreurs de prévision.

Pour l'étude des séries chronologiques univariées, la théorie est extrêmement bien développée. Par exemple, la méthode de Box–Jenkins (Box & Jenkins, 1970), basée sur l'unification des modèles dits *autorégressifs* et à *moyenne mobile*, est très bien documentée. On peut se référer aux ouvrages de Brockwell & Davis (2002) et Tsay (2005) pour un aperçu complet. Des modèles plus complexes ont également été introduits. On peut mentionner la famille des modèles à composantes hétéroscédastiques, dans lesquels la volatilité peut varier en fonction du temps (voir Gouriéroux, 1997), et les modèles non-linéaires (voir Fan & Yao, 2003).

L'étude simultanée de plusieurs séries chronologiques peut s'avérer intéressante lorsqu'on suspecte de la dépendance entre les séries. Par exemple, soient les débits hebdomadaires de deux rivières situées dans une même région. À l'aide d'un modèle adéquat, on pourrait anticiper un débordement du lit d'une des rivières à la suite de l'observation d'un débit anormalement élevé sur l'autre rivière (suite à des pluies diluviennes, par exemple). Dans la même veine, considérons les fluctuations des titres boursiers de la compagnie pétrolière *Advantage Oil & Gas Ltd.* et de la compagnie *Plastic Omnium SA*, qui oeuvre dans le domaine du plastique. Comme le plastique est à base de pétrole, on s'attend à ce que les séries associées à ces titres soient dépendantes. Il serait intéressant d'étudier les effets d'une augmentation de la valeur de *Advantage Oil & Gas Ltd.* sur celle de *Plastic Omnium SA*.

Dans les deux exemples précédents, la justesse des analyses qui peuvent être faites est tributaire de la qualité du modèle multivarié utilisé. La littérature concernant les modèles de séries chronologiques multivariés est assez bien développée. Parmi les modèles les plus populaires, on retrouve les modèles vectoriels autorégressifs (VAR), les modèles vectoriels autorégressifs conditionnellement hétéroscédastiques (VARCH) et les modèles vectoriels autorégressifs conditionnellement hétéroscédastiques généralisés (VGARCH) (voir Juselius, 2006). Cependant, ces modèles imposent que la structure de dépendance qui lie les séries soit de type *gaussienne*, c'est-à-dire basée sur la loi Normale multidimensionnelle. Cette hypothèse est périlleuse car beaucoup de phénomènes, notamment en finance, ont un comportement qui est mal représenté par cette loi (voir Breymann et al., 2003). L'utilisation des modèles *classiques* peut ainsi conduire à d'importantes erreurs de prévision.

Dans ce mémoire, on propose de relâcher cette supposition restrictive sur la forme de la dépendance en utilisant les copules. Plus précisément, on construira une famille de modèles de séries chronologiques multivariés dont la structure de dépendance peut être choisie afin de refléter le mieux possible les observations. La force principale des modèles à base de copules est leur grande flexibilité. En effet, il est possible, grâce à la théorie des copules, de modéliser séparément les comportements marginaux et la dépendance. Dans le contexte des séries chronologiques multivariées, les copules vont permettre une modélisation des séries individuelles et, indépendamment, un choix d'une structure de dépendance appropriée. Les contributions principales de ce travail sont les suivantes :

- (i) construction d'un modèle à base de copules dans un contexte de séries chronologiques à plusieurs variables;
- (ii) mise en oeuvre et étude de l'efficacité de méthodes d'estimation des paramètres;
- (iii) élaboration et étude empirique d'une méthode d'adéquation pour le choix de la copule;
- (iv) description d'une méthode de prévision.

Au chapitre 2, plusieurs notions de base en séries chronologiques sont abordées et quelques modèles de séries chronologiques sont présentés. Au chapitre 3, la procédure d'identification d'un modèle de série chronologique de Box & Jenkins (1970) est décrite en détail, alors qu'au chapitre 4, un résumé des résultats les plus pertinents concernant la théorie des copules est offert. Une famille de modèles de séries chronologiques à base de copules est introduite au chapitre 5, et des méthodes d'estimation de paramètres et d'adéquation sont décrites. Au chapitre 6, la performance de ces techniques inférentielles est étudiée à l'aide de simulations, alors qu'au chapitre 7, les méthodes développées dans ce travail sont appliquées à de vraies séries boursières. Le travail se termine par une brève conclusion.

CHAPITRE 2

QUELQUES MODÈLES DE SÉRIES CHRONOLOGIQUES

2.1 Notions de base en séries chronologiques

Une série chronologique est un processus stochastique $(X_t)_{t \in \mathbb{Z}}$, où t représente le temps. Dans ce chapitre, plusieurs concepts importants liés à l'analyse de séries chronologiques seront abordés. Parmi ceux-ci, on retrouve les notions d'autocorrélation, de stationnarité et de *bruit blanc*. On énoncera également un résultat de Wold (1948) concernant les séries chronologiques stationnaires.

2.1.1 Fonctions d'autocovariance et d'autocorrélation

Dans un processus $(X_t)_{t \in \mathbb{Z}}$, la variable aléatoire au temps t dépend généralement, d'une certaine façon, des observations précédentes X_s , $s < t$. Une manière de comprendre les liens d'interdépendance entre les termes d'une telle série consiste à considérer les fonctions d'autocovariance et d'autocorrélation. Ces fonctions mesurent, respectivement, la covariance et la corrélation

entre les termes d'une série chronologique. De façon générale, la covariance entre des variables aléatoires V_1 et V_2 est définie par

$$\text{cov}(V_1, V_2) = E \{ (V_1 - \mu_1)(V_2 - \mu_2) \},$$

où $\mu_1 = E(V_1)$ et $\mu_2 = E(V_2)$. On peut montrer facilement que $\text{cov}(V_1, V_2) = E(V_1 V_2) - \mu_1 \mu_2$. Pour sa part, la corrélation est une version standardisée de la covariance. Spécifiquement,

$$\text{cor}(V_1, V_2) = \frac{\text{cov}(V_1, V_2)}{\sigma_1 \sigma_2},$$

où σ_1 et σ_2 sont les écart-types de V_1 et V_2 , respectivement. Quelques propriétés de la covariance et de la corrélation, dont les preuves s'obtiennent aisément, sont énoncées dans la suite.

- (i) Si V_1 et V_2 sont stochastiquement indépendantes, alors $\text{cov}(V_1, V_2) = \text{cor}(V_1, V_2) = 0$;
- (ii) Si le lien entre V_1 et V_2 est de la forme $V_2 = AV_1 + B$, où $A > 0$, alors $\text{cor}(V_1, V_2) = +1$; on parle alors de lien positif linéaire très fort entre V_1 et V_2 ;
- (iii) Si le lien entre V_1 et V_2 est de la forme $V_2 = AV_1 + B$, où $A < 0$, alors $\text{cor}(V_1, V_2) = -1$; on parle alors de lien négatif linéaire très fort entre V_1 et V_2 .

À noter que le contraire de la propriété (i) n'est pas nécessairement vrai. En effet, il se peut que la corrélation entre deux variables aléatoires soit nulle, sans que cela ne signifie qu'elles soient indépendantes.

Pour une série chronologique $(X_t)_{t \in \mathbb{Z}}$, la fonction d'autocovariance de délai ℓ , pour tout $\ell \in \mathbb{Z}$, est définie par

$$\gamma(\ell) = \text{cov}(X_t, X_{t-\ell}).$$

En particulier, on a $\gamma(0) = \text{var}(X_t)$. La fonction d'autocorrélation de délai ℓ est définie à partir de la fonction d'autocovariance, à savoir

$$\rho(\ell) = \text{cor}(X_t, X_{t-\ell}) = \frac{\gamma(\ell)}{\gamma(0)}.$$

À noter que les fonctions d'autocovariance et d'autocorrélation sont symétriques au sens où $\gamma(-\ell) = \gamma(\ell)$ et $\rho(-\ell) = \rho(\ell)$.

2.1.2 Autocorrélation partielle

L'autocorrélation partielle mesure la corrélation entre X_t et X_{t-h} sans toutefois prendre en considération l'influence des variables antérieures à X_{t-h} . Ainsi, on peut montrer que la fonction d'autocorrélation partielle d'un processus $(X_t)_{t \in \mathbb{Z}}$ est donnée par

$$p(\ell) = \frac{|P(\ell)^*|}{|P(\ell)|}, \quad \ell \in \mathbb{N},$$

où

$$P(\ell)^* = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(\ell-2) & \rho(1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(\ell-3) & \rho(2) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho(\ell-1) & \rho(\ell-2) & \rho(\ell-3) & \cdots & \rho(1) & \rho(\ell) \end{pmatrix}$$

et

$$P(\ell) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(\ell-2) & \rho(\ell-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(\ell-3) & \rho(\ell-2) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho(\ell-1) & \rho(\ell-2) & \rho(\ell-3) & \cdots & \rho(1) & 1 \end{pmatrix}.$$

Ici, $|A|$ est le déterminant d'une matrice carrée A . Ainsi, les trois premières autocorrélations partielles sont déterminées par les relations

$$\begin{aligned} p(1) &= \rho(1), \\ p(2) &= \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}, \\ p(3) &= \frac{\rho(1)^3 - \rho(1)\rho(2)(2 - \rho(2)) + \rho(3)(1 - \rho(1)^2)}{1 - \rho(2)^2 - 2\rho(1)^2(1 - \rho(2))}. \end{aligned}$$

2.1.3 Stationnarité

La stationnarité est une caractéristique d'une série chronologique qui implique que le comportement de la série ne dépend pas du temps. En particulier, on dit qu'une série $(X_t)_{t \in \mathbb{Z}}$ est stable si elle ne comporte pas de tendance saisonnière, ni de tendance à la hausse ou à la baisse. Plus formellement, on distingue deux types de stationnarité, à savoir forte et faible.

Définition 2.1. Une série $(X_t)_{t \in \mathbb{Z}}$ est dite

- (i) *fortement stationnaire (ou stationnaire d'ordre 1) si pour tout $t \in \mathbb{Z}$, la loi de (X_1, X_2, \dots, X_t) est la même que la loi de $(X_{1+\ell}, X_{2+\ell}, \dots, X_{t+\ell})$ pour tout $\ell \in \mathbb{Z}$;*

- (ii) *faiblement stationnaire (ou stationnaire d'ordre 2) si $E(X_t)$ et $\text{cov}(X_t, X_{t-\ell})$, pour tout $\ell \in \mathbb{Z}$, sont indépendantes de t .*

La notion de stationnarité d'ordre 1 est généralement beaucoup trop forte pour les applications. Dans la pratique, il est habituellement suffisant de considérer des modèles faiblement stationnaires.

2.1.4 Processus bruits blancs

Le *bruit blanc* fait partie de la classe des processus stationnaires. Spécifiquement, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un *bruit blanc* si $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2 < \infty$ et

$$\rho(\ell) = \text{cor}(\varepsilon_t, \varepsilon_{t+\ell}) = 0, \quad \text{pour tout } \ell \geq 1.$$

Par conséquent, le comportement d'un bruit blanc au temps t n'a aucune incidence sur celui-ci au temps $t + \ell$. On parle de *bruit blanc gaussien* lorsque $\varepsilon_1, \varepsilon_2, \dots$ sont i.i.d. $\mathcal{N}(0, \sigma_\varepsilon^2)$.

2.1.5 Théorème de Wold

Le théorème de Wold (1948) est fondamental pour l'analyse de séries temporelles stationnaires. Ce résultat se veut, d'une certaine façon, un préalable aux processus ARMA qui seront abordés au chapitre 3.

Théorème 2.1. *Tout processus $(X_t)_{t \in \mathbb{Z}}$ faiblement stationnaire peut s'écrire sous la forme*

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \kappa_t, \tag{2.1}$$

où les paramètres ψ_1, ψ_2, \dots sont des réels tels que $\psi_0 = 1$ et $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien et κ_t est une composante linéaire telle que $\text{cov}(\kappa_t, \varepsilon_{t-j}) = 0$ pour tout $j \in \mathbb{Z}$.

Selon le théorème de Wold (1948), tout processus stationnaire d'ordre 2 peut être représenté comme une somme pondérée infinie de chocs passés, caractérisés par un *bruit blanc*.

Dans ce qui suit, deux exemples de séries chronologiques stationnaires sont présentés. On obtiendra, dans chaque cas, leur représentation selon la formule (2.1) du théorème de Wold (1948).

Exemple 2.1. *Soit le processus*

$$X_t = \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \nu_{t-j+1},$$

où $(\nu_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance unitaire. Afin que la condition $\psi_0 = 1$ du théorème de Wold (1948) soit satisfaite, il s'agit de poser

$$\varepsilon_t = \frac{1}{2} \nu_t.$$

On a alors

$$X_t = \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^{j-1} \varepsilon_{t-j+1} = \varepsilon_t + \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-j},$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc de variance $\text{var}(\varepsilon_t) = \text{var}(\nu_t/2) = 1/4$. Ainsi, $\psi_0 = 1$ et $\psi_j = (1/2)^{j-1}$. On note que la condition $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ est également satisfaite car

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^{2j} = \sum_{j=0}^{\infty} \left(\frac{1}{4}\right)^j = \frac{4}{3},$$

où le résultat

$$\sum_{j=0}^{\infty} a^j = \frac{1}{1-a}, \quad |a| < 1,$$

a été utilisé.

Exemple 2.2. Soit le processus

$$X_t = \rho X_{t-1} + \varepsilon_t,$$

où $|\rho| < 1$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus bruit blanc de variance $\sigma_\varepsilon^2 = 1$. Comme $X_{t-1} = \rho X_{t-2} + \varepsilon_{t-1}$, on peut écrire

$$X_t = \rho(\rho X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \rho^2 X_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t.$$

Par itération, on obtient la représentation

$$X_t = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j},$$

qui est précisément un cas particulier de la formule (2.1) avec $\psi_j = \rho^j$. En utilisant le résultat sur la somme géométrique infinie, on note que

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{\infty} (\rho^j)^2 = \sum_{j=0}^{\infty} (\rho^2)^j = \frac{1}{1-\rho^2} < \infty.$$

2.1.6 Opérateur délai

L'opérateur *délai* permet d'écrire simplement un processus à partir des réalisations du passé. Spécifiquement, pour un processus stochastique $(X_t)_{t \in \mathbb{Z}}$, l'opérateur délai, noté \mathcal{B} , est défini par

$$\mathcal{B}X_t = X_{t-1}.$$

Lemme 2.1. *Soit une série chronologique $(X_t)_{t \in \mathbb{Z}}$. L'opérateur délai possède les propriétés suivantes :*

- (i) $\mathcal{B}^j X_t = X_{t-j}$; en particulier, $\mathcal{B}^0 X_t = X_t$;
- (ii) Si $X_t = c \in \mathbb{R}$ pour tout $t \in \mathbb{Z}$, alors $\mathcal{B}^j X_t = \mathcal{B}^j c = c$ pour tout $j \in \mathbb{Z}$;
- (iii) $\mathcal{B}^j (\mathcal{B}^k X_t) = \mathcal{B}^{j+k} X_t = X_{t-j-k}$;
- (iv) $\mathcal{B}^{-j} X_t = X_{t+j}$;
- (v) $(\mathcal{B}^j + \mathcal{B}^k) X_t = \mathcal{B}^j X_t + \mathcal{B}^k X_t = X_{t-j} + X_{t-k}$;
- (vi) Si $|a| < 1$, alors

$$(1 - a\mathcal{B})^{-1} X_t = \sum_{j=0}^{\infty} a^j X_{t-j}.$$

Remarque 2.1. *La propriété (vi) est une conséquence de la somme géométrique. En effet, on a*

$$(1 - a\mathcal{B})^{-1} = \sum_{j=0}^{\infty} a^j \mathcal{B}^j.$$

Il est possible de représenter une série temporelle sous la forme d'un polynôme, à l'aide de l'opérateur délai. Par exemple, soit le polynôme d'ordre q

$$\Psi(\mathcal{B}) = \Psi_0 \mathcal{B}_0 + \Psi_1 \mathcal{B}_1 + \Psi_2 \mathcal{B}_2 + \cdots + \Psi_q \mathcal{B}_q.$$

En appliquant ce polynôme à un processus aléatoire $(\varepsilon_t)_{t \in \mathbb{Z}}$, on a

$$\begin{aligned} \Psi(\mathcal{B})\varepsilon_t &= \Psi_0 \mathcal{B}_0 \varepsilon_t + \Psi_1 \mathcal{B}_1 \varepsilon_t + \cdots + \Psi_q \mathcal{B}_q \varepsilon_t \\ &= \Psi_0 \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \cdots + \Psi_q \varepsilon_{t-q}. \end{aligned}$$

Exemple 2.3. À l'aide de la notation fournie par l'opérateur délai, la représentation d'un processus stationnaire d'ordre deux $(X_t)_{t \in \mathbb{Z}}$ énoncée dans le théorème de Wold (1948) peut s'écrire sous la forme

$$X_t = \Psi(\mathcal{B})\varepsilon_t + \kappa_t,$$

où $\Psi(\mathcal{B})$ est un polynôme de degré infini défini par

$$\Psi(\mathcal{B}) = \sum_{j=0}^{\infty} \psi_j \mathcal{B}^j,$$

où $\psi_0 = 1$ et $\sum_{j=0}^{\infty} \psi_j^2 < \infty$.

2.2 Processus aléatoires stationnaires

2.2.1 Modèle à moyenne mobile : processus MA

Le processus $(X_t)_{t \in \mathbb{Z}}$ satisfait une représentation d'un processus *moyenne mobile* d'ordre q , noté MA(q), si et seulement si

$$X_t = \mu + \Theta(\mathcal{B})\varepsilon_t. \quad (2.2)$$

où

$$\Theta(\mathcal{B}) = \sum_{j=0}^q \theta_j \mathcal{B}^j$$

et $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont i.i.d. $\mathcal{N}(0, \sigma^2)$. Ici, $\theta_0, \theta_1, \dots, \theta_q$ sont des paramètres réels. À noter que $\theta_q \neq 0$ et que, par convention, on pose $\theta_0 = 1$. Une représentation équivalente pour le processus (2.2) est

$$X_t = \mu + (1 - \theta_1 \mathcal{B} - \theta_2 \mathcal{B}^2 - \dots - \theta_q \mathcal{B}^q) \varepsilon_t.$$

À noter que $E(X_t) = \mu$. La fonction d'autocorrélation est donnée par

$$\rho_\ell = \begin{cases} 1, & \ell = 0 \\ \frac{-\theta_\ell + \theta_1\theta_{\ell+1} + \cdots + \theta_{q-\ell}\theta_q}{\sum_{j=0}^q \theta_j^2}, & 0 < \ell \leq q \\ 0, & \ell > q. \end{cases} \quad (2.3)$$

Exemple 2.4. Le processus $X_t = \varepsilon_t + 2\varepsilon_{t-3}$ est un processus $MA(3)$ de moyenne nulle, c'est-à-dire que $\mu = 0$. Le processus $X_t = 4 + \varepsilon_t + 0.8\varepsilon_{t-1}$ est un processus $MA(1)$ de moyenne $\mu = 4$. Des réalisations de taille $T = 100$ de ces séries sont présentées à la figure 2.1.

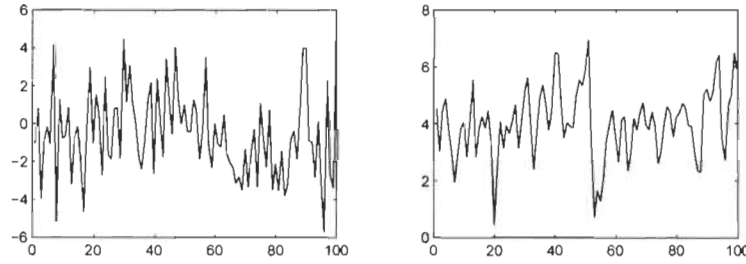


Figure 2.1: Réalisations des processus $X_t = \varepsilon_t + 2\varepsilon_{t-3}$ et $X_t = 4 + \varepsilon_t + 0.8\varepsilon_{t-1}$

Le résultat suivant stipule qu'un processus MA est toujours stationnaire. Il indique également les conditions afin qu'un processus MA soit inversible, c'est-à-dire représentable en fonction des observations passées. La preuve peut se retrouver dans les ouvrages de base en séries chronologiques, par exemple dans Brockwell & Davis (2002) et Tsay (2005).

Théorème 2.2. Un processus $(X_t)_{t \in \mathbb{Z}}$ qui satisfait une représentation $MA(q)$ est toujours stationnaire. Il est inversible lorsque toutes les racines de $\Theta(B)$, notées $\beta_j \in \mathbb{C}$, $\forall j \leq q$, sont de module strictement supérieur à 1.

Exemple 2.5. Soit un processus $(X_t)_{t \in \mathbb{Z}}$ qui satisfait une représentation $MA(1)$, c'est-à-dire que

$$X_t = (1 - \theta \mathcal{B})\varepsilon_t.$$

Ici, le polynôme est $\Theta(\mathcal{B}) = 1 - \theta \mathcal{B}$ et son unique racine est $\beta = 1/\theta$. Ainsi, si $|\theta| < 1$, le processus est inversible. En effet, on peut alors écrire

$$\varepsilon_t = \frac{1}{1 - \theta \mathcal{B}} X_t = \sum_{k=0}^{\infty} (\theta \mathcal{B})^k X_t = \sum_{k=0}^{\infty} \theta^k X_{t-k}.$$

On voit alors que le processus peut s'écrire sous la forme $\Phi(\mathcal{B})X_t = \varepsilon_t$, où

$$\Phi(\mathcal{B}) = \sum_{k=0}^{\infty} \theta^k \mathcal{B}^k.$$

Ainsi, une observation au temps t dépend du passé via la représentation

$$X_t = - \sum_{k=1}^{\infty} \theta^k X_{t-k} + \varepsilon_t.$$

2.2.2 Modèle autorégressif : processus AR

Le processus $(X_t)_{t \in \mathbb{Z}}$ satisfait une représentation d'un processus *autorégressif* d'ordre p , notée $AR(p)$, si et seulement si

$$\Phi(\mathcal{B})X_t = c + \varepsilon_t, \tag{2.4}$$

où $c \in \mathbb{R}$,

$$\Phi(\mathcal{B}) = \sum_{j=0}^p \phi_j \mathcal{B}^j$$

et $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont i.i.d. $\mathcal{N}(0, \sigma^2)$. Ici, $\phi_0, \phi_1, \dots, \phi_p$ sont des paramètres réels. À noter que $\phi_p \neq 0$ et que, par convention, on pose $\phi_0 = 1$.

Contrairement au processus MA, la constante c ne correspond pas à l'espérance de X_t . On a plutôt

$$E(X_t) = c \left/ \sum_{j=0}^p \phi_j \right.$$

En effet, à partir de la représentation (2.4), on a

$$\Phi(\mathcal{B})E(X_t) = c,$$

car $E(\varepsilon_t) = 0$. La conclusion découle en posant

$$E(X_t) = E(X_{t-1}) = \cdots = E(X_{t-p}).$$

La fonction d'autocorrélation est donnée par

$$\rho_\ell = \begin{cases} 1, & \ell = 0 \\ \sum_{j=1}^{\ell-p+1} \phi_j \rho_{\ell-j}, & \ell \in \mathbb{Z} \setminus \{0\}. \end{cases} \quad (2.5)$$

Exemple 2.6. *Le processus $X_t = -0.5X_{t-1} + 0.2X_{t-3} + \varepsilon_t$ est un $AR(3)$ avec $c = 0$. Le processus $X_t = 0.8 + 0.7X_{t-1} + 0.25X_{t-2} + \varepsilon_t$ est un $AR(2)$ avec $c = 0.8$. Dans ce cas, on obtient que $E(X_t) = 0.8/(1 - 0.7 - 0.25) = 16$. Des réalisations de taille $T = 200$ de ces séries sont présentées à la figure 2.2.*

Un processus AR est inversible, c'est-à-dire qu'il peut toujours s'écrire sous la forme d'un processus MA. En particulier, pour un processus $AR(1)$ de la forme $X_t = \rho X_{t-1} + \varepsilon_t$, on a

$$X_t = \rho X_{t-1} + \varepsilon_t = \frac{1}{1 - \rho\mathcal{B}} \varepsilon_t.$$

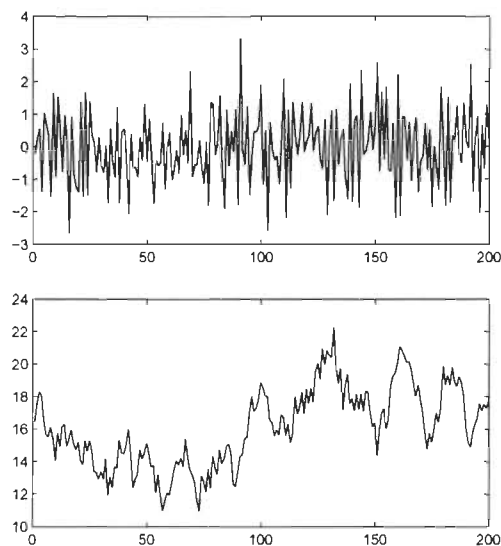


Figure 2.2: Réalisations des processus $X_t = -0.5X_{t-1} + 0.2X_{t-3} + \varepsilon_t$ et $X_t = 0.8 + 0.7X_{t-1} + 0.25X_{t-2} + \varepsilon_t$

Comme

$$\frac{1}{1 - \rho\mathcal{B}} = \sum_{i=0}^{\infty} (\rho\mathcal{B})^i,$$

on obtient

$$X_t = \sum_{i=0}^{\infty} \rho^i \mathcal{B}^i \varepsilon_t = \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i},$$

ce qui correspond à un processus MA d'ordre infini.

Un processus AR n'est cependant pas toujours stationnaire. D'abord, pour que les trois conditions de stationnarité du second ordre d'un processus AR(p) soient satisfaites, il faut que

$$\sum_{j=0}^{\infty} \phi^{2j} < \infty \quad \text{et} \quad \gamma(\ell) = \sigma_{\varepsilon}^2 \sum_{j=0}^{\infty} \phi^{2j+\ell}.$$

En effet, un tel processus satisfait les conditions de stationnarité en autant que toutes les racines du polynôme $\Phi(\mathcal{B})$ aient un module strictement supérieur à 1. Pour le montrer, notons les racines de $\Phi(\mathcal{B})$ par $\lambda_1, \dots, \lambda_p$. Celles-ci peuvent être des nombres complexes. Ainsi, comme ce sont des racines, on a pour chaque $k \in \{1, \dots, p\}$ que $\Phi(\lambda_k) = 0$. Ainsi,

$$\Phi(\mathcal{B}) = \prod_{j=1}^p \left(1 - \frac{1}{\lambda_j} \mathcal{B}\right).$$

Exemple 2.7. Soit un processus $(X_t)_{t \in \mathbb{Z}}$ qui satisfait une représentation $AR(1)$. Il est possible d'écrire

$$X_t = \frac{1}{1 - \phi \mathcal{B}} \varepsilon_t = \sum_{k=0}^{\infty} (\phi \mathcal{B})^k \varepsilon_t = \sum_{k=0}^{\infty} \phi^k \varepsilon_{t-k}.$$

Autrement dit, $(X_t)_{t \in \mathbb{Z}}$ peut se représenter par un processus MA d'ordre infini de paramètres $\psi_j = \phi^j$, $j \in \{0, 1, \dots\}$. Dans ce cas particulier, $\Phi(\mathcal{B}) = 1 - \phi \mathcal{B}$, ce qui fait que $\lambda = 1/\phi$ est l'unique racine de ce polynôme. Dans ce cas, la condition de stationnarité $|\lambda| > 1$ est équivalente à $|\phi| < 1$.

Exemple 2.8. Le polynôme associé au processus

$$X_t = -0.5X_{t-1} + 0.3X_{t-2} + \varepsilon_t$$

est

$$\Phi(\mathcal{B}) = 1 + 0.5\mathcal{B} - 0.3\mathcal{B}^2.$$

On trouve alors que les racines de $\Phi(\mathcal{B})$ sont $\lambda_1 = 2.84$ et $\lambda_2 = -1.1736$. Comme $|\lambda_1| > 1$ et $|\lambda_2| > 1$, ce processus est stationnaire.

Exemple 2.9. Le polynôme associé au processus

$$X_t = 2X_{t-1} - 3X_{t-2} + 2X_{t-3} + \varepsilon_t$$

est

$$\Phi(\mathcal{B}) = 1 - 2\mathcal{B} + 3\mathcal{B}^2 - 2\mathcal{B}^3.$$

Ses racines, obtenu avec le logiciel Maple, sont

$$\lambda_1 = 1, \quad \lambda_2 = 0.25 - 0.25\sqrt{-7} \quad \text{et} \quad \lambda_3 = 0.25 + 0.25\sqrt{-7}.$$

Comme

$$|\lambda_2| = |\lambda_3| = \sqrt{(0.25)^2 + (0.25\sqrt{7})^2} = 0.0625 < 1,$$

ce processus n'est pas stationnaire.

Exemple 2.10. Le polynôme associé au processus

$$X_t = 0.1X_{t-1} - 0.5X_{t-2} + \varepsilon_t$$

est

$$\Phi(\mathcal{B}) = 1 - 0.1\mathcal{B} + 0.5\mathcal{B}^2.$$

Ses racines sont

$$\lambda_1 = 0.1 - \sqrt{-2} \quad \text{et} \quad \lambda_2 = 0.1 + \sqrt{-2}.$$

Comme

$$|\lambda_1| = |\lambda_2| = \sqrt{(0.1)^2 + (\sqrt{2})^2} > \sqrt{2} > 1,$$

ce processus est stationnaire.

2.2.3 Combinaison des modèles AR et MA : processus ARMA

Le processus $(X_t)_{t \in \mathbb{Z}}$ satisfait une représentation d'un processus *autorégressif moyenne mobile* d'ordres p et q , noté ARMA(p, q), si et seulement si

$$\Phi(\mathcal{B})X_t = c + \Theta(\mathcal{B})\varepsilon_t,$$

où $c \in \mathbb{R}$,

$$\Phi(\mathcal{B}) = \sum_{j=0}^p \phi_j \mathcal{B}^j \quad \text{et} \quad \Theta(\mathcal{B}) = \sum_{j=0}^q \theta_j \mathcal{B}^j.$$

Par convention, on pose $\theta_0 = \phi_0 = 1$. À l'instar des processus MA et AR, $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont i.i.d. $\mathcal{N}(0, \sigma^2)$.

Notons aussi que les processus AR(p) et MA(q) sont des cas particuliers du processus ARMA(p, q). En effet, le processus AR(p) correspond à un processus ARMA($p, 0$), alors que le processus MA(q) correspond à ARMA($0, q$).

Deux exemples de processus *autorégressifs moyenne mobile* sont présentés dans la suite.

Exemple 2.11. Soit le processus ARMA($1, 1$) défini par

$$X_t = \rho X_{t-1} + \alpha \varepsilon_{t-1} + \varepsilon_t.$$

Une façon équivalente d'écrire ce processus est

$$(1 - \rho B)X_t = (1 + \alpha B)\varepsilon_t.$$

Il en découle

$$\begin{aligned}
X_t &= (1 + \alpha B) \left(\frac{1}{1 - \rho B} \right) \varepsilon_t \\
&= (1 + \alpha B) \left(\sum_{j=0}^{\infty} \rho^j B^j \right) \varepsilon_t \\
&= \left(\sum_{j=0}^{\infty} \rho^j B^j + \sum_{j=0}^{\infty} \alpha \rho^j B^{j+1} \right) \varepsilon_t \\
&= \left(1 + \sum_{j=1}^{\infty} \rho^j B^j + \sum_{j=1}^{\infty} \alpha \rho^{j-1} B^j \right) \varepsilon_t \\
&= \sum_{j=1}^{\infty} (\rho^j + \alpha \rho^{j-1}) \varepsilon_{t-j} + \varepsilon_t.
\end{aligned}$$

Ainsi, le processus $ARMA(1,1)$ a une formulation équivalente en terme d'un processus $AR(\infty)$ de paramètres $\phi_j = \rho^j + \alpha \rho^{j-1}$, $j \in \{1, 2, \dots\}$. On peut montrer que

$$\text{var}(X_t) = 1 + \frac{(\alpha + \rho)^2}{1 - \rho^2}.$$

Exemple 2.12. Le processus

$$X_t = 2 - 0.5X_{t-1} + 0.2X_{t-3} + \varepsilon_t - 0.5\varepsilon_{t-1}$$

est un processus $ARMA(3,1)$. Sa moyenne est

$$E(X_t) = \frac{2}{1 + 0.5 - 0.2} = 1.5385.$$

Exemple 2.13. Le processus

$$X_t - 0.2X_{t-1} - 1.5X_{t-2} = \varepsilon_t + 0.5\varepsilon_{t-1} + 0.6\varepsilon_{t-4}$$

est un processus $ARMA(2,4)$. Sa moyenne est nulle.

Les conditions de stationnarité d'un processus ARMA sont déterminées par les racines du polynôme associé à sa composante AR, alors que son inversibilité dépend des racines du polynôme associé à sa composante MA.

Exemple 2.14. *Soit le processus*

$$X_t = 2 - 0.5X_{t-1} + 0.2X_{t-3} + \varepsilon_t - 0.5\varepsilon_{t-1}.$$

Les polynômes associés à la partie autorégressive et à la partie moyenne mobile sont respectivement

$$\Theta(\mathcal{B}) = 1 + 0.5\mathcal{B} - 0.2\mathcal{B}^3 \quad \text{et} \quad \Phi(\mathcal{B}) = 1 - 0.5\mathcal{B}.$$

Les racines de $\Theta(\mathcal{B})$ sont

$$\beta_1 = 2.18, \quad \beta_2 = -1.09 - 1.04i \quad \text{et} \quad \beta_3 = -1.09 + 1.04i.$$

Ce processus est donc stationnaire. De même, comme la racine de $\Phi(\mathcal{B})$ est $\lambda = 2 > 1$, le processus est inversible.

2.2.4 Modèle ARMA à composantes saisonnières : processus SARMA

Le processus $(X_t)_{t \in \mathbb{Z}}$ satisfait une représentation d'un processus *autorégressif moyenne mobile saisonnier*, noté SARMA_{s,s'}(p,q), si et seulement si

$$\sum_{j=0}^p \phi_{js} X_{t-js} = c + \sum_{j=0}^q \theta_{js'} \varepsilon_{t-js'},$$

où $c \in \mathbb{R}$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont i.i.d. $\mathcal{N}(0, \sigma^2)$. Ici, s désigne la période de saisonnalité de la composante AR, alors que s' représente la période de saisonnalité de la composante MA. Par exemple, un processus SARMA_{12,6}(2,2) s'écrit

$$X_t + \phi_{12}X_{t-12} + \phi_{24}X_{t-24} = c + \varepsilon_t + \theta_6\varepsilon_{t-6} - \theta_{12}\varepsilon_{t-12}.$$

2.3 Processus aléatoires non stationnaires

2.3.1 Généralités

À la section précédente, quelques processus aléatoires stationnaires ont été présentés et étudiés. Toutefois, dans la réalité, il arrive souvent que des séries observées montrent un comportement non stationnaire. Dans ces circonstances, il convient alors d'identifier le type de non stationnarité auquel la série est contrainte. Cela permet d'utiliser un modèle approprié de façon à rendre stationnaire la série en question. Dans la suite, on présentera quelques exemples de processus non stationnaires, notamment les processus TS et DS.

Un exemple de processus non stationnaire est donné par $(X_t)_{t \in \mathbb{Z}}$ tel que

$$X_t = \begin{cases} 4 + \varepsilon_t, & \text{si } t < t_0; \\ 10 + \varepsilon_t, & \text{si } t \geq t_0. \end{cases}$$

La série présentée à la figure 2.3 consiste en $T = 200$ réalisations de ce processus avec une rupture abrupte au temps $t_0 = 100$.

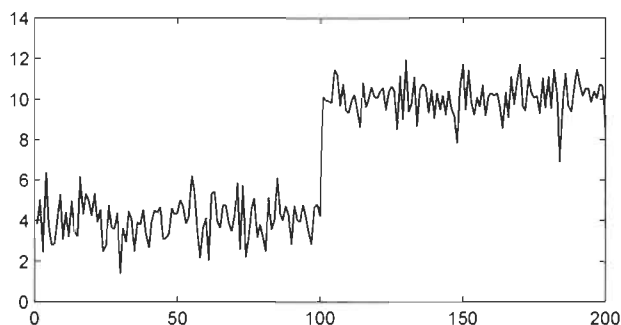


Figure 2.3: Modèle avec rupture abrupte de moyenne

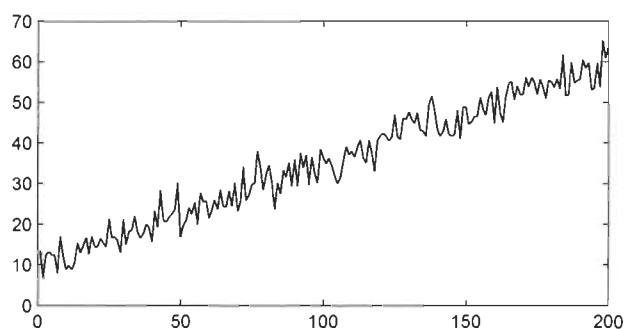


Figure 2.4: Modèle à tendance linéaire déterministe

Quelquefois, la moyenne d'une série observée semble augmenter en fonction du temps. C'est le cas de la série présentée à la figure 2.4, qui consiste en $T = 200$ réalisations du modèle $X_t = 10 + .25t + \varepsilon_t$, où $\varepsilon_t \sim \mathcal{N}(0, 9)$.

Un autre exemple de processus non stationnaire est donné par la *marche aléatoire sans dérive* $X_t = X_{t-1} + \varepsilon_t$. La figure 2.5 présente $T = 200$ réalisations de ce processus. On peut montrer que celui-ci est non stationnaire car les conditions de stationnarité faible ne sont pas satisfaites.

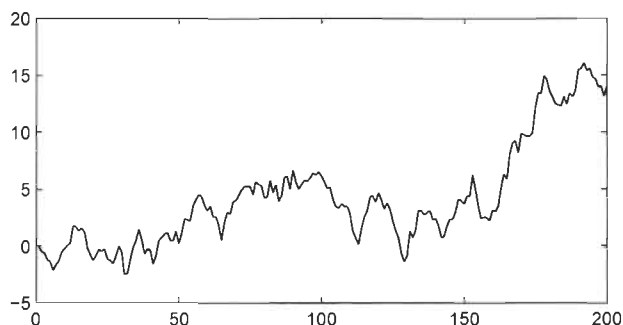


Figure 2.5: Marche aléatoire sans dérive

Les processus non stationnaires sont de deux types : (i) déterministe ou (ii) stochastique. Les modèles de type (i) font partie des processus à *tendance stationnaire* (TS). Dans ce cas, la non stationnarité est fonction du temps, comme pour les modèles avec rupture de moyenne et à tendance linéaire. Un processus fait partie de la classe des processus à *différence stationnaire* (DS); leur non stationnarité provient de l'accumulation des chocs stochastiques ε_t , comme pour la marche aléatoire sans dérive.

2.3.2 Modèle à tendance stationnaire: processus TS

Un processus $(X_t)_{t \in \mathbb{Z}}$ fait partie de la famille des modèles à tendance saisonnière (TS) s'il peut s'écrire sous la forme

$$X_t = f(t) + Z_t,$$

où $f(t)$ est une certaine fonction et $(Z_t)_{t \in \mathbb{Z}}$ est un processus stochastique stationnaire. Si $f'(t) \neq 0$, alors le processus $(X_t)_{t \in \mathbb{Z}}$ est non stationnaire car

$$E(X_t) = f(t) + E(Z_t) = f(t) + \mu_Z$$

dépend de t de par la fonction f .

Exemple 2.15. *Un cas particulier du processus TS est le modèle à tendance linéaire. Dans ce cas, $f(t) = A + Bt$ et $Z_t = \varepsilon_t$. Ce processus est clairement non stationnaire puisque $E(X_t) = A + Bt$ dépend de t .*

Un processus TS peut contenir une composante autorégressive. À titre d'exemple, soit le modèle

$$X_t = A + Bt + Z_t,$$

où $Z_t = \rho X_{t-1} + \varepsilon_t$.

2.3.3 Modèle à différence stationnaire: processus DS

Un processus non stationnaire $(X_t)_{t \in \mathbb{Z}}$ fait partie de la famille des modèles à différence stationnaire (DS) d'ordre d si le processus *filtré*

$$F_t = (1 - \mathcal{B})^d X_t$$

est stationnaire. Par exemple, un processus X_t est DS(2) si

$$F_t = (1 - \mathcal{B})^2 X_t = X_t - 2X_{t-1} + X_{t-2}$$

est stationnaire. Comme cas particuliers des modèles DS, on retrouve les modèles ARIMA et SARIMA. On présente ces modèles dans la suite.

Modèle ARIMA

Le processus stationnaire $(X_t)_{t \geq 0}$ satisfait une représentation ARIMA(p,q,d) si et seulement si

$$\Phi(\mathcal{B})(1 - \mathcal{B})^d X_t = c + \Theta(\mathcal{B})\varepsilon_t,$$

où $c \in \mathbb{R}$,

$$\Phi(\mathcal{B}) = \sum_{j=0}^p \phi_j \mathcal{B}^j \quad \text{et} \quad \Theta(\mathcal{B}) = \sum_{j=0}^q \theta_j \mathcal{B}^j.$$

On suppose que toutes les racines des polynômes $\Theta(\mathcal{B})$ et $\Phi(\mathcal{B})$ sont situées à l'extérieur du cercle unité.

Modèle ARIMA à composantes saisonnières : processus SARIMA

Le processus SARIMA est un modèle ARIMA dans lequel une composante saisonnière est ajoutée. Ainsi, $(X_t)_{t \geq 0}$ est un processus SARIMA(p,q,d) de périodes s et s' si et seulement si

$$\Phi(\mathcal{B}^s)(1 - \mathcal{B}^s)^d X_t = c + \Theta(\mathcal{B}^{s'})\varepsilon_t,$$

où $c \in \mathbb{R}$,

$$\Phi(\mathcal{B}) = \sum_{j=0}^p \phi_j \mathcal{B}^j \quad \text{et} \quad \Theta(\mathcal{B}) = \sum_{j=0}^q \theta_j \mathcal{B}^j.$$

La période de saisonnalité s est associée à la composante AR, alors que la période de saisonnalité s' est associée à la composante MA. On suppose que toutes les racines des polynômes $\Theta(\mathcal{B})$ et $\Phi(\mathcal{B})$ sont situées à l'extérieur du cercle unité.

Marche aléatoire

Soit la *marche aléatoire* $X_t = X_{t-1} + c + \varepsilon_t$. En écrivant la formulation équivalente

$$(1 - \mathcal{B})X_t = c + \varepsilon_t,$$

on voit que $(X_t)_{t \geq 0}$ est un processus AR(1) intégré d'ordre $d = 1$.

2.4 Processus stochastiques non linéaires

2.4.1 Modèles autorégressifs conditionnellement hétéroscédastiques : processus ARCH

Le processus $(X_t)_{t \geq 0}$ satisfait une représentation d'un processus *autorégressif conditionnellement hétéroscédastique* d'ordre $p \geq 1$, noté ARCH(p), si et seulement si

$$X_t = \sigma_t \varepsilon_t \quad \text{et} \quad \sigma_t^2 = \alpha_0 + \sum_{j=1}^p \alpha_j X_{t-j}^2.$$

Ici, $\alpha_0, \alpha_1, \dots, \alpha_p$ sont des paramètres à valeurs dans \mathbb{R}^+ et le processus $(\varepsilon_t)_{t \geq 0}$ est un bruit blanc tel que ε_t est indépendant de $\{X_{t-\ell}, \ell \geq 1\}$ pour tout t . Contrairement aux modèles ARMA, les modèles ARCH permettent à la volatilité, c'est-à-dire la variance, de varier en fonction du temps.

2.4.2 Modèles ARCH généralisés : processus GARCH

Le modèle GARCH(p,q) pour un processus $(X_t)_{t \geq 0}$ s'écrit sous la forme

$$X_t = \sigma_t \varepsilon_t \quad \text{et} \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2. \quad (2.6)$$

Ici, $\alpha_0, \alpha_1, \dots, \alpha_p$ et β_1, \dots, β_q sont des paramètres à valeurs dans \mathbb{R}^+ et le processus $(\varepsilon_t)_{t \geq 0}$ est un bruit blanc tel que ε_t est indépendant de $\{X_{t-\ell}, \ell \geq 1\}$ pour tout t . Les modèles GARCH généralisent les modèles ARCH en permettant à la variance conditionnelle de suivre un processus ARMA(p,q) via l'introduction d'une composante autorégressive.

CHAPITRE 3

MÉTHODE DE BOX & JENKINS

Box & Jenkins (1970) ont élaboré une méthodologie pour identifier un modèle adéquat pour une série chronologique. Leur méthode est fondée sur les modèles ARIMA. Dans ce chapitre, les principales étapes de cette technique sont présentées. Pour les méthodes d'inférence présentées dans la suite, on supposera que T réalisations d'une série chronologique univariée, notées X_1, \dots, X_T ont été observées.

3.1 Validation des données

Il est important de se familiariser avec le jeu de données. Cela consiste à vérifier si les observations sont homogènes par rapport au temps, à s'assurer de leur exactitude et à détecter de possibles valeurs extrêmes. Cette étape s'effectue essentiellement par l'examen visuel de la série. Cela permet, en outre, d'identifier une possible tendance ou la présence d'une variation cyclique, par exemple une saisonnalité.

3.2 Identification du modèle

De façon générale, l'étape d'identification du modèle consiste à identifier le modèle qui représente au mieux la série étudiée. En d'autres mots, il s'agit de trouver un modèle stationnaire qui tient compte de la variabilité dans le temps et pour lequel il y a absence d'autocorrélation des résidus. Plus particulièrement, cette étape implique les méthodes d'estimation du paramètre d'intégration d , l'estimation des ordres p et q , les tests d'hétéroscédasticité, les tests de non stationnarité ou de racine unitaire et les transformations non linéaires. Ces thèmes sont abordés dans ce chapitre.

3.2.1 Estimation du paramètre d'intégration

Lorsqu'un processus possède une non stationnarité de type DS, on parle alors d'un modèle *intégré*. Dans ce cas, il convient de déterminer l'*ordre d'intégration* d du processus filtré $(1 - \mathcal{B})^d X_t$ pour lequel le processus est stationnaire, d'où le processus $I(d)$.

Pour estimer le paramètre d'intégration d , on peut comparer les variances pour différentes séries différenciées telles que $\{X_t\}$, $\{\Delta X_t\}$, $\{\Delta_s X_t\}$, $\{\Delta\Delta_s X_t\}$, $\{\Delta^2 X_t\}$, où s désigne la période. La série ayant la plus petite variance conduit souvent à la modélisation la plus simple. Il est également possible de vérifier les autocorrélations de chaque série. Habituellement, une différence première est nécessaire lorsque les autocorrélations décroissent de manière linéaire, alors qu'une différence saisonnière est caractérisée par

des autocorrélations de périodes $s, 2s, 4s, \dots$ qui sont proches de 1 et qui décroissent de façon linéaire. On peut aussi utiliser l'approche empirique par l'autocorrélogramme qui consiste à appliquer diverses différenciations et à retenir celle qui semble présenter la meilleure stabilisation au niveau de la moyenne.

Remarque 3.1. *L'objectif visé par les différenciations est de stabiliser la moyenne d'une série.*

Exemple 3.1. *Soit $(Z_t)_{t \in \mathbb{Z}}$, un processus ARMA(2,2) défini par*

$$\Phi(\mathcal{B})X_t = \Theta(\mathcal{B})\varepsilon_t,$$

où

$$\Phi(\mathcal{B}) = 1 - 2.5\mathcal{B} + 1.5\mathcal{B}^2 \quad \text{et} \quad \Theta(\mathcal{B}) = 1 - 0.5\mathcal{B}.$$

On sait que la stationnarité ou non stationnarité d'un processus s'établit à l'aide de la partie AR d'un ARMA. Les racines du polynôme à considérer sont donc $\lambda_1 = 1$ et $\lambda_2 = 2/3$, de sorte que

$$\Phi(\mathcal{B}) = \left(1 - \frac{1}{\lambda_1}\mathcal{B}\right) \left(1 - \frac{1}{\lambda_2}\mathcal{B}\right) = (1 - \mathcal{B}) \left(1 - \frac{3}{2}\mathcal{B}\right).$$

Comme $(Z_t)_{t \in \mathbb{Z}}$ n'est pas stationnaire, on trouvera alors l'ordre d'intégration $d = 1$ pour lequel il sera stationnaire.

3.2.2 Estimation des ordres p et q

Identification d'un processus MA

Soit un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ satisfaisant une représentation MA(q) définie à l'équation (2.2). Pour déterminer la valeur de q , on se base sur la

fonction d'autocorrélation du processus MA donnée à l'équation (2.3). En fait, q correspondra au plus grand délai tel que l'autocorrélation n'est pas statistiquement égale à 0.

Identification d'un processus AR

Soit un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ satisfaisant une représentation AR(p) définie à l'équation (2.4). Globalement, l'identification d'un processus AR(p) s'effectue de la même façon que celle d'un processus MA(q). La seule différence réside dans le fait que c'est l'autocorrélation partielle, plutôt que l'autocorrélation proprement dite, qui est utilisée.

Identification d'un processus ARMA

Il existe plusieurs méthodes pour déterminer les ordres p et q d'un processus ARMA. L'une d'elles est la *méthode du coin* élaborée par Beguin et al. (1980). On peut également se baser sur les autocorrélations et les autocorrélations partielles. Pour obtenir l'ordre de la composante MA, il faut identifier l'autocorrélation significative dont l'ordre est le plus élevé; pour la composante AR, il faut identifier l'autocorrélation partielle significative dont l'ordre est le plus élevé.

3.2.3 Tests d'hétéroscédasticité

Une des hypothèses de base des modèles ARMA est que la variance est indépendante du temps. On dit alors que la série est homoscedastique. À l'inverse, il arrive que des séries observées montrent une certaine forme d'*hétéroscédasticité*, c'est-à-dire une variabilité qui se modifie au cours du temps. Il est donc important d'identifier lorsque cela se produit et, éventuellement, tenter de trouver une transformation qui stabilise la variance.

Bien qu'il existe plusieurs méthodes pour identifier de l'*hétéroscédasticité*, seul le test de Engle (1982) sera décrit ici. L'idée de ce test consiste à effectuer une régression à partir des résidus au carré selon le modèle

$$\varepsilon_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 + \eta_t,$$

où $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$. Les observations sont les résidus au carré, à savoir ε_t^2 , obtenus à partir d'un modèle ARMA. Les hypothèses nulle et alternative à confronter sont

$$\mathcal{H}_0 : \alpha_1 = \dots = \alpha_q = 0 \quad \text{vs} \quad \mathcal{H}_1 : \alpha_j \neq 0 \text{ pour un certain } j \in \{1, \dots, q\}.$$

L'hypothèse \mathcal{H}_0 correspond à l'homoscedasticité.

La statistique de test proposée par Engle (1982) est définie par $E = T R^2$, où T est le nombre d'observations et R^2 est le coefficient de détermination défini par

$$R^2 = \frac{\text{Variation expliquée}}{\text{Variation totale}}.$$

Dans le cas présent, on a

$$R^2 = 1 - \frac{\sum_{t=1}^T \hat{\eta}_t^2}{\sum_{t=1}^T (\hat{\varepsilon}_t - \bar{\varepsilon})},$$

où

$$\hat{\eta}_t = \hat{\varepsilon}_t + \hat{\theta}_1 \hat{\varepsilon}_{t-1} + \cdots + \hat{\theta}_{q-1} \hat{\varepsilon}_{t-q+1}.$$

La loi asymptotique de E est une khi-carré à q degrés de liberté, notée χ_q^2 . Ainsi, l'hypothèse \mathcal{H}_0 d'homoscédasticité est rejetée, au seuil α , si $E > \chi_{\alpha;q}^2$, où le nombre $\chi_{\alpha;q}^2$ est tel que $P(\chi_q^2 > \chi_{\alpha;q}^2) = \alpha$.

3.2.4 Tests de non stationnarité

Il existe plusieurs tests de non stationnarité ou de racine unitaire, ces tests ayant pour objet de déterminer la présence de non stationnarité. Ces tests permettent aussi d'identifier le type de non stationnarité (TS ou DS) d'un processus. Parmi les tests de non stationnarité les plus populaires, on a ceux de Dickey & Fuller (1979) et de Phillips & Perron (1988). Cependant, on n'abordera pas davantage ces tests dans ce travail.

Remarque 3.2. *Il existe également des tests de stationnarité tel que celui de Kwiatkowski & Shin (1992), nommé aussi test de KPSS. Contrairement aux tests de non stationnarité, les tests de stationnarité soutiennent la stationnarité comme hypothèse nulle.*

3.2.5 Transformations non linéaires

Les transformations non linéaires permettent de stabiliser la variance d'une série chronologique. Celles-ci sont utiles quand il est difficile de trouver un modèle qui *blanchit* correctement les résidus. Ainsi, un modèle ARMA ayant subi une transformation logarithmique s'écrit

$$\Phi(\mathcal{B})(\ln X_t) = c + \Theta(\mathcal{B})\varepsilon_t.$$

Afin d'identifier une transformation appropriée, une idée consiste à définir

$$\tilde{X}_t = \begin{cases} X_t^{1/\lambda}, & \lambda \neq 0; \\ \exp(X_t), & \lambda = 0. \end{cases}$$

Le Tableau 3.1 présente quelques-unes des principales transformations.

Tableau 3.1: Quelques exemples de la transformation de Box–Cox

λ	Type de transformation
-1	Inverse
-0.5	Racine carrée inverse
0	Logarithmique
1/3	Racine cubique
0.5	Racine carrée
1	\emptyset
2	carrée
3	cubique

Remarque 3.3. *Il est plus facile de choisir la transformation appropriée après avoir appliqué les opérateurs de différence adéquats.*

3.3 Estimation des paramètres

3.3.1 Maximum de vraisemblance

Une méthode populaire pour estimer les paramètres d'un modèle est le maximum de vraisemblance. La fonction de vraisemblance associée à un échantillon X_1, \dots, X_T i.i.d. d'une loi dont la densité est $f(x|\theta)$, avec $\theta = (\theta_1, \dots, \theta_k) \in \mathbb{R}^k$, est définie par

$$L(\theta) = \prod_{t=1}^T f(X_t|\theta).$$

L'estimateur du maximum de vraisemblance (EMV) est la valeur $\hat{\theta}_{\text{EMV}}$ qui maximise $L(\theta)$. Parfois, il est possible de déduire cet estimateur en dérivant $L(\theta)$ par rapport à chacun des paramètres $(\theta_1, \dots, \theta_k)$ et de résoudre le système à k équations

$$\frac{\partial L(\theta)}{\partial \theta_j} = 0, \quad \text{où } j = 1, \dots, k.$$

Pour illustrer l'utilisation de la méthode du maximum de vraisemblance dans l'étude des séries chronologiques, supposons un processus stationnaire ARMA(p,q) dont les innovations sont de loi $\mathcal{N}(0, \sigma^2)$. On rappelle que la fonction de densité de la loi normale d-variée est donnée par

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu) \right\}.$$

où $\mu = (\mu_1, \dots, \mu_d)$ est le vecteur des moyennes, Σ est la matrice de variance-covariance, et $\mathbf{x} = (x_1, \dots, x_d)$. La fonction de vraisemblance du processus ARMA est donc

$$L(\phi, \theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{T/2}} \frac{1}{\sqrt{|\Omega(\theta, \phi)|}} \exp \left\{ -\frac{1}{2\sigma^2} \mathbf{X}^\top \Omega^{-1}(\theta, \phi) \mathbf{X} \right\},$$

où $\mathbf{X} = (X_1, \dots, X_T)^\top$ et σ^2 est la variance des innovations et $\Omega(\theta, \phi) \in \mathbb{R}^{T \times T}$ est la matrice de variance-covariance de \mathbf{X} , qui est fonction des paramètres $\theta = (\theta_1, \dots, \theta_q)$ et $\phi = (\phi_1, \dots, \phi_p)$.

3.3.2 Application au processus AR(1)

Soit un modèle AR(1) de la forme

$$X_t = c + \phi X_{t-1} + \varepsilon_t,$$

où $\varepsilon_1, \varepsilon_2, \dots$ sont i.i.d. $\mathcal{N}(0, \sigma^2)$ et $|\phi| < 1$. Par conséquent, la loi conditionnelle de X_t étant donnée X_{t-1} est $\mathcal{N}(c + \phi X_{t-1}, \sigma^2)$. La densité conditionnelle de $X_t | X_{t-1}$ est donc

$$f_{X_t | X_{t-1}=a}(y | c, \phi, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y - c - \phi a)^2 \right\}.$$

Ainsi, la fonction de vraisemblance conditionnelle du modèle AR(1) s'écrit

$$L(c, \phi, \sigma^2) = \prod_{t=2}^T \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2} (X_t - c - \phi X_{t-1})^2 \right\}.$$

Pour les modèles ARMA, la vraisemblance conditionnelle et la vraisemblance inconditionnelle fournissent des estimateurs qui sont asymptotiquement équivalents. On peut néanmoins trouver la fonction de vraisemblance inconditionnelle, en supposant que la loi inconditionnelle de X_t est normale avec une

certaine moyenne $\tilde{\mu}$ et une certaine variance $\tilde{\sigma}^2$. Pour ce faire, il est bon de rappeler que pour toute paire de variables aléatoires (W_1, W_2) , on a

$$(i) \quad E(W_1) = E\{E(W_1|W_2)\};$$

$$(ii) \quad \text{var}(W_1) = E\{\text{var}(W_1|W_2)\} + \text{var}\{E(W_1|W_2)\}.$$

De la propriété (i), on déduit que

$$\begin{aligned} E(X_t) &= E_{X_{t-1}}\{E_{X_t}(X_t|X_{t-1})\} \\ &= E_{X_{t-1}}(c + \phi X_{t-1}) \\ &= c + \phi E(X_{t-1}). \end{aligned}$$

En posant $\tilde{\mu} = E(X_t) = E(X_{t-1})$, il s'ensuit facilement que

$$\tilde{\mu} = \frac{c}{1 - \phi}.$$

De même, en utilisant la propriété (ii), on a

$$\begin{aligned} \text{var}(X_t) &= E_{X_{t-1}}\{\text{var}_{X_t}(X_t|X_{t-1})\} + \text{var}_{X_{t-1}}\{E_{X_t}(X_t|X_{t-1})\} \\ &= E_{X_{t-1}}(\sigma^2) + \text{var}_{X_{t-1}}(c + \phi X_{t-1}) \\ &= \sigma^2 + \phi^2 \text{var}(X_{t-1}). \end{aligned}$$

En posant $\tilde{\sigma}^2 = \text{var}(X_t) = \text{var}(X_{t-1})$, on déduit

$$\tilde{\sigma}^2 = \frac{\sigma^2}{1 - \phi^2}.$$

Donc, la loi inconditionnelle de X_t est une normale de moyenne $\tilde{\mu}$ et de variance $\tilde{\sigma}^2$.

3.4 Adéquation et validation du modèle

Cette étape consiste à s'assurer que le modèle retenu décrit bien la série d'observations. Par exemple, on peut vérifier si la présence ou non d'une tendance dans le modèle est justifiée considérant la représentation graphique de la série. Aussi, il convient de s'assurer que les résidus ε_t ne possèdent pas d'information permettant de prétendre à un meilleur modèle. Pour cela, il s'agit de vérifier les trois conditions d'un processus *bruit blanc* à l'aide de tests appropriés. Dans cette optique, regardons maintenant quelques tests permettant de vérifier certaines de ces conditions.

3.4.1 Test sur la moyenne des résidus

Le test sur la moyenne des résidus fait référence à la première condition d'un processus *bruit blanc*, à savoir une moyenne égale à 0. Ainsi, pour vérifier que les résidus d'un certain modèle sont de moyenne nulle, comme il se doit, il s'agit de confronter les hypothèses

$$\mathcal{H}_0 : E(\varepsilon_t) = 0 \quad \text{et} \quad \mathcal{H}_1 : E(\varepsilon_t) \neq 0.$$

Une statistique de test est définie par

$$\xi = \frac{\bar{\varepsilon}}{S_\varepsilon / \sqrt{T}},$$

où

$$\bar{\varepsilon} = \frac{1}{T} \sum_{t=1}^T \varepsilon_t \quad \text{et} \quad S_\varepsilon^2 = \frac{1}{T} \sum_{t=1}^T \varepsilon_t^2.$$

La statistique ξ est approximativement distribuée selon la loi normale centrée réduite. Ainsi, on rejette \mathcal{H}_0 si $|\xi| > z_{\alpha/2}$, où $z_{\alpha/2}$ est le percentile d'ordre $\alpha/2$ de la distribution normale centrée réduite.

3.4.2 Tests sur les autocorrélations

Une étape préliminaire à l'ajustement d'un modèle consiste à vérifier si le coefficient d'autocorrélation d'un certain ordre donné ℓ est nul ou pas. Il s'agit donc de confronter les hypothèses

$$\mathcal{H}_0 : \rho_\ell = 0 \quad \text{et} \quad \mathcal{H}_1 : \rho_\ell \neq 0.$$

Une statistique de test est définie par

$$\xi_\ell = \frac{\hat{\rho}_\ell}{S_{\rho_\ell}},$$

où

$$S_{\rho_\ell}^2 = \frac{1}{T} \sum_{j=-L}^L \hat{\rho}_j^2 = \frac{1}{T} \left(1 + 2 \sum_{j=1}^L \hat{\rho}_j^2 \right).$$

La statistique ξ_ℓ converge en loi vers une distribution normale centrée réduite. Ainsi, on rejette \mathcal{H}_0 si $\xi_\ell > z_{\alpha/2}$.

Plutôt que de tester pour un seul coefficient d'autocorrélation, on pourrait plutôt tester la nullité de plusieurs de ces coefficients. Un test global consiste ainsi à vérifier s'il y a de l'autocorrélation pour les L premiers délais. Les hypothèses à confronter sont donc

$$\mathcal{H}_0 : \rho_1 = \dots = \rho_\ell = 0 \quad \text{et} \quad \mathcal{H}_1 : \rho_\ell \neq 0 \text{ pour un certain } \ell \in \{1, \dots, L\}.$$

Une statistique de test proposée par Box & Pierce (1970) est

$$Q_L = T \sum_{\ell=1}^L \hat{\rho}_\ell^2.$$

Une version asymptotiquement équivalente a été introduite par Ljung & Box (1978), à savoir

$$Q'_L = T(T+2) \sum_{\ell=1}^L \frac{\hat{\rho}_\ell^2}{T-\ell}.$$

La loi asymptotique des statistiques Q_L et Q'_L est la khi-carré à L degrés de liberté. On rejette donc \mathcal{H}_0 si $Q_L > \chi_{\alpha;L}^2$ ou, de façon équivalente, si $Q'_L > \chi_{\alpha;L}^2$. Selon Mélard (1990), il est préférable d'utiliser Q'_L , plutôt que Q , lorsque l'échantillon est de petite taille.

Quelques recommandations s'imposent quant à l'utilisation et l'interprétation des tests portant sur la fonction d'autocovariance γ_ℓ . Tout d'abord, il convient de mentionner qu'aucun des deux tests n'est supérieur à l'autre. Effectivement, chacun d'eux présente des inconvénients. Par exemple, lorsque le test individuel est appliqué à plusieurs délais L , il y a de fortes chances de conclure à des autocorrélations significatives pour certains délais, même si la série est un bruit blanc. En fait, la probabilité de rejeter une hypothèse pour l'un des délais L alors que le processus est véritablement un *bruit blanc* est égale à $1 - (1 - \alpha)^L$. Cette probabilité est donc plus élevée que son seuil de probabilité α car pour $L > 1$, on a $1 - (1 - \alpha)^L > \alpha$. À l'inverse, le principal inconvénient de la méthode globale est sa faible puissance, qui réside dans le fait qu'il est possible d'accepter l'hypothèse de *bruit blanc* alors qu'il existe une autocorrélation relativement importante pour un délai particulier.

Pour éviter les risques évoqués précédemment, les tests individuels pourraient

s'effectuer pour tous les délais de 1 à L , avec un L de taille raisonnable par rapport à la longueur de la série, tout en apportant une attention particulière aux délais suspects *a priori*. Parallèlement, un test global pourrait être effectué pour les mêmes délais, rendant possible la comparaison des résultats du test global avec ceux obtenus lors des tests individuels.

3.4.3 Test de normalité des résidus

Plusieurs des modèles de séries chronologiques supposent que les termes d'innovation sont indépendants et distribués selon la loi Normale. Une façon de vérifier cette hypothèse consiste à étudier les résidus. Un des tests permettant de vérifier la normalité des résidus est celui de Jarque & Bera (1980). Les hypothèses à confronter sont

$$\mathcal{H}_0 : \varepsilon_t \sim \mathcal{N}(0, 1) \quad \text{et} \quad \mathcal{H}_1 : \varepsilon_t \not\sim \mathcal{N}(0, 1).$$

Avant de décrire le test pour les résidus d'un modèle de séries chronologiques, prenons le cas de n observations X_1, \dots, X_n indépendantes et de même loi. Dans ce cas, la statistique du test de Jarque & Bera (1980), de loi asymptotique khi-carré à deux degrés de liberté, est définie par

$$\text{JB} = \frac{n}{6} \beta_1^2 + \frac{n}{24} (\beta_2 - 3)^2,$$

où

$$\beta_1 = \frac{1}{n} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{S} \right)^3 \quad \text{et} \quad \beta_2 = \frac{1}{n} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{S} \right)^4$$

sont respectivement les coefficients d'asymétrie et d'aplatissement. Ici, \bar{X} et S sont respectivement la moyenne et l'écart-type empiriques. Sous l'hypo-

thèse de normalité, on peut montrer que

$$\frac{\sqrt{n}\beta_1}{\sqrt{6}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad \text{et} \quad \frac{\sqrt{n}(\beta_2 - 3)}{\sqrt{24}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

et que asymptotiquement, ces deux variables aléatoires sont indépendantes.

Il s'ensuit que

$$\frac{n\beta_1^2}{6} \xrightarrow{\mathcal{L}} \chi_1^2 \quad \text{et} \quad \frac{n(\beta_2 - 3)^2}{24} \xrightarrow{\mathcal{L}} \chi_1^2,$$

où χ_ν^2 représente la loi khi-carré à ν degrés de liberté. Ainsi, la loi asymptotique de JB est la distribution khi-carré à deux degrés de liberté. On rejette donc l'hypothèse de normalité, au seuil α , si $JB > \chi_{\alpha;2}^2$.

Pour appliquer le test aux résidus d'un modèle paramétrique de séries chronologiques, il faut supposer, à tort, que ceux-ci sont indépendants. On applique donc le test à $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_T$ en utilisant la valeur critique asymptotique sous l'hypothèse d'un échantillon de données indépendantes.

3.5 Prédiction

Une fois qu'on a trouvé un modèle acceptable pour la série chronologique étudiée, il est possible de calculer les prévisions. Pour aborder ce sujet, partons dans un premier temps du théorème de Wold (1948). Ainsi, on note \hat{X}_{T+h} la prévision de X_{T+h} au temps $T+h$ où T est la taille de l'échantillon des observations X_t et h l'horizon de la prévision. Selon Wold (1948), le processus X_{T+h} prend en quelque sorte la forme d'un MA(∞):

$$X_{T+h} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T+h-j} = \varepsilon_{T+h} + \psi_1 \varepsilon_{T+h-1} + \psi_2 \varepsilon_{T+h-2} + \dots$$

alors que la prévision \hat{X}_{T+h} s'écrit

$$\hat{X}_{T+h} = \sum_{j=1}^{\infty} \psi_j \varepsilon_{T+h-j} = \psi_1 \varepsilon_{T+h-1} + \psi_2 \varepsilon_{T+h-2} + \cdots$$

La prévision \hat{X}_{T+h} peut aussi s'écrire sous la forme d'une espérance conditionnelle:

$$\hat{X}_{T+h} = E(X_{T+h} | X_{T+h-1}, X_{T+h-2}, \cdots) = E(X_{T+h} | \varepsilon_{T+h-1}, \varepsilon_{T+h-2}, \cdots)$$

L'erreur de prévision est déterminée par

$$X_{T+h} - \hat{X}_{T+h} = \sum_{j=0}^{h-1} \psi_j \varepsilon_{T+h-j} = \varepsilon_{T+h}$$

Dans la pratique, on ne peut pas appliquer la décomposition de Wold (1948) pour effectuer des prévisions puisque cette représentation implique qu'on estime une infinité de paramètres (les ψ_j et le σ_ε^2). Or, on sait que lorsqu'on cherche à modéliser une série chronologique, il conviendra de choisir, pour une même qualité d'ajustement, la représentation présentant le minimum de paramètres. C'est pourquoi la décomposition de Wold (1948) ne s'avèrera jamais être la représentation optimale parmi toutes les représentations possibles d'un même processus. Cependant, étant donné qu'il existe toujours pour tout processus stationnaire, un ordre fini, il est possible de représenter convenablement n'importe quelle série stationnaire par une somme pondérée finie de chocs passés jusqu'à cet ordre. Par exemple, un processus MA(q) pourrait s'écrire à la manière d'une décomposition de Wold tronquée:

$$X_t = \mu + \sum_{j=0}^q \psi_j \varepsilon_{t-j} = \mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \cdots + \psi_q \varepsilon_{t-q}$$

où ε_t *i.i.d.* $\mathcal{N}(0, \sigma_\varepsilon^2)$, $q > 0$ et μ représente la moyenne. Ainsi, en estimant les coefficients de pondération (ψ_j) et la variance du *bruit blanc* (σ_ε^2), cela permet de proposer des prévisions pour un processus stationnaire quelconque. Pour illustrer, soit le processus ARIMA(4,1,5) de la forme

$$X_t - X_{t-1} - X_{t-4} + X_{t-5} = \varepsilon_t - 0.4\varepsilon_{t-1} - 0.5\varepsilon_{t-4} + 0.2\varepsilon_{t-5}.$$

L'estimation des résidus pour un horizon de prévision $h = 6$ sont

$$\begin{aligned}\hat{\varepsilon}_{T+1} &= -0.4\varepsilon_T - 0.5\varepsilon_{T-3} + 0.2\varepsilon_{T-4}, \\ \hat{\varepsilon}_{T+2} &= -0.5\varepsilon_{T-2} + 0.2\varepsilon_{T-3}, \\ \hat{\varepsilon}_{T+3} &= -0.5\varepsilon_{T-1} + 0.2\varepsilon_{T-2}, \\ \hat{\varepsilon}_{T+4} &= -0.5\varepsilon_T + 0.2\varepsilon_{T-1}, \\ \hat{\varepsilon}_{T+5} &= 0.2\varepsilon_T, \\ \hat{\varepsilon}_{T+6} &= 0.\end{aligned}$$

Comme $X_t = X_{t-1} + X_{t-4} - X_{t-5} + \varepsilon_t - 0.4\varepsilon_{t-1} - 0.5\varepsilon_{t-4} + 0.2\varepsilon_{t-5}$, les prévisions pour un horizon $h = 6$ sont

$$\begin{aligned}\hat{X}_{T+1} &= X_T + X_{T-3} - X_{T-4} + \hat{\varepsilon}_{T+1}, \\ \hat{X}_{T+2} &= \hat{X}_{T+1} + X_{T-2} - X_{T-3} + \hat{\varepsilon}_{T+2}, \\ \hat{X}_{T+3} &= \hat{X}_{T+2} + X_{T-1} - X_{T-2} + \hat{\varepsilon}_{T+3}, \\ \hat{X}_{T+4} &= \hat{X}_{T+3} + X_T - X_{T-1} + \hat{\varepsilon}_{T+4}, \\ \hat{X}_{T+5} &= \hat{X}_{T+4} + \hat{X}_{T+1} - X_T + \hat{\varepsilon}_{T+5}, \\ \hat{X}_{T+6} &= \hat{X}_{T+5} + \hat{X}_{T+2} - \hat{X}_{T+1} + \hat{\varepsilon}_{T+6}.\end{aligned}$$

Pour établir un intervalle de confiance sur les prévisions, il suffit d'appliquer le Théorème de Wold (1948) au processus $(X_t)_{t \in \mathbb{Z}}$ et considérer la forme MA(∞) correspondante, tel que vu précédemment. Ainsi, en supposant la normalité des résidus ε_t , on montre que

$$\frac{X_{T+h} - \hat{X}_{T+h}}{\text{var} \left(X_{T+h} - \hat{X}_{T+h} \right)^{1/2}}$$

converge en loi vers une distribution $\mathcal{N}(0, 1)$. Or, comme

$$\text{E}\{(X_{T+h} - \hat{X}_{T+h})^2\} = \text{E} \left\{ \left(\sum_{j=0}^{h-1} \psi_j \varepsilon_{T+h-j} \right)^2 \right\} = \sum_{j=0}^{h-1} \psi_j^2 \sigma_\varepsilon^2$$

on a que

$$\frac{X_{T+h} - \hat{X}_{T+h}}{\sigma_\varepsilon \left(\sum_{j=0}^{h-1} \psi_j^2 \right)^{1/2}}$$

converge en loi vers une distribution $\mathcal{N}(0, 1)$. On en déduit l'intervalle de confiance

$$IC = \left\{ \hat{X}_{T+h} \pm t_{\alpha/2} \left(\sum_{j=0}^{h-1} \psi_j^2 \right)^{1/2} \hat{\sigma}_\varepsilon \right\}.$$

CHAPITRE 4

THÉORIE DES COPULES

4.1 Introduction

Depuis quelques années, les copules s'avèrent être des outils extrêmement populaires pour modéliser la dépendance entre deux ou plusieurs variables aléatoires. Notamment, les copules sont beaucoup utilisées en finance, en actuariat et en hydrologie. L'utilité principale des copules réside dans la flexibilité qu'elles permettent lors de la construction de modèles multivariés.

La théorie des copules repose sur le Théorème de Sklar (1959). Ce résultat stipule que si H est une fonction de répartition conjointe de marges continues F_1 et F_2 , alors il existe une unique fonction $C : [0, 1]^2 \rightarrow [0, 1]$ telle que

$$H(x_1, x_2) = C \{F_1(x_1), F_2(x_2)\}.$$

La fonction C est appelée la copule de H . Ce résultat se généralise facilement au cas à $d > 2$ variables. En effet, pour une loi $H : \mathbb{R}^d \rightarrow \mathbb{R}$ de marges continues F_1, \dots, F_d , il existe une unique copule $C : [0, 1]^d \rightarrow [0, 1]$ telle que

$$H(x_1, \dots, x_d) = C \{F_1(x_1), \dots, F_d(x_d)\}.$$

Si (X_1, X_2) est un vecteur aléatoire de loi H , alors on peut distinguer trois composantes qui constituent son comportement aléatoire, à savoir

- (i) la loi marginale de X_1 , représentée par F_1 ;
- (ii) la loi marginale de X_2 , représentée par F_2 ;
- (iii) la structure de dépendance entre X_1 et X_2 , représentée par C .

Ainsi, il est possible de modéliser ces trois composantes par des étapes indépendantes. Pour ce qui est des marges (étapes (i) et (ii) ci-dessus), il existe de nombreuses méthodes classiques d'estimation de paramètres et d'adéquation. Voir, par exemple, Stephens (1976) et Cheng & Stephens (1989). Pour ce qui est du choix d'une copule, quelques méthodes existent, dont celles de Genest et al. (2006) et Mesfioui et al. (2009).

Le résultat suivant, qu'on peut retrouver dans Nelsen (2006), décrit les conditions pour qu'une fonction bivariable soit une copule.

Lemme 4.2. *Toute fonction $C : [0, 1]^2 \rightarrow [0, 1]$ qui satisfait les conditions suivantes est une copule :*

$$(i) \ C(u_1, 0) = C(0, u_2) = 0, \ C(u_1, 1) = u_1 \text{ et } C(1, u_2) = u_2;$$

$$(ii) \ \text{Pour tout } 0 \leq u_1 \leq v_1 \leq 1 \text{ et pour tout } 0 \leq u_2 \leq v_2 \leq 1,$$

$$C(v_1, v_2) - C(v_1, u_2) - C(u_1, v_2) + C(u_1, u_2) \geq 0.$$

Pour déduire la copule associée à un couple (X_1, X_2) dont la loi conjointe est H , il s'agit de poser $u_1 = F_1(x_1)$ et $u_2 = F_2(x_2)$. De là, on obtient

$$C(u_1, u_2) = H \{ F_1^{-1}(u_1), F_2^{-1}(u_2) \}. \quad (4.1)$$

Exemple 4.1. Soit la fonction de répartition

$$H(x_1, x_2) = \sqrt{\frac{x_1 x_2}{x_1 + x_2 - x_1 x_2}},$$

où $0 \leq x_1 \leq 1$ et $0 \leq x_2 \leq 1$. Les marges de cette loi sont

$$F_1(x_1) = H(x_1, 1) = \sqrt{x_1} \quad \text{et} \quad F_2(x_2) = H(1, x_2) = \sqrt{x_2}.$$

Ainsi, $F_1^{-1}(u_1) = F_2^{-1}(u_2) = u_1^2$. Par une application de la formule (4.1), on déduit que la copule de H est

$$C(u_1, u_2) = H(u_1^2, u_2^2) = \frac{u_1 u_2}{\sqrt{u_1^2 + u_2^2 - u_1^2 u_2^2}}.$$

Par quelques manipulations algébriques, on peut également écrire

$$C(u_1, u_2) = (u_1^{-2} + u_2^{-2} - 1)^{-1/2}.$$

Cette copule appartient à la famille de Clayton, qui sera décrite plus loin.

Lorsqu'une copule C est absolument continue, on peut écrire

$$C(u_1, u_2) = \int_0^{u_1} \int_0^{u_2} c(s_1, s_2) ds_1 ds_2,$$

où

$$c(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1, u_2)$$

est la densité associée à C .

4.2 Présentation de quelques copules

4.2.1 Copule d'Ali–Mikhail–Haq

La copule d'Ali–Mikhail–Haq est la structure de dépendance associée au modèle décrit par Ali et al. (1978), c'est-à-dire

$$C_{\theta}^{\text{AMH}}(u_1, u_2) = \frac{u_1 u_2}{1 - \theta(1 - u_1)(1 - u_2)}, \quad -1 \leq \theta \leq 1.$$

Deux des densités que l'on peut déduire de ce modèle sont présentées à la figure 4.1.

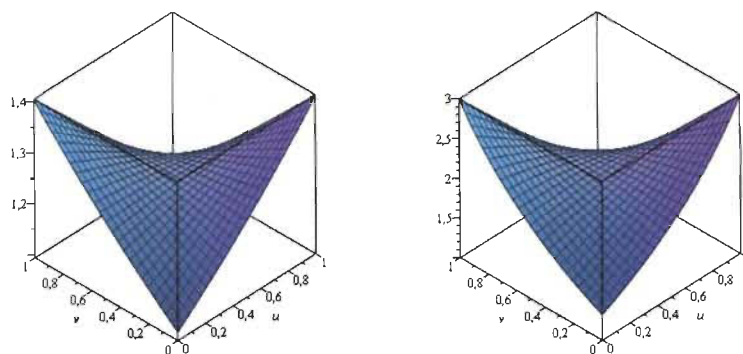


Figure 4.1: Densité de la copule de Ali–Mikhail–Haq pour $\theta = 0.10$ (à gauche) et $\theta = 0.25$ (à droite)

4.2.2 Copule de Farlie–Gumbel–Morgenstern

L'expression de la copule de Farlie–Gumbel–Morgenstern est donnée par

$$C_{\theta}^{\text{FGM}}(u_1, u_2) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad -1 \leq \theta \leq 1.$$

En dérivant, on obtient que la densité associée est de la forme

$$c_{\theta}^{\text{FGM}}(u_1, u_2) = 1 + \theta(1 - 2u_1)(1 - 2u_2).$$

Deux de ces densités sont présentées à la figure 4.2, à savoir lorsque $\theta = 0.25$ et $\theta = 0.75$.

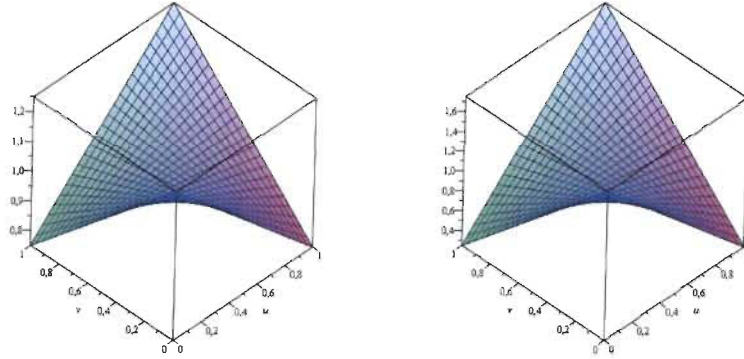


Figure 4.2: Densité de la copule de Farlie–Gumbel–Morgenstern pour $\theta = 0.25$ (à gauche) et $\theta = 0.75$ (à droite)

4.2.3 Copule de Clayton

La copule de Clayton se déduit du modèle étudié par Clayton (1978). Sa forme est donnée par

$$C_{\theta}^{\text{CL}}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \quad \theta \neq 0.$$

Comme on peut le constater sur la figure 4.3, cette fonction de dépendance accorde un poids important aux faibles valeurs de u_1 et u_2 .

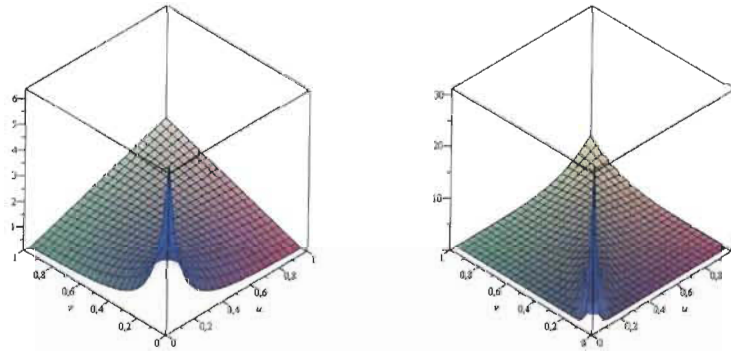


Figure 4.3: Densité de la copule de Clayton pour $\theta = 1$ (à gauche) et $\theta = 5$ (à droite)

4.2.4 Copule de Gumbel–Hougaard

La copule de Gumbel-Hougaard est

$$C_{\theta}^{\text{GH}}(u_1, u_2) = \exp \left[- \left\{ (-\log u_1)^{\theta} + (-\log u_2)^{\theta} \right\}^{1/\theta} \right], \quad \theta \geq 1.$$

Ce modèle appartient à la famille générale des copules de valeurs extrêmes. Les densités associées à cette copule ont une forme compliquée. Néanmoins, il est possible d'obtenir leur expression à l'aide d'un logiciel de calcul symbolique comme Maple. Deux de ces densités sont présentées à la figure 4.4.

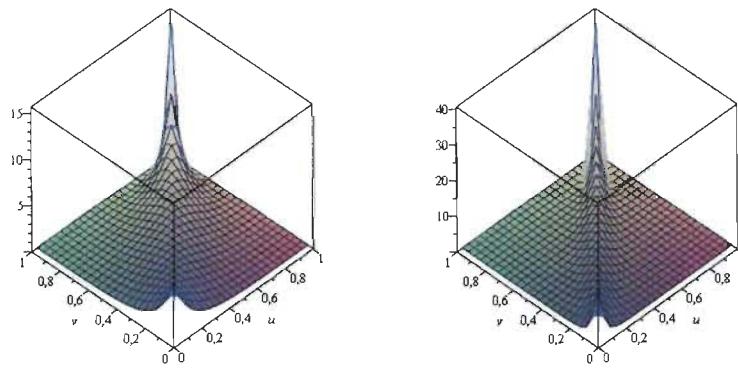


Figure 4.4: Densité de la copule de Gumbel–Hougaard pour $\theta = 3$ (à gauche) et $\theta = 7$ (à droite)

4.2.5 Copule de Frank

L'expression de la copule de Frank (voir Genest, 1987) est donnée par

$$C_{\theta}^F(u_1, u_2) = -\frac{1}{\theta} \ln \left\{ 1 - \frac{(1 - e^{-\theta u_1})(1 - e^{-\theta u_2})}{1 - e^{-\theta}} \right\}, \quad \theta \neq 0.$$

À l'instar de la copule de Clayton, cette fonction de dépendance couvre tout le spectre de la dépendance négative et positive. Deux densités de Frank sont présentées à la figure 4.5.

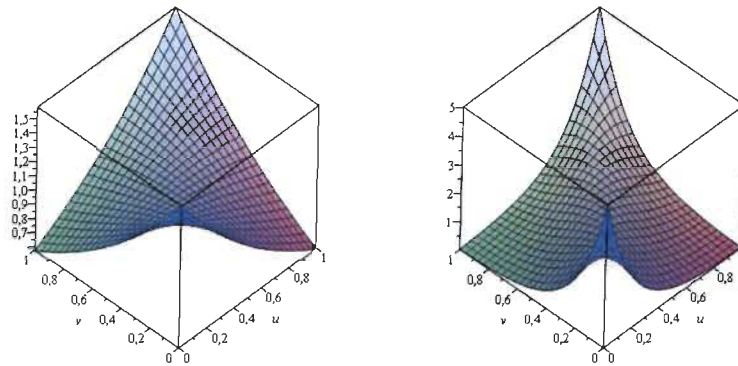


Figure 4.5: Densité de la copule de Frank pour $\theta = 1$ (à gauche) et $\theta = 5$ (à droite)

4.2.6 Copule Normale

La copule Normale est la fonction de dépendance que l'on extrait de la loi Normale bivariée. Spécifiquement, on a

$$C_{\rho}^N(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} h_{\rho}(x_1, x_2) dx_1 dx_2,$$

où

$$h_{\rho}(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)} \right\}$$

est la densité de la loi Normale à deux variables de moyennes nulles, de variances unitaires et de corrélation ρ . Ici, Φ est la fonction de répartition

de la loi $\mathcal{N}(0, 1)$. La densité associée à la copule Normale est

$$\begin{aligned}
 c_{\rho}^N(u_1, u_2) &= \frac{\partial^2}{\partial u_1 \partial u_2} C_{\rho}(u_1, u_2) \\
 &= \frac{\partial}{\partial u_1} \left\{ \frac{\partial}{\partial u_2} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} h_{\rho}(s, t) ds dt \right\} \\
 &= \frac{\partial}{\partial u_1} \left\{ \int_{-\infty}^{\Phi^{-1}(u_1)} h_{\rho}(s, \Phi^{-1}(u_2)) ds d\Phi^{-1}(u_2) \right\} \\
 &= h_{\rho} \{ \Phi^{-1}(u_1), \Phi^{-1}(u_2) \} d\Phi^{-1}(u_1) d\Phi^{-1}(u_2).
 \end{aligned}$$

À noter que

$$d\Phi^{-1}(u) = \frac{1}{\Phi' \{ \Phi^{-1}(u) \}},$$

où $\Phi'(x) = \exp(-x^2/2)/\sqrt{2\pi}$ est la densité Normale standard. À la figure 4.6, deux densités normales sont présentées, à savoir pour $\rho = 0.25$ et $\rho = 0.75$.

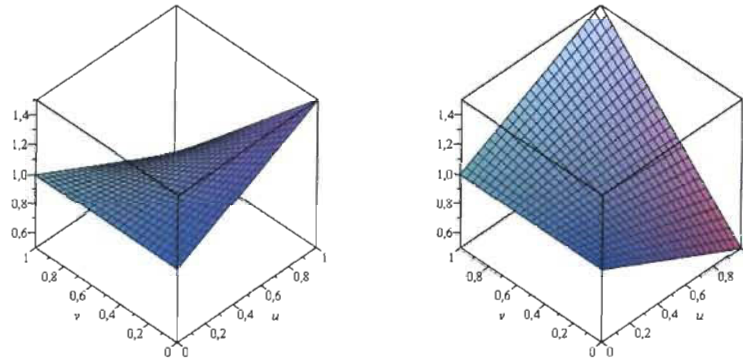


Figure 4.6: Densité de la copule Normale pour $\rho = 0.25$ (à gauche) et $\rho = 0.75$ (à droite)

4.3 Mesures de dépendance

Tout cela amène à parler des limites de la corrélation lors de l'utilisation des modèles multivariés classiques. Tout d'abord, l'utilisation des copules permet de s'assurer qu'on n'acceptera pas à tort l'hypothèse d'indépendance entre des variables aléatoires ce que la corrélation ne permet assurément pas. D'autre part, rappelons que la corrélation pour des modèles multivariés classiques est définie par

$$\rho_{X_1 X_2} = \frac{E(X_1 X_2) - \mu_{X_1} \mu_{X_2}}{\sigma_{X_1} \sigma_{X_2}}$$

où $\mu_{X_1} = E(X_1)$, $\mu_{X_2} = E(X_2)$, $\sigma_{X_1}^2 = \text{var}(X_1)$ et $\sigma_{X_2}^2 = \text{var}(X_2)$.

Remarque 4.1. *L'identité de Hoeffding (1948) stipule que*

$$E(X_1 X_2) - \mu_{X_1} \mu_{X_2} = \int_{\mathcal{X}_1} \int_{\mathcal{X}_2} \{H_{X_1, X_2}(x_1, x_2) - F_{X_1}(x_1)G_{X_2}(x_2)\} dx_2 dx_1.$$

Comme $H_{X_1, X_2}(x_1, x_2) = C\{F_{X_1}(x_1), G_{X_2}(x_2)\}$, on a

$$\begin{aligned} \rho_{X, Y} &= \frac{1}{\sigma_{X_1} \sigma_{X_2}} \int_{\mathcal{X}_1} \int_{\mathcal{X}_2} [C\{F_{X_1}(x_1), G_{X_2}(x_2)\} - F_{X_1}(x_1)G_{X_2}(x_2)] dx_1 dx_2 \\ &= \frac{1}{\sigma_{X_1} \sigma_{X_2}} \int_0^1 \int_0^1 \{C(u_1, u_2) - u_1 u_2\} dF_{X_1}^{-1}(u_1) dG_{X_2}^{-1}(u_2). \end{aligned}$$

Ainsi, la corrélation dépend des lois marginales.

Il existe des indices qui permettent de mesurer la dépendance sans que ceux-ci soient affectés par les lois marginales. Une de ces mesures est le tau de Kendall. Pour le définir, soient (X_{11}, X_{12}) et (X_{21}, X_{22}) , des couples aléatoires indépendants de loi H_{X_1, X_2} . On dit que ces couples sont concordants (\mathcal{C}) si

$(X_{11} - X_{21})(X_{21} - X_{22}) > 0$ et discordants (\mathcal{D}) si $(X_{11} - X_{21})(X_{21} - X_{22}) < 0$.

La valeur théorique du tau de Kendall est définie par

$$\tau_{X_1, X_2} = P(\mathcal{C}) - P(\mathcal{D}).$$

On a $\tau_{X_1, X_2} = 2P(\mathcal{C}) - 1$. Par la définition de concordance,

$$\begin{aligned} P(\mathcal{C}) &= P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} \\ &= P(X_1 < X_2, Y_1 < Y_2) + P(X_2 < X_1, Y_2 < Y_1) \\ &= 2P(X_1 < X_2, Y_1 < Y_2) \\ &= 2 \int_{\mathbb{R}^2} P\{X_1 < X_2, Y_1 < Y_2 | (X_2, Y_2) = (x_1, x_2)\} dH_{X_1, X_2}(x_1, x_2) \\ &= 2 \int_{\mathbb{R}^2} P(X_1 < x_1, Y_1 < y_1) dH_{X_1, X_2}(x_1, x_2) \\ &= 2 \int_{\mathbb{R}^2} H_{X_1, X_2}(x_1, x_2) dH_{X_1, X_2}(x_1, x_2) \\ &= 2E\{H_{X_1, X_2}(X_1, X_2)\}, \end{aligned}$$

où $(X_1, X_2) \sim H_{X_1, X_2}$. Ainsi, $\tau_{X_1, X_2} = 4E\{H_{X_1, X_2}(X_1, X_2)\} - 1$. À noter qu'en effectuant le changement de variables $U_1 = F_{X_1}(X_1)$ et $U_2 = F_{X_2}(X_2)$, on montre que $\tau_{X_1, X_2} = 4E\{C(U_1, U_2)\} - 1$, où $(U_1, U_2) \sim C$. Ainsi, le tau de Kendall ne dépend que de la copule associée à H_{X_1, X_2} .

Le rho de Spearman est défini par la corrélation entre U_1 et U_2 . Puisque U_1 et U_2 sont uniformément distribuées sur $(0, 1)$, on a $E(U_1) = E(U_2) = 1/2$ et $\text{var}(U_1) = \text{var}(U_2) = 1/12$. Ainsi, le rho de Spearman s'écrit

$$\rho_{X_1, X_2} = 12 E(U_1 U_2) - 3.$$

Remarque 4.2. *Sous l'indépendance, on montre que $\rho_{X_1, X_2} = \tau_{X_1, X_2} = 0$. Également, on a $\rho_{X_1, X_2} = \tau_{X_1, X_2} = 1$ en présence de dépendance positive parfaite. Sous la dépendance négative parfaite, $\rho_{X_1, X_2} = \tau_{X_1, X_2} = -1$.*

Soit $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$, un échantillon de données bivariées. La version empirique de $\tau_{X,Y}$ est définie par

$$\tau_n = \frac{4Q_n}{n(n-1)} - 1,$$

où Q_n est le nombre de paires concordantes dans l'échantillon: deux paires (X_{i1}, X_{i2}) et (X_{j1}, X_{j2}) sont dites concordantes si $(X_{i1} - X_{j1})(X_{i2} - X_{j2}) > 0$. La version empirique de ρ_{X_1, X_2} est définie par

$$\rho_n = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (R_{i1} - R_{i2})^2$$

où pour $i \in \{1, \dots, n\}$ et $j = 1, 2$, R_{ij} représente le rang de X_{ij} parmi X_{1j}, \dots, X_{nj} .

4.4 Estimation des paramètres

Il existe plusieurs méthodes pour estimer les paramètres d'une copule. On présentera ici la méthode pseudo maximum de vraisemblance. Celle-ci peut s'appliquer aux situations à plusieurs paramètres.

Soit un couple (X_1, X_2) de copule C_θ et de marges F et G . La fonction de répartition jointe est donc $H(x_1, x_2) = C_\theta\{F(x_1), G(x_2)\}$ et la densité

$$h_\theta(x_1, x_2) = c_\theta\{F(x_1), G(x_2)\}f(x_1)g(x_2),$$

où $f = dF$, $g = dG$ et $c_\theta(u_1, u_2) = \partial^2 C_\theta(u_1, u_2) / \partial u_1 \partial u_2$ sont les densités associées à F , G et C_θ respectivement. La méthode classique du maximum de

vraisemblance consiste à maximiser en fonction de θ la fonction de vraisemblance. Cependant, il est souvent plus facile de maximiser le logarithme de cette fonction. Ainsi, lorsque les marges sont connues, on a

$$m(\theta) = \sum_{i=1}^n \log h_{\theta}(X_i, Y_i) = \sum_{i=1}^n \log c_{\theta}\{F(X_i), G(Y_i)\} + K,$$

où $K = \sum_{i=1}^n \log f(X_i)g(Y_i)$ ne dépend pas de θ . La solution $\hat{\theta}_{MV}$ est l'estimateur à maximum de vraisemblance classique de θ .

L'intérêt principal de l'utilisation des copules tient au fait qu'il n'est pas nécessaire de connaître les marges pour effectuer de l'inférence sur la fonction de dépendance. Ainsi, lorsque les marges sont inconnues, on remplace habituellement les marges $F_{X_1}(x_1)$ et $G_{X_2}(x_2)$ par leurs estimations naturelles

$$F_n(x_1) = \frac{1}{n+1} \sum_{i=1}^n I(X_i \leq x_1) \quad \text{et} \quad G_n(x_2) = \frac{1}{n+1} \sum_{i=1}^n I(Y_i \leq x_2)$$

c'est-à-dire les marges empiriques. Ainsi, pour des marges inconnues, Genest et al. (1995) ont proposé de maximiser la pseudo log-vraisemblance

$$\ell(\theta) = \sum_{i=1}^n \log c_{\theta}\{F_n(X_i), G_n(Y_i)\}.$$

À remarquer que $nF_n(X_i) = R_i$ et $nG_n(Y_i) = S_i$ où R_i et S_i sont les rangs de X_i et Y_i dans leurs échantillons univariés respectifs. La méthode du pseudo maximum de vraisemblance revient donc à maximiser

$$\ell(\theta) = \sum_{i=1}^n \log c_{\theta}\left(\frac{R_i}{n}, \frac{S_i}{n}\right).$$

En posant $\dot{c}_{\theta} = \partial c_{\theta} / \partial \theta$, on voit que l'estimateur $\hat{\theta}_n$ basé sur les rangs des observations sera, en général, la solution de $\ell'(\theta) = 0$, c'est-à-dire que

$$\sum_{i=1}^n \frac{\dot{c}_{\hat{\theta}_n}}{c_{\hat{\theta}_n}}\left(\frac{R_i}{n}, \frac{S_i}{n}\right) = 0.$$

CHAPITRE 5

UN MODÈLE MULTIVARIÉ À BASE DE COPULES

Les chapitres 2 et 3 de ce travail consistait à faire un survol des différents modèles et méthodes concernant les séries chronologiques à une variable. Maintenant, à l'aide des outils de la théorie des copules, nous entamons l'objectif initial de ce mémoire, à savoir la construction d'un nouveau modèle multivarié de séries chronologiques. Ce chapitre permettra de faire le lien entre les modèles de prévision présentés au chapitre 2 et la théorie des copules abordée au chapitre 4. En d'autres mots, la théorie des copules sera exploitée afin de construire un modèle susceptible d'améliorer la capacité de prédiction par rapport aux modèles existants.

Tout d'abord, quelques modèles de séries chronologiques multivariés seront décrits. On mettra l'emphasis sur les limites de ces modèles, ce qui motivera la construction d'un modèle plus flexible dont les ingrédients de base seront les copules. Ensuite, des méthodes d'estimation des paramètres, d'adéquation et de prévision seront présentées.

5.1 Inconvénients des modèles classiques

Une série chronologique multivariée est un vecteur composé de plusieurs séries chronologiques. Les modèles classiques pour ce type de séries tiennent compte, d'une certaine façon, de la possible dépendance entre les composantes du vecteur. Pour beaucoup de ces modèles, le comportement marginal des séries est représenté par l'un des nombreux modèles de séries chronologiques présentés au chapitre 2.

Parmi les nombreux modèles de séries chronologiques multivariés, on retrouve le processus *vecteur autorégressif* (VAR). Les modèles VAR constituent une généralisation des processus AR au cas à plusieurs séries. La modélisation de tels processus repose sur l'hypothèse que le comportement des différentes variables dépend linéairement de leurs propres valeurs passées mais aussi des valeurs passées et présentes des autres variables. Ainsi, un processus stationnaire $(Y_t)_{t \in \mathbb{Z}}$ satisfait une représentation d'un processus à vecteur autorégressif d'ordre p à d variables, notée $VAR(p)$, si et seulement si

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \cdots + \Phi_p Y_{t-p} + \varepsilon_t,$$

où

$$Y_t = \begin{pmatrix} Y_{1t} \\ \vdots \\ Y_{dt} \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{dt} \end{pmatrix}, \quad \Phi_0 = \begin{pmatrix} a_1^0 \\ \vdots \\ a_d^0 \end{pmatrix}$$

et

$$\Phi_j = \begin{pmatrix} a_{11}^j & a_{12}^j & \cdots & a_{1d}^j \\ \vdots & \vdots & & \vdots \\ a_{d1}^j & a_{d2}^j & \cdots & a_{dd}^j \end{pmatrix}, \quad j \in \{1, \dots, p\}.$$

Ici, les innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ constituent un bruit blanc, c'est-à-dire qu'elles sont indépendantes, de moyenne nulle et de structure de covariance Σ_ε . On a l'expression équivalente

$$\Phi(\mathcal{B})Y_t = \Phi_0 + \varepsilon_t,$$

où

$$\Phi(\mathcal{B}) = I - \sum_{j=1}^p \Phi_j \mathcal{B}^j$$

et I est la matrice identité de dimension $p \times p$.

Une limitation de beaucoup de modèles classiques, incluant le processus VAR, est qu'en général, on suppose que la loi des innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ est normale. Autrement dit, on impose que $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma_\varepsilon)$. Ainsi, la dépendance entre les séries est de type normal. Cette hypothèse est dangereuse puisqu'il a été observé que beaucoup de phénomènes, notamment en finance, sont très mal représentés par cette loi.

Comme les copules offrent davantage de flexibilité pour modéliser la structure de dépendance, il serait intéressant de construire un modèle de séries chronologiques multivarié en permettant que la dépendance soit représentée par une copule quelconque. Dans la section suivante, un tel modèle sera présenté; quelques-unes de ses propriétés seront ensuite étudiées. Ultérieurement, on s'attend qu'un tel modèle améliore la précision des prévisions par rapport aux processus classiques.

5.2 Description d'un nouveau modèle

Par souci de simplicité, on décrira d'abord, aux Sections 5.1–5.5, un modèle et des méthodes d'inférence pour des séries chronologiques à deux variables; la généralisation à $d > 2$ variables est immédiate; elle sera présentée à la Section 5.6.

Soient T réalisations $(X_{1,1}, X_{1,2}), \dots, (X_{T,1}, X_{T,2})$ d'une série chronologique à deux variables. Afin de modéliser le comportement d'une telle série, un modèle à base de copules sera proposé. Grâce au Théorème de Sklar (1959), on pourra modéliser les séries individuelles $X_{1,1}, \dots, X_{T,1}$ et $X_{1,2}, \dots, X_{T,2}$ séparément – par des modèles ARMA, par exemple – puis de modéliser, à l'aide d'une copule, la structure de dépendance qui lie les deux séries.

Pour débiter, supposons que les séries individuelles sont générées par des processus autorégressifs d'ordre un. Ainsi, on a

$$X_{t,1} = \rho_1 X_{t-1,1} + \varepsilon_{t,1} \quad \text{et} \quad X_{t,2} = \rho_2 X_{t-1,2} + \varepsilon_{t,2},$$

où les processus $(\varepsilon_{t,1})_{t \in \mathbb{Z}}$ et $(\varepsilon_{t,2})_{t \in \mathbb{Z}}$ sont des bruits blancs tels que $\varepsilon_{t,1} \sim \mathcal{N}(0, \sigma_1^2)$ et $\varepsilon_{t,2} \sim \mathcal{N}(0, \sigma_2^2)$. Si les deux séries sont indépendantes, il est raisonnable de supposer que les processus d'innovations $(\varepsilon_{t,1})_{t \in \mathbb{Z}}$ et $(\varepsilon_{t,2})_{t \in \mathbb{Z}}$ sont indépendants.

Toutefois, si on pense qu'il y a une certaine forme de dépendance entre les deux séries, alors on peut supposer que le processus $(\varepsilon_{t,1}, \varepsilon_{t,2})_{t \in \mathbb{Z}}$ est tel que la loi conjointe de $(\varepsilon_{t,1}, \varepsilon_{t,2})$ est une certaine fonction de répartition H , fonction autre que celle représentant l'indépendance. Par les hypothèses faites sur les

lois marginales de $\varepsilon_{t,1}$ et $\varepsilon_{t,2}$, on a

$$F_{1,\sigma_1}(x_1) = P(\varepsilon_t \leq x_1) = \Phi\left(\frac{x_1}{\sigma_1}\right) \quad \text{et} \quad F_{2,\sigma_2}(x_2) = P(\varepsilon_t \leq x_2) = \Phi\left(\frac{x_2}{\sigma_2}\right),$$

où Φ est la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Ainsi, selon le Théorème de Sklar (1959), il existe une unique copule C telle que

$$H(x_1, x_2) = C\{F_{1,\sigma_1}(x_1), F_{2,\sigma_2}(x_2)\}. \quad (5.1)$$

Pour tenir compte de la dépendance entre deux séries chronologiques autorégressives, on propose le modèle

$$\begin{aligned} X_{t,1} &= \rho_1 X_{t-1,1} + \varepsilon_{t,1}, \\ X_{t,2} &= \rho_2 X_{t-1,2} + \varepsilon_{t,2}. \end{aligned} \quad (5.2)$$

où $(\varepsilon_{t,1}, \varepsilon_{t,2})$ suit une loi dont les marges sont normales et la copule appartient à une certaine famille paramétrique $\mathcal{C} = \{C_\theta; \theta \in \Theta\}$, où $\Theta \subset \mathbb{R}^p$ est l'espace des paramètres de dépendance.

Plus généralement, soient des processus $(X_{t,1})_{t \in \mathbb{Z}}$ et $(X_{t,2})_{t \in \mathbb{Z}}$ qui peuvent s'écrire sous la forme $X_{t,j} = R_{j,\alpha_j}(X_{t-1,j}, \dots, X_{t-p_j,j}) + \varepsilon_{t,j}$, $j = 1, 2$, où $\varepsilon_{t,j} \sim F_{j,\beta_j}$. Ici, α_1 et α_2 sont les paramètres inconnus des modèles de séries chronologiques individuelles R_{1,α_1} et R_{2,α_2} , alors que β_1 et β_2 sont les paramètres des lois des erreurs. Maintenant, en supposant que la structure de dépendance des innovations appartient à la famille \mathcal{C} , alors la loi conjointe de $(\varepsilon_{t,1}, \varepsilon_{t,2})$ s'écrit

$$H(x_1, x_2) = C_\theta\{F_{1,\beta_1}(x_1), F_{2,\beta_2}(x_2)\}, \quad (5.3)$$

où $C_\theta \in \mathcal{C}$. Ainsi, le modèle proposé est

$$\begin{aligned} X_{t,1} &= R_{1,\alpha_1}(X_{t-1,1}, \dots, X_{t-p_1,1}) + \varepsilon_{t,1} \\ X_{t,2} &= R_{2,\alpha_2}(X_{t-1,2}, \dots, X_{t-p_2,2}) + \varepsilon_{t,2}, \end{aligned} \quad (5.4)$$

où le vecteur $(\varepsilon_{t,1}, \varepsilon_{t,2})$ est distribué selon la loi H décrite en (5.3).

Remarque 5.1. *Souvent, on suppose que les lois individuelles des innovations sont normales et de moyenne nulle. On a alors*

$$F_{1,\beta_1}(x_1) = \Phi\left(\frac{x_1}{\beta_1}\right) \quad \text{et} \quad F_{2,\beta_2}(x_2) = \Phi\left(\frac{x_2}{\beta_2}\right).$$

Exemple 5.1. *Soient les processus $X_{t,1} = 0.2X_{t-1,1} + 1.5X_{t-2,1} + 0.5\varepsilon_{t-1,1} + 0.6\varepsilon_{t-4,1} + \varepsilon_{t,1}$ et $X_{t,2} = 0.8\varepsilon_{t-1,2} + \varepsilon_{t,2}$. Ainsi, $X_{t,1}$ est un processus $ARMA(2,4)$ et $X_{t,2}$ est un processus $MA(1)$. On peut montrer que le processus $X_{t,2}$ a une représentation $AR(\infty)$. En effet, puisque $X_{t,2} = (0.8\mathcal{B} + 1)\varepsilon_{t,2}$, on a*

$$\begin{aligned} \varepsilon_{t,2} &= \frac{X_{t,2}}{1 - (-0.8\mathcal{B})} \\ &= \sum_{j=0}^{\infty} (-0.8\mathcal{B})^j X_{t,2} \\ &= X_{t,2} + \sum_{j=1}^{\infty} (-0.8)^j X_{t-j,2}. \end{aligned}$$

Donc,

$$X_{t,2} = - \sum_{j=1}^{\infty} (-0.8)^j X_{t-j,2} + \varepsilon_{t,2}$$

est processus $AR(\infty)$. Ainsi, selon la notation introduite en (5.4), on a

$$R_{2,\alpha_2}(x_{t-1}, x_{t-2}, \dots) = - \sum_{j=1}^{\infty} (-0.8)^j x_{t-j}.$$

Pour ce qui est du processus $X_{t,1}$, une représentation AR est possible mais difficile à obtenir. On peut néanmoins établir que $X_{t,1}$ a une écriture de la forme (5.4).

Remarque 5.2. Pour utiliser notre modèle, il n'est pas nécessaire que les processus soient de la forme AR.

Supposons que l'on veuille simuler des observations à partir du modèle $X_{t,1} = \rho_1 X_{t-1,1} + \varepsilon_{t,1}$ et $X_{t,2} = \rho_2 X_{t-1,2} + \varepsilon_{t,2}$, où la loi conjointe de $(\varepsilon_{t,1}, \varepsilon_{t,2})$ est donnée par

$$H(x_1, x_2) = P(\varepsilon_{t,1} \leq x_1, \varepsilon_{t,2} \leq x_2) = C \left\{ \Phi \left(\frac{x_1}{\beta_1} \right), \Phi \left(\frac{x_2}{\beta_2} \right) \right\}.$$

Ainsi, on suppose que la dépendance des erreurs est C , et leurs lois marginales sont des normales centrées. Fixons un point de départ $(X_{0,1}, X_{0,2})$ quelconque. Ensuite, on pose

$$X_{1,1} = \rho_1 X_{0,1} + \varepsilon_{1,1} \quad \text{et} \quad X_{1,2} = \rho_2 X_{0,2} + \varepsilon_{1,2},$$

où $(\varepsilon_{1,1}, \varepsilon_{1,2}) \sim H$. À remarquer que pour simuler des observations à partir de la loi H , on simule d'abord $(U_{1,1}, U_{1,2}) \sim C$; on pose ensuite

$$\varepsilon_{1,1} = \beta_1 \Phi^{-1}(U_{1,1}) \quad \text{et} \quad \varepsilon_{1,2} = \beta_2 \Phi^{-1}(U_{1,2}).$$

On génère ensuite $(X_{2,1}, X_{2,2})$ à partir de $(X_{1,1}, X_{1,2})$, et ainsi de suite.

Remarque 5.3. En pratique, il est préférable de réchauffer les séries afin d'amenuiser l'influence du choix du point de départ. Ainsi, pour avoir une série de taille T , on génère $T_0 + T$ observations, à savoir

$$(X_{1,1}, X_{1,2}), \dots, (X_{T_0+T,1}, X_{T_0+T,2}),$$

et on conserve seulement les T derniers couples.

5.3 Estimation des paramètres

Selon le modèle (5.4), la loi conjointe de

$$\varepsilon_{t,1} = X_{t,1} - R_{1,\alpha_1}(X_{t-1,1}, \dots, X_{t-p_1,1})$$

et

$$\varepsilon_{t,2} = X_{t,2} - R_{2,\alpha_2}(X_{t-1,2}, \dots, Y_{t-p_2,2})$$

est $H(x_1, x_2) = C_\theta\{F_{1,\beta_1}(x_1), F_{2,\beta_2}(x_2)\}$. Ainsi, leur densité conjointe est

$$h(x_1, x_2) = c_\theta\{F_{1,\beta_1}(x_1), F_{2,\beta_2}(x_2)\} f_{1,\beta_1}(x_1) f_{2,\beta_2}(x_2),$$

où

$$c_\theta(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C_\theta(u_1, u_2)$$

est la densité de C_θ et

$$f_{j,\beta_j}(x_j) = \frac{d}{dx_j} F_{j,\beta_j}(x_j), \quad j = 1, 2,$$

sont les densités marginales. Ainsi, la fonction de vraisemblance basée sur

$(\varepsilon_{1,1}, \varepsilon_{1,2}), \dots, (\varepsilon_{T,1}, \varepsilon_{T,2})$ est

$$L(\alpha_1, \beta_1, \alpha_2, \beta_2, \theta) = L^{(C)}(\alpha_1, \beta_1, \alpha_1, \beta_1, \theta) \times L^{(1)}(\alpha_1, \beta_1) \times L^{(2)}(\alpha_2, \beta_2),$$

où

$$\begin{aligned} & L^{(C)}(\alpha_1, \beta_1, \alpha_2, \beta_2, \theta) \\ &= \prod_{t=p_1 \vee p_2 + 1}^T c_\theta [F_{1,\beta_1}\{X_{t,1} - R_{1,\alpha_1}(X_{t-1,1}, \dots, X_{t-p_1,1})\}, \\ & \quad F_{2,\beta_2}\{X_{t,2} - R_{2,\alpha_2}(X_{t-1,2}, \dots, X_{t-p_2,2})\}]. \end{aligned}$$

$$L^{(1)}(\alpha_1, \beta_1) = \prod_{t=p_1+1}^T f_{1,\alpha_1} \{X_{t,1} - R_{1,\alpha_1}(X_{t-1,1}, \dots, X_{t-p_1,1})\}$$

et

$$L^{(2)}(\alpha_2, \beta_2) = \prod_{t=p_2+1}^T f_{2,\alpha_2} \{X_{t,2} - R_{2,\alpha_2}(X_{t-1,2}, \dots, X_{t-p_2,2})\}.$$

Ici, $a \vee b = \max(a, b)$. Une méthode d'estimation globale consisterait à résoudre directement pour tous les paramètres à la fois. Ainsi,

$$(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\theta}) = \arg \max_{\alpha_1, \beta_1, \alpha_2, \beta_2, \theta} L(\alpha_1, \beta_1, \alpha_2, \beta_2, \theta).$$

Toutefois, cette méthode est, en général, extrêmement lente du point de vue calculatoire. Une façon plus efficace consiste à appliquer une méthode en deux étapes. D'abord, on obtient les estimateurs des séries individuelles par

$$(\hat{\alpha}_1, \hat{\beta}_1) = \arg \max_{\alpha_1, \beta_1} L^{(1)}(\alpha_1, \beta_1)$$

et

$$(\hat{\alpha}_2, \hat{\beta}_2) = \arg \max_{\alpha_2, \beta_2} L^{(2)}(\alpha_2, \beta_2).$$

Ensuite, conditionnellement aux valeurs $\hat{\alpha}_1$, $\hat{\beta}_1$, $\hat{\alpha}_2$ et $\hat{\beta}_2$, on déduit l'estimateur du paramètre de la copule, à savoir θ , par

$$\hat{\theta} = \arg \max_{\theta} L^{(C)}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \theta).$$

Remarque 5.4. Pour le modèle décrit à l'équation (5.2), on a

$$\begin{aligned}
L^{(C)}(\alpha_1, \beta_1, \alpha_2, \beta_2, \theta) &= \prod_{t=2}^T c_\theta \left\{ \Phi \left(\frac{X_{t,1} - \alpha_1 X_{t-1,1}}{\beta_1} \right), \Phi \left(\frac{X_{t,2} - \alpha_2 X_{t-1,2}}{\beta_2} \right) \right\}, \\
L^{(1)}(\alpha_1, \beta_1) &= \prod_{t=2}^T \frac{1}{\beta_1} \phi \left(\frac{X_{t,1} - \alpha_1 X_{t-1,1}}{\beta_1} \right), \\
L^{(2)}(\alpha_2, \beta_2) &= \prod_{t=2}^T \frac{1}{\beta_2} \phi \left(\frac{X_{t,2} - \alpha_2 X_{t-1,2}}{\beta_2} \right),
\end{aligned}$$

où ϕ est la densité de la loi normale standard.

5.4 Procédure d'adéquation pour la copule

Une question intéressante est de savoir si la copule choisie dans le modèle (5.4) est appropriée. Ainsi, on souhaite confronter les hypothèses nulle et alternative

$$\mathcal{H}_0 : C \in \mathcal{C} \quad \text{vs} \quad \mathcal{H}_1 : C \notin \mathcal{C},$$

où $\mathcal{C} = \{C_\theta; \theta \in \Theta\}$ est une famille paramétrique de copules.

On remarque d'abord que dans le modèle proposé, la loi conjointe de

$$U_{t,1} = F_{1,\beta_1}(\varepsilon_{t,1}) \quad \text{et} \quad U_{t,2} = F_{2,\beta_2}(\varepsilon_{t,2})$$

est la copule C_θ . Ainsi, si les variables aléatoires $(U_{t,1}, U_{t,2})$ étaient observables, on pourrait estimer la copule inconnue par

$$C'_T(u_1, u_2) = \frac{1}{T} \sum_{t=1}^T \mathbf{1}(U_{t,1} \leq u_1, U_{t,2} \leq u_2).$$

Malheureusement, cette estimation n'est pas possible puisque

- (i) les erreurs $\varepsilon_{t,1}$ et $\varepsilon_{t,2}$ ne sont pas observables;
- (ii) les lois F_{1,β_1} et F_{2,β_2} ont habituellement des paramètres inconnus que l'on doit estimer.

Ainsi, l'estimation de la copule sera plutôt basée sur les pseudo-observations $(\hat{U}_{t,1}, \hat{U}_{t,2})$, $t \in \{2, \dots, T\}$, où

$$\hat{U}_{t,1} = F_{1,\hat{\beta}_1}(\hat{\varepsilon}_{t,1}) \quad \text{et} \quad \hat{U}_{t,2} = F_{2,\hat{\beta}_2}(\hat{\varepsilon}_{t,2}).$$

Ici, $\hat{\varepsilon}_{t,1}$ et $\hat{\varepsilon}_{t,2}$ sont les résidus tirés du modèle (5.4), c'est-à-dire que

$$\begin{aligned} \hat{\varepsilon}_{t,1} &= X_{t,1} - R_{1,\hat{\alpha}_1}(X_{t-1,1}, \dots, X_{t-p_1,1}) \\ \hat{\varepsilon}_{t,2} &= X_{t,2} - R_{2,\hat{\alpha}_2}(X_{t-1,2}, \dots, X_{t-p_2,2}). \end{aligned}$$

Une estimation semi-paramétrique de C_θ est alors

$$C_T(u_1, u_2) = \frac{1}{T} \sum_{t=1}^T I(\hat{U}_{t,1} \leq u_1, \hat{U}_{t,2} \leq u_2).$$

Remarque 5.5. Dans le cas où les erreurs sont supposées normales, on a

$$\hat{U}_{t,1} = \Phi\left(\frac{\hat{\varepsilon}_{t,1}}{\hat{\beta}_1}\right) \quad \text{et} \quad \hat{U}_{t,2} = \Phi\left(\frac{\hat{\varepsilon}_{t,2}}{\hat{\beta}_2}\right).$$

Pour vérifier si le modèle de copule appartient vraisemblablement à \mathcal{C} , on comparera C_T avec $C_{\hat{\theta}}$, où $\hat{\theta}$ est l'estimateur de θ présenté à la section précédente. La statistique de test proposée sera basée sur le processus empirique

$$S'_T(u_1, u_2) = \sqrt{T} \{C_T(u_1, u_2) - C_{\hat{\theta}}(u_1, u_2)\}.$$

La convergence de ce processus dans un espace de fonctions appropriées est un problème compliqué. Ici, on se contentera d'énoncer sa convergence en loi comme une conjecture.

Conjecture 5.1. *Sous des conditions appropriées, le processus empirique $S'_T(u_1, u_2)$ converge vers un certain processus Gaussien $S(u_1, u_2)$.*

Une statistique possible pour vérifier l'adéquation à la famille \mathcal{C} est la fonctionnelle de Cramér-von Mises, c'est-à-dire

$$W'_T = \int_0^1 \int_0^1 \{S'_T(u_1, u_2)\}^2 du_1 du_2.$$

De la Conjecture 5.1, une application du Théorème sur les applications continues (*Continuous Mapping Theorem*) permet de déduire que

$$W'_T \rightsquigarrow W = \int_0^1 \int_0^1 \{S(u_1, u_2)\}^2 du_1 du_2,$$

où le symbole \rightsquigarrow dénote la convergence en loi. Pour plus de détails, voir van der Vaart & Wellner (1996).

Il est généralement impossible d'avoir une expression explicite pour W'_T , ce qui complique son implémentation. Une façon de contourner ce problème consiste à remplacer $C_{\hat{\theta}}$ par une copule empirique C_N calculée à partir d'observations $(U_{1,1}, U_{1,2}), \dots, (U_{N,1}, U_{N,2})$ générées artificiellement à partir de la loi $C_{\hat{\theta}}$. Ainsi, $C_{\hat{\theta}}$ est remplacée par

$$C_N(u_1, u_2) = \frac{1}{N} \sum_{t=1}^N \mathbf{1}(U_{t,1} \leq u_1, U_{t,2} \leq u_2).$$

La statistique qui sera utilisée dans la suite sera donc

$$W_T = \int_0^1 \int_0^1 \{S_{T,N}(u_1, u_2)\}^2 du_1 du_2,$$

où

$$S_{T,N}(u_1, u_2) = \sqrt{T} \{C_T(u_1, u_2) - C_N(u_1, u_2)\}.$$

Proposition 5.1. *On a*

$$\begin{aligned} W_T = & \frac{1}{T} \sum_{s=1}^T \sum_{t=1}^T \left(1 - \widehat{U}_{s,1} \vee \widehat{U}_{t,1}\right) \left(1 - \widehat{U}_{s,2} \vee \widehat{U}_{t,2}\right) \\ & - \frac{2}{N} \sum_{s=1}^N \sum_{t=1}^T \left(1 - U_{s,1} \vee \widehat{U}_{t,1}\right) \left(1 - U_{s,2} \vee \widehat{U}_{t,2}\right) \\ & + \frac{n}{N^2} \sum_{s=1}^N \sum_{t=1}^N (1 - U_{s,1} \vee U_{t,1}) (1 - U_{s,2} \vee U_{t,2}), \end{aligned}$$

où $a \vee b = \max(a, b)$.

Démonstration. En développant le carré, on a

$$\begin{aligned} & \{C_T(u_1, u_2) - C_N(u_1, u_2)\}^2 \\ &= \left\{ \frac{1}{T} \sum_{s=1}^T \mathbf{1}(\widehat{U}_{s,1} \leq u_1, \widehat{U}_{s,2} \leq u_2) - \frac{1}{N} \sum_{s=1}^N \mathbf{1}(U_{s,1} \leq u_1, U_{s,2} \leq u_2) \right\}^2 \\ &= \left\{ \frac{1}{T} \sum_{s=1}^T \mathbf{1}(\widehat{U}_{s,1} \leq u_1, \widehat{U}_{s,2} \leq u_2) \right\} \left\{ \frac{1}{T} \sum_{t=1}^T \mathbf{1}(\widehat{U}_{t,1} \leq u_1, \widehat{U}_{t,2} \leq u_2) \right\} \\ &\quad - 2 \left\{ \frac{1}{T} \sum_{s=1}^T \mathbf{1}(\widehat{U}_{s,1} \leq u_1, \widehat{U}_{s,2} \leq u_2) \right\} \left\{ \frac{1}{N} \sum_{t=1}^N \mathbf{1}(U_{t,1} \leq u_1, U_{t,2} \leq u_2) \right\} \\ &\quad + \left\{ \frac{1}{N} \sum_{s=1}^N \mathbf{1}(U_{s,1} \leq u_1, U_{s,2} \leq u_2) \right\} \left\{ \frac{1}{N} \sum_{t=1}^N \mathbf{1}(U_{t,1} \leq u_1, U_{t,2} \leq u_2) \right\} \\ &= \frac{1}{T^2} \sum_{s=1}^T \sum_{t=1}^T \mathbf{1}(\widehat{U}_{s,1} \vee \widehat{U}_{t,1} \leq u_1, \widehat{U}_{s,2} \vee \widehat{U}_{t,2} \leq u_2) \\ &\quad - \frac{2}{TN} \sum_{s=1}^T \sum_{t=1}^N \mathbf{1}(\widehat{U}_{s,1} \vee U_{t,1} \leq u_1, \widehat{U}_{s,2} \vee U_{t,2} \leq u_2) \\ &\quad + \frac{1}{N^2} \sum_{s=1}^N \sum_{t=1}^N \mathbf{1}(U_{s,1} \vee U_{t,1} \leq u_1, U_{s,2} \vee U_{t,2} \leq u_2). \end{aligned}$$

Ainsi,

$$\begin{aligned}
W_T &= T \int_0^1 \int_0^1 \{C_n(u_1, u_2) - C_N(u_1, u_2)\}^2 du_1 du_2 \\
&= \frac{1}{T} \sum_{s=1}^T \sum_{t=1}^T \int_{\widehat{U}_{s,1} \vee \widehat{U}_{t,1}}^1 \int_{\widehat{U}_{s,2} \vee \widehat{U}_{t,2}}^1 du_1 du_2 \\
&\quad - \frac{2}{N} \sum_{s=1}^T \sum_{t=1}^N \int_{U_{s,1} \vee U_{t,1}}^1 \int_{U_{s,2} \vee U_{t,2}}^1 du_1 du_2 \\
&\quad + \frac{T}{N^2} \sum_{s=1}^N \sum_{t=1}^N \int_{U_{s,1} \vee U_{t,1}}^1 \int_{U_{s,2} \vee U_{t,2}}^1 du_1 du_2 \\
&= \frac{1}{T} \sum_{s=1}^T \sum_{t=1}^T (1 - \widehat{U}_{s,1} \vee \widehat{U}_{t,1}) (1 - \widehat{U}_{s,2} \vee \widehat{U}_{t,2}) \\
&\quad - \frac{2}{N} \sum_{s=1}^N \sum_{t=1}^T (1 - U_{s,1} \vee \widehat{U}_{t,1}) (1 - U_{s,2} \vee \widehat{U}_{t,2}) \\
&\quad + \frac{T}{N^2} \sum_{s=1}^N \sum_{t=1}^N (1 - U_{s,1} \vee U_{t,1}) (1 - U_{s,2} \vee U_{t,2}),
\end{aligned}$$

tel qu'annoncé. \diamond

Sous l'hypothèse nulle, $C_T \rightarrow C_\theta$ et $C_N \rightarrow C_\theta$ presque sûrement. Ainsi, $W_T \rightarrow 0$ presque sûrement lorsque la vraie copule appartient à la famille \mathcal{C} . À l'opposé, si C n'appartient pas à \mathcal{C} , alors on a $C = D \notin \mathcal{C}$. Dans ce cas, pour un certain $\theta \in \Theta$, on a

$$\frac{W_T}{T} \rightarrow \int_0^1 \int_0^1 \{D(u_1, u_2) - C_\theta(u_1, u_2)\}^2 du_1 du_2 > 0.$$

Par conséquent, un test basé sur la statistique W_T sera convergent, c'est-à-dire que pour de grandes tailles d'échantillons, il rejettera toujours l'hypothèse nulle si $C \notin \mathcal{C}$.

Pour élaborer un test formel basé sur W_T , il faut trouver une valeur critique

q_α telle que

$$P(W_T > q_\alpha | \mathcal{H}_0 \text{ est vraie}) = \alpha.$$

Deux aspects empêchent d'obtenir q_α directement :

- La loi de W_T a une forme extrêmement compliquée;
- Le paramètre θ est inconnu.

Pour pallier ces problèmes, on utilisera une méthode de *bootstrap* paramétrique. Cette technique, qui sera décrite dans la suite, a été utilisée avec succès par Genest et al. (2006) et par Mesfioui et al. (2009) pour tester l'adéquation de copules. L'idée consiste à générer artificiellement des échantillons à partir de la loi C_θ et de calculer la statistique de test à chaque fois. On obtient ainsi $W_T^{(1)}, \dots, W_T^{(M)}$ qui, sous des conditions appropriées, sont des copies indépendantes, asymptotiquement, de W_T .

Conjecture 5.2. *Sous des conditions appropriées,*

$$(W_T, W_T^{(1)}, \dots, W_T^{(M)}) \rightsquigarrow (W, W^{(1)}, \dots, W^{(M)}),$$

où $W^{(1)}, \dots, W^{(M)}$ sont des copies indépendantes de W .

Pour montrer formellement que la Conjecture 5.2 est vraie, il faudrait utiliser des arguments semblables à ceux de Genest & Rémillard (2008), mais appliqués dans un contexte de résidus de séries chronologiques. La complexité de ce problème est au-delà des objectifs de ce mémoire.

La Conjecture 5.2 permet de calculer une valeur critique asymptotiquement valide, à savoir \hat{q}_α telle que

$$\frac{1}{M} \sum_{j=1}^M \mathbf{1}(W_T^{(j)} > \hat{q}_\alpha) \approx \alpha.$$

Autrement dit, \hat{q}_α est le $(1 - \alpha)$ -ème percentile de $W_T^{(1)}, \dots, W_T^{(M)}$. De façon équivalente, une p -valeur approximative du test est donnée par

$$\hat{p} = \frac{1}{M} \sum_{j=1}^M \mathbf{1} \left(W_T^{(j)} > W_T \right).$$

5.5 Prédiction

Pour le modèle de séries chronologiques bivarié à base de copules dont les innovations sont normales, la loi conditionnelle de $(X_{t,1}, X_{t,2})$ étant donné $(X_{t-1,1}, X_{t-1,2})$ est

$$H(x_1, x_2) = C_\theta \left\{ \Phi \left(\frac{x_1 - \alpha_1 X_{t-1,1}}{\beta_1} \right), \Phi \left(\frac{x_2 - \alpha_2 X_{t-1,2}}{\beta_2} \right) \right\}.$$

De là, on peut montrer que la loi conditionnelle de $X_{T+1,1}$ étant donné $X_{T+1,2} = x_2$ est

$$P(X_{T+1,1} \leq x_1 | X_{T+1,2} = x_2) = C_{\theta,01} \left\{ \Phi \left(\frac{x_1 - \alpha_1 X_{T,1}}{\beta_1} \right), \Phi \left(\frac{x_2 - \alpha_2 X_{T,2}}{\beta_2} \right) \right\},$$

où $C_{\theta,01}(u, v) = \partial C_\theta(u, v) / \partial v$. La densité conditionnelle associée est donc

$$c_\theta \left\{ \Phi \left(\frac{x_1 - \alpha_1 X_{T,1}}{\beta_1} \right), \Phi \left(\frac{x_2 - \alpha_2 X_{T,2}}{\beta_2} \right) \right\} \frac{1}{\beta_1} \phi \left(\frac{x_1 - \alpha_1 X_{T,1}}{\beta_1} \right),$$

où c_θ est la densité associée à C_θ . Ainsi, pour prédire la valeur de $X_{T+1,1}$, on peut considérer l'espérance conditionnelle $E(X_{T+1,1} | X_{T+1,2} = x_2)$ en remplaçant les paramètres inconnus par leurs estimations. Ainsi, la prédiction de $\hat{X}_{T+1,1}$ est donnée par

$$\begin{aligned}
\hat{X}_{T+1,1} &= E\{X_{T+1,1}|X_{T+1,2}\} \\
&= \int_0^1 x_1 c_{\hat{\theta}} \left\{ \Phi \left(\frac{x_1 - \hat{\alpha}_1 X_{T,1}}{\hat{\beta}_1} \right), \Phi \left(\frac{X_{T+1,2} - \hat{\alpha}_2 X_{T,2}}{\hat{\beta}_2} \right) \right\} \\
&\quad \times \frac{1}{\hat{\beta}_1} \phi \left(\frac{x_1 - \hat{\alpha}_1 X_{T,1}}{\hat{\beta}_1} \right) dx_1. \tag{5.5}
\end{aligned}$$

Cette expression est généralement difficile à calculer; il est alors nécessaire d'utiliser un logiciel de calcul symbolique comme Maple pour obtenir des prévisions.

Remarque 5.6. Dans le cas particulier de l'indépendance, c'est-à-dire lorsque $c_{\theta}(u, v) = 1$, on a

$$\hat{X}_{T+1,1} = \int_0^1 x_1 \frac{1}{\hat{\beta}_1} \phi \left(\frac{x_1 - \hat{\alpha}_1 X_{T,1}}{\hat{\beta}_1} \right) dx_1 = \hat{\alpha}_1 X_{T,1}.$$

Cette prévision est celle qu'on obtiendrait d'une modélisation classique à partir du modèle autorégressif d'ordre 1.

5.6 Extension multivariée

Pour $j \in \{1, \dots, d\}$, soient des séries chronologiques

$$X_{t,j} = R_{j,\alpha_j} (X_{t-1,j}, \dots, X_{t-p_j,j}) + \varepsilon_{t,j}.$$

Afin de tenir compte de la possible dépendance entre ces séries, on suppose que pour chaque $t \in \mathbb{Z}$, $(\varepsilon_{t,1}, \dots, \varepsilon_{t,d}) \sim H$, où

$$H(x_1, \dots, x_d) = C_{\theta} \{F_{1,\beta_1}(x_1), \dots, F_{d,\beta_d}(x_d)\}.$$

Ici, $C_\theta : [0, 1]^d \rightarrow [0, 1]$ est une copule à d dimensions et $F_{1,\beta_1}, \dots, F_{d,\beta_d}$ sont les marges de H . Par des arguments similaires au cas $d = 2$, la densité conjointe de $(\varepsilon_{t,1}, \dots, \varepsilon_{t,d})$ est

$$h(x_1, \dots, x_d) = c_\theta \{F_{1,\beta_1}(x_1), \dots, F_{d,\beta_d}(x_d)\} f_{1,\beta_1}(x_1) \times \dots \times f_{d,\beta_d}(x_d),$$

où

$$c_\theta(u_1, \dots, u_d) = \frac{\partial^d}{\partial u_1 \dots \partial u_d} C_\theta(u_1, \dots, u_d) \quad \text{et} \quad f_{j,\beta_j}(x_j) = \frac{d}{dx_j} F_{j,\beta_j}(x_j).$$

La vraisemblance basée sur $(\varepsilon_{1,1}, \dots, \varepsilon_{1,d}), \dots, (\varepsilon_{T,1}, \dots, \varepsilon_{T,d})$ est donc

$$\begin{aligned} L(\alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d, \theta) \\ = L^{(C)}(\alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d, \theta) \times \prod_{j=1}^d L^{(j)}(\alpha_j, \beta_j), \end{aligned}$$

où $L^{(C)}(\alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d, \theta)$

$$\begin{aligned} = \prod_{t=p_1 \vee \dots \vee p_d + 1}^T c_\theta [F_{1,\beta_1} \{X_{t,1} - R_{1,\alpha_1}(X_{t-1,1}, \dots, X_{t-p_1,1})\}, \dots, \\ F_{d,\beta_d} \{X_{t,d} - R_{d,\alpha_d}(X_{t-1,d}, \dots, X_{t-p_d,d})\}] \end{aligned}$$

et pour $j \in \{1, \dots, d\}$,

$$L^{(j)}(\alpha_j, \beta_j) = \prod_{t=p_j+1}^T f_{j,\beta_j} \{X_{t,j} - R_{j,\alpha_j}(X_{t-1,j}, \dots, X_{t-p_j,j})\}$$

sont les vraisemblances marginales. Une procédure d'estimation en deux étapes consiste d'abord à estimer les paramètres des modèles marginaux par

$$(\hat{\alpha}_j, \hat{\beta}_j) = \arg \max_{\alpha_j, \beta_j} L^{(j)}(\alpha_j, \beta_j), \quad j \in \{1, \dots, d\}.$$

Ensuite, le paramètre de la copule est estimé par

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L^{(C)} \left(\hat{\alpha}_1, \dots, \hat{\alpha}_d, \hat{\beta}_1, \dots, \hat{\beta}_d, \theta \right).$$

Afin de construire une procédure d'adéquation, on remarque que la loi conjointe de $(U_{t,1}, \dots, U_{t,d})$, où pour $j \in \{1, \dots, d\}$, $U_{t,j} = F_{j,\beta_j}(\varepsilon_{t,j})$, est la copule C_θ . Ces vecteurs aléatoires n'étant pas observables, l'estimation de la copule sera plutôt basée sur les vecteurs de pseudo-observations

$$\left(\hat{U}_{1,1}, \dots, \hat{U}_{1,d} \right), \dots, \left(\hat{U}_{T,1}, \dots, \hat{U}_{T,d} \right),$$

où $\hat{U}_{t,j} = F_{j,\hat{\beta}_j}(\hat{\varepsilon}_{t,j})$ et

$$\hat{\varepsilon}_{t,j} = X_{t,j} - R_{j,\hat{\alpha}_j}(X_{t-1,j}, \dots, X_{t-p_j,j}).$$

Une estimation semi-paramétrique de C_θ est alors

$$C_T(u_1, \dots, u_d) = \frac{1}{T} \sum_{T=1}^n \mathbf{1} \left(\hat{U}_{t,1} \leq u_1, \dots, \hat{U}_{t,d} \leq u_d \right).$$

La statistique pour vérifier l'adéquation à la famille \mathcal{C} sera la fonctionnelle de Cramér-von Mises

$$W_{T,d} = \int_0^1 \dots \int_0^1 \{S_{T,N}(u_1, \dots, u_d)\}^2 du_1, \dots, du_d,$$

où

$$S_{T,N}(u_1, \dots, u_d) = \sqrt{T} \{C_T(u_1, \dots, u_d) - C_N(u_1, \dots, u_d)\}$$

et C_N est la copule empirique construite à partir d'un échantillon de taille N tiré de la loi $C_{\hat{\theta}}$. Similairement au cas $d = 2$, on obtient une expression explicite pour la statistique de test $W_{T,d}$. C'est le sujet de la proposition suivante, dont la preuve est similaire à celle de la Proposition 5.1.

Proposition 5.2. *On a*

$$\begin{aligned} W_{T,d} = & \frac{1}{T} \sum_{s=1}^T \sum_{t=1}^T \prod_{j=1}^d \left(1 - \widehat{U}_{s,j} \vee \widehat{U}_{t,j} \right) \\ & - \frac{2}{N} \sum_{s=1}^N \sum_{t=1}^T \prod_{j=1}^d \left(1 - U_{s,j} \vee \widehat{U}_{t,j} \right) \\ & + \frac{T}{N^2} \sum_{s=1}^N \sum_{t=1}^N \prod_{j=1}^d \left(1 - U_{s,j} \vee U_{t,j} \right). \end{aligned}$$

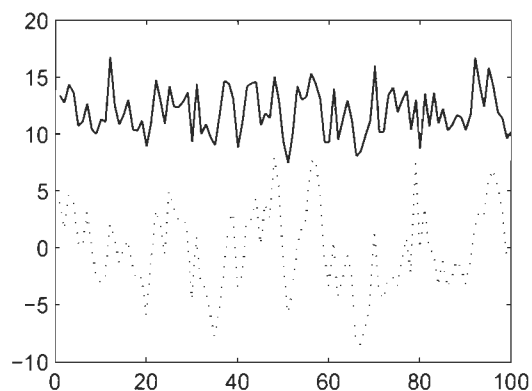
5.7 Illustration sur des données simulées

Afin d'illustrer la méthode de sélection de copule, soit le processus $(X_{t,1}, X_{t,2})$, $t \in \mathbb{N}$, où $X_{t,1} = 0,3X_{t-1,1} + \varepsilon_{t,1}$ et $X_{t,2} = 0,7X_{t-1,2} + \varepsilon_{t,2}$. Ici, $\varepsilon_{t,1} \sim \mathcal{N}(0, 4)$, $\varepsilon_{t,2} \sim \mathcal{N}(0, 9)$, et la dépendance entre $\varepsilon_{t,1}$ et $\varepsilon_{t,2}$ est caractérisée par la copule de Clayton de paramètre $\theta = 6$. La loi conjointe de $(\varepsilon_{t,1}, \varepsilon_{t,2})$ est donc

$$H(x_1, x_2) = C_6^{\text{CL}} \left\{ \Phi \left(\frac{x_1}{2} \right), \Phi \left(\frac{x_2}{3} \right) \right\}.$$

À la figure 5.1, on peut voir une réalisation de taille $T = 100$ de ce processus. La série en traits continus correspond à $X_{t,1}$, alors que la série en traits discontinus correspond à $X_{t,2}$. À noter que la série $X_{t,1}$ a subi une translation pour des fins de présentation.

D'abord, il s'agit d'estimer les paramètres des séries marginales sous l'hypothèse que celles-ci sont des réalisations de processus AR(1). On obtient $\widehat{\alpha}_1 = 0,2880$, $\widehat{\beta}_1 = 1,9320$, $\widehat{\alpha}_2 = 0,6976$ et $\widehat{\beta}_2 = 2,9763$. Ensuite, des tests d'adéquation pour six modèles de copules ont été effectués. Au Tableau 5.1, on retrouve la valeur de la statistique W_T , ainsi que la valeur critique obtenue

Figure 5.1: Deux séries AR(1) de dépendance Clayton de paramètre $\theta = 6$

par la méthode du *bootstrap* paramétrique avec $M = 1\,000$ répétitions. On retrouve, de plus, l'estimation du paramètre pour chaque modèle. À noter que le tau de Kendall des résidus vaut 0.7737. On constate que seule la copule de Clayton est acceptée au seuil de signification $\alpha = 0.05$. On voit donc que notre méthode de sélection de copule a identifié le bon modèle.

Tableau 5.1: Résultats des tests d'adéquation pour les données simulées

Famille de copules	W_T	Valeur critique	$\hat{\theta}$
Ali-Mikhail-Haq	2,4931	0.0394	1,0000
Farlie-Gumbel-Morgenstern	4,2065	0.0470	1,0000
Clayton	0,0037	0.0112	6,8361
Gumbel-Hougaard	0,0978	0.0109	0,7737
Frank	0,0511	0.0109	15,8461
Normale	0,0526	0.0103	0,9375

CHAPITRE 6

ÉTUDES DE SIMULATION

Dans ce chapitre, les propriétés à tailles finies des procédures développées au chapitre 5 seront étudiées à l'aide de simulations. Toutes les méthodes ont été implantées à l'aide du logiciel Matlab, exploité sous Windows. Les programmes sont présentés dans leur intégralité à l'annexe A.

6.1 Performance des estimateurs

6.1.1 Estimateurs MV du processus AR(1)

Pour valider la méthode d'estimation en deux étapes présentée au Chapitre 5, on a d'abord voulu vérifier si les paramètres α et β du processus auto-régressif $X_t = \alpha X_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, \beta^2)$, sont bien estimés par la méthode du maximum de vraisemblance. Pour ce faire, l'erreur quadratique moyenne (EQM) de chacun des paramètres a été estimée à partir de N échantillons de taille T générés du modèle (5.2). On rappelle que si $\hat{\theta}_T$ est un estimateur de θ , alors

l'erreur quadratique moyenne est donnée par

$$\text{EQM}(\hat{\theta}_T) = \text{E} \left\{ \left(\hat{\theta}_T - \theta \right)^2 \right\}.$$

Quand une expression explicite pour $\text{EQM}(\hat{\theta}_T)$ n'est pas disponible, comme c'est le cas ici, on l'estime par

$$\widehat{\text{EQM}}(\hat{\theta}_T) = \frac{1}{N} \sum_{i=1}^N \left(\hat{\theta}_T^{(i)} - \theta \right)^2,$$

où $\theta_T^{(1)}, \dots, \theta_T^{(N)}$ sont les estimateurs obtenus des différentes simulations. Dans notre cas, le nombre d'échantillons générés a été fixé à $N = 10\,000$ et les tailles d'échantillons considérées furent $T = 25, 50, 100$. Les résultats sont présentés au Tableau 6.1.

Tableau 6.1: Erreurs quadratiques moyennes ($\times 10^2$) des estimateurs à maximum de vraisemblance des paramètres du modèle autorégressif d'ordre un $X_t = \alpha X_{t-1} + \varepsilon_t$, où $\varepsilon_t \sim \mathcal{N}(0, \beta^2)$, estimées à l'aide de 10 000 répétitions.

α	β^2	$T = 25$		$T = 50$		$T = 100$	
		$\text{EQM}(\hat{\alpha})$	$\text{EQM}(\hat{\beta}^2)$	$\text{EQM}(\hat{\alpha})$	$\text{EQM}(\hat{\beta}^2)$	$\text{EQM}(\hat{\alpha})$	$\text{EQM}(\hat{\beta}^2)$
1/4	1	3,79	2,15	1,86	1,03	0,91	0,51
1/2		3,27	2,13	1,61	1,01	0,78	0,52
3/4		2,58	2,15	1,08	1,05	0,50	0,50
1/4	2	3,69	8,60	1,86	4,12	0,92	2,02
1/2		3,34	8,38	1,58	4,14	0,76	2,06
3/4		2,55	8,61	1,11	4,10	0,49	2,05

À la vue des résultats du Tableau 6.1, on constate que la méthode d'estimation des paramètres α et β est généralement efficace. En effet, les erreurs quadratiques moyennes de $\hat{\alpha}$ et $\hat{\beta}$ sont petites dans tous les cas considérés. On

remarque aussi que les estimateurs deviennent davantage précis à mesure que T augmente. Également, on constate que la valeur du paramètre autorégressif α ne semble pas exercer une grande influence sur la précision de $\hat{\alpha}$. Ce n'est pas le cas de la variance : en effet, les EQM de $\hat{\beta}$ sont beaucoup plus grandes pour $\beta = 2$ que pour $\beta = 1$.

6.1.2 Estimation du paramètre de dépendance de la copule

Pour vérifier si le paramètre de dépendance est bien estimé par la méthode à deux niveaux présentée au Chapitre 5, on a supposé que les séries individuelles étaient des réalisations de deux processus autorégressifs d'ordre un, à savoir $X_{t,1} = \alpha_1 X_{t-1,1} + \varepsilon_{t,1}$ et $X_{t,2} = \alpha_2 X_{t-1,2} + \varepsilon_{t,2}$, où la loi conjointe de $(\varepsilon_{t,1}, \varepsilon_{t,2})$ possède des marges $\mathcal{N}(0, \beta_1^2)$ et $\mathcal{N}(0, \beta_2^2)$, et une fonction de dépendance représentée par une copule paramétrique C_θ . Une fois que les paramètres α_1 , β_1 , α_2 et β_2 sont estimés par la méthode du maximum de vraisemblance, la deuxième étape consiste à estimer θ *conditionnellement* aux valeurs de $\hat{\alpha}_1$, $\hat{\beta}_1$, $\hat{\alpha}_2$ et $\hat{\beta}_2$ obtenues dans une première étape.

Au Chapitre 5, on montre comment estimer θ par la méthode du maximum de vraisemblance. Cet estimateur est noté $\hat{\theta}_{MV}$. Pour des fins de comparaison, on a aussi considéré l'estimation de θ par une inversion du tau de Kendall, qui peut s'écrire en fonction de θ par la formule

$$\tau(\theta) = 4 \int_{(0,1)^2} C_\theta(u_1, u_2) dC_\theta(u_1, u_2).$$

Cet estimateur, noté $\hat{\theta}_K$, est tel que $\tau(\hat{\theta}_K) = \tau_n$, où τ_n est le tau de Kendall

empirique. C'est, en quelque sorte, une version non-paramétrique de la méthode des moments classique. Dans le cas présent, τ_n est calculé à partir des résidus des modèles autorégressifs d'ordre un; sa valeur sera donc conditionnelle à $\hat{\alpha}_1$, $\hat{\beta}_1$, $\hat{\alpha}_2$ et $\hat{\beta}_2$ obtenues à la première étape d'estimation.

Les Tableaux 6.2 à 6.7 présentent les EQM estimées de ces deux estimateurs pour le paramètre de dépendance des copules de Ali-Mikhail-Haq, Farlie-Gumbel-Morgenstern, Clayton, Gumbel-Hougaard, Frank et Normale, respectivement. Le nombre de répétitions pour estimer les EQM a été fixé à $N = 10\,000$ et les tailles d'échantillons considérées sont $T = 25, 50, 100$. Pour uniformiser la divulgation des résultats, ceux-ci sont présentés en fonction de $\tau(\theta)$, plutôt qu'en fonction de θ . Les paramètres des séries individuelles sont tels que $\alpha_1 = \alpha_2 = \alpha$ et $\beta_1^2 = 1$, $\beta_2^2 = 2$.

Tableau 6.2: Erreurs quadratiques moyennes ($\times 10^2$) de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Ali-Mikhail-Haq

τ	α	$\hat{\theta}_{MV}$			$\hat{\theta}_K$		
		$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
0.10	1/4	46,05	49,17	51,33	11,74	7,90	4,89
	1/2	45,20	47,97	50,36	11,73	7,92	4,85
	3/4	43,68	45,66	47,90	11,59	7,84	4,87
0.15	1/4	25,80	27,69	28,29	12,14	7,95	4,3
	1/2	25,35	27,06	28,05	12,12	7,95	4,27
	3/4	24,42	25,86	27,03	12,11	7,85	4,27
0.20	1/4	12,69	13,41	13,09	11,48	6,63	3,32
	1/2	12,53	13,31	13,30	11,44	6,65	3,26
	3/4	12,10	12,93	13,35	11,45	6,37	3,24

Tableau 6.3: Erreurs quadratiques moyennes ($\times 10^2$) de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Farlie–Gumbel–Morgenstern

τ	α	$\hat{\theta}_{MV}$			$\hat{\theta}_K$		
		$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
0.10	1/4	35,13	15,87	8,44	29,33	15,96	8,60
	1/2	30,86	15,90	8,24	29,22	16,60	8,91
	3/4	32,57	16,69	8,63	30,20	16,20	8,73
0.15	1/4	25,63	12,16	6,88	25,63	13,21	7,03
	1/2	25,14	12,62	7,06	25,33	13,17	7,09
	3/4	24,90	13,32	7,14	25,21	12,87	7,01
0.20	1/4	20,09	9,39	4,31	21,50	9,67	4,65
	1/2	19,52	9,04	4,38	21,82	9,45	4,56
	3/4	20,18	9,34	4,77	21,32	9,70	4,68

Tableau 6.4: Erreurs quadratiques moyennes ($\times 10^2$) de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Clayton

τ	α	$\hat{\theta}_{MV}$			$\hat{\theta}_K$		
		$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
0.25	1/4	17,27	7,86	3,58	30,74	12,77	5,73
	1/2	17,13	7,79	3,66	30,63	12,56	5,69
	3/4	21,09	8,95	3,84	31,24	12,68	5,87
0.50	1/4	51,03	24,43	11,52	116,49	43,83	19,86
	1/2	51,26	24,88	11,81	110,38	44,84	19,52
	3/4	50,55	24,37	11,92	112,98	44,10	19,75
0.75	1/4	369,46	193,07	92,38	745,22	270,52	121,06
	1/2	390,92	196,64	92,07	755,67	256,61	116,82
	3/4	397,57	201,23	94,81	746,09	266,88	114,78

Tableau 6.5: Erreurs quadratiques moyennes ($\times 10^2$) de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Gumbel–Hougaard

τ	α	$\hat{\theta}_{MV}$			$\hat{\theta}_K$		
		$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
0.25	1/4	1,94	0,97	0,49	1,88	0,92	0,44
	1/2	2,37	1,30	0,66	1,85	0,91	0,45
	3/4	3,41	2,13	1,27	1,84	0,91	0,45
0.50	1/4	1,43	0,65	0,32	1,53	0,64	0,30
	1/2	1,78	0,85	0,43	1,43	0,63	0,30
	3/4	2,76	1,51	0,81	1,43	0,63	0,30
0.75	1/4	0,52	0,23	0,11	0,58	0,24	0,10
	1/2	0,66	0,30	0,14	0,57	0,24	0,10
	3/4	1,07	0,51	0,27	0,57	0,23	0,10

Tableau 6.6: Erreurs quadratiques moyennes ($\times 10^2$) de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule de Frank

τ	α	$\hat{\theta}_{MV}$			$\hat{\theta}_K$		
		$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
0.25	1/4	275,68	118,87	59,63	227,00	99,41	45,32
	1/2	356,11	159,88	79,98	228,33	98,82	45,22
	3/4	528,41	280,53	154,55	226,07	100,95	43,18
0.50	1/4	532,00	222,46	103,67	432,55	170,99	77,06
	1/2	695,07	294,46	146,74	434,93	172,00	78,27
	3/4	1220,66	574,23	294,25	444,90	171,57	78,10
0.75	1/4	2307,87	817,65	365,47	1912,74	659,78	280,27
	1/2	2846,48	1118,02	580,31	1897,53	652,96	270,59
	3/4	4951,58	2195,21	1196,06	1872,19	663,45	274,59

Tableau 6.7: Erreurs quadratiques moyennes ($\times 10^2$) de $\hat{\theta}_{MV}$ et de $\hat{\theta}_K$ pour deux processus AR(1) de copule Normale

τ	α	$\hat{\theta}_{MV}$			$\hat{\theta}_K$		
		$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
0.25	1/4	3,26	1,52	0,78	4,01	1,78	0,87
	1/2	3,30	1,58	0,75	3,97	1,82	0,86
	3/4	3,30	1,58	0,76	4,01	1,81	0,88
0.50	1/4	1,39	0,59	0,27	1,78	0,72	0,33
	1/2	1,40	0,60	0,27	1,82	0,71	0,33
	3/4	1,38	0,59	0,28	1,75	0,74	0,33
0.75	1/4	0,15	0,06	0,02	0,23	0,08	0,03
	1/2	0,16	0,06	0,02	0,23	0,08	0,03
	3/4	0,17	0,06	0,02	0,23	0,08	0,03

Selon les résultats des Tableaux 6.2 à 6.7, on voit que les estimateurs du paramètre de dépendance performant bien pour chacune des copules considérées. De plus, ces estimateurs deviennent davantage précis à mesure que la taille des échantillons augmente. Également, on voit que la performance des estimateurs a tendance à diminuer lorsque le niveau de dépendance est élevé. Voici maintenant quelques remarques pour chacun des six modèles :

- (i) Ali-Mikhail-Haq. L'estimateur $\hat{\theta}_K$ est bien meilleur que $\hat{\theta}_{MV}$;
- (ii) Farlie-Gumbel-Morgenstern. Les estimateurs $\hat{\theta}_K$ et $\hat{\theta}_{MV}$ sont équivalents;
- (iii) Clayton. L'estimateur $\hat{\theta}_{MV}$ est bien meilleur que $\hat{\theta}_K$;
- (iv) Gumbel-Hougaard. Les EQM sont généralement très petites. L'estimateur $\hat{\theta}_K$ est meilleur que $\hat{\theta}_{MV}$. L'estimateur $\hat{\theta}_{MV}$ varie beaucoup selon les valeurs de α ; ce n'est pas le cas de $\hat{\theta}_K$;

- (v) Frank. Les EQM sont élevées, surtout lorsque la valeur de τ est élevée. L'estimateur $\hat{\theta}_K$ est bien meilleur que $\hat{\theta}_{MV}$, surtout quand la dépendance est forte. L'estimateur $\hat{\theta}_{MV}$ varie beaucoup selon les valeurs de α ; ce n'est pas le cas de $\hat{\theta}_K$;
- (vi) Normale. L'estimateur $\hat{\theta}_{MV}$ est meilleur que $\hat{\theta}_K$. Contrairement aux autres modèles, les EQM diminuent lorsque τ augmente.

En résumé, les deux méthodes d'estimation sont efficaces. La supériorité de l'une par rapport à l'autre dépend fortement du modèle considéré. Ainsi, on choisirait

- $\hat{\theta}_K$ pour les copules de Ali–Mikhail–Haq, Gumbel–Hougaard et Frank;
- $\hat{\theta}_{MV}$ pour les copules de Clayton et Normale.

Dans le doute, on pourrait favoriser $\hat{\theta}_K$ puisque cet estimateur est plus stable vis-à-vis des paramètres des séries marginales.

6.2 Performance de la méthode d'adéquation

6.2.1 Plan de simulation

Pour évaluer la performance de la méthode d'adéquation décrite au Chapitre 5, on estimera sa puissance pour différentes tailles d'échantillon. En d'autres termes, on souhaite estimer la probabilité de rejet d'un modèle de copules

donné sous une certaine contre-hypothèse. À cette fin, on a généré des séries bivariées $(X_{1,1}, X_{1,2}), \dots, (X_{T,1}, X_{T,2})$ selon le modèle (5.2), avec une certaine copule C . Si l'hypothèse nulle est $\mathcal{H}_0 : C \in \mathcal{C}$, alors le test d'adéquation devrait la rejeter avec une forte probabilité lorsque $C \notin \mathcal{C}$. Pour quelques modèles de copules, on a donc généré $IT = 1\,000$ échantillons et calculé la proportion de rejet de \mathcal{H}_0 .

6.2.2 Résultats

Les Tableaux 6.8, 6.9 et 6.10 présentent les résultats de puissance de la méthode d'adéquation pour les tailles d'échantillon $T = 50$, $T = 100$ et $T = 250$, respectivement. Les familles de copules utilisées sont les mêmes que celles considérées dans l'évaluation des estimateurs, à savoir Ali-Mikhail-Haq (C^{AMH}), Farlie-Gumbel-Morgenstern (C^{FGM}), Clayton (C^{CL}), Gumbel-Hougaard (C^{GH}), Frank (C^{F}) et Normal (C^{N}). Pour chaque modèle de contre-hypothèse, on a considéré différents niveaux de dépendance mesurée par le tau de Kendall, noté τ . Pour les modèles de Ali-Mikhail-Haq et de Farlie-Gumbel-Morgenstern, on a $\tau \in \{0.10, 0.15, 0.20\}$, alors que pour les autres modèles, $\tau \in \{0.25, 0.50, 0.75\}$.

Tableau 6.8: Estimation de la puissance de la procédure d'adéquation, basée sur 1 000 répétitions, pour des tailles d'échantillon $T = 50$

Modèle sous \mathcal{H}_1	τ	Modèle sous \mathcal{H}_0					
		C^{AMH}	C^{FGM}	C^{CL}	C^{GH}	C^{F}	C^{N}
C^{AMH}	0.10	2,6	6,4	7,9	2,8	2,3	5,8
	0.15	2,1	5,6	9,9	4,6	2,8	5,3
	0.20	3,6	7,6	9,1	6,0	4,2	5,6
C^{FGM}	0.10	1,9	5,3	10,7	3,3	3,2	7,4
	0.15	4,2	6,3	11,7	4,1	3,7	6,2
	0.20	3,8	10,2	17,7	5,4	4,0	5,2
C^{CL}	0.25	1,8	14,2	6,9	15,8	6,5	6,8
	0.50	44,9	90,5	4,6	35,3	13,6	10,4
	0.75	99,9	100,0	5,5	22,5	7,0	8,2
C^{GH}	0.25	19,9	22,7	38,6	5,8	13,1	11,1
	0.50	76,4	91,7	67,0	4,3	20,4	13,1
	0.75	100,0	100,0	76,1	2,8	15,7	10,6
C^{F}	0.25	5,6	6,2	8,5	12,4	6,3	7,6
	0.50	4,7	4,7	11,5	13,5	2,9	5,6
	0.75	6,3	4,8	8,4	15,0	3,8	7,4
C^{N}	0.25	7,1	13,8	20,9	3,8	5,7	4,3
	0.50	62,0	90,6	40,0	3,8	7,0	4,4
	0.75	100,0	100,0	52,2	1,5	9,0	2,1

Tableau 6.9: Estimation de la puissance de la procédure d'adéquation, basée sur 1 000 répétitions, pour des tailles d'échantillon $T = 100$

Modèle sous \mathcal{H}_1	τ	Modèle sous \mathcal{H}_0					
		C^{AMH}	C^{FGM}	C^{CL}	C^{GH}	C^{F}	C^{N}
AMH	0.10	2,6	4,0	6,3	4,4	4,0	4,5
	0.15	3,3	5,1	8,8	6,8	4,0	4,1
	0.20	3,4	9,1	9,3	15,1	4,9	7,2
FGM	0.10	2,2	5,3	9,3	4,5	2,6	4,7
	0.15	3,7	5,5	15,1	5,8	3,1	5,3
	0.20	8,4	7,6	21,6	8,5	4,0	7,7
CL	0.25	4,2	26,0	5,4	37,8	18,5	16,2
	0.50	73,1	99,9	5,7	76,3	42,5	39,0
	0.75	100,0	100,0	5,8	75,1	36,5	49,5
GH	0.25	35,0	27,1	52,1	4,1	14,4	12,6
	0.50	94,9	99,3	94,1	4,8	27,2	17,3
	0.75	100,0	100,0	94,9	4,5	34,5	21,4
F	0.25	4,8	5,9	19,9	16,7	6,4	6,5
	0.50	7,2	5,4	44,0	24,5	4,5	9,4
	0.75	18,4	5,3	39,5	34,3	4,9	16,1
N	0.25	15,2	15,5	26,6	5,6	6,2	5,0
	0.50	86,0	99,9	64,0	9,6	11,8	5,3
	0.75	100,0	100,0	75,4	4,0	18,0	7,5

Tableau 6.10: Estimation de la puissance de la procédure d'adéquation, basée sur 1 000 répétitions, pour des tailles d'échantillon $T = 250$

Modèle sous \mathcal{H}_1	τ	Modèle sous \mathcal{H}_0					
		C^{AMH}	C^{FGM}	C^{CL}	C^{GH}	C^{F}	C^{N}
AMH	0.10	2,5	5,2	12,8	8,2	4,9	5,2
	0.15	4,6	4,5	15,3	15,8	6,6	5,8
	0.20	4,3	9,5	12,6	28,5	6,8	11,5
FGM	0.10	2,3	5,7	15,1	5,9	4,2	5,7
	0.15	5,1	4,9	29,4	9,5	5,1	7,0
	0.20	11,6	4,5	38,7	13,4	4,3	8,8
CL	0.25	5,9	42,5	6,9	74,0	37,2	34,7
	0.50	95,1	100,0	2,5	99,8	87,1	83,2
	0.75	100,0	100,0	3,8	99,5	89,9	91,9
GH	0.25	55,1	27,1	79,9	4,2	18,8	18,0
	0.50	100,0	100,0	98,6	7,0	46,9	28,8
	0.75	100,0	100,0	100,0	7,9	61,3	29,9
F	0.25	3,6	5,9	36,7	18,2	4,1	7,3
	0.50	7,6	5,6	86,0	46,1	5,5	15,5
	0.75	37,2	5,1	91,4	63,5	9,4	41,7
N	0.25	18,8	18,7	42,0	13,3	5,7	5,0
	0.50	98,9	100,0	79,2	20,7	16,7	6,0
	0.75	100,0	100,0	93,1	21,6	36,5	8,8

6.2.3 Interprétation des résultats

D'abord, on constate que les tests conservent bien leur seuil sous l'hypothèse nulle. En effet, les probabilités de rejet de \mathcal{H}_0 , lorsque celle-ci est vraie, sont très près du seuil de signification des tests, fixé à 0.05 pour l'ensemble des résultats de simulation présentés.

Du côté de la puissance, on constate que les tests sont davantage aptes à détecter des écarts à l'hypothèse nulle lorsque le niveau de dépendance, mesuré par τ , est élevé. En contre-partie, les tests sont moins puissants lorsque τ est petit; cela s'explique par le fait que dans ce cas, on est près de la copule d'indépendance, cette dernière faisant partie de toutes les familles de copules considérées. La méthode d'adéquation est particulièrement puissante lorsque la vraie copule appartient à la famille de Clayton. Par ailleurs, lorsque l'hypothèse nulle concerne l'appartenance à la famille de Frank, le test a relativement de la difficulté à rejeter \mathcal{H}_0 quand la vraie copule est plutôt membre de la classe de Clayton, de Gumbel–Hougaard ou Normale. Finalement, on voit que la puissance, dans tous les cas de figure étudiés, augmente sensiblement lorsque la taille des échantillons augmente.

Les résultats peuvent sembler décevants pour certaines copules; cela s'explique de différentes façons. D'abord, il est normal que certaines copules soient plus facilement détectables que d'autres. Pour un modèle donné, on peut même constater que les puissances varient selon le niveau de dépendance. En fait, les résultats ressemblent largement à ceux obtenus par Genest et al. (2009) dans un contexte non sériel.

CHAPITRE 7

ANALYSE DES INDICES BOURSIERS *STANDARD & POOR 500 ET NASDAQ COMPOSITE*

Dans ce chapitre, la méthode de construction d'un modèle de séries chronologiques à base de copules présentée au Chapitre 5 sera illustrée sur la série des valeurs journalières de fermeture de deux indices boursiers, à savoir

- (i) l'indice *Standard & poor 500*, noté SP500;
- (ii) l'indice *Nasdaq composite*, noté NASDAQ.

Les séries analysées seront basées sur les valeurs à la fermeture des $T = 1759$ jours ouvrables pendant la période allant du 1er juin 2000 au 1er juin 2007.

D'abord, à la Section 7.1, on présentera chacun des indices boursiers et on effectuera une modélisation univariée à l'aide du logiciel *Eviews*. Ensuite, à la Section 7.2, on tentera de sélectionner une copule qui explique bien la dépendance entre les résidus des deux séries. Enfin, on fera des prévisions des valeurs futures à l'aide du modèle global choisi.

7.1 Présentation des données et analyses univariées

7.1.1 Série SP500: indice *Standard & Poor 500*

L'indice boursier *Standard & Poor 500*, noté SP500, est basé sur les valeurs boursières de cinq cents grandes sociétés cotées sur les bourses américaines. Les $T = 1759$ valeurs de cette série sont présentées à la figure 7.1. On constate que la série présente une tendance à la baisse pour environ les cinq cents premières valeurs; il y a ensuite une stabilisation (environ deux cents jours), puis une hausse constante.

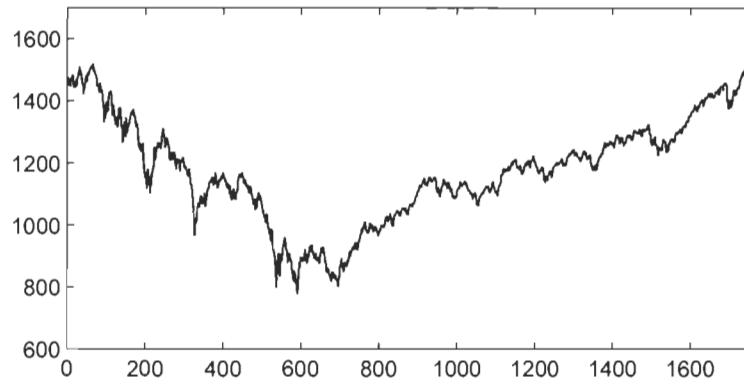


Figure 7.1: Valeurs journalières de fermeture de l'indice *Standard & poor 500* pour la période du 1er juin 2000 au 1er juin 2007

La figure 7.2 montre les fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour l'indice *Standard & poor 500*. On voit que toutes les autocorrélations sont significativement différentes de zéro; c'est le cas

seulement pour l'autocorrélation partielle d'ordre un. Ce comportement est typique d'une série autorégressive d'ordre un.

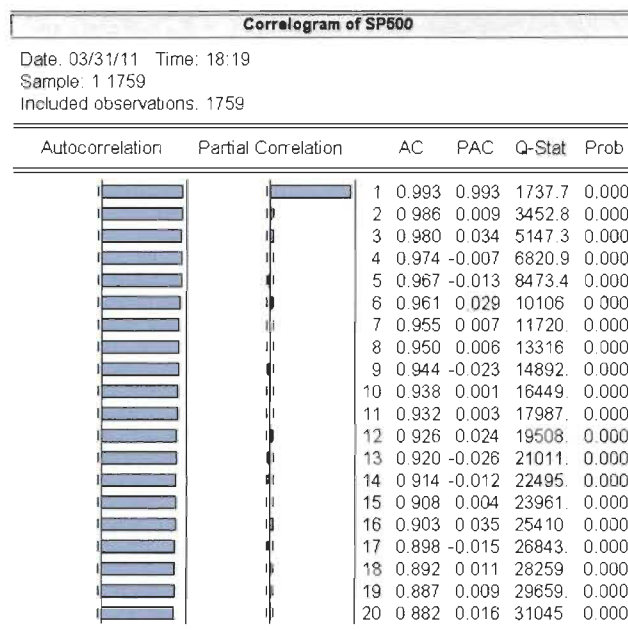


Figure 7.2: Fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour l'indice *Standard & poor 500*

La figure 7.3 présente l'analyse de la série différenciée *Standard & poor 500* en supposant que le modèle adéquat est AR(1), c'est-à-dire un processus autorégressif d'ordre un. Dans ce cas, le modèle est de la forme $(X_t - X_{t-1} - \mu) = \alpha(X_{t-1} - \mu) + \varepsilon_t$, où $\varepsilon_t \sim \mathcal{N}(0, \beta^2)$. Selon la sortie de *Eviews*, les paramètres estimés de la série dérivée sont $\hat{\mu} = 0.033625$ et $\hat{\alpha} = -0.031091$. Sur la base des statistiques de Student (*t-Statistic*), le coefficient $\hat{\alpha}$ n'est pas très significatif puisque la *p*-valeur associée est très petite, ce qui indique que le processus est un bruit blanc. Aussi, on a $\hat{\beta} = 11.77626$.

Dependent Variable: SP500				
Method: Least Squares				
Date: 03/31/11 Time: 18:32				
Sample (adjusted): 2 1759				
Included observations: 1758 after adjustments				
Convergence achieved after 4 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1203.302	152.3226	7.899690	0.0000
AR(1)	0.998131	0.001658	601.9731	0.0000
R-squared	0.995178	Mean dependent var	1176.710	
Adjusted R-squared	0.995175	S.D. dependent var	169.7500	
S.E. of regression	11.79149	Akaike info criterion	7.773771	
Sum squared resid	244153.1	Schwarz criterion	7.779996	
Log likelihood	-6831.145	F-statistic	362371.6	
Durbin-Watson stat	2.056400	Prob(F-statistic)	0.000000	
Inverted AR Roots	1.00			

Figure 7.3: Analyse de la série *Standard & poor 500* sous l'hypothèse d'un modèle autorégressif d'ordre un

Une façon de s'assurer que le modèle choisi est bon consiste à analyser les résidus. Idéalement, ceux-ci devraient se comporter comme un bruit blanc. La figure 7.4 montre les résidus du modèle autorégressif d'ordre un obtenu des résultats de la figure 7.3. La figure 7.5 montre les autocorrélations et les autocorrélations partielles des résidus; on voit qu'ils sont tous près de zéro.

Pour vérifier formellement l'hypothèse de bruit blanc des résidus, on base un test sur la Q-statistique. Pour les vingt premiers délais, celle-ci vaut 29.690 et la p -valeur associée est 0.056. Au seuil $\alpha = 0.05$, on peut donc rejeter l'hypothèse et affirmer qu'il est raisonnable de modéliser la série SP500 par un processus autorégressif d'ordre un.

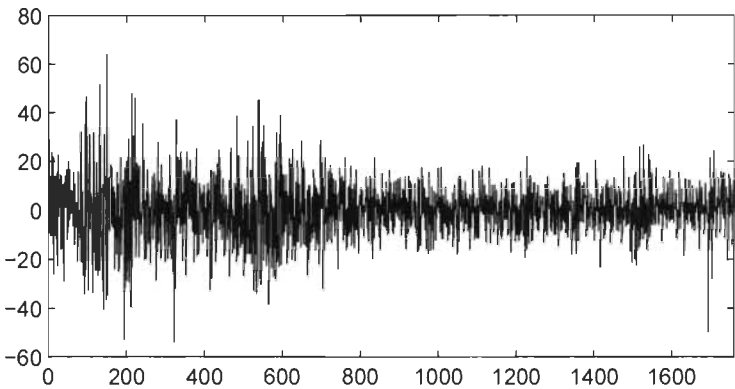


Figure 7.4: Résidus du modèle AR(1) pour la série SP500

Correlogram of Residuals					
Date: 03/31/11 Time: 18:34					
Sample: 2 1759					
Included observations: 1758					
Q-statistic probabilities adjusted for 1 ARMA term(s)					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.030	-0.030	1.5870	
		2 -0.058	-0.059	7.5277	0.006
		3 0.011	0.007	7.7322	0.021
		4 0.019	0.016	8.3377	0.040
		5 -0.042	-0.040	11.460	0.022
		6 -0.011	-0.012	11.693	0.039
		7 0.002	-0.004	11.697	0.069
		8 -0.000	-0.001	11.697	0.111
		9 -0.004	-0.003	11.723	0.164
		10 0.010	0.008	11.886	0.220
		11 -0.019	-0.020	12.559	0.249
		12 0.039	0.038	15.186	0.174
		13 0.051	0.051	19.751	0.072
		14 -0.019	-0.012	20.424	0.085
		15 -0.032	-0.027	22.289	0.073
		16 0.041	0.034	25.219	0.047
		17 0.007	0.007	25.305	0.065
		18 -0.019	-0.009	25.945	0.075
		19 -0.055	-0.055	31.272	0.027
		20 0.003	-0.006	31.291	0.038

Figure 7.5: Fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour les résidus de l'indice *Standard & poor 500*

7.1.2 Série NASDAQ: indice *Nasdaq composite*

L'indice boursier *Nasdaq composite* comprend toutes les sociétés américaines et étrangères cotées sur le marché électronique d'instruments financiers aux États-Unis; cela totalise plus de cinq milles sociétés. Les $T = 1759$ valeurs de cette série sont présentées à la figure 7.6. On voit que l'indice a subi une baisse très marquée pour les deux cents premiers jours; cette tendance a persisté, mais de façon plus légère, jusqu'à la six centième journée. Par la suite, on a enregistré une légère tendance à la hausse.

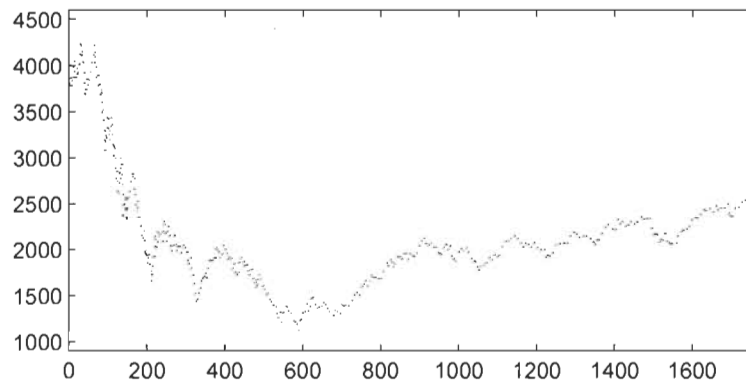


Figure 7.6: Valeurs journalières de fermeture de l'indice *Nasdaq composite* pour la période du 1er juin 2000 au 1er juin 2007

À la figure 7.7, on présente les fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour cette série. On voit que le modèle autorégressif d'ordre un n'est pas un bon modèle, car les autocorrélations partielles des délais 1, 2, 3, 5, 13 et 14 sont significativement différentes de zéro. Il faudra donc chercher à modéliser la série avec un autre modèle.

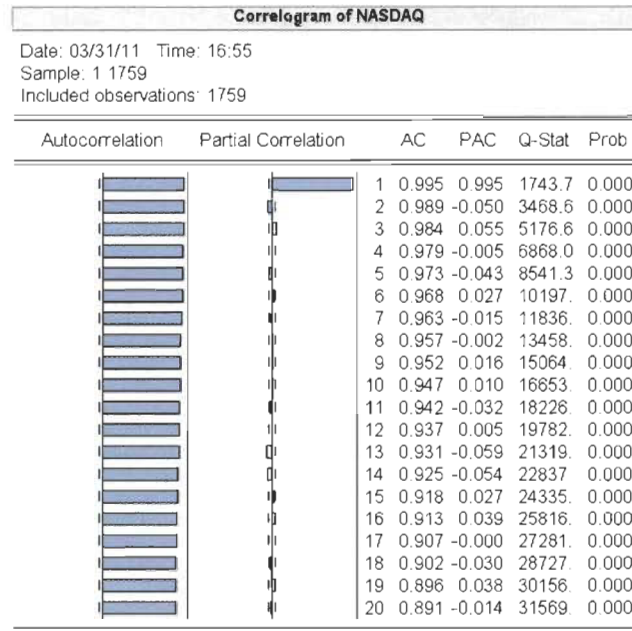


Figure 7.7: Fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour l'indice *Nasdaq composite*

Après de nombreuses étapes d'essais et erreurs, on a trouvé que le modèle GARCH(1,1), appliqué à la série dérivée, était approprié. Dans ce cas, le processus est un cas particulier de (2.6), à savoir

$$BX_t = \sigma_t \varepsilon_t, \quad \text{où } \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2) \quad \text{et} \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

La figure 7.8 montre l'analyse de la série *Nasdaq composite* sous l'hypothèse d'un tel modèle. Pour vérifier la qualité de l'ajustement, il s'agit d'analyser les résidus. Comme la Q-statistique vaut 20.674 et que sa p -valeur est 0.418 (voir la figure 7.9), on accepte l'hypothèse d'absence d'autocorrélation des résidus, confirmant que le processus GARCH(1,1) est un modèle acceptable.

Dependent Variable: D(NASDAQ)
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 03/31/11 Time: 17:38
Sample (adjusted): 2 1759
Included observations: 1758 after adjustments
Convergence achieved after 21 iterations
Variance backcast: ON
GARCH = C(1) + C(2)*RESID(-1)^2 + C(3)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
Variance Equation				
C	2.870644	0.745355	3.851379	0.0001
RESID(-1)^2	0.021502	0.004846	4.436990	0.0000
GARCH(-1)	0.971784	0.005077	191.4114	0.0000
R-squared	-0.000190	Mean dependent var	-0.550956	
Adjusted R-squared	-0.001329	S.D. dependent var	40.02928	
S.E. of regression	40.05587	Akaike info criterion	9.498370	
Sum squared resid	2815850.	Schwarz criterion	9.507707	
Log likelihood	-8346.067	Durbin-Watson stat	2.006700	

Figure 7.8: Analyse de la série *Nasdaq composite* sous l'hypothèse d'un modèle GARCH(1,1)

Correlogram of Standardized Residuals					
Date: 03/31/11 Time: 17:40					
Sample: 2 1759					
Included observations: 1758					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.000	0.000	8.E-06	0.998
		2 -0.050	-0.050	4.4831	0.106
		3 0.007	0.007	4.5592	0.207
		4 0.014	0.011	4.8898	0.299
		5 -0.020	-0.020	5.6100	0.346
		6 -0.023	-0.022	6.5474	0.365
		7 -0.020	-0.022	7.2618	0.402
		8 -0.015	-0.017	7.6563	0.468
		9 -0.017	-0.019	8.1972	0.514
		10 -0.001	-0.003	8.2004	0.609
		11 0.021	0.019	8.9551	0.626
		12 0.035	0.034	11.105	0.520
		13 0.039	0.040	13.830	0.386
		14 0.001	0.003	13.831	0.462
		15 -0.015	-0.014	14.248	0.507
		16 0.042	0.040	17.337	0.364
		17 0.005	0.004	17.376	0.429
		18 -0.018	-0.010	17.957	0.458
		19 -0.037	-0.033	20.373	0.372
		20 0.013	0.013	20.674	0.417

Figure 7.9: Fonctions d'autocorrélation et d'autocorrélation partielle empiriques pour les résidus de l'indice *Nasdaq composite*

7.2 Analyse multivariée

7.2.1 Sélection d'une copule

Dans cette section, on effectuera une analyse multivariée des indices *Standard & poor 500* et *Nasdaq composite*. Les deux séries sont montrées simultanément à la figure 7.10.

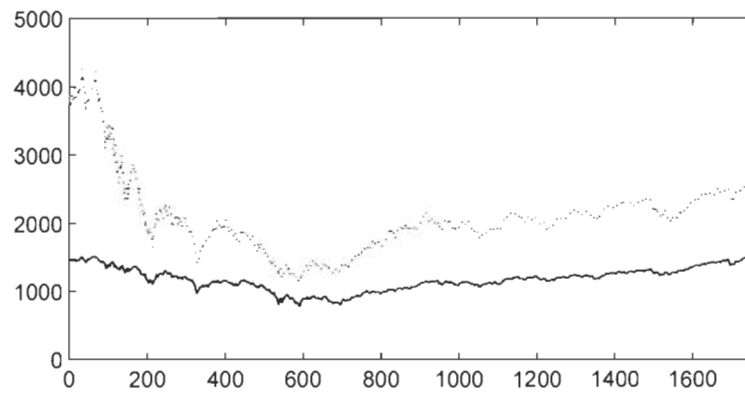


Figure 7.10: Valeurs journalières de fermeture des indices *Standard & poor 500* et *Nasdaq composite* pour la période du 1er juin 2000 au 1er juin 2007

On peut présumer qu'il y a une dépendance importante entre les deux séries, puisqu'elles tendent à se comporter de façon similaire. En effet, les périodes de croissance et de décroissance des deux indices coïncident. On peut constater ce lien en examinant le nuage de points des résidus présenté à la figure 7.11. On voit que la dépendance est fortement positive. Cela est confirmé par la valeur du coefficient de corrélation, qui vaut 0.8207; cette valeur est significativement supérieure à zéro.

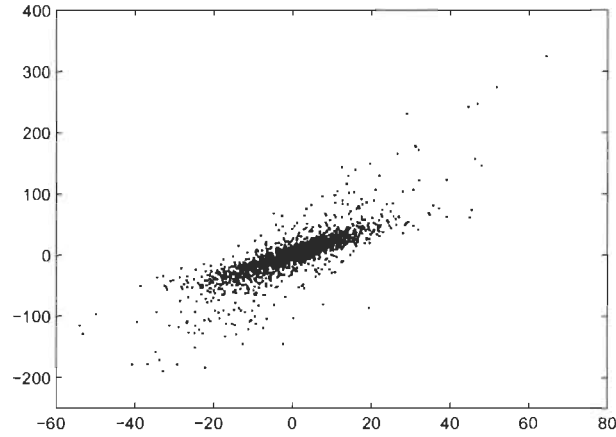


Figure 7.11: Nuage de points des résidus de la série bivariable SP500–Nasdaq

Le nuage de points de la figure 7.11 est influencé par trois éléments, à savoir les comportements marginaux des résidus des deux séries et la forme de la dépendance, c'est-à-dire la copule. Pour voir plus clairement la forme de la copule sous-jacente aux résidus, il faut les *normaliser* en considérant leurs rangs normalisés. Le nuage de points correspondant est offert à la figure 7.12.

Afin d'identifier formellement un modèle de copules adéquat pour ces résidus, on appliquera le test d'adéquation qui a été développé au Chapitre 5. Pour cela, six modèles de copules ont été considérés : Ali–Mikhail–Haq, Farlie–Gumbel–Morgenstern, Clayton, Gumbel–Hougaard, Frank et Normale. Les résultats sont présentés au Tableau 7.1. Strictement parlant, on voit qu'aucun des six modèles n'est accepté au seuil de signification $\alpha = 0.05$ car dans tous les cas, la valeur de la statistique W_T est plus grande que la valeur critique estimée par le *bootstrap* paramétrique. Cependant, on voit que le rejet de la copule Normale n'est pas sans appel : la valeur de la statistique n'excède pas

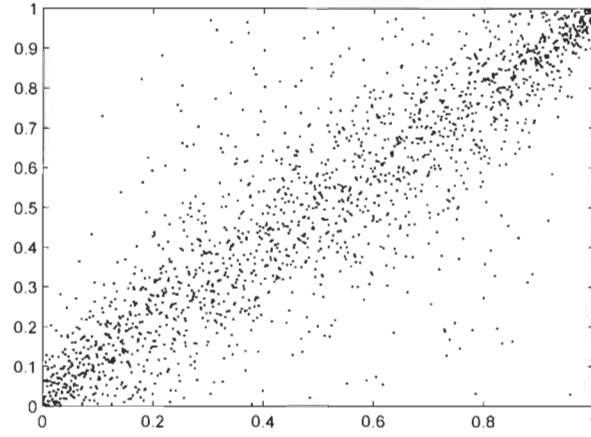


Figure 7.12: Nuage de points des résidus normalisés de la série bvariée SP500–Nasdaq

de beaucoup la valeur critique. On peut donc considérer que ce modèle est raisonnable. D'ailleurs, pour fins de comparaison, la figure 7.13 montre le nuage de points des résidus standardisés ainsi que celui de $T = 1757$ paires de points générées de la copule Normale dont le tau de Kendall est le même que le tau de Kendall des résidus, soit $\tau = 0.6994$.

7.2.2 Prédiction

Soient $\tilde{X}_{t,1}$ et $\tilde{X}_{t,2}$, les valeurs *différenciées* des indices SP500 et NASDAQ au temps t , c'est-à-dire $\tilde{X}_{t,1} = \mathcal{B}X_{t,1}$ et $\tilde{X}_{t,2} = \mathcal{B}X_{t,2}$. En se basant sur les analyses univariées des sous-sections 7.1.1 et 7.2.2, ainsi que sur la sélection

Tableau 7.1: Résultats des tests d'adéquation de copules pour les séries *Standard & poor 500* et *Nasdaq composite*

Famille de copules	W_T	Valeur critique	$\hat{\theta}$
Ali-Mikhail-Haq	3,4572	0,0507	1,0000
Farlie-Gumbel-Morgenstern	5,8125	0,0647	1,0000
Clayton	0,1499	0,0168	4,6523
Gumbel-Hougaard	0,0590	0,0154	0,6994
Frank	0,0567	0,0158	11,3686
Normale	0,0258	0,0154	0,8905

de la copule faite à la sous-section 7.2.1, le modèle retenu est

$$\begin{aligned}\tilde{X}_{t,1} &= -0.0311X_{t-1,1} + \varepsilon_{t,1} \\ \tilde{X}_{t,2} &= \sigma_t \varepsilon_{t,2},\end{aligned}$$

où $\sigma_t^2 = 2.8706 + 0.9718 \tilde{X}_{t-1,2}^2 + 0.0215 \sigma_{t-1}^2$ et la copule sous-jacente à $(\varepsilon_{t,1}, \varepsilon_{t,2})$ est la Normale de paramètre $\hat{\theta} = 0.8895$. À noter que l'équation pour la variance hétéroscédastique σ_t^2 peut s'écrire différemment à l'aide de l'opérateur délai. En effet, on a

$$(1 - 0.0215 \mathcal{B}) \sigma_t^2 = 2.8706 + 0.9718 \tilde{X}_{t-1}^2,$$

ce qui amène

$$\begin{aligned}\sigma_t^2 &= 2.8706 + 0.9718 \sum_{i=0}^{\infty} (0.0215)^i \tilde{X}_{t-i-1}^2 \\ &\approx 2.8706 + 0.9718 \tilde{X}_{t-1}^2 + 0.9718(0.0215) \tilde{X}_{t-2}^2.\end{aligned}$$

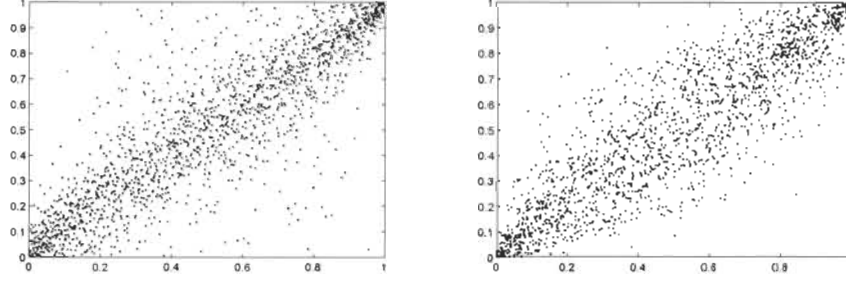


Figure 7.13: Nuages des résidus (à gauche) et de $T = 1757$ observations indépendantes tirées de la copule Normale avec $\tau = 0.6994$ (à droite)

Ainsi, on déduit que conditionnellement aux valeurs passées $(\tilde{X}_{t,1}, \tilde{X}_{t,2})$, $t \in \{1, \dots, T\}$, les valeurs futures $\tilde{X}_{T+1,1}$ et $\tilde{X}_{T+1,2}$ sont telles que

$$\begin{aligned}\tilde{X}_{T+1,1} &\sim \mathcal{N}\left(-0.0311\tilde{X}_{T,1}, 0.0336\right), \\ \tilde{X}_{T+1,2} &\sim \mathcal{N}\left(0, 2.8706 + 0.9718\tilde{X}_{T,2}^2 + 0.9718(0.0215)\tilde{X}_{T-1,2}^2\right).\end{aligned}$$

Comme la copule de $(\varepsilon_{t,1}, \varepsilon_{t,2})$ est Normale de paramètre $\hat{\theta} = 0.8895$, la loi conjointe de $(\tilde{X}_{T+1,1}, \tilde{X}_{T+1,2})$ est Normale bivariée de moyennes $\mu_1 = -0.0311\tilde{X}_{T,1}$ et $\mu_2 = 0$, de variances $\sigma_1^2 = 0.0336$ et $\sigma_2^2 = 2.8706 + 0.9718\tilde{X}_{T,2}^2 + 0.9718(0.0215)\tilde{X}_{T-1,2}^2$, et de corrélation $\rho = 0.8895$.

Un résultat classique en statistique mathématique permet de déduire les lois conditionnelles dans le modèle Normal bivarié. Spécifiquement, supposons que le couple (X_1, X_2) est distribué selon la loi Normale à deux variables de moyennes μ_1, μ_2 , de variances σ_1^2, σ_2^2 , et de corrélation ρ . Alors

$$\begin{aligned}(X_1|X_2 = x) &\sim \mathcal{N}\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(x - \mu_2), (1 - \rho^2)\sigma_1^2\right), \\ (X_2|X_1 = x) &\sim \mathcal{N}\left(\mu_2 + \frac{\sigma_2}{\sigma_1}\rho(x - \mu_1), (1 - \rho^2)\sigma_2^2\right).\end{aligned}$$

Par conséquent,

$$E(X_1|X_2 = x) = \mu_1 + \frac{\sigma_1}{\sigma_2} \rho(x - \mu_2) \quad \text{et} \quad (X_2|X_1 = x) = \mu_2 + \frac{\sigma_2}{\sigma_1} \rho(x - \mu_1).$$

Dans notre cas, on a déjà $\sigma_1^2 = 0.033625$, $\mu_2 = 0$ et $\rho = 0.8895$. Ensuite, puisque $\tilde{X}_{T,1} = 0.39$, $\tilde{X}_{T,2} = 9.4$ et $\tilde{X}_{T-1,2} = 11.9$, des calculs directs permettent d'obtenir $\mu_1 = -0.0121$ et $\sigma_2^2 = 91.70$. Ainsi,

$$\begin{aligned} g_1(x) &= E(\tilde{X}_{T+1,1}|\tilde{X}_{T+1,2} = x) = 0.017x - 0.012, \\ g_2(x) &= E(\tilde{X}_{T+1,2}|\tilde{X}_{T+1,1} = x) = 46.7561x + 0.5657. \end{aligned}$$

Ces courbes sont tracées à la figure 7.14. Pour illustrer, supposons que les deux dernières valeurs de l'indice NASDAQ sont telles que $\tilde{X}_{T,2} = 10$; la valeur prédite de l'indice SP500 sera donc $\hat{\tilde{X}}_{T+1,1} = g_1(10) = 0.158$. À l'inverse, si $\tilde{X}_{T,1} = 0.5$, alors $\hat{\tilde{X}}_{T+1,2} = g_2(0.5) = 23.94$.

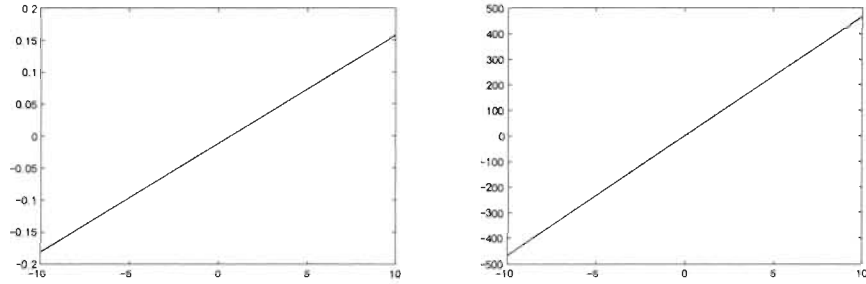


Figure 7.14: Espérances conditionnelles $E(\tilde{X}_{T+1,1}|\tilde{X}_{T+1,2} = x)$ (à gauche) et $E(\tilde{X}_{T+1,2}|\tilde{X}_{T+1,1} = x)$ (à droite)

CONCLUSION

Ce mémoire de maîtrise s'est intéressé à la modélisation de séries chronologiques. Dans la première partie du travail, on a effectué une revue des principales notions de base concernant les séries chronologiques à une variable; cela a permis d'introduire, entre autres, la famille des processus ARMA ainsi que la procédure de sélection de modèles de Box & Jenkins. Ensuite, pour pallier au manque de flexibilité de la plupart des modèles de séries chronologiques à plusieurs variables, on a développé un nouveau modèle à base de copules. Cette partie constitue l'apport principal de ce mémoire.

Le modèle que l'on a proposé offre un avantage important face à ses concurrents *classiques*, qui supposent généralement que la structure de dépendance est de type *normal*. Le modèle que l'on a introduit est beaucoup plus flexible car il permet de modéliser la forme de la dépendance entre plusieurs séries chronologiques par une panoplie de scénarios possibles. Cela fut rendu faisable grâce à l'utilisation judicieuse de la théorie des copules.

Le fait que la forme de la copule ne soit pas spécifiée dans notre modèle a posé le problème de la sélection d'une famille appropriée. Pour choisir un bon modèle, on a proposé une méthode d'adéquation formelle permettant de décider si une famille de copules est adéquate ou non pour expliquer la

dépendance entre plusieurs séries observées. L'étude de puissance exhaustive qui a été élaborée permet de conclure que la procédure proposée est efficace: d'abord, les tests maintiennent leur niveau sous l'hypothèse nulle; de plus, ils sont très puissants. La méthode n'est évidemment pas parfaite : il arrive que certaines copules soient plus difficiles à détecter.

Le but ultime de la modélisation des séries chronologiques est de prévoir les valeurs futures. Dans cette veine, on a élaboré une méthode de prévision à partir du modèle de séries chronologiques à base de copules que l'on a proposé. Tel qu'on l'a vu, cette prévision dépend à la fois des modèles choisis pour les séries individuelles et de la copule sélectionnée. Cette prévision devrait généralement être plus précise qu'une autre basée sur l'hypothèse d'une structure de dépendance normale. Dans le cadre de recherches futures, il serait d'ailleurs intéressant d'évaluer l'efficacité de notre méthode de prévision par rapport aux méthodes *classiques*. Il serait également intéressant d'étudier la possibilité d'apporter davantage de flexibilité à notre modèle en permettant au paramètre de la copule de changer dans le temps.

Références

- ALI, M. M., MIKHAIL, N. N. & HAQ, M. S. (1978). A class of bivariate distributions including the bivariate logistic. *J. Multivariate Anal.* **8**, 405–412.
- BEGUIN, J.-M., GOURIÉROUX, C. & MONFORT, A. (1980). Identification of a mixed autoregressive-moving average process: the corner method. In *Time series (Proc. Internat. Meeting, Univ. Nottingham, Nottingham, 1979)*. Amsterdam: North-Holland, pp. 423–436.
- BOX, G. E. P. & JENKINS, G. M. (1970). *Times series analysis. Forecasting and control*. San Francisco, Calif.: Holden-Day.
- BOX, G. E. P. & PIERCE, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.* **65**, 1509–1526.
- BREYMANN, W., DIAS, A. & EMBRECHTS, P. (2003). Dependence structures for multivariate high-frequency data in finance. In *Selected Proceedings from Quantitative Methods in Finance, 2002 (Cairns/Sydney)*, vol. 3.

- BROCKWELL, P. J. & DAVIS, R. A. (2002). *Introduction to time series and forecasting*. Springer Texts in Statistics. New York: Springer-Verlag, 2nd ed. With 1 CD-ROM (Windows).
- CHENG, R. C. H. & STEPHENS, M. A. (1989). A goodness-of-fit test using Moran's statistic with estimated parameters. *Biometrika* **76**, 385–392.
- CLAYTON, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.
- DICKEY, D. A. & FULLER, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74**, 427–431.
- ENGLE, R. F. (1982). A general approach to Lagrange multiplier model diagnostics. *J. Econometrics* **20**, 83–104.
- FAN, J. & YAO, Q. (2003). *Nonlinear time series*. Springer Series in Statistics. New York: Springer-Verlag. Nonparametric and parametric methods.
- GENEST, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549–555.
- GENEST, C., GHOUDI, K. & RIVEST, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.
- GENEST, C., QUESSY, J.-F. & RÉMILLARD, B. (2006). Goodness-of-fit

- procedures for copula models based on the probability integral transformation. *Scand. J. Statist.* **33**, 337–366.
- GENEST, C. & RÉMILLARD, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Inst. Henri Poincaré Probab. Stat.* **44**, 1096–1127.
- GENEST, C., RÉMILLARD, B. & BEAUDOIN, D. (2009). Goodness-of-fit tests for copulas: a review and a power study. *Insurance Math. Econom.* **44**, 199–213.
- GOURIÉROUX, C. (1997). *ARCH models and financial applications*. Springer Series in Statistics. New York: Springer-Verlag.
- HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19**, 293–325.
- JARQUE, C. M. & BERA, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econom. Lett.* **6**, 255–259.
- JUSELIUS, K. (2006). *The cointegrated VAR model: methodology and applications*. Advanced Texts in Econometrics. Oxford: Oxford University Press.
- KWIATKOWSKI, PHILLIPS, S. & SHIN (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* **54**, 159–178.

- LJUNG, G. & BOX, G. (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 297.
- MÉLARD, G. (1990). *Méthodes de prévision à court terme*. Editions de l'Université de Bruxelles, Bruxelles.
- MESFIOUI, M., QUESSY, J.-F. & TOUPIN, M.-H. (2009). On a new goodness-of-fit process for families of copulas. *Canad. J. Statist.* **37**, 80–101.
- NELSEN, R. B. (2006). *An introduction to copulas*, 2nd edition. New York: Springer.
- PHILLIPS, P. C. B. & PERRON, P. (1988). Testing for a unit root in time series regression. *Biometrika* **75**, 335–346.
- SKLAR, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.
- STEPHENS, M. A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *Ann. Statist.* **4**, 357–369.
- TSAY, R. S. (2005). *Analysis of financial time series*. Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley-Interscience [John Wiley & Sons], 2nd ed.
- VAN DER VAART, A. W. & WELLNER, J. A. (1996). *Weak convergence and empirical processes*. Springer Series in Statistics. New York: Springer-Verlag. With applications to statistics.

- WOLD, H. O. A. (1948). On prediction in stationary time series. *Ann. Math. Statistics* **19**, 558–567.

ANNEXE A

PROGRAMMES EN MATLAB

A.1 Simulation de nombre aléatoires

A.1.1 Paires indépendantes à partir d'une copule

```
1 function X = sim_copula(n,model,tau)
2
3 % Simulation d'un processus bivarié de taille n suivant une copule :
4 % 1. Clayton
5 % 3. Ali-Mikhail-Haq
6 % 4. Gumbel-Hougaard
7 % 5. Frank
8 % 7. Farlie-Gumbel-Morgenstern
9 % 11. Normal
10
11 theta = tau_theta(model,tau);
12
13 %Clayton
14 if (model == 1)
15     if (theta == 0)
16         X = rand(n,2);
17     else
18         for i=1:n
19             v = rand;
20             t = rand;
21             a = v^(-theta);
22             b = t^(-theta/(theta+1)) - 1;
23             X(i,1) = (a*b+1)^(-1/theta);
24             X(i,2) = v;
25         end
26     end
27
28
```

```

29 %Ali-Mikhail-Haq
30 elseif (model == 3)
31     if (theta == 0)
32         X = rand(n,2);
33     else
34         for i=1:n
35             u = rand; t = rand;
36             a = -theta - 4*t*theta^2*u + 2*theta*t*u - 2*theta*t + 2*theta^2*t + 2*theta^2
37                 b = 1 - 2*theta + 4*t*theta^2*u^2 + 4*theta*t*u - 4*theta^2*t*u +
38                 c = -2*theta^2*t*u + theta^2*u^2*t - theta + theta^2*t;
39             X(i,1) = u;
40             X(i,2) = (a-sqrt(b))/(2*c);
41         end
42     end
43
44 %Gumbel-Hougaard
45 elseif (model == 4)
46     if (theta == 0)
47         X = rand(n,2);
48     else
49         for i=1:n
50             U0 = rand; U1 = rand; U2 = rand;
51             z = (U0^(1-theta))/(U0^(1-theta)+(1.0-U0)^(1-theta));
52             if (rand < 1-theta)
53                 w = U1*U2;
54             else
55                 w = U2;
56             end
57             t1 = ((1-z)^(1/(1-theta)) + z^(1/(1-theta)))^(1-theta);
58             X(i,1) = w^(z/t1);
59             X(i,2) = w^((1.0-z)/t1);
60         end
61     end
62
63 %Frank
64 elseif (model == 5)
65     if (theta == 0)
66         X = rand(n,2);
67     else
68         for i=1:n
69             u = rand; t = rand;
70             t1 = t*exp(-theta) + (1-t)*exp(-theta*u);
71             t2 = t + (1-t)*exp(-theta*u);
72             X(i,1) = u;
73             X(i,2) = log(t2/t1)/theta;
74         end
75     end
76
77
78
79
80

```

```

81 %Farlie-Gumbel-Morgenstern
82 elseif (model == 7)
83     if (theta == 0)
84         for i=1:n
85             X(i,1) = rand; X(i,2) = rand;
86         end
87     else
88         for i=1:n
89             u = rand; t = rand;
90             b = theta*(1-2*u);
91             root = sqrt((1+b)^2 - 4*t*b);
92             X(i,1) = u;
93             X(i,2) = 2*t/(1+b+root);
94         end
95     end
96
97 %Normal
98 elseif (model == 11)
99     mu = [0 0]; sigma = [1 theta ; theta 1];
100     for i=1:n
101         X(i,:) = mvnrnd(mu,sigma);
102     end
103     X = normcdf(X);
104
105 end

```


A.1.2 Deux processus AR(1) liés par une copule

```

1  function H = sim_copula_ts(n,model,tau,rhoX,sigmaX,rhoY,sigmaY)
2
3  % Étape 1 : Simulation de deux séries réchauffées
4
5  % Longueur du réchauffement
6  N = 100;
7
8  % Création de l'espace mémoire pour les deux séries réchauffées (SR)
9  W = zeros(N,1); Z = zeros(N,1);
10
11 % Détermination aléatoire de la valeur de départ de chacune des SR
12 W(1) = rand(); Z(1) = rand();
13
14 % Simulation d'un vecteur bivarié suivant une certaine copule
15 U = sim_copula(N,model,tau);
16
17 % Calcul des résidus standardisés au seuil de 5E = norminv(U(:,1),0,sigmaX);
18 e = norminv(U(:,2),0,sigmaY);
19
20 % Simulation des séries réchauffées
21 for t=2:N
22     W(t) = rhoX*W(t-1) + E(t);
23     Z(t) = rhoY*Z(t-1) + e(t);
24 end
25
26 % Étape 2 : Simulation de deux vraies séries à partir des SR
27
28 % Création de l'espace mémoire pour les deux vraies séries
29 X = zeros(n,1); Y = zeros(n,1);
30
31 % Détermination de la valeur de départ de chacune des vraies séries
32 X(1) = W(N); Y(1) = Z(N);
33
34 % Simulation d'un vecteur bivarié suivant une certaine copule
35 U = sim_copula(n,model,tau);
36
37 % Calcul des résidus standardisés au seuil de 5E = norminv(U(:,1),0,sigmaX);
38 e = norminv(U(:,2),0,sigmaY);
39
40 % Simulation des vraies séries
41 for t=2:n
42     X(t) = rhoX*X(t-1) + E(t);
43     Y(t) = rhoY*Y(t-1) + e(t);
44 end
45
46 % Regroupement des deux vraies séries dans un vecteur
47 H = [X,Y];

```

A.2 Estimation de paramètres

A.2.1 Processus AR(1)

```
1 function [rho, sigma] = EMV_AR(X)
2
3 % Estimation des paramètres d'une série auto-régressive d'ordre un
4
5 % Points de départ des paramètres rho et sigma pour le calcul des EMV
6 tx0 = [0, std(X)];
7
8 % Estimation des paramètres d'une série AR(1) par la méthode du maximum de vraisem-
9   blance
10 B = fminsearch(@(A)(-likelihood_AR(X,A)), tx0);
    rho = B(1); sigma = B(2);
```

A.2.2 Processus AR(1) liés par une copule

```

1  function [theta , rhoX , sigmaX , rhoY , sigmaY] = EMV(X, model)
2
3  % Estimation des paramètres de deux séries auto-régressives d'ordre un et
4  % du paramètre de la copule selon une certaine copule et un vecteur X
5
6  % Définition de X1 et X2 à partir du vecteur X pour le traitement univarié
7  X1 = X(:,1);
8  X2 = X(:,2);
9
10 % Points de départ des paramètres des deux séries pour le calcul de l'EMV
11 tx0 = [0, std(X1)];
12 ty0 = [0, std(X2)];
13
14 % Étape 1 : estimation des paramètres de la première série AR(1)
15 B = fminsearch(@(A)(-likelihood_AR(X1,A)), tx0);
16 rhoX = B(1); sigmaX = B(2);
17
18 % Étape 2 : estimation des paramètres de la deuxième série AR(1)
19 C = fminsearch(@(A)(-likelihood_AR(X2,A)), ty0);
20 rhoY = C(1); sigmaY = C(2);
21
22
23 % Étape 3 : estimation du paramètre de la copule
24
25 % Regroupement des paramètres des deux séries AR(1) en un vecteur
26 P = [rhoX , sigmaX , rhoY , sigmaY];
27
28 % Points de départ du paramètre theta pour le calcul de l'EMV
29 tk0 = tau_theta(model , tau(X));
30
31 % Restriction du nombre d'évaluations permises
32 options = optimset('MaxFunEvals',500);
33
34 % Estimation du paramètre theta par la méthode du maximum de vraisemblance
35 theta = fminsearch(@(T)(-Likelihood_copula(X,model,P,T)),tk0,options);

```

A.2.3 EQM du maximum de vraisemblance

```

1  function [EQM_theta ,EQM_rhoX ,EQM_sigmaX ,EQM_rhoY ,EQM_sigmaY]=EQM_EMV( model , n , IT )
2
3  % Valeurs du paramètre de dépendance de la copule (tau de Kendall)
4  tau = [0.25 , 0.5 , 0.75]; J=3;
5
6  % Valeurs des paramètres des processus autorégressifs d'ordre un
7  sigmaX=1; sigmaY=2;
8  rho = [.25 ,.5 ,.75]; K = 3;
9
10 % Création de l'espace mémoire pour les EQM des paramètres
11 EQM_theta = zeros(J,K);
12 EQM_rhoX = zeros(J,K);
13 EQM_sigmaX = zeros(J,K);
14 EQM_rhoY = zeros(J,K);
15 EQM_sigmaY = zeros(J,K);
16
17 % Boucle pour les valeurs du tau de Kendall
18 for j=1:J
19     theta = tau_theta( model , tau(j) );
20
21     % Boucle pour les différentes valeurs de rhoX, rhoY
22     for k=1:K
23         rhoX = rho(k); rhoY = rho(k);
24
25         % Estimation des erreurs quadratiques avec IT répétitions
26         for i=1:IT
27
28             % Simulation de 2 processus AR(1) selon une certaine copule
29             X = sim_copula_ts( n , model , tau(j) , rhoX , sigmaX , rhoY , sigmaY );
30
31             % Estimation des paramètres à partir de la simulation
32             [A,B,C,D,E] = EMV(X, model );
33
34             % Incrémentations des erreurs quadratiques pour les différentes
35             % valeurs de tau et de rho
36             EQM_theta(j , k) = EQM_theta(j , k) + (theta -A)^2;
37             EQM_rhoX(j , k) = EQM_rhoX(j , k) + (rhoX -B)^2;
38             EQM_sigmaX(j , k) = EQM_sigmaX(j , k) + (sigmaX -C)^2;
39             EQM_rhoY(j , k) = EQM_rhoY(j , k) + (rhoY -D)^2;
40             EQM_sigmaY(j , k) = EQM_sigmaY(j , k) + (sigmaY -E)^2;
41         end
42     end
43 end
44
45 % Estimation des erreurs quadratiques moyennes (EQM)
46 EQM_theta = EQM_theta / IT;
47 EQM_rhoX = EQM_rhoX / IT;
48 EQM_sigmaX = EQM_sigmaX / IT;
49 EQM_rhoY = EQM_rhoY / IT;
50 EQM_sigmaY = EQM_sigmaY / IT;

```

A.2.4 EQM de l'inversion du tau de Kendall

```

1  function EQM_theta = EQM_Kendall(model,n,IT)
2
3  % Valeurs du paramètre de dépendance de la copule (tau de Kendall)
4  tau = [0.25 , 0.5 , 0.75]; J=3;
5
6  % Valeurs des paramètres des processus autorégressifs d'ordre un
7  sigmaX=1; sigmaY=2;
8
9  rho = [.25 ,.5 ,.75]; K = 3;
10
11 % Création de l'espace mémoire pour les EQM du paramètre de dépendance
12 EQM_theta = zeros(J,K);
13
14 % Boucle pour les valeurs du tau de Kendall
15 for j=1:J
16     theta = tau_theta(model,tau(j));
17
18     % Boucle pour les différentes valeurs de rhoX, rhoY
19     for k=1:K
20         rhoX = rho(k); rhoY = rho(k);
21
22         % Estimation des erreurs quadratiques avec IT répétitions
23         for i=1:IT
24
25             % Simulation de 2 processus AR(1) selon une certaine copule
26             X = sim_copula_ts(n,model,tau(j),rhoX,sigmaX,rhoY,sigmaY);
27
28             % Estimation des paramètres à partir de la simulation
29             [A,B] = EMV_AR(X(:,1));
30             [C,D] = EMV_AR(X(:,2));
31
32             % Calcul des résidus à l'aide des estimations de paramètres
33             % autorégressifs
34             U_hat = zeros(n-1,2);
35             for l=1:n-1
36                 U_hat(l,1) = ( X(l+1,1) - A*X(l,1) ) / B;
37                 U_hat(l,2) = ( X(l+1,2) - C*X(l,2) ) / D;
38             end
39
40             % Calcul du paramètre de dépendance theta
41             E = tau_theta(model,tau(U_hat));
42
43             % Incrémentation des erreurs quadratiques pour les différentes
44             % valeurs de tau et de rho
45             EQM_theta(j,k) = MSE_theta(j,k) + (theta-E)^2;
46         end
47     end
48 end
49 EQM_theta = EQM_theta/IT;

```

A.3 Procédure d'adéquation

A.3.1 Calcul de la statistique W_T

```

1 function [Tn,kendall,theta] = CvM_TS(X,model,N)
2
3 % Calcul de la longueur du vecteur bivarié X
4 n = length(X);
5
6 % Estimation des paramètres des deux processus AR(1) à partir du vecteur X
7 [rho1,sigma1] = EMV_AR(X(:,1));
8 [rho2,sigma2] = EMV_AR(X(:,2));
9
10 % Calcul des résidus standardisés
11 U_hat1 = zeros(n-1,2);
12 for i=1:n-1
13     U_hat1(i,1) = ( X(i+1,1) - rho1*X(i,1) ) / sigma1;
14     U_hat1(i,2) = ( X(i+1,2) - rho2*X(i,2) ) / sigma2;
15 end
16
17 % Calcul du rang des résidus standardisés
18 U_hat = [rankit(U_hat1(:,1))/n,rankit(U_hat1(:,2))/n];
19
20 % Calcul du tau de kendall et estimation de theta
21 kendall = tau(U_hat);
22 theta = tau_theta(model,kendall);
23
24 % Simulation d'un processus bivarié de taille N suivant une certaine copule
25 Y = sim_copula(N,model,kendall);
26
27 % Calcul du rang des données issues de la simulation
28 U = [rankit(Y(:,1))/(N+1),rankit(Y(:,2))/(N+1)];
29
30 % Calcul de la statistique de test Tn
31 sum1=0; sum2=0; sum3=0;
32 for i=1:n-1
33     for j=1:n-1
34         sum1 = sum1 + (1-max(U_hat(i,1),U_hat(j,1)))*(1-max(U_hat(i,2),U_hat(j,2)));
35     end
36 end
37 for i=1:n-1
38     for j=1:N
39         sum2 = sum2 + (1-max(U_hat(i,1),U(j,1)))*(1-max(U_hat(i,2),U(j,2)));
40     end
41 end
42 for i=1:N
43     for j=1:N
44         sum3 = sum3 + (1-max(U(i,1),U(j,1)))*(1-max(U(i,2),U(j,2)));
45     end
46 end
47 Tn = (1/(n-1))*sum1 - (2/N)*sum2 + ((n-1)/N^2)*sum3;

```

A.3.2 Estimation des valeurs critiques

```

1  function CV = CV_CvM(n, model, N, IT)
2
3  % Création de l'espace mémoire pour les valeurs critiques (VC)
4  CV = zeros(10,1);
5
6  % Boucle pour les valeurs du tau de Kendall
7  for j=1:10
8
9      % Détermination des valeurs de tau de Kendall
10     kendall = (j - 1)/10;
11
12     % Création de l'espace mémoire pour les statistiques de test Tn
13     T = zeros(IT,1);
14
15     % Boucle pour le nombre de statistiques de test Tn à calculer
16     for i=1:IT
17
18         % Simulation de 2 processus AR(1) selon une certaine copule
19         X = sim_copula_ts(n, model, kendall, .2, 1, .2, 1);
20
21         % Calcul et enregistrement des statistiques de test Tn
22         T(i) = CvM_TS(X, model, N);
23     end
24
25     % Calcul et enregistrement de la VC pour un tau de Kendall donné selon
26     % le (1-0.05)-ème percentile de Tn(1),...,Tn(IT)
27     CV(j) = prctile(T,95);
28
29     % Suppression des enregistrements du vecteur de statistiques de test Tn
30     clear T;
31 end

```

A.3.3 Puissance

```

1  function pow = power_TS(n,modelHo,modelH1,tauH1,N,IT)
2
3  % Valeurs des paramètres du processus autorégressif d'ordre un
4  rhoX = .5; sigmaX = 1;
5  rhoY = .5; sigmaY = 1;
6
7  % Détermination de la colonne du fichier de la valeur critique (VC) à lire
8  % selon la taille n
9  if (n==50)
10     C = 1;
11 elseif (n==100)
12     C = 2;
13 elseif (n==250)
14     C = 3;
15 end
16
17 % Détermination du fichier à ouvrir selon le modèle sous H0 à tester
18 if (modelHo == 1)
19     a = 'VC_Clayton.txt';
20 elseif (modelHo == 3)
21     a = 'VC_AMH.txt';
22 elseif (modelHo == 4)
23     a = 'VC_Gumbel.txt';
24 elseif (modelHo == 5)
25     a = 'VC_Frank.txt';
26 elseif (modelHo == 7)
27     a = 'VC_FGM.txt';
28 elseif (modelHo == 11)
29     a = 'VC_Normale.txt';
30 end
31
32 % Ouverture du fichier de valeurs critiques à lire
33 fp1 = fopen(a, 'r');
34
35 % Création de l'espace mémoire pour l'enregistrement des valeurs critiques
36 VC = zeros(10,3);
37
38 % Enregistrement des valeurs critiques dans un vecteur
39 for i=1:10
40     for j=1:3
41         VC(i,j) = fscanf(fp1, '%f ', 1);
42     end
43 end
44
45 % Fermeture du fichier de valeurs critiques
46 fclose(fp1);
47
48 % Valeur initiale de la puissance
49 pow = 0;
50

```



```

51 % Boucle du nombre d'évaluations de la statistique de test par rapport à la VC
52 for k=1:IT
53
54     % Simulation de 2 processus AR(1) avec une dépendance selon une
55     % certaine copule sous H1
56     X = sim_copula_ts(n,modelH1,tauH1,rhoX,sigmaX,rhoY,sigmaY);
57
58     % Calcul de la statistique de test et du tau de kendall à partir du
59     % modèle sous H0
60     [Tn,kendall] = CvM_TS(X,modelHo,N);
61
62     % Calcul de la position de la VC à retenir dans le vecteur en fonction
63     % du tau de Kendall
64     L = max(int16(10*kendall),0) + 1;
65
66     % Évaluation de la statistique de test par rapport à la VC dans la
67     % table
68     if ( Tn > VC(L,C))
69         pow = pow + 1;
70     end
71 end
72
73 % Estimation de la puissance en fonction du nombre d'itérations
74 pow = pow / IT;

```

A.4 Procédures utilitaires

A.4.1 Vraisemblance d'un processus AR(1)

```
1 function L = likelihood_AR(X,A)
2
3 % Calcul de la taille du vecteur X
4 a = size(X); n = a(1);
5
6 % Calcul de la log-vraisemblance d'une série autorégressive d'ordre un
7 L = 0;
8 for t=2:n
9     L = L - log(A(2)) + log(normpdf((X(t) - A(1)*X(t-1))/A(2)));
10 end
```

A.4.2 Vraisemblance de processus AR(1) liés par une copule

```

1 function L = Likelihood_copula(X,model,P,theta)
2
3 a = size(X); n = a(1);
4
5 L = 1;
6
7 %Clayton
8 if (model == 1)
9     if (theta==0)
10         L = 0;
11     else
12         for t=2:n
13             u = normcdf((X(t,1) - P(1)*X(t-1,1))/P(2));
14             v = normcdf((X(t,2) - P(3)*X(t-1,2))/P(4));
15             L = L + log(theta+1) - (2+(1/theta))*log(u^(-theta)+v^(-theta)-1) - (theta+1)*
16         end
17     end
18 end
19
20 %Ali-Mikhail-Haq
21 if (model == 3)
22     if (theta <=0)
23         L = 0;
24     else
25         for t=2:n
26             u = normcdf((X(t,1) - P(1)*X(t-1,1))/P(2));
27             v = normcdf((X(t,2) - P(3)*X(t-1,2))/P(4));
28             if (theta >1)
29                 L = L + log(2*u*v) - 3*log(v-u*v);
30             else
31                 L = L + log(1 + theta*u*v + theta*v - 2*theta + theta^2*u*v + theta^2 + th
32             end
33         end
34     end
35 end
36
37 %Gumbel-Hougaard
38 if (model == 4)
39     if (theta==0)
40         L = 0;
41     else
42         for t=2:n
43             u = normcdf((X(t,1) - P(1)*X(t-1,1))/P(2));
44             v = normcdf((X(t,2) - P(3)*X(t-1,2))/P(4));
45             L = 0;
46         end
47     end
48 end

```

```

49 %Frank
50 if (model == 5)
51     if (theta==0)
52         L = 0;
53     else
54         for t=2:n
55             u = normcdf((X(t,1) - P(1)*X(t-1,1))/P(2));
56             v = normcdf((X(t,2) - P(3)*X(t-1,2))/P(4));
57             L = L + log(theta) - theta*u - theta*v + log(exp(-theta)-1) - 2*log(-exp(-thet.
58         end
59     end
60 end
61
62 %Farlie-Gumbel-Morgenstern
63 if (model == 7)
64     if (theta==0)
65         L = 0;
66     else
67         for t=2:n
68             u = normcdf((X(t,1) - P(1)*X(t-1,1))/P(2));
69             v = normcdf((X(t,2) - P(3)*X(t-1,2))/P(4));
70             if (theta>1)
71                 L = L + log(1 + (1-2*u)*(1-2*v));
72             else
73                 L = L + log(1 + theta*(1-2*u)*(1-2*v));
74             end
75         end
76     end
77 end
78
79 %Normal
80 if (model == 11)
81     if (theta==0)
82         L = 0;
83     else
84         for t=2:n
85             u = normcdf((X(t,1) - P(1)*X(t-1,1))/P(2));
86             v = normcdf((X(t,2) - P(3)*X(t-1,2))/P(4));
87             x = norminv(u); y = norminv(v);
88             L = L - (1/2)*log(1-theta^2) - (x^2+y^2-2*theta*x*y)/(2*(1-theta^2));
89         end
90     end
91 end

```

A.4.3 Tau de Kendall

```
1 function t = tau(X)
2
3 % Calcul de la taille du vecteur bivarié X
4 a = size(X); n = a(1); d = a(2);
5
6 % Algorithme estimant la valeur du tau de Kendall
7 c = 0;
8 for i=1:n
9     for j=1:n
10         mark = 0;
11         for k=1:d
12             if ( X(i,k) < X(j,k) )
13                 mark = mark + 1;
14             end
15             if ( mark==d )
16                 c = c+1;
17             end
18         end
19     end
20 end
21
22 b = c/(n*(n-1));
23
24 t = (2^d*b-1)/(2^(d-1)-1);
```

A.4.4 Calcul de rangs

```
1 function R = rankit(X)
2
3 % Calcul de la taille du vecteur X
4 a = size(X); n = a(1);
5
6 % Création de l'espace mémoire pour le vecteur X
7 R = zeros(n,1);
8
9 % Calcul du rang des Xi pour X1 à Xn
10 for i=1:n
11     for j=1:n
12         if X(j) <= X(i)
13             R(i) = R(i) + 1;
14         end
15     end
16 end
```

A.4.5 Inversion du tau de Kendall

```

1 function theta = tau_theta(model,tau)
2
3 %Estimation du paramètre d'une copule par inversion du tau de Kendall
4
5 %Clayton
6 if (model == 1)
7     if tau == 1
8         theta = 1000;
9     else
10        theta = 2*tau/(1-tau);
11    end
12
13 %Ali-Mikhail-Haq
14 elseif (model == 3)
15     if (tau <= 0.0)
16         theta = 0;
17     elseif (tau >= 1/3)
18         theta = 1;
19     else
20         s = 0.00001; v = 0.9; v1 = 0.1;
21         while ( abs(v-v1) > s )
22             t2 = (3*v^2-2*v-2*(1-v)^2*log(1-v))/(3*v^2) - tau;
23             t1 = (3*v1^2-2*v1-2*(1-v1)^2*log(1-v1))/(3*v1^2) - tau;
24             v2 = v1 - t1*(v1-v)/(t1-t2);
25             if (v2 <= 0.0)
26                 v2 = 0.001;
27             elseif (v2 >= 1)
28                 v2 = 0.999;
29             end
30             v = v1;
31             v1 = v2;
32             if ( abs(v1) > 100000 )
33                 v = tau;
34                 v1 = tau-0.01;
35             end
36         end
37         theta = v1;
38     end
39
40 %Gumbel
41 elseif (model == 4)
42     if (tau <= 0)
43         theta = 0;
44     else
45         theta = tau;
46     end
47
48
49
50

```

```
51 %Frank
52 elseif (model == 5)
53     fp1 = fopen('tau_Frank.txt', 'r');
54     b2 = -1; t2 = 0;
55     while (tau > b2)
56         t1 = t2;
57         b1 = b2;
58         a = fscanf(fp1, '%f %f', 2);
59         b2 = a(1); t2 = a(2);
60     end
61     fclose(fp1);
62     theta = ((t1 - t2)*tau + (b1*t2 - b2*t1))/(b1 - b2);
63
64 %Farlie-Gumbel-Morgenstern
65 elseif (model == 7)
66     if (tau < -2/9)
67         theta = -1;
68     elseif (tau > 2/9)
69         theta = 1;
70     else
71         theta = 9*tau/2;
72     end
73
74 %Normal
75 elseif (model == 11)
76     theta = sin((pi/2)*tau);
77
78 end
```