

DEPARTEMENT DES MATHÉMATIQUES

Master Mathématique et Application au Calcul Scientifique
(MACS)

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du Diplôme de Master Sciences et Techniques
(MST)

**Le Processus ARMA Localement
Stationnaire avec des innovations α -
stable**

Réalisé par: EL FAHFOUHI Hanae

Encadré par: CHAIBI Ghizlane

Soutenu le 16 juillet 2021

Devant le jury composé de:

-Pr. EL KHOMSSI Mohammed

Examineur

-Pr. HILALI Abdelmajid

Examineur

-Pr. EL AYADI Rachid

Examineur

-Pr. CHAIBI Ghizlane

Encadrante

Année Universitaire 2020 / 2021

FACULTE DES SCIENCES ET TECHNIQUES FES – SAISS

✉ B.P. 2202 – Route d'Imouzzer – FES

☎ 212 (0)5 35 61 16 86 – Fax : 212 (0)5 35 60 82 14

Site web : <http://www.fst-usmba.ac.ma>

Remerciements

Au terme de ce travail, je rends louanges à Allah pour la volonté, la santé, la patience qui m'a donné durant toutes les années d'études et le pouvoir de parachever ce travail.

Mes remerciements et ma profonde gratitude vont à l'endroit à tous ceux qui ont contribué de prêt ou de loin à la réalisation de ce projet en particulier :

Madame CHAIBI Ghizlane mon encadrante qui a accepté en toute modestie de m'accompagner et de me fournir un encadrement adéquat tout au long de ce projet, pour ses remarques et ses suggestions qui ont fait que ce rapport a vu ce jour. Elle m'a apporté énormément de connaissances mathématiques, elle a été toujours à l'écoute de mes questions, et s'est toujours intéressée à l'avancée de mon travail. Cela qui m'a permis d'enrichir mes connaissances dans ce domaine, et de m'approcher du monde des mathématiques et de la recherche scientifique.

Je tiens également à exprimer toute ma gratitude et tout mon sentiment de reconnaissance aux professeurs EL KHOMSSI Mohammed, HILALI Abdelmajid et EL AYADI Rachid pour l'honneur qu'ils me font d'accepter d'être membres de jury de ce mémoire. Un grand merci à Monsieur Coulibaly Bakary D qui n'a pas cessé de me donner des conseils et de me guider durant la réalisation de ce projet.

J'adresse mes sincères remerciements et ma grande reconnaissance à tous mes professeurs de Master Mathématique et Applications aux Calculs Scientifiques.

Je n'oublie pas de remercier ma famille, mes amis et surtout mes parents, pour leurs sacrifices et soutiens inconditionnels durant mes études.

Table des matières

Remerciements	2
Annexes des notations	9
Résumé	10
Introduction générale	11
1 Généralités sur les séries chronologiques	14
Introduction	14
1.1 Abstraction	14
1.2 Processus stochastiques	15
1.2.1 Processus stationnaires	15
1.2.2 Processus linéaires	16
1.2.3 Fonctions d'auto-covariation et d'auto-corrélation	18
1.2.4 Estimations des fonctions caractéristiques	20
1.3 Modélisation des séries stationnaires	22
1.3.1 Processus moyenne mobile	22
1.3.2 Processus autorégressif	25
1.3.3 Les processus mixtes ARMA	30
1.4 Méthode de Box-Jenkins	34
1.4.1 Transformation	34
1.4.2 Identification du P et du Q du ARMA(P,Q)	34
1.4.3 Estimation des paramètres des modèles sélectionnés	34
1.4.4 Validation	34
1.4.5 Choix d'un modèle parmi plusieurs	35
Conclusion	35
2 Les distributions α-stable	36
Introduction	36
2.1 Caractérisation des lois α -stable	36
2.1.1 Quelques définitions et propriétés	36
2.2 Comportement des queues et moments	40
2.3 Fonctions de densité et de distribution stables	41
2.4 Algorithme de simulation d'une variable aléatoire α -stable	42

Conclusion	46
3 Méthodes d'estimation et Tests statistiques	47
Introduction	47
3.1 Le maximum de Vraisemblance	47
3.1.1 Fonction de vraisemblance	47
3.1.2 Estimateurs du maximum de vraisemblance pour (x_1, \dots, x_n)	48
3.1.3 Fonction de Log-vraisemblance	48
3.1.4 Équation de vraisemblance	48
3.1.5 Expression analytique de MLE	48
3.1.6 Pratique de MLE	49
3.2 Méthode de Whittle	50
3.2.1 L'ajustement de modèles autorégressifs variant dans le temps	51
3.3 Inférence indirecte	53
3.3.1 Le principe de l'approche	54
3.4 La méthode de Monte-Carlo	55
3.5 Les tests d'adéquation	55
3.5.1 Les coefficients Skewness et Kurtosis	56
3.5.2 Test Jarque-Bera (ou test de Bowman-Shelton)	57
3.5.3 Test de Shapiro-Wilk	58
3.5.4 Diagramme Q-Q plot	59
3.5.5 Box plot ou Boite à moustache	60
3.6 L'analyse de résidus	63
3.6.1 MAE (Erreur absolue moyenne)	63
3.6.2 MSE (erreur quadratique moyenne)	63
3.6.3 RMSE (Racine carré de l'erreur quadratique moyenne)	64
Conclusion	64
4 Les Processus tvARMA avec des innovations α-stable	65
Introduction	65
4.1 Les processus localement stationnaire	65
4.2 tvARMA avec des innovations stables	67
4.2.1 L'existence et l'unicité d'une solution	68
4.2.2 Localement stationnaire	72
4.3 Simulation	75
4.3.1 le cas de α connu	76
4.3.2 Le cas α inconnu	80
4.4 Application	86
Conclusion	92
Conclusion générale et perspectives	93
Bibliographie	95

Table des figures

3.1	Distributions avec différentes valeurs d'asymétrie et d'aplatissement	56
3.2	Test de Jarque-Bera	57
3.3	Diagramme QQ-plot	60
3.4	Boite à moustache	61
3.5	Boite à moustache de la série	62
4.1	Estimations de θ_0, θ_1 et γ pour différentes tailles d'échantillon, sur la base de $R = 1000$ réplifications de tvAR(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.9, 0.9, -0.3, 0.8, 1)$, en utilisant l'inférence indirecte.	77
4.2	Estimations de θ_0, θ_1 et γ pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications de tvMA(1) α -stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.1, -0.2, 0.35, -0.6, 1.2)$ en utilisant l'inférence indirecte.	78
4.3	MC moyenne et erreur standard pour différentes tailles d'échantillon T , en utilisant des estimateurs indirects et BWE en supposant que α et β sont connus, à partir de tvARMA(1,1) stable avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.8, 0.3, -0.4, 0.1, 0.1, 0.3, 1)$, sur la base de $R = 1000$ réplifications.	79
4.4	Kurtosis et asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T , en supposant que α et β connus à partir de α -stable tvARMA(1,1) avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.8, 0.3, -0.4, 0.1, 0.1, 0.3, 1)$, sur la base de $R = 1000$ réplifications.	80
4.5	Estimations de $\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}$ et γ pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications de tvARMA(1,1) stable avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.8, 0.3, -0.4, 0.1, 0.1, 0.3, 1)$ en utilisant l'inférence indirecte.	80
4.6	Estimations de $\theta_0, \theta_1, \alpha, \gamma_0$ et γ_1 pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications pour tvAR(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, -0.6, 0.5, 0.1)$ en utilisant l'inférence indirecte.	82
4.7	Estimations de θ_0, θ_1 , pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications de tvMA(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$ en utilisant l'inférence indirecte.	83
4.8	Estimations de la densité de $\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha$ et γ pour différentes tailles d'échantillon basées sur $R = 1000$ réplifications de tvARMA(1,1) α -stable avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.3, 0, -0.2, -0.4, 0.2, 0.3, 1.1)$ en utilisant l'inférence indirecte.	85

4.9	Évolution de l'énergie éolienne fonction du temps	86
4.10	Évolution de la série stationnaire de l'énergie éolienne fonction du temps	87
4.11	Graphe des Auto-corrélations	87
4.12	Graphe des Auto-corrélations partielles	88
4.13	Graphe des Auto-corrélogrammes partiels des résidus	89
4.14	Variogramme des résidus	90
4.15	Variogramme de la différenciation	91
4.16	Histogramme résiduel et Boite à moustache	91

Liste des tableaux

2.1	Moyenne et variance empirique calculées sur 5000 réalisations	46
3.1	Les notes de 24 élèves d'une classe lors d'un examen noté sur 100 points	61
3.2	Tableau trié	61
4.1	Les moyennes MC et les erreurs standard pour différentes tailles d'échantillon T, en utilisant des estimateurs indirects et BWE en supposant que α et β sont connus, à partir de tvAR(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.9, 0.9, 0.3, 0.8, 1)$, basé sur $R = 1000$ répétitions.	76
4.2	Le Kurtosis et l'asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T, en supposant que α et β sont connus, pour α -stable tvAR(1) avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.9, 0.9, -0.3, 0.8, 1)$, basé sur $R = 1000$ répétitions.	77
4.3	Les moyennes MC et l'erreur standard pour différentes tailles d'échantillon T, en utilisant des estimateurs indirects et BWE en supposant que α et β sont connus, pour tvMA(1) α -stable avec $(\alpha, \theta_0, \theta_1, \gamma) = (1.1, 0.2, 0.35, -0.6, 1.2)$, basé sur $R = 1000$ répétitions	78
4.4	Kurtosis et asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T en supposant que α et β connus à partir de tvMA(1) α -stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.1, 0.2, 0.35, 0.6, 1.2)$ basé sur $R = 1000$ répétitions.	78
4.5	La moyenne MC et erreur type pour différentes tailles d'échantillon T à l'aide d'estimateurs indirects en supposant $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, 0.6, 0.5, 0.1)$ avec β connu pour tvAR(1) α -stable basé sur $R = 1000$ répétitions	81
4.6	L'amplitude et l'asymétrie des estimations indirectes pour différentes tailles d'échantillon T en supposant que $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, -0.6, 0.5, 0.1)$ avec β connu pour tvAR(1) α -stable sur la base de $R = 1000$ répétitions.	81
4.7	La moyenne MC et erreur standard pour différentes tailles d'échantillon T en utilisant des estimateurs indirects en supposant $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$ avec β connu à partir de α -stable tvMA(1) basé sur $R = 1000$ répétitions.	82
4.8	Kurtosis et asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T en supposant que $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$ avec β connu pour tvMA(1) α -stable basé sur $R = 1000$ répétitions.	83

4.9	La moyenne MC et l'erreur standard pour différentes tailles d'échantillon T en utilisant des estimateurs indirects en supposant $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha, \gamma) = (1.3, 0, 0.2, 0.4, 0.2, 0.3, 1.1)$ avec β connu pour tvARMA(1,1) α -stable basé sur $R = 1000$ réplifications	84
4.10	Le Kurtosis et skewness des estimations indirectes et BWE pour différentes tailles d'échantillon ($T = 500, 1000, 1500$) en supposant que $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha, \gamma) = (1.3, 0, -0.2, -0.4, 0.2, 0.3, 1.1)$ avec β connu à partir de α tvARMA(1,1) stable basé sur $R = 1000$ réplifications.	84
4.11	Erreurs résiduelles des différents modèles	89

Annexes des notations

Notations	Significations
ACF	La fonction d'auto-corrélation
$PACF$	La fonction d'auto-corrélation partielle
$sign(.)$	Fonction signe
MC	La méthode de Monte Carlo
BWE	la méthode de whittle par block
$Cov(X, Y)$	La covariance de deux variables aléatoires X et Y
$E(X)$	L'espérance de la variable aléatoire X
$Var(X)$	la variance de la variable aléatoire X
$X \equiv S_\alpha(\beta, \sigma, \mu)$	variable aléatoire de loi stable de paramètres α, β, σ et μ
α	paramètre de stabilité ou Exposant caractéristique d'une loi stable
β	Paramètre d'asymétrie d'une loi stable
σ	paramètre du Facteur d'échelle d'une loi stable
μ	Paramètre de localisation d'une loi stable
$S_\alpha S$	Loi α -stable symétrique
$\stackrel{d}{=}$	A la même distribution que
$p.s$	presque sûrement
\approx	Approximation
\sim	suit une loi
$AR(p)$	Processus autorégressif d'ordre p
$MA(q)$	processus moyenne mobile (Moving Average) d'ordre q
$ARMA(p, q)$	Processus auto-régressif moyenne mobile d'ordre (p, q)
$tvARMA$	Processus auto-régressif moyenne mobile d'ordre (p, q) avec des coefficients variant dans le temps
$\mathcal{N}(\mu, \sigma)$	Loi normale μ et de variance σ^2
$i.i.d$	Indépendantes et identiquement distribuées.
\bar{X}	La moyenne empirique
ε_t	Le bruit blanc
IM	Le modèle d'intérêt
AM	Le modèle auxiliaire
MLE	Maximum de vraisemblance
$[\cdot]_{11}$	L'élément 11
$[\cdot]$	la partie entière
MAE	Erreur absolue moyenne
MSE	Erreur quadratique moyenne
$RMSE$	Racine carré de l'erreur quadratique moyenne
AIC	Critère d'information d'Akaike
BIC	Critère d'information Bayésienne

Résumé

Ce travail s'inspire d'un article existant intitulé "Indirect inference for locally stationary ARMA processes with stable innovations" des deux mathématiciens Shu Wei Chou-Chen & Pedro A. Morettin basé sur des processus ARMA Localement Stationnaires qui font partie de la classe des modèles de prévisions pour les séries chronologiques.

L'idée consiste à étudier les processus tvARMA dans lesquels les coefficients varient en fonction du temps avec des innovations α -stable. Pour cela nous commençons notre étude par le cas où les coefficients sont constants et en faisant quelques rappels sur les lois α -stables ainsi que certaines méthodes d'estimations à savoir : la méthode de Whittle et la méthode de l'inférence indirecte qui sont dérivées de la méthode de maximum de vraisemblance. Cette dernière ne fonctionne pas puisque les lois α -stable n'ont pas une formule générale de densité .

Ensuite, la partie simulation traite les modèles de type tvARMA dans les cas où α connu et inconnu. En terminant avec une application suivant la méthodologie de Box & Jenkins pour la recherche d'un modèle tvARMA optimal.

Introduction générale

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.

L'analyse des séries chronologiques occupe une place prépondérante dans l'étude des phénomènes aléatoires, à savoir l'économie, l'automatique, la météorologie, l'hydrologie... Celle-ci rassemble un ensemble de techniques et de méthodes mathématiques permettant la construction de modèles stochastiques à travers des séries de données indexée dans le temps. Cette construction se fait généralement selon trois étapes essentielles : l'identification du modèle paramétrique, l'estimation de ses paramètres et les tests permettant la validation de cette construction.

Le modèle ainsi obtenu peut être exploité entre autres dans la prévision ou le contrôle. Bien qu'ils ne représentent souvent qu'une approximation brutale de la réalité, les modèles stationnaires (ou ceux qui peuvent être rendus stationnaires) tels que les modèles ARMA, ARIMA, SARIMA, etc., sont les plus répandus et les plus utilisés dans l'analyse des séries temporelles. Nous intéressons au processus ARMA.

Un exemple populaire est le modèle ARMA (Auto Regressive Moving Average), pour prévoir le PIB Marocaine en 2020 par exemple, il ne s'agit plus d'utiliser un modèle structurel qui explique le PIB (exemple par l'intermédiaire de la consommation, de l'investissement, des dépenses publiques et du solde commercial etc.) et de projeter ensuite les tendances passées. Mais avec le modèle ARMA, il s'agit de prévoir le PIB en 2020 en exploitant les propriétés statistiques du PIB (moyenne, variance etc.) . Ce modèle utilise souvent des valeurs retardées du PIB (d'où le terme Auto Regressive) et de chocs aléatoires qui sont en général de moyenne nulle, de variance constante et non auto-corrélés (bruit blanc), quand la variable qui représente ces chocs est retardée, on parle (de moyenne mobile).

George Box et Gwilym Jenkins sont deux statisticiens qui ont contribué, dans les années

1970, à populariser la théorie des séries temporelles univariées. Les procédures de modélisation sont présentées dans leur célèbre ouvrage intitulé " Time Series Analysis : Forecasting and control ". Ils ont proposé une démarche générale de prévision pour les séries chronologiques. Cette démarche est fondée sur la notion de processus ARMA et elle comprend quatre phases : l'identification a priori, l'estimation du modèle ARMA identifié, la validation et la prévision.

La majorité des processus étudiés ont été supposés stationnaires en raison de leurs propriétés importantes telles que l'invariance de la moyenne et de la variance et la structure de covariance en fonction de la différence de temps. Cependant, la plupart des données de séries temporelles du monde réel ne sont pas stationnaires et l'ajustement d'un processus stationnaire à une série temporelle non stationnaire peut être inapproprié et conduit généralement à des conclusions erronées.

Pour traiter la non-stationnarité, il existe quelques techniques bien connues qui convertissent les séries temporelles non stationnaires à des séries stationnaires, comme la différenciation, la suppression de la tendance et l'analyse de régression basée sur d'autres variables d'entrée.

Une façon de généraliser le processus stationnaire est l'idée de stationnarité locale. C'est le cas où un processus stochastique X_t peut être stationnaire sur de petites périodes de temps [3], [4] .

La plupart des résultats des processus localement stationnaires supposent des innovations avec un second moment fini. Cependant, différentes zones ont observé des phénomènes avec des distributions de queue lourdes ou une variance infinie. La classe des processus localement stationnaires est considérée comme des outils clés de la généralisation de pas mal des méthodes comme la méthode de Whittle .

La généralisation due à Cramér (1961), de la décomposition de Wold (1938), aux processus non stationnaires, a permis d'élargir la classe des modèles linéaires autorégressifs moyenne mobile, à coefficients constants (ARMA), à la classe des modèles linéaires autorégressifs moyenne mobile à coefficients dépendant du temps (tvARMA).

Dans ce travail, nous considérons que les innovations du processus localement stationnaire suivent des distributions α -stables . L'avantage de supposer des distributions α -stables est sa flexibilité pour l'asymétrie et les queues épaisses. de plus , il est fermé sous des combinaisons linéaires et inclut la distribution gaussienne comme cas particulier. Cependant, son estimation est difficile puisque la fonction de densité n'a pas de forme fermée. Par conséquent, les méthodes d'estimation habituelles telles que le maximum de vraisemblance et la méthode des moments ne fonctionnent pas.

La stratégie pour estimer ce type de processus est la méthode d'inférence indirecte proposée par Gourieroux et al. [12] et Gallant et Tauchen [11]. Comme les distributions α -stable qui peuvent être facilement simulées, l'approche indirecte, qui est une méthode intensive basée sur la simulation informatique, peut être une solution pour surmonter le problème d'estimation .

Ce mémoire est partagé en quatre axes principaux à savoir :

Le premier chapitre consacré à donner quelques rappels des processus stochastiques en particulier ARMA(p,q) et afin de traiter ce type de processus avec la méthodologie de Box & Jenkins .

dans **le deuxième chapitre** nous introduisons les lois α -stable avec certaines de ses propriétés.

Le troisième chapitre regroupe des outils indispensables pour la simulation ainsi que dans la procédure de la méthode de Box-Jenkins, c'est à dire ce chapitre a été conçu aux méthodes d'estimations en particuliers .

En fin nous terminons ce travail avec **un quatrième chapitre** présente les processus localement en particulier tvARMA et leurs propriétés .Dans ce chapitre nous traitons une application des modèles tvARMA avec des innovations α -stable pour faire la prévision de l'énergie éolienne en utilisant la méthode de Box & Jenkins pour la sélection d'un modèle optimal. Nous clôturons ce travail par une conclusion générale et perspectives.

Généralités sur les séries chronologiques

Introduction

L'étude des séries temporelles, ou séries chronologiques, correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Elles ont été utilisées en astronomie ('on the periodicity of sunspots', 1906), en météorologie ('time-series regression of sea level on weather', 1968), en théorie du signal ('Noise in FM receivers', 1963), en biologie ('the autocorrelation curves of schizophrenic brain waves and the power spectrum', 1960), en économie ('time-series analysis of imports, exports and other economic variables', 1971) ...etc.

L'objectif de l'étude de ce chapitre est de faire quelques rappels sur les aspects d'une série temporelle de façon générale qui seront très utiles dans la suite de notre travail.

1.1 Abstraction

La théorie des séries chronologiques (ou temporelles) est appliquée de nos jours dans des domaines variés. nous nous intéressons à l'évolution au cours du temps d'un phénomène, dans le but de décrire, expliquer puis prévoir ce phénomène dans le futur. On dispose ainsi d'observations à des dates différentes, c'est à dire d'une suite de valeurs numériques indicées par le temps.

Exemple : On peut songer par exemple à l'évolution du nombre de voyageurs utilisant le train, à l'accroissement relatif mensuel de l'indice des prix ou encore à l'occurrence d'un phénomène naturel (comme le nombre de taches solaires).

Cette suite d'observations d'une famille de variables aléatoires réelles notées $(X_t)_{t \in \Theta}$ est appelée **série chronologique** (ou temporelle). où l'ensemble Θ est appelé espace des temps qui peut être :

- **discret** (nombre de voyageurs SNCF quotidien, température maximale...). Dans ce cas, $\Theta \subset \mathbb{Z}$ (le plus souvent $\Theta = \mathbb{Z}$).
- **continu** (signal radio, résultat d'un électrocardiogramme...). L'indice de temps est à valeurs dans un intervalle de \mathbb{R} et on dispose (au moins potentiellement) d'une infinité d'observations issues d'un processus $(X_t)_{t \in \Theta}$ où Θ est un intervalle de \mathbb{R} .

Définition 1.1.1. On dit que la série admet une tendance (ou trend) si on peut écrire $X_t = f(t) + \varepsilon_t$ avec f une fonction fixée et (ε_t) des bruits aléatoires.

Si $f(t) = \alpha t + \beta$, on dit que la tendance est linéaire, plus généralement, si $X_t = \sum_{i=0}^p \alpha_i t^i$, on dit que la tendance est polynomiale.

Si $f(t)$ est périodique, on dit que la tendance est périodique.

Si $f(t) = s(t) + \alpha t + \beta$ avec s une fonction périodique on dit que la série a une tendance linéaire et une composante périodique (saisonnrière).

1.2 Processus stochastiques

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Un processus stochastique X est une famille de variables aléatoires réelles $(X_t)_{t \in \Theta}$, où Θ est appelé l'espace des temps.

Définition 1.2.1. Bruit blanc

Un processus de bruit blanc est une suite de variables aléatoires (X_t) , d'espérance et de variance constante (c'est à dire $E(X_t)$ et $V(X_t)$ ne dépendent pas de t) et tel que $\text{Cov}(X_t, X_{t+h}) = 0$ si $h \neq 0$.

- Si l'espérance est nulle, le bruit blanc est dit centré.
- Si les variables sont gaussiennes, le bruit blanc est dit gaussien.

Bruits blancs faible et fort

On appelle bruit blanc faible tout processus $\varepsilon = (\varepsilon_t)_{t \in \mathbb{Z}}$ tel que

$$\mathbb{E}(\varepsilon_t) = 0 \text{ et } \mathbb{E}(\varepsilon_t \varepsilon_{t'}) = \sigma^2 \delta_{tt'}, \quad t, t' \in \mathbb{Z}$$

où $\delta_{tt'}$ est le symbole de Kronecker

$$\delta_{tt'} = \begin{cases} 1 & \text{si } t = t' \\ 0 & \text{si } t \neq t' \end{cases}$$

On dira qu'il s'agit d'un bruit blanc fort s'il est faible de plus les ε_t sont indépendants et identiquement distribués (i.i.d)

1.2.1 Processus stationnaires

Définition 1.2.2. Un processus $X = (X_t)_{t \in \mathbb{Z}}$ est stationnaire du second ordre (ou faiblement stationnaire) si et seulement si :

i) $\mathbb{E}(X_t) = \mu, \forall t \in \mathbb{Z}$

ii) X_t est de carré intégrable pour tout $t \in \mathbb{Z} : \mathbb{E}(X_t^2) < \infty ;$

iii) $\text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_{t-1}, X_{t-1+h}) = \dots = \text{Cov}(X_0, X_h), \forall t, h \in \mathbb{Z}$

Exemple 1.2.1. *Un bruit blanc est un processus stationnaire.*

Exemple 1.2.2. *Un exemple de processus non stationnaire : la marche aléatoire. Le processus $(X_t)_{t \in \mathbb{N}}$ est une marche aléatoire (random walk en anglais) lorsque :*

$$X_t = X_{t-1} + \varepsilon_t, \quad t \in \mathbb{N}^*$$

où $(\varepsilon_t)_{t \in \mathbb{N}^*}$ est une suite de variables indépendantes et identiquement distribuées de loi $\mathcal{N}(0, \sigma^2)$.

Preuve 1.2.1. *Le processus $(X_t, t \in \mathbb{Z})$ peut se réécrire sous la forme :*

$$\begin{aligned} X_t &= \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \dots + \varepsilon_{t-j} + \dots \\ &= \sum_{j=0}^{\infty} \varepsilon_{t-j} \end{aligned}$$

Dès lors, connaissant les propriétés du bruit blanc ε_t , on montre que :

$$E(X_t) = E\left(\sum_{j=0}^{\infty} \varepsilon_{t-j}\right) = \sum_{j=0}^{\infty} E(\varepsilon_{t-j}) = 0$$

Donc le processus $X_t, \forall t \in \mathbb{Z}$ a une espérance nulle et donc satisfait la première condition de la définition de la stationnarité. Mais il ne satisfait pas la seconde condition puisque :

$$\text{Var}(X_t) = \text{Var}\left(\sum_{j=0}^{\infty} \varepsilon_{t-j}\right) = \sum_{j=0}^{\infty} \text{Var}(\varepsilon_{t-j}) = \sum_{j=0}^{\infty} \sigma_\varepsilon^2 \equiv \infty$$

La variance de X_t est non finie. De plus, si l'on avait définie une condition initiale X_0 , alors la variance de X_t , définie par $\text{Var}(X_t) = \sum_{j=0}^{t-1} \sigma_\varepsilon^2 = t\sigma_\varepsilon^2$ serait fonction de t .

La troisième condition de la stationnarité faible est alors violée. Le processus $(X_t, t \in \mathbb{Z})$ est donc un processus non stationnaire.

Définition 1.2.3. *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus de carré intégrable et X_t^* la régression affine de X_t sur $(X_s, s < t)$. On appellera **innovation** du processus à la date t la variable aléatoire réelle : $\varepsilon_t^* := X_t - X_t^*$.*

Exemple 1.2.3. *Lorsque $(X_t)_{t \in \mathbb{Z}}$ est stationnaire, la suite des innovations $(\varepsilon_t^*)_{t \in \mathbb{Z}} := (X_t - X_t^*)_{t \in \mathbb{Z}}$ est un bruit blanc.*

1.2.2 Processus linéaires

Un processus linéaire est un processus stochastique X_t formé par une combinaison linéaire (non nécessairement finie) de bruits blancs forts. On définit également la classe des processus linéaires généraux, qui sont constitués de combinaisons linéaires de bruits blancs faibles. Introduisons formellement ces deux types de processus.

Définition 1.2.4. $(X_t)_{t \in \mathbb{Z}}$ est un **processus linéaire** (resp. **linéaire général**) de moyenne μ s'il peut être écrit sous la forme :

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

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est **un bruit blanc fort** (resp. **faible**), de variance σ^2 et où la suite des coefficients b_i est supposée telle que

$$\sum_{i=-\infty}^{+\infty} b_i^2 < +\infty$$

Théorème 1.2.1. (Wold)

Pour tout processus stationnaire centré X , il existe (de manière unique) un processus purement innovant (ou nul) Y et un processus purement déterministe η non corrélés [1], tels que :

$$X = Y + \eta$$

De plus, la partie purement innovante s'exprime (de manière unique) comme moyenne mobile en fonction du bruit blanc d'innovation ε de X :

$$Y_t = \sum_{j \geq 0} c_j \varepsilon_{t-j} \quad \forall t$$

avec $c_0 = 1$ et $\sum_{j \geq 0} c_j^2 < \infty$. Le bruit blanc d'innovation de X est également bruit blanc d'innovation pour Y .

Définition 1.2.5. 1. Un processus $(X_t)_{t \in \mathbb{Z}}$ admet **une représentation inversible** s'il peut s'écrire comme combinaison linéaire des valeurs d'un autre processus, c'est-à-dire qu'il existe une suite $(\psi_i, i \in \mathbb{Z})$ et un processus $(Y_t)_{t \in \mathbb{Z}}$ tels que

$$\forall t \in \mathbb{Z}, \quad X_t = \sum_{i \in \mathbb{Z}} \psi_i Y_{t-i}$$

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

$$\forall t \in \mathbb{Z}, \quad X_t = \sum_{i \in \mathbb{N}} \psi_i Y_{t-i}$$

Théorème 1.2.2. Soit $(X_t, t \in \mathbb{Z})$ le processus stationnaire solution de l'équation suivante :

$$\forall t \in \mathbb{Z}, \quad \Phi(B)X_t = Y_t$$

où $(Y_t, t \in \mathbb{Z})$ est un processus stationnaire et $\Phi(B)$ le polynôme en B de la forme

$$\Phi(B) = I - \phi_1 B - \dots - \phi_p B^p$$

avec $p \in \mathbb{N}^*$ et $BX_t = X_{t-1}, \forall t \in \mathbb{Z}$ (B : appelé l'opérateur de retard)

Alors :

1. Si Φ n'a pas de racine de module égal à 1, alors il existe une représentation inversible du processus $(X_t, t \in \mathbb{Z})$, c'est-à-dire qu'il existe une suite $(\psi_i, i \in \mathbb{Z})$ telle que

$$\forall t \in \mathbb{Z}, \quad X_t = \sum_{i \in \mathbb{Z}} \psi_i Y_{t-i}$$

2. Si de plus toutes les racines de Φ sont de module supérieur à 1, alors il existe une représentation causale du processus $(X_t, t \in \mathbb{Z})$, c'est-à-dire qu'il existe une suite $(\psi_i, i \in \mathbb{Z})$ telle que

$$\forall t \in \mathbb{Z}, \quad X_t = \sum_{i \in \mathbb{N}} \psi_i Y_{t-i}$$

Autrement dit, dans le premier cas, on garantit que X_t puisse s'écrire en fonction du processus $(Y_t, t \in \mathbb{Z})$ et dans le second cas, que X_t puisse s'écrire en fonction seulement du passé de Y_t .

1.2.3 Fonctions d'auto-covariation et d'auto-corrélation

La fonction d'auto-covariance

Pour un processus $(X_t)_{t \geq 0}$ stationnaire, on définit la fonction d'auto-covariance par :

$$h \in \mathbb{Z} \rightarrow \gamma(h) = \text{Cov}(X_t, X_{t+h})$$

Le graphe de cette fonction est appelé variogramme.

La densité spectrale

(i) la fonction d'auto-covariance $\gamma(h)$ d'un processus stationnaire peut être écrit sous la forme

$$\gamma(h) = \int_{-\pi}^{+\pi} \exp(i\omega h) dF(\omega)$$

où $F(\omega)$ est une fonction de répartition,

(ii) tout processus stationnaire peut se mettre sous la forme $X_t = \int_{-\pi}^{+\pi} \exp(i\omega t) dz(\omega)$ où $z(\omega)$ est une fonction aléatoire, complexe, à accroissements non corrélés. Cette représentation est appelée représentation de Cramér.

Définition 1.2.6. Soit (X_t) un processus stationnaire de fonction d'autocovariance $\gamma(\cdot)$, la densité spectrale de (X_t) s'écrit

$$f(\omega) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma(h) \exp(i\omega h)$$

Proposition 1.2.1. Réciproquement, si $f(\cdot)$ est la densité spectrale de (X_t) alors

$$\gamma(h) = \int_{-\pi}^{+\pi} f(\omega) \exp(i\omega h) d\omega$$

La fonction d'auto-corrélation

La fonction d'auto-corrélation par :

$$h \in \mathbb{N}^* \rightarrow \rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

Le graphe de cette fonction est appelé *corrélogramme*.

Propriété 1.2.1. La fonction d'auto-covariance d'un processus stationnaire vérifie :

1. elle est paire ; $\gamma(-h) = \gamma(h), \forall h \in \mathbb{Z}$.

2. elle est de type positif ;

$$\forall (a_i) \in \mathbb{R}^N, \forall (t_i) \in \mathbb{Z}^N, \quad \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(t_i - t_j) > 0, \forall n \in \mathbb{N}.$$

3. $\gamma(0) = \text{Var}(X_t)$;

4. $|\gamma(h)| \leq \gamma(0), \forall h$.

Preuve 1.2.2. 1. *Parité :*

$$\gamma(h) = \text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_{t-h}, X_{(t-h)+h}) = \text{Cov}(X_{t-h}, X_t) = \gamma(-h)$$

2. *Positivité :*

$$\begin{aligned} \text{Var}\left(\sum_i a_i X_{t_i}\right) &= \text{Cov}\left(\sum_i a_i X_{t_i}, \sum_j a_j X_{t_j}\right) \\ &= \sum_{i,j} a_i a_j \text{Cov}(X_{t_i}, X_{t_j}) \\ &= \sum_{i,j} a_i a_j \gamma(t_i - t_j) \geq 0 \end{aligned}$$

On suppose toujours qu'il n'y a pas de relations linéaires entre les X_t . En effet, si on avait $\text{Var}(\sum_i a_i X_{t_i}) = 0$ alors $\sum_i a_i X_{t_i} = \text{constante}$ presque sûrement.

3.

$$\gamma(0) = \text{Cov}(X_t, X_t) = \text{Var}(X_t)$$

4. *L'inégalité de Cauchy-Schwarz appliquée à $\gamma(k)$ implique*

$$|\gamma(h)| = |\mathbb{E}[(X_{t+h} - \mu_X)(X_t - \mu_X)^*]| \leq \sqrt{\mathbb{E}[|X_{t+h} - \mu_X|^2] \mathbb{E}[|X_t - \mu_X|^2]} = \gamma(0)$$

la dernière inégalité découlant de l'hypothèse de stationnarité ($\text{Cov}(X_{t+h}, X_{t+h}) = \text{Cov}(X_t, X_t)$)

Propriété 1.2.2. La fonction d'autocorrélation d'un processus stationnaire vérifie :

1. elle est paire ; $\rho(-h) = \rho(h), \forall h \in \mathbb{Z}$,

2. $\rho(0) = 1$

3. $|\rho(h)| \leq 1, \forall h \in \mathbb{Z}$.

La fonction $\rho(h)$ est l'expression du lien linéaire entre X_t et X_{t-h} . Si t est l'instant présent et $h > 0$, $\rho(h)$ est l'expression du lien linéaire entre le présent et le passé d'ordre h . Par parité de la fonction γ entre le présent et le futur d'ordre h ; plus $|\rho(h)|$ est proche de 1 et plus ce lien est fort.

La fonction d'auto-corrélation partielle

La d'auto-corrélation partielle d'ordre h d'un processus stationnaire se calcule de la manière suivante :

$$\tau(h) = \frac{|R(h)^*|}{|R(h)|}$$

avec

$$R(h) = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ \vdots & & \ddots & \\ \rho_{h-1} & \rho_{h-2} & \dots & 1 \end{pmatrix}$$

$R(h)^*$ est la matrice $R(h)$ dans laquelle on a remplacé la colonne h par $\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_h \end{pmatrix}$

$$R(h)^* = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & & \ddots & \\ \rho_{h-1} & \rho_{h-2} & \dots & \rho_h \end{pmatrix}$$

1.2.4 Estimations des fonctions caractéristiques

la moyenne empirique

Définition 1.2.7. Soit $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire de moyenne μ

($\mu = \mathbb{E}(X_t) \forall t$), de fonctions d'auto-corrélation ρ et d'auto-corrélation partielle τ . On dispose en pratique de l'observation du processus jusqu'à l'instant T de X_1, \dots, X_T .

On estime généralement la moyenne du processus par

$$\bar{X} := \frac{1}{T} \sum_{i=1}^T X_i$$

Propriété 1.2.3. .

1. \bar{X} est un estimateur sans biais de μ : $\mathbb{E}(\bar{X}) = \mu$

2. $\text{Var}(\bar{X}) = \frac{1}{T} \sum_{|h| < T} \left(1 - \frac{|h|}{T}\right) \gamma(h)$.

Preuve 1.2.3. .

1.

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{1}{T} \sum_{i=1}^T X_i\right) = \frac{1}{T} \sum_{i=1}^T \mathbb{E}(X_i) = \frac{1}{T} \sum_{i=1}^T \mu = \mu$$

2.

$$\begin{aligned} \text{Var}(\bar{X}) &= \frac{1}{T^2} \sum_{i=1}^T \sum_{j=1}^T \text{Cov}(X_i, X_j) \\ &= \frac{1}{T^2} \sum_{i-j=-n}^n (n - |i - j|) \gamma(i - j) \\ &= \frac{1}{T} \sum_{h=-T}^T \left(1 - \frac{|h|}{T}\right) \gamma(h) \end{aligned}$$

Estimation de la fonction d'auto-covariance

Pour construire un estimateur de la fonction d'auto-covariance, rappelons que si X_1, \dots, X_T sont des observations de même loi de X et Y_1, \dots, Y_T sont des observations de même loi de Y sont i.i.d. de variance finie, un estimateur de la covariance entre X et Y est donné par :

$$\frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})$$

On estime donc naturellement la fonction d'auto-covariance par

$$\hat{\gamma}(h) := \frac{1}{T - |h|} \sum_{t=1}^{T-|h|} (X_t - \bar{X})(X_{t+|h|} - \bar{X})$$

Il est appelé *l'auto-covariance empirique*.

L'estimation de la densité spectrale

Pour estimer la densité spectrale de $\{X_t\}$, il est naturel de s'intéresser au périodogramme, défini comme le module au carré de la transformée de Fourier discrète des observations $\{X_1, X_2, \dots, X_T\}$:

$$I_T(x) = \frac{1}{2\pi T} \left| \sum_{t=1}^T X_t \exp(itx) \right|^2 = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \hat{\gamma}_T(h) \exp(i\omega x) \quad (1.1)$$

Estimation des fonctions d'auto-corrélation

On estime ensuite naturellement la fonction d'auto-corrélation à partir de **l'auto-covariance empirique** de la façon suivante

$$\hat{\rho}(h) := \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{T-|h|} (X_t - \bar{X})(X_{t+|h|} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

Il est appelé l'auto-corrélation empirique.

Propriété 1.2.4. .

1. $\hat{\rho}(h)$ est de type positif.
2. $\forall h \in \mathbb{Z}, -1 \leq \hat{\rho}(h) \leq 1$.

1.3 Modélisation des séries stationnaires

De façon générale, quand on parle de séries stationnaires, on a en tête une représentation de la forme X_t , où $t \in \mathbb{Z}$, représentant les observations du processus, dont on peut définir un ensemble d'auto-covariance $\gamma(t; s) = E([X_t - m][X_s - m])$ qui ne dépend que de la distance entre t et s , $\gamma(t; s) = \gamma(t + h; s + h)$ pour tout $h \in \mathbb{Z}$ (notion 'faible' de stationnarité). On demande généralement à cette auto-covariance $\gamma(t; s)$ de tendre vers 0 quand la différence entre t et s tend vers l'infini : la covariance entre des éléments très éloignés dans la série tend vers 0.

Cette approche, basée sur l'utilisation des corrélations, correspond à l'analyse de type 'temporelle' : elle consiste à étudier les corrélations croisées de fonctions de la série (X_t). Ces méthodes sont généralement paramétriques de type moyenne-mobiles (moving average MA) ou autorégressives (AR) - voire les deux (ARMA). Toutes ces méthodes consistent à estimer des paramètres peuvent généralement être vus comme des généralisations de la régression linéaire.

1.3.1 Processus moyenne mobile

Ces processus forment une classe flexible de modèles pour de nombreux phénomènes observés. Ils sont construits à partir de l'idée que l'observation au temps t s'explique linéairement par les observations d'un bruit blanc ; ils sont donc définis par :

Définition 1.3.1. On appelle **processus moyenne mobile** [Moving Average] d'ordre q tout processus $(X_t)_{t \in \mathbb{Z}}$ stationnaire tel que :

$$\forall t \in \mathbb{Z}, \quad X_t = \eta_t - \theta_1 \eta_{t-1} - \dots - \theta_q \eta_{t-q}$$

où $\theta_1, \theta_2, \dots, \theta_q$ sont des réels fixés et $(\eta_t)_{t \in \mathbb{Z}}$ est un bruit blanc de variance σ^2 .

Posons : $\Theta(B) = I - \theta_1 B - \dots - \theta_q B^q$, on peut écrire :

$$X_t = \Theta(B)\eta_t, \quad t \in \mathbb{Z}$$

On suppose bien évidemment que q est inférieur au nombre d'observations.

L'observation au temps t est donc la somme d'un choc aléatoire η_t à l'instant t et d'une fonction linéaire du passé de ce choc $-\sum_{i=1}^q \theta_i \eta_{t-i}$.

Théorème 1.3.1. .

1. Si Θ n'a pas de racine de module égal à 1, alors il existe un choix unique de coefficients π_j pour lesquels on a

$$\eta_t = \sum_{j \in \mathbb{Z}} \pi_j X_{t-j}$$

avec $\sum_{i=-\infty}^{+\infty} |\pi_i| < \infty$. Dans ce cas, on dit que le processus est **inversible**.

2. Si de plus les racines de Θ sont toutes de module strictement supérieur à 1, alors il existe un choix unique de coefficients π_j pour lesquels on a :

$$\eta_t = \sum_{j \in \mathbb{N}} \pi_j X_{t-j}$$

Les coefficients π_j convergent rapidement vers 0 lorsque $j \rightarrow \infty$, en effet ce sont les coefficients du filtre associé au polynôme $1/\Theta$. Dans ce cas, on dit que le processus est **causal** et η_t est l'innovation du processus à la date t . Dans le cas où Θ n'a pas de racine multiple, on a la formule explicite suivante :

$$\forall j \geq 0, \quad \pi_j = \sum_{1 \leq k \leq q} m_k \mu_k^j$$

avec

$$m_k = \frac{1}{\prod_{l \neq k} \left(1 - \frac{\mu_l}{\mu_k}\right)}$$

et

$$\Theta(z) = 1 - \theta_1 z - \dots - \theta_q z^q = \prod_j (1 - \mu_j z)$$

Par contre, si Θ s'annule sur le cercle unité, on a le théorème suivant :

Théorème 1.3.2. Si Θ s'annule en un point du cercle unité, alors il n'existe aucun choix de coefficients π_j pour lesquels $\sum_{j \geq 0} \pi_j X_{t-j}$ converge et coïncide avec η_t . Cependant, on peut représenter η_t comme limite de telles combinaisons linéaires des X_u , $u \leq t$.

Bruit blanc d'innovation d'un MA(q)

Comme nous l'avons dit précédemment, le bruit blanc d'innovation $(\varepsilon_t)_{t \in \mathbb{Z}}$ du processus MA(q) $(X_t)_{t \in \mathbb{Z}}$ n'est pas nécessairement le processus $(\eta_t)_{t \in \mathbb{Z}}$ c'est-à-dire que η_t ne représente pas l'information ajoutée au passé pour obtenir la valeur présente de X_t . Mais nous avons le résultat suivant

Théorème 1.3.3. *La relation d'un MA(q) avec son bruit blanc d'innovation est aussi du type $X_t = \tilde{\Theta}(B)\varepsilon_t$, $t \in \mathbb{Z}$. Le polynôme $\tilde{\Theta}$ de degré q , dit polynôme canonique, s'obtient à partir de Θ en remplaçant toutes les racines éventuelles à l'intérieur du cercle unité par leur inversion $(\mu \mapsto \frac{1}{\bar{\mu}})$, les déplaçant ainsi à l'extérieur du cercle unité. Notons μ_1, \dots, μ_q les co-racines de $\Theta(z)$:*

$$\Theta(z) = \prod_j (1 - \mu_j z) = 1 - \theta_1 z - \dots - \theta_q z^q$$

Le polynôme $\tilde{\Theta}$ qui relie X à son bruit blanc d'innovation est donné par

$$\tilde{\Theta}(z) = \prod_{j, |\mu_j| \leq 1} (1 - \mu_j z) \prod_{|\mu_j| > 1} \left(1 - \frac{1}{\bar{\mu}_j} z\right)$$

Proposition 1.3.1. *La variance du bruit blanc d'innovation ε est donnée par :*

$$\text{Var}(\varepsilon_t) = \gamma_\varepsilon(0) = \gamma_\eta(0) \prod_{j, |\mu_j| > 1} |\mu_j|^2$$

A partir de maintenant, nous ne considérerons plus que des processus MA(q) donnés sous leur représentation canonique i.e. vérifiant l'équation

$$X_t = \Theta(B)\eta_t$$

avec $\Theta(z) = 1 - \theta_1 z - \dots - \theta_q z^q = \prod_{i=1}^q (1 - \mu_i z)$ où $|\mu_i| \leq 1 \forall i$ puisque nous venons de voir qu'il est toujours possible de se ramener à ce cas quitte à changer de bruit blanc.

Proposition 1.3.2. .

1. La variance de X_t est donnée par

$$\text{Var}(X_t) = \gamma^X(0) = (1 + \theta_1^2 + \dots + \theta_q^2) \gamma^\eta(0) > \gamma^\eta(0) = \sigma^2, \quad \forall t$$

Ayant de plus $\mathbb{E}(X_t) = 0$, on en déduit que tout processus moyenne mobile est un processus stationnaire.

2. Plus généralement, la fonction d'auto-covariance est donnée par

$$\gamma(h) = \begin{cases} (-\theta_h + \theta_{h+1}\theta_1 + \dots + \theta_q\theta_{q-h}) \sigma^2 & \text{si } 0 < |h| \leq q \\ 0 & \text{si } |h| > q \end{cases}$$

3. La fonction d'auto-corrélation est donnée par

$$\rho(h) = \begin{cases} \frac{-\theta_h + \theta_{h+1}\theta_1 + \dots + \theta_q\theta_{q-h}}{1 + \theta_1^2 + \dots + \theta_q^2} & \text{si } 0 < |h| \leq q \\ 0 & \text{si } |h| > q \end{cases}$$

4. L'expression de la fonction d'auto-corrélation partielle est compliquée. Notons simplement qu'elle ne s'annule pas à partir d'un certain rang mais qu'il existe un nombre $r \in]0, 1[$ tel que

$$|\tau(h)| \leq r^h, \quad h \geq 2$$

En particulier, $\rho(q) = -\theta_q / (1 + \theta_1^2 + \dots + \theta_q^2)$ ne s'annule que si $\theta_q = 0$, ce qui revient à dire que le processus n'est pas d'ordre q .

Insistons sur le fait que si $h > q$ la fonction d'autocorrélation d'un processus MA(q) s'annule. Cette observation sera utile pour la modélisation : si à partir de données X_1, \dots, X_T , la fonction d'autocorrélation empirique n'est pas significativement différente de zéro au-delà d'un certain nombre q_0 , on sera alors guidé pour choisir d'ajuster un modèle MA(q_0) aux données.

Proposition 1.3.3. Soit (X_t) un processus linéaire stationnaire corrélé d'ordre q , c'est à dire dont $\gamma(h) = 0$ pour tout $|h| > q$. Alors (X_t) possède une représentation comme processus MA(q).

Prédiction d'un MA(q)

Soit X un MA(q) de polynôme Φ et de bruit blanc d'innovation η (on suppose que η est causal). La relation qui le lie à son bruit blanc d'innovation η :

$$X_t = \eta_t - \theta_1 \eta_{t-1} - \dots - \theta_q \eta_{t-q}$$

On a donc

Théorème 1.3.4. Les prévisions de X_t sont :

$$\begin{aligned} \hat{X}_t(1) &= -\theta_1 \eta_t - \dots - \theta_q \eta_{t-(q-1)} \\ \hat{X}_t(2) &= -\theta_2 \eta_t - \dots - \theta_q \eta_{t-(q-2)} \\ &\dots \\ \hat{X}_t(q) &= -\theta_q \eta_t \\ \hat{X}_t(j) &= 0, \quad j > q \end{aligned}$$

Les MA(q) ne sont pas bien adaptés à la prévision puisque d'une part $j > q \Rightarrow \hat{X}_t(j) = 0$ et d'autre part, les prévisions nécessitent la combinaison de toutes les valeurs du passé de X .

1.3.2 Processus autorégressif

Ces processus forment une classe flexible de modèles pour de nombreux phénomènes observés. Ils sont construits à partir de l'idée que l'observation au temps t s'explique linéairement par les observations précédentes ; ils sont donc définis implicitement :

Définition 1.3.2. On appelle **processus autorégressif** d'ordre p tout processus $(X_t)_{t \in \mathbb{Z}}$ stationnaire tel que

$$\forall t \in \mathbb{Z}, \quad X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \eta_t \quad (1.2)$$

où $\phi_1, \phi_2, \dots, \phi_p$ sont des réels fixés et $(\eta_t)_{t \in \mathbb{Z}}$ est un bruit blanc de variance σ^2 .

Posons $\Phi(B) = I - \phi_1 B - \dots - \phi_p B^p$, on peut écrire :

$$\Phi(B)X_t = \eta_t, \quad t \in \mathbb{Z} \quad (1.3)$$

L'observation au temps t est donc la somme d'un choc aléatoire η_t à l'instant t (indépendant de l'historique) et d'une fonction linéaire de son passé $\sum_{i=1}^p \phi_i X_{t-i}$ (qui peut-être vue comme la prédiction de X_t à partir des p dernières valeurs observées. Nous reviendrons en détail sur cette propriété dans la section consacrée à la prédiction).

Contrairement aux processus MA, leur définition pose quelques problèmes :

- de leur définition ne découle pas naturellement la stationnarité. C'est pourquoi nous ajoutons l'hypothèse de stationnarité dans la définition suivante.
- ils sont définis de manière implicite. Il va donc s'agir de déterminer une forme explicite.

Proposition 1.3.4. *En prenant l'espérance de 1.2 et en utilisant la stationnarité du processus, on obtient que l'espérance μ du processus vérifie $\mu(1 - \sum_{i=1}^p \phi_i) = 0$. Donc lorsque 1 n'est pas racine de Φ , on a nécessairement $\mu = 0$.*

Écriture moyenne mobile infinie d'un AR(p)

Dans le cas régulier où le polynôme Φ ne s'annule pas sur le cercle unité, on a le théorème

Théorème 1.3.5. *i) Si Φ n'a pas de racine de module égal à 1, $\Phi(B)$ est inversible et on en déduit que l'équation (1.2) a une solution unique, avec une écriture moyenne mobile infinie*

$$\forall t \in \mathbb{Z}, \quad X_t = (\Phi(B))^{-1} \eta_t = \sum_{j \in \mathbb{Z}} \psi_j \eta_{t-j}$$

avec $\sum_{i=-\infty}^{+\infty} |\psi_i| < \infty$. Les coefficients ψ_j convergent rapidement vers 0 lorsque $j \rightarrow \infty$, en effet ce sont les coefficients du filtre associé au polynôme $1/\Phi$.

Ainsi le processus X est bien stationnaire, déterminé de manière unique par la relation précédente et la valeur présente de X dépend à la fois du passé, du présent et du futur du bruit blanc.

ii) Si de plus les racines de Φ sont toutes de module strictement supérieur à 1, alors l'opérateur inverse $\Phi^{-1}(B)$ admet un développement ne faisant intervenir que les puissances positives de B :

$$\forall t \in \mathbb{Z}, \quad X_t = (\Phi(B))^{-1} \eta_t = \sum_{i=0}^{+\infty} \psi_i \eta_{t-i}$$

avec $\sum_{i=0}^{+\infty} |\psi_i| < +\infty$. C'est-à-dire que le processus X_t s'exprime en fonction de $\eta_s, s \leq t$ et d'après la définition, on voit que η_t n'est pas corrélé avec X_{t-1}, X_{t-2}, \dots . La variable η_t est donc l'innovation du processus à la date t et $\sum_{i=0}^{+\infty} \psi_i \eta_{t-i}$ est la régression affine de X_t sur $(X_s, s \leq t-1)$; il s'agit de la représentation canonique d'un processus AR(p).

Bruit blanc d'innovation d'un AR(p)

Comme pour les MA, le bruit blanc η n'est pas nécessairement le bruit blanc d'innovation.

Théorème 1.3.6. *Si X est un AR(p) alors X est purement innovant et il existe un unique polynôme $\tilde{\Phi}$ (dit canonique) tel que le bruit blanc d'innovation de X soit de la forme $\varepsilon = \tilde{\Phi}(B)X$. Ce polynôme est de degré p et s'obtient à partir de Φ par la même règle que pour les MA(q). Il est donc de la forme $\tilde{\Phi}(z) = \prod_{1 \leq k \leq p} (1 - \mu_k z)$ avec $\forall k |\mu_k| < 1$*

Fonctions caractéristiques d'un AR(p)

Proposition 1.3.5. Soit X un AR(p) de polynôme générateur Φ et de bruit blanc η .

1. La fonction d'auto-covariance d'un AR(p) est donnée par

$$\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i)$$

2. La fonction d'auto-corrélation se déduit aisément $\rho(h) = \sum_{i=1}^p \phi_i \rho(h-i)$.

3. La fonction d'auto-corrélation partielle est nulle pour $|h| > p$ et vaut ϕ_p pour $|h| = p$.

Prédiction d'un AR(p)

Les AR(p) sont bien adaptés à la prédiction. Les prévisions optimales $\widehat{X}_t(h)$ pour $h > 0$ sont des combinaisons linéaires de $X_t, X_{t-1}, \dots, X_{t-p+1}$ comme d'habitude. On les calcule de façon récurrente.

A l'horizon 1 : on a

$$X_{t+1} = \phi_1 X_t + \phi_2 X_{t-1} + \dots + \phi_p X_{t+1-p} + \eta_{t+1}$$

que l'on prédit par

$$\widehat{X}_t(1) = \phi_1 X_t + \phi_2 X_{t-1} + \dots + \phi_p X_{t+1-p}$$

A l'horizon 2 : on a

$$X_{t+2} = \phi_1 X_{t+1} + \phi_2 X_t + \dots + \phi_p X_{t+2-p} + \eta_{t+2}$$

que l'on prédit par

$$\widehat{X}_t(2) = \phi_1 \widehat{X}_t(1) + \phi_2 X_t + \dots + \phi_p X_{t+2-p}$$

et ainsi de suite. Ainsi on a le théorème suivant :

Théorème 1.3.7. Soit X un AR(p) de bruit blanc d'innovation η et de polynôme canonique $\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$. Les prévisions optimales $\widehat{X}_t(h)$ pour $h > 0$ sont des combinaisons linéaires de $X_t, X_{t-1}, \dots, X_{t-p+1}$:

$$\widehat{X}_t(h) = \sum_{0 \leq j \leq p-1} a_j(h) X_{t-j}$$

qui s'obtiennent par récurrence selon

$$\widehat{X}_t(h) = \sum_{1 \leq k \leq p} \phi_k \widehat{X}_t(h-k)$$

avec les conditions initiales $\widehat{X}_t(-j) = X_{t-j}$ pour $j \geq 0$

On peut donner des formules explicites pour les coefficients $a_j(h)$ mais c'est un peu compliqué, surtout si Φ a des racines multiples. Du point de vue du calcul par un ordinateur, il est en général plus simple, et tout aussi efficace, d'évaluer $\widehat{X}_t(h)$ par récurrence comme ci-dessus. Lorsque Φ n'a pas de racines multiples, on peut aussi procéder selon la remarque suivante.

Remarque 1.3.1. Dans le cas des processus autorégressifs causaux, le bruit blanc coïncide avec l'erreur de prédiction un pas en avant. En effet, dans ce cas,

$$\begin{aligned}\widehat{X}_t(1) &= \sum_{1 \leq k \leq p} \phi_k \widehat{X}_t(1-k) \\ &= \phi_1 \widehat{X}_t(0) + \phi_2 \widehat{X}_t(-1) + \dots + \phi_p \widehat{X}_t(1-p) \\ &= \phi_1 X_t + \phi_2 X_{t-1} + \dots + \phi_p X_{t-p} \\ &= X_{t+1} - \eta_{t+1}\end{aligned}$$

D'où l'on déduit que $X_{t+1} - \widehat{X}_t(1) = \eta_{t+1}$.

Proposition 1.3.6. Soit X un $AR(p)$ de bruit blanc d'innovation η et de polynôme canonique $\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p = \prod_{1 \leq k \leq p} (1 - \lambda_k z)$, sans racine multiple. Soit $\Phi_k(z) = \prod_{j \neq k} (1 - \lambda_j z)$ et $Y^k = \Phi_k(B)X$. Alors les Y^k sont des $AR(1)$ qui partagent avec X le même bruit blanc d'innovation. On a

$$\widehat{X}_t(h) = \sum_{1 \leq k \leq p} m_k \lambda_k^h Y_t^k$$

avec

$$m_k = \frac{1}{\prod_{j \neq k} \left(1 - \frac{\lambda_j}{\lambda_k}\right)}$$

Equations de Yule-Walker d'un $AR(p)$

Soit X un $AR(p)$ causal de bruit blanc d'innovation η et de polynôme canonique associé $\Phi(x) = 1 - \phi_1 x - \dots - \phi_p x^p$

Pour des raisons théoriques, on veut pouvoir calculer les autocovariances de X lorsque les ϕ_k et $\sigma^2 = \gamma^\eta(0)$ sont connus.

Pour des raisons pratiques, on veut pouvoir calculer ϕ_k et $\gamma^\eta(0)$ lorsque les autocovariances de X sont connues. Dans le cas d'un processus autorégressif, on peut utiliser les équations de **Yule-Walker**.

Etudions le cas particulier des processus $AR(2)$.

Exemple 1.3.1. Soit le processus $AR(2)$ de la forme :

$$\forall t \in \mathbb{Z}, \quad X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = \eta_t \quad (1.4)$$

On le suppose de plus centré et causal comme usuellement. En multipliant l'équation (1.4) par X_{t-1} et X_{t-2} puis en prenant l'espérance de ces équations et enfin en divisant par $\gamma(0)$, on obtient les équations de **Yule-Walker** :

$$\begin{pmatrix} \rho(1) \\ \rho(2) \end{pmatrix} = \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

On obtient donc la solution suivante :

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \frac{1}{1 - \rho(1)^2} \begin{pmatrix} \rho(1)(1 - \rho(2)) \\ \rho(2) - \rho(1)^2 \end{pmatrix}$$

Quant à la variance, partant de l'équation (1.4), on obtient que :

$$X_t^2 - \phi_1 X_{t-1} X_t - \phi_2 X_{t-2} X_t = \eta_t X_t$$

et par passage à l'espérance

$$\gamma(0) - \phi_1 \gamma(1) - \phi_2 \gamma(2) = \mathbb{E}(\eta_t X_t)$$

Mais

$$\begin{aligned} \mathbb{E}(\eta_t X_t) &= \mathbb{E}(\eta_t (\eta_t + \phi_1 X_{t-1} + \phi_2 X_{t-2})) \\ &= \sigma^2 + \phi_1 \mathbb{E}(\eta_t X_{t-1}) + \phi_2 \mathbb{E}(\eta_t X_{t-2}) \\ &= \sigma^2 \end{aligned}$$

D'où $\sigma^2 = \gamma(0) - \phi_1 \gamma(1) - \phi_2 \gamma(2)$. Comme on a des estimateurs convergents de la fonction d'autocovariance et donc aussi de la fonction d'autocorrélation, on pourra en déduire des estimateurs convergents de ϕ_1, ϕ_2 et σ^2 .

Sur cet exemple, on voit donc bien comment dériver les équations de Yule-Walker pour un AR(2). En procédant de façon analogue, on obtient le théorème suivant

Théorème 1.3.8. (Yule-Walker) On a les équations suivantes :

$$\begin{aligned} (YW_0) \quad & \gamma^X(0) - \phi_1 \gamma^X(-1) - \dots - \phi_p \gamma^X(-p) = \gamma^\eta(0) \\ (YW_1) \quad & \gamma^X(1) - \phi_1 \gamma^X(0) - \dots - \phi_p \gamma^X(1-p) = 0 \\ & \dots \\ (YW_j) \quad & \gamma^X(j) - \phi_1 \gamma^X(j-1) - \dots - \phi_p \gamma^X(j-p) = 0 \end{aligned}$$

dites **équations de Yule-Walker**.

Les équations de Yule-Walker peuvent ensuite être utilisées pour l'estimation des paramètres d'un processus autorégressif. Notons $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)$ les estimateurs (dits de Yule-Walker) de $\phi = (\phi_1, \dots, \phi_p)$ et $\hat{\sigma}^2$ l'estimateur de σ^2 par cette méthode. Le théorème suivant donne la normalité asymptotique de ces estimateurs :

Théorème 1.3.9. Si $(X_t)_{t \in \mathbb{Z}}$ est un processus AR(p) avec toutes ses racines de module supérieur à 1, alors :

$$\begin{aligned} \sqrt{T}(\hat{\phi} - \phi) &\underset{d, \rightarrow +\infty}{\longrightarrow} \mathcal{N}(0, \sigma^2 \Gamma_p^{-1}), \\ \text{avec : } \Gamma_p &= (\gamma(i-j))_{(i,j) \in [1,p]^2} \text{ De plus, } \hat{\sigma}^2 \underset{\mathbb{P}, T \rightarrow \infty}{\longrightarrow} \sigma^2. \end{aligned}$$

Note Rappelons que $\gamma^X(-j) = \gamma^X(j)$

Théorème 1.3.10. Etant donnés ϕ_1, \dots, ϕ_p et $\gamma^\eta(0)$, le système des $p+1$ équations linéaires en les $p+1$ inconnues $\gamma^X(0), \dots, \gamma^X(p)$ possède une unique solution. Les $\gamma^X(j)$ pour $j > p$ sont alors calculés par récurrence avec les $YW_j, j > p$.

1.3.3 Les processus mixtes ARMA

Les processus AR et MA ont des caractéristiques qui se révèlent grâce à leurs fonctions d'auto-corrélations et leurs fonctions d'auto-corrélations partielles. Pour un processus AR, nous avons vu que la fonction d'auto-corrélation partielle possède un point de rupture après un certain nombre d'écart; ce dernier détermine l'ordre du polynôme AR. Pour un processus MA, nous avons vu que c'est la fonction d'auto-corrélation qui possède un point de rupture après un certain nombre d'écart; ce dernier détermine l'ordre du polynôme MA. Cependant pour certains processus, ni la fonction d'auto-corrélation, ni la fonction d'auto-corrélation partielle ne possèdent de point de rupture. Dans de tels cas, il faut construire un modèle mixte. Nous définissons dans cette section les séries ARMA qui sont des généralisations directes des deux exemples introductifs, la combinaison des processus auto-régressifs et moyennes mobiles. Cette classe de processus ARMA est encore un cas particulier de processus linéaires et jouera un rôle prépondérant dans la modélisation concrète des processus stationnaires. Elle présente l'avantage d'être plus souple à l'utilisation et de fournir généralement de bonnes approximations des séries réelles avec moins de paramètres que les modèles purs.

Définition d'un ARMA(p,q)

Définition 1.3.3. processus auto-régressif moyenne mobile

On appelle processus auto-régressif moyenne mobile d'ordre (p, q) (AutoRegressive Moving Average of order (p, q)) tout processus $(X_t)_{t \in \mathbb{Z}}$ stationnaire tel que :

$$\forall t \in \mathbb{Z}, \quad X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \eta_t - \theta_1 \eta_{t-1} - \dots - \theta_q \eta_{t-q} \quad (1.5)$$

où $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ sont des réels fixés et $(\eta_t)_{t \in \mathbb{Z}}$ est un bruit blanc de variance σ^2 .

Posons comme précédemment :

$$\Phi(B) = I - \phi_1 B - \dots - \phi_p B^p$$

et

$$\Theta(B) = I - \theta_1 B - \dots - \theta_q B^q.$$

On peut alors écrire :

$$\forall t \in \mathbb{Z}, \quad \Phi(B)X_t = \Theta(B)\eta_t$$

Propriété 1.3.1. (i) Si le polynôme Φ a toutes ses racines de module différent de 1, l'opérateur $\Phi(B)$ est inversible et la relation admet une solution stationnaire donnée par

$$\forall t \in \mathbb{Z}, X_t = \frac{\Theta(B)}{\Phi(B)} \eta_t = \sum_{i \in \mathbb{Z}} \psi_i \eta_{t-i} \text{ avec } \sum_{i \in \mathbb{Z}} |\psi_i| < +\infty$$

c'est à dire sous une forme moyenne mobile infinie.

(ii) Si de plus les racines du polynôme Φ sont de module strictement supérieur à 1, seules les valeurs présente et passées du bruit interviennent dans cette écriture $MA(\infty)$. Dans ce cas, les ψ_i de la représentation causale vérifient $\psi_0 = 1$ et

$$\psi_k = \begin{cases} -\theta_k + \sum_{i=1}^k \phi_i \psi_{k-i} & \text{pour } 1 \leq k < \max(p, q+1) \\ \sum_{i=1}^p \phi_i \psi_{k-i} & \text{pour } k \geq \max(p, q+1) \end{cases}$$

en posant $\theta_0 = -1, \theta_i = 0$ si $i > q$ et $\phi_i = 0$ si $i > p$. Ces équations peuvent se résoudre successivement pour $\psi_0, \psi_1 \dots$. Ainsi

$$\begin{aligned}\psi_0 &= 1 \\ \psi_1 &= -\theta_1 + \psi_0\phi_1 = -\theta_1 + \phi_1 \\ \psi_2 &= -\theta_2 + \psi_0\phi_2 + \psi_1\phi_1 = -\theta_2 + \phi_2 - \theta_1\phi_1 + \theta_1^2 \\ &\dots\end{aligned}$$

Une démarche analogue peut-être suivie pour le polynôme Θ :

Propriété 1.3.2. (i) Supposons que le polynôme Θ a toutes ses racines de module différent de 1, l'opérateur $\Theta(B)$ est inversible et on obtient la forme autorégressive infinie

$$\forall t \in \mathbb{Z}, \quad \eta_t = \frac{\Phi(B)}{\Theta(B)}X_t = \sum_{i \in \mathbb{Z}} \pi_i X_{t-i} \text{ avec } \sum_{i \in \mathbb{Z}} |\pi_i| < +\infty$$

(ii) Si de plus les racines du polynôme Θ sont de module strictement supérieur à 1, cette représentation $AR(\infty)$ ne fait intervenir que les valeurs présente et passées du processus. Dans ce cas, les π_i de la représentation causale vérifient $\pi_0 = 1$ et

$$\pi_k = \begin{cases} -\phi_k + \sum_{i=1}^k \theta_i \pi_{k-i} & \text{pour } 1 \leq k < \max(p+1, q) \\ \sum_{i=1}^q \theta_i \pi_{k-i} & \text{pour } k \geq \max(p+1, q) \end{cases}$$

Bruit blanc d'innovation d'un ARMA (p, q)

Comme pour les AR et les MA, le bruit blanc η n'est pas nécessairement le bruit blanc d'innovation :

Théorème 1.3.11. Pour que η soit le bruit blanc d'innovation de X il faut et il suffit que

$$\Phi(z) = \prod_{1 \leq k \leq p} (1 - \lambda_k z) \text{ et } \Theta(z) = \prod_{1 \leq j \leq q} (1 - \mu_j z)$$

avec $\forall k, \quad |\lambda_k| < 1$ et $\forall j, \quad |\mu_j| \leq 1$.

Théorème 1.3.12. Si X est un ARMA (p, q) donné selon $\Phi(B)X = \Theta(B)\eta$, alors la relation (dite minimale) qui le lie à son bruit blanc d'innovation est aussi du type $\tilde{\Phi}(B)X = \tilde{\Theta}(B)\varepsilon$, ou les polynômes $\tilde{\Phi}$ et $\tilde{\Theta}$ s'obtiennent à partir de Φ et Θ selon les mêmes règles que pour les AR (p) et les MA (q) puis en supprimant tout facteur commun éventuel.

A partir de maintenant, nous ne considérerons plus que des processus ARMA (p, q) donnés sous leur représentation canonique i.e. vérifiant l'équation

$$\Phi(B)X_t = \Theta(B)\eta_t$$

avec

$$\begin{aligned}\Phi(z) &= 1 - \phi_1 z - \dots - \phi_p z^p = \prod_{i=1}^p (1 - \lambda_i z) \text{ où } |\lambda_i| \leq 1 \forall i \\ \Theta(z) &= 1 - \theta_1 z - \dots - \theta_q z^q = \prod_{i=1}^q (1 - \mu_i z) \text{ où } |\mu_i| \leq 1 \forall i.\end{aligned}$$

puisque nous venons de voir qu'il est toujours possible de se ramener à ce cas quitte à changer de bruit blanc.

Proposition 1.3.7. *La fonction d'auto-covariance d'un processus ARMA(p, q) vérifie pour $h \geq \max(p, q+1)$*

$$\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i) \quad (1.6)$$

et pour $0 \leq h < \max(p, q+1)$,

$$\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i) + \sigma^2 (\psi_0 - \psi_1 \theta_{h+1} - \dots - \psi_{q-h} \theta_q) \quad (1.7)$$

Proposition 1.3.8. *La somme d'un ARMA(p_1, q_1) et d'un ARMA(p_2, q_2) (non corrélés) est un ARMA($\leq \max(p_1, p_2), \leq \max(q_1, q_2)$)*

Prévision des modèles ARMA(p, q)

Supposons que l'on dispose d'un grand nombre d'observations consécutives de X jusqu'à l'instant T . Une fois le modèle choisi et ses paramètres estimés, il va être possible de faire de la prévision. Pour estimer les prévisions optimales $\hat{X}_T(h)$, il faut disposer :

- ⊗ des coefficients ψ_j qui lient X à son bruit blanc d'innovation η : $X_t = \sum_{j \geq 0} \psi_j \eta_{t-j}$,
- ⊗ des coefficients π_j qui lient η à X : $\eta_t = \sum_{j \geq 0} \pi_j X_{t-j}$,
- ⊗ des valeurs numériques observées des $X_u, u \leq T$,
- ⊗ des valeurs numériques observées des $\eta_u, u \leq T$.

Formule déduite de la forme autorégressive

Le théorème suivant donne les prévisions optimales obtenues grâce à la représentation AR(∞) du processus lorsque celle-ci existe :

Théorème 1.3.13. .

1. Les prévisions optimales s'expriment comme des combinaisons linéaires

$$\hat{X}_T(h) = \sum_{j \geq 0} \alpha_j(h) X_{T-j}$$

avec pour chaque h

$$\sum_{j \geq 0} \alpha_j(h) z^j = \left[\frac{1}{z^h} \frac{\Theta(z)}{\Phi(z)} \right] \frac{\Phi(z)}{\Theta(z)}$$

où $[[]]$, signifie que l'on ne retient que les termes en z^k pour $k \geq 0$.

2. Les $\hat{X}_T(h)$ s'obtiennent par récurrence selon

$$\hat{X}_T(h) = - \sum_{j \geq 1} \pi_j \hat{X}_T(h-j)$$

avec les conditions initiales pour $t \leq T$: $\widehat{X}_T(t - T) = X_t$.

Remarquons que pour $h = 1$, on a $\alpha_j(1) = -\pi_{j+1}$ et il est inutile de calculer les $\alpha_j(h)$ pour $h \geq 2$. Autrement dit on procède comme pour un AR(p) à partir du moment où l'on dispose de la représentation "AR(∞)" qui exprime le bruit blanc d'innovation à partir de X .

La deuxième équation du théorème permet de calculer récursivement les prévisions $\widehat{X}_T(h)$ à partir des prévisions précédentes. Il faut d'abord remarquer que, pour $t \leq T$, $\widehat{X}_T(t - T) = X_t$. Ainsi, pour la prévision à l'horizon 1, on a la formule suivante :

$$\widehat{X}_T(1) = - \sum_{j=1}^{\infty} \pi_j X_{T+1-j}$$

qui ne fait intervenir que des valeurs observées de la série temporelle. Pour la prévision à l'horizon 2, on a une formule basée sur les observations et sur la prévision donnée ci-dessus :

$$\widehat{X}_T(2) = -\pi_1 \widehat{X}_T(1) - \sum_{j=2}^{\infty} \pi_j X_{T+2-j}$$

Et ainsi de suite pour tout $h \in \mathbb{N}^*$. Cependant, ces prévisions font intervenir des valeurs non observées, à savoir X_t pour $t \leq 0$. Il faut alors effectuer une approximation en tronquant la série. On obtient alors la prévision suivante :

$$\widehat{X}_T^*(h) = - \sum_{j=1}^{T+h-1} \pi_j X_{T+h-j}$$

avec toujours $\widehat{X}_T^*(t - T) = X_t$ pour $t \leq T$.

Formule déduite de la forme moyenne mobile

Le théorème suivant donne les prévisions optimales obtenues grâce à la représentation MA(∞) du processus lorsque celle-ci existe :

Théorème 1.3.14. .

(i) Les $\widehat{X}_T(h)$ s'expriment comme des combinaisons linéaires des valeurs passées du bruit blanc d'innovation

$$\widehat{X}_T(h) = \sum_{j \geq h} \psi_j \eta_{T+h-j}$$

(ii) Enfin, la variance de l'erreur de prédiction est donnée par

$$\mathbb{E} \left(\left(X_{T+h} - \widehat{X}_{T+h} \right)^2 \right) = \sigma^2 \sum_{j=0}^{h-1} \psi_j^2$$

Les erreurs de prédiction $X_{T+h} - \widehat{X}_T(h)$ ne sont donc pas non-corrélées. En effet, comme on a :

$$X_{T+h} - \widehat{X}_T(h) = \sum_{j=0}^{h-1} \psi_j \eta_{T+h-j}$$

alors la covariance entre l'erreur de prévision à l'horizon k et l'erreur de prévision à l'horizon h avec $h \geq k$ vaut :

$$\mathbb{E} \left(\left(X_{T+h} - \widehat{X}_T(h) \right) \left(X_{T+k} - \widehat{X}_T(k) \right) \right) = \sigma^2 \sum_{j=0}^{k-1} \psi_j \psi_{j+h-k}$$

Proposition 1.3.9. .

1. L'erreur de prévision à horizon 1 pour un processus ARMA est le bruit d'innovation η_{T+1} .
2. La variance de l'erreur de prévision à horizon h pour un processus ARMA croît depuis la variance du bruit d'innovation (valeur prise pour $h = 1$) jusqu'à la variance du processus lui-même.

1.4 Méthode de Box-Jenkins

Box & Jenkins (1976) ont promu une méthodologie consistant à modéliser les séries temporelles univariées au moyen des processus ARMA qui donnent souvent des bons résultats en prévision. la méthode de Box & Jenkins possède cinq étapes principales.

1.4.1 Transformation

Les modèles ARMA ne sont adaptés qu'aux données stationnaires de paramètres constants et d'erreurs de variance constante. Pour se ramener dans ces conditions, il faut souvent transformer les données. La transformation logarithmique permet de stabiliser la variance. Pour stabiliser les séries qui croient avec le temps il existe deux principaux techniques. On peut retirer le trend temporel de cette série à l'aide la régression linéaire, ou d'utiliser un filtre du type $(1 - T)$ qui différencie la série.

1.4.2 Identification du P et du Q du ARMA(P,Q)

L'identification de p et q est essentiellement fondée sur l'analyse conjointe des auto-corrélations et des auto-corrélations partielles . On retire le P_{max} du graphe de l'auto-corrélation partielle et Q_{max} du graphe de l'auto-corrélation, et toutes les combinaisons de zéro à P_{max} et de zéro à Q_{max} sont considérés comme modèles candidats.

1.4.3 Estimation des paramètres des modèles sélectionnés

L'estimation de la partie auto-régressif se fait à l'aide de l'équation de Yule- Walker, et la partie moyenne mobile peut être estimée par la méthode du maximum de vraisemblance.

1.4.4 Validation

Il s'agit de vérifier si les résidus du modèle ARMA estimé, vérifient les propriétés requises pour que l'estimation soit valide. Ces propriétés sont : les résidus doivent être un bruit blanc, non auto corrélés, de même variances, et ils suivent une loi normale de moyenne nulle. Si au moins l'une de ces hypothèses est rejetée, on peut alors mener des tests sur les paramètres du

modèle estimer.

Pour tester la validité du modèle nous allons utiliser quelques tests principaux à savoir le test de normalité de Jarque-Bera et le test d'autocorrélation de Ljung Box.

1.4.5 Choix d'un modèle parmi plusieurs

Après avoir suivi les étapes précédant si on a plusieurs modèles candidats, il va falloir choisir un optimal. Pour ce faire il y'a plusieurs critères, mais dans ce rapport nous allons faire rappels sur quelques uns à savoir :

1. Critère d'Akaike (AIC) :

$$AIC(p, q) = \log \hat{\sigma}^2 + \frac{2(p+q)}{T}$$

2. Critère de Schwarz (SBC ou BIC) :

$$SBC(p, q) = \log \hat{\sigma}^2 + \frac{2(p+q) \log T}{T}$$

3. Critère de Hannan :

$$\varphi(p, q) = \log \hat{\sigma}^2 + \frac{c(p+q) \log \log T}{T}$$

4. Critère de l'erreur :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \hat{X}_i)^2}$$

avec $c > 2$, T est le maximum de la fonction de vraisemblance du modèle ARMA(p, q) estimé et $p+q$ est le nombre de paramètres à estimer du modèle.

Conclusion

Au cours de notre étude dans ce chapitre, nous en concluons que pour les processus ARMA(p, q), la situation est moins simple puisque les auto-corrélations ou les auto-corrélations partielles convergent vers 0 à vitesse géométrique mais ne sont pas nulles à partir d'un certain rang. Pour une représentation ARMA(p, q) minimale, il existe une caractérisation des ordres p et q à l'aide de certains déterminants de matrices de corrélations qui peuvent être estimés. Une autre approche proposée par les logiciels est de sélectionner un modèle ARMA à partir des critères d'information du type AIC ou BIC etc....

Introduction

La théorie générale des distributions stables a été initiée par Paul Lévy en 1924. Dans son livre Calcul des probabilités, ce dernier consacra tout un chapitre à l'étude et la caractérisation de ces lois. En 1937 P. Lévy développa une approche nouvelle et plus simple de toute la théorie avec la découverte des lois infiniment divisibles. A partir de ses travaux et ceux d'autres collaborateurs, les lois α -stables se sont progressivement vulgarisées.

Les lois stables sont une famille de lois qui présentent un grand intérêt dans la modélisation de nombreux problèmes physiques. La loi la plus connue de cette famille est la loi normale. Nous donnerons plus de détails sur cette notion de lois stable dans ceux qui suivent.

L'objectif de l'étude de ce chapitre est de comprendre ces distributions dans leurs aspects théoriques et pratiques qui nous serviront des outils très pertinents dans la réalisation de notre mémoire.

2.1 Caractérisation des lois α -stable

L'intérêt des mathématiciens pour les sommes de variables aléatoires indépendantes se voit clairement dans des résultats bien connus de la théorie des probabilités, tels que la loi des grands nombres ou le théorème limite centrale. C'est aussi l'étude de ces sommes qui a permis la caractérisation des lois dites α -stables (ou simplement stables).

Dans cette section, nous allons présenter quelques définitions et propriétés ainsi que les fonctions caractéristiques d'une distribution α -Stable.

2.1.1 Quelques définitions et propriétés

Définition 2.1.1. (Propriété de stabilité) *Une variable aléatoire réelle X est dite stable (ou a une distribution stable) si et seulement si pour tous réels positifs a et b , il existe un réel positif c et un réel k tels que :*

$$aX_1 + bX_2 \stackrel{d}{=} cX + k \quad (2.1)$$

L'écriture $\stackrel{d}{=}$ indique une égalité en distribution. où X_1 et X_2 sont des variables aléatoires indépendantes qui ont chacune la même distribution

que X .

Si $k = 0$ on dit que la distribution est strictement stable .

La distribution stable peut être décrite plus convenable par sa fonction caractéristique. La formule suivante est dérivée de la représentation dite de Lévy du fonction caractéristique d'une loi infiniment divisible [14] .

Définition 2.1.2. une variable aléatoire X est dite α -stable si et seulement si sa fonction caractéristique est donnée par :

$$\log \phi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \left\{ 1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2} \right\} + i\mu t, & \alpha \neq 1 \\ -\sigma |t| \left\{ 1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \log |t| \right\} + i\mu t, & \alpha = 1 \end{cases} \quad (2.2)$$

où $\alpha \in (0, 2]$, $\beta \in [-1, 1]$, $\sigma > 0$, $\mu \in R$

et

$$\operatorname{sign}(t) = \begin{cases} 1 & \text{si } t > 0 \\ 0 & \text{si } t = 0 \\ -1 & \text{si } t < 0 \end{cases}$$

Puisque 2.2 est caractérisé par quatre paramètre on note α -stable distributions par : $S_\alpha(\sigma, \beta, \mu)$ et nous écrivons :

$$X \sim S_\alpha(\sigma, \beta, \mu)$$

pour indiquer que X a une distribution stable avec l'exposant caractéristique (index) α , le paramètre d'échelle σ , l'asymétrie β et le paramètre de localisation μ .

Mais cette représentation de la fonction caractéristique, qui s'appelle paramétrisation standard, a le désavantage de ne pas être continue en ses paramètres. En fait, il y a discontinuité en les points où $\alpha = 1$ et $\beta = 0$.

Par ailleurs il existe d'autres paramétrisations de la fonction plus adaptées aux différents problèmes. Ici, on fait référence à deux paramétrisations proposées par Zolotarev.

$$\log \phi_0(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \left\{ \left\{ 1 + i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2} \right\} [(\sigma|t|)^{1-\alpha} - 1] \right\} + i\mu_0 t, & \alpha \neq 1 \\ -\sigma |t| \left\{ 1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \log |t| \right\} + i\mu_0 t, & \alpha = 1 \end{cases} \quad (2.3)$$

Cette représentation S^0 se note $X \sim S_\alpha^0(\sigma, \beta, \mu_0)$. Les paramètres α, β et σ de la paramétrisation S^0 sont les mêmes que ceux de la paramétrisation standard, mais μ et μ_0 sont reliés par :

$$\mu = \begin{cases} \mu_0 - \beta\sigma \tan \frac{\pi\alpha}{2}, & \alpha \neq 1 \\ \mu_0 - \beta \frac{2}{\pi} \sigma \log \sigma, & \alpha = 1 \end{cases} \quad (2.4)$$

Cette paramétrisation est très importante par ce que la fonction caractéristique, la densité et la fonction cumulative de répartition sont continues par rapport aux quatre paramètres. Donc, elle est bien conditionnée numériquement pour le calcul.

L'interprétation des quatre paramètres

- ◇ Le paramètre α est appelé **exposant caractéristique** ou **index de stabilité**. Il détermine la vitesse de décroissance de la queue de distribution, c'est-à-dire plus α tend vers 0, plus la queue est lourde. Feller montre que pour cette famille de lois de probabilités, on ne peut pas avoir $\alpha > 2$.
- ◇ Le paramètre β est le paramètre **d'asymétrie**. Lorsque $\beta > 0$ ($\beta = 1$) la distribution de X est dite asymétrique (totalement asymétrique) à droite et lorsque $\beta < 0$ ($\beta = -1$) on dit que la distribution de X est asymétrique (totalement asymétrique) à gauche. Quand $\beta = 0$, la distribution de X est symétrique par rapport à μ .
- ◇ Le paramètre μ est le paramètre de **localisation**. En terme de fonction caractéristique, une distribution est symétrique autour de 0 si et seulement si sa fonction caractéristique est réelle. Dans 2.2 cela correspond à $\beta = \mu = 0$.
- ◇ Le paramètre σ est le paramètre **d'échelle**. Lorsque $\sigma = 1$ et $\mu = 0$ la distribution est appelée standard stable.

Les cas particuliers :

Les distributions α -stable les plus connues, et les seules dont nous disposons d'une forme explicite pour les densités, sont les suivantes :

- ⊗ La distribution gaussienne de densité

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

a pour fonction caractéristique

$$\varphi_G(t) = \exp\left\{-\frac{\sigma^2}{2}t^2 + i\mu t\right\}.$$

Ce qui correspond à la fonction caractéristique d'une loi $S_2\left(\frac{\sigma}{\sqrt{2}}, 0, \mu\right)$.

- ⊗ La distribution de Cauchy généralisée de densité

$$\frac{\sigma}{\pi((x-\mu)^2 + \sigma^2)}$$

a pour fonction caractéristique

$$\varphi_C(t) = \exp\{-\sigma|t| + i\mu t\},$$

ce qui correspond à une loi $S_1(\sigma; 0; \mu)$

⊗ La distribution de Lévy a pour densité

$$\left(\frac{\sigma}{2\pi}\right)^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp\left\{-\frac{\sigma}{2(x-\mu)}\right\}$$

sur (μ, ∞) , avec σ le paramètre d'échelle. Sa fonction caractéristique correspond à celle d'une loi $S_{1/2}(\sigma, 1, \mu)$.

Remarque 2.1.1. Comme la loi normale, toutes les distributions stables restent stables sous des transformations linéaires. On peut ainsi, pour simplifier les calculs, choisir des valeurs standards pour les paramètres d'échelle et de localisation.

Définition 2.1.3. Une variable aléatoire X est symétrique α -stable si et seulement si sa fonction caractéristique est donnée par :

$$\varphi_X(t) = E(\exp\{itX\}) = e^{-\sigma^\alpha |t|^\alpha} \quad (2.5)$$

où σ est le paramètre d'échelle de X .

Preuve 2.1.1. Supposons que nous connaissons la forme générale de la fonction caractéristique d'une loi α -stable générale. Dans ce cas, il suffit d'utiliser la propriété suivante :

– Une variable aléatoire est symétrique si et seulement si sa fonction caractéristique est réelle. Ce qui est le cas si et seulement si dans 2.2 on pose $\beta = \mu = 0$ (il s'agit d'une symétrie par rapport à l'origine).

Notation 2.1.1. Une variable aléatoire X distribuée suivant une loi stable de paramètres α, β, σ et μ sera notée $X \sim S_\alpha(\sigma, \beta, \mu)$ ou $X_\alpha(\sigma, \beta, \mu)$.

Lorsque la variable X sera symétrique autour de 0, elle sera simplement notée $X \sim S_\alpha S(\sigma)$ ou $X_\alpha(\sigma)$.

Proposition 2.1.1. Soient X_1 et X_2 deux variables aléatoires indépendantes de loi stable $S_\alpha(\sigma_1; \beta_1; \mu_1)$ et $S_\alpha(\sigma_2; \beta_2; \mu_2)$ alors $X_1 + X_2$ suit une loi stable $S_\alpha(\sigma; \beta; \mu)$ avec :

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{\frac{1}{\alpha}} \quad (2.6)$$

$$\beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha} \quad (2.7)$$

$$\text{et } \mu = \mu_1 + \mu_2 \quad (2.8)$$

Notons que si $\beta_1 = \beta_2$ alors $\beta = \beta_1 = \beta_2$.

Cette propriété d'additivité est très intéressante en finance, car deux titres ayant les mêmes valeurs des paramètres α peuvent être considérés ensemble et la loi qui résultera de cette association conservera les mêmes valeurs du paramètre α mais les autres paramètres seront modifiés.

Propriété 2.1.1. Soient X_1 et X_2 deux variables aléatoires qui suivent respectivement des distributions α -stables $S_\alpha(\sigma_1, \beta, \mu_1)$ et $S_\alpha(\sigma_2, \beta, \mu_2)$. Les réels $a > 0$ et b tels que :

$$X_1 \stackrel{d}{=} aX_2 + \sigma_1 b \quad (2.9)$$

sont déterminés de façon unique par

$$\begin{aligned} a &= (\sigma_1/\sigma_2)^{1/\alpha} \\ b &= \begin{cases} \mu_1 - \mu_2 (\mu_1/\mu_2)^{1/\alpha}, & \alpha \neq 1 \\ \mu_1 - \mu_2 + (2/\pi)\beta \ln(\sigma_1/\sigma_2), & \alpha = 1 \end{cases} \end{aligned} \quad (2.10)$$

En posant $\mu_2 = 0$ et $\sigma_2 = 1$ dans 2.10, on obtient le cas particulier suivant :

$$X_1 \stackrel{d}{=} \begin{cases} \sigma_1^{1/\alpha} X_2 + \sigma_1 \mu_1, & \alpha \neq 1 \\ \sigma_1^{1/\alpha} X_2 + \sigma_1 [\mu_1 + (2/\pi) \ln \sigma_1], & \alpha = 1 \end{cases} \quad (2.11)$$

2.2 Comportement des queues et moments

La loi normale ($\alpha = 2$) se distingue nettement des autres lois stables par le comportement asymptotique de sa queue de distribution. En effet, si la variable aléatoire X est distribuée suivant une loi normale centrée et de variance $2\sigma^2$, Feller montre que [10] :

$$P(X < -x) = P(X > x) \sim \frac{1}{2\sqrt{\pi}\sigma x} e^{-x^2/(4\sigma^2)} \quad (2.12)$$

quand $x \rightarrow \infty$.

Cependant, lorsque $\alpha < 2$ les queues sont asymptotiquement de type Pareto. Une distribution de Pareto est définie par la fonction de survie :

$$P(Y > x) = x^\nu L(x) \quad (2.13)$$

où $\nu > 0$ et $L(x)$ est une fonction à variation lente, c'est-à-dire que pour $t > 0$, $\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1$. D'où la propriété suivante.

Proposition 2.2.1. Soit X une variable aléatoire de loi $S_\alpha(\sigma, \beta, \mu)$ avec $0 < \alpha < 2$. Alors :

$$\begin{cases} \lim_{t \rightarrow \infty} t^\alpha \mathbb{P}(X > t) = C_\alpha \frac{1 + \beta}{2} \sigma^\alpha \\ \lim_{t \rightarrow \infty} t^\alpha \mathbb{P}(X < -t) = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha \end{cases} \quad (2.14)$$

où C_α est une constante donnée par :

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx \right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha) \cos(\frac{\pi\alpha}{2})} & \text{si } \alpha \neq 1, \\ 2/\pi & \text{si } \alpha = 1. \end{cases}$$

où $\Gamma(\theta)$ est la fonction de gamma définie pour $\theta > 0$, par :

$$\Gamma(\theta) = \int_0^{+\infty} x^{\theta-1} e^{-x} dx \quad (2.15)$$

La loi normale est la seule loi α -stable qui possède des moments finis de tous ordres. Pour toutes les autres lois α -stables, nous avons la propriété ci-dessous.

Proposition 2.2.2. Soit $X \sim S_\alpha(\sigma, \beta, \mu)$ avec $0 < \alpha < 2$. Alors

$$\begin{aligned} E|X|^p &< \infty && \text{pour tout } p < \alpha \\ E|X|^p &= \infty && \text{pour tout } p \geq \alpha \end{aligned} \quad (2.16)$$

Conséquence : Toutes les lois stables non-gaussiennes ont une variance infinie ou non définie. L'existence d'une variance finie pour la loi normale est simplement liée à une plus grande décroissance de queue par rapport aux autres lois stables.

Remarque 2.2.1. .

1. Dès que α est strictement inférieur à 2, la variance d'une loi α -stable est infinie. Dès que α est strictement inférieur à 1, c'est la moyenne qui devient infinie
2. Si α est plus grand que 1, la moyenne d'une loi α -stable est μ .
3. La loi de Cauchy ($\alpha = 1$) et celle de Lévy ($\alpha = 1/2$) ont chacune une espérance mathématique infinie. Dans ces deux cas, la variance n'est pas définie.

2.3 Fonctions de densité et de distribution stables

L'absence de formules générales pour la plupart des densités et des fonctions de distribution stables a des conséquences négatives. Par exemple, lors de l'estimation de la probabilité par méthode de maximum de vraisemblance, des approximations numériques lourdes à calculer doivent être utilisées.

Il existe généralement deux approches à ce problème. Soit la transformée de Fourier rapide doit être appliquée à la fonction caractéristique ou bien l'intégration numérique directe numérique directe.

En posant : $\zeta = \beta \tan \frac{\pi\alpha}{2}$. Alors la densité $f(x, \alpha, \beta)$ d'une variable aléatoire standard α -stable dans la représentation S^0 c'est à dire $X \sim S_\alpha^0(1, \beta, 0)$ peut être exprimée comme :

- pour $\alpha \neq 1$ et $x > \zeta$

$$f(x, \alpha, \beta) = \frac{\alpha(x-\zeta)^{\frac{1}{\alpha}-1}}{\pi|\alpha-1|} \int_{-\zeta}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp\{(x-\zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta)\} d\theta \quad . \quad (2.17)$$

- pour $\alpha \neq 1$ et $x = \zeta$

$$f(x, \alpha, \beta) = \frac{\Gamma(1+\frac{1}{\alpha}) \cos(\zeta)}{\pi(1+\zeta^2)^{\frac{1}{2\alpha}}} \quad . \quad (2.18)$$

- pour $\alpha \neq 1$ et $x < \zeta$

$$f(x, \alpha, \beta) = f(-x, \alpha, -\beta) \quad . \quad (2.19)$$

- pour $\alpha = 1$

$$f(x, 1, \beta) = \begin{cases} \frac{e^{-\frac{\pi x}{2\beta}}}{2|\beta|} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta; 1, \beta) \exp\{e^{-\frac{\pi x}{2\beta}} V(\theta; 1, \beta)\} d\theta; & \beta \neq 0 \\ \frac{1}{\pi(1+x^2)}; & \beta = 0 \end{cases} \quad (2.20)$$

où

$$\epsilon = \begin{cases} \frac{1}{\alpha} \arctan(-\zeta), & \alpha \neq 1 \\ \frac{\pi}{2}, & \alpha = 1 \end{cases} \quad (2.21)$$

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos \alpha \epsilon)^{\frac{1}{\alpha-1}} \left(\frac{\cos \theta}{\sin \alpha(\epsilon+\theta)} \right)^{\frac{\alpha}{\alpha-1}} \frac{\cos\{\alpha\epsilon+(\alpha-1)\theta\}}{\cos \theta}, & \alpha \neq 1 \\ \frac{2}{\pi} \frac{(\frac{\pi}{2}+\beta\theta)}{\cos \theta} \exp\{\frac{1}{\beta}(\frac{\pi}{2} + \beta\theta) \tan \theta\}; & \alpha = 1, \beta \neq 0. \end{cases} \quad (2.22)$$

La distribution $F(x; \alpha; \beta)$ d'une variable aléatoire standard α -stable dans la représentation S^0 peut être exprimée comme suit :

- pour $\alpha \neq 1$ et $x > \zeta$

$$F(x, \alpha, \beta) = c_1(\alpha, \beta) + \frac{\text{sign}(1-\alpha)}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\{(x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta)\} d\theta \quad . \quad (2.23)$$

où

$$c_1(\alpha, \beta) = \begin{cases} \frac{1}{\pi}(\frac{\pi}{2} - \epsilon), & \alpha < \zeta \\ 1, & \alpha > 1 \end{cases} \quad (2.24)$$

- pour $\alpha \neq 1$ et $x = \zeta$

$$F(x, \alpha, \beta) = \frac{1}{\pi}(\frac{\pi}{2} - \epsilon) \quad . \quad (2.25)$$

- pour $\alpha \neq 1$ et $x < \zeta$

$$F(x, \alpha, \beta) = 1 - F(-x, \alpha, -\beta) \quad . \quad (2.26)$$

- pour $\alpha = 1$

$$F(x, 1, \beta) = \begin{cases} \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\{e^{-\frac{\pi x}{2\beta}} V(\theta; 1, \beta)\} d\theta; & \beta > 0 \\ \frac{1}{2} + \frac{1}{\pi}; & \beta = 0 \\ 1 - F(x, 1, -\beta), & \beta < 0 \end{cases} \quad (2.27)$$

2.4 Algorithme de simulation d'une variable aléatoire α -stable

Pour simuler les lois stables, il existe un algorithme développé par Chambers et al. (1976). Celui-ci permet de générer une loi $S_\alpha(1, \beta, 0)$. Pour obtenir une loi $S_\alpha(\sigma, \beta, \mu)$, avec $\alpha \in]0; 2]$ et $\beta \in [-1; 1]$, nous allons effectuer quelques changements de variables dans les étapes à venir :

Première étape

Elle consiste à générer une loi Φ uniforme sur $[-\frac{\pi}{2}; \frac{\pi}{2}]$ et une loi W exponentielle de paramètre 1. Pour cela, il faut d'abord générer deux v.a.r. uniformes sur $]0; 1[$ (notées U_1 et U_2). Puis en utilisant le changement de variables suivant :

$$\Phi = \pi U_1 - \frac{\pi}{2} \quad (2.28)$$

$$W = -\log(1 - U_2) \quad (2.29)$$

Deuxième étape

Elle consiste à calculer différentes quantités (fonction de Φ et de W).

On pose :

$$\left\{ \begin{array}{l} \epsilon = 1 - \alpha \\ a = \tan\left(\frac{\Phi}{2}\right) \\ b = \tan\left(\frac{\epsilon\Phi}{2}\right) \\ \tau = -\epsilon \tan(\alpha\Phi_0) \\ B = \frac{b}{\frac{\epsilon\Phi}{2}} \\ z = \frac{\cos(\epsilon\Phi) - \tan(\alpha\Phi_0) \sin(\epsilon)\Phi}{W \cos \Phi} \\ d = \frac{z^{1-\epsilon} - 1}{\epsilon} \end{array} \right.$$

Avec :

$$\Phi_0 = \frac{\pi\beta}{2} \cdot \frac{1-|1-\alpha|}{\alpha}$$

Troisième étape

Elle consiste à générer une loi Y stable $S_\alpha(1, \beta, 0)$. Pour obtenir cela, il faut utiliser la proposition suivante :

Proposition 2.4.1. Soit une loi uniforme sur $[-\frac{\pi}{2}; \frac{\pi}{2}]$ et W une loi exponentielle de paramètre 1, si nous posons :

- pour $\alpha \neq 1$

$$Y = (\cos(\alpha\Phi_0))^{\frac{1}{\alpha}} \left\{ \frac{2(a-b)(1+ab) - \Phi\tau B[b(1-a^2) - 2a]}{(1-a^2)(1+b^2)} \cdot (1 + \epsilon d) + \tau(d + \frac{1}{\epsilon}) \right\} \quad (2.30)$$

Ce qui entraîne que :

$$Y = \frac{\sin \alpha(\Phi - \Phi_0)}{(\cos \Phi)^{\frac{1}{\alpha}}} \cdot \left(\frac{\cos(\Phi - \alpha(\Phi - \Phi_0))}{W} \right)^{\frac{1-\alpha}{\alpha}} \quad (2.31)$$

- pour $\alpha = 1$

$$Y = \frac{2}{\pi} \left(\left(\frac{1}{2}\pi + \beta\Phi \right) \tan \Phi - \beta \log \left(\frac{\frac{1}{2}\pi W \cos \Phi}{\frac{1}{2}\pi + \beta\Phi} \right) \right) \quad (2.32)$$

alors la v.a.r. Y suit une loi $S_\alpha(1, \beta, 0)$.

Pour le cas $\alpha \neq 1$, la démonstration du passage de la première ligne seconde ligne, on a :

$$\begin{aligned} Y &= (\cos(\alpha\Phi_0))^{\frac{1}{\alpha}} \left\{ \frac{2(a-b)(1+ab) - \Phi\tau B[b(1-a^2) - 2a]}{(1-a^2)(1+b^2)} \cdot (1 + \epsilon d) + \tau(d + \frac{1}{\epsilon}) \right\} \\ &= \frac{\sin \alpha(\Phi - \Phi_0)}{(\cos \Phi)^{\frac{1}{\alpha}}} \cdot \left(\frac{\cos(\Phi - \alpha(\Phi - \Phi_0))}{W} \right)^{\frac{1-\alpha}{\alpha}} \end{aligned}$$

Tout d'abord, quelques rappels sur la fonction tangente :

$$\tan(2x) - \tan(2y) = \frac{2(\tan x - \tan y)(1 + \tan x \tan y)}{(1 - \tan^2 x)(1 - \tan^2 y)} \quad (2.33)$$

$$\frac{1 - \tan^2}{1 + \tan^2} = \cos(2x) \quad (2.34)$$

$$\tan(2x) = \frac{2 \tan x}{1 - \tan^2 x} \quad (2.35)$$

$$\frac{2 \tan x}{1 + \tan^2 x} = \sin(2x) \quad (2.36)$$

$$\frac{2 \tan^2 x}{1 + \tan^2 x} = 1 - \cos(2x) \quad (2.37)$$

Posons :

$$\begin{aligned} A &= \frac{2(a-b)(1+ab)}{(1-a^2)(1+b^2)} \\ C &= -\frac{\Phi\tau B[b(1-a^2) - 2a]}{(1-a^2)(1+b^2)} \\ D &= 1 + \epsilon d \\ G &= \tau(d + \frac{1}{\epsilon}) \end{aligned}$$

Alors A s'écrit sous la forme :

$$\begin{aligned} A &= \frac{2(\tan \frac{\Phi}{2} - \tan \frac{\epsilon\Phi}{2})(1 + \tan \frac{\Phi}{2} \tan \frac{\epsilon\Phi}{2})}{(1 - \tan^2 \frac{\Phi}{2})(1 + \tan^2 \frac{\epsilon\Phi}{2})} \\ &= \frac{2(\tan \frac{\Phi}{2} - \tan \frac{\epsilon\Phi}{2})(1 + \tan \frac{\Phi}{2} \tan \frac{\epsilon\Phi}{2})}{(1 - \tan^2 \frac{\Phi}{2})(1 + \tan^2 \frac{\epsilon\Phi}{2})} \cdot \left(\frac{1 - \tan^2 \frac{\epsilon\Phi}{2}}{1 - \tan^2 \frac{\epsilon\Phi}{2}} \right) \end{aligned}$$

D'après les relations 2.33 et 2.34, on a :

$$\begin{aligned} A &= (\tan \Phi - \tan \epsilon\Phi) \cos \epsilon\Phi \\ &= \tan \Phi \cos \epsilon\Phi - \sin \epsilon\Phi \\ &= \tan \Phi (\cos \Phi \cos \alpha\Phi + \sin \Phi \sin \alpha\Phi) - \sin \Phi \cos \alpha\Phi + \cos \Phi \sin \alpha\Phi \\ &= \sin \Phi \cos \alpha\Phi + \frac{\sin^2 \Phi \sin \alpha\Phi}{\cos \Phi} - \sin \Phi \cos \alpha\Phi + \frac{\cos^2 \Phi \sin \alpha\Phi}{\cos \Phi} \\ &= \frac{\sin \alpha\Phi}{\cos \Phi} \end{aligned}$$

De plus,

$$C = \frac{-\Phi (-\varepsilon \tan(\alpha\phi_0)) \frac{\tan \frac{\varepsilon\Phi}{2}}{\varepsilon \frac{\Phi}{2}} \left[\tan \frac{\varepsilon\Phi}{2} (1 - \tan^2 \frac{\Phi}{2}) - 2 \tan \frac{\Phi}{2} \right]}{(1 - \tan^2 \frac{\Phi}{2}) (1 + \tan^2 \frac{\varepsilon\Phi}{2})}$$

$$C = -\tan(\alpha\phi_0) \frac{2 \tan \frac{\varepsilon\Phi}{2} \left[-\tan \frac{\varepsilon\Phi}{2} (1 - \tan^2 \frac{\Phi}{2}) + 2 \tan \frac{\Phi}{2} \right]}{(1 - \tan^2 \frac{\Phi}{2}) (1 + \tan^2 \frac{\varepsilon\Phi}{2})}$$

$$C = -\tan(\alpha\phi_0) \frac{\left[-2 \tan^2 \frac{\varepsilon\Phi}{2} (1 - \tan^2 \frac{\Phi}{2}) + 4 \tan \frac{\Phi}{2} \tan \frac{\varepsilon\Phi}{2} \right]}{(1 - \tan^2 \frac{\Phi}{2}) (1 + \tan^2 \frac{\varepsilon\Phi}{2})}$$

D'après les relations 2.35, 2.36 et 2.37, on a

$$C = -\tan(\alpha\phi_0) (-1 + \cos \varepsilon\Phi + \tan \Phi \sin \varepsilon\Phi)$$

$$C = -\tan(\alpha\phi_0) \left(-1 + \frac{\cos \Phi \cos \varepsilon\Phi}{\cos \Phi} + \frac{\sin \varepsilon\Phi \sin \Phi}{\cos \Phi} \right)$$

$$C = -\tan(\alpha\phi_0) \left(\frac{\cos \alpha\Phi}{\cos \Phi} - 1 \right)$$

Enfin,

$$D = 1 + \varepsilon \left(\frac{z^{\varepsilon/1-\varepsilon} - 1}{\varepsilon} \right) = z^{\frac{1-\alpha}{\alpha}}$$

$$G = -\varepsilon \tan(\alpha\phi_0) \left(\left(\frac{z^{\varepsilon/1-\varepsilon} - 1}{\varepsilon} \right) + \frac{1}{\varepsilon} \right) = -\tan(\alpha\phi_0) z^{\frac{1-\alpha}{\alpha}}$$

Ce qui entraîne que

$$Y = (\cos(\alpha\phi_0))^{1/\alpha} \{(A + C) \cdot D + G\}$$

$$Y = (\cos(\alpha\phi_0))^{1/\alpha} \left[\left(\frac{\sin \alpha\Phi}{\cos \Phi} - \tan(\alpha\phi_0) \left(\frac{\cos \alpha\Phi}{\cos \Phi} - 1 \right) \right) \cdot z^{\frac{1-\alpha}{\alpha}} - \tan(\alpha\phi_0) z^{\frac{1-\alpha}{\alpha}} \right]$$

$$Y = (\cos(\alpha\phi_0))^{1/\alpha} z^{\frac{1-\alpha}{\alpha}} \left(\frac{\sin \alpha\Phi}{\cos \Phi} - \tan(\alpha\phi_0) \frac{\cos \alpha\Phi}{\cos \Phi} \right)$$

$$Y = (\cos(\alpha\phi_0))^{1/\alpha} z^{\frac{1-\alpha}{\alpha}} \frac{1}{\cos \Phi \cos \alpha\phi_0} (\sin \alpha\Phi \cos \alpha\phi_0 - \sin \alpha\phi_0 \cos \alpha\Phi)$$

$$Y = (\cos(\alpha\phi_0))^{1/\alpha} \left(\frac{\cos(\varepsilon\Phi) \cos(\alpha\phi_0) - \sin(\alpha\phi_0) \sin(\varepsilon\Phi)}{W \cos \Phi \cos(\alpha\phi_0)} \right)^{\frac{1-\alpha}{\alpha}} \frac{1}{\cos \Phi} (\sin \alpha(\Phi - \phi_0))$$

$$Y = \frac{\sin \alpha(\Phi - \phi_0)}{\cos \Phi^{\frac{1}{\alpha}}} \cdot \left(\frac{\cos(\varepsilon\Phi + \alpha\phi_0)}{W} \right)^{\frac{1-\alpha}{\alpha}}$$

$$Y = \frac{\sin \alpha(\Phi - \phi_0)}{\cos \Phi^{\frac{1}{\alpha}}} \cdot \left(\frac{\cos(\Phi - \alpha(\Phi - \phi_0))}{W} \right)^{\frac{1-\alpha}{\alpha}}$$

Cas particuliers :

- Dans le cas d'une loi $S\alpha S$ (c'est-à-dire $\beta = 0$), nous avons :

$$Y = \frac{\sin(\alpha\Phi)}{(\cos \Phi)^{\frac{1}{\alpha}}} \cdot \left(\frac{\cos(1-\alpha)\Phi}{W} \right)^{\frac{1-\alpha}{\alpha}} \quad (2.38)$$

- Plus particulièrement, dans le cas où α est égal à 2, nous avons :

$$\begin{aligned} Y &= \frac{\sin(2\Phi)}{\sqrt{\cos \Phi}} \cdot \left(\frac{\cos \Phi}{W}\right)^{\frac{-1}{2}} \\ &= 2\sqrt{W} \sin \Phi \end{aligned} \quad (2.39)$$

- Enfin, dans le cas où α est égal à 1 et β égal à 0, nous avons :

$$Y = \tan(\Phi) \quad (2.40)$$

formule connue, qui permet de simuler une loi de Cauchy.

Le cas $\alpha = 1$ se résout de manière similaire.

Donc on a :

- $Y = \frac{\sin \alpha(\Phi - \Phi_0)}{(\cos \Phi)^{\frac{1}{\alpha}}} \cdot \left(\frac{\cos(\Phi - \alpha(\Phi - \Phi_0))}{W}\right)^{\frac{1-\alpha}{\alpha}}$ suit une loi stable $S_\alpha(1, \beta, 0)$, pour $\alpha \neq 1$.
- $Y = \frac{2}{\pi} \left(\left(\frac{1}{2}\pi + \beta\Phi \right) \tan \Phi - \beta \log\left(\frac{\frac{1}{2}\pi W \cos \Phi}{\frac{1}{2}\pi + \beta\Phi}\right) \right)$, suit une loi stable $S_\alpha(1, \beta, 0)$, pour $\alpha = 1$.

De plus, si Y suit une loi stable $S_\alpha(1, \beta, 0)$, Ainsi

$$X = \begin{cases} \sigma Y + \mu & , \alpha \neq 1 \\ \sigma Y + \frac{2}{\pi} \beta \ln \sigma + \mu & , \alpha = 1 \end{cases} \quad (2.41)$$

suit une loi stable $S_\alpha(\sigma, \beta, \mu)$.

Donc on peut ainsi simuler des variables X de loi stable $S_\alpha(\sigma, \beta, \mu)$.

Exemple 2.4.1. Le tableau suivant représente la moyenne et la variance empirique de 5000 réalisations d'une loi $S_\alpha S$ pour différentes valeurs de α .

Ces résultats confirment l'équation sur le calcul des moments. En effet, lorsque α décroît vers 1, la variance explose et lorsque α devient plus petit que 1, c'est la moyenne qui commence à exploser.

α	2	1,7	1,5	1,2	1	0,9
moyenne	0,02	0,02	0,03	0,01	-0,33	27,58
variance	2,02	6,70	36,96	150,97	2071,94	3214.206,71

TABLE 2.1 – Moyenne et variance empirique calculées sur 5000 réalisations

Conclusion

Au cours de l'étude de ce chapitre, nous avons fait une description générale concernant les lois stables qui ont été proposées par Paul Lévy en (1924), une étude de l'interprétation et l'influence de chaque paramètre de cette loi, l'algorithme qui permet de les générer et nous avons pu en déduire quelques constats à savoir que :

parmi les paramètres d'une distribution α -stable, on peut dire que l'indice de stabilité ($0 < \alpha \leq 2$) semblerait un paramètre très intéressant lorsqu'on utilise ces distributions pour des éventuels travaux car il nous aide à déterminer la forme ou le degré d'épaisseur de la queue de distribution chose à ne pas négliger et suivre de très prêt.

Méthodes d'estimation et Tests statistiques

Introduction

L'objectif de l'étude de ce chapitre est dans un premier temps de présenter quelques méthodes d'estimation des paramètres d'un modèle. A noter que la liste des méthodes présentes dans ce chapitre n'est pas exhaustive. En suite nous ferons un aperçu général sur quelques tests de normalité ainsi que certains graphiques statistiques. En fin nous terminerons ce chapitre avec l'analyse des erreurs qui nous aidera à bien choisir le modèle le plus adéquat parmi tant d'autres.

3.1 Le maximum de Vraisemblance

C'est une technique qui, sous l'hypothèse que les variables ont une distribution connue, usuellement la distribution normale, permet d'estimer les paramètres d'un modèle (d'une équation ou d'un système, linéaire ou non linéaire) avec des restrictions sur les paramètres (coefficients, matrice de variances et covariances). Plus spécifiquement la technique consiste à construire une fonction appelée fonction de vraisemblance (construite à partir de la fonction de densité) et à maximiser son logarithme par rapport aux paramètres inconnus.

3.1.1 Fonction de vraisemblance

On appelle fonction de vraisemblance pour (x_1, \dots, x_n) la fonction de θ :

$$L_n(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta) \quad (3.1)$$

Avec :

$$f(x; \theta) = \begin{cases} \mathbb{P}_\theta(X = x) & \text{si } X \text{ est une variable discrète,} \\ f_\theta(x) & \text{si } X \text{ est une variable continue de densité } f_\theta \end{cases} \quad (3.2)$$

La fonction de vraisemblance n'est intéressante que si θ et x_i vérifient $f(x_i; \theta) \neq 0$ pour tout $i \in \{1, \dots, n\}$.

3.1.2 Estimateurs du maximum de vraisemblance pour (x_1, \dots, x_n)

On appelle estimateur du maximum de vraisemblance de MLE de θ pour (x_1, \dots, x_n) est un réel θ^* qui maximise la fonction de vraisemblance $L_n(x_1, \dots, x_n; \theta)$ en θ , i.e. pour tout θ ,

$$L_n(x_1, \dots, x_n; \theta) \leq L_n(x_1, \dots, x_n; \theta^*) \quad (3.3)$$

Une expression alternative est :

$$\theta^* \in \operatorname{argmax}_{\theta} L_n(x_1, \dots, x_n; \theta) \quad (3.4)$$

où argmax désigne l'argument du maximum qui est l'ensemble des points en lesquels une expression atteint sa valeur maximale.

3.1.3 Fonction de Log-vraisemblance

On appelle fonction de log-vraisemblance pour (x_1, \dots, x_n) la fonction de θ définie par :

$$\ell_n(x_1, \dots, x_n; \theta) = \ln(L_n(x_1, \dots, x_n; \theta)) \quad (3.5)$$

Elle n'a de sens que si θ vérifie $L_n(x_1, \dots, x_n; \theta) > 0$. La fonction logarithme népérien étant croissante, MLE; θ^* de θ pour (x_1, \dots, x_n) vérifie :

$$\theta^* \in \operatorname{argmax}_{\theta} L_n(x_1, \dots, x_n; \theta) = \operatorname{argmax}_{\theta} \ell_n(x_1, \dots, x_n; \theta). \quad (3.6)$$

3.1.4 Équation de vraisemblance

On appelle équation de vraisemblance l'équation en θ :

$$\frac{\partial}{\partial \theta} \ell_n(x_1, \dots, x_n; \theta) = 0. \quad (3.7)$$

3.1.5 Expression analytique de MLE

Pour envisager d'avoir une expression analytique de θ^* MLE de θ pour (x_1, \dots, x_n) , une idée est d'exprimer $L_n(x_1, \dots, x_n; \theta)$ en fonction de produits de termes exponentiels/puissances, puis de considérer la fonction de log-vraisemblance $\ell_n(x_1, \dots, x_n; \theta)$. Si cette dernière est dérivable en θ , une condition nécessaire que doit vérifier θ^* est d'être solution de l'équation de vraisemblance. Il faut ensuite vérifier que θ^* est bien un maximum pour $\ell_n(x_1, \dots, x_n; \theta)$:

— soit en étudiant les variations de $\ell_n(x_1, \dots, x_n; \theta)$

— soit en montrant que

$$\frac{\partial^2}{\partial \theta^2} \ell_n(x_1, \dots, x_n; \theta^*) < 0 \quad (3.8)$$

3.1.6 Pratique de MLE

Bien souvent, la résolution de l'équation de vraisemblance n'amène pas une expression analytique pour MLE θ^* de θ pour (x_1, \dots, x_n) . En tant que problème d'optimisation, on peut quand même approcher la valeur de θ^* avec précision à l'aide d'algorithmes itératifs efficaces. Il y a notamment :

- ⊗ l'algorithme de Newton-Raphson,
- ⊗ l'algorithme de Gauss-Newton.

Par exemple, partant de l'équation de vraisemblance et d'un développement de Taylor au premier ordre de $\frac{\partial}{\partial \theta} \ell_n(x_1, \dots, x_n; \theta)$, l'algorithme de Newton-Raphson est caractérisé par la m -ème itération :

$$\theta^{(m+1)} = \theta^{(m)} - \left(\frac{\partial^2}{\partial \theta^2} \ell_n(x_1, \dots, x_n; \theta^{(m)}) \right)^{-1} \left(\frac{\partial}{\partial \theta} \ell_n(x_1, \dots, x_n; \theta^{(m)}) \right) \quad (3.9)$$

Dès qu'il y a stabilisation pour un m_* , i.e. $\theta^{(m_*+1)} = \theta^{(m_*)}$, on considère la valeur $\theta^* = \theta^{(m_*)}$.

Exemple 3.1.1. On suppose à présent que l'échantillon x_1, \dots, x_n est tiré de manière normale avec une espérance μ et un écart-type σ , mais μ et σ sont inconnus. On modélise donc le problème par une loi normale $\mathcal{N}(\mu, \sigma)$ dont la densité est :

$$f_{(\mu, \sigma)}(x) := \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.10)$$

On va chercher un estimateur de $\theta = (\mu, \sigma)$ par la méthode du maximum de vraisemblance. La vraisemblance de l'échantillon x_1, \dots, x_n est donc :

$$\begin{aligned} L_n(x_1, \dots, x_n; \theta) &:= \prod_{i=1}^n f_{\theta}(x_i) = \prod_{i=1}^n \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ &= \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^n e^{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}} \end{aligned}$$

Ici, il est une nouvelle fois plus agréable de considérer la log-vraisemblance

$$l_n(x_1, \dots, x_n; \theta) := \ln(L_n(x_1, \dots, x_n; \theta)) = -n(\ln(\sigma) + \ln(\sqrt{2\pi})) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \quad (3.11)$$

Pour que $\theta^* = (\mu^*, \sigma^*)$ soit un extremum sur $\mathbb{R} \times \mathbb{R}_*^+$ il faut que les deux dérivées :

$$\frac{\partial}{\partial \mu} l_n(x_1, \dots, x_n; \theta) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = \frac{1}{\sigma^2} (s - n\mu)$$

et

$$\frac{\partial}{\partial \sigma} l_n(x_1, \dots, x_n; \theta) = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2$$

s'annulent pour $\theta = \theta^*$, où $s = \sum_{i=1}^n x_i$, ce qui implique que :

$$\mu = \frac{s}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

et

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

Ceci nous conduit donc à envisager l'estimateur

$$\hat{\theta} = (\hat{\mu}, \hat{\sigma}) = \left(\frac{1}{n} \sum_{i=1}^n X_i, \left(\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \right)^{\frac{1}{2}} \right) \quad (3.12)$$

En ce qui concerne la première composante $\hat{\mu}$, nous retrouvons une nouvelle fois la moyenne comme estimateur de l'espérance $\mu = \mathbb{E}(X_i)$, quant-à la seconde composante, nous trouvons :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \quad (3.13)$$

3.2 Méthode de Whittle

La méthode d'estimation des paramètres Whittle est basé sur la minimisation de la fonction de vraisemblance suivante :

$$L_n(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \log f_{\theta}(\lambda) + \frac{I_n(\lambda)}{f_{\theta}(\lambda)} \right\} d\lambda \quad (3.14)$$

où $f_{\theta}(\lambda)$ est la fonction de densité spectrale du modèle et $I_n(\lambda)$ est le périodogramme .

Cette estimation est asymptotiquement efficace et $L_n(\theta)$ converge en probabilité vers la fonction de vraisemblance gaussienne quand $n \rightarrow \infty$. Dans le cas des processus AR stationnaires, les estimateurs obtenus par cette méthode sont exactement les mêmes que les estimateurs de Yule-Walker et leurs asymptotiques sont les mêmes [4]. $L_n(\theta)$ peut être interprétée comme la distance entre la densité spectrale paramétrique $f_{\theta}(\lambda)$ et l'estimation du périodogramme non paramétrique $I_n(\lambda)$. Par conséquent, l'estimateur de Whittle est appelé distance minimale .

La théorie d'estimation des processus localement stationnaires est motivée par la généralisation de la fonction de vraisemblance de Whittle pour les processus stationnaires (Dahlhaus, 1997), où le périodogramme habituel est remplacé par des périodogrammes locaux sur des segments de données stationnaires pouvant se chevaucher. La version généralisée de la méthode Whittle pour un processus localement stationnaire remplace $I_n(\lambda)$ dans $L_n(\theta)$ par la version locale et l'intègre au long du temps. Le résultat est encore un estimateur efficace [4] .

Généralisation de la méthode de Whittle aux processus localement stationnaires

Soit $X_{t,n}(t = 0, \dots, n-1)$ un processus localement stationnaire, et soit $h : \mathbb{R} \rightarrow \mathbb{R}$ est une donnée efficace avec $h(u) = 0$ pour $u \notin [0, 1)$. Ensuite , pour une longueur de segment uniforme $N \leq n$ et mise à l'échelle du temps $u = \frac{t}{n}$ on définit :

$$J_N(u, \lambda) = \sum_{j=0}^{N-1} h\left(\frac{j}{N}\right) X_{\lfloor un \rfloor - (N/2) + j + 1, n} \exp(-i\lambda j) \quad (3.15)$$

$$H_{k,N}(\lambda) = \sum_{j=0}^{N-1} h \left(\frac{j}{N} \right)^k \exp(-i\lambda j) \quad (3.16)$$

$$I_N(u, \lambda) = \frac{1}{2\pi H_{2,N}(0)} |J_N(u, \lambda)|^2 \quad (3.17)$$

où $\lfloor un \rfloor$ est la partie entière de un .

Ainsi, $I_N(u, \lambda)$ est le périodogramme sur un segment de longueur N avec milieu $\lfloor un \rfloor$.

Le passage d'un segment à l'autre est noté S ; c'est-à-dire que nous calculons I_N sur des segments avec des points médians $t_j := S(j-1) + N/2$ ($j = 1, \dots, M$) où $n = S(M-1) + N$ et M est le nombre total de segments, ou, écrit en temps redimensionné, à des moments temporels $u_j := t_j/N$.

Notez que, si S est le décalage d'un segment à l'autre, alors le chevauchement entre les segments est $N - S$.

Nous définissons maintenant la généralisation de la fonction de vraisemblance Whittle à un processus localement stationnaire comme suivant :

$$\mathcal{L}_n(\theta) = \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \left\{ \log f_{\theta}(u_j, \lambda) + \frac{I_N(u_j, \lambda)}{f_{\theta}(u_j, \lambda)} \right\} d\lambda \quad (3.18)$$

avec une estimation de contraste minimum :

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \mathcal{L}_n(\theta). \quad (3.19)$$

On définit $\mathcal{L}(\theta)$:

$$\mathcal{L}(\theta) = \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left(\log f_{\theta}(u, \lambda) + \frac{f_{\theta}(u, \lambda)}{f(u, \lambda)} \right) d\lambda du \quad (3.20)$$

il est Claire que , $\mathcal{L}_n(\theta)$ devrait converger vers $\mathcal{L}(\theta)$ quand $n \rightarrow \infty$.

3.2.1 L'ajustement de modèles autorégressifs variant dans le temps

Soit $X_{t,n}$ la solution du système de différence équations :

$$\sum_{j=0}^p a_j \left(\frac{t}{n} \right) X_{t-j,n} = \sigma \left(\frac{t}{n} \right) \varepsilon_t \quad \text{pour } t \in \mathbb{Z} \quad (3.21)$$

Où $a_0(u) = 1$ et $\{\varepsilon_t\}$ sont des variables aléatoires indépendantes avec une moyenne nulle et une variance unité. Ensuite, $X_{t,n}$ est localement stationnaire théorème 4.2.4 avec la densité spectrale variante dans le temps :

$$f(u, \lambda) = \frac{\sigma^2(u)}{2\pi} \left| \sum_{j=0}^p a_j(u) \exp(i\lambda j) \right|^{-2} \quad (3.22)$$

Exemple de fonction d'autocovariance de $X_{t,n}$:

$$\begin{aligned} c_N(u, j) &= \int_{-\Pi}^{\Pi} I_N(u, \lambda) \exp(i\lambda j) d\lambda \\ &= H_{2,N}(0)^{-1} \sum_{\substack{s,t=0 \\ s-t=j}}^{N-1} h\left(\frac{s}{N}\right) h\left(\frac{t}{N}\right) X_{[un]-(N/2)+s+1,T} X_{[un]-(N/2)+t+1,T} \end{aligned} \quad (3.23)$$

Soit

$$C_N(u) = (c_N(u, 1), \dots, c_N(u, p))' \quad (3.24)$$

$$\Sigma_N(u) = \{c_N(u, j-k)\}_{j,k=1,\dots,p}. \quad (3.25)$$

Pour plus de commodité, nous avons divisé la fonction de vraisemblance de Whittle généralisée \mathcal{L}_n en deux parties \mathcal{L}_n^1 et \mathcal{L}_n^2 , avec :

$$\begin{aligned} \mathcal{L}_n^1 &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \log f_{\theta}(u_j, \lambda) d\lambda \\ \mathcal{L}_n^2 &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \left\{ \frac{I_N(u_j, \lambda)}{f_{\theta}(u_j, \lambda)} \right\} d\lambda \end{aligned} \quad (3.26)$$

Supposons que (avec un abus de notation) $a(u) = (a_1(u), \dots, a_p(u))'$ et

$$\mathcal{L}_n^2(u) = \int_{-\pi}^{\pi} \left\{ \frac{I_N(u, \lambda)}{f_{\theta}(u, \lambda)} \right\} d\lambda \quad (3.27)$$

Ensuite

$$\begin{aligned} \mathcal{L}_n^2(u) &= \frac{2\pi}{\sigma^2(u)} \int_{-\pi}^{\pi} I_N(u, \lambda) \left| \sum_{j=0}^p a_j(u) \exp(i\lambda j) \right|^2 d\lambda \\ &= \frac{2\pi}{\sigma^2(u)} \int_{-\pi}^{\pi} I_N(u, \lambda) \left(\sum_{j,k=0}^p a_j(u) a_k(u) \exp(i\lambda(j-k)) \right) d\lambda \\ &= \frac{2\pi}{\sigma^2(u)} \left(\sum_{j=0}^p a_j^2(u) c_N(u, 0) + \sum_{\substack{j,k=0 \\ j \neq k}}^p a_j(u) a_k(u) c_N(u, (j-k)) \right) \\ &= \frac{2\pi}{\sigma^2(u)} (a'(u) \Sigma_N(u) a(u) + 2a'(u) C_N(u) + c_N(u, 0)) \end{aligned}$$

Donc on aura :

$$\begin{aligned} \mathcal{L}_n^2(u) &= \frac{2\pi}{\sigma^2(u)} \{ a'(u) \Sigma_N(u) a(u) + a'(u) C_N(u) + C_N'(u) a(u) + \\ &\quad + c_N(u, 0) + C_N'(u) \Sigma_N^{-1}(u) C_N(u) - C_N'(u) \Sigma_N^{-1}(u) C_N(u) \} \end{aligned}$$

Et finalement :

$$\mathcal{L}_n^2(u) = \frac{2\pi}{\sigma^2(u)} \{(\Sigma_N(u)a'(u) + C_N(u))' \Sigma_N^{-1} (\Sigma_N(u)a'(u) + C_N(u)) + c_N(u, 0) - C_N'(u) \Sigma_N^{-1} C_N(u)\}$$

Soit $\theta = \{(a_1(u_j), \dots, a_p(u_j), \sigma^2(u_j))\}$, où $j = 1, \dots, M$. en combinant \mathcal{L}_n^1 et \mathcal{L}_n^2 , Nous obtenons :

$$\begin{aligned} \mathcal{L}_n(\theta) = & \frac{1}{2} \frac{1}{M} \sum_{j=1}^M \left\{ \log \sigma^2(u_j) + \frac{1}{\sigma^2(u_j)} \times [(\Sigma_N(u_j) a(u_j) + \right. \\ & \left. + C_N(u_j))' \Sigma_N^{-1} (\Sigma_N(u_j) a(u_j) + C_N(u_j)) + c_N(u_j, 0) - C_N'(u_j) \Sigma_N^{-1} C_N(u_j)] \right\} \end{aligned} \quad (3.29)$$

En utilisant l'équation ci-dessus, on peut obtenir les dérivées partielles suivantes :

$$\begin{aligned} \frac{\partial \mathcal{L}_n(\theta)}{\partial a(u_j)} &= \frac{2a(u_j) \Sigma_N(u_j) + 2C_N(u_j)}{\sigma^2(u_j)} \\ \frac{\partial \mathcal{L}_n(\theta)}{\partial \sigma^2(u_j)} &= \frac{1}{\sigma^2(u_j)} - \frac{a'(u_j) \Sigma_N(u_j) a(u_j) + 2a'(u_j) C_N(u_j) + c_N(u, 0)}{\sigma^4(u_j)} \end{aligned}$$

lorsqu'on met les dérivées partielles ci-dessus égale à zéro on obtient les estimateurs suivants :

$$\begin{aligned} \hat{a}(u_j) &= -\Sigma_N^{-1}(u_j) C_N(u_j) \\ \hat{\sigma}^2(u_j) &= c_N(u, 0) - C_N'(u_j) \Sigma_N^{-1}(u_j) C_N(u_j) \end{aligned} \quad (3.30)$$

3.3 Inférence indirecte

L'avantage d'utiliser la classe des distributions α -stables est leur flexibilité pour l'asymétrie et les queues lourdes de plus, elles sont fermées sous les combinaisons linéaires, ce qui inclut la distribution gaussienne comme cas particulier. Cependant, son estimation est difficile car sa fonction de densité n'a pas de forme générale et les moments d'ordre supérieur à deux n'existent pas.

Par conséquent, les méthodes d'estimation habituelles, telles que le maximum de vraisemblance et la méthode des moments, ne fonctionnent pas.

Les approches indirectes proposées par *Gourieroux et al. (1993)* et *Gallant et Tauchen (1996)* pourraient être la solution pour des modèles plus complexes impliquant des distributions stables. L'inférence indirecte a été proposée par *Gourieroux et al. (1993)* dans le contexte de modèles économétriques avec des variables latentes, mais elle s'est avérée utile dans des situations où la maximisation directe de la fonction de vraisemblance n'est pas disponible.

3.3.1 Le principe de l'approche

Dans cette partie, nous présentons le principe de l'inférence indirecte en se basant sur la quatrième chapitre de Gouriéroux et Monfort [12]. Supposons que nous avons un échantillon de T observations \mathbf{y} et un modèle d'intérêt IM dont la fonction de vraisemblance $\mathcal{L}_T^*(\mathbf{y}; \theta)$ est difficile à manipuler et à maximiser. Par conséquent, le maximum de vraisemblance de $\theta \in \Theta$, donné par :

$$\hat{\theta}_T = \operatorname{argmax}_{\theta \in \Theta} \mathcal{L}_T^*(\theta; \mathbf{y}) = \operatorname{argmax}_{\theta \in \Theta} \sum_{t=1}^T \ln \ell^*(\theta; y_t), \quad (3.31)$$

n'est pas disponible. En outre, on considère un modèle alternatif, dépendant d'un vecteur de paramètres $\lambda \in \Lambda$, appelé modèle auxiliaire (AM). Supposons que la fonction de vraisemblance $\mathcal{L}_T(\mathbf{y}; \lambda)$ du AM est plus facile à manipuler.

Cependant, son estimateur :

$$\hat{\lambda}_T = \operatorname{argmax}_{\lambda \in \Lambda} \mathcal{L}_T(\lambda; \mathbf{y}) = \operatorname{argmax}_{\lambda \in \Lambda} \sum_{t=1}^T \ln \ell_T(\lambda; y_t) \quad (3.32)$$

L'idée est d'effectuer des simulations sous le IM pour corriger le biais de l'estimateur $\hat{\lambda}$. Pour continuer, nous décrivons la procédure générale de l'inférence indirecte.

L'algorithme de la méthode

- ◇ **Étape 1** : Calculer l'estimation du maximum de vraisemblance de λ sur la base de T observations \mathbf{y} , qui sera désignée par $\hat{\lambda}_T$.
- ◇ **Étape 2** : Simuler un ensemble de S vecteurs de taille T à partir de l'IM sur la base d'un vecteur paramètre arbitraire $\hat{\theta}^{(0)}$.
Désignons chacun de ces vecteurs par $\tilde{\mathbf{y}}^s(\hat{\theta}^{(0)}) = \{\tilde{y}_1^s(\hat{\theta}^{(0)}), \dots, \tilde{y}_T^s(\hat{\theta}^{(0)})\}$ pour $s = 1, \dots, S$.

- ◇ **Étape 3** : Ensuite, estimer les paramètres de l'AM en utilisant les valeurs simulées de l'IM

$$\tilde{\lambda}_T^s(\hat{\theta}^{(0)}) = \operatorname{argmax}_{\lambda \in \Lambda} \mathcal{L}_T(\lambda; \tilde{\mathbf{y}}^s(\hat{\theta}^{(0)})) \quad (3.33)$$

pour chaque $s = 1, \dots, S$ et calculez la moyenne

$$\tilde{\lambda}_{TS}(\hat{\theta}^{(0)}) = \frac{1}{S} \sum_{s=1}^S \tilde{\lambda}_T^s(\hat{\theta}^{(0)}) \quad (3.34)$$

- ◇ **Étape 4** : Actualiser numériquement l'estimation initiale $\hat{\theta}^{(0)}$ afin de minimiser la distance

$$\left[\hat{\lambda}_T - \tilde{\lambda}_{TS}(\theta) \right]' \Omega \left[\hat{\lambda}_T - \tilde{\lambda}_{TS}(\theta) \right] \quad (3.35)$$

où Ω est une matrice symétrique non négative définissant la métrique. En conséquence, l'estimateur par inférence indirecte est défini par

$$\hat{\theta}_{TS} = \hat{\theta}_{TS}(\Omega) = \operatorname{argmin}_{\theta \in \Theta} \left[\hat{\lambda}_T - \tilde{\lambda}_{TS}(\theta) \right]' \Omega \left[\hat{\lambda}_T - \tilde{\lambda}_{TS}(\theta) \right]. \quad (3.36)$$

En général, l'étape d'estimation est effectuée à l'aide d'un algorithme numérique, par exemple Newton-Raphson. Ensuite, pour une estimation donnée $\hat{\theta}^{(p)}$, la procédure donne $\hat{\theta}^{(p+1)}$ et le processus sera répété jusqu'à ce que la série $\hat{\theta}^{(p)}$ converge. L'estimateur est alors donné par :

$$\hat{\theta} = \lim_{p \rightarrow \infty} \hat{\theta}^{(p)}. \quad (3.37)$$

3.4 La méthode de Monte-Carlo

Les méthodes de Monte Carlo permettent d'estimer des quantités en utilisant la simulation de variables aléatoires. Les problèmes pouvant être rencontrés comprennent le calcul d'intégrales, les problèmes d'optimisation et la résolution de systèmes linéaires. La simplicité, la flexibilité et l'efficacité pour les problèmes en grande dimension de la méthode en font un outil intéressant, pouvant servir d'alternative ou de référence pour d'autres méthodes numériques.

Supposons que nous disposions de réalisations indépendantes $X_1(\omega), X_2(\omega), \dots$ (dans \mathbb{R}) d'une même loi (c'est à dire que les variables X_1, X_2, \dots sont i.i.d. et ont toute la même fonction de répartition, que nous noterons F).

Soit $t \in \mathbb{R}$. Pour tout i , nous introduisons la variable $U_i = \mathbf{1}_{]-w;t]}(X_i)$. Nous avons, en utilisant la propriété de croissance de l'espérance .

$$E((U_i^2)) \leq E(1) = 1 < \infty.$$

Donc, par la loi des grands nombres :

$$S_n := \frac{U_1 + \dots + U_n}{n} \xrightarrow[n \rightarrow +\infty]{} E(U_1) \quad (3.38)$$

où $E(U_1) = \mathbb{P}(X \leq t) \times 1 + \mathbb{P}(X > t) \times 0 = \mathbb{P}(X \leq t)$.

Cette remarque nous donne un moyen de calculer $F(t)$. Ce genre de méthode s'appelle la méthode de Monte-Carlo.

3.5 Les tests d'adéquation

Un test d'adéquation permet de statuer sur la compatibilité d'une distribution observée avec une distribution théorique associée à une loi de probabilité. Il s'agit de modélisation. Nous résumons une information brute, une série d'observations, à l'aide d'une fonction analytique paramétrée. L'estimation des valeurs des paramètres est souvent un préalable au test de conformité.

Au delà de la simplification, ce test permet de valider une appréhension du processus de formation des données, il permet de savoir si notre perception du réel est compatible avec ce que nous observons.

3.5.1 Les coefficients Skewness et Kurtosis

Pour caractériser la forme d'une distribution, c'est-à-dire pour préciser l'allure de la courbe des fréquences, il existe des coefficients permettant d'évaluer l'asymétrie d'une distribution et son aplatissement.

Le coefficient d'asymétrie (skewness)

Ce coefficient permet de mesurer le degré d'asymétrie de la distribution d'une variable aléatoire réelle et est donné par :

$$S_{sk} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{1}{n\sigma^3} \sum_{i=1}^n (X_i - \bar{X})^3. \quad (3.39)$$

$S_{sk} > 0$: La distribution est asymétrique vers la droite.

$S_{sk} = 0$: La distribution est qualifiée de normale et symétrique.

$S_{sk} < 0$: La distribution est asymétrique vers la gauche.

Plus le coefficient en valeur absolue est grand, plus l'asymétrie est marquée .

Mesure d'aplatissement (kurtosis)

Ce coefficient sert à mesurer le degré d'aplatissement de la distribution d'une variable aléatoire réelle et défini par :

$$S_{ku} = \frac{1}{n\sigma^4} \sum_{i=1}^n (X_i - \bar{X})^4 \quad (3.40)$$

$S_{ku} > 3$: La distribution est dite pointue et donc leptokurtotique.

$S_{ku} = 3$: La distribution est qualifiée de normale.

$S_{ku} < 3$: La distribution est dite écrasée et donc platykurtotique.

Plus le coefficient en valeur absolue est grand, plus l'aplatissement (ou le non aplatissement) est marqué.

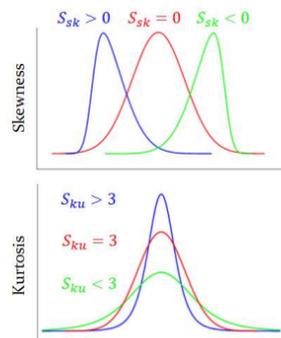


FIGURE 3.1 – Distributions avec différentes valeurs d'asymétrie et d'aplatissement

3.5.2 Test Jarque-Bera (ou test de Bowman-Shelton)

Le test de normalité de Jarque-Bera est également fondé sur les coefficients d'asymétrie et d'aplatissement. Il évalue les écarts simultanés de ces coefficients avec les valeurs de référence de la loi normale.

Le test de Jarque-Bera est un test de normalité. A noter que pour une loi normale le coefficient $S_{sk} = 0$ et $S_{ku} = 3$.

On peut traduire les hypothèses sous la forme :

$$H_0 : S_{sk} = 0 \text{ et } S_{ku} = 3$$

$$H_1 : S_{sk} \neq 0 \text{ ou } S_{ku} \neq 3$$

On remarque ainsi que si on rejettent H_0 , le test ne permet pas de connaître la raison principale du rejet (asymétrie ou aplatissement). On calcule

$$JB = \frac{n}{6} \left(S_{sk}^2 + \frac{(S_{ku} - 3)^2}{4} \right) \quad (3.41)$$

où n est le nombre d'observations, sachant que n doit être suffisamment grand ($n > 50$).

La statistique JB suit asymptotiquement une loi du χ^2 à 2 degrés de liberté.

Si les données suivent une loi normale, le test s'approche alors de 0 et on accepte (ne rejette pas) H_0 au seuil α .

La figure 3.2 représente les coefficients d'asymétries et aplatissement ainsi que le test de Jarque-Bera. Les valeurs pour lesquelles la variable y est le plus concentrée tournent autour de 38.

A partir de cet donné statistique :

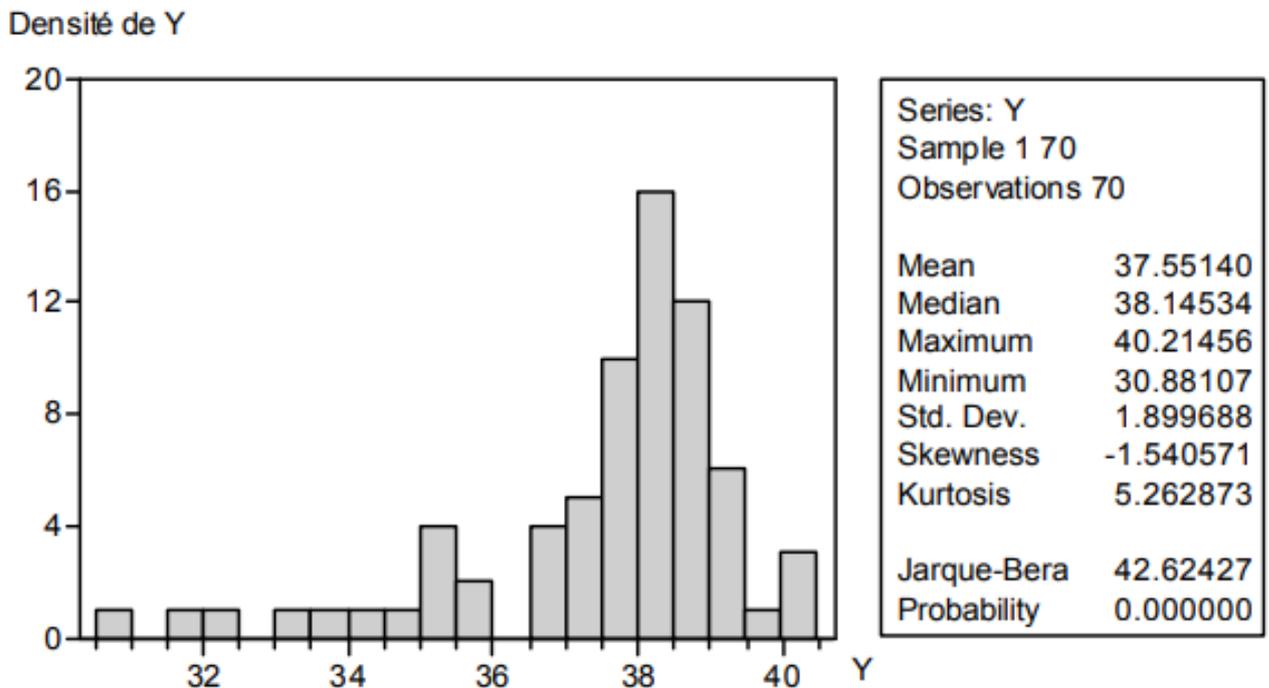


FIGURE 3.2 – Test de Jarque-Bera

1. la valeur de Skewness est de $-1,540571$ ce qui signifie que la dispersion est étalée à gauche (voir la figure 3.2)
2. La valeur de Kurtosis de la distribution normale est de 3. Dans notre cas, la valeur de Kurtosis est de $5,262873$ (la distribution est dite pointue).
3. pour le test de normalité de jarque-Bera nous allons vérifier les hypothèses suivantes :

$$H_0 : S_{sk} = 0 \text{ et } S_{ku} - 3 = 0$$

$$H_1 : S_{sk} \neq 0 \text{ ou } S_{ku} \neq 3$$

Cette série statistique est distribuée selon une loi de Khi deux à 2 degrés de liberté.

La table numérique de Khi deux à 2 degré de liberté nous donne la valeur critique de $5,99$ à 5% d'erreur. Or dans la figure 3.2, la valeur de Jarque-Bera est de $42,62427$ donc elle est supérieur à la valeur critique alors on rejette l'hypothèse H_0 . Ce qui signifie que notre distribution de la variable Y n'est pas normale.

3.5.3 Test de Shapiro-Wilk

Le test de Shapiro, aussi appelé test de Shapiro et Wilk, est un test de normalité d'une série d'observations d'une variable quantitative. Soit X une variable aléatoire réelle et X_1, X_2, \dots, X_n , n réalisations de X . On range alors ces observations par ordre croissant :

$$X_1 \leq X_2 \leq X_3 \leq \dots \leq X_n \quad (3.42)$$

Le test de Shapiro se réalise alors en 6 étapes :

1. On calcule la moyenne de cette série de mesures :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3.43)$$

2. On calcule le nombre T_n défini par :

$$T_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (3.44)$$

3. On calcule les différences suivantes :

$$\begin{aligned} d_1 &= X_n - X_1 \\ d_2 &= X_{n-1} - X_2 \\ &\vdots \\ d_i &= X_{n-i+1} - X_i \end{aligned} \quad (3.45)$$

4. On calcule alors le nombre W défini par :

$$W = \frac{1}{T_n} \sum_{j=1}^p (a_j d_j)^2 \quad (3.46)$$

Les coefficients a_j étant donnés en fonction de n par une table [23].

5. On choisit un risque (5% ou 1%) et on compare la valeur de W à une valeur W_{crit} , dite valeur critique, lue dans la table de Shapiro et Wilk.

6. La règle du test est alors la suivante :

- Si $W > W_{crit}$ on accepte, au risque choisi, l'hypothèse de normalité de la série de mesure.
- Si $W < W_{crit}$ on rejette l'hypothèse de normalité de la série de mesure.

Les valeurs seuils W_{crit} pour différents risques α et effectifs n sont lues dans la table de Shapiro-Wilk (voir [23]).

3.5.4 Diagramme Q-Q plot

Le "diagramme Quantile-Quantile" ou "diagramme Q-Q" ou "Q-Q plot" est un outil graphique permettant d'évaluer la pertinence de l'ajustement d'une distribution donnée à un modèle théorique. à partir de la série statistique observée, on calcule alors un certain nombre de quantiles. Si la série statistique suit bien la distribution théorique choisie, on devrait avoir les quantiles observés égaux aux quantiles associés au modèle théorique.

Les graphiques Quantile-Quantile Plot (QQ-Plot) permettent de tester graphiquement l'adéquation d'une famille de lois à des données. L'idée consiste à regarder si les quantiles de la famille de loi testée ($Q(p)$) et les quantiles de l'échantillon des $X_i, i = 1, \dots, n$, ($Q_n(p)$) sont linéairement liés, pour des valeurs de $p \in [0, 1]$. Autrement dit, on souhaite vérifier si les points $(F^{-1}(p), \widehat{F}_n^{-1}(p))$ sont alignés pour différentes valeurs de $p \in [0, 1]$, où F est la fonction de répartition de la vraie loi et \widehat{F}_n est la fonction de répartition empirique. En pratique, on représente graphiquement les points $(F^{-1}(\frac{i}{n+1}), x_{(i)})$ pour $i = 1, \dots, n$ où $x_{(i)}$ est la i -ème valeur de l'échantillon ordonnée.

Le QQ-Plot le plus utilisé dans le domaine de Théorie des Valeurs Extrêmes est celui de la loi exponentielle, où $F(x) = 1 - \exp(-x/\lambda)$ pour $x \geq 0$. La fonction quantile est donnée par $F^{-1}(p) = -\lambda \ln[1 - p]$, $p \in (0, 1)$.

Le QQ-Plot exponentiel consiste à tracer les points.

$$\left(-\ln \left(1 - \frac{i}{n+1} \right), x_{(i)} \right) \quad i = 1, \dots, n \quad (3.47)$$

Le diagramme QQ-plot s'interprète de la manière suivante :

- Si l'échantillon des x_i est un échantillon indépendant et identiquement distribué, issu d'une loi exponentielle, alors les points sont alignés selon une droite dont la pente est donnée par $1/\lambda$.

- le graphe a une forme concave, alors la distribution des x_i est à queue plus épaisse.

A l'inverse, si le graphe a une forme convexe, la distribution des x_i est à queue plus fine.

Nous illustrons dans la figure ci-dessus, les QQ-Plot exponentiel pour des échantillons de loi exponentielle, log-normale, Weibull ($\alpha < 1$) et Pareto. Comme nous voyons, l'échantillon de loi exponentielle suit une droite par rapport à la loi de référence (Loi exponentielle); l'échantillon de loi Pareto a la forme la plus concave parmi tous ces échantillons, cela correspond à nos

connaissances que la loi Pareto est à queue très épaisse ; la loi Log-normale est à queue moyenne, sa forme est donc légèrement concave. Pour une loi de Weibull avec $\alpha < 1$, il est difficile de la distinguer de la forme de loi Log-normale.

Le QQ-Plot peut être adapté à une multitude de lois pour tester graphiquement l'adéquation d'un échantillon à ces lois. Ainsi, dans le cas des lois de Pareto de paramètres (α, a) , le QQ-Plot Pareto consiste à tracer les points :

$$\left(-\ln \left(1 - \frac{i}{n+1} \right), \ln(x_{(i)}) \right) \quad i = 1, \dots, n \quad (3.48)$$

On s'attend alors à obtenir une droite de pente $1/\alpha$. Pour la loi log-normale, nous utilisons

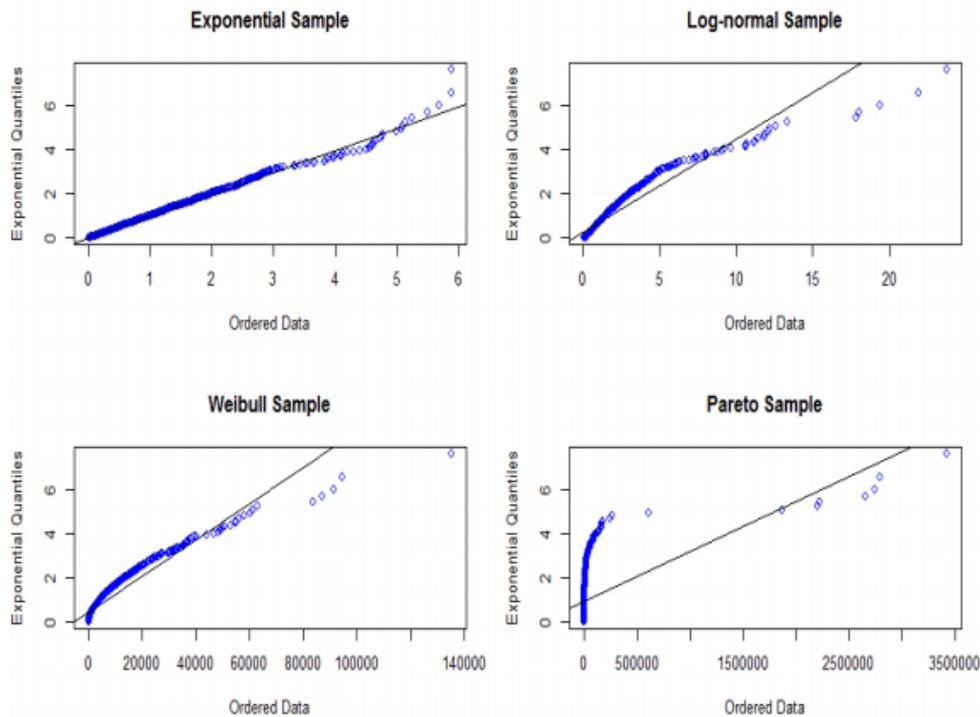


FIGURE 3.3 – Diagramme QQ-plot

la propriété que si l'échantillon suit une loi log-normale, alors son logarithme suit une loi normale. Le QQ-Plot log-normal consiste à tracer les points :

$$\left(\phi^{-1} \left(\frac{i}{n+1} \right), \ln(x_{(i)}) \right) \quad i = 1, \dots, n \quad (3.49)$$

où ϕ est la fonction de répartition de $\mathcal{N}(\mu, \sigma^2)$.

Enfin, pour la loi de Weibull, le QQ-Plot consiste à tracer les points :

$$\left(\ln \left(-\ln \left(1 - \frac{i}{n+1} \right) \right), \ln(x_{(i)}) \right) \quad i = 1, \dots, n \quad (3.50)$$

3.5.5 Box plot ou Boite à moustache

Box-plot que l'on appelle aussi boîte à moustache pour sa forme originale. Ce graphique tout simple permet de résumer une variable de manière simple et visuel, d'identifier les valeurs extrêmes et de comprendre la répartition des observations.

Définition 3.5.1. On appelle *diagramme en boîte* ou *boîte à moustache* d'une série, la représentation graphique suivante. Elle est composée de deux rectangles et de deux segments dont les longueurs correspondent aux paramètres de la série, représentés sur un axe gradué

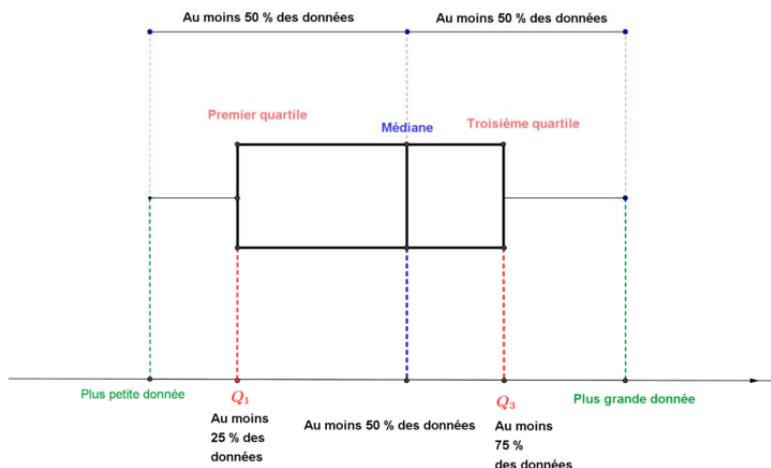


FIGURE 3.4 – Boîte à moustache

Remarque 3.5.1. .

- Un tel diagramme est aussi appelé « diagramme en boîte », « boîte à pattes » ou encore « diagramme de Tuckey » du nom de son concepteur.
- Les boîtes à moustaches sont un moyen simple pour comparer un même caractère sur plusieurs séries statistiques.

Exemple 3.5.1.

Le tableau 3.1 représente les notes de 24 élèves d'une classe lors d'un examen noté sur 100 points. Pour cela nous avons calculé les médiane et quartiles de cette série et représenté la boîte

78	79	77	59	57	65	65	67
68	67	59	54	64	68	72	74
72	72	76	77	76	74	77	76

TABLE 3.1 – Les notes de 24 élèves d'une classe lors d'un examen noté sur 100 points

à moustache 3.5. Tout d'abord nous avons trié cette série par ordre croissant comme l'illustre le tableau 3.2 Comme il y a 24 valeurs la médiane est la moyenne entre la 12^{ème} et la 13^{ème}

54	57	59	59	64	65	65	67
67	68	68	72	72	72	74	74
76	76	76	77	77	77	78	79

TABLE 3.2 – Tableau trié

valeur soit $M = \frac{72+72}{2} = 72$ le premier quartile est la 6^{ème} valeur soit $Q_1 = 65$ et le troisième quartile est la 18^{ème} valeur $Q_3 = 76$

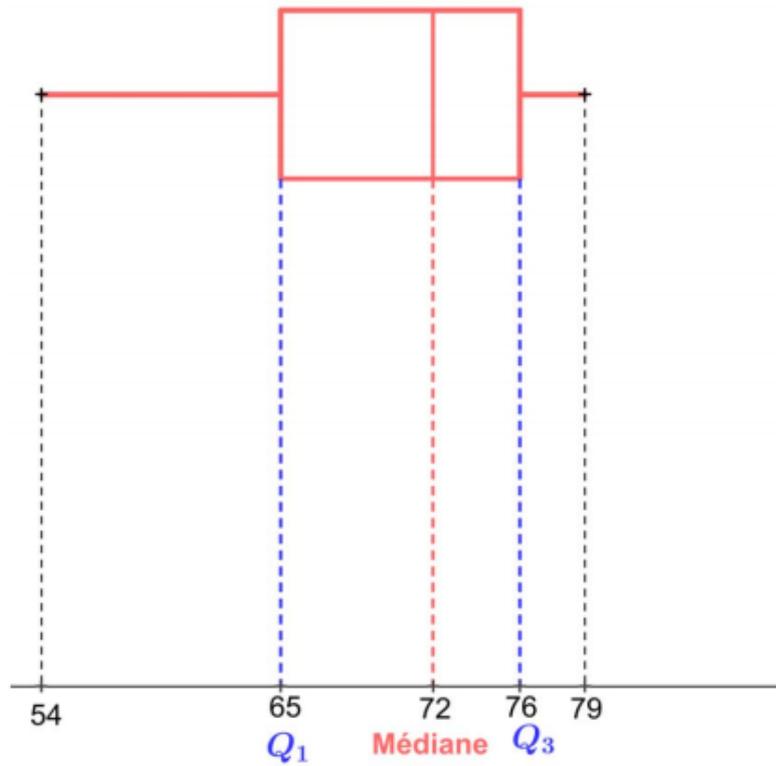


FIGURE 3.5 – Boite à moustache de la série

3.6 L'analyse de résidus

Les résidus ou (erreurs observées) sont définis comme étant les différences entre les valeurs observées et les valeurs estimées par un modèle de régression .

On les note communément comme suit :

$$e_i = X_i - \hat{X}_i$$

L'analyse des résidus est un outil important pour étudier la qualité et les défauts d'un modèle. De façon générale un peu décrire toute démarche statistique visant à expliquer une variable dépendante comme une décomposition de la variabilité en une partie attribuable aux variables explicatives (expliquée par le modèle) et une partie non-attribuables aux variables explicatives (reste).

$$Y = f(X_1, X_2, \dots, X_n) + \text{reste} \quad (3.51)$$

L'analyse des résidus se concentre sur :

$$\text{Reste} = Y - f(X_1, X_2, \dots, X_n) \quad (3.52)$$

L'étude des résidus permet de voir les défaut dans les restes, car une structure, une information intéressantes dans les résidus signifient que le modèle n'en as pas tenu compte. Dans ce contexte on trouve tous les problèmes, tels que :

Modèle linéaire non-approprié : Les défauts de linéarité seront visibles dans les résidus. Indicateurs qui ne respectent pas les hypothèses de normalité (notamment l'aspect symétrie)

3.6.1 MAE (Erreur absolue moyenne)

cette erreur est la forme la plus simple de métrique de régression que nous utilisons. Nous prenons la moyenne des résidus de la valeur réelle et de la valeur prédite. Nous prenons les valeurs absolues de chaque résidu, de sorte que les résidus positifs et négatifs ne s'annulent pas. La formule de MAE est :

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - \hat{X}_i| \quad (3.53)$$

Dans ce cas, chaque résidu sera proportionnel à l'erreur totale, ce qui signifie que les erreurs importantes seront également linéaires à l'erreur totale (il traite toutes les erreurs de la même manière). Un petit MAE suggérera que notre modèle est excellent pour les prédictions, tandis qu'un grand MAE suggère que notre modèle est mauvais pour les prédictions. Contrairement à MSE, nous ne mettons pas au carré les résidus, donc MAE est plus robuste aux valeurs aberrantes.

3.6.2 MSE (erreur quadratique moyenne)

dans MSE (Mean Square Error), nous mesurons la proximité d'une droite de régression par rapport à un ensemble de points.

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{X}_i)^2 \quad (3.54)$$

En raison de la quadrature des résidus, le MSE sera toujours plus grand que le MAE. Par conséquent, nous ne pouvons pas comparer le MSE avec celui du MAE. Cependant, nous pouvons comparer ces valeurs aux métriques d'un modèle de régression différent. Cela nous aidera à choisir le meilleur modèle pour les données.

Contrairement à MAE où chaque résidu est proportionnel à l'erreur totale, l'erreur augmente de façon quadratique dans MSE. Cela signifie finalement que MSE aura une erreur totale plus élevée en raison de la présence de valeurs aberrantes que dans MAE. Un MSE plus élevé signifie que le modèle sera pénalisé pour avoir fait des prédictions qui diffèrent considérablement de la valeur réelle. Cela signifie qu'une grande différence entre les valeurs prévues et réelles sera plus punie dans MSE que dans MAE.

3.6.3 RMSE (Racine carré de l'erreur quadratique moyenne)

La RMSE (Root Mean Square Error), comme son nom l'indique, n'est que la racine carrée de MSE. MSE a une valeur plus élevée parce que nous la mettons au carré. Afin d'amener cette valeur à la même échelle que celle de l'erreur de prédiction, nous utilisons RMSE. Cela facilite l'interprétation. Comme les valeurs MSE et RMSE quadrillent les résidus, elles sont également affectées par les valeurs aberrantes. La formule de RMSE :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \hat{X}_i)^2} \quad (3.55)$$

Conclusion

Nous retenons au cours de la présentation de ce chapitre quelques points principaux à savoir l'intérêt des méthodes d'estimation des paramètres d'un modèle donné et qu'il n'existe pas d'outils permettant de choisir au préalable une meilleure méthode d'estimation des paramètres mais que l'idéal sera de considérer celle qui ajuste au mieux les coefficients du modèle. Les tests de normalité, ainsi que les coefficients d'asymétrie et d'aplatissement joueront un grand rôle pour savoir si les données suivent une loi normale ou non, et ce qui est l'un des points forts de ce projet.

Le prochain chapitre fera l'objet de voir la pertinence de ces outils présents ici.

Les Processus tvARMA avec des innovations α -stable

Introduction

La classe des processus localement stationnaires a été introduite par [8] puis étudiée dans [4] et [9]. L'objectif est de définir une densité spectrale locale $f(\cdot, t)$ en tout point t du temps.

La première idée consiste à rendre le temps continu afin que le sens de "locale" corresponde au resserrement du temps autour d'un instant donné ; la seconde est d'ajouter une dimension afin de pouvoir définir des processus stationnaires locaux à chaque instant.

Dans ce chapitre, nous commencerons de faire une caractérisation sur les processus tvARMA selon la projection des lois α -stables et en terminerons avec quelques simulations et une application.

4.1 Les processus localement stationnaire

Une façon de généraliser le processus stationnaire est l'idée de stationnarité locale. C'est le cas où un processus stochastique X_t peut être stationnaire sur de petites périodes de temps, mais cette propriété de stationnarité change lentement sur une période de temps plus longue. Priestley (1965) a introduit les processus avec une représentation spectrale variant dans le temps :

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t} A_t(\lambda) d\xi(\lambda), \quad t \in \mathbb{Z} \quad (4.1)$$

où $\xi(\lambda)$ est un processus d'incrémentations orthogonal et $A_t(\lambda)$ est une fonction de transfert variant dans le temps. Notons que lorsque $A_t(\lambda)$ est constant par rapport à t , on obtient le cas particulier où X_t est globalement stationnaire.

Puisque les observations futures d'un processus non stationnaire peuvent ne contenir aucune information sur la structure probabiliste du processus au moment présent, la théorie des processus localement stationnaires est basée sur une approche asymptotique de remplissage. Comme en statistique non-paramétrique, l'idée de l'asymptotique de remplissage est que les fonctions dans le temps sont rédimensionnées sur l'intervalle unitaire afin d'obtenir une théorie asymptotique significative .

Processus linéaires localement stationnaires

La définition formelle d'un processus linéaire localement stationnaire est la suivante.

Définition 4.1.1. (Processus linéaires localement stationnaires)

On dit que la suite de processus stochastiques $X_{t,T}(t = 1, \dots, T)$ est un processus linéaire localement stationnaire si $X_{t,T}$ a une représentation :

$$X_{t,T} = \mu \left(\frac{t}{T} \right) + \sum_{j=-\infty}^{\infty} a_{t,T}(j) \varepsilon_{t-j} \quad (4.2)$$

où certaines conditions de régularité sont satisfaites pour μ , $a_{t,T}$ et ε_t .

Notez que si les ε_t sont stationnaires, ils peuvent être représentés comme suivant :

$$\varepsilon_t = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{i\lambda t} d\xi(\lambda) \quad (4.3)$$

où $\xi(\lambda)$ est un processus à moyenne nulle et à incréments orthonormés. Alors, la représentation 4.2) est fondamentalement équivalente à

$$X_{t,T} = \mu \left(\frac{t}{T} \right) + \int_{-\pi}^{\pi} e^{i\lambda t} A_{t,T}(\lambda) d\xi(\lambda) \quad (4.4)$$

Avec la fonction de transfert $A_{t,T}(\lambda) := \sum_{j=-\infty}^{\infty} a_{t,T}(j) e^{-i\lambda j}$ Pour continuer avec les conditions de régularité du processus linéaire localement stationnaire défini dans (4.2),

Soit $V(g)$ la variation totale d'une fonction g sur $[0, 1]$:

$$V(g) = \sup \left\{ \sum_{k=1}^m |g(x_k) - g(x_{k-1})| : 0 \leq x_0 < \dots < x_m \leq 1, m \in \mathbb{N} \right\}, \quad (4.5)$$

et pour un certain $\kappa > 0$ soit

$$\ell(j) := \begin{cases} 1, & |j| \leq 1 \\ |j| \log^{1+\kappa} |j| & |j| > 1 \end{cases} \quad (4.6)$$

Hypothèse 4.1.1. Supposons que la suite de processus stochastiques $X_{t,T}$ ait une représentation comme dans (4.2) et satisfasse les conditions suivantes :

(i)

$$\sup_t |a_{t,T}(j)| \leq \frac{K}{\ell(j)}, \quad (4.7)$$

avec K indépendant de T ;

(ii) il existe des fonctions $a(\cdot, j) : (0, 1] \rightarrow \mathbb{R}$ avec

$$\sup_u |a(u, j)| \leq \frac{K}{\ell(j)} \quad (4.8)$$

$$\sup_j \sum_{t=1}^T \left| a_{t,T}(j) - a \left(\frac{t}{T}, j \right) \right| \leq K, \quad (4.9)$$

$$V(a(\cdot, j)) \leq \frac{K}{\ell(j)} \quad (4.10)$$

(iii) μ a une variation totale finie, et les ε_t sont i.i.d. avec $E[\varepsilon_t] = 0$, $E[\varepsilon_s, \varepsilon_t] = 0$ pour $s \neq t$ et $E[\varepsilon_t^2] = 1$.

Pour certains résultats locaux, des hypothèses plus fortes doivent être imposées. Par exemple, pour certains i :

$$\sup_u \left| \frac{\partial^i \mu(u)}{\partial u^i} \right| \leq K \quad (4.11)$$

$$\sup_u \left| \frac{\partial^i a(u, j)}{\partial u^i} \right| \leq \frac{K}{\ell(j)}, \quad \text{for } j = 0, 1, \dots \quad (4.12)$$

et une hypothèse plus forte que (4.9) est

$$\sup_{t, T} \left| a_{t, T}(j) - a\left(\frac{t}{T}, j\right) \right| \leq \frac{K}{T\ell(j)} \quad (4.13)$$

Par conséquent, l'approximation stationnaire de (4.2) peut être construite

$$\tilde{X}_t(u) = \mu(u) + \sum_{j=-\infty}^{\infty} a(u, j) \varepsilon_{t-j} \quad (4.14)$$

Définition 4.1.2. (Densité spectrale et covariance variant dans le temps) Soit $X_{t, T}$ un processus stochastique dont la représentation est donnée par (4.2)

La fonction

$$f(u, \lambda) := \frac{1}{2\pi} |A(u, \lambda)|^2 \quad (4.15)$$

est la densité spectrale variable dans le temps de $X_{t, T}$, où $A(u, \lambda) := \sum_{j=-\infty}^{\infty} a(u, j) e^{-i\lambda j}$ est la fonction de transfert variable dans le temps, et

$$c(u, k) := \int_{-\pi}^{\pi} f(u, \lambda) e^{i\lambda k} d\lambda = \sum_{j=-\infty}^{\infty} a(u, k+j) a(u, j) \quad (4.16)$$

est la covariance variable dans le temps du retard k au temps recalculé u .

Sous l'hypothèse 4.1.1 avec la condition (4.13), on peut montrer que

$$\text{cov}(X_{[uT], T}, X_{[uT]+k, T}) = c(u, k) + O(T^{-1}) \quad (4.17)$$

uniformément en u et k . De plus, la condition (4.13) implique que

$$\sup_{t, \lambda} \left| A_{t, T}(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1} \quad (4.18)$$

4.2 tvARMA avec des innovations stables

Dans la suite de notre travail nous intéresserons au processus ARMA variable dans le temps (on le note tvARMA) avec des innovations α -stable. Considérons le modèle suivant :

$$\sum_{j=0}^p \alpha_j \left(\frac{t}{T}\right) X_{t-j, T} = \sum_{k=0}^q \beta_k \left(\frac{t}{T}\right) \gamma \left(\frac{t-k}{T}\right) \varepsilon_{t-k} \quad (4.19)$$

où ε_t i.i.d. et $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ avec $\alpha \in (0, 2)$.

Supposons que : $\alpha_0(u) \equiv \beta_0(u) \equiv 1$ et $\alpha_j(u) = \alpha_j(0), \beta_k(u) = \beta_k(0)$ pour $u < 0$. de plus supposons $\alpha_j(\cdot)$ et $\beta_k(\cdot)$, aussi bien que $\gamma^2(\cdot)$, sont de variation bornée.

L'équation (4.19) peut être écrite comme :

$$\Phi_{t,T}(B)X_{t,T} = \Theta_{t,T}(B)z_{t,T} \quad (4.20)$$

où

$$z_{t,T} = \gamma(t/T)\varepsilon_t$$

$$\Phi_{t,T}(B) = 1 + \alpha_1(t/T)B + \dots + \alpha_p(t/T)B^p$$

et

$$\Theta_{t,T}(B) = 1 + \beta_1(t/T)B + \dots + \beta_q(t/T)B^q$$

sont respectivement les polynômes caractéristiques autorégressifs (AR) et de moyenne mobile (MA).

La densité spectrale variant dans le temps est donné par :

$$f(u, \lambda) = \frac{\sigma^2(u) \left| \sum_{k=0}^q \beta_k(u) \exp(i\lambda k) \right|^2}{2\pi \left| \sum_{j=0}^p \alpha_j(u) \exp(i\lambda j) \right|^2} \quad (4.21)$$

4.2.1 L'existence et l'unicité d'une solution

Avant d'étudier les conditions de stationnarité locale des coefficients variant dans le temps, nous présentons un ensemble de conditions d'existence de régularité et d'unicité de solution du système basé sur les concepts définis par Shelton Peiris et Thavaneswaran [17], [18]

Définition 4.2.1. .

1. Le processus (4.20) est AR régulier (ou causal) s'il existe $a_{t,T}(j)$ tel que :

$$X_{t,T} = \sum_{j=0}^{\infty} a_{t,T}(j)\varepsilon_{t-j}. \quad (4.22)$$

satisfaisant $\sum_{j=0}^{\infty} |a_{t,T}(j)|^\delta < \infty$ pour tout t et $\delta = \min\{1, \alpha\}$.

2. Le processus (4.20) est MA régulier (ou inversible) s'il existe $b_{t,T}(j)$ tel que :

$$\varepsilon_t = \sum_{j=0}^{\infty} b_{t,T}(j)X_{t-j,T} \quad (4.23)$$

satisfaisant $\sum_{j=0}^{\infty} |b_{t,T}(j)|^\delta < \infty$ pour tout t et $\delta = \min\{1, \alpha\}$.

La série aléatoire dans (4.22) converge p.s. si et seulement si $\sum_{j=0}^{\infty} |a_{t,T}(j)|^\alpha < \infty$, et en appliquant la proposition 13.3.1 dans [2], il converge absolument si et seulement si $\sum_{j=0}^{\infty} |a_{t,T}(j)|^\delta < \infty$ avec $\delta = \min\{1, \alpha\}$.

Des arguments similaires sont appliqués à la représentation MA dans (4.23).

Proposition 4.2.1. *La série aléatoire de (4.22) converge p.s. si et seulement si*

$$\sum_{j=0}^{\infty} |a_{t,T}(j)|^{\alpha} < \infty \quad (4.24)$$

Preuve 4.2.1. *Supposons que : $\sum_{j=0}^{\infty} |a_{t,T}(j)|^{\alpha} < \infty$. en utilisant les propositions 3.3 et 3.8 dans [20], nous avons :*

$$a_{t,T}(j)\varepsilon_{t-j} \sim S_{\alpha} \left(\left(\frac{1}{\sqrt{2}} |a_{t,T}(j)|, \text{sign} [a_{t,T}(j)] \right), \beta, 0 \right), \quad (4.25)$$

et pour $0 < p < \alpha$, on a :

$$E |a_{t,T}(j)\varepsilon_{t-j}|^p = c_{\alpha,\beta}(p)^p \left(\frac{1}{\sqrt{2}} \right)^p |a_{t,T}(j)|^p \quad (4.26)$$

où $c_{\alpha,\beta}(p)$ est une constante . alors,

$$\sum_{j=0}^{\infty} E |a_{t,T}(j)\varepsilon_{t-j}|^p = c_{\alpha,\beta}(p)^p \left(\frac{1}{\sqrt{2}} \right)^p \sum_{j=0}^{\infty} |a_{t,T}(j)|^p < \infty \quad (4.27)$$

Puisque $\sum_{j=0}^{\infty} E |a_{t,T}(j)\varepsilon_{t-j}|^p < \infty$, il s'en suit que (4.22) converge p.s.

Réciproquement, la série aléatoire (4.22) converge p.s. implique que $a_{t,T}(j)\varepsilon_{t-j} \xrightarrow{p.s.} 0$. puis, puisque $\{a_{t,T}(j)\varepsilon_{t-j}\}$ est une suite indépendante, d'après le lemme de Borel-Cantelli, il existe $K_1 > 0$ tel que :

$$\sum_{j=0}^{\infty} P (|a_{t,T}(j)\varepsilon_{t-j}| > K_1) < \infty \quad (4.28)$$

D'après la proposition 2.2.1, nous obtenons

$$\lim_{\lambda \rightarrow \infty} \lambda^{\alpha} \sum_{j=1}^{\infty} P (|a_{t,T}(j)\varepsilon_{t-j}| > \lambda) = C_{\alpha} \left(\frac{1}{\sqrt{2}} \right) \sum_{j=1}^{\infty} |a_{t,T}(j)|^{\alpha} . \quad (4.29)$$

En d'autres termes, pour tout $\varepsilon > 0$, il existe K_2 tel que pour tout $\lambda > K_2$:

$$\frac{C_{\alpha} \left(\frac{1}{\sqrt{2}} \right) \sum_{j=1}^{\infty} |a_{t,T}(j)|^{\alpha} - \varepsilon}{\lambda^{\alpha}} < \sum_{j=1}^{\infty} P (|a_{t,T}(j)\varepsilon_{t-j}| > \lambda) < \frac{C_{\alpha} \left(\frac{1}{\sqrt{2}} \right) \sum_{j=1}^{\infty} |a_{t,T}(j)|^{\alpha} + \varepsilon}{\lambda^{\alpha}} \quad (4.30)$$

Soit $K = \max(K_1, K_2)$, nous avons :

$$\frac{C_{\alpha} \left(\frac{1}{\sqrt{2}} \right) \sum_{j=1}^{\infty} |a_{t,T}(j)|^{\alpha} - \varepsilon}{K^{\alpha}} < \sum_{j=1}^{\infty} P (|a_{t,T}(j)\varepsilon_{t-j}| > K) < \infty \quad (4.31)$$

Nous concluons donc que : $\sum_{j=1}^{\infty} |a_{t,T}(j)|^{\alpha} < \infty$.

Pour continuer, nous omettons l'indice T de la notation ci-dessus. Considérons l'équation de différence homogène :

$$\Phi_t(B)u_t = 0. \quad (4.32)$$

Si $\alpha_p(t/T) \neq 0$ pour tout t , il existe p solution linéairement indépendante $\psi_{1,t}, \psi_{2,t}, \dots, \psi_{p,t}$ tel que :

$$\Psi(t) = \begin{bmatrix} \psi_{1,t} & \cdots & \cdots & \psi_{p,t} \\ \psi_{1,t-1} & \ddots & & \psi_{p,t-1} \\ \vdots & & \ddots & \vdots \\ \psi_{1,t-p+1} & \cdots & \cdots & \psi_{p,t-p+1} \end{bmatrix} \quad (4.33)$$

est inversible pour tout t [15]. Par conséquent, nous pouvons définir :

$$G(t, s) = \Psi(t)[\Psi(s)]', \quad (4.34)$$

la matrice de fonction de Green unilatérale associée à l'opérateur $\Phi_t(B)$ de AR. On peut montrer que $G(t, s)$ est unique et invariant sous différentes solutions $\Psi(t)$ obtenues à partir de l'équation aux différences homogènes (4.32). De plus, la fonction unilatérale de Green associée à l'opérateur $\Phi_t(B)$ de AR est définie comme l'élément supérieur gauche dans (4.34)

$$g(t, s) = [G(t, s)]_{11} \quad (4.35)$$

qui est également unique et invariant. Maintenant, nous sommes prêts à établir les conditions de régularité AR et de régularité MA.

Théorème 4.2.1. Soit $\{X_{t,T}\}$ une suite de processus stochastique qui satisfait (4.20). Supposons que $\alpha_p(t/T) \neq 0$ pour tout t et $g(t, s)$ les fonctions unilatérales de Green associées à $\Phi_t(B)$, tel que $\sum_{s=-\infty}^t |g(t, s)|^\delta < \infty$, pour tout t .

Supposons aussi que $\sum_{s=-0}^q |\beta_j(\cdot)|^2 < \infty$ pour tout t et $\Phi_t(z)$ ($\Phi_t(z) \neq 0$ pour $|z| \leq 1$) et $\Theta_t(z)$ n'ont pas de racines communes. Ensuite, il y a une solution valide, donnée par :

$$X_{t,T} = \sum_{j=0}^{\infty} a_{t,T}(j) \varepsilon_{t-j} \quad (4.36)$$

à (4.20) avec des coefficients déterminés uniquement par :

$$a_{t,T}(j) = \begin{cases} 0, & j < 0, \\ \gamma\left(\frac{t-j}{T}\right), & j = 0, \\ \gamma\left(\frac{t-j}{T}\right) \sum_{k=0}^j \beta_k\left(\frac{t-j+k}{T}\right) g(t, t-j+k), & 0 \leq j \leq q, \\ \gamma\left(\frac{t-j}{T}\right) \sum_{k=0}^q \beta_k\left(\frac{t-j+k}{T}\right) g(t, t-j+k), & j > q \end{cases} \quad (4.37)$$

Preuve 4.2.2. En définissant $z_{t,T} = \gamma(t/T)\varepsilon_t$, avec les conditions de convergence absolue ci-dessus, la preuve est similaire à [18].

Théorème 4.2.2. Soit $\{X_{t,T}\}$ une séquence de processus stochastique qui satisfait (4.20). Supposons que $\beta_q(t/T) \neq 0$ pour tout t , et $h(t, s)$, la fonction de Green unilatérale associée à $\Theta_t(B)$, tel que $\sum_{s=-\infty}^t |h(t, s)|^\delta < \infty$, pour tout t . $\sum_{s=-0}^p |\alpha_j(\cdot)|^2 < \infty$ pour tout t , $\Phi_t(z)$ et

$\Theta_t(z)$ ($\Theta_t(z) \neq 0$ pour $|z| \leq 1$) n'ont pas de racines communes. Ensuite, le processus (4.20) est inversible et son inversion explicite est donnée par :

$$\varepsilon_t = \sum_{j=0}^{\infty} b_{t,T}(j) X_{t-j,T} \quad (4.38)$$

où $X_{t,T}$ désigne une solution arbitraire et les coefficients sont déterminés uniquement par :

$$b_{t,T}(j) = \begin{cases} 0, & j < 0, \\ \frac{1}{\gamma(\frac{t}{T})}, & j = 0, \\ \frac{1}{\gamma(\frac{t}{T})} \sum_{l=0}^k \alpha_k \left(\frac{t-j+k}{T}\right) h(t, t-j+k), & 0 \leq j \leq p, \\ \frac{1}{\gamma(\frac{t}{T})} \sum_{l=0}^q \alpha_k \left(\frac{t-j+k}{T}\right) h(t, t-j+k), & j > p \end{cases} \quad (4.39)$$

Théorème 4.2.3. Soit $\{X_{t,T}\}$ une suite de processus stochastique qui satisfait (4.19) c'est-à-dire AR régulier. La solution $X_{t,T}$ de la forme (4.22) est strictement stable et $X_{t,T} \sim S_{\alpha}(\sigma^*, \beta^*, 0)$, avec

$$\sigma^* = \left(\frac{1}{\sqrt{2}}\right) \left\{ \sum_{j=0}^{\infty} |a_{t,T}(j)|^{\alpha} \right\}^{1/\alpha} \quad \text{et} \quad \beta^* = \beta \left\{ \frac{\sum_{j=0}^{\infty} \text{sign}[a_{t,T}(j)] |a_{t,T}(j)|^{\alpha}}{\sum_{j=0}^{\infty} |a_{t,T}(j)|^{\alpha}} \right\} \quad (4.40)$$

Preuve 4.2.3. .

La forme explicite de la solution est simple puisque la combinaison linéaire de distributions stables est également stable la proposition 2.1.1 dans le deuxième chapitre et la propriété 1.2.6 de [16] implique que pour chaque t , la solution $X_{t,T}$ est strictement stable puisque chacun d'eux a un paramètre de localisation égal à 0.

Cherchons σ^* et β^*

Soit $\varepsilon_t \sim S_{\alpha}(1/\sqrt{2}, \beta, 0)$ et $X_{t,T} \sim S_{\alpha}(\sigma, \beta, \mu)$

Pour $\alpha \neq 1$ on a :

$$\phi_x(t) = \exp \left\{ -\sigma^{\alpha} |t|^{\alpha} \left\{ 1 - i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2} \right\} + i\mu t \right\}$$

On a :

$$\begin{aligned} X_{t,T} &= \sum_{j=0}^{\infty} a_{t,T}(j) \varepsilon_{t-j} \\ \phi_x(t) &= E(e^{i\theta X_{t,T}}) \\ &= E(e^{i\theta(\sum_{j=0}^{\infty} a_{t,T}(j) \varepsilon_{t-j})}) \\ &= E(e^{i\theta a_1 \varepsilon_{t-1}}) E(e^{i\theta a_2 \varepsilon_{t-2}}) \dots \\ &= \phi_{\varepsilon_{t-1}}(a_1 v) \phi_{\varepsilon_{t-2}}(a_2 v) \dots \end{aligned}$$

En appliquant Log sur l'équation on obtient :

$$\begin{aligned}
\text{Log}(\phi_{X_t}(v)) &= \sum_{j=0}^{\infty} \text{Log}(\phi_{\varepsilon_{t-j}}(a_j v)) \\
&= \sum_{j=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^{\alpha} |a_j v|^{\alpha} \left(1 - i\beta \text{sign}(a_j v) \tan\left(\frac{\alpha\pi}{2}\right)\right) \\
&= \sum_{j=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^{\alpha} |a_j v|^{\alpha} - i \sum_{j=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^{\alpha} |a_j v|^{\alpha} \beta \text{sign}(a_j v) \tan\left(\frac{\alpha\pi}{2}\right)
\end{aligned}$$

Comme $\text{sign}(a_j v) = \text{sign}(a_j) \text{sign}(v)$, On obtient :

$$\begin{aligned}
\text{Log}(\phi_{X_t}) &= \sum_{j=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^{\alpha} |a_j|^{\alpha} |v|^{\alpha} - i \sum_{j=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^{\alpha} |a_j|^{\alpha} |v|^{\alpha} \beta \text{sign}(a_j) \text{sign}(v) \tan\left(\frac{\alpha\pi}{2}\right) \\
&= \sum_{j=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^{\alpha} |a_j|^{\alpha} |v|^{\alpha} \left(1 - \frac{i\beta \sum_{j=0}^{\infty} |a_j|^{\alpha} \text{sign}(a_j) \tan\left(\frac{\alpha\pi}{2}\right)}{\sum_{j=0}^{\infty} |a_j|^{\alpha}}\right)
\end{aligned}$$

Par identification avec $\text{Log}(\phi_{X_t})$ on obtient les valeurs de β^* et σ^*

$$\sigma^* = \left(\frac{1}{\sqrt{2}}\right) \left\{ \sum_{j=0}^{\infty} |a_{t,T}(j)|^{\alpha} \right\}^{1/\alpha} \quad \text{et} \quad \beta^* = \beta \left\{ \frac{\sum_{j=0}^{\infty} \text{sign}[a_{t,T}(j)] |a_{t,T}(j)|^{\alpha}}{\sum_{j=0}^{\infty} |a_{t,T}(j)|^{\alpha}} \right\}$$

4.2.2 Localement stationnaire

Semblable à la proposition 2.4 dans [6], nous pouvons présenter la version correspondante pour les innovations stables. Puisqu'il ne s'agit pas d'un processus de second ordre, la densité spectrale variant dans le temps n'existe pas

Théorème 4.2.4. *Considérons le système d'équations de différence dans (4.19) satisfaisant les conditions régulières AR énoncées ci-dessus.*

Supposons que tous les $\alpha_j(\cdot)$ et $\beta_k(\cdot)$, ainsi que $\gamma^2(\cdot)$ sont de variation bornée. Ensuite, il existe une solution de la forme :

$$X_{t,T} = \sum_{j=0}^{\infty} a_{t,T}(j) \varepsilon_{t-j}, \tag{4.41}$$

qui réalise (4.8), (4.9) et (4.10).

Preuve 4.2.4. *Nous donnons la preuve que pour les processus tvAR ($q = 0$). L'extension aux processus tvARMA est alors simple. La preuve est similaire à celle de Künsch [13], qui a prouvé l'existence d'une solution de la forme (4.2) sous l'hypothèse que les fonctions $\alpha_i(u)$ sont continus.*

Soit :

$$\boldsymbol{\alpha}(u) = \begin{pmatrix} -\alpha_1(u) & -\alpha_2(u) & \dots & \dots & -\alpha_p(u) \\ 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

et $\alpha(u) = \alpha(0)$ pour $u < 0$. Puisque $\det(\lambda I_p - \alpha(u)) = \lambda^p \left(\sum_{j=0}^p \alpha_j(u) \lambda^{-j} \right)$, il s'ensuit que $\delta(\alpha(u)) \leq \frac{1}{1+\delta}$ pour tout u où $\delta(A) := \max\{|\lambda| : \lambda \text{ valeur propre de } A\}$.

Soit :

$$a_{t,T}(j) = \left(\prod_{\ell=0}^{j-1} \alpha \left(\frac{t-\ell}{T} \right) \right)_{11} \sigma \left(\frac{t-j}{T} \right)$$

et

$$X_{t,T} = \sum_{j=0}^{\infty} a_{t,T}(j) \varepsilon_{t-j}.$$

Il est facile de vérifier que $X_{t,T}$ est une solution de (4.19) à condition que les coefficients soient absolument sommables.

Pour le prouver, nous notons que pour tout $\varepsilon > 0$ et $u \in [0, 1]$, il existe une matrice $M(u)$ avec :

$$\|\alpha(u)\|_{M(u)} \leq \delta(\alpha(u)) + \varepsilon,$$

où $\|A\|_M := \sup \{\|Ax\|_M : \|x\|_M = 1\}$ et $\|x\|_M = \|M^{-1}x\|_1 = \sum_{i=1}^p |(M^{-1}x)_i|$.

comme les $\alpha_i(u)$ sont des fonctions de variation bornée (c'est-à-dire la différence de deux fonctions monotones), il existe pour tout $\varepsilon > 0$ une partition finie d'intervalles $I_1 \cup \dots \cup I_m = [0, 1]$ tel que : $|\alpha_i(u) - \alpha_i(v)| < \varepsilon$ pour tout i lorsque u, v sont dans le même I_k .

Soit $M_k := M(u_k)$ pour un arbitraire $u_k \in I_k$. Par conséquent, m (une partition) peut être choisi de telle sorte que :

$$\|\alpha(v)\|_{M_k} \leq \rho := \left(1 + \frac{\delta}{2}\right)^{-1} < 1 \quad \text{pour tout } v \in I_k.$$

Nous remplaçons maintenant le premier intervalle I_1 par $I_1 \cup (-\infty, 0)$ (n'oubliez pas que $\alpha(u) = \alpha(0)$ pour $u < 0$). il existe une constante c_0 tel que : $\|B\|_1 := \sum_{i,j} |B_{i,j}| \leq c_0 \|B\|_{M_k}$ pour tout k . pour t et n fixés, définissons maintenant $L_k := \{\ell \geq 0 : \frac{t-\ell}{n} \in I_k\}$ et $L_{k,j} := L_k \cap \{0, \dots, j-1\}$. Alors

$$\begin{aligned} |a_{t,T}(j)| &= \left| \left(\prod_{\ell=0}^{j-1} \alpha \left(\frac{t-\ell}{T} \right) \right)_{11} \sigma \left(\frac{t-j}{T} \right) \right| \leq \left\| \prod_{\ell=0}^{j-1} \alpha \left(\frac{t-\ell}{T} \right) \right\|_{11} \sigma \left(\frac{t-j}{T} \right) \\ &\leq \prod_{k=1}^m \left\| \prod_{\ell \in L_{k,j}} \alpha \left(\frac{t-\ell}{T} \right) \right\|_{11} \sigma \left(\frac{t-j}{T} \right) \\ &\leq c_0^m \prod_{k=1}^m \left\| \prod_{\ell \in L_{k,j}} \alpha \left(\frac{t-\ell}{T} \right) \right\|_{M_k} \sigma \left(\frac{t-j}{T} \right) \\ &\leq c_0^m \sup_u \sigma(u) \prod_{k=1}^m \rho^{|L_{k,j}|} = K \rho^j \quad (\text{puisque } m \text{ est fixé}) \end{aligned}$$

c'est-à-dire que nous avons prouvé (4.7).

comme $\|\alpha \left(\frac{t-k}{T} \right) - \alpha \left(\frac{t}{T} \right)\|_1 = \sum_{i=1}^p |\alpha_i \left(\frac{t-k}{T} \right) - \alpha_i \left(\frac{t}{T} \right)|$ et

$$a(u, j) := (\alpha(u)^j)_{11} \sigma(u),$$

$$\begin{aligned} & \sum_{t=1}^T |a_{t,T}(j) - a\left(\frac{t}{T}, j\right)| \\ & \leq \sum_{t=1}^T \sum_{k=1}^{j-1} \left\| \alpha\left(\frac{t}{T}\right)^k \left(\alpha\left(\frac{t-k}{T}\right) - \alpha\left(\frac{t}{T}\right) \right) \prod_{\ell=k+1}^{j-1} \alpha\left(\frac{t-\ell}{T}\right) \right\|_1 \sigma\left(\frac{t-j}{T}\right) \end{aligned}$$

Nous obtenons :

$$\begin{aligned} & + \sum_{t=1}^T \left\| \alpha\left(\frac{t}{T}\right)^j \right\|_1 \left| \sigma\left(\frac{t-j}{T}\right) - \sigma\left(\frac{t}{T}\right) \right| \\ & \leq \sum_{t=1}^T \sum_{k=1}^{j-1} c_0 \rho^k \sum_{i=1}^p \left| \alpha_i\left(\frac{t-k}{T}\right) - \alpha_i\left(\frac{t}{T}\right) \right| c_0^m \rho^{j-1-k} \\ & \leq K j^2 \rho^{j-1}, \end{aligned}$$

de la même manière pour (4.8), (4.9) et (4.10) .

Remarque 4.2.1. .

(1) Notez que $a_{t,T}(j) \approx a(t/T, j)$, $X_{t,T}$ peut être approximé par

$$\tilde{X}_{t,T} = \sum_{j=0}^{\infty} a\left(\frac{t}{T}, j\right) \varepsilon_{t-j} \quad (4.42)$$

qui converge p.s si et seulement si $\sum_{j=0}^{\infty} |a(t/T, j)|^\alpha < \infty$.

De plus, $\tilde{X}_{t,T} \sim S_\alpha(\sigma^+, \beta^+, 0)$, avec

$$\begin{aligned} \sigma^+ &= \frac{1}{\sqrt{2}} \left\{ \sum_{j=0}^{\infty} \left| a\left(\frac{t}{T}, j\right) \right|^\alpha \right\}^{1/\alpha} \\ \beta^+ &= \beta \left\{ \frac{\sum_{j=0}^{\infty} \text{sign} \left[a\left(\frac{t}{T}, j\right) \right] \left| a\left(\frac{t}{T}, j\right) \right|^\alpha}{\sum_{j=0}^{\infty} \left| a\left(\frac{t}{T}, j\right) \right|^\alpha} \right\} \end{aligned}$$

(2) $X_{t,T}$ dans (4.22) peut être exprimé comme une combinaison linéaire de variables aléatoires α -stable et $X_{t,T}$ est strictement stable avec le même indice de stabilité α .

(3) Remarquez que $X_{t,T}$ n'est pas strictement stationnaire, mais il peut être approché par $\tilde{X}_{t,T}$ qui est localement (strictement) stationnaire et strictement stable avec le même indice de stabilité.

(4) Une faible stationnarité n'a pas de sens puisque le second moment n'existe pas. Par conséquent, la représentation spectrale variable dans le temps n'existe pas.

(5) Soit $X_{1,T}, \dots, X_{T,T}$ la suite de solutions définies dans (4.22) et $\tilde{X}_{1,T}, \dots, \tilde{X}_{T,T}$ une suite du processus stochastique défini dans (4.42). Les deux processus sont strictement α -stables, puisque toutes les combinaisons linéaires sont strictement stables avec le même indice de stabilité. Cela signifie que la faible stationnarité est perdue mais elle est remplacée par le même comportement de queue tout au long du temps. C'est la raison pour laquelle nous appelons ce processus α -stable processus localement (strictement) stationnaire.

Si nous considérons les innovations symétriques α -stable (S), c'est-à-dire $\beta = 1$, la forme la plus simple est obtenue.

Corollaire 4.2.1. (*tvARMA avec des innovations symétriques stables*)

Soit $X_{t,T}$ être une séquence de processus stochastique qui satisfait (4.19) avec i.i.d. Innovations $S\alpha S$, c'est-à-dire $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, 0, 0)$ avec $\alpha \in (0, 2)$. Alors, il existe une solution de la forme (4.22). cette solution $X_{t,T}$ est symétrique, α -stable and $X_{t,T} \sim S_\alpha(\sigma^*, 0, 0)$, avec

$$\sigma^* = \frac{1}{\sqrt{2}} \left\{ \sum_{j=0}^{\infty} |a_{t,T}(j)|^\alpha \right\}^{1/\alpha} \quad (4.43)$$

De même que le cas général, $X_{t,T}$ peut être approximé par $\tilde{X}_{t,T} \sim S_\alpha(\sigma^+, 0, 0)$ comme dans (4.42), avec

$$\sigma^+ = \left\{ \sum_{j=0}^{\infty} \left| a\left(\frac{t}{T}, j\right) \right|^\alpha \right\}^{1/\alpha} \quad (4.44)$$

Remarque 4.2.2. .

(i) Notons que dans le cas d'innovations $S\alpha S$, puisque $X_{t,T}$ dans (4.22) peut être exprimé comme une combinaison linéaire de variables aléatoires $S\alpha S$, $X_{t,T}$ est $S\alpha S$ avec le même indice de stabilité α .

(ii) En même temps, $X_{t,T}$ n'est pas strictement stationnaire, mais il peut être approximé par $\tilde{X}_{t,T}$ qui est localement (strictement) stationnaire et $S\alpha S$ avec le même indice de stabilité.

(iii) La stationnarité faible n'a pas de sens puisque le second moment n'existe pas. Par conséquent, 4.21 n'existe pas.

(iv) De la même manière, considérons $X_{1,T}, \dots, X_{T,T}$ la séquence de solutions définie en (4.22) et $\tilde{X}_{1,T}, \dots, \tilde{X}_{T,T}$ la séquence du processus stochastique défini en (4.42). Les deux processus sont symétriques α -stables, puisque toutes les combinaisons linéaires sont symétriques stables avec le même indice de stabilité. Cela signifie que la stationnarité faible est perdue mais qu'elle est remplacée par le même comportement de queue tout au long du temps.

4.3 Simulation

Dans cette section nous présentons la simulation de Monte Carlo (MC) afin d'étudier les propriétés d'estimation par inférence indirecte. Tous les programmes et les routines de simulation ont été implémentés en langage R. Nous présentons un scénario pour chacun des modèles suivants, mais des valeurs différentes de α ont été sélectionnées. D'autres scénarios ont été réalisés pour chaque cas et des résultats similaires ont été obtenus. Pour chaque scénario, des simulations ont été effectuées pour $T = 500, 1000$ et 1500 observations basées sur $R = 1000$ répliquions indépendantes. L'inférence indirecte a été réalisée en utilisant $S = 100$. Pour α connu, nous avons également effectué l'estimation de Whittle par bloc (BWE), proposée par Dahlhaus [4], pour comparer l'estimation de la structure temporelle du modèle avec l'inférence indirecte. La suggestion de taille de bloc $N = \lfloor T^{0.8} \rfloor$ et de décaler chaque bloc de $S = \lfloor 0.2N \rfloor$ unités de temps à partir de [7] est utilisée.

4.3.1 le cas de α connu

tvAR(1) α -stable

Considérons $tvAR(p)$ dans (4.19) avec $p = 1$ et $\gamma(t/T) = \gamma$:

$$X_{t,T} + \alpha_1\left(\frac{t}{T}\right)X_{t-1,T} = \gamma\varepsilon_t \quad (4.45)$$

où $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ avec α et β connus. Nous illustrons comment l'inférence indirecte peut être utilisée pour le $tvAR(1)$ avec la forme paramétrique linéaire du coefficient variant dans le temps $\alpha_1(u) = \theta_0 + \theta_1 u$, et on considère que $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ pour α connu. Par conséquent, les paramètres de l'IM sont $\theta = (\theta_0, \theta_1, \gamma)$.

La simulation a été effectuée en supposant que les paramètres α et β sont connus ($\alpha = 1.9$ et $\beta = 0.9$) et inconnus $(\theta_0, \theta_1, \gamma) = (-0.3, 0.8, 1)$. Il est important de signaler que puisque α est proche de 2, toutes les réplifications pour le BWE ont convergé. Ce résultat est attendu puisque les distributions d'innovation se rapprochent de la distribution gaussienne pour α proche de 2.

Le tableau 4.1 présente la moyenne MC et l'erreur pour les deux méthodes d'estimation. Notez que la moyenne MC des estimations indirectes semble consistente, c'est-à-dire qu'elles se rapprochent des paramètres réels et présentent des erreurs standard plus faibles lorsque T augmente. D'autre part, les MC moyennes du BWE sont différentes des paramètres réels et présentent des erreurs standard plus élevées par rapport à notre approche d'estimation.

T	Indirect estimates			BWE		
	θ_0	θ_1	γ	$\theta_0^{(W)}$	$\theta_1^{(W)}$	$\gamma^{(W)}$
500	-0.2952 (0.0881)	0.7897 (0.1523)	0.9966 (0.0366)	-0.2880 (0.1172)	0.7825 (0.2216)	1.2086 (0.6352)
1000	-0.2975 (0.0585)	0.7926 (0.1028)	0.9996 (0.0260)	-0.2917 (0.0811)	0.7845 (0.1545)	1.2197 (0.4734)
1500	-0.2974 (0.0494)	0.7958 (0.0793)	0.9997 (0.0209)	-0.2940 (0.0639)	0.7926 (0.1162)	1.2709 (0.8738)

TABLE 4.1 – Les moyennes MC et les erreurs standard pour différentes tailles d'échantillon T , en utilisant des estimateurs indirects et BWE en supposant que α et β sont connus, à partir de $tvAR(1)$ stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.9, 0.9, 0.3, 0.8, 1)$, basé sur $R = 1000$ réplifications.

Le tableau 4.2 présente les coefficients d'aplatissement et d'asymétrie des deux méthodes. En général, toutes les estimations indirectes présentent un kurtosis plus faible et une asymétrie proche de 0.

T		Indirect estimates			BWE		
		θ_0	θ_1	γ	$\theta_0^{(W)}$	$\theta_1^{(W)}$	$\gamma^{(W)}$
500	Kur	3.0330	2.8565	3.1076	3.3783	3.0944	375.8563
	Skw	0.1388	-0.1354	0.1241	-0.0129	-0.0875	16.6778
1000	Kur	3.1678	3.2835	2.7390	2.8722	2.9543	95.2261
	Skw	0.0341	-0.0260	0.0437	0.0057	-0.1047	8.4487
1500	Kur	3.1024	3.0645	2.9329	3.7487	6.1799	187.9560
	Skw	-0.0299	0.0248	0.0026	0.1707	-0.5935	12.4661

TABLE 4.2 – Le Kurtosis et l’asymétrie des estimations indirectes et BWE pour différentes tailles d’échantillon T , en supposant que α et β sont connus, pour α -stable tvAR(1) avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.9, 0.9, -0.3, 0.8, 1)$, basé sur $R = 1000$ répétitions.

La figure 4.1 montre que les estimations de densité de chaque paramètre. Ils montrent aussi que la norme de l’erreur devient plus petite lorsque T augmente. Avec les résultats des tableaux 4.1 et 4.2, nous pouvons conclure que les estimations indirectes se comportent mieux que le BWE en termes de moyenne, d’erreur standard, d’asymétrie et d’aplatissement. Par conséquent, les résultats de la simulation indiquent que l’inférence indirecte fonctionne bien.

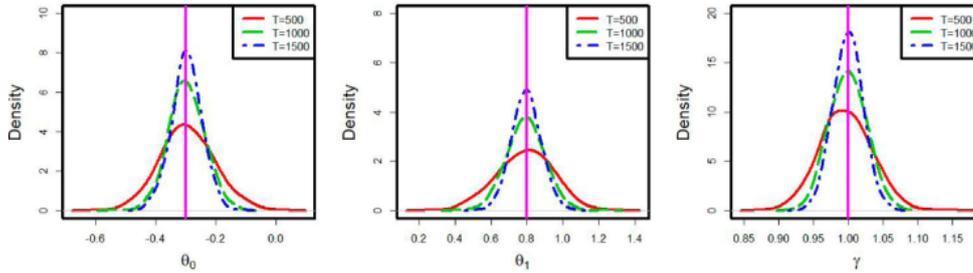


FIGURE 4.1 – Estimations de θ_0 , θ_1 et γ pour différentes tailles d’échantillon, sur la base de $R = 1000$ réplifications de tvAR(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.9, 0.9, -0.3, 0.8, 1)$, en utilisant l’inférence indirecte.

tvMA(1) α -stable

Dans cette section, nous avons effectué des simulations pour un tvMA(q) dans (4.19) avec $q = 1$ et $\gamma(t/T) = \gamma$:

$$X_{t,T} = \gamma \left\{ \varepsilon_t + \beta_1 \left(\frac{t}{T} \right) \varepsilon_{t-1} \right\} \quad (4.46)$$

où $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ avec α et β connu. L’inférence indirecte est employée pour la forme paramétrique linéaire du coefficient variant dans le temps $\beta_1(u) = \theta_0 + \theta_1 u$, et nous considérons que $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ pour les α et β connus. Par conséquent, le vecteur de paramètres du modèle d’intérêt est $\theta = (\theta_0, \theta_1, \gamma)$.

On suppose que $\alpha = 1.1$ et $\beta = -0.2$ et inconnus $(\theta_0, \theta_1, \gamma) = (0.35, -0.6, 1.2)$. Pour le BWE, nous considérons seulement $R = 939, 978$ et 978 réplifications avec des estimations converges pour $T = 500, 1000$ et 1500 , respectivement. Ce résultat est attendu car BWE suppose que le

second moment est fini. La moyenne MC, l'erreur standard, l'aplatissement et l'asymétrie des estimations de la simulation sont présentées dans les tableaux 4.3 et 4.4 et les estimations de densité dans la figure 4.2 .

T	Indirect estimates			BWE ^a		
	θ_0	θ_1	γ	$\theta_0^{(W)}$	$\theta_1^{(W)}$	$\gamma^{(W)}$
500	0.3561 (0.0298)	-0.5888 (0.0577)	1.1989 (0.0600)	0.3424 (0.1418)	-0.5427 (0.3084)	18.7932 (38.4343)
1000	0.3545 (0.0186)	-0.5953 (0.0352)	1.1986 (0.0412)	0.3386 (0.0870)	-0.5532 (0.1955)	47.5752 (232.0620)
1500	0.3536 (0.0131)	-0.5982 (0.0244)	1.1986 (0.0331)	0.3357 (0.0747)	-0.5555 (0.1690)	49.4572 (178.3984)

TABLE 4.3 – Les moyennes MC et l'erreur standard pour différentes tailles d'échantillon T , en utilisant des estimateurs indirects et BWE en supposant que α et β sont connus, pour tvMA(1) α -stable avec $(\alpha, \theta_0, \theta_1, \gamma) = (1.1, 0.2, 0.35, -0.6, 1.2)$, basé sur $R = 1000$ répétitions

T		Indirect estimates			BWE ^a		
		θ_0	θ_1	γ	$\theta_0^{(W)}$	$\theta_1^{(W)}$	$\gamma^{(W)}$
500	Kur	7.9023	6.3460	2.9050	6.9952	7.3434	233.1652
	Skw	1.2950	0.0800	0.2117	0.1961	1.1869	13.5324
1000	Kur	9.8633	10.4926	2.8616	11.7454	8.6755	385.9194
	Skw	1.6510	0.7121	0.0841	0.8633	1.5096	17.6374
1500	Kur	8.1466	20.6873	2.8140	9.7011	10.5014	156.1843
	Skw	1.5194	1.4762	0.1493	-0.1837	1.9926	11.4943

TABLE 4.4 – Kurtosis et asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T en supposant que α et β connus à partir de tvMA(1) α -stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.1, 0.2, 0.35, 0.6, 1.2)$ basé sur $R = 1000$ répétitions.

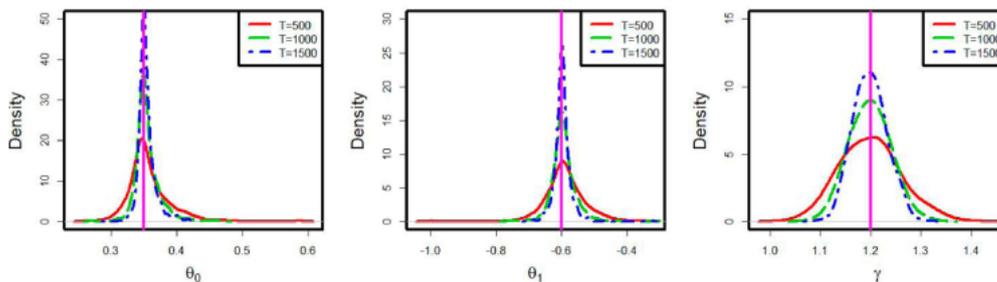


FIGURE 4.2 – Estimations de θ_0, θ_1 et γ pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications de tvMA(1) α -stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.1, -0.2, 0.35, -0.6, 1.2)$ en utilisant l'inférence indirecte.

Comme dans le cas précédent, les estimations indirectes semblent consistant et l'erreur standard diminue lorsque T augmente. Dans ce cas, puisque α est plus petit, la distribution des estimations indirectes a des queues plus lourdes, et elles ont une aplatissement et une asymétrie similaires à celles de l'estimations BWE, à l'exception du paramètre γ , où l'estimation indirecte se comporte mieux. De plus, en terme d'erreur standard et de moyenne MC, ils se comportent toujours mieux que le BWE. Nous concluons que l'inférence indirecte a une bonne performance

ARMA(1,1) α -stable

La troisième simulation a été réalisée avec le cas de $tvARMA(p,q)$ avec $p = 1, q = 1$ et $\gamma(t/T) = \gamma$:

$$X_{t,T} + \alpha_1 \left(\frac{t}{T} \right) X_{t-1,T} = \gamma \left\{ \varepsilon_t + \beta_1 \left(\frac{t}{T} \right) \varepsilon_{t-1} \right\} \quad (4.47)$$

où $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ avec α et β connu. On suppose une forme paramétrique linéaire des coefficients variant dans le temps $\alpha_1(u) = \theta_{a0} + \theta_{a1}u$ et $\beta_1(u) = \theta_{b0} + \theta_{b1}u$. Par conséquent, le paramètre de l'IM est $\theta = (\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma)$. La simulation a été faite en supposant que $\alpha = 1.8, \beta = 0.3$ et $(\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (-0.4, 0, 1, 0, 1, 0.3, 1, 1)$. Pour BWE, $R = 989, 996$ et 994 réplifications avec des estimateurs convergents sont incluses pour $T = 500, 1000$ et 1500 , respectivement.

La moyenne MC, l'erreur standard, l'aplatissement et l'asymétrie des estimations de la simulation $tvARMA(1,1)$ sont présentées dans les tableaux 4.3 et 4.4 et les estimations de densité dans la figure 4.5.

T	Indirect estimates					BWE				
	θ_{00}	θ_{a1}	θ_{b0}	θ_{b1}	γ	$\theta_{a0}^{(W)}$	$\theta_{a1}^{(W)}$	$\theta_{b0}^{(W)}$	$\theta_{b1}^{(W)}$	$\gamma^{(W)}$
500	-0.4000 (0.1360)	0.1061 (0.2222)	0.0987 (0.1501)	0.3097 (0.2395)	0.9976 (0.0386)	-0.3917 (0.1952)	0.1021 (0.3522)	0.1078 (0.2130)	0.3049 (0.3810)	1.4151 (0.7603)
1000	-0.3921 (0.1001)	0.0881 (0.1617)	0.1064 (0.1053)	0.2905 (0.1652)	0.9982 (0.0290)	-0.3850 (0.1409)	0.0815 (0.2535)	0.1105 (0.1470)	0.2880 (0.2599)	1.4919 (0.5806)
1500	-0.3992 (0.0754)	0.1021 (0.1269)	0.0988 (0.0793)	0.3060 (0.1285)	0.9982 (0.0232)	-0.3939 (0.1085)	0.0926 (0.1955)	0.1055 (0.1155)	0.2964 (0.2040)	1.5538 (0.8009)

FIGURE 4.3 – MC moyenne et erreur standard pour différentes tailles d'échantillon T , en utilisant des estimateurs indirects et BWE en supposant que α et β sont connus, à partir de $tvARMA(1,1)$ stable avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.8, 0.3, -0.4, 0.1, 0.1, 0.3, 1)$, sur la base de $R = 1000$ réplifications.

T		Indirect estimates					BWE ^a				
		θ_{a0}	θ_{a1}	θ_{b0}	θ_{b1}	γ	$\theta_{a0}^{(W)}$	$\theta_{a1}^{(W)}$	$\theta_{b0}^{(W)}$	$\theta_{b1}^{(W)}$	$\gamma^{(W)}$
500	Kur	3.3650	3.1791	3.3657	3.4297	2.9935	2.9112	3.0692	3.2168	3.2822	203.2823
	Skw	0.2754	-0.2426	-0.0746	-0.1274	0.1839	0.1699	-0.1329	-0.2024	-0.0422	12.0737
1500	Kur	3.3964	3.5054	3.4470	3.4690	3.0327	3.4242	3.2166	3.2601	3.0064	41.6803
	Skw	0.2002	-0.1558	0.0100	-0.1184	0.2460	0.3149	-0.2253	0.0267	-0.1713	4.9608
1500	Kur	3.5817	3.1935	3.7052	3.3930	2.9790	2.9176	2.8801	3.3685	3.3091	96.2273
	Skw	0.2895	-0.1097	0.0137	-0.0730	0.0718	0.0809	0.0268	-0.1083	0.0372	7.7396

FIGURE 4.4 – Kurtosis et asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T, en supposant que α et β connus à partir de α -stable tvARMA(1,1) avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.8, 0.3, -0.4, 0.1, 0.1, 0.3, 1)$, sur la base de $R = 1000$ réplifications.

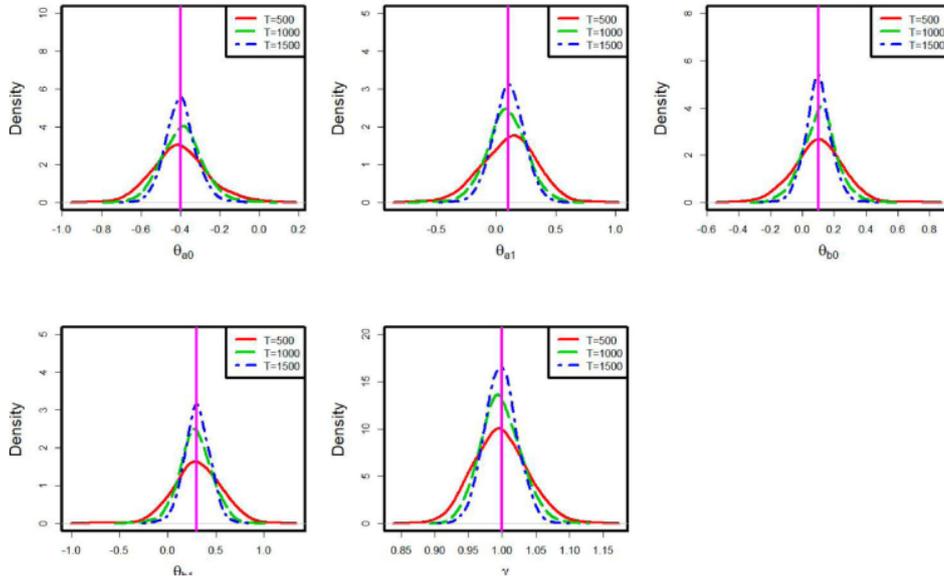


FIGURE 4.5 – Estimations de $\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}$ et γ pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications de tvARMA(1,1) stable avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.8, 0.3, -0.4, 0.1, 0.1, 0.3, 1)$ en utilisant l'inférence indirecte.

En général, la distribution des estimations indirectes a des queues plus lourdes, et le kurtosis et l'asymétrie sont similaires au BWE (à l'exception du paramètre γ , les estimations indirectes se comportent mieux).

Cependant, en termes d'erreur standard et de moyenne MC, ils se comportent bien mieux que le BWE. Par conséquent, l'inférence indirecte fonctionne bien pour tvARMA(1,1)

4.3.2 Le cas α inconnu

tvAR(1) α -stable

Considérons le modèle tvAR(1)

$$X_{t,T} + \alpha_1 \left(\frac{t}{T} \right) X_{t-1,T} = \gamma \left(\frac{t}{T} \right) \varepsilon_t \quad (4.48)$$

où $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ avec β connu . Ici, l'inférence indirecte est employée au tvAR(1) dans 4.45 avec la forme paramétrique linéaire du coefficient variant dans le temps $\alpha_1(u) = \theta_0 + \theta_1 u$, et $\gamma(u) = \gamma_0 + \gamma_1 u$. Les paramètres de IM sont $\theta = (\theta_0, \theta_1, \alpha, \gamma_0, \gamma_1)$. Pour AM, la même forme paramétrique avec t -distribution en supposant que v inconnu est utilisé, c'est-à-dire $\lambda = (\theta_0^{(A)}, \theta_1^{(A)}, v, \gamma_0^{(A)}, \gamma_1^{(A)})$

La simulation a été effectuée en supposant que $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, -0.6, 0.5, 0.1)$. Le tableau 4.5 présente la moyenne MC et l'erreur standard des estimations. Notez que la moyenne MC des estimations indirectes semble consistant. Le tableau 4.6 présente l'aplatissement et l'asymétrie des estimations indirectes. Toutes les estimations indirectes ne présentent pas de kurtosis proche de 3 et le skewness proche de 0. En effet, ils sont similaires au cas où α est connu.

T	Indirect estimates									
	θ_0	θ_1	α	γ_0	γ_1	$\theta_0^{(A)}$	$\theta_1^{(A)}$	v	$\gamma_0^{(A)}$	$\gamma_1^{(A)}$
500	0.3482	-0.5980	1.4083	0.4922	0.1111	0.3482	-0.5980	1.8853	0.3994	0.0897
	(0.0406)	(0.0715)	(0.0737)	(0.0527)	(0.0960)	(0.0407)	(0.0716)	(0.2351)	(0.0446)	(0.0778)
1000	0.3492	-0.5986	1.4037	0.4974	0.1033	0.3492	-0.5986	1.8622	0.4033	0.0834
	(0.0244)	(0.0430)	(0.0520)	(0.0370)	(0.0661)	(0.0244)	(0.0429)	(0.1570)	(0.0311)	(0.0533)
1500	0.3498	-0.5988	1.4000	0.4976	0.1011	0.3499	-0.5988	1.8478	0.4030	0.0818
	(0.0187)	(0.0323)	(0.0417)	(0.0305)	(0.0546)	(0.0187)	(0.0323)	(0.1244)	(0.0255)	(0.0441)

TABLE 4.5 – La moyenne MC et erreur type pour différentes tailles d'échantillon T à l'aide d'estimateurs indirects en supposant $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, 0.6, 0.5, 0.1)$ avec β connu pour tvAR(1) α -stable basé sur $R = 1000$ réplifications

T		Indirect estimates				
		θ_0	θ_1	α	γ_0	γ_1
500	kur	3.7767	3.6800	3.1078	2.8924	3.0343
	skw	-0.1369	0.1213	0.2730	0.1583	0.0156
1000	kur	4.7209	3.8513	2.7680	3.1397	3.0710
	skw	-0.1548	0.1008	0.0889	0.0672	-0.0654
1500	kur	4.2029	3.8664	2.7385	3.0881	3.0266
	skw	0.1274	-0.0192	0.0967	0.0973	0.0108

TABLE 4.6 – L'amplitude et l'asymétrie des estimations indirectes pour différentes tailles d'échantillon T en supposant que $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, -0.6, 0.5, 0.1)$ avec β connu pour tvAR(1) α -stable sur la base de $R = 1000$ réplifications.

Enfin, la figure 4.6 montre les estimations de densité de chaque paramètre. On remarque que l'erreur standard deviennent plus petits quand T augmente. Nous concluons que la distribution des estimations indirectes semble être consistant pour la longueur de ces échantillons.

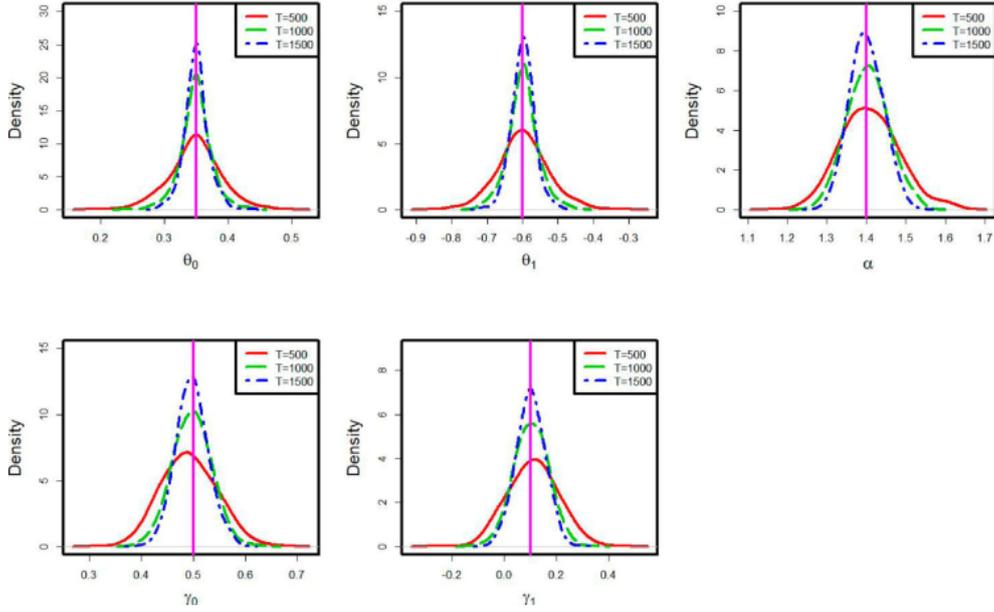


FIGURE 4.6 – Estimations de $\theta_0, \theta_1, \alpha, \gamma_0$ et γ_1 pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications pour tvAR(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma_0, \gamma_1) = (1.4, 0, 0.35, -0.6, 0.5, 0.1)$ en utilisant l'inférence indirecte.

tvMA(1) α stable

L'inférence indirecte pour le modèle 4.46 avec α inconnu est illustrée. Les paramètres de IM est $\theta = (\theta_0, \theta_1, \alpha, \gamma)$ et les paramètres de AM est $\lambda = (\theta_0^{(A)}, \theta_1^{(A)}, v, \gamma^{(A)})$. La simulation a été effectuée en supposant que $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$. La moyenne MC et l'erreur standard des estimations des deux modèles (IM et AM) sont présentées dans le tableau 4.7, et l'aplatissement et l'asymétrie sont présentés dans le tableau 4.8. Avec les estimations de densité présentées à la figure 4.7, les estimations indirectes semblent être consistant avec des longueur de chemin d'échantillon. Un résultat intéressant est que si $\alpha < 2$, alors l'IM a une variance infinie, l'AM a été estimée avec $v > 2$, c'est-à-dire une variance finie.

T	Indirect estimates							
	Model of interest				Auxiliary model			
	θ_0	θ_1	α	γ	$\theta_0^{(A)}$	$\theta_1^{(A)}$	v	$\gamma^{(A)}$
500	-0.3518 (0.0699)	0.4016 (0.1245)	1.7566 (0.0739)	0.7008 (0.0296)	-0.3518 (0.0694)	0.4016 (0.1237)	3.9795 (1.0183)	0.3810 (0.0390)
1000	-0.3487 (0.0446)	0.3987 (0.0787)	1.7527 (0.0559)	0.6999 (0.0229)	-0.3486 (0.0445)	0.3987 (0.0788)	3.8307 (0.6414)	0.3776 (0.0299)
1500	-0.3504 (0.0375)	0.4009 (0.0663)	1.7525 (0.0457)	0.7003 (0.0187)	-0.3502 (0.0373)	0.4007 (0.0661)	3.7874 (0.4852)	0.3785 (0.0242)

TABLE 4.7 – La moyenne MC et erreur standard pour différentes tailles d'échantillon T en utilisant des estimateurs indirects en supposant $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$ avec β connu à partir de α -stable tvMA(1) basé sur $R = 1000$ réplifications.

T		Indirect estimates			
		θ_0	θ_1	α	γ_0
500	kur	3.8667	3.2718	2.8731	3.3445
	skw	-0.0413	-0.0030	-0.1140	0.0003
1000	kur	3.7260	3.4565	2.8049	2.9412
	skw	0.0436	0.0582	-0.0081	0.1446
1500	kur	3.6876	3.4211	3.0489	3.0002
	skw	-0.0043	0.0187	-0.2133	0.0557

TABLE 4.8 – Kurtosis et asymétrie des estimations indirectes et BWE pour différentes tailles d'échantillon T en supposant que $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$ avec β connu pour tvMA(1) α -stable basé sur $R = 1000$ réplifications.

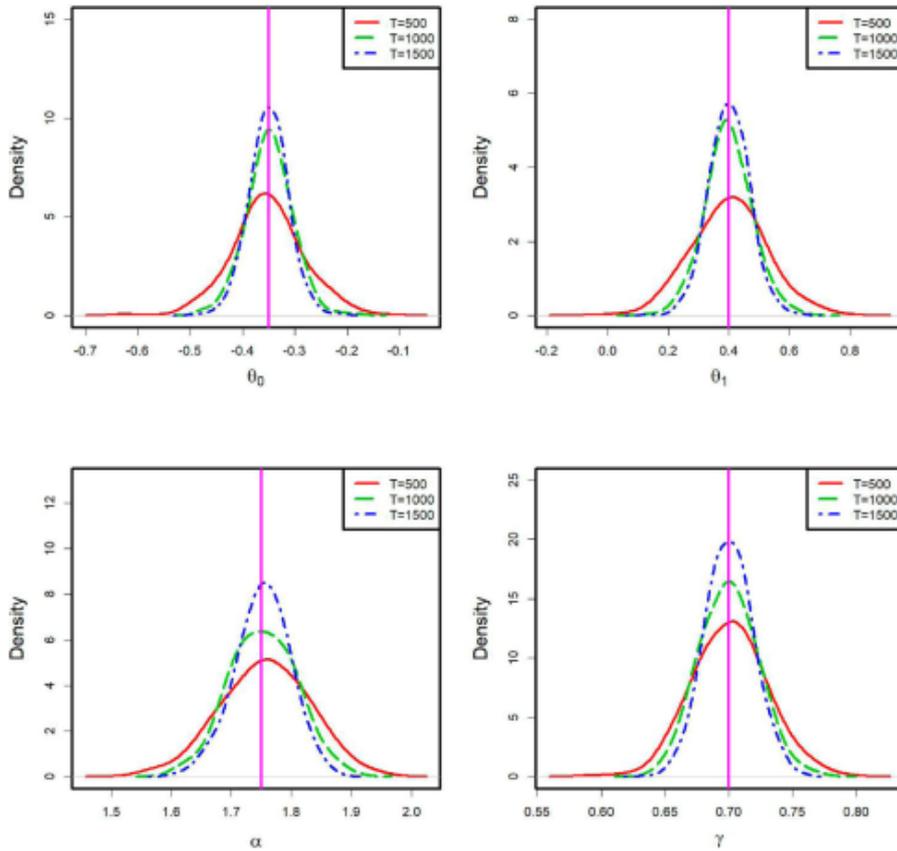


FIGURE 4.7 – Estimations de θ_0, θ_1 , pour différentes tailles d'échantillons basées sur $R = 1000$ réplifications de tvMA(1) stable avec $(\alpha, \beta, \theta_0, \theta_1, \gamma) = (1.75, 0.2, -0.35, 0.4, 0.7)$ en utilisant l'inférence indirecte.

tvARMA(1,1) α -stable

Enfin, la simulation a été faite pour le cas de tvARMA(1,1) dans 4.47, mais α est supposé inconnu. Les coefficients variant dans le temps sont supposés linéaires, c'est-à-dire $\alpha_1(u) =$

$\theta_{a0} + \theta_{a1}u$ et $\beta_1(u) = \theta_{b0} + \theta_{b1}u$, et $\varepsilon_t \sim S_\alpha(1/\sqrt{2}, \beta, 0)$ pour β connu. Par conséquent, les paramètres de IM sont $\theta = (\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha, \gamma)$, tandis que AM a le paramètre $\lambda = (\theta_{a0}^{(A)}, \theta_{a1}^{(A)}, \theta_{b0}^{(A)}, \theta_{b1}^{(A)}, v, \gamma^{(A)})$. La simulation a été effectuée en supposant $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha, \gamma) = (1.3, 0, -0.2, -0.4, 0.2, 0.3, 1.1)$

La moyenne MC et l'erreur standard des estimations de l'IM et de l'AM sont présentées dans le tableau 4.9, l'aplatissement et l'asymétrie sont présentés dans le tableau 4.10.

	T	θ_{a0}	θ_{a1}	θ_{b0}	θ_{b1}	α	γ
Model of interest	500	-0.2036	-0.3932	0.1971	0.3064	1.3018	1.0923
		(0.0585)	(0.0869)	(0.0587)	(0.0891)	(0.0698)	(0.0587)
	1000	-0.2005	-0.3986	0.2003	0.3004	1.3045	1.0976
		-(0.0319)	(0.0489)	(0.0329)	(0.0504)	(0.0471)	(0.0433)
	1500	-0.1998	-0.3999	0.2012	0.2983	1.2998	1.0953
		(0.0233)	(0.0359)	(0.0250)	(0.0374)	(0.0390)	(0.0347)
	T	$\theta_{a0}^{(A)}$	$\theta_{a1}^{(A)}$	$\theta_{b0}^{(A)}$	$\theta_{b1}^{(A)}$	v	$\gamma^{(A)}$
Auxiliary model	500	-0.2036	-0.3936	0.1971	0.3062	1.5904	0.7465
		(0.0584)	(0.0864)	(0.0586)	(0.0889)	(0.1731)	(0.0940)
	1000	-0.2006	-0.3986	0.2003	0.3006	1.5917	0.7542
		(0.0319)	(0.0483)	(0.0329)	(0.0505)	(0.1160)	(0.0668)
	1500	-0.1998	-0.4009	0.2012	0.2983	1.5772	0.7487
		(0.0232)	(0.0344)	(0.0250)	(0.0374)	(0.0947)	(0.0540)

TABLE 4.9 – La moyenne MC et l'erreur standard pour différentes tailles d'échantillon T en utilisant des estimateurs indirects en supposant $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha, \gamma) = (1.3, 0, 0.2, 0.4, 0.2, 0.3, 1.1)$ avec β connu pour tvARMA(1,1) α -stable basé sur $R = 1000$ réplifications

T		Indirect estimates					
		θ_{a0}	θ_{a1}	θ_{b0}	θ_{b1}	α	γ
500	kur	5.2893	4.5345	6.0807	5.5860	3.0705	3.3815
	skw	0.2593	-0.1945	-0.1951	0.1297	0.1406	0.2709
1000	kur	4.6288	4.1073	4.5984	4.1306	3.4796	2.9077
	skw	-0.1406	0.1144	0.0360	-0.0328	0.1167	-0.0713
1500	kur	4.9301	4.0790	5.1471	4.5091	3.1301	3.0701
	skw	0.0236	-0.1964	0.0964	-0.2378	0.1004	0.0833

TABLE 4.10 – Le Kurtosis et skewness des estimations indirectes et BWE pour différentes tailles d'échantillon ($T = 500, 1000, 1500$) en supposant que $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha, \gamma) = (1.3, 0, -0.2, -0.4, 0.2, 0.3, 1.1)$ avec β connu à partir de α tvARMA(1,1) stable basé sur $R = 1000$ réplifications.

Les estimations de densité sont présentées dans la figure 4.8. Encore une fois, les estimations indirectes semblent être consistant. De plus, si nous comparons ces résultats avec les résultats de

simulation du α connu, ils présentent une erreur standard, un aplatissement et une asymétrie similaires.

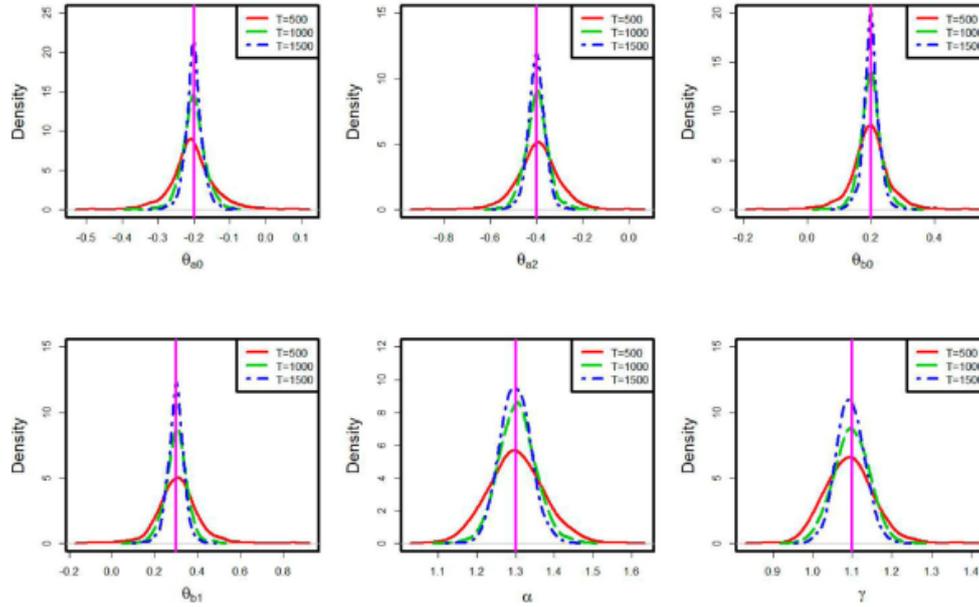


FIGURE 4.8 – Estimations de la densité de $\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \alpha$ et γ pour différentes tailles d'échantillon basées sur $R = 1000$ répliquions de tvARMA(1,1) α -stable avec $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.3, 0, -0.2, -0.4, 0.2, 0.3, 1.1)$ en utilisant l'inférence indirecte.

4.4 Application

L'énergie éolienne est l'énergie du vent, dont le principe de fonctionnement repose sur la transformation de l'énergie cinétique en énergie électrique : le vent fait tourner des pales qui font elles même tourner le générateur de l'éolienne. À son tour le générateur transforme l'énergie mécanique du vent en énergie électrique. Le courant électrique est ensuite transformé et injecté dans le réseau électrique pour alimenter nos foyers. il peut être stocké pour être utilisé plus tard.

Dans cette partie, nous illustrons une application pour la puissance éolienne totale produite dans les parcs éoliens offshore en Allemagne du 16/06/2015 à 00 :00 au 27 / 07/2015 à 00 :00 ($T=1008$ heures), obtenue à partir des ensembles de données EMHIRES (European Meteorological High resolution RES time series) . La raison de la sélection d'un petit segment des données est due au fait que la série temporelle entière a une structure plus complexe, comme la saisonnalité .

Notre objectif c'est de déterminer le modèle optimal pour la prévision de l'énergie éolienne en Allemagne avec la méthode de Box & Jenkins.

Transformation

Avant d'effectuer des transformations sur la série nous allons faire une description graphique. La figure 4.9 représente l'évolution de l'énergie éolienne du 16 juin 2015 au 27 juillet

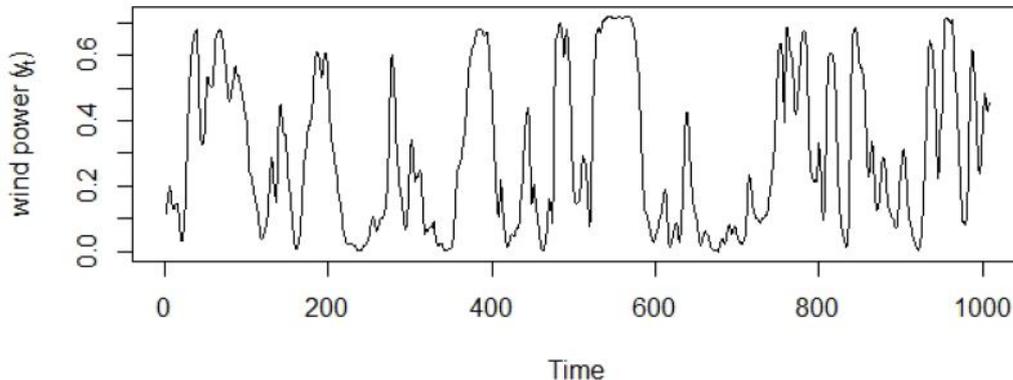


FIGURE 4.9 – Évolution de l'énergie éolienne fonction du temps

2015 nous permet de voir rapidement que les séries d'énergie éolienne ne sont pas stationnaires.

Maintenant nous allons utiliser la fonction *decompose* de R pour visualiser les différentes composantes de la série. La décomposition de l'historique des données permet de mettre en évidence graphiquement en plus d'une tendance mais également la présence d'une saisonnalité. On voit maintenant après les transformations que le graphe 4.10 de la série est stationnaire

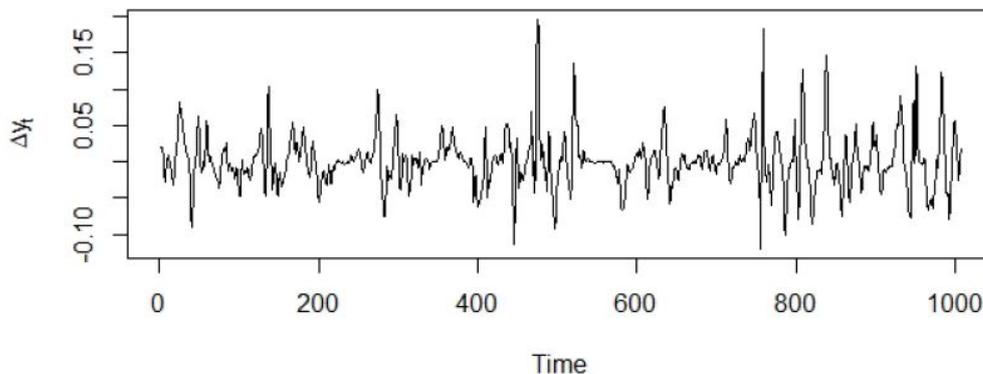


FIGURE 4.10 – Évolution de la série stationnaire de l'énergie éolienne fonction du temps

sans tendance. Et ainsi on s'intéresse à l'étape suivante dans la méthode de Box & Jenkins qui est la détermination du P_{max} et du Q_{max} .

Détermination du P_{max} et du Q_{max}

Pour déterminer le P_{max} et Q_{max} , on observe le graphe des auto-corrélations et des auto-corrélations partielles. Les graphes des auto-corrélations et auto-corrélations partielles nous

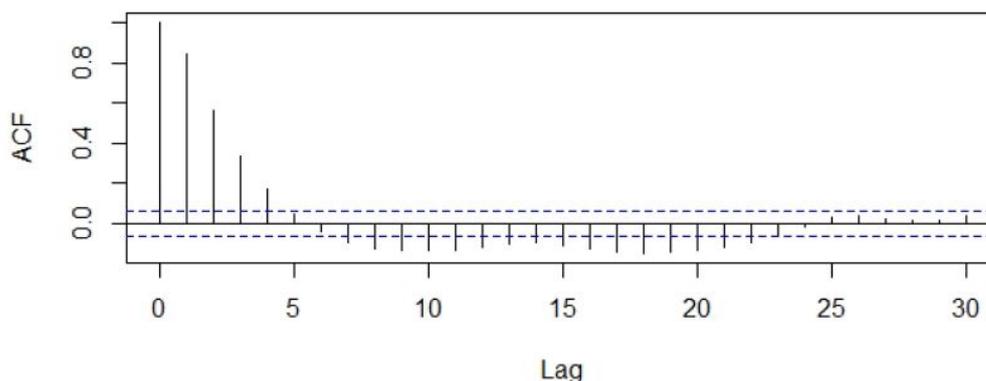


FIGURE 4.11 – Graphe des Auto-corrélations

donnent respectivement $Q_{max} = 5$ et $P_{max} = 4$.

Donc on doit simuler 31 modèles avec constante et 30 modèles sans constante. Mais on doit éliminer les modèles dont les résidus ne sont pas normalement distribués et n'ont pas de moyenne nul ou proches de zéro.

Ainsi à l'aide d'un programme simple du logiciel R on élimine 59 modèles et on se retrouve avec 2 modèles dont les résidus sont distribués normalement, de moyenne nulle, non corrélés

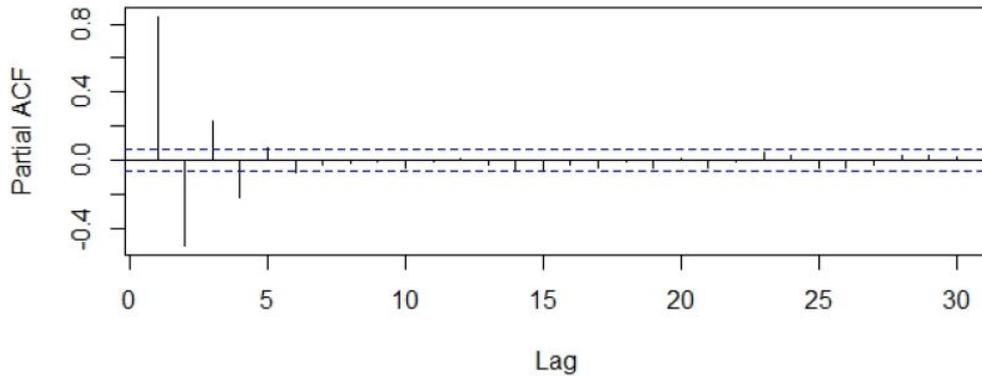


FIGURE 4.12 – Graphe des Auto-corrélations partielles

et homogènes. Et comme nous voulons faire une comparaison entre deux distributions dans ce mémoire alors nous allons prendre l'équivalent de ces deux modèles retenus en supposant que leurs innovations sont α -stable ce qui nous fait 4 maintenant ainsi les modèles avec des résidus Valides sont : $tvAR(4)$, $tvARMA(1,1)$, α - $tvAR(4)$, α - $tvARMA(1,1)$.

C'est à dire :

1. Le modèle $tvARMA(1,1)$ avec des coefficient linéaire, $\alpha_1(u) = \theta_{a0} + \theta_{a1}(u)$, $\beta_1(u) = \theta_{b0} + \theta_{b1}(u)$ et $\gamma(u) = \gamma_0 + \gamma_1(u)$.
2. Le modèle $tvAR(4)$ avec des coefficients, $\alpha_1(u) = \theta_{a0} + \theta_{a1}(u)$, $\alpha_2(u) = \theta_{b0} + \theta_{b1}(u)$, $\alpha_3(u) = \theta_{c0} + \theta_{c1}(u)$, $\alpha_4(u) = \theta_{d0} + \theta_{d1}(u)$ et $\gamma(u) = \gamma_0 + \gamma_1(u)$.

Donc nous allons passer à l'étape suivante de la méthode Box & Jenkins qui est la détermination des coefficients des modèles sélectionnés.

Détermination des coefficients des modèles retenus

A l'aide d'un simple commande du logiciel R, on trouve que :

1. Pour le modele $tvARMA(1,1)$ les coefficients sont :
 $(\alpha, \beta, \theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \gamma) = (1.3, 0, -0.2, -0.4, 0.2, 0.3, 1.1)$
2. et pour le $tvAR(4)$ les coefficients sont :

$$(\theta_{a0}, \theta_{a1}, \theta_{b0}, \theta_{b1}, \theta_{c0}, \theta_{c1}, \theta_{d0}, \theta_{d1}, \gamma_0, \gamma_1) = (-1.598, 0.33, 0.91, 0.02, -0.057, -0.714, -0.13, 0.55, 0.008, 0.015)$$

Ainsi nous allons passer à l'étape suivante de la méthode Box & Jenkins qui est la sélection du modèle optimal parmi tant d'autres.

Sélection du modèle optimal

Pour sélectionner le modèle optimal, nous allons utiliser le critère des erreurs pour ce choix de sélection. Le tableau 4.11 représente des erreurs résiduelles de chacun de ces modèles. En regardant la MAE, on peut voir que les modèles α -stable $tvARMA(1,1)$ et $tvAR(4)$ sont très proches mais la MAE de α -stable $tvAR(4)$ est plus petite que celle des autres modèles donc si on ne considérait que la MAE des modèles on aurait choisi α -stable $tvAR(4)$ comme le modèle optimal pour faire la prévision de l'énergie éolienne. Comme par précaution nous avons calculé aussi la RMSE et la MSE de chacun des modèles alors on trouve qu'en tenant compte de tous

résidus le modèle $tvAR(4)$ donne des résidus très petits par rapport aux autres modèles. Finalement nous allons retenir le modèle $tvAR(4)$ comme le modèle optimal pour faire la prévision de l'énergie éolienne.

Model	MSE	RMSE	MAE
tvARMA(1,1)	0.000248	0.015739	0.009675
α -stable tvARMA(1,1)	0.000257	0.016028	0.009469
tvAR(4)	0.000242	0.015542	0.009468
α -stable tvAR(4)	0.000256	0.015993	0.009094

TABLE 4.11 – Erreurs résiduelles des différents modèles

Corrélogramme des résidus du modèle optimal

On observe à partir du graphe des auto-corrélogramme 4.13 que toutes les autocorrélations

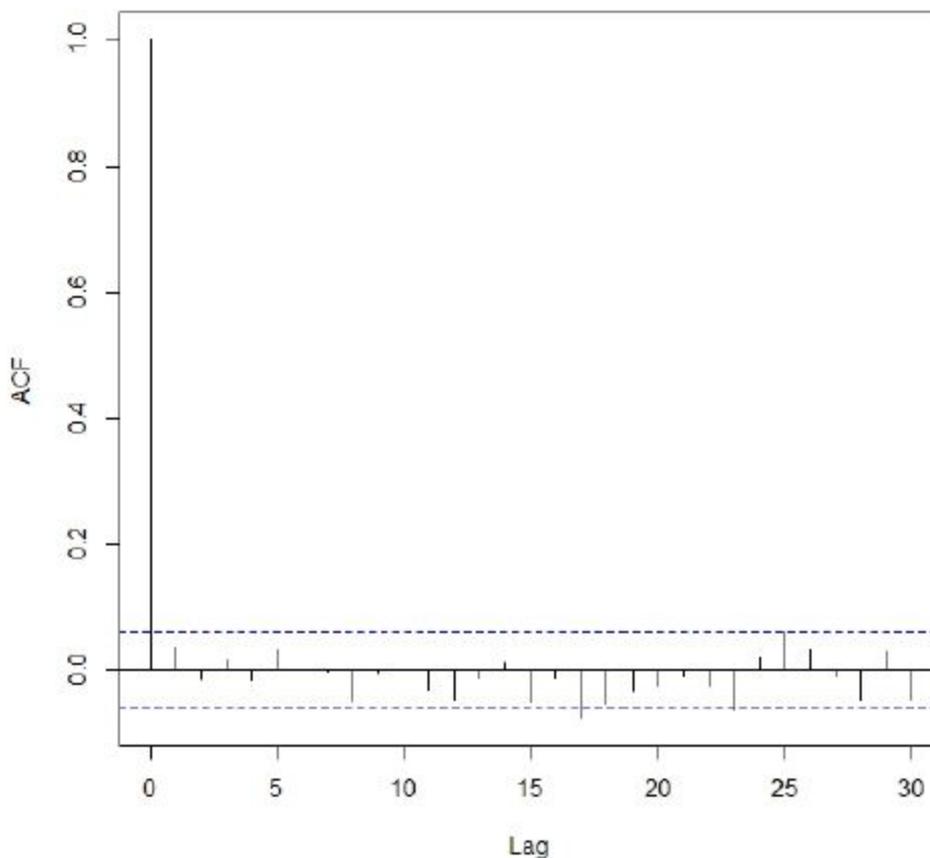


FIGURE 4.13 – Graphe des Auto-corrélogrammes partiels des résidus

partielles sont à l'intérieur de l'intervalles de confiances donc ce qui assure notre choix de sélection de modèle. Donc finalement, on retient le modèle $tvAR(4)$.

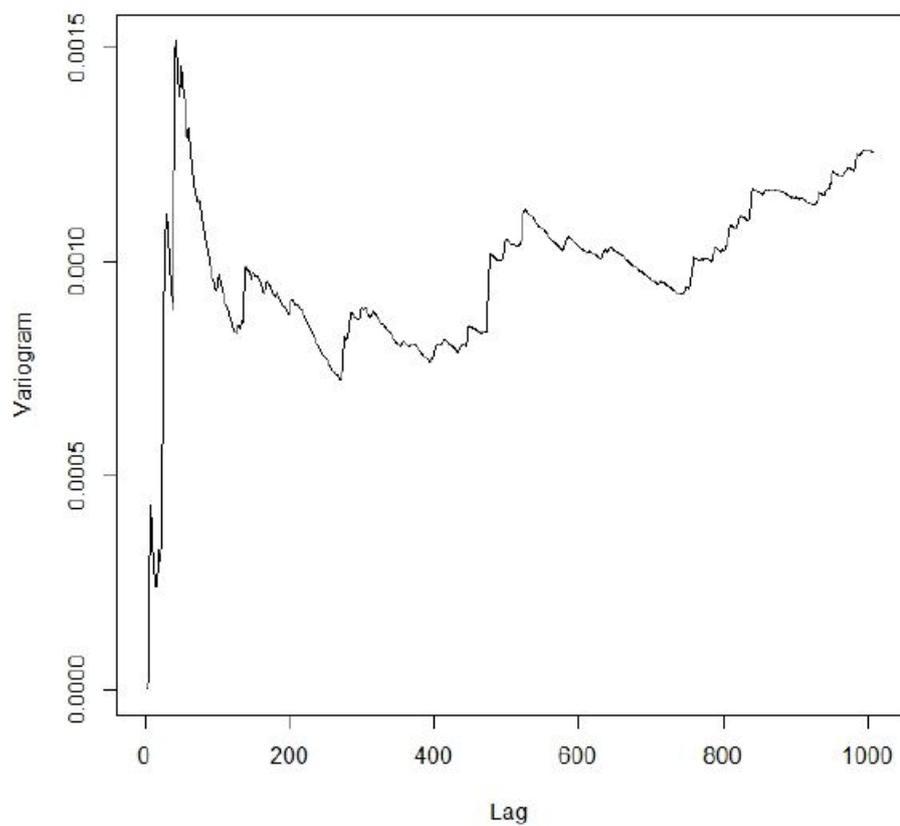


FIGURE 4.14 – Variogramme des résidus

Le variogramme de la différenciation 4.15 nous montre clairement que la série est stationnaire et celui des résidus 4.14 confirme que les données sont normalement distribuées et de moyenne nulle en utilisant le modèle optimal $tvAR(4)$.

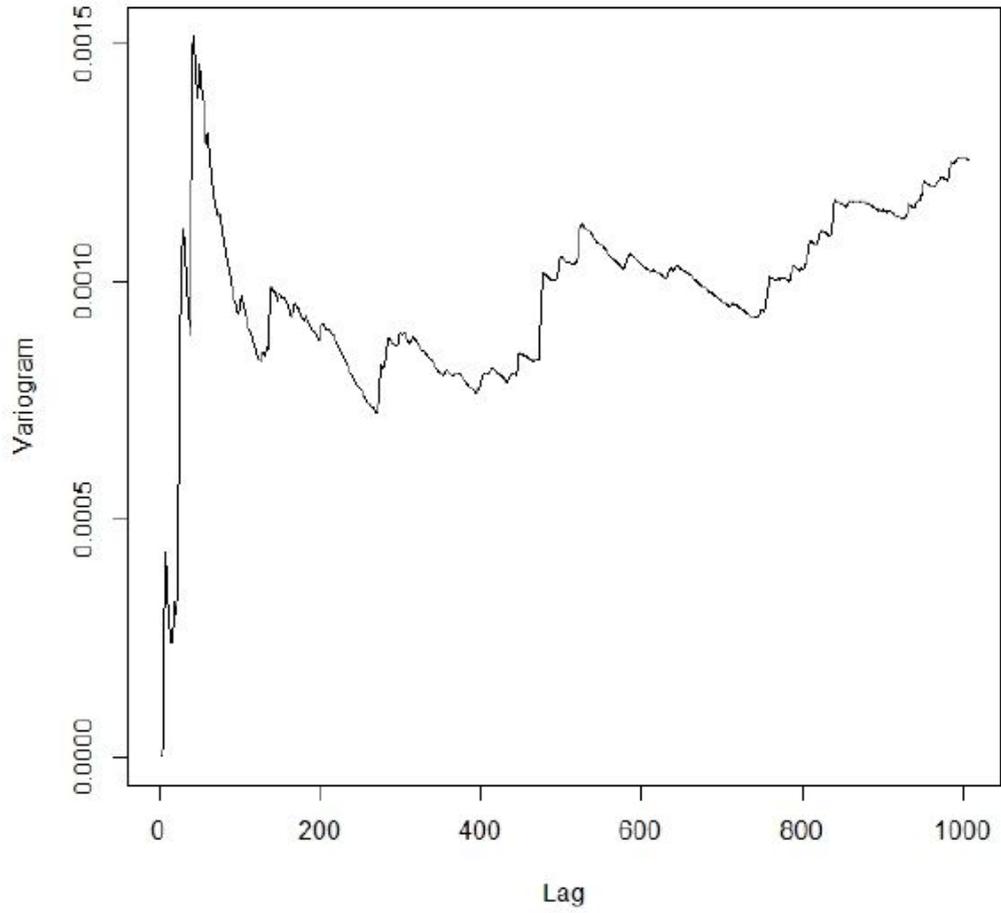


FIGURE 4.15 – Variogramme de la différenciation

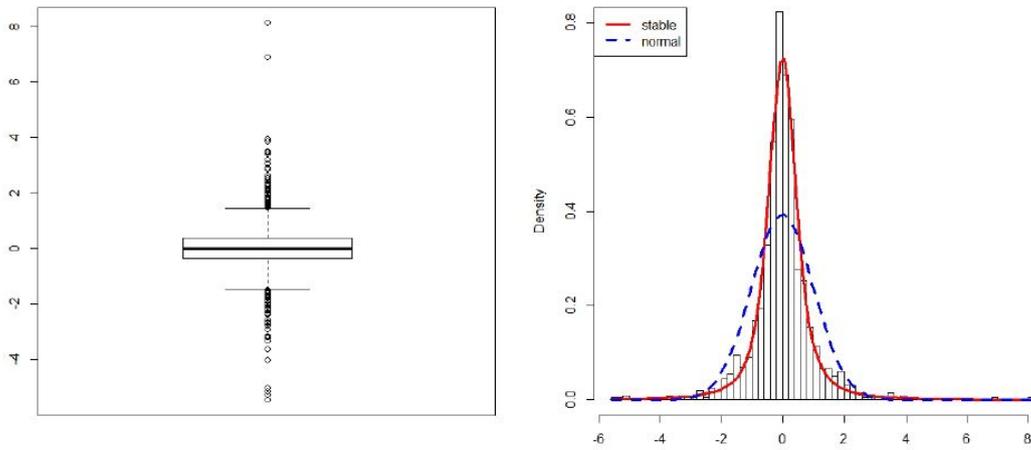


FIGURE 4.16 – Histogramme résiduel et Boite à moustache

Conclusion

L'étude de ce chapitre a été bénéfique pour nous dans les deux sens théorique et pratique . Premièrement ça nous a permis de définir les processus localement stationnaire . Du coté pratique ça nous a permis de définir quelques modèles processus tvARMA et d'interpréter les résultats de la simulation des méthodes d'estimation qu'on avait abordées dans le chapitre précédent. Ensuite de voir la consistance de ces méthodes en augmentant ou en diminuant la taille de l'échantillon. En fin nous avons terminé ce chapitre avec une application des modèles tvARMA appliqués à l'énergie éolienne chose qui nous a servi de voir l'intérêt de chaque étape de la méthode de Box-Jenkins pour la sélection d'un modèle optimal pour la prévision.

Conclusion générale et perspectives

Ce mémoire de master s'est intéressé à la modélisation de séries chronologiques en particulier les processus localement stationnaires à coefficients dépendant du temps.

Dans les premières parties 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 utilisé des lois α -stable comme bruit blanc. Cette dernière 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. Cela fut rendu faisable grâce à l'utilisation judicieuse de la théorie des lois α -stables ainsi que d'autres outils qu'on a fait référence dans ce mémoire.

Notre contribution à ce travail est double. Tout d'abord, nous avons présenté la partie qui concerne l'étude théorique des séries chronologique, le Processus auto-régressif moyenne mobile d'ordre p et q : $ARMA(p, q)$ un exemple populaire de ses séries qui est une généralisation directe des deux exemples introductifs ; la combinaison des processus auto-régressifs d'ordre p et moyennes mobiles d'ordre q .

Nous avons aussi traité les processus localement stationnaires avec des innovations stables et présentons leurs propriétés théoriques. Nous justifions également la raison pour laquelle nous les appelons processus localement stationnaires α -stables. nous proposons l'inférence indirecte et la méthode de whittle afin d'estimer les modèles à coefficient linéaire variable dans le temps $tvARMA$.

Deuxièmement la partie appliquée qui était consacré aux simulation et l'application, la première effectuée par deux méthodes célèbres (l'inférence indirect et la méthode de whittle), nous avons conclu que la méthode d'inférence indirect fonctionne bien.

A la fin nous avons appliqué la méthodologie de Box & Jenkins pour prédire la puissance éolienne totale produite dans les parcs éoliens offshore en Allemagne.

Comme **perspectives**, premièrement certaines limitations doivent encore être résolues à l'avenir :

- ◇ *Des simulations doivent être effectuées pour des modèles plus complexes et envisager également la possibilité de modèles non paramétriques.*
- ◇ *L'inférence indirecte prend du temps, mais elle est appropriée lorsque des innovations à queue lourde sont présentes.*
- ◇ *Les simulations suggèrent que lorsque α est proche de 2, des innovations gaussiennes peuvent être supposées. Le choix du modèle reste une question ouverte. De plus, il existe peu de travaux liés à la prédiction.*
- ◇ *Enfin, nous proposons d'étudier les processus GARCH dans le cas où les coefficients dépendent du temps avec des innovations α -stable en introduisant les notions de la co-variation et de la co-différence pour les lois α -stable aux modèles de GARCH qui sera certainement un très bon axe de recherche dans le futur.*

Bibliographie

- [1] Agnès Iagnoux. *Renforcement statistique des séries chronologique*, 2013.
- [2] Brockwell PJ, Davis RA. *Time series : theory and methods*. New York : Springer-Verlag ; 1991.
- [3] Dahlhaus R. *Maximum likelihood estimation and model selection for locally stationary processes*. *J Nonparametr Stat.* 1996 ;6(2-3) :171-191.
- [4] Dahlhaus R. *Fitting time series models to nonstationary processes*. 1997 ;25(1) :1-37 .
- [5] Dahlhaus R. *Locally stationary processes*. In : Tata Subba Rao SSR, Rao C, editors. *Time series analysis : methods and applications*. (Handbook of Statistics ; vol.30). North Holland. Elsevier ; 2012.p.351-413 .
- [6] Dahlhaus R, Polonik W. *Empirical spectral processes for locally stationary time series*. *Bernoulli*. 2009 ;15(1) :1-39 .
- [7] Dahlhaus R, Giraitis L. *On the optimal segment length for parameter estimates for locally stationary time series*. *J Time Ser Anal.* 1998 ;19(6) :629-655.
- [8] R. Dahlhaus. *Asymptotic statistical inference for nonstationary processes with evolutionary spectra*. In *Athens Conference on Applied Probability and Time Series Analysis, Vol. II (1995)*, volume 115 of *Lecture Notes in Statist.*, pages 145-159. Springer, New York, 1996.
- [9] R. Dahlhaus. *On the Kullback-Leibler information divergence of locally stationary processes*. *Stochastic Process. Appl.*, 62(1) :139-168, 1996.
- [10] Feller, W .(1971) *An Introduction to Probability Theory and Applications Vol.2, Second Edition*, New York :John Wiley & Sons .
- [11] Gallant AR, Tauchen G. *Which moments to match* *Econom Theory*. 1996 ;12(4) :657-681.
- [12] Gouriéroux C, Monfort A, Renault E. *Indirect inference*. 1993 ;8(S1) :S85-S118 .
- [13] Kunsch, H.R. (1995). *A note on causal solutions for locally stationary AR processes*. Preprint, ETH Zurich.
- [14] Ludovic d'estampes . *Traitement statistique des processus α -stables mesures de dépendance et identification des AR stables tests séquentiels tronqués* ,2003.
- [15] Miller K. *Linear difference equations*. New York (NY) : W.A.Benjamin ; 1968 .

- [16] Samorodnitsky G, Taqqu M. *Stable non-Gaussian random processes : stochastic models with infinite variance*. Boca Raton (FL) : Taylor Francis ; 1994.
- [17] Shelton Peiris M, Thavaneswaran A. *On the properties of some nonstationary ARMA processes with infinite variance*. *Int J Modell Simul*. 2001 ;21(4) :301–304.
- [18] Shelton Peiris M, Thavaneswaran A. *Multivariate stable ARMA processes with time dependent coefficients*. *Metrika*.2001 Nov ;54(2) :131–138.
- [19] Shu Wei Chou-Chen & Pedro A. Morettin. *Indirect inference for locally stationary ARMA processes with stable innovations* .
- [20] Shu Wei Chou-Chen. *Locally Stationary Processes with Stable and Tempered* .
- [21] Sylvain Rubenthaler. *Séries chronologiques (avec R)* .
- [22] Van Bellegem S, von Sachs R. *Forecasting economic time series with unconditional time-varying variance*. *Int J Forecast*.2004 ;20(4) :611–627 .
- [23] *Tables de valeurs* Christophe Chesneau <https://chesneau.users.lmno.cnrs.fr/> .