

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie «Houari Boumediene»



Faculté de Mathématiques
Département de Recherche Opérationnelle

MÉMOIRE

Présenté pour l'obtention du diplôme de **Magister**

En : MATHEMATIQUES

Spécialité : **Recherche Opérationnelle** : *Méthodes Stochastiques*

Par

Wissam BENTARZI

Thème

Instabilité dans les Modèles de Séries Chronologiques

Soutenu, publiquement, le 21/07/2010, devant le Jury composé de :

Melle BOUCHEMAKHE Isma	Professeur	à l'U. S. T. H. B.,	Présidente
M. AKNOUCHE Abdelhakim	M. C. A	à l'U. S. T. H. B.,	Directeur de Mémoire
M. MOULAI Mustapha	Professeur	à l'U. S. T. H. B.,	Examineur
M. KHALDI Khaled	M. C. A	à l'U. M. B. B.,	Examineur
M. BOUKHEBOUZE Ahmed	M. A. A	à l'U. S. T. H. B.,	Invité

Table des matières

Introduction	ii
0.1 Présentation du mémoire	iv
1 Bagage mathématique	1
1.1 Introduction	1
1.2 Convergence stochastique de suite de variables aléatoires	4
1.2.1 Inégalités de probabilité	5
1.2.2 Convergence en probabilité	6
1.2.3 Convergence presque sûre	7
1.2.4 Convergence complète	11
1.2.5 Convergence en moyenne quadratique et d'ordre r	11
1.2.6 Convergence en loi	12
1.2.7 Relations entre les modes de convergence	17
1.2.8 Stabilité des modes de convergence	20
1.3 Théorèmes limites : cas d'indépendance	21
1.3.1 Loi des grands nombres	21
1.3.2 Théorème limite central (Principe d'invariance)	25
1.4 Dépendance : généralités sur les processus stochastiques	29
1.4.1 Introduction	29
1.4.2 Notions de base sur les processus aléatoires	30

1.4.3	Processus stationnaires	30
1.4.4	Ergodicité	33
1.4.5	Deux exemples de processus aléatoires : Martingales et Processus de Wiener	34
1.5	Théorèmes limites : cas de dépendance	38
1.5.1	Loi des grands nombres : cas de dépendance	38
1.5.2	Théorème limite central : cas de dépendance	39
1.6	Equations aux récurrences stochastique	42
2	Instabilité et MCO pour AR(1)	46
2.1	Introduction	46
2.2	Regions de stabilités d'un processus AR(1)	47
2.2.1	Existence d'une solution strictement stationnaire	48
2.2.2	Existence d'une solution stationnaire au second ordre	49
2.2.3	Existence d'une solution strictement stationnaire ayant des moments d'ordre supérieurs finis	51
2.2.4	Régions de stabilité et régions d'instabilité	52
2.3	Distribution asymptotique de l'estimateur MCO d'un AR(1)	53
2.3.1	Estimateur des moindres carrés	53
2.3.2	Distribution asymptotique quand le paramètre est à l'intérieur de la région de stabilité ($ \varphi < 1$)	54
2.3.3	Distribution asymptotique quand le paramètre est sur la frontière de la région de stabilité ($\varphi = 1$)	56
2.3.4	Distribution asymptotique quand le paramètre est en dehors de la région de stabilité ($\varphi > 1$)	63
2.4	Simulation	68
2.4.1	Cas où $X_0 = 0$	68

2.4.2	Cas où $X_0 = 1$	76
3	Instabilité et MCO pour $PAR_S(1)$	84
3.1	Introduction	84
3.2	Processus périodiquement corrélé	84
3.2.1	Modèles autorégressifs moyennes mobiles périodiques	87
3.2.2	Conditions de causalité et d'inversibilité	89
3.2.3	Décomposition de Wold-Cramér d'un $PAR_S(1)$	96
3.3	Propriétés asymptotiques des estimateurs des paramètres d'un $PAR_S(1)$. .	98
3.3.1	Estimateurs du MCO des paramètres d'un $PAR_S(1)$	98
3.3.2	Étude des Propriétés Asymptotiques des Estimateurs du MCO d'un $PAR_S(1)$	104
3.4	Simulation	108
4	Instabilité et MCO pour $ARCH(1)$	123
4.1	Introduction	123
4.2	Hétéroscédasticité Conditionnelle	124
4.3	Etude des Modèles ARCH	130
4.3.1	Etude d'un processus ARCH(1)	131
4.4	Stationnarité d'un modèle $ARCH(1)$	137
4.4.1	Stationnarité stricte	137
4.4.2	Stationnarité au second ordre	140
4.5	Propriétés asymptotiques des estimateurs des paramètres d'un $ARCH(1)$. .	140
4.5.1	Estimation des paramètres d'un $ARCH(1)$ par la méthode des MCO en deux étapes ($2S - LS$)	140
4.5.2	Estimation des paramètres d'un $ARCH(1)$ par la méthode des MCO en deux étapes ($2S - WLS$)	148

4.6 Simulation 154

Alà barakati Allah

Introduction

La modélisation des séries chronologiques occupe une place prépondérante dans l'analyse probabiliste et statistique des phénomènes évoluant dans le temps. L'objectif en est la représentation d'un phénomène évolutif au moyen d'un processus stochastique, modèle mathématique dont la loi du mouvement tente de mimer le mieux possible les caractéristiques empiriques du phénomène sous-jacent. La mise en œuvre d'un processus adéquat pour représenter une série chronologique est fondée sur une méthodologie qui peut se résumer en quatre étapes essentielles, étapes qui ne sont pas forcément effectuées de manière séquentielle et ordonnée.

La première consiste en le postulat d'une équation aux différences (ou différentielle dans le cas d'observation en temps continu) stochastique dépendant d'un paramètre et dont la solution lorsqu'elle existe représente le processus d'intérêt. Une telle équation est caractérisée par : i) un processus entrée dit d'innovation qui représente le moteur de la dynamique stochastique du processus d'intérêt et qui vérifie certaines hypothèses de structure, ii) une forme spécifique qui peut être linéaire ou non par rapport aux termes du processus d'intérêt et qui dépend des caractéristiques des séries que l'on souhaite représenter, iii) un support d'évolution par rapport auquel est définie l'équation et qui peut être l'ensemble des nombres entiers naturels, celui des entiers relatifs ou un ensemble fini, et enfin iv) un vecteur de paramètres qui représente la marge de manœuvre, l'étendue des possibilités d'ajustement du processus solution aux données réelles et le but crucial sera d'en estimer la valeur en se basant sur ces données.

Une fois l'équation spécifiée, on étudie dans une deuxième étape l'existence et unicité d'un processus solution de l'équation postulée ainsi que diverses de ses propriétés probabilistes et ce en fonction des valeurs du paramètres. Parmi les propriétés probabilistes auxquelles on s'intéresse figure : la stationnarité stricte, l'ergodicité, l'ergodicité géométrique, la stationna-

rité au second ordre, l'existence des moments d'ordre supérieur, le comportement des queues de la distribution marginale... Pour l'établissement de chacune de ces propriétés, il existe un domaine spécifique des valeurs du paramètre et le but de cette étape est d'identifier les différentes régions. Pour le cas d'équations de formes linéaires soumises à certaines hypothèses sur le processus d'innovation, les domaines sont confondus mais il n'en est pas de même dans le cas de formes non linéaires.

La troisième étape consiste d'une part en le développement des méthodes d'estimation du paramètre et des tests d'hypothèses le concernant (méthode du maximum de vraisemblance, méthode des moindres carrés, test du rapport de vraisemblance, test de racine unitaire,...) et d'autre part en l'étude des propriétés statistiques des procédures d'inférence (convergence, normalité asymptotique, efficacité asymptotique,...), étude qui repose sur les propriétés probabilistes étudiées dans l'étape précédente.

Enfin, dans la dernière étape, le modèle étant mis en œuvre, peut être exploité pour résoudre des problématiques posées au préalable telles que la prédiction, le filtrage, le contrôle, la gestion de risque...

Dans ce long processus de mise en œuvre d'un modèle sur la base d'une série finie de données, on a toujours besoin de valeurs initiales pour démarrer l'équation stochastique puisqu'on se base sur des données finies. Les hypothèses que l'on formule à propos de ces valeurs peuvent affecter les propriétés probabilistes du processus et par là même les propriétés statistiques des procédures d'inférence associées. L'équation est alors dite stable pour une propriété probabiliste donnée si, à la limite, dans le long terme, le processus atteint cette propriété quelles que soient les valeurs de démarrage. Dans le cas contraire, l'équation est dite instable pour cette propriété. Il est important de noter que la stabilité est relative à la propriété probabiliste à laquelle on s'intéresse dans le sens qu'une équation peut être stable pour une propriété et instable pour une autre. De plus, la stabilité, pour une propriété donnée, dépend des valeurs des paramètres de l'équation et il en ressort des régions (domaines) de stabilité et d'instabilité.

Bien que la stabilité se manifeste de façon unique sur tout son domaine et ses effets sur les propriétés probabilistes et statistiques du processus sont les mêmes d'une sous-région à une autre, il n'en est pas de même pour l'instabilité qui peut en revanche se manifester de

plusieurs façons différentes et dont les effets peuvent être différents d'une sous-région du domaine d'instabilité à une autre¹. En particulier, pour certaines équations, la position du paramètre sur la frontière ou à l'intérieur de la région d'instabilité influence considérablement les propriétés probabilistes du processus solution ainsi que les propriétés asymptotiques de certains estimateurs des paramètres de l'équation.

Dans ce mémoire, nous nous intéressons à l'étude de l'instabilité, de ses différentes régions et de son influence sur les propriétés probabilistes et statistiques de trois classes de modèles (équations) de séries chronologiques. La première classe d'équations connues sous le nom "autorégressif moyenne mobile" (*ARMA*) est celle dont la forme est linéaire par rapport au processus d'intérêt et au processus d'innovation et dont les paramètres sont invariants dans le temps. La seconde, dite *ARMA* périodique, est encore de forme linéaire mais dont les paramètres évoluent en fonction du temps de façon périodique. Enfin la troisième classe d'équations est de forme non linéaire en le processus d'intérêt, mais qui peut se mettre sous forme linéaire par rapport au carré du processus d'intérêt (modèles *ARCH*). Pour chacune de ces équations, nous étudions les régions de stabilité et d'instabilité par rapport à chacune des propriétés probabilistes suivantes : stationnarité stricte, ergodicité, stationnarité au second ordre et stationnarité aux moments d'ordre supérieurs. Nous en étudions également l'influence de la position du paramètre dans chacune des régions d'instabilité sur les propriétés asymptotiques des estimateurs des moindres carrés ordinaire et généralisé ainsi que celles de l'estimateur du quasi-maximum de vraisemblance.

0.1 Présentation du mémoire

Ce mémoire de Magistère en Mathématiques, Spécialité Recherche Opérationnelle :
Option Méthodes Stochastiques, dont l'intitulé est

"Instabilité dans les Modèles de Séries Chronologiques"

porte, essentiellement, sur l'étude de l'instabilité de quelques modèles de séries chronologiques (linéaires classiques et périodiques et non linéaires), l'influence des régions sur les comportements des lois asymptotiques des estimateurs de la méthode des moindres carrés

¹"Les familles heureuses se ressemblent toutes ; les familles malheureuses sont malheureuses chacune à leur façon". Leon Tolstoi, *Anna Karenine*.

ordinaire et généralisée. Il est composé de six chapitres dont nous passons en revue leurs continus.

Chapitre 1 : Bagage mathématique

Une des propriétés importantes d'une suite d'estimateurs d'un certain paramètre scalaire (ou vectoriel), en analyse des séries chronologiques ou plus généralement en statistique, est bien la propriété de consistance. Brièvement, ce concept intéressant assure la croissance de la *fiabilité* de l'estimateur ou encore la *proximité* de cet estimateur et le paramètre lui est associé, selon un certain sens, avec la croissance de la taille de la série temporelle sous étude. Formellement, pour mesurer cette proximité on est amené à définir une certaine norme avec laquelle on mesure la fluctuation stochastique entre l'estimateur et son contrepartie. La consistance s'interprète, dans ce sens, par la convergence de cette norme vers 0. Il est alors intéressant de d'inférer sur le comportement d'une suite d'estimateurs et/ou d'une suite de statistiques de test, $\{T_n, n \geq 1\}$, quand $n \rightarrow \infty$.

Ce premier chapitre composé, en outre de la section d'introduction, de six sections, est consacré, essentiellement, à de brèves présentations portant, d'une part, sur des notions fondamentales des modes de convergence des suites de variables aléatoires, de ces propriétés, de ces diverses propriétés de stabilité et de ces implications et d'autre part, sur les concepts de base de la théorie des lois de grands nombres classique à savoir la loi faible des grands nombres et la loi forte des grands nombres ainsi que les principaux théorèmes centraux limites.

Dans la deuxième section, nous passons en revue les modes de convergence usuels, à savoir la convergence en probabilité, la convergence presque sûre, la convergence en moyenne d'ordre r ($r > 0$) et enfin le mode de convergence en loi. Par la suite, nous étudions les diverses propriétés, les implications éventuelles de ces modes de convergence. Nous terminons cette section d'une part par la présentation de quelques opérations sur les suites de variables aléatoires convergentes selon quelques un de ces modes de convergence et d'autre part par traiter la question de convergence stochastique de certaines fonctions (mesurable de Borel satisfaisant quelque conditions) d'une suite de variables aléatoires convergente selon un certain mode.

La loi faible des grands nombres et la loi forte des grands nombres dans le cas de dépendance

sont exposées dans la quatrième section. La cinquième section est consacrée à la présentation des connaissances de base concernant le concept du théorème central limite. En effet, nous traitons quelques versions de ce fameux théorème dans le cas de suites de variables aléatoires indépendantes et dépendantes.

Il est bien connu que la modélisation et l'étude probabiliste et statistique d'une série chronologique, qui n'est autre qu'une réalisation (ou une trajectoire) particulière, parmi une infinité de réalisations (trajectoires) possibles, d'un certain processus aléatoire $\{X_t; t \in \mathbb{T}\}$, consiste en l'identification du processus stochastique générateur de cette série. De ce fait, il nous semble intéressant de terminer le présent chapitre de ce mémoire, par la présentation des concepts de base de la théorie des processus aléatoires et d'évoquer quelques résultats existant dans la littérature concernant la stationnarité (stricte et faible), l'ergodicité (stricte et faible) et les lois des grands nombres de quelques processus à temps discret rencontrés dans notre mémoire. Le processus de Wiener, qui est un processus à temps continu, est aussi évoqué. Quelques théorèmes fondamentaux portant sur ses lois des grands nombres sont aussi présentés.

Chapitre 2 : Propriétés des Estimateurs des MCO d'un AR(1)

Le comportement probabiliste d'un processus stochastique $\{X_t, t \in \mathbb{Z}\}$, solution d'une équation stochastique paramétrique donnée, autrement dit satisfaisant à un modèle stochastique, dépend essentiellement de l'appartenance des paramètres de l'équation à des régions bien spécifiées induites par ces paramètres pour satisfaire à une ou plusieurs conditions imposées par un certain concept.

Par conséquent, les comportements de différentes statistiques d'intérêt, liées à l'étude théorique et pratique du modèle, en particulier des estimateurs et des statistiques de tests, différent d'une de ces régions à autre, montrant l'importance et l'utilité de spécifier, sans ambiguïté, ces différentes régions. Cependant, le concept de stabilité stochastique est essentiellement lié à la forme de l'équation aux différences d'intérêt mais aussi aux propriétés probabilistes que l'on désire mettre en évidence. En général pour les modèles linéaires le domaine de stabilité est invariant des propriétés probabilistes à considérer. A l'encontre, nous verrons qu'il n'en est pas de même pour les modèles non linéaires et que pour chaque

propriété probabiliste donnée, il en existe un domaine de stabilité propre.

Brièvement, dans ce deuxième chapitre du mémoire, nous étudions les régions de stabilité, qui sont de nombre trois régions, d'un processus autorégressif d'ordre 1 ainsi que les propriétés distributionnelles asymptotiques de l'estimateur des moindres carrés selon les trois régions de stabilité. Plus précisément, nous établissons en premier lieu l'estimateur des moindres carrés (*MCO*) et l'estimateur des moindres carrés pondérés (généralisé) (*MCG*) du paramètre φ d'un *AR*(1) et en deuxième étape nous obtenons les lois asymptotiques de ces estimateurs basés sur un échantillon de taille n , $\widehat{\varphi}_n$, selon l'appartenance du paramètre à une des trois régions suivantes : $|\varphi| < 1$, (dite région stable), $|\varphi| = 1$, (dite région de racine unitaire) et enfin $|\varphi| > 1$, (dite région explosive).

Nous terminons le chapitre par une étude de simulation intensive à travers laquelle nous étudions les comportements des lois asymptotiques des suites $\{\widehat{\varphi}_n, n \in \mathbb{N}^*\}$ des estimateurs selon les trois régions de stabilité. Les résultats de simulations correspondant à des valeurs différentes de φ , selon la taille n , ont bien renforcé les résultats théoriques étudiés.

Chapitre 3 : Propriétés des Estimateurs des MCO d'un *PAR_S*(1)

Le concept de processus périodiquement corrélé a été introduit, dans la littérature de l'étude des processus aléatoires, par *Gladychew* (1961). Diverses sont les classes de processus non stationnaires, parmi lesquelles nous citons la classe de processus dont les fonctions d'autocovariance sont des fonctions périodiques dans le temps. En effet, les processus périodiquement corrélés sont rencontrés dans plusieurs domaines, entre autre, l'étude environnementale et économique, en particulier financière.

La classe des modèles de séries chronologiques linéaires ou non linéaires à coefficients périodiques, dans le temps, a montré sa capacité et sa performance pour capturer et décrire ce type de phénomènes qui ne peuvent être décrits par les modèles saisonniers de *Box* et *Jenkins* (1970, 1976). En effet, ces modèles ont reçu, pendant les trois dernières décennies, un développement important que ce soit du point de vue théorique que pratique ; et ils continuent de recevoir, dans nos jours, un intérêt considérable de la part des analystes théoriciens et praticiens.

Dans ce chapitre composé de trois sections, nous nous intéressons, essentiellement, d'une part, à l'obtention des estimateurs des moindres carrés ordinaire et pondéré, *MCO*, des paramètres d'un modèle autorégressif d'ordre un périodique, de période quelconque S et, d'autre part, à l'étude de l'influence de chacune des trois régions déterminées par les paramètres du modèle (à savoir la région de stabilité, la région de racine unitaire et la région explosive), sur le comportement asymptotique de la loi de ces estimateurs. Cette étude est une extension de l'étude faite pour un modèle autorégressif classique, d'ordre un, au cas du modèle périodique. Dans la section suivante, nous passons en revue les notions fondamentales concernant les processus périodiquement corrélés et les modèles autorégressifs moyenne mobiles, à coefficients périodiques de période S , $ARMA_S(p_t, q_t)$. Nous abordons en particulier l'existence de la décomposition de *Wold-Cramér* ainsi que les propriétés de causalité et d'inversibilité de ces modèles, tout en suivant les deux approches existant dans la littérature de l'analyse des séries temporelles à coefficients périodiques.

La troisième section, de ce chapitre, est consacrée à l'obtention des estimateurs, $\widehat{\Phi}_n$, du moindres carrés ordinaire et pondérée des paramètres inconnus, Φ , d'un modèle autorégressif d'ordre un, périodique de période S . Par la suite, nous nous intéressons au problème intéressant qui consiste en l'étude de l'influence des trois régions, qui peuvent être induites par ces paramètres, à savoir la région de stabilité, région de racine unitaire et région explosive données respectivement par $\left| \prod_{s=1}^S \varphi_s \right| < 1$, $\left| \prod_{s=1}^S \varphi_s \right| = 1$ et $\left| \prod_{s=1}^S \varphi_s \right| > 1$, sur le comportement de la loi asymptotique de ces estimateurs. En effet, nous obtenons les lois asymptotiques $\sqrt{n} \left(\widehat{\Phi} - \Phi \right)$ dans les trois précédentes régions.

Chapitre 4 : Propriétés des Estimateurs des MCO d'un ARCH(1)

Parmi les faits stylisés des séries temporelles rencontrées dans le domaine environnemental, économique et financier (entre autre comme l'excès de *Kurtosis* (*Excess of Kurtosis*), l'asymétrie (*Skewness*), la persistance dans la structure dynamique, la multimodalité de la densité du processus d'innovation) nous nous intéressons particulièrement dans ce chapitre au regroupement de volatilité (*Volatility Clustering*) instantanée. L'évidence empirique de la volatilité instantanée qui caractérise plusieurs séries temporelles en particulier les financières

remonte aux années soixante. En effet, Mandelbort 1963 a bien mis en évidence ce fait par la description :

"Large changes tend to be followed by large changes of either sign—and small changes tend to be followed by small changes".

Les formulations classiques (de type *ARMA*) sont avérées inadéquates car centrées sur la structure d'autocovariance des processus. Pour remédier aux insuffisances des modèles classiques, une nouvelle classe de séries chronologiques a été introduite dans la littérature des séries chronologiques. Il s'agit de la classe des modèles Autorégressifs Conditionnellement Hétéroscédastiques (*Autoregressive Conditional Heteroscedasticity : ARCH*), introduit pour la première fois dans la littérature de l'analyse des séries chronologiques par Engle en 1982 et ses diverses extensions.

Pour l'importance de la classe des modèles autorégressifs conditionnellement hétéroscédastiques et ses extensions, nous passons en revue les caractéristiques essentielles de plusieurs phénomènes, en particulier la volatilité, qui ne peuvent pas être capturées et décrites par les modèles *ARMA*. Par la suite nous étudions plusieurs propriétés importantes concernant des processus *ARCH* (q) en particulier le processus *ARCH* (1). Les estimateurs des deux paramètres d'un modèle *ARCH* (1) sont obtenus par la méthode des moindres carrés ordinaire (*MCO*) en deux étapes. Les propriétés des lois asymptotiques sont établies selon le type de la région de stabilité. Nous terminons le chapitre par une étude de simulation. Nous terminons le chapitre par une étude de simulation intensive à travers laquelle nous étudions les comportements des lois asymptotiques des suites des estimateurs obtenus, selon les trois régions de stabilité. Les résultats de simulations correspondant à des valeurs différentes des deux paramètres du modèle, selon la taille n , ont bien renforcé les résultats théoriques étudiés.

Chapitre 1

Bagage mathématique

1.1 Introduction

L'objectif essentiel de la statistique inférence paramétrique est, d'une part, la recherche d'estimateurs scalaires ou vectoriels d'un certain nombre de paramètres réels scalaires ou vectoriels inconnus $\underline{\phi}$ et/ou de statistiques de test pour tester une hypothèse et ; d'autre part, l'étude des diverses propriétés probabilistes et/ou statistiques (exactes ou asymptotiques) de ces estimateurs fondés sur un échantillon aléatoire de taille n . Ces estimateurs et statistiques de test, ou plus généralement, des règles de décision statistique, selon la terminologie de la théorie de décision, ne sont que des fonctions mesurables de l'échantillon aléatoire, $\{T_n, n \geq 1\}$, et ne dépendent d'aucun paramètre inconnu.

Une des propriétés exigées d'une suite de statistiques $\{T_n, n \geq 1\}$ est la propriété fondamentale dite la propriété de *consistance*. Ce concept à une suite de statistiques, assure la croissance de la *fiabilité* ou encore la *proximité*, selon un certain sens, avec la croissance de la taille de l'échantillon n . Ainsi, on peut, brièvement, dire qu'un estimateur consistant devient de plus en plus proche, dans un certain sens, de son contrepartie théorique (de la population statistique) quand le volume de l'échantillon augmente.

Formellement, pour mesurer cette proximité on est amené à définir une certaine norme avec laquelle on mesure la fluctuation stochastique entre l'estimateur et l'estimé. La consistance s'interprète, dans ce sens, par la convergence de cette norme vers 0. Il est alors intéressant d'inférer sur le comportement d'une suite de statistiques ou plus généralement d'une suite de fonctions stochastiques $\{T_n, n \geq 1\}$, quand $n \rightarrow \infty$.

Le présent chapitre est consacré, essentiellement, à de brèves présentations portant, d'une part, sur des notions fondamentales des modes de convergence des suites de variables aléatoires, de ses propriétés et de ses implications et d'autre part, sur les concepts de base des lois de grands nombres à savoir la loi faible des grands nombres, et la loi forte des grands nombres ainsi que les principaux théorèmes centraux limites. Ce chapitre comporte deux sections, dans la première nous passons en revue les modes de convergence usuels, à savoir la convergence en probabilité, la convergence presque sûre, la convergence en moyenne d'ordre r ($r > 0$) et enfin le mode de convergence en loi. Par la suite, nous étudions les diverses propriétés, les implications éventuelles de ces modes de convergence. Nous terminons cette partie d'une part par la présentation de quelques opérations sur les suites de variables aléatoires convergentes selon quelques uns de ces modes de convergence et d'autre part par traiter la question de convergence stochastique de certaines fonctions (mesurable de Borel satisfaisant quelque conditions) d'une suite de variables aléatoires convergente selon un certain mode.

La deuxième section traite, le concept des lois de grands nombres à savoir la loi faible des grands nombres et la loi forte des grands nombres, les principaux théorèmes centraux limites classiques et le théorème central limite fonctionnel qui est primordial en analyse des séries chronologiques. L'étude du théorème limite classique qui est très importante aussi bien du point de vue théorique que pratique, diffère du point de vue de complexité selon les conditions exigées par la suite de variables aléatoires sous-jacente; autrement dit, les

énoncées de ces théorèmes diffèrent selon les hypothèses vérifiées par la suite de variables aléatoires. De ce fait, on se limite aux énoncées qui correspondent aux hypothèses imposées ultérieurement dans le présent mémoire.

1.2 Convergence stochastique de suite de variables aléatoires

Cette section est consacrée à la notion de la convergence stochastique de suites de variables aléatoires. Brièvement, les modes de convergence stochastique sont des approximations des variables aléatoires, selon certain sens à préciser ultérieurement. Ces modes de convergence, leurs implications et leurs diverses propriétés jouent un rôle prédominant en analyse des séries chronologiques, comme nous verrons lors de l'étude de la théorie des grands échantillons et plus généralement dans l'étude des comportements asymptotiques des statistiques d'intérêt (estimateurs, statistiques de tests, ...) fondées sur des séries chronologiques de grandes tailles. Notre objectif est loin d'être l'étude probabiliste profonde et détaillée de la notion de la convergence stochastique, mais seulement de présenter brièvement les notions fondamentales et essentielles concernant ces modes de convergence. Nous ne manquerons pas de souligner que la littérature concernant ce sujet est abondante. Toutefois, nous ne manquerons pas de citer les fameux ouvrages de Billingsley (1968) et Serfling (1980) dans lesquels ces notions sont étudiées avec rigueur et détail.

Dans la suite de cette section, nous passons en revue les modes de convergence usuels, à savoir la convergence en probabilité, la convergence presque sûre, la convergence en moyenne d'ordre r ($r > 0$) et enfin le mode de convergence en loi. Par la suite, nous étudions les diverses propriétés, les implications éventuelles de ces modes de convergence. Nous terminons cette partie d'une part par la présentation de quelques opérations sur les suites de variables aléatoires convergentes selon quelques un de ces modes de convergence et d'autre part par le traitement de la stabilité de ces modes de convergence, autrement, la question de convergence stochastique de certaines fonctions (mesurable de Borel satisfaisant quelque conditions) d'une suite de variables aléatoires convergente selon un certain mode, en particulier, selon la convergence en probabilité et presque sûre que nous utilisons fréquemment dans ce travail.

1.2.1 Inégalités de probabilité

Les inégalités de probabilités sont, d'une part utiles en elles mêmes pour s'informer sur les comportements des variables aléatoires et, d'autre part, elles sont des outils importants pour démontrer plusieurs lois faibles et lois fortes des grands nombres. En effet, ces inégalités de probabilité trouvent une diversité d'applications importantes lors de l'étude des implications dans les modes de convergence et les lois faibles et les lois fortes des grands nombres que nous exposerons ultérieurement. L'inégalité (de la probabilité) suivante, dite de *Chebyshev*, est le précurseur de plusieurs d'autres formes d'inégalités (de probabilité) plus générales.

Théorème 1.2.1. (*Inégalité de Chebyshev*)

Soit X une variable aléatoire, non négative de moyenne finie $\mathbb{E}(X) = \mu$. Alors, pour quelque soit $\varepsilon (> 0)$, on a l'inégalité de probabilité suivante dite de *Chebyshev* :

$$P(X > \mu \varepsilon) \leq \frac{1}{\varepsilon}.$$

A partir de cette inégalité, on peut, facilement, déduire une autre forme de *Techybeshev* correspondant au cas d'une suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ dont les moyennes et les variances $\mu_n = \mathbb{E}(X_n)$ et $\sigma_n^2 = \mathbb{E}((X_n - \mu_n)^2)$ existent et ils sont finis.

Théorème 1.2.2. (*Inégalité de Chebyshev*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, dont $\mu_n = \mathbb{E}(X_n)$ et $\sigma_n^2 = \mathbb{E}((X_n - \mu_n)^2)$ existent et ils sont finis. Alors, pour quelque soit $\varepsilon (> 0)$, on a l'inégalité de probabilité suivante dite aussi de *Chebyshev* :

$$P(|X_n - \mu_n| > \sigma_n \varepsilon) \leq \frac{1}{\varepsilon^2}.$$

Cette inégalité assure qu'une variable aléatoire $X_n, n \in \mathbb{N}^*$, sous ces conditions qui sont fréquemment vérifiées, que l'écart absolu entre la variable aléatoire X_n et son espérance mathématique μ_n dépasse la valeur (positive) $\sigma_n \varepsilon$ avec une probabilité au plus égale à la borne majorant ε^{-2} . Il est utile de noter que l'information apportée par cette inégalité ne dépend de la distribution de $X_n, n \in \mathbb{N}^*$, que par le biais de sa moyenne, μ_n , et sa variance σ_n^2 .

Une autre inégalité qui peut être déduite de l'inégalité de Chebyshev donnée par le *Théorème (1.2.1)* et qui peut donner des intervalles d'étendues plus courtes est l'inégalité suivante dite de Markov.

Théorème 1.2.3. (*Inégalité de Markov*)

Soit X une variable aléatoire, non négative de moment d'un certain ordre $r (> 0)$ fini, noté $\mu_r = \mathbb{E}(X^r)$. Alors, pour quelque soit $\varepsilon (> 0)$, on a l'inégalité de probabilité suivante dite de Markov :

$$P(X > \varepsilon) \leq \frac{\mu_r}{\varepsilon^r}.$$

Théorème 1.2.4. (*Kolmogorov Maximal Inequality*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires indépendantes dont les moyennes $\mu_n = \mathbb{E}(X_n)$ et les variances $\sigma_n^2 = \mathbb{E}((X_n - \mu_n)^2)$, existent pour $\forall n \in \mathbb{N}^*$. Et considérons la somme partielle $T_k = \sum_{i=1}^k (X_i - \mu_i)$, $k = 1, 2, \dots, n$. Alors, pour quelque soit $t (> 0)$, on a l'inégalité de probabilité suivante dite Kolmogorov Maximal Inequality :

$$P\left(\max_{1 \leq k \leq n} |T_k| > t\right) \leq t^{-2} \text{Var}(T_n) = t^{-2} \sum_{i=1}^n \sigma_i^2$$

1.2.2 Convergence en probabilité

Un des modes de convergence de suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ est la convergence en probabilité dont la définition est la suivante.

Définition 1.2.1

On dit que la suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ (définies sur un même espace de probabilité (Ω, \mathcal{F}, P)) converge, en probabilité, quand $n \rightarrow \infty$, vers la variable aléatoire réelle (peut être dégénérée) X (définie sur ce même espace de probabilité), et on note

$$X_n \xrightarrow{P} X, \text{ ou encore}$$

$$X_n - X \xrightarrow{P} 0,$$

si

$$\forall \varepsilon, \eta > 0, \exists n_0 = n_0(\varepsilon, \eta) \in \mathbb{N} : n > n_0, \quad P(|X_n - X| > \varepsilon) < \eta.$$

Cette définition peut être réécrite sous la forme équivalente suivante :

$$\forall \varepsilon, \eta > 0, \exists n_0 = n_0(\varepsilon, \eta) \in \mathbb{N} : n > n_0, P(|X_n - X| > \varepsilon) > 1 - \eta.$$

Brièvement, la convergence en probabilité de la suite de variables aléatoires X_n vers la variable aléatoire X , ou encore $X_n - X$ converge vers zéro, quand $n \rightarrow \infty$, signifie que pour quelques soient les constantes réelles strictement positives fixées à priori ε et h , on peut trouver un entier positif n_0 (dépendant évidemment de ε et η) tel que l'événement aléatoire

$$\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\},$$

se réalise avec une probabilité strictement supérieure à $1 - \eta$, et ce pour tout n fixé à partir du rang déterminé n_0 .

Dans le paragraphe suivant, nous présentons un autre type de convergence, de suite de variables aléatoires, plus forte et plus stricte que la convergence en probabilité à savoir la convergence presque sûre.

1.2.3 Convergence presque sûre

Brièvement, nous disons qu'une variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ convergence presque sûrement, converge vers une variable aléatoire X (définie sur ce même espace de probabilité), si pour $\forall \varepsilon, \eta > 0$ on peut trouver un certain rang n_0 tel que tous les événements aléatoires $\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}$, correspondant à tout n supérieure à ce rang se réalisent mutuellement (tout ensemble). Il est clair que la condition de cette dernière convergence est plus exigeante et plus stricte que celle prise en compte par la convergence en probabilité. Cette dernière condition est tenue en considération par la convergence presque sûre dont la définition est la suivante.

Soient la suite de variables aléatoires réelles (non nécessairement indépendantes) $\{X_n, n \geq 1\}$ et la variable aléatoire réelle X , définies sur un même espace de probabilité (Ω, \mathcal{F}, P) . Pour tout réel positif donné ε , considérons la suite des événement associée $\{A_n(\varepsilon), n \geq 1\}$ dont le terme général est donné par

$$A_n(\varepsilon) = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\},$$

Notant $A(\varepsilon)$ l'événement consistant en la réalisation d'un nombre infini des événements $A_n(\varepsilon)$ Nous rappelons que cet événement est, par définition, la limite supérieure de la suite

$A_n(\varepsilon)$:

$$A(\varepsilon) = \limsup A_n(\varepsilon) = \bigcap_{n \geq 1} \bigcap_{m \geq n} A_n(\varepsilon).$$

Une définition formelle de la convergence presque sûre est donnée par la suivante.

Définition 1.2.2

On dit que la suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ (définies sur un même espace de probabilité (Ω, \mathcal{F}, P)) converge, presque sûre, quand $n \rightarrow \infty$, vers la variable aléatoire réelle (peut être dégénérée) X (définie sur ce même espace de probabilité), et on note

$$X_n - X \xrightarrow{p.s.} 0, \quad \text{ou encore } X_n \xrightarrow{p.s.} X,$$

si pour $\forall \varepsilon > 0$, on a

$$P(A(\varepsilon)) = 0. \tag{1.2.1}$$

La convergence presque sûre peut être donnée par une autre définition équivalente à celle présentée. En effet, la limite supérieure de la suite d'événements $\{A_n(\varepsilon), n \geq 1\}$, peut s'exprimer sous la forme suivante

$$A(\varepsilon) = \limsup A_n(\varepsilon) = \bigcap_{n \geq 1} \bigcup_{m \geq n} A_n(\varepsilon) = \bigcap_{n \geq 1} B_n(\varepsilon),$$

où la suite d'événements $\{B_n(\varepsilon), n \geq 1\}$ est donnée par son terme général

$$\begin{aligned} B_n(\varepsilon) &= \sup_{m \geq n} A_m(\varepsilon) = \bigcup_{m \geq n} A_m(\varepsilon), \\ &= \{\omega \in \Omega : \sup_{m \geq n} (|X_m(\omega) - X(\omega)|) > \varepsilon\}. \end{aligned}$$

Il est clair que cette dernière suite d'événements est décroissante, alors, selon les propriétés d'une suite décroissante d'ensembles, on a

$$A(\varepsilon) = \limsup A_n(\varepsilon) = \lim_{n \rightarrow \infty} B_n(\varepsilon).$$

Ce qui permet d'écrire la probabilité de l'événement, consistant en la réalisation d'une infinité d'événements, $A_n(\varepsilon)$; c'est-à-dire $P(A(\varepsilon))$ sous la forme suivante

$$P(A(\varepsilon)) = P\left(\lim_{n \rightarrow \infty} B_n(\varepsilon)\right),$$

ce qui peut s'écrire, tenant compte de la continuité de la probabilité, sous la forme :

$$\begin{aligned} P(A(\varepsilon)) &= \lim_{n \rightarrow \infty} P\left(B_n(\varepsilon)\right), \\ &= \lim_{n \rightarrow \infty} P\left(\sup_{m \geq n} (|X_m(\omega) - X(\omega)|) > \varepsilon\right). \end{aligned}$$

En vertu de cette formulation, nous pouvons donner la définition de la convergence presque sûre équivalente suivante.

Définition 1.2.3

La suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ (définies sur un même espace de probabilité (Ω, \mathcal{F}, P)) est dite converge, presque sûre, lorsque $n \rightarrow \infty$, vers la variable aléatoire réelle (peut être dégénérée) X (définie sur ce même espace de probabilité), et on note

$$X_n - X \xrightarrow{p.s} 0,$$

ou encore $X_n \xrightarrow{p.s} X$, si on a :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P \left(\sup_{m \geq n} (|X_m(\omega) - X(\omega)|) > \varepsilon \right) = 0. \quad (1.2.2)$$

Il est utile de noter que la convergence presque sûre est un mode de convergence stochastique plus stricte que la convergence de probabilité. En effet,

$$\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\} \subseteq \left\{ \omega \in \Omega : \sup_{m \geq n} (|X_m(\omega) - X(\omega)|) > \varepsilon \right\}$$

ce qui implique

$$P(|X_n(\omega) - X(\omega)| > \varepsilon) \geq P \left(\sup_{m \geq n} (|X_m(\omega) - X(\omega)|) > \varepsilon \right),$$

ceci montre que la convergence presque sûre entraîne la convergence en probabilité. Il est facile de montrer que la condition nécessaire et suffisante (1.2.2) de convergence presque sûrement de la suite X_n vers X , quand $n \rightarrow \infty$, est équivalente aux deux conditions équivalentes données par la définition équivalente suivante.

Définition 1.2.4

Une suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ (définies sur un même espace de probabilité (Ω, \mathcal{F}, P)) est dite converge, presque sûre, lorsque $n \rightarrow \infty$, vers la variable aléatoire réelle (peut être dégénérée) X (définie sur ce même espace de probabilité) si on a :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P \left(\bigcup_{m \geq n} (|X_m - X|) > \varepsilon \right) = 0,$$

ou encore de façon équivalente

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P \left(\bigcap_{m \geq n} (|X_m - X|) \geq \varepsilon \right) = 1.$$

La propriété suivante établit une condition suffisante pour la convergence presque sûrement.

Propriété 1.2.1

Si la série

$$\sum_{n \geq 1} P(|X_n(\omega) - X(\omega)| > \varepsilon),$$

est convergente pour un ε positif quelconque, alors

$$X_n - X \xrightarrow{p.s.} 0.$$

La propriété suivante donne une condition suffisante pour la convergence *p.s.*

Propriété 1.2.2

S'il existe un ordre r tel que la suite numérique $E(|\bar{X}_n - \pi|^r)$ existe et de plus la série

$$\sum_{n \geq 1} E(|\bar{X}_n - \pi|^r),$$

est convergente, alors, on a :

$$X_n - X \xrightarrow{p.s.} 0.$$

On ne manquera pas de remarquer qu'à partir la propriété 1.2.1 et de la propriété 1.2.2 on peut aisément voir encore la différence entre la convergence en probabilité et la convergence presque sûre de point de vue l'exigence requise pour chacune d'elles. la propriété 1.2.2 montre que la convergence de la suite numérique $\{|X_n - X|^r, n \geq 1\}$ (pour un certain $r, r > 0$) vers 0, lorsque $n \rightarrow \infty$, assure la convergence en probabilité de la suite de variables aléatoires $\{X_n - X, n \geq 1\}$, vers 0, quand $n \rightarrow \infty$. Cependant, cette dernière condition ne suffit pas, comme le montre la propriété 1.2.2, pour établir la convergence presque sûre.

En effet, pour établir cette convergence forte il faut que le moment absolu d'ordre r de $X_n - X, E(|\bar{X}_n - \pi|^r)$, soit le terme général d'une série convergente. Ceci est une condition suffisante pour la convergence presque sûre et elle est évidemment une condition plus forte que celle qu'exige la convergence en probabilité.

1.2.4 Convergence complète

Définition 1.2.5

Une suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ (définies sur un même espace de probabilité (Ω, \mathcal{F}, P)) telle que la somme

$$\sum_n P(|X_n(\omega) - X(\omega)| > \varepsilon),$$

est convergente, $\forall \varepsilon > 0$, où X est une variable aléatoire (peut être dégénérée en un point a) définie sur ce même espace de probabilité est dite converger complètement, lorsque $n \rightarrow \infty$, vers X , et on note

$$X_n - X \xrightarrow{c} 0.$$

1.2.5 Convergence en moyenne quadratique et d'ordre r

Un troisième mode de convergence stochastique important et qui joue un rôle prédominant dans les critères de la convergence en probabilité est la convergence en moyenne d'ordre r , en particulier la convergence en moyenne quadratique correspondant à $r = 2$.

Définition 1.2.6

Une suite de variables aléatoires réelles (ou vectorielles) $\{X_n, n \in \mathbb{N}^*\}$ (définies sur un même espace de probabilité (Ω, \mathcal{F}, P)) est dite converge, en moyenne d'ordre r ($r \in \mathbb{N}^*$), lorsque $n \rightarrow \infty$, vers la variable aléatoire réelle (peut être dégénérée) X (définie sur ce même espace de probabilité), et on note

$$X_n - X \xrightarrow{m,r} 0,$$

ou encore

$$X_n \xrightarrow{m,r} X,$$

si

$$\forall \varepsilon > 0, \exists n_0 = n_0(\varepsilon) \in \mathbb{N} : n > n_0 \longrightarrow \mathbb{E}(\|X_n - X\|^r) < \varepsilon, \quad (1.2.3)$$

ou encore, d'une façon équivalente

$$\mathbb{E}(\|X_n - X\|^r) \longrightarrow 0, \text{ quand } n \rightarrow \infty,$$

où $\|\cdot\|$ désigne la norme Euclidien et où $\mathbb{E}(\cdot)$ est l'opérateur d'espérance mathématique.

1.2.6 Convergence en loi

Fonctions de répartition

Rappelons que la loi de probabilité qui régit le comportement d'une variable aléatoire scalaire X (resp. vecteur aléatoire \underline{X} de dimension n) peut être exprimée de différentes manières. Une manière générale et universelle pour décrire et caractériser une variable aléatoire est le concept fondamental de la fonction de répartition $F_X(\cdot)$ (resp. vecteur aléatoire $F_{\underline{X}}(\cdot, \dots, \cdot)$). Brièvement, cette fonction est une application de la droite réelle \mathbb{R} (resp. \mathbb{R}^n) dans l'intervalle des nombres réels $[0, 1]$ qui associe, pour chaque nombre réel $x \in \mathbb{R}$, (resp. $\underline{x} \in \mathbb{R}^n$) la valeur $F_X(x)$ (resp. $F_{\underline{X}}(\underline{x})$) qui est la probabilité de l'événement aléatoire $\{\omega : X(\omega) \leq x\}$ ($\{\omega : \underline{X}(\omega) \leq \underline{x}\}$). Autrement dit, la fonction de répartition conjointe $F_{\underline{X}}(\cdot)$ est définie comme suit.

Définition 1.2.7

La fonction de répartition, notée $F_{\underline{X}}(\underline{x})$, d'un vecteur aléatoire \underline{X} , définie sur \mathbb{R}^n et à valeur dans $[0, 1]$, est donnée par :

$$F_{\underline{X}}(\underline{x}) = P(\{\omega : \underline{X}(\omega) \leq \underline{x}\}) = P_{\underline{X}}(\underline{X} \leq \underline{x}), \quad \forall \underline{x} \in \mathbb{R}^n.$$

Le théorème ci-dessous établit les conditions ou les propriétés que doit satisfaire une fonction de répartition $F_X(\cdot)$ d'une variable aléatoire scalaire X .

Théorème 1.2.5

La fonction de répartition d'une variable aléatoire X , satisfait les propriétés suivantes :

- 1) $F_X(\cdot)$ est non décroissante,
- 2) $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$,
- 3) $F_X(\cdot)$ est continue à droite en tout point $x \in \mathbb{R}$.

On ne manquera pas de noter que les trois propriétés précédentes caractérisent une fonction de répartition, c'est-à-dire chaque application de \mathbb{R} dans $[0, 1]$ satisfaisant ces trois propriétés est une fonction de répartition d'une certaine variable aléatoire X .

La propriété suivante établit les propriétés qui caractérisent de la fonction de répartition conjointe $F_{\underline{X}}(\cdot)$ d'une variable aléatoire vectorielle \underline{X} .

Théorème 1.2.6 (*Propriétés de la fonction de répartition conjointe $F_{\underline{X}}(\cdot)$*)

La fonction de répartition conjointe $F_{\underline{X}}(\underline{x})$ satisfait les propriétés suivantes :

- 1) $\forall \underline{x} \in \mathbb{R}^n$, on a $0 \leq F_{\underline{X}}(\underline{x}) \leq 1$,
- 2) $\forall \underline{x}, \underline{y} \in \mathbb{R}^n$, tels que $\underline{x} \leq \underline{y}$ on a $F(\underline{x}) \leq F(\underline{y})$,
- 3) $\lim_{\substack{\underline{x} \rightarrow -\infty \\ \text{pour au moins un } i}} F_{\underline{X}}(x_1, x_2, \dots, x_n) = 0$,
- $\lim_{\underline{x} \rightarrow \infty} F_{\underline{X}}(x_1, x_2, \dots, x_n) = 1$,
- 4) $F(\underline{x})$ est continue à droite,
- 5) $\Delta_{I_1} \Delta_{I_2} \dots \Delta_{I_n} F_{\underline{X}}(\underline{x}) \geq 0$.

où l'opérateur Δ_{I_k} agissant sur $F_{\underline{X}}(\underline{x})$, est donné, pour un intervalle de \mathbb{R} de la forme $I_k =]a_k, b_k]$, comme suit :

$$\begin{aligned} \Delta_{I_k} F_{\underline{X}}(\underline{x}) &= F_{\underline{X}}(x_1, \dots, x_{k-1}, b_k, x_{k+1}, \dots, x_n) \\ &\quad - F_{\underline{X}}(x_1, \dots, x_{k-1}, a_k, x_{k+1}, \dots, x_n) \end{aligned}$$

plus généralement, pour chacun des intervalles $I_k =]a_k, b_k]$, $k = 1, 2, \dots, n$, l'opérateur $\Delta_{I_1} \Delta_{I_2} \dots \Delta_{I_n}$ est défini par :

$$\Delta_{I_1} \Delta_{I_2} \dots \Delta_{I_n} F_{\underline{X}}(\underline{x}) = \Delta_{I_1} \left(\Delta_{I_2} \left(\dots \left(\Delta_{I_{n-1}} \left(\Delta_{I_n} F_{\underline{X}}(\underline{x}) \right) \right) \dots \right) \right).$$

Remarque

On remarque que les quatre premières conditions de la propriété précédente sont des généralisations naturelles de leurs analogues concernant une fonction de répartition d'une variable aléatoire scalaire, données par le *Théorème 1.2.1*. Par contre la cinquième condition est indispensable seulement pour la fonction de répartition conjointe d'un vecteur aléatoire. En effet, on peut avoir des fonctions conjointes qui satisfont les quatre premières propriétés sont qu'elles soient des fonctions de répartition conjointe d'une variable aléatoire scalaire.

Convergence en loi

Un autre mode de convergence très répandu et qui n'implique pas directement la suite de variables aléatoires $\{X_n, n \geq 1\}$ mais plutôt leur suite de fonctions de répartition (distributions) $\{F_n(\cdot), n \geq 1\}$ est connu sous le nom convergence en distribution.

Définition 1.2.8

On dit qu'une suite de fonctions de répartition (distributions) $\{F_n(\cdot), n \geq 1\}$ converge faiblement vers une fonction de répartition limite $F(\cdot)$ d'ensemble de points de continuité \mathcal{C} , si, on a :

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \quad \forall x \in \mathcal{C} \subseteq \mathbb{R},$$

i.e., pour quelque soit $\varepsilon (> 0)$ donné, il existe un entier positif $n_0 = n_0(\varepsilon)$, tel que pour quelque soit x point de continuité de F , on a :

$$|F_n(x) - F(x)| < \varepsilon, \quad \forall n > n_0(\varepsilon),$$

et on note

$$F_n(\cdot) \xrightarrow{w} F(\cdot).$$

Soient une suite de variables aléatoires, définie sur un même espace de probabilité, $\{X_n, n \geq 1\}$, dont la suite de fonctions de répartition est $\{F_n(\cdot), n \geq 1\}$ et une variable aléatoire X , définie sur la même espace de probabilité, dont la fonction de répartition est $F(\cdot)$.

Définition 1.2.9 (Convergence en distribution)

On dit la suite de v.a. $\{X_n, n \geq 1\}$ converge en loi (ou en distribution) vers la variable aléatoire (peut être dégénérée) X (n'est pas nécessairement définie sur le même espace de probabilité), si

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \quad \forall x \in \mathcal{C} \subseteq \mathbb{R},$$

et on note

$$X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X \text{ ou encore } X_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} X.$$

La définition de la convergence en loi peut être généralisée au cas de suite de vecteurs aléatoires. Le théorème suivant établit une condition nécessaire et suffisante de convergence d'une vecteurs vecteur aléatoires en termes de la convergence d'une suite variables aléatoires scalaires.

Théorème 1.2.7 (*Cramér-Wold devise*)

Soit $(\underline{X}_n, n \in \mathbb{N}^*)$ une suite de vecteurs aléatoires à valeurs dans \mathbb{R}^m , définis sur un même espace (Ω, \mathcal{F}, P) . Alors

$$\underline{X}_n \xrightarrow{\mathcal{L}} \underline{X} \quad \text{ssi} \quad \lambda' \underline{X}_n \xrightarrow{\mathcal{L}} \lambda' \underline{X}, \quad \forall \lambda = (\lambda_1, \dots, \lambda_m)' \in \mathbb{R}^m.$$

Quelques conséquences de la convergence en loi

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires à valeurs dans \mathbb{R} , définies sur un même espace (Ω, \mathcal{F}, P) . Alors, on a les résultats suivants.

Théorème 1.2.8

Si la suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ converge en loi, vers la variable aléatoire X alors, on a : $\mathbb{E}(|X|) = \liminf_{n \rightarrow \infty} \mathbb{E}(|X_n|)$.

Théorème 1.2.9

Si la suite de variables aléatoires uniformément intégrables $\{X_n, n \in \mathbb{N}^*\}$ converge en loi, vers la variable aléatoire X , alors, on a :

$$\mathbb{E}(|X|) < \infty \text{ et } \mathbb{E}(X_n) \longrightarrow \mathbb{E}(X).$$

Théorème 1.2.10

La suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ converge en loi, vers la variable aléatoire X , si, et seulement si, si et seulement si

$$\lim_{n \rightarrow \infty} \mathbb{E}(\phi(X_n)) = \mathbb{E}(\phi(X)),$$

pour quelque soit la fonction mesurable absolument continue et bornée $\phi(\cdot)$.

Fonction caractéristique d'une variable aléatoire

On rappelle que la fonction caractéristique $\varphi_X(t)$ d'une variable aléatoire scalaire X est donnée, par définition, par l'espérance mathématique du variable aléatoire complexe $e^{i t X}$, i.e.,

$$\varphi_X(t) = \mathbb{E}(e^{i t X}), \quad \forall t \in \mathbb{R} \text{ où } i^2 = -1.$$

La notion de la fonction caractéristique d'une variable aléatoire scalaire X peut facilement être étendue au cas d'une variable aléatoire vectoriel $\underline{X} = (X_1, X_2, \dots, X_n)'$.

En effet, la fonction caractéristique $\varphi_{\underline{X}}(t_1, t_2, \dots, t_n)$ est donnée par la définition suivante.

Définition 1.2.10

La fonction caractéristique, qui existe toujours, notée $\varphi_{\underline{X}}(t_1, t_2, \dots, t_n)$ ou de façon abrégée $\varphi_{\underline{X}}(\underline{t})$, d'une variable aléatoire à n dimensions \underline{X} , est donnée par :

$$\varphi_{\underline{X}}(\underline{t}) = \mathbb{E}(e^{it'X}) = \int e^{t'x} dF_{\underline{X}}(x), \quad \underline{t} = (t_1, t_2, \dots, t_n) \in \mathbb{R}^n.$$

Le théorème suivant établit une condition nécessaire et suffisante d'indépendance stochastique des variables aléatoires X_1, X_2, \dots, X_n , en termes de leur fonction caractéristique conjointe et leurs fonctions caractéristiques marginales.

Unicité de la fonction caractéristique et théorème d'inversion

On a vu que la fonction caractéristique d'une variable aléatoire caractérise et spécifie complètement la loi de probabilité de cette variable aléatoire. Et par conséquent, la fonction caractéristique peut être utilisée pour reconnaître les lois des variables aléatoires et/ou pour déduire leurs moments, s'ils existent. De plus, il est montré, grâce au théorème d'inversion de la fonction caractéristique, qu'à partir de cette fonction caractéristique on peut trouver mathématiquement la loi de probabilité représentée par la fonction répartition.

Théorème 1.2.11 (Unicité de la fonction caractéristique)

Soient $\underline{X} = (X_1, X_2, \dots, X_n)$ et $\underline{Y} = (Y_1, Y_2, \dots, Y_n)$ deux variables aléatoires multidimensionnelles de fonctions de répartition conjointes $F_{\underline{X}}(\cdot)$ et $F_{\underline{Y}}(\cdot)$, respectivement, dont les fonctions caractéristiques sont, respectivement, $\varphi_{\underline{X}}(t_1, t_2, \dots, t_n)$ et $\varphi_{\underline{Y}}(t_1, t_2, \dots, t_n)$. Alors $F_{\underline{X}}(\cdot)$ et $F_{\underline{Y}}(\cdot)$ sont équivalentes i. e., les deux vecteurs aléatoires \underline{X} et \underline{Y} sont équivalents, si, et seulement si, les fonctions caractéristiques $\varphi_{\underline{X}}(t_1, t_2, \dots, t_n)$ et $\varphi_{\underline{Y}}(t_1, t_2, \dots, t_n)$ leurs correspondantes sont équivalentes.

Théorème 1.2.12. (Théorème d'inversion)

Si X est une variable aléatoire de fonction caractéristique $\varphi_X(t)$. Alors, la fonction de répartition $F_X(\cdot)$ de X est donnée, pour quelques soient les points de continuité de $F_X(\cdot)$, a et b ($a < b$), par :

$$F_X(b) - F_X(a) = \frac{1}{2\pi} \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt.$$

On ne manquera pas de noter que l'énoncé du théorème d'inversion qu'on vient de donner, est adéquat seulement aux cas d'une variable aléatoire discrète ou absolument continue.

Théorème 1.2.13. (*Levy-Cramér*)

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires réelles de fonctions caractéristiques correspondantes $\{\varphi_n(t), n \geq 1\}$ et une variable aléatoire X de fonction caractéristique $\varphi(t)$, continue en $t = 0$. Alors, une condition nécessaire et suffisante pour que $X_n \xrightarrow{\mathcal{L}} X$, est que, pour tout réel t , $\varphi_n(t)$ tend simplement, lorsque $n \rightarrow \infty$, vers la fonction caractéristique $\varphi_X(t)$, de la variable aléatoire X .

Théorème 1.2.14. (*Théorème de continuité de Lévy*)

Soit $\{F_n(\cdot), n \in \mathbb{N}^*\}$ une suite de fonctions de répartition (distributions) faiblement convergente vers une fonction $F(\cdot)$ non négative, bornée, non décroissante et continue à droite, alors, si

$$\varphi_n(t) = \int_{\mathbb{R}} e^{itx} dF_n \longrightarrow \varphi(t),$$

et où $\varphi(t)$ est continue en $t = 0$, alors $F(\cdot)$ est une fonction de répartition et $\varphi(\cdot)$ est la fonction caractéristique lui est associée.

1.2.7 Relations entre les modes de convergence

Après avoir passé en revue les modes de convergence stochastique les plus usuels et les plus fréquemment rencontrés en théorie de probabilité ainsi qu'en statistique mathématique, nous nous intéressons dans cette section à l'étude de l'implication mutuelle de ces divers modes de convergence. Un des intérêts particuliers de l'étude des relations existantes entre les divers modes de convergence est de fournir des moyens et des outils mathématiques pour établir la convergence selon un mode de convergence précis. En effet, il y a des situations où la preuve directe d'une certaine convergence est trop compliquée, par contre elle peut être sensiblement simple en passant par la convergence selon un autre mode de convergence qui implique la convergence en question.

Propriété 1.2.3

La convergence d'une suite de variables aléatoires, $\{X_n, n \in \mathbb{N}^*\}$, pour un certain ordre r ($r > 0$), vers X (peut être dégénérant) implique la convergence en probabilité vers X . Autrement dit, s'il existe un ordre r ($r > 0$) tel que

$$X_n \xrightarrow{m.r} X, \text{ alors, on a } X_n \xrightarrow{P} X.$$

De cette propriété on conclut que la convergence en moyenne d'un certain ordre r ($r > 0$) est une condition suffisante pour la convergence en probabilité. En revanche, l'inverse n'est pas, généralement, correct. En effet, une suite de variables aléatoires réelles $\{X_n, n \in \mathbb{N}^*\}$ peut converger en probabilité, lorsque $n \rightarrow \infty$, vers une variable aléatoire (peut être dégénérée en un point) X sans qu'il n'existe un ordre r pour lequel cette suite converge en moyenne de cet ordre. De plus, la suite peut même converger en probabilité sans l'existence de n'importe quel moment absolu de la suite $X_n - X$. Ce dernier cas peut être éclairci par l'exemple classique, d'une suite de variables aléatoires de Cauchy, suivant. En effet, la suite de Cauchy converge en probabilité vers la constante 0, quand $n \rightarrow \infty$, malgré que tous les moments de cette suite n'existent pas. En effet, on peut facilement montrer que la suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ données par la suite de densités de Cauchy

$$f_n(x) = \frac{n}{\pi} \frac{1}{1 + n^2 x^2}, \quad x \in \mathbb{R},$$

est telle que $X_n \xrightarrow{P} 0$ et que X_n ne possède aucun moment.

Néanmoins, il existe des situations où on possède une condition suffisante sous laquelle la convergence en probabilité entraîne la convergence en moyenne d'ordre r ($r > 0$) quelconque, comme le montre le théorème suivant.

Théorème 1.2.15

La convergence en probabilité d'une suite de variables aléatoires presque sûrement bornées $\{X_n, n \in \mathbb{N}^\}$, vers une variable aléatoire, presque sûrement bornée X , implique la convergence en moyenne d'ordre quelconque r ($r > 0$), vers X .*

La réciproque n'est pas vraie à l'exception du cas où X est une v. a. certaine. Le théorème suivant établit les conditions sur lesquelles la convergence en probabilité, d'une suite de variables aléatoires non nécessairement presque sûrement bornées $\{X_n, n \in \mathbb{N}^*\}$, implique la convergence en moyenne d'un certain ordre r ($r \in \mathbb{N}^*$).

Théorème 1.2.16

Soient $\{X_n, n \in \mathbb{N}^\}$ une suite de variables aléatoires et la variable aléatoire X , définie sur le même espace de probabilité. Alors, pour toute fonction continue croissante et bornée $\varphi(\cdot)$ avec $\varphi(0) = 0$, on a :*

$$X_n - X \xrightarrow{p} 0, \text{ si, et seulement si } \varphi(|X_n - X|) \xrightarrow{m,r} 0, \text{ pour un certain ordre } r.$$

Théorème 1.2.17

Si la suite de variables aléatoires (scalaires ou vectoriels) $\{X_n, n \in \mathbb{N}^*\}$, converge en loi, vers une variable aléatoire dégénérant en un point c (constante), alors cette suite converge en probabilité, vers c .

Théorème 1.2.18

La convergence presque sûre implique la convergence en probabilité.

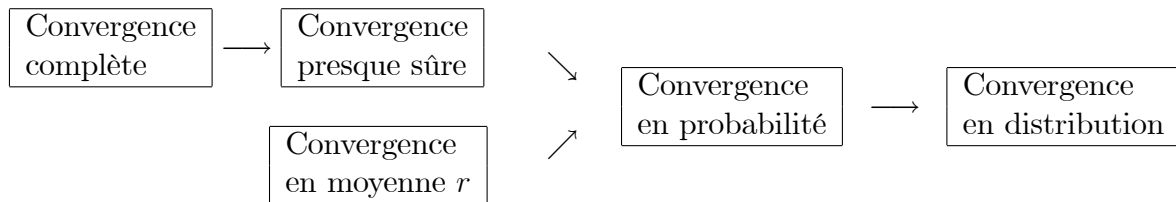
$$X_n \xrightarrow{p.s.} X, \implies X_n \xrightarrow{P} X.$$

Théorème 1.2.19 (Slutsky)

Soient $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{P} c$ (une constante finie). Alors, on a :

- i) $X_n + Y_n \xrightarrow{\mathcal{L}} cX$,
- ii) $X_n Y_n \xrightarrow{\mathcal{L}} X + c, \quad X_n + Y_n \xrightarrow{\mathcal{L}} X + c$,
- iii) $X_n / Y_n \xrightarrow{\mathcal{L}} X / c$ pour $c \neq 0$.

Le schéma suivant fait apparaître les implications des différents modes de convergences étudiés.



1.2.8 Stabilité des modes de convergence

Connaissions qu'une suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ converge, lorsque $n \rightarrow \infty$, vers une variable aléatoire (peut être dégénérée en un point) X selon un certain mode de convergence, il sera donc intéressant et utile de savoir si pour une certaine fonction mesurable (satisfaisant éventuellement quelques conditions) $g(\cdot)$ la suite de variables aléatoires, transformée de la suite originale par $g(\cdot)$, c'est-à-dire la suite de variables aléatoires $\{g(X_n), n \geq 1\}$ converge selon le même type de convergence vers la variable aléatoire (peut être dégénérée) $g(X)$. La réponse à cette préoccupation est affirmative pour quelques modes et pour des fonctions continues $g(\cdot)$. En effet, les théorèmes suivants montrent que les modes de convergences en probabilité, presque sûre et en loi possèdent cette propriété importante de stabilité relativement aux fonctions continues.

Théorème 1.2.20. (*Théorème de Slutsky*)

Soient une suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ et une variable aléatoire (peut être dégénérée) X définies sur un même espace de probabilité. Alors, si

$$X_n \xrightarrow{P} X, \text{ on a, pour quelque soit la fonction continue } g(\cdot)$$

$$g(X_n) \xrightarrow{P} g(X),$$

Théorème 1.2.21. (*Théorème de Slutsky*)

Soient une suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ et une variable aléatoire (peut être dégénérée) X définies sur un même espace de probabilité. Alors, si

$$X_n \xrightarrow{p.s.} X, \text{ on a, pour quelque soit la fonction continue } g(\cdot),$$

$$g(X_n) \xrightarrow{p.s.} g(X).$$

Théorème 1.2.22

Soient une suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ et une variable aléatoire (peut être dégénérée) X (peuvent ne pas être définies sur un même espace de probabilité). Alors, si

$$X_n \xrightarrow{\mathcal{L}} X, \text{ on a, pour quelque soit la fonction continue } g(\cdot),$$

$$g(X_n) \xrightarrow{\mathcal{L}} g(X).$$

1.3 Théorèmes limites : cas d'indépendance

1.3.1 Loi des grands nombres

Introduction

En analyse des séries chronologiques ou plus généralement en statistique inférence, on est souvent confronté au problème de l'étude des propriétés probabilistes de quelques statistiques qui s'expriment comme moyennes empiriques, \bar{X}_n , et/ou comme fonctions (mesurables) de ces moyennes empiriques, $g(\bar{X}_n)$, de variables aléatoires formant un échantillon aléatoire, de taille n , prélevé dans une population statistique donnée par un modèle statistique paramétrique de moyenne μ . Dans le cas où une telle expression n'existe pas ou est trop complexe, il existe des situations où on peut, par le biais d'un développement adéquat, de la statistique sous-jacente sous forme d'une somme de la moyenne empirique, \bar{X}_n , et d'une certaine autre variable aléatoire convergente, selon un certain mode, vers 0.

Dans la présente section, nous commençons par rappeler quelques concepts, définitions et résultats concernant les lois faibles et les lois fortes des grands nombres dans le cas d'indépendance et de même loi; ainsi que dans le cas d'indépendance sans l'hypothèse de même loi.

Lois faibles et lois fortes des grands nombres : cas d'indépendance

Dans la présente section, nous passons en revue les résultats fondamentaux existants concernant la convergence stochastique en probabilité et presque sûrement de la moyenne empirique \bar{X}_n ainsi que de quelques fonctions particulières d'intérêt de cette statistique, quand la taille n augmente indéfiniment. Ces deux types de convergence stochastique, à savoir la convergence en probabilité et presque sûrement, de \bar{X}_n quand $n \rightarrow \infty$, qui sont dits lois faibles et lois fortes des grands nombres, respectivement, seront présentés dans le cas d'indépendance et de même loi (*i.i.d*) et dans le cas d'indépendance seulement qui est un cas moins restreignant. L'étude de la loi faible des grands nombres (*WLLN* : *Weak Law of Large Numbers*) et la loi forte des grands nombres (*SLLN* : *Strong Law of Large Numbers*) dans le cas de dépendance sera l'objet de la section suivante.

Lois faibles des grands nombres de variables aléatoires indépendantes

Dans ce paragraphe, nous présentons la convergence en probabilité, quand $n \rightarrow \infty$, de la suite des moyennes empiriques $\{\bar{X}_n, n \in \mathbb{N}^*\}$, qui n'est que la suite des moyennes empiriques partielles $\bar{X}_n = \sum_{i=1}^n X_i/n$ de la suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$, dans le cas où les variables aléatoires sont indépendantes et identiquement distribuées (*i.i.d*) et dans le cas où elles sont indépendantes et non nécessairement de même loi de probabilité.

A1) Lois faibles des grands nombres cas *i.i.d*

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires indépendantes de même loi de probabilité, alors le théorème suivant établit la convergence en probabilité, quand $n \rightarrow \infty$, vers la moyenne théorique, $\mu = \mathbb{E}(X_1)$, commune de ces variables aléatoires. Dans ce cas la convergence en probabilité de la suite $\{\bar{X}_n, n \in \mathbb{N}^*\}$, vers la constante μ , est donnée par le théorème suivant.

Théorème 1.3.1. (*Khintchine WLLN*)

Si $\{X_n, n \in \mathbb{N}^\}$ est une suite de variables aléatoires indépendantes identiquement distribuées d'espérance commune $\mu = \mathbb{E}(X_1)$, alors la suite des moyennes empiriques $\{\bar{X}_n, n \in \mathbb{N}^*\}$ converge en probabilité, vers la constante μ , quand $n \rightarrow \infty$, i.e.,*

$$\bar{X}_n \xrightarrow{p} \mu.$$

A2) Lois faibles des grands nombres cas d'indépendance

Le théorème suivant établit la loi faible des grands nombres dans le cas où les variables aléatoires de la suite $\{X_n, n \in \mathbb{N}^*\}$ sont indépendantes mais non nécessairement de même loi de probabilité. Cette version de la loi faible des grands nombres qui n'exige pas l'indépendance des variables aléatoires de la suite adjacente mais autres conditions fréquemment vérifiées dans plusieurs modèles statistiques, trouve des applications importantes en analyse des séries chronologiques où le processus générateur d'une série temporelle est évidemment corrélé.

Théorème 1.3.2. (*Markov WLLN*)

Si $\{X_n, n \in \mathbb{N}^*\}$ est une suite de variables aléatoires indépendantes, non nécessairement de même loi de probabilité, dont les espérances, $\mu_n = \mathbb{E}(X_n)$, existent et les moments centrés absolus $\mathbb{E}\left(|X_n - \mu_n|^{1+\delta}\right)$ existent pour une certaine valeur de $\delta \in (0, 1]$. En supposant de plus que

$$n^{-(1+\delta)} \sum_{i=1}^n |X_n - \mu_n|^{1+\delta} = \varphi_n(\delta) \longrightarrow 0, \text{ quand } n \rightarrow \infty,$$

alors, on a la convergence en probabilité de la suite $\{\bar{X}_n, n \in \mathbb{N}^*\}$ vers la constante $\mathbb{E}(\bar{X}_n)$, quand $n \rightarrow \infty$, i.e., $\bar{X}_n \xrightarrow{p} \mathbb{E}(\bar{X}_n)$.

Loi forte des grands nombres de variables aléatoires indépendantes

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires indépendantes de même loi de probabilité, alors le théorème suivant établit la convergence en probabilité, quand $n \rightarrow \infty$, vers la moyenne théorique, $\mu = \mathbb{E}(X_1)$, commune de ces variables aléatoires. Dans ce cas la convergence presque sûrement de la suite $\{\bar{X}_n, n \in \mathbb{N}^*\}$, vers la constante μ , quand $n \rightarrow \infty$, est donnée par le théorème de Khintchine suivant.

B1) Loi forte des grands nombres cas i.i.d**Théorème 1.3.3.** (*Khintchine SLLN*)

Si $\{X_n, n \in \mathbb{N}^*\}$ est une suite de variables aléatoires indépendantes identiquement distribuées d'espérance commune $\mu = \mathbb{E}(X_1)$, alors la suite des moyennes empiriques $\{\bar{X}_n, n \in \mathbb{N}^*\}$ converge presque sûrement, vers la constante μ , quand $n \rightarrow \infty$, i.e., $\bar{X}_n \xrightarrow{p.s} \mu$.

Ce théorème de la loi forte des grands nombres dû à Khintchine énoncé pour une suite de variables aléatoires scalaires peut être étendu au cas d'une suite de variables aléatoires vectorielles, de dimensions m . Soit $\{Y_n, n \in \mathbb{N}^*\}$ où $Y_n = (Y_{1,n}, Y_{2,n} \dots Y_{m,n})'$, une suite de variables aléatoires vectorielles, indépendantes identiquement distribuées, dont le vecteur moyen $\underline{\mu} = \mathbb{E}(Y_1) = (\mathbb{E}(Y_{1,n}), \mathbb{E}(Y_{2,n}) \dots \mathbb{E}(Y_{m,n}))'$, alors, en appliquant le théorème Khintchine (1.3.3) sur les composantes $Y_{i,n}$, $i = 1, 2, \dots, m$, on peut énoncer le corollaire suivant.

Corollaire 1.3.1. (*Khintchine SLLN*)

Si $\{Y_n, n \in \mathbb{N}^*\}$ est une suite de variables aléatoires vectorielles, indépendantes identiquement distribuées, de vecteur moyen commun $\underline{\mu} = \mathbb{E}(Y_1)$, alors la suite des vecteurs des moyennes empiriques $\{\bar{Y}_n, n \in \mathbb{N}^*\}$ converge presque sûrement, vers le vecteur constant $\underline{\mu}$, quand $n \rightarrow \infty$, i.e., $\bar{Y}_n \xrightarrow{p.s} \underline{\mu}$.

Plus généralement, soit $g(\cdot)$ une application mesurable, de \mathbb{R}^m dans \mathbb{R} , continue au point $\underline{\mu}$, alors, on a le résultat suivant qui trouve une variété d'applications importantes.

Théorème 1.3.4

Si $\{Y_n, n \in \mathbb{N}^*\}$ est une suite de variables aléatoires vectorielles, indépendantes identiquement distribuées, de vecteur moyen commun $\underline{\mu} = \mathbb{E}(Y_1)$, alors la suite des variables aléatoires scalaires $\{g(Y_n), n \in \mathbb{N}^*\}$ converge presque sûrement, vers la constante scalaire $g(\underline{\mu})$, quand $n \rightarrow \infty$, i.e., $g(\bar{Y}_n) \xrightarrow{p.s} g(\underline{\mu})$.

B2) Loi forte des grands nombres cas d'indépendance

Les résultats précédents sont énoncés sous l'hypothèse d'indépendance et de même loi de la suite de variables aléatoires adjacente. Dans ce qui suit, on expose brièvement, les résultats fondamentaux existants concernant la loi forte des grands nombres des suites de variables aléatoires indépendantes mais non nécessairement de même loi de probabilité.

Théorème 1.3.5

Si $\{X_n, n \in \mathbb{N}^*\}$ est une suite de variables aléatoires indépendantes dont les espérances sont nulles ($\mathbb{E}(X_n) = 0, \forall n \in \mathbb{N}^*$) et de variances σ_n^2 telles que $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$, alors la suite des sommes partielles $\{S_n, n \in \mathbb{N}^*\}$ où $S_n = \sum_{i=1}^n X_i$, converge presque sûrement, vers 0, quand $n \rightarrow \infty$, i.e., $S_n \xrightarrow{p.s} \mu$.

Théorème 1.3.6. (*Kolmogorov SLLN*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires indépendantes, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , dont les moyennes, $\mu_i = \mathbb{E}(X_i)$, et les variances, $\sigma_i^2 = \mathbb{E}((X_i - \mu_i)^2)$, existent, $\forall i \in \mathbb{N}^*$. Alors, si $\sum_{n=1}^{\infty} i^{-2} \sigma_i^2 < \infty$, on a la convergence presque sûre de la suite des moyennes empirique partielles, $\{\bar{X}_n, n \in \mathbb{N}^*\}$, suivante : $\bar{X}_n \xrightarrow{p.s} \bar{\mu}_n$ où $\bar{\mu}_n = \sum_{i=1}^n \mu_i / n$.

Théorème 1.3.7. (*Corollaire de Kolmogorov SLLN*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de v.a. indépendantes, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , de moyennes nulles et telles que $\text{Var}(X_i) = \sigma_i^2 < \infty$. Alors, si $\sum_{n=1}^{\infty} i^{-2} \sigma_i^2 < \infty$, on a, , pour $n \rightarrow \infty$, la convergence presque sûre suivante $\overline{X}_n \xrightarrow{p.s.} 0$.

1.3.2 Théorème limite central (Principe d'invariance)

Introduction

Un estimateur d'un paramètre scalaire (ou vectoriel) inconnu et/ou une statistique de test, liés à un modèle de statistique paramétrique en statistique inférence ou en analyse des séries temporelle sur laquelle porte notre travail, ne sont que des fonctions mesurables de l'échantillon prélevé dans une population statistique caractérisée par cette structure de probabilité. Les estimateurs et les statistiques de test sont, souvent des fonctions de certaines sommes de variables aléatoires, indépendantes ou dépendantes. En effet, une classe particulière importante de ces estimateurs est celle des estimateurs qui sont fonctions de sommes partielles et/ou de moyennes partielles de n variables aléatoires d'une suite de variables aléatoires (non nécessairement indépendantes) $\{X_n, n \in \mathbb{N}^*\}$. Il est utile de noter que dans le cas où une telle expression n'existe pas ou est elle trop complexe, il reste possible d'exprimer, par le biais d'un développement adéquat, la statistique en question sous forme d'une somme de la moyenne empirique, \overline{X}_n (ou de la somme partielle $n \overline{X}_n$, et d'une certaine autre variable aléatoire convergente, selon un certain mode, vers 0. Après avoir présenté la convergence stochastique et la loi faible et la loi forte des grands nombres de ce type de statistiques dans les sections précédentes, nous présentons brièvement les notions fondamentales de la convergence en loi des sommes partielles et/ou des moyennes empiriques partielles des suites de variables aléatoires satisfaisant des hypothèses différentes. Les théorèmes établissant cette convergence sont des versions, qui se diffèrent l'une de l'autre selon des conditions satisfaites par la suite adjacente, du théorème très utilisé en statistique asymptotique à savoir le *Théorème Central Limite*.

Soit $\{\widehat{\phi}_n, n \in \mathbb{N}^*\}$ une suite de estimateurs basés sur un échantillon de taille n , i.e., $\widehat{\phi}_n(X_1, X_2, \dots, X_n)$, $n \in \mathbb{N}^*$. L'objectif de la convergence, en loi, de cette suite d'estimateurs est la re-

cherche d'une certaine fonction de répartition $F(\cdot)$, ne dépendant pas de n , vers laquelle converge la fonction de répartition empirique, $F_n(\cdot)$, de cet estimateur normalisé adéquatement dans le sens de trouver des constants (dépendant de n) α_n et β_n ($\beta_n > 0$) tels que :

$$F_n(x) = P\left(\frac{\hat{\phi}_n - \alpha_n}{\beta_n} \leq x\right) \longrightarrow F(x), \forall x \text{ de l'ensemble de continuité } F(\cdot).$$

A titre d'exemple, considérons une suite de variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli, $B(1, p)$, $\{X_n, n \in \mathbb{N}^*\}$, alors les constantes α_n et β_n adéquates pour la convergence, en loi, de la suite des moyennes empiriques partielles de ces variables aléatoires, \bar{X}_n , sont, respectivement, p et $\frac{pq}{n}$. Plus précisément, on a :

$$F_n(x) = P\left(\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{pq}} \leq x\right) \xrightarrow{n \rightarrow \infty} F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}x^2} dx, \forall x \in \mathbb{R} - \left\{0, \frac{1}{n}, \frac{2}{n} \dots \frac{n}{n}\right\}.$$

C'est-à-dire que la variable aléatoire $\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{pq}}$ où la statistique \bar{X}_n n'est que l'estimateur du maximum de la vraisemblance du paramètre inconnu p , converge en loi, quand $n \rightarrow \infty$, vers une loi normale centrée réduite.

Théorème limite central classique : cas i.i.d.

La plus simple version et la plus utilisée du théorème limite central est le *théorème limite central classique* correspondant à une suite de variables aléatoires indépendantes et identiquement distribuées, *i.i.d.*, dont l'énoncé est donné par le théorème suivant.

Théorème 1.3.8. (Théorème limite central classique)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , de moyenne $\mu = \mathbb{E}(X_n)$ et de variance finie $\sigma^2 = \mathbb{E}((X_n - \mu)^2)$. Alors, la suite des

variables aléatoires $Y_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$, $n \in \mathbb{N}^*$, où \bar{X}_n est la moyenne empirique partielle, converge en loi, quand $n \rightarrow \infty$, vers une variable aléatoire, X , de loi normale centrée et réduite $N(0, 1)$, et on note :

$$Y_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, 1) \quad \text{ou encore} \quad Y_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N(0, 1).$$

Ce théorème peut être donné (en utilisant les sommes partielles à la place des moyennes empiriques partielles) sous la forme équivalente suivante.

Théorème 1.3.9. (*Théorème limite central classique*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , de moyenne $\mu = \mathbb{E}(X_n)$ et de variance finie $\sigma^2 = \mathbb{E}((X_n - \mu)^2)$. Alors, la suite des variables aléatoires $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, $n \in \mathbb{N}^*$, où S_n est la somme partielle, converge en loi, quand $n \rightarrow \infty$, vers une variable aléatoire, X , de loi normale centrée et réduite $N(0, 1)$, et on note :

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, 1) \quad \text{ou encore} \quad Z_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N(0, 1).$$

Théorème limite central cas d'indépendance non i.i.d.

Dans ce paragraphe, nous présentons les versions du théorème limite central correspondant à une suite de variables aléatoires indépendantes mais non nécessairement de même loi, dont l'énoncé de celle dû à Lyapunov est suivant.

Théorème 1.3.10. (*Théorème limite central de Lyapunov*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , de moyennes $\mu_n = \mathbb{E}(X_n)$ et de variances $\sigma_n^2 = \mathbb{E}((X_n - \mu_n)^2)$ et supposons l'existence d'un certain nombre réel δ , $0 < \delta \leq 1$, tel que les moments centrés absolus d'ordre $2 + \delta$ sont finis, i.e.,

$$v_n^{2+\delta} = \mathbb{E}(|X_n - \mu_n|^{2+\delta}) < \infty, \forall n \in \mathbb{N}^*,$$

Soient la suite de variables aléatoires centrées réduites $\{Z_n, n \in \mathbb{N}^*\}$ où $Z_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{S_n}$

avec $S_n^2 = \sum_{i=1}^n \sigma_i^2$, $\forall n \in \mathbb{N}^*$ et la suite de numérique $\{\rho_n, n \in \mathbb{N}^*\}$ où $\rho_n = S_n^{-(2+\delta)} v_n^{2+\delta}$, $\forall n \in \mathbb{N}^*$. Alors, si $\lim_{n \rightarrow \infty} \rho_n = 0$, on a la convergence de la suite $\{Z_n, n \in \mathbb{N}^*\}$, quand $n \rightarrow \infty$,

vers une variable aléatoire de loi normale centrée réduite $N(0, 1)$, et on note :

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{S_n} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, 1) \quad \text{ou encore} \quad Z_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N(0, 1).$$

Lindeberg et Feller ont montré que la condition de l'existence $v_n^{2+\delta}$ n'est pas nécessaire. En effet, ils établissent deux conditions plus faibles et qui sont nécessaires et suffisantes pour que

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{S_n} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, 1).$$

Théorème 1.3.11. (*Théorème limite central de Lindeberg-Feller*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , de moyennes $\mu_n = \mathbb{E}(X_n)$ et de variances $\sigma_n^2 = \mathbb{E}((X_n - \mu_n)^2)$

Soient la suite de variables aléatoires centrées réduites $\{Z_n, n \in \mathbb{N}^*\}$ où $Z_n = \sum_{i=1}^n Y_{n,i}$ avec

$Y_{n,i} = \frac{X_i - \mu_i}{S_n}$ et $S_n^2 = \sum_{i=1}^n \sigma_i^2$. Et considérons les deux conditions suivantes :

A) $\max_{1 \leq i \leq n} \frac{\sigma_i^2}{S_n^2} \longrightarrow 0$, quand $n \rightarrow \infty$, (*condition de l'Uniform Asymptotic Negligibility*)

B) $\forall \varepsilon < 0, \frac{1}{S_n^2} \sum_{i=1}^n \mathbb{E}((X_i - \mu_i)^2 I_{\{|X_i - \mu_i| > \varepsilon S_n\}}) \longrightarrow 0$, quand $n \rightarrow \infty$, (*condition d'intégrabilité uniforme de Lindeberg-Feller*).

Alors, la réalisation simultanée de A) et de B) est une condition nécessaire et suffisant pour la convergence de la suite $\{F_n(z), n \in \mathbb{N}^*\}$, où $F_n(z)$ est la fonction de répartition de Z_n , vers la fonction de répartition $\Phi(x)$ d'une variable aléatoire normale centrée et réduite.

1.4 Dépendance : généralités sur les processus stochastiques

1.4.1 Introduction

Dans les sections précédentes, nous avons exposé les notions fondamentales liées à des suites de variables aléatoires (indépendantes et/ou dépendantes) représentant mathématiquement des phénomènes aléatoires statiques, dans le sens où en fixant un issu $\omega \in \Omega$ de l'expérience, la valeur prise par cette variable aléatoire ne dépendant d'aucun autre fait. Cependant, en analyse des séries chronologiques sur laquelle porte notre thème d'initialisation à la recherche, la variable aléatoire représentant un phénomène étudié est dépendante, outre de l'issu, d'un autre facteur à savoir l'instant, t , de l'observation de ce phénomène ; et donc le modèle mathématique convenable à cette situation est dynamique, c'est-à-dire il évolue, dans le temps. En effet, une série chronologique de taille ou encore de longueur N est un ensemble de N mesures, prises à des instants (généralement) équidistants, x_1, x_2, \dots, x_N , par une variable aléatoire notée $X_t(\omega)$, désignant une grandeur liée à un phénomène sous étude, dépend d'une part de l'issu ω et, d'autre part, de l'instant t de l'observation. Autrement, dit une variable aléatoire scalaire (ou vectoriel de dimension m) $X_t(\omega)$ (notée de façon abrégée X_t) est, dans ce contexte, une fonction mesurable de $T \times \Omega$ dans \mathbb{R} (respectivement dans \mathbb{R}^m) où T est le domaine des valeurs de t .

Une suite de ce type de variables aléatoires, notée $\{X_t; t \in T\}$, est dite processus aléatoire. L'étude probabiliste des processus aléatoires est primordial en analyse des séries chronologiques, car la modélisation et l'étude probabiliste et statistique d'une série chronologique de taille N , qui n'est autre qu'une réalisation (ou une trajectoire) particulière, parmi une infinité de réalisations (trajectoires) possibles, d'un certain processus aléatoire $\{X_t; t \in T\}$, consiste en l'identification du processus générateur de cette série. Il est connu qu'un processus aléatoire $\{X_t; t \in T\}$ est complètement caractérisé par la donnée de la loi conjointe de probabilité des variables aléatoires $X_{t_1}, X_{t_2}, \dots, X_{t_k}$, $k \in \mathbb{N}$. Cependant, la connaissance complète des lois de probabilité de telle sous-famille finie quelconque de variables aléatoires est généralement impossible. Néanmoins, l'analyse théorique permet de s'informer partiellement sur ce processus générateur. A cet effet, on verra ultérieurement que l'analyste des séries

chronologiques impose quelques conditions et hypothèses sur le processus associé à la série chronologique en question. Du fait que l'étude des séries chronologiques est fondée essentiellement sur la notion de processus aléatoire, il semble utile de rappeler quelques définitions et notions de base sur les processus stochastiques.

1.4.2 Notions de base sur les processus aléatoires

Processus aléatoires au second ordre

Un processus aléatoire scalaire ou vectoriel $\{X_t; t \in \mathbb{T}\}$, peut être classé, selon l'ensemble de temps \mathbb{T} , en deux catégories à savoir les processus à temps discret et les processus à temps continu selon que \mathbb{T} est \mathbb{Z} (ou une partie \mathbb{T}) et \mathbb{R} (ou une partie \mathbb{R}), respectivement. Dorénavant, nous considérons dans cette section, sauf mention contraire, les processus stochastiques discrets. Un processus aléatoire $\{X_t; t \in \mathbb{Z}\}$ est dit du second ordre (ou encore de carré intégrable) si, et seulement si, ses deux premiers moments existent et sont finis, i. e.,

$$\mathbb{E}(X_t) < \infty \text{ et } \mathbb{E}(X_t^2) < \infty, \forall t \in \mathbb{T}.$$

La fonction d'autocovariance d'un processus stochastique (du second ordre) $\{X_t; t \in \mathbb{Z}\}$ est définie par :

$$\gamma(t, s) = Cov(X_t, X_s) = \mathbb{E}\left((X_t - \mu_t)(X_s - \mu_s)\right), \forall t, s \in \mathbb{Z},$$

où μ_t est la moyenne, à l'instant t , du processus, i.e., $\mu_t = \mathbb{E}(X_t)$.

En particulier, la variance, à l'instant t , de ce processus, notée $\gamma_0(t)$, est donnée par

$$Var(X_t) = \gamma(t, t) = \gamma_0(t) = Cov(X_t, X_t) = \mathbb{E}(X_t - \mu_t)^2, \forall t \in \mathbb{Z}.$$

1.4.3 Processus stationnaires

Processus strictement stationnaire

Considérons un processus stochastique $\{X_t, t \in \mathbb{Z}\}$ et soit k un entier naturel quelconque alors la fonction de répartition conjointe des k variables aléatoires de ce processus est définie par :

$$F(X_{t_1}, X_{t_2}, \dots, X_{t_k}) = P(X_{t_1} \leq x_1, \dots, X_{t_k} \leq x_m),$$

$$\forall (t_i, x_i) \in \mathbb{Z} \times \mathbb{R}, \forall k \in \mathbb{N}^* \text{ et } \forall t_i \in \mathbb{Z}, i = 1, \dots, k.$$

Définition 1.4.1

Un processus stochastique (du second ordre) $\{X_t, t \in \mathbb{Z}\}$ est dit strictement stationnaire si, et seulement si, on a

$$F(X_{t_1+m}, X_{t_2+m}, \dots, X_{t_k+m}) = F(X_{t_1}, X_{t_2}, \dots, X_{t_k}), \quad \forall k \in \mathbb{N}^* \text{ et } \forall m, t_i \in \mathbb{Z},$$

où $F(\cdot)$ désigne la fonction de répartition du processus.

On remarque que de l'hypothèse de stationnarité stricte d'un processus stochastique découle que la moyenne de ce processus est indépendante du temps et que sa fonction d'autocovariance ne dépend que de la différence entre les deux instants, autrement dit du décalage ou encore du retard $h = t_2 - t_1$. Aussi, il est clair que ce concept est très exigeant d'où la difficulté d'étudier sa vérification. Un autre concept de stationnarité moins exigeant et qu'on peut rencontrer, en pratique, sa vérification est le concept suivant.

Processus faiblement stationnaire**Définition 1.4.2**

Un processus stochastique (du second ordre) $\{X_t, t \in \mathbb{Z}\}$ est dit stationnaire du second ordre (ou faiblement stationnaire) si, et seulement si, on a :

$$\mathbb{E}(X) = \mu, \text{ (une constante ne dépend pas de } t),$$

$$\gamma(t_1, t_2) = \gamma(h), \text{ où } h = t_2 - t_1, \text{ (dépend seulement du retard } h).$$

Fonction d'autocorrélation d'un processus faiblement stationnaire

La fonction d'autocorrélation, de décalage h , $h \in \mathbb{Z}$, notée ρ_h , d'un processus du second ordre stationnaire au sens large est définie comme suit :

$$\rho_h = \frac{\gamma(h)}{\gamma(0)} = \frac{\mathbb{E}(X_t - \mu)(X_{t-h} - \mu)}{\mathbb{E}(X_t - \mu)^2}, \quad h \in \mathbb{Z}.$$

Notons que

$$\rho_{-h} = \rho_h,$$

alors on peut toujours se restreindre aux valeurs entières positives de h . On note aussi que ρ_h est telle que

$$\rho_0 = 1 \text{ et } |\rho_h| \leq 1, \quad h \in \mathbb{Z}.$$

Estimations empiriques des fonctions d'autocovariance et d'autocorrélation

Les estimateurs empiriques C_h et r_h des fonctions d'autocovariance et d'autocorrélation, respectivement, sont donnés comme suit :

1/ Estimateur empirique de la fonction d'autocovariance

La fonction d'autocovariance $\gamma(h)$, $h = 0, 1, \dots$, d'un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ est, habituellement, estimée par l'estimateur empirique C_h (ou encore la fonction d'autocovariance empirique) donné par :

$$C_h = \frac{\sum_{t=h+1}^n (X_t - \bar{X})(X_{t-h} - \bar{X})}{n}, \quad \forall h \in \mathbb{N}, \text{ où } \bar{X} \text{ est l'estimateur empirique de } \mu.$$

2/ Estimateur empirique de la fonction d'autocorrélation

L'estimateur habituel de la fonction d'autocorrélation ρ_h , $h = 0, 1, \dots$, d'un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ notée ρ_h (dite fonction d'autocorrélation empirique) est donné par :

$$r_h = \frac{C_h}{C_0} = \frac{\sum_{t=h+1}^n (X_t - \bar{X})(X_{t-h} - \bar{X})}{\sum_{t=h+1}^n (X_{t-h} - \bar{X})^2}, \quad h \in \mathbb{N}.$$

Processus bruit blanc

Le processus (du second ordre) faiblement stationnaire le simple et le plus utilisable en analyse des séries chronologique est le processus non corrélé dit, en particulier par les ingénieurs, processus bruit blanc. Le rôle primordial que joue le concept d'un processus bruit blanc sera mis en évidence lors de la présentation de la fameuse décomposition d'un processus faiblement stationnaire établie par Wold et son généralisation au cas d'un processus non nécessairement stationnaire dû à *Cramér* dont la définition et la suivante.

Définition 1.4.3

Un processus stochastique, du second ordre, dont la fonction d'autocovariance est donnée par :

$$\gamma(h) = \begin{cases} \sigma^2, & \text{si } h = 0, \\ 0, & \text{ailleurs,} \end{cases}$$

est dit processus non corrélé ou encore processus bruit blanc. Par conséquent, la fonction d'autocovariance de ce processus est donnée par :

$$\rho_h = \begin{cases} 1, & \text{si } h = 0, \\ 0, & \text{ailleurs.} \end{cases}$$

1.4.4 Ergodicité

Une des notions importantes, outre que la stationnarité stricte ou faible, qui joue un rôle prédominant, en analyse des séries chronologiques, en particulier pour l'obtention des théorèmes des grands nombres (faibles et/ou fortes) dans les cas des processus sériellement corrélés, analogues à ceux dans le cas d'indépendance est bien l'ergodicité.

Brièvement, le concept d'ergodicité d'un processus assure qu'il peut prendre n'importe quelle valeur appartenant à son espace de valeurs indépendamment de sa valeur initiale ou encore de son démarrage, de plus il ne peut être absorbé par une partie strictement incluse dans son espace d'état, autrement dit le processus satisfait la propriété d'irréductibilité. Afin de présenter des définitions formelles de la notion d'ergodicité, il semble utile de rappeler le concept d'invariance d'un espace par une transformation donnée \mathcal{T} .

Définition 1.4.4

Soit $\mathbb{R}^{\mathbb{Z}}$ l'espace de toutes les suites de nombres $\{X_t, t \in \mathbb{Z}\}$, alors, la transformation $\mathcal{T} : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$ est :

- i) dite opérateur de translation si $\mathcal{T}(\dots, Y_0, Y_1, Y_2, \dots) = (\dots, Y_1, Y_2, Y_3, \dots)$,
- ii) dite \mathcal{T} -invariant si $\mathcal{T}A = A$, pour quelque soit le sous ensemble $A \subset \mathbb{R}^{\mathbb{Z}}$.

En utilisant la définition de l'invariance, par rapport à une transformation, nous pouvons définir le concept d'ergodicité.

Définition 1.4.5 (Ergodicité)

Un processus stochastique stationnaire $\{X_t, t \in \mathbb{Z}\}$ est dit ergodique si $P((Y_0, Y_1, \dots) \in A) = 0$ ou 1 pour chaque ensemble A \mathcal{T} -invariant.

1.4.5 Deux exemples de processus aléatoires : Martingales et Processus de Wiener

Rappel sur l'espérance conditionnelle par rapport à une tribu

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires intégrables et considérons l'ensemble de toutes informations apportées par la σ -algèbre engendrée par les variables aléatoires $X_1, X_2, X_3, \dots, X_{n-1}, I_{\underline{n-1}}$, (tous le passé du processus jusqu'à l'instant $n-1$), i.e., X_1, X_2, \dots, X_{n-1} engendrent la σ -algèbre $\sigma(X_1, X_2, \dots, X_{n-1})$. Alors l'espérance conditionnelle de la variable aléatoire X_n conditionnée par X_1, X_2, \dots, X_{n-1} , est notée $\mathbb{E}(X_n/I_{\underline{n-1}})$ ou encore plus explicitement $\mathbb{E}(X_n/X_1, X_2, \dots, X_{n-1})$. Il est clair de la définition de la quantité d'information, contenue dans $I_{\underline{t-1}}$, augmente au fur et à mesure que t augmente. En effet, on peut écrire la relation suivante :

$$\dots I_{\underline{t-2}} \subseteq I_{\underline{t-1}} \subseteq I_{\underline{t}} \subseteq I_{\underline{t+1}} \subseteq I_{\underline{t+2}} \dots \subseteq I_{\underline{t+h}} \dots$$

Lors du calcul des espérances conditionnelles, on fait, très souvent, appel aux propriétés fondamentales des espérances conditionnelles (linéaires ou non linéaires) suivantes :

Considérons deux ensembles d'informations, portant sur un processus stochastique $\{X_t, t \in \mathbb{Z}\}$, à savoir I_τ et I_s , alors on peut énoncer les deux propriétés essentielles suivantes :

Propriété 1.4.1 :

$$\mathbb{E}(X_t/I_s) = X_t, \quad \forall t \leq s.$$

Propriété 1.4.2 :

$$\mathbb{E}(X_t/I_\tau) = \mathbb{E}\left(\mathbb{E}(X_t/I_s) / I_\tau\right), \quad \text{pour } s \leq \tau, \text{ autrement dit pour } I_s \subseteq I_\tau,$$

en particulier, on a $\mathbb{E}(X_t) = \mathbb{E}\left(\mathbb{E}(X_t/I_s)\right)$.

Martingales et différences de martingales

La définition le concept d'une suite de martingale est donnée par la définition suivante.

Définition 1.4.6

Une suite de variables aléatoires, $\{X_n, n \in \mathbb{N}^*\}$, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , dont la moyenne commune, $\mu = \mathbb{E}(X_n)$ existe, $\forall n \in \mathbb{N}^*$, est dite une **martingale**, si elle vérifie la condition suivante : $\mathbb{E}(X_n/X_i, \forall i \leq n-1) = X_{n-1}$, presque partout (p.p).

Les suites de submartingale et supmartingale sont définies comme suit.

Définition 1.4.7

Une suite de variables aléatoires, $\{X_n, n \in \mathbb{N}^*\}$, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , dont les espérances, $\mu_n = \mathbb{E}(X_n)$ existent, $\forall i \in \mathbb{N}^*$, est dite

- a) une **submartingale**, si $:\mathbb{E}(X_n / X_i, \forall i \leq n-1) \geq X_{n-1}$, presque partout (p.p),
- b) un **supermartingale**, si $:\mathbb{E}(X_n / X_i, \forall i \leq n-1) \leq X_{n-1}$, presque partout (p.p).

Propriété 1.4.3

Soient $\{X_n, n \in \mathbb{N}^*\}$ est une suite de martingales et $\{Y_n, n \in \mathbb{N}^*\}$ suite de variables aléatoires données par $Y_n = g(X_n)$, pour une certaine fonction $g(\cdot)$, alors la suite $\{Y_n, n \in \mathbb{N}^*\}$ est :

- a) une **submartingale**, si la fonction $g(\cdot)$ est convexe,
- b) une **supermartingale**, si la fonction $g(\cdot)$ est concave.

On peut, facilement, voir qu'une submartingale $\{X_n, n \in \mathbb{N}^*\}$ peut être exprimée sous la forme $X_n = X_{n,0} + X_{n,1}$ où les processus $X_{n,0}$ et $X_{n,1}$ sont, respectivement, donnés par :

$$X_{n,0} = \sum_{j \leq n-1} [X_{j+1} - \mathbb{E}(X_{j+1} / X_i, \forall i \leq j)],$$

et

$$X_{n,1} = \sum_{j \leq n-1} [\mathbb{E}(X_{j+1} / X_i, \forall i \leq j) - X_j].$$

D'où la propriété suivante.

Propriété 1.4.4

Chaque submartingale $\{X_n, n \in \mathbb{N}^*\}$ peut être décomposé en une martingale, $X_{n,0}$, et un résidu, $X_{n,1}$, non négatif non décroissant dépendant seulement de $X_j, j \leq n-1$.

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite martingale de moyenne nulle et définissons les variables aléatoires $Y_n = \nabla X_n = X_n - X_{n-1}$ (en prenant $X_1 = 0$ ou encore $Y_1 = X_1$). On peut facilement vérifier que $X_n = Y_1 + Y_2 + \dots + Y_n$ et que l'ensemble de variables aléatoires $X_1, X_2, X_3, \dots, X_n$ et l'ensemble $Y_1, Y_2, Y_3, \dots, Y_n$ engendrent la même σ -algèbre i.e., $\sigma(X_1, X_2, \dots, X_n) = \sigma(Y_1, Y_2, \dots, Y_n)$. Alors, en tenant compte de la linéarité de l'opérateur de l'espérance conditionnelle, on peut réécrire la condition du martingale $\{X_n, n \in \mathbb{N}^*\}$ sous la forme :

$$\mathbb{E}(X_n / X_i, \forall i \leq n-1) = X_{n-1}, \quad p.p$$

$$\mathbb{E}((X_n - X_{n-1}) / X_i, \forall i \leq n-1) = \mathbb{E}(Y_n / Y_i, \forall i \leq n-1) = 0, \quad p.p$$

D'où la définition d'une suite de différence de martingale dont le rencontre dans l'étude des processus intégré, comme à titre d'exemple une marche aléatoire.

Définition 1.4.8

Une suite de variables aléatoires, $\{X_n, n \in \mathbb{N}^*\}$, définies sur (Ω, \mathcal{F}, P) à valeurs dans \mathbb{R} , de moyenne nulle, est dite une **différence de martingale**, si elle vérifie la condition suivante :

$$\mathbb{E}(X_n / X_i, \forall i \leq n - 1) = 0, \text{ presque partout (p.p).}$$

Martingale double array et martingale array triangulaire

Définition 1.4.9

Une suite de variables aléatoires (non nécessairement indépendantes) qui peuvent être arrangées sous la forme suivante

$$\begin{array}{cccc} X_{11}, & X_{12}, & X_{13}, & \dots, & X_{1k_1}, \\ X_{21}, & X_{22}, & X_{23}, & \dots, & X_{2k_2}, \\ \vdots & \vdots & \ddots & & \vdots \\ X_{n1}, & X_{n2}, & X_{n3}, & \dots, & X_{nk_n}, \end{array}$$

tel que chaque variable d'une ligne est indépendante de chaque élément d'une autre différente ligne, est dite suite de variables aléatoires doublement arrangées (double array random variables).

Dans le cas où $k_n = n, n \geq 1$, les variables sont dites *triangular array random variables*.

Processus de Wiener (Processus de mouvement Brownien)

Le processus de Wiener est un processus stochastique à temps continu dont la définition générale est la suivante.

Définition 1.4.10.

Un processus stochastique (scalaire), à temps continu, $\{W_t, t \geq 0\}$, défini sur un espace de probabilité $(\Omega, \mathcal{F}, P(\cdot))$, est dit mouvement Brownien ou processus de Wiener, si les trois propriétés suivantes sont satisfaites :

- 1) La valeur initial du processus est $W_0 = 0$, i.e.,

$$P(W_0 = 0) = 1,$$

- 2) Les incréments $W_{t_i} - W_{t_{i-1}}$ où $0 \leq t_{i-1} \leq t_i, i = 0, 1, 2, \dots, k$, sont indépendants $\forall i \leq k$, i.e.,

$$P\left(W_{t_i} - W_{t_{i-1}} \in H_i, \forall i \leq k\right) = \prod_{i \leq k} P\left(W_{t_i} - W_{t_{i-1}} \in H_i\right),$$

- 3) L'incrément $W_t - W_s$, pour $\forall t, s$ tels que $0 \leq s < t$, suit la loi normale de moyenne nulle et de variance égale à $t - s$, i.e.,

$$P\left(W_t - W_s \in H\right) = \frac{1}{\sqrt{2\pi}\sqrt{t-s}} \int_H e^{-\frac{1}{2}x^2} dx.$$

Il est montré qu'un cheminement d'un processus de Wiener est continu. La définition générale précédente permet de donner la définition standard d'un processus de Wiener.

Définition 1.4.11.

Un processus stochastique de Wiener standard est un processus scalaire, à temps continu, (scalaire), à temps continu, $\{W(t), t \in [0, 1] \geq 0\}$, tel que :

- 1) $W(0) = 0$,

- 2) Les incréments $W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_k) - W(t_{k-1})$ où $0 \leq t_1 < t_2 < \dots < t_k \leq 1$ sont indépendantes de lois de Gauss de moyennes nulles $\mathbb{E}(W(t_i) - W(t_{i-1})) = 0$ et de variances $\text{Var}(W(t_i) - W(t_{i-1})) = t_i - t_{i-1}, \forall i = 0, 1, 2, \dots, k, \forall i \leq k$, i.e.,

$$W(t) - W(s) \mapsto N(0, t - s), \forall s, t \text{ tels que } 0 \leq s < t \leq 1.$$

- 3) $W(t)$ est continue, en $t, \forall \omega \in \Omega$.

1.5 Théorèmes limites : cas de dépendance

1.5.1 Loi des grands nombres : cas de dépendance

Dans la section précédente, nous avons exposé certaines lois (faibles et fortes) des grands nombres pour quelques statistiques, en particulier, pour la moyenne empirique \bar{X}_n sous l'hypothèse d'indépendance des variables aléatoires $X_1, X_2 \dots X_n$. Cependant, dans beaucoup de problèmes de la statistique inférence paramétrique, en particulier dans l'étude du problème d'estimation ponctuelle et/ou par région de confiance et du problème de test des hypothèses, concernant quelques paramètres inconnus, les estimateurs et/ou les statistiques des tests sont exprimés comme somme ou moyenne empirique des variables aléatoires dépendantes.

Dans la présente section, nous présentons quelques résultats existants concernant les lois (faibles et fortes) des grands nombres pour la moyenne empirique de quelques variables aléatoires dépendantes. L'idée derrière l'obtention de ces résultats est de remplacer l'hypothèse d'indépendance d'une suite de variables aléatoires, qui est trop restreignant, par quelques conditions faibles dans la structure de l'espérance conditionnelle.

Loi forte des grands nombres pour les martingales

Parmi les suites de variables aléatoires (non nécessairement indépendantes) qui jouent un rôle primordial dans le problème de l'étude des lois des grands nombres de variables aléatoires dépendantes, que nous présentons dans ce paragraphes, sont les suites de Martingale, de différence de Martingale, de submartingale et de supermartingale.

Le théorème suivant, dû à *Kolmogorov*, établit la loi forte pour une martingale.

Théorème 1.5.1. (*Kolmogorov SLLN pour une martingale*)

Soit $\{Y_n = \sum_{i \leq n} X_i, n \in \mathbb{N}^*\}$ une martingale, telle que les espérances $\mathbb{E}(|X_i|^p)$, pour certain p ($1 \leq p \leq 2$) existent pour $\forall i \in \mathbb{N}^*$. Supposons de plus l'existence d'une suite numérique positive croissante $\{\alpha_n, n \in \mathbb{N}^*\}$ tels que $\alpha_n \rightarrow \infty$, quand $n \rightarrow \infty$ et

$$\sum_{n \geq 2} \alpha_n^{-2} \mathbb{E} \left(|X_n|^p / X_j, j \leq n-1 \right) < \infty,$$

alors, on a la convergence presque sûre suivante : $\alpha_n^{-1} Y_n \xrightarrow{p.s.} 0$.

Théorème ergodique pour les processus stationnaires

Le théorème ergodique est une version de la loi des grands nombres pour la classe des processus strictement stationnaire intégrables.

Théorème 1.5.2 (Théorème ergodique)

Soit $\{X_t, t \in \mathbb{Z}\}$ un processus strictement stationnaire dont l'espérance existe. Alors :

i) Il existe une variable aléatoire X intégrable vers laquelle converge presque sûrement la suite des moyennes empiriques $\{\bar{X}_t, t \in \mathbb{Z}\}$, quand $n \rightarrow \infty$ où $\bar{X}_t = \frac{1}{n} \sum_{s=1}^t X_s$.

ii) Si de plus $\{X_t, t \in \mathbb{Z}\}$ est ergodique alors X n'est autre que l'espérance de X_t .

1.5.2 Théorème limite central : cas de dépendance

Théorème central limite des martingales

Il est encore possible d'établir quelques inégalités de probabilité et une version de la loi faible des grands nombres, tout en détendant la supposition d'indépendance des variables aléatoires adjacentes satisfaisant la structure d'une martingale. En effet, nous présentons, dans ce paragraphe, les résultats existants concernant les lois faibles des grands nombres des suites de variables aléatoires dépendantes satisfaisant une structure de martingale. Le théorème suivant établit une inégalité de probabilité d'une suite de martingale (de moyenne nulle) dû à Kolmogorov qui étend l'inégalité de probabilité donnée par le Théorème (1.2.4), pour une suite de variables aléatoires indépendantes, dû aussi à Kolmogorov.

Théorème 1.5.3

Soit $\{X_n, n \in \mathbb{N}^\}$ une suite de martingales (de moyenne nulle) et de variances finies, ($\text{Var}(X_n) = \mathbb{E}(X_n^2) < \infty, \forall n \in \mathbb{N}^*$). Alors, on a l'inégalité suivante :*

$$P\left(\max_{1 \leq i \leq n} |X_i| > t\right) \leq t^{-2} \mathbb{E}(X_n^2), \forall t > 0.$$

Le théorème suivant établit la loi faible des grands nombres pour la suite des sommes partielles d'une suite de martingale.

Théorème 1.5.4. (*Théorème limite central d'une martingale dû à White (1984)*)

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de différences de martingale et considérons la suite de ses moyennes empiriques partielles $\{\bar{X}_n, n \in \mathbb{N}^*\}$, $\bar{X}_n = \sum_{i=1}^n X_i/n$. En supposant que les conditions suivantes sont vérifiées :

$$a) \sigma_i^2 = \mathbb{E}(X_i^2) < \infty, \forall i \in \mathbb{N}^* \text{ et } \sum_{i=1}^n \sigma_i^2/n \longrightarrow \sigma^2 > 0,$$

$$b) \sigma_i^2 = \mathbb{E}(|X_i|^r) < \infty, \text{ pour certain } r > 2 \text{ et } \forall i \in \mathbb{N}^*,$$

$$c) \sum_{i=1}^n X_i^2/n \xrightarrow{p} \sigma^2, \text{ pour certain } r > 2 \text{ et } \forall i \in \mathbb{N}^*,$$

alors, on a la convergence en loi suivante $\sqrt{n} \bar{X}_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, \sigma^2)$.

Le théorème suivant étend le théorème limite central de Lindeberg-Lévy, au cas d'un processus martingale array.

Théorème 1.5.5

Soit $\{S_{ni}, \mathcal{F}_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ un processus de martingale array à différences X_{ni} , de moyenne nulle et carré intégrable où $S_{nk_n} = \sum_{i=1}^{k_n} X_{ni}$ et où les σ -algèbres sont telles que $\mathcal{F}_{ni} \subseteq \mathcal{F}_{n+1i} \forall 1 \leq i \leq k_n, n \geq 1$. Soit η^2 une variable aléatoire presque sûrement finie et supposons que :

$$\max_i |X_{ni}| \xrightarrow{p} 0, \tag{1.5.1a}$$

$$\sum_{i=1}^{k_n} X_{ni}^2 \xrightarrow{p} \eta^2, \tag{1.5.1b}$$

$$E\left(\max_i (X_{ni}^2)\right) \text{ est borné en } n. \tag{1.5.1c}$$

Alors, on a la convergence en loi suivante :

$$S_{nk_n} = \sum_{i=1}^{k_n} X_{ni} \xrightarrow{\mathcal{L}} Z,$$

où la variable aléatoire Z a comme fonction caractéristique $\mathbb{E}(\exp(-\frac{1}{2}\eta^2 t^2))$.

Le corollaire suivant assure que la convergence, en loi, de la suite des sommes partielles $\{S_{nk_n}, n \geq 1\}$ établie par le théorème précédent reste valide si les deux conditions (1.5.1a) et (1.5.1c) sont remplacées par la condition de Lindeberg .

Corollaire 1.5.1

Si les deux conditions du théorème précédent (1.5.1a) et (1.5.1c) sont remplacées par la condition de Lindeberg suivante :

$$\forall \varepsilon > 0, \sum_{i=1}^{k_n} \mathbb{E} [X_{ni}^2 \mathbb{I}_{(|X_{ni}| > \varepsilon)} / \mathcal{F}_{ni-1}] \xrightarrow{p} 0 \quad (1.5.2)$$

alors la conclusion du théorème ci-dessus reste vraie.

Théorème central limite fonctionnel

Soit $C_{[0,1]}$ des fonction continues et bornée sur $[0,1]$ et considérons la suite de processus stochastique $\{X_n\}$ dans $C_{[0,1]}$. Soient les variables aléatoires dans $(\Omega, \mathcal{F}, \mathcal{P})$ $\varepsilon_1, \varepsilon_2, \dots$, définissons la somme partielle :

$$\begin{aligned} S_j &= S_{j-1} + \varepsilon_j & (S_0 &= 0) \\ &= \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_j & (j = 1, \dots, n) \end{aligned} \quad (1.5.3)$$

pour $\frac{j-1}{n} \leq t \leq \frac{j}{n}$ on construit :

$$\begin{aligned} X_n(t) &= \frac{1}{\sqrt{n}} S_{j-1} + n \left(t - \frac{j-1}{n} \right) \frac{1}{\sqrt{n}} \varepsilon_j \\ &= X_n \left(\frac{j-1}{n} \right) + n \left(t - \frac{j-1}{n} \right) (X_n \left(\frac{j}{n} \right) - X_n \left(\frac{j-1}{n} \right)) \\ &= \frac{1}{\sqrt{n}} S_j + n \left(t - \frac{j}{n} \right) \frac{1}{\sqrt{n}} \varepsilon_j \\ &= X_n \left(\frac{j}{n} \right) + n \left(t - \frac{j}{n} \right) (X_n \left(\frac{j}{n} \right) - X_n \left(\frac{j-1}{n} \right)) \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} \varepsilon_j + (nt - [nt]) \frac{1}{\sqrt{n}} \varepsilon_{[nt]+1} \end{aligned} \quad (1.5.4)$$

où $X_n(0) = 0$ et $X_n(1) = \frac{1}{\sqrt{n}} S_n$

Le théorème central limite fonctionnel de *Donsker* est donné par :

Théorème 1.5.6 (Donsker)

Supposons que la somme partielle $\{X_n(t)\}$ définie dans (1.5.4) avec $\{\varepsilon_j\}$ iid(0, σ^2) où $\sigma^2 > 0$.

Alors

$$\frac{X_n}{\sigma} \xrightarrow{\mathcal{L}} W.$$

où $W = \{W(t)\}$ est un mouvement brownien unidimensionnel dans $[0, 1]$

Théorème de stabilité par transformation fonctionnelle continue (Continuous mapping theorem) Un résultat très lié au théorème central limite fonctionnel est celui de la stabilité par fonctionnelle continue.

Théorème 1.5.7 (Billingsley (1968))

Soit la fonction continue $h(x)$ définie sur C . Si $X_n \xrightarrow{\mathcal{L}} X$, alors

$$h(X_n) \xrightarrow{\mathcal{L}} h(X).$$

1.6 Equations aux récurrences stochastique

Comme nous l'avons déjà signalé dans l'introduction de ce chapitre, nous pouvons associer aux modèles de séries chronologiques une équation aux récurrences stochastique du type suivant

$$Y_{n+1} = A_{n+1}Y_n + B_{n+1} \quad n \in \mathbb{Z} \quad (1.6.1)$$

Notre premier objectif dans cette section est de déterminer les conditions d'existence d'une unique solution stationnaire de l'équation (1.6.1), dans le cas où la séquence (A_n, B_n) est stationnaire et ergodique.

Les propriétés de stationnarité de la quasi totalité des modèles de série chronologique sont liées directement aux propriétés de stationnarité de la solution de l'équation (1.6.1). Il est suffisant donc d'étudier cette équation. Brandt (1986) a établi un théorème dans lequel il a donné les conditions d'existence d'une solution strictement stationnaire pour (1.6.1). Il l'a utilisé pour le cas unidimensionnel ($d = 1$), mais la démonstration est valable pour tout d .

Théorème 1.6.1 (Brandt, 1986)

Soit $\{(A_n, B_n), n \in \mathbb{Z}\}$ un processus strictement stationnaire et ergodique tel que $E(\log^+(A_0))$ et $E(\log^+(B_0))$ sont toutes les deux finies. Supposons que l'exposant de Lyapounov γ défini par

$$\gamma = \inf\left\{E\left(\frac{1}{n+1} \log \|A_0 A_{-1} \dots A_{-n}\|\right), n \in \mathbb{N}\right\}$$

est strictement négatif. Alors, pour tout $n \in \mathbb{N}$, la série

$$Y_n = \sum_{k=0}^{+\infty} A_n A_{n-1} \dots A_{n-k+1} B_{n-k}$$

converge presque sûrement, et le processus $(Y_n, n \in \mathbb{Z})$ est l'unique solution strictement stationnaire de (1.6.1).

On se place dans le cas où (A_n, B_n) est i.i.d. et on donne quelques définitions relatives à l'équation (1.6.1).

Définition 1.6.1

Un modèle autorégressif généralisé avec des coefficients i.i.d. est un modèle

$$Y_{n+1} = A_{n+1} Y_n + B_{n+1}, \quad n \in \mathbb{Z}, \quad (1.6.2)$$

avec $\{(A_n, B_n), n \in \mathbb{Z}\}$ est une séquence de variables aléatoires i.i.d. définies dans l'espace de probabilité (Ω, A, P) qui prennent leurs valeurs dans $M(d) \times \mathbb{R}^d$. La solution de cette équation est une séquence $(Y_n, n \in \mathbb{Z})$ de \mathbb{R}^d variables aléatoires pour laquelle (1.6.2) est vérifiée.

Dans la suite nous allons définir une solution stationnaire qui ne dépend pas du futur.

Définition 1.6.2 (Solution non anticipative)

Une solution strictement stationnaire non anticipative de (1.6.2) est un processus strictement stationnaire $(Y_n, n \in \mathbb{Z})$ qui est une solution de (1.6.2), tel que pour chaque $p \in \mathbb{Z}$, Y_p est indépendant des variables aléatoires $\{(A_n, B_n), n \in \mathbb{Z}\}$.

Passant à la définition suivante qui introduit l'irréductibilité de (1.6.2).

Définition 1.6.3

Soit H un sous espace affine de \mathbb{R}^d . Il serait dit invariant sous le modèle (1.6.2) si $\{A_0y + B_0; y \in H\}$ est contenue dans H presque sûrement. Le modèle (1.6.2) est appelé irréductible si \mathbb{R}^d est le seul sous espace affine invariant.

Théorème 1.6.2 (Bougerol et Picard, 1992)

Considérons le modèle (1.6.2) avec des coefficients i.i.d. Supposons que ce modèle est irréductible et qu'il a une solution strictement stationnaire et non anticipative $(Y_n, n \in \mathbb{Z})$. Donc il s'ensuit que :

- i) $A_0A_{-1}\dots A_{-k}$ converge vers 0 presque sûrement quand $k \rightarrow +\infty$,
- ii) Pour tout entier n ,

$$Y_n = \sum_{k=0}^{+\infty} A_n A_{n-1} \dots A_{n-k+1} B_{n-k}$$

où la série converge presque sûrement,

- iii) Cette solution est l'unique solution strictement stationnaire de (1.6.2).

Le théorème que nous allons donner par la suite stipule que, sous la condition d'irréductibilité, s'il existe une solution strictement stationnaire non anticipative de (1.6.2) alors l'exposant de Lyapounov γ est strictement négatif.

Théorème 1.6.3 (Bougerol et Picard, 1992)

Supposons que le modèle (1.6.2) avec les coefficients i.i.d est irréductible et que $E(\log^+ \|A_0\|)$ et $E(\log^+ \|B_0\|)$ sont tous les deux finies. Alors (1.6.2) a une solution strictement stationnaire non anticipative si et seulement si l'exposant de Lyapounov est strictement négatif.

Proposition 1.6.1 (Bougerol et Picard, 1992)

Supposons que le modèle (1.6.2) a une solution strictement stationnaire non anticipative $(Y_n, n \in \mathbb{Z})$. Soit H le sous espace affine minimal de \mathbb{R}^d tel que $P(Y_0 \in H) = 1$. Alors H est invariant sous le modèle et n'importe quel sous espace invariant de H porte une solution non anticipative strictement stationnaire.

Le corollaire suivant est une conséquence immédiate du théorème 1.6.1, théorème 1.6.3 et de la proposition 1.6.1.

Corollaire 1.6.1 (Bougerol et Picard, 1992)

Considérons le modèle (1.6.2) avec $E(\log^+ \|A_0\|)$ et $E(\log^+ \|B_0\|)$ sont finies. Supposons qu'il existe une solution non anticipative strictement stationnaire qui n'est pas portée par l'hyperplan affine. Alors les conditions suivantes sont équivalentes :

- i) L'exposant de Lyapounov est strictement négatif,*
- ii) Le modèle est irréductible,*
- iii) Il existe une unique solution stationnaire.*

Dans le cas critique $\gamma = 0$, l'équation (1.6.2) a été étudiée par Babillot et al (1997) pour (A_n) scalaire.

Chapitre 2

Instabilité et Propriétés Asymptotiques de l'estimateur des *MCO* d'un $AR(1)$

2.1 Introduction

Dans le présent chapitre nous étudions les lois de probabilité asymptotiques de l'estimateur des Moindres Carrés Ordinaire (*MCO*) d'un processus autorégressif d'ordre un ($AR(1)$) lorsque les paramètres du modèle sont à l'intérieur, sur la frontière et à l'extérieur de la région de stabilité. Ces distributions sont d'une grande utilité pour la mise en oeuvre statistique d'un modèle, et en particulier pour la construction des intervalles de confiances, des tests de significativité des paramètres, des tests de spécifications particulières... Nous verrons que la distribution limite varie en fonction de la position du paramètre à l'intérieur, sur la frontière ou à l'extérieur du domaine de stabilité.

Dans un premier lieu nous spécifions cette région de stabilité et son influence sur les propriétés probabilistes du modèle autorégressif, à savoir, la stationnarité stricte et au second ordre, l'ergodicité, l'ergodicité géométrique, l'existence des moments d'ordres supérieurs et la structure d'autocovariance, ensuite nous présentons l'estimateur des moindres carrés ordi-

naire (MCO) du paramètre autorégressif φ d'un processus $AR(1)$ que nous notons $\hat{\varphi}$. Enfin, nous présentons les distributions limites de l'estimateur MCO pour chacune des situations. L'intérêt de cet estimateur est qu'il est simple à calculer et qu'il a les mêmes propriétés (optimales) asymptotiques que l'estimateur du maximum de vraisemblance.

2.2 Régions de stabilités d'un processus $AR(1)$

Le comportement probabiliste d'un processus stochastique $\{X_t, t \in \mathbb{Z}\}$, solution d'une équation stochastique paramétrique donnée dépend essentiellement de l'appartenance des paramètres de l'équation à des régions bien spécifiques qui, dans le cas linéaire, sont, comme nous verrons ultérieurement, de trois types. Par conséquent, les comportements de différentes statistiques d'intérêt, liées à l'étude théorique et pratique du modèle, en particulier des estimateurs et des statistiques de tests, diffèrent d'une région à autre, montrant l'importance et l'utilité de spécifier, sans ambiguïté, ces différentes régions. Cependant, le concept de stabilité stochastique est essentiellement lié à la forme de l'équation aux différences d'intérêt mais aussi aux propriétés probabilistes que l'on désire mettre en évidence. En général pour les modèles linéaires le domaine de stabilité est invariant des propriétés probabilistes à considérer. A l'encontre, nous verrons qu'il n'en est pas de même pour les modèles non linéaires et que pour chaque propriété probabiliste donnée, il en existe un domaine de stabilité propre.

Dans cette section, nous nous intéressons à la détermination des trois régions (*région de stabilité*, *région de racines unitaires* et *région d'explosion* du processus) engendrées par le paramètre autorégressif, φ , d'une équation autorégressive d'ordre 1, $AR(1)$, que nous considérons pour tout le reste de ce chapitre. Ainsi, un processus stochastique $\{X_t, t \in \mathbb{Z}\}$ est dit admettre une représentation $AR(1)$ s'il est solution de l'équation aux différences stochastique d'ordre 1 suivante

$$X_t = \varphi X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (2.2.1)$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*), de moyenne nulle et de variance finie σ^2 . On suppose également que $\{\varepsilon_t, t \in \mathbb{Z}\}$ constitue un processus d'innovation pour $\{X_t, t \in \mathbb{Z}\}$ et on a donc $\mathbb{E}(\varepsilon_t X_{t-j}) = 0, \forall t$ et $\forall j \in \mathbb{N}^*$. Dans certains cas, on s'intéresse à la même équation mais avec $t \in \mathbb{N}$.

2.2.1 Existence d'une solution strictement stationnaire

Nous avons vu dans le chapitre 1 un cas plus général d'équations aux différences stochastiques Markoviennes à coefficients aléatoires (i.i.d.) et dont on a étudié l'existence d'une solution (non anticipative) strictement stationnaire. Nous utiliserons donc directement ce résultat pour le cas particulier d'un AR(1) et nous étudions également les propriétés d'ergodicité simple et géométrique, de stationnarité au second-ordre et d'existence des moments d'ordre supérieurs. Le modèle (2.2.1) est un cas particulier de l'équation aux récurrences stochastique

$$Y_{t+1} = A_{t+1}Y_t + B_{t+1}, \quad t \in \mathbb{Z},$$

étudiée par Bougerol et Picard (1992) et décrite dans le Chapitre 1 (section 1.6) où $\{(A_t, B_t)\}$ avec $(A_t, B_t) = (\varphi, \varepsilon_t)$ forme une suite *i.i.d.* telle que $\mathbb{E}(\log^+ \|A_t\|) < \infty$ et $\mathbb{E}(\log^+ \|B_t\|) < \infty$. De plus le modèle (2.2.1) est *irréductible* dans le sens de Bougerol et Picard (1992) puisqu'il est contrôlable par construction.

Théorème 2.2.1. (Bougerol et Picard, 1992)

Une condition nécessaire et suffisante pour que l'équation (2.2.1) admette une solution non-anticipative strictement stationnaire $\{X_t, t \in \mathbb{Z}\}$ de la forme

$$X_t = \sum_{j=0}^{\infty} \varphi^j X_{t-j}, \quad t \in \mathbb{Z}, \quad (2.2.2)$$

où la dernière série converge absolument presque sûrement, est que son plus grand exposant de Lyapunov donné par

$$\gamma = \inf_n \left(\frac{1}{n} \mathbb{E}(\log \|A_n A_{n-1} \cdots A_1\|) \right) = \log |\varphi|,$$

soit strictement négatif. De plus la solution est unique et ergodique.

Remarque 2.2.1

i) La condition du Théorème 2.2.1 sur φ est évidemment équivalente à ce que

$$|\varphi| < 1. \quad (2.2.3)$$

ii) Il n'est pas nécessaire que la variance de ε_t soit finie. En revanche, il suffit que

$\mathbb{E} \log(|\varepsilon_t|)$ le soit et dans ce cas, il se pourrait que la solution $\{X_t, t \in \mathbb{Z}\}$ n'ait ni de moyenne ni de variance finies.

iii) Ce résultat n'a été obtenu que grâce à l'hypothèse clé que $\{\varepsilon_t, t \in \mathbb{Z}\}$ soit un processus *i.i.d.* (i.e. un bbruit blanc fort). Dans le cas de l'hypothèse classique que $\{\varepsilon_t, t \in \mathbb{Z}\}$

soit un bruit blanc faible, on ne peut conclure au sujet de la stationnarité stricte d'une solution éventuelle.

iv) La condition (2.2.3) est donc une condition de stabilité par rapport à la propriété de stationnarité stricte.

v) On peut considérer sous les mêmes hypothèses l'équation

$$X_t = \varphi X_{t-1} + \varepsilon_t, \quad t \in \mathbb{N}, \quad (2.2.4)$$

dont le support est l'ensemble \mathbb{N} et non \mathbb{Z} . Si $\{X_t, t \in \mathbb{N}\}$ est une solution nonanticipative strictement stationnaire de (2.2.4) alors puisque $\{\varepsilon_t, t \in \mathbb{N}\}$ est i.i.d., il est facile de vérifier que le processus $\{(X_t, \varepsilon_t, \varphi), t \in \mathbb{N}\}$ est aussi strictement stationnaire. Donc il peut être étendu sur \mathbb{Z} à n'importe quel nouveau processus stationnaire $\{(X_t, \varepsilon_t, \varphi), t \in \mathbb{Z}\}$ (voir par exemple Billingsley, 1995) qui, de plus, vérifie l'équation $X_t = \varphi X_{t-1} + \varepsilon_t$, pour tout $t \in \mathbb{Z}$. Donc $\{X_t, t \in \mathbb{Z}\}$ est une solution nonanticipative strictement stationnaire de (2.2.1), montrant que l'étude de l'équation (2.2.4) se réduit à celle de (2.2.1).

2.2.2 Existence d'une solution stationnaire au second ordre

On s'intéresse maintenant à l'existence d'une solution de (2.2.1), toujours nonanticipative, mais stationnaire au second ordre seulement. Dans ce cas, le processus solution (s'il existe) doit posséder des moments d'ordre deux et il n'est pas nécessaire que le processus bruit blanc $\{\varepsilon_t, t \in \mathbb{Z}\}$ associé soit fort. Il suffit qu'il soit faible de variance finie et c'est l'hypothèse qu'on considère. Le résultat suivant, classique et pouvant être trouvé dans la plus part des livres sur l'analyse des séries chronologiques, est tiré de Brockwell et Davis (1991).

Théorème 2.2.2. (Brockwell et Davis, 1992)

Une condition nécessaire et suffisante pour que l'équation (2.2.1) admette une unique solution $\{X_t, t \in \mathbb{Z}\}$, nonanticipative, stationnaire au second ordre de la forme,

$$X_t = \sum_{j=0}^{\infty} \varphi^j X_{t-j}, \quad t \in \mathbb{Z},$$

où la dernière série converge en moyenne quadratique, est que (2.2.3) soit vérifiée. Si de plus le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ définie dans (2.2.1) est i.i.d. alors l'unique solution de (2.2.1) est également strictement stationnaire et ergodique.

Il est clair que la moyenne du processus donné par (2.2.1) est nulle puisqu'elle est solution de l'équation

$$\mathbb{E}(X_t) = \varphi \mathbb{E}(X_{t-1}), \quad t \in \mathbb{Z},$$

dont la solution est $\mathbb{E}(X_t) = 0$, $t \in \mathbb{Z}$. La variance, $Var(X_t)$, est par hypothèse finie et vérifie l'équation suivante

$$\begin{aligned} Var(X_t) &= Var(\varphi X_{t-1} + \varepsilon_t)^2, \\ &= \varphi^2 Var(X_{t-1}) + Var(\varepsilon_t) + 2 Cov(X_{t-1}, \varepsilon_t). \end{aligned}$$

Tenant compte, d'une part, du fait que le dernier terme de l'expression précédente est nul, par l'hypothèse de la non corrélation de ε_t avec le passé de X_t , et d'autre part du fait que $Var(\varepsilon_t) = \sigma^2$, on peut réécrire l'équation précédente sous la forme de l'équation aux récurrences (ordinaire) non homogène, d'ordre 1, donnée par :

$$Var(X_t) = \varphi^2 Var(X_{t-1}) + \sigma^2.$$

Cette équation aux différences est stable et possède une solution, $Var(X_t) = \gamma_0$, si et seulement si les solutions de l'équation caractéristique associée à son équation aux différences homogène (cf, Benès, (1967)) :

$$z = \frac{1}{\varphi^2},$$

sont à l'intérieur du cercle unité qui est la condition (2.2.3). Sous une telle condition la variance est donnée par

$$\gamma_0 = \frac{\sigma^2}{1 - \varphi^2}.$$

Remarque 2.2.2

- i) La condition (2.2.3) est connue dans la littérature sous le nom de condition de causalité. En effet sous cette condition il est nécessaire et suffisant que l'équation (2.2.1) soit causale, i.e., on peut exprimer chaque membre X_t du processus comme combinaison linéaire infinie, convergente en moyenne quadratique, du passé du processus $\{\varepsilon_t, t \in \mathbb{Z}\}$.
- ii) On peut également s'intéresser à l'existence d'une solution stationnaire au second ordre de l'équation (2.2.4) dont le support est \mathbb{N} plutôt que \mathbb{Z} . Dans ce cas, les conditions dépendent de l'hypothèse faite sur la distribution de X_0 . Par exemple pour calculer l'espérance sous l'hypothèse de stationnarité au second ordre, on a à résoudre

l'équation aux différences suivante $\mathbb{E}(X_t) = \varphi \mathbb{E}(X_{t-1})$, $t \geq 1$, où $\mathbb{E}(X_0)$ est donnée dont la solution est donnée par $\mathbb{E}(X_t) = \varphi^t \mathbb{E}(X_0)$ et donc une condition nécessaire de stationnarité au premier ordre est que $\mathbb{E}(X_0) = 0$.

2.2.3 Existence d'une solution strictement stationnaire ayant des moments d'ordre supérieurs finis

A priori, les conditions précédentes d'existence de solution stationnaire stricte ou au second ordre ne garantissent pas nécessairement que ces solutions aient des moments d'ordres supérieurs finis. Des conditions sur l'existence de solutions ayant des moments finis sont très importantes en particulier pour l'établissement des propriétés statistiques asymptotiques des estimateurs qui le plus souvent nécessitent l'existence de certains moments. On s'intéresse donc à l'existence d'une solution de (2.2.1), toujours nonanticipative, mais stationnaire au $r^{\text{ème}}$ ordre, r étant un nombre entier strictement positif. Dans ce cas, il est nécessaire de supposer que le processus bruit blanc $\{\varepsilon_t, t \in \mathbb{Z}\}$ associé ait des moments d'ordre r finis et c'est l'hypothèse qu'on retiendra.

$$H_0 : \{\varepsilon_t, t \in \mathbb{Z}\} \text{ est une suite i.i.d., telle que } \mathbb{E}(\varepsilon_t^r) \text{ soit finie.}$$

Le résultat suivant donne des conditions nécessaires et suffisantes pour l'existence d'une solution stationnaire au $r^{\text{ème}}$ ordre.

Théorème 2.2.3 (Feigin et Tweedie, 1985)

Une condition nécessaire et suffisante pour que l'équation (2.2.1) soumise à l'hypothèse H_0 , admette une unique solution $\{X_t, t \in \mathbb{Z}\}$, nonanticipative, stationnaire au $r^{\text{ème}}$ ordre de la forme

$$X_t = \sum_{j=0}^{\infty} \varphi^j X_{t-j}, \quad t \in \mathbb{Z},$$

où la dernière série converge en moyenne d'ordre r , est que (2.2.3) soit vérifiée. Si de plus le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ définie dans (2.2.1) est i.i.d. alors l'unique solution de (2.2.1) est également strictement stationnaire et ergodique.

Remarque 2.2.3

- i) La condition $|\varphi| < 1$ est nécessaire et suffisante pour trois types de stabilité :*
 — Stabilité conduisant à la stationnarité stricte quand le processus des innovations est *i.i.d.*

- Stabilité conduisant à la stationnarité au second ordre lorsque le processus des innovations est un bruit blanc faible.
 - Stabilité conduisant à l'existence des moments d'ordre r quand le processus d'innovation est i.i.d. de moments d'ordre r finis.
- ii)* La dernière condition est également suffisante pour la stabilité entraînant la propriété d'ergodicité géométrique (cf, Feigin et Tweedie, 1985).

2.2.4 Régions de stabilité et régions d'instabilité

Nous avons étudié diverses propriétés probabilistes du modèle donné par l'équation (2.2.1) et on a vu que toutes ces propriétés sont vérifiées sous la condition commune de stabilité (2.2.3). Lorsque le paramètre φ n'est pas dans cette région, il existe en fait deux situations différentes d'instabilité qu'il faudrait spécifier :

- i)* cas où $|\phi| = 1$, i.e. sur la frontière de stabilité.
- ii)* cas où $|\phi| > 1$: à l'extérieur de la région de stabilité (cas explosif).

Dans le premier cas, le plus grand exposant de Lyapunov γ est nul, signifiant qu'on est dans une situation de Chaos dans laquelle il existe une dépendance sensitive aux conditions initiales. La valeur limite d'un processus dépend sensiblement de la valeur initiale. Dans le second par contre, le processus, tend à exploser de plus en plus qu'on avance dans le temps. C'est pourquoi dans ces deux situations on s'intéresse à l'équation de la forme (2.2.4) dont le support est dans \mathbb{N} .

Dans le reste de ce chapitre, nous nous intéressons aux trois régions suivantes :

- i)* région de stabilité $|\phi| < 1$,
- ii)* frontière de stabilité $\phi = 1$,
- iii)* cas explosif : $\phi > 1$.

2.3 Distribution asymptotique de l'estimateur MCO d'un AR(1)

2.3.1 Estimateur des moindres carrés

Considérons une série X_1, X_2, \dots, X_n supposée générée à partir d'un processus $\{X_t, t \in \mathbb{Z}\}$ autorégressif d'ordre un, donné par (2.2.1) et vérifiant les hypothèses suivantes :

$H1 : \{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite *i.i.d.*

$H2 : \mathbb{E}(\varepsilon_t^2) < \infty$.

$H3 : \varphi \in \Phi =]-1, 1[$.

L'hypothèse $H1$ permet d'aspirer à une solution (si elle existe) strictement stationnaire et ergodique. L'hypothèse $H2$ est nécessaire pour l'existence des moments d'ordre 2 du processus autorégressif. Enfin, $H3$ est la condition nécessaire et suffisante de l'existence d'une telle solution quand $H1$ et $H2$ sont vérifiées. Ainsi, d'après les *Propositions* 2.2.1 – 2.2.3, sous les hypothèses $H1$ - $H3$, l'équation (2.2.1) admet une unique solution $\{X_t, t \in \mathbb{Z}\}$ nonanticipative, strictement stationnaire, ergodique et ayant des moments d'ordre 2 finis.

Le but est d'estimer la valeur inconnue φ sur la base de X_1, X_2, \dots, X_n en appliquant la méthode des moindres carrés ordinaire MCO dont le principe consiste à minimiser la somme des carrés des fonctions erreurs $\varepsilon_t(\varphi)$ données par

$$X_t = \varphi X_{t-1} + \varepsilon_t(\varphi), \quad t \in \{1, \dots, n\}, \quad (2.3.1)$$

conditionnellement à la valeur x_0 qu'on suppose connue et fixée. La somme des carrés des erreurs est donnée par :

$$\begin{aligned} Q_n(\varphi; X_0) &= \sum_{t=1}^n \varepsilon_t^2(\varphi), \\ &= \sum_{t=1}^n (X_t - \varphi X_{t-1})^2. \end{aligned} \quad (2.3.2)$$

L'estimateur des moindres carrés $\hat{\varphi}_n$ de φ est une solution mesurable du problème mathématique suivant

$$\hat{\varphi}_n = \arg \min_{\varphi \in]-1, 1[} Q_n(\varphi; X_0). \quad (2.3.3)$$

La résolution du problème (2.2.3) est équivalente à la résolution de l'équation normale

$$\frac{\partial Q_n(\varphi; X_0)}{\partial \varphi} = 0, \quad (2.3.4)$$

qu'on résout par rapport à φ . Or,

$$\frac{\partial Q_n(\varphi; X_0)}{\partial \varphi} = -2 \sum_{t=1}^n X_{t-1} X_t + 2\varphi \sum_{t=1}^n X_{t-1}^2,$$

et donc l'équation (2.2.4) devient

$$\varphi \sum_{t=1}^n X_{t-1}^2 = \sum_{t=1}^n X_{t-1} X_t.$$

D'où l'estimateur $\hat{\varphi}_n$ de φ , basé sur la série X_1, X_2, \dots, X_n de taille n est donné par :

$$\hat{\varphi}_n = \frac{\sum_{t=1}^n X_{t-1} X_t}{\sum_{t=1}^n X_{t-1}^2}. \quad (2.3.5)$$

2.3.2 Distribution asymptotique quand le paramètre est à l'intérieur de la région de stabilité ($|\varphi| < 1$)

Nous nous intéressons dans cette section seulement à la normalité asymptotique de $\hat{\varphi}_n$ de φ du modèle (2.3.4) lorsque la valeur absolue du paramètre φ appartienne à l'intervalle $] -1, 1[$. Le résultat suivant est du à Mann et Wald (1943).

Théorème 2.3.1 (Mann et Wald, 1943)

Sous les hypothèses H1 – H3, on a

$$\sqrt{n} (\hat{\varphi}_n - \varphi) \xrightarrow{\mathcal{L}} N(0, 1 - \varphi^2). \quad (2.3.6)$$

Démonstration.

En remplaçant X_t par sa valeur donnée par (2.2.1) dans (2.3.5) on obtient

$$\hat{\varphi}_n = \frac{\sum_{t=1}^n X_{t-1} (\varphi X_{t-1} + \varepsilon_t)}{\sum_{t=1}^n X_{t-1}^2} = \varphi + \frac{\sum_{t=1}^n X_{t-1} \varepsilon_t}{\sum_{t=1}^n X_{t-1}^2}.$$

Donc

$$\sqrt{n} (\hat{\varphi}_n - \varphi) = \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^n X_{t-1} \varepsilon_t}{\frac{1}{n} \sum_{t=1}^n X_{t-1}^2}. \quad (2.3.7)$$

Étudions en premier lieu la convergence du dénominateur $\frac{1}{n} \sum_{t=1}^n X_{t-1}^2$ de la fraction donnée par (2.3.7). Sous $H1$ et $H3$ on a vu que le processus $\{X_t, t \in \mathbb{Z}\}$ est strictement stationnaire et ergodique. De plus sous $H2$, les moments d'ordre 2 du processus sont finis. Donc par le théorème ergodique, on a

$$\frac{1}{n} \sum_{t=1}^n X_{t-1}^2 \xrightarrow{p.s.} \mathbb{E}(X_0^2) = \frac{\sigma^2}{(1 - \varphi^2)}.$$

Passons à présent à l'étude du numérateur $\frac{1}{\sqrt{n}} \sum_{t=1}^n X_{t-1}\varepsilon_t$, de la fraction donnée par (2.3.6), qui peut s'écrire comme un tableau de martingales (double array martingale) :

$$S_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n X_{t-1}\varepsilon_t = \sum_{t=1}^n X_{n,t}.$$

avec $X_{n,t} = \frac{X_{t-1}\varepsilon_t}{\sqrt{n}}$. Il est clair que $\{S_{n,t}, \mathcal{F}_{n,t}, 1 \leq t \leq n, n \geq 1\}$ est une martingale par rapport à la filtration $\mathcal{F}_{n,t} = \sigma(\{\varepsilon_t, \varepsilon_{t-1}, \dots\})$, i.e. $(\mathcal{F}_{n,t} \subseteq \mathcal{F}_{n+1,t} \forall 1 \leq t \leq n, n \geq 1)$. Vérifions que $\{S_{n,t}, \mathcal{F}_{n,t}, 1 \leq t \leq n, n \geq 1\}$ satisfait les conditions suivantes du *Corollaire 3.1* de Hall et Heyde (voir chapitre 1) :

i) $\max_n |X_{n,t}| \xrightarrow{p} 0$.

ii) $\sum_{t=1}^n X_{n,t}^2$ converge en probabilité vers une certaine v.a. $\eta^2 < \infty$.

iii) La condition de Lindberg :

$$\forall \varepsilon > 0, \sum \mathbb{E} \left[X_{n,t}^2 \mathbb{I}_{(|X_{n,t}| > \varepsilon)} / \mathcal{F}_{n,t-1} \right] \xrightarrow{p} 0.$$

— La condition i) est vérifiée puisque $\mathbb{E}(X_{t-1}\varepsilon_t) = 0 < \infty$ et donc $\frac{X_{t-1}\varepsilon_t}{\sqrt{n}}$ converge presque sûrement vers 0 (voir exercice 2.11 dans Francq et Zakoïan, (2009)).

— Concernant la condition ii), par le théorème ergodique on a sous $H1 - H3$

$$\sum_{t=1}^n \left(\frac{X_{t-1}\varepsilon_t}{\sqrt{n}} \right)^2 = \frac{1}{n} \sum_{t=1}^n X_{t-1}^2 \varepsilon_t^2 \xrightarrow{p.s.} \mathbb{E}(X_{t-1}^2 \varepsilon_t^2) = \sigma^2 \mathbb{E}(X_0^2) = \eta^2 < \infty.$$

— La condition de Lindberg est vérifiée puisque toujours par le théorème ergodique, $\forall \varepsilon > 0$:

$$\frac{1}{n} \sum \mathbb{E} \left[X_{t-1}^2 \varepsilon_t^2 \mathbb{I}_{\left(\left| \frac{X_{t-1}\varepsilon_t}{\sqrt{n}} \right| > \varepsilon\right)} / \mathcal{F}_{nt-1} \right] \xrightarrow{p} \mathbb{E} \left[\mathbb{E}(X_{t-1}^2 \varepsilon_t^2 \mathbb{I}_{\left(\left| \frac{X_{t-1}\varepsilon_t}{\sqrt{n}} \right| > \varepsilon\right)} / \mathcal{F}_{nt-1}) \right] = 0.$$

Donc les conditions du théorème central limite des martingales étant vérifiée, on conclue que

$$\frac{X_{t-1}\varepsilon_t}{\sqrt{n}} \xrightarrow{\mathcal{L}} N(0, \sigma^2 \mathbb{E}(X_0^2)).$$

D'où le théorème est prouvé en appliquant le lemme de Slutsky. ■

Remarque 2.3.1

On peut établir la convergence presque sûre de l'estimateur des moindres carrés $\widehat{\varphi}_n$ vers la vraie valeur du paramètre φ quand celui-ci vérifie : $|\varphi| < 1$. En effet, on peut réécrire (2.3.6) comme suit

$$\widehat{\varphi}_n - \varphi = \frac{\frac{1}{n} \sum_{t=1}^n X_{t-1}\varepsilon_t}{\frac{1}{n} \sum_{t=1}^n X_{t-1}^2},$$

et par le théorème ergodique, on a

$$\frac{1}{n} \sum_{t=1}^n X_{t-1}\varepsilon_t \xrightarrow{p.s.} \mathbb{E}(X_{t-1}\varepsilon_t) = 0. \quad (2.3.8)$$

et

$$\frac{1}{n} \sum_{t=1}^n X_{t-1}^2 \xrightarrow{p.s.} \mathbb{E}(X_{t-1}^2). \quad (2.3.9)$$

En combinant (2.3.7) et (2.3.8) on conclue via le lemme de Slutsky que

$$\widehat{\varphi}_n \xrightarrow{p.s.} \varphi. \quad (2.3.9a)$$

2.3.3 Distribution asymptotique quand le paramètre est sur la frontière de la région de stabilité ($\varphi = 1$)

On a vu que dans le cas de la stabilité de l'équation (2.2.1), les propriétés asymptotiques de l'estimateur des moindres carrés ont été établies en se basant d'une part sur le théorème ergodique et la stationnarité stricte du processus et d'autre part sur le théorème central limite des martingales qui lui même est fondé sur le théorème ergodique. Lorsque $\varphi = 1$ le processus (2.3.4) qui constitue une marche aléatoire ne peut être stationnaire ni ergodique et

les arguments précédents ne peuvent être utilisés et une approche totalement différente est à utiliser. Une telle approche a été le fruit d'un long parcours. Erdos et Kac (1946) furent les premiers à considérer la distribution limite de certaines fonctions d'une marche aléatoire. White (1958) conjectura une distribution limite (fonction d'intégrales de mouvements browniens) de l'estimateur des moindres carrés lorsque $\varphi = 1$ et attira l'attention sur l'application du théorème central limite fonctionnel de Donsker (1951) comme moyen pour prouver son assertion dont la démonstration restera un problème ouvert. Rao (1978, 1980) en déterminant une autre forme équivalente de la distribution limite mais en utilisant des techniques différentes (voir aussi Evans et Savin, (1981)). Enfin, motivé par les travaux de Dickey et Fuller (1979, 1981), Phillips (1987) donna une démonstration rigoureuse de la conjecture de White (1958) en utilisant le théorème central limite fonctionnel (*TCLF*) tout en se basant sur une machinerie mathématique élaborée par Billingsley (1968).

La plus simple illustration de l'approche de Phillips (1987) peut être vue sur la marche aléatoire suivante

$$X_t = \varphi X_{t-1} + \varepsilon_t, \quad t \geq 0, \quad (2.3.10)$$

avec $\varphi = 1$ et $\{\varepsilon_t\} \sim iid(0, \sigma^2)$. En itérant successivement (2.3.9) on obtient

$$X_t = X_0 + \sum_{t=1}^n \varepsilon_t = X_0 + S_t, \quad t \geq 0.$$

Phillips (1987) suppose que $X_0 = 0$ et l'équation précédente devient alors

$$X_t = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t, \quad t \geq 0. \quad (2.3.11)$$

Pour $\varphi = 1$, la distribution asymptotique donnée par le *Théorème 2.3.1* (voir relation (2.3.5)) devient

$$\sqrt{n}(\hat{\varphi}_n - \varphi) \xrightarrow{\mathcal{L}} N(0, 0),$$

i.e.

$$\sqrt{n}(\hat{\varphi}_n - \varphi) \xrightarrow{p} 0.$$

La distribution limite est ainsi dégénérée et il faudrait multiplier $\hat{\varphi}_n - \varphi$ par un coefficient plus grand que \sqrt{n} , signifiant que la vitesse de convergence de $\hat{\varphi}_n - \varphi$ vers 0 est plus grande

dans le cas où $\varphi = 1$ que celle dans le cas de stabilité. En revanche, on vérifie facilement (voir Hamilton, 1994, p. 476) qu'il faut multiplier $\widehat{\varphi}_n - \varphi$ plutôt par n pour éviter d'avoir une distribution limite dégénérée et la relation (3.3.6) peut alors être écrite comme suit

$$n(\widehat{\varphi}_n - \varphi) = \frac{\frac{1}{n} \sum_{t=1}^n X_{t-1} \varepsilon_t}{\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2}. \quad (2.3.12)$$

L'idée fondamentale de Phillips (1987) qui repose sur (2.3.11) peut être résumée en trois étapes essentielles :

Première étape : En premier lieu on exprime chacun des numérateur et dénominateur de (2.3.11) comme fonction d'un processus aléatoire $\{X_n(r), r \in [0, 1]\}$ à temps continu (dit aussi fonction aléatoire) indexé dans $[0, 1]$ et défini par

$$X_n(r) = \begin{cases} 0 & 0 \leq r < \frac{1}{n}, \\ \frac{1}{n} \varepsilon_1 & \frac{1}{n} \leq r < \frac{2}{n}, \\ \frac{1}{n} (\varepsilon_1 + \varepsilon_2) & \frac{2}{n} \leq r < \frac{3}{n}, \\ \vdots & \\ \frac{1}{n} (\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n) & r = 1, \end{cases} \quad (2.3.13)$$

ou encore de manière condensée

$$X_n(r) = \frac{1}{n} X_{[nr]}, \quad \frac{j-1}{n} \leq r < \frac{j}{n} \quad j = 1, \dots, n.$$

Il est clair que l'aire délimitée par l'axe des abscisses, les axes menés par les droites $r = 0$ et $r = 1$ et la courbe représentative de cette fonction en escalier est d'une part donnée par

$$\int_0^1 X_n(r) dr,$$

et d'autre part peut être écrite comme la somme des aires de n rectangles (dont le t -ième

est de longueur $\frac{1}{n}$ et de largeur $\frac{X_{t-1}}{n}$ et donc d'aire $\frac{X_{t-1}^2}{n}$) et est alors égale à

$$n^{-2} \sum_{t=1}^n X_{t-1}^2.$$

D'où, on peut écrire

$$n^{-2} \sum_{t=1}^n X_{t-1} = \int_0^1 X_n(r) dr. \quad (2.3.14)$$

Pour exprimer le dénominateur de (2.3.11) en fonction de $\{X_n(r), r \in [0, 1]\}$, s'inspirant de (2.3.13), on a besoin d'introduire la fonction stochastique suivante $\{S_n(r), r \in [0, 1]\}$ définie par

$$S_n(r) = \begin{cases} 0 & 0 \leq r < \frac{1}{n}, \\ \frac{1}{n} X_1^2 & \frac{1}{n} \leq r < \frac{2}{n}, \\ \frac{1}{n} X_2^2 & \frac{2}{n} \leq r < \frac{3}{n}, \\ \vdots & \\ \frac{1}{n} X_n^2 & r = 1, \end{cases} \quad (2.3.15)$$

et pouvant s'exprimer en fonction de $\{X_n(r), r \in [0, 1]\}$ comme suit

$$S_n(r) = n (X_n(r))^2. \quad (2.3.16)$$

En évaluant l'aire délimitée par l'axe des abscisses, la courbe de $S_n(r)$ et les axes menés par les droites $r = 0$ et $r = 1$, il s'ensuit que

$$\int_0^1 S_n(r) dr = \frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2,$$

ou encore

$$\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2 = n \int_0^1 X_n(r)^2 dr, \quad (2.3.17)$$

exprimant ainsi le dénominateur de (2.3.11) en fonction de $\{X_n(r), r \in [0, 1]\}$. Concernant le numérateur de (2.3.11), on a d'après (2.3.9)

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

ou encore

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

Donc

$$\begin{aligned}
 \sum_{t=1}^n X_{t-1} \varepsilon_t &= \frac{1}{2} \sum_{t=1}^n (X_t^2 - X_{t-1}^2) - \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2, \\
 &= \frac{1}{2} [(X_1^2 - X_0^2) + (X_2^2 - X_1^2) + \dots + (X_n^2 - X_{n-1}^2)] - \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2, \\
 &= \frac{1}{2} X_n^2 - \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2,
 \end{aligned}$$

et ainsi,

$$n^{-1} \sum_{t=1}^n X_{t-1} \varepsilon_t = \frac{1}{2} n^{-1} X_n^2 - n^{-1} \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2.$$

En utilisant (2.3.14) ($n^{-1} X_n^2 = S_n(1)$) on trouve enfin

$$\begin{aligned}
 n^{-1} \sum_{t=1}^n X_{t-1} \varepsilon_t &= \frac{1}{2} S_n(1) - n^{-1} \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2, \\
 &= \frac{1}{2} n X_n^2(1) - n^{-1} \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2,
 \end{aligned} \tag{2.3.18}$$

exprimant enfin le numérateur de (2.3.11) en fonction de $\{X_n(r), r \in [0, 1]\}$.

Deuxième étape :

La deuxième étape dans l'approche de Phillips (1987) consiste à appliquer d'abord le théorème central limite fonctionnel (**Théorème 1.5.5**) sur la fonction aléatoire $\{X_n(r), r \in [0, 1]\}$, i.e.

$$\sqrt{n} X_n(\cdot) \xrightarrow{\mathcal{L}} \sigma W(\cdot), \tag{2.3.19}$$

(voir aussi chapitre 1) puis utiliser le théorème de stabilité par transformation fonctionnelle continue (continuous mapping theorem, de Billingsley, (1968)) selon lequel si

$$S_n(\cdot) \xrightarrow{\mathcal{L}} S(\cdot)$$

et $g(\cdot)$ est une fonctionnelle continue alors

$$g(S_n(\cdot)) \xrightarrow{\mathcal{L}} g(S(\cdot)). \tag{2.3.20}$$

Par exemple, gardant en vue (2.3.18) et puisque la fonctionnelle qui pour toute fonction aléatoire $\sqrt{n}X_n(\cdot)$ associe $\sigma^{-1}\sqrt{n}X_n(\cdot)$ est continue, donc

$$\sqrt{n} \sigma^{-1} X_n(\cdot) \xrightarrow{\mathcal{L}} W(\cdot). \quad (2.3.21)$$

De même, la fonctionnelle qui pour toute fonction $X_n(\cdot)$ associe $\int_0^1 X_n(r) dr$ est continue et par conséquent on en déduit à partir de (2.3.18)

$$\sqrt{n} \int_0^1 X_n(r) dr \xrightarrow{\mathcal{L}} \sigma \int_0^1 W(r) dr.$$

En utilisant ces deux théorèmes clés (théorème central limite fonctionnel et théorème de stabilité par continuité), on peut trouver les distributions limites du numérateur et du dénominateur de (2.3.11). En effet, concernant le dénominateur de (2.3.11) en utilisant (2.3.13), (2.3.18) et (2.3.19) on trouve alors

$$\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2 = \int_0^1 (\sqrt{n} X_n(r))^2 dr, \xrightarrow{\mathcal{L}} \sigma^2 \int_0^1 W(r)^2 dr. \quad (2.3.22)$$

Pour le numérateur de (2.3.11), par la loi des grands nombres on a

$$n^{-1} \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2 \xrightarrow{p} \frac{1}{2} \mathbb{E}(\varepsilon_t^2) = \frac{1}{2} \sigma^2.$$

et en utilisant (2.3.17) – (2.3.19), (2.3.21) et le lemme de Slutsky on trouve

$$\begin{aligned} n^{-1} \sum_{t=1}^n X_{t-1} \varepsilon_t &= \frac{1}{2} (\sqrt{n} X_n(1))^2 - n^{-1} \frac{1}{2} \sum_{t=1}^n \varepsilon_t^2, \\ &\xrightarrow{\mathcal{L}} \frac{1}{2} \sigma^2 W(1)^2 - \frac{1}{2} \sigma^2. \end{aligned} \quad (2.3.23)$$

Troisième étape :

La troisième étape consiste à conclure le résultat en utilisant de nouveau le théorème de stabilité par fonctionnelle continue à (2.3.11) tout en utilisant (2.3.21) – (2.3.22). Avant d'exposer la distribution limite de l'estimateur des moindres carrés dans le cas instable, nous jugeons utile de rappeler le lemme suivant qui résume la démarche précédente et dont la démonstration a été déjà donnée plus haut.

Lemme 2.3.1 (Phillips, (1987))

Supposons que $\{X_t, t \geq 0\}$ suit une marche aléatoire sans dérive

$$X_t = X_{t-1} + \varepsilon_t, t \geq 0,$$

avec $X_0 = 0$ et $\{\varepsilon_t\}$ est une suite i.i.d. de moyenne nulle et de variance σ^2 . Alors, les assertions suivantes sont vraies

$$\begin{aligned} i) \quad & n^{-1} \sum_{t=1}^n X_{t-1} \varepsilon_t \xrightarrow{\mathcal{L}} \frac{1}{2} \sigma^2 [(W(1))^2 - 1], \\ ii) \quad & n^{-2} \sum_{t=1}^n X_{t-1}^2 \xrightarrow{\mathcal{L}} \sigma^2 \int_0^1 [W(r)]^2 dr. \end{aligned}$$

Il est possible à présent de donner la loi asymptotique de l'estimateur des MCO du paramètre φ lorsque ce dernier est égale à l'unité.

Théorème 2.3.2 (Phillips, (1987))

Sous les mêmes hypothèses du Lemme 2.3.1, on a

$$n(\widehat{\varphi}_n - 1) \xrightarrow{\mathcal{L}} \frac{\frac{1}{2} [(W(1))^2 - 1]}{\int_0^1 [W(r)]^2 dr}. \quad (2.3.24)$$

Démonstration. La preuve du théorème découle du Lemme 2.3.1 et de l'application du théorème de stabilité des fonctionnelles continues sur (2.3.11). ■

Remarque 3.3.2

- i) Le terme initial X_0 n'affecte pas la distribution asymptotique et peut être pris arbitrairement.
- ii) La loi asymptotique $n(\widehat{\varphi}_n - 1)$ n'est pas symétrique comme dans le cas de la stabilité.
- iii) D'après le théorème précédent, $\widehat{\varphi}_n$ est sur-consistent dans le sens où sa vitesse de convergence qui est n est plus grande que celle dans le cas stable (\sqrt{n}).

2.3.4 Distribution asymptotique quand le paramètre est en dehors de la région de stabilité ($\varphi > 1$)

Soit le processus autorégressif d'ordre un défini par :

$$X_t = \varphi X_{t-1} + \varepsilon_t, \quad t \geq 1 \quad (2.3.25)$$

avec $|\varphi| > 1$ et $\{\varepsilon_t, t = -\infty, +\infty\} \sim iid(0, \sigma^2)$. Le processus ci-dessus peut être écrit sous la forme :

$$X_t = \sum_{j=0}^{n-1} \varphi^j \varepsilon_{t-j} \quad (2.3.26)$$

Il s'ensuit que pour $|\varphi| > 1$, on a :

$$\begin{aligned} Var(X_t) &= \left(E \left(\sum_{j=0}^{n-1} \varphi^j \varepsilon_{t-j} \right)^2 \right) - \left(\mathbb{E} \left(\sum_{j=0}^{n-1} \varphi^j \varepsilon_{t-j} \right) \right)^2 = \left(\sum_{j=0}^{n-1} \varphi^{2j} \mathbb{E}(\varepsilon_{t-j})^2 \right) - 0, \\ &= \sigma^2 (1 + \varphi^2 + \dots + \varphi^{2(t-1)}) = \frac{(\varphi^{2t} - 1)}{(\varphi^2 - 1)} \sigma^2 \quad |\varphi| > 1, t \geq 1. \end{aligned}$$

Donc $Var(X_t)$ augmente exponentiellement avec t . d'autre part, on a

$$X_t = \sum_{i=1}^{t-1} \varphi^i \varepsilon_{t-i} = \sum_{i=1}^n \varphi^{t-i} \varepsilon_i = \varphi^t \sum_{i=1}^n \varphi^{-i} \varepsilon_i, \quad (2.3.27)$$

$$X_t = \begin{cases} \varphi_0 + \varphi_1 X_{t-1} + \varepsilon_t & t \geq 1 \\ X_0 & t = 0. \end{cases}$$

Alors, on a : $Q(\varphi_0, \varphi_1) = \sum_{t=1}^n \varepsilon_t^2 = \sum_{t=1}^n (X_t - (\varphi_0 + \varphi_1 X_{t-1}))^2$.

Donc et après calcul on aura :

$$\begin{aligned} \frac{\partial Q(\varphi_0, \varphi_1)}{\partial \varphi_0} &= \sum_{t=1}^n -2(X_t - \varphi_0 - \varphi_1 X_{t-1}) \\ \frac{\partial Q(\varphi_0, \varphi_1)}{\partial \varphi_0} &= 0 \implies n\varphi_0 = \sum_{t=1}^n X_t - \sum_{t=1}^n \varphi_1 X_{t-1} \\ \varphi_0 &= \frac{1}{n} X_t - \varphi_1 \frac{1}{n} X_{t-1} \end{aligned}$$

Posons $\bar{X}_{(i)} = \frac{1}{n} \bar{X}_{t+i}$, $i = 0, -1$, on obtient donc :

$$\hat{\varphi}_0 = \bar{X}_{(0)} - \varphi_1 \bar{X}_{(-1)}. \quad (2.3.28)$$

En suivant les mêmes étapes on obtient l'estimateur $\hat{\varphi}_1$ en fonction de $\hat{\varphi}_0$

$$\hat{\varphi}_1 = \frac{\sum_{t=1}^n (X_{t-1} - \bar{X}_{(-1)}) (X_t - \bar{X}_{(0)})}{\sum_{t=1}^n (X_{t-1} - \bar{X}_{(-1)})^2} \quad \text{avec } \hat{\varphi}_0 = \bar{X}_{(0)} - \hat{\varphi}_1 \bar{X}_{(-1)} \quad (2.3.29)$$

Hasza (1977) a étudié les estimateurs du modèle (2.3.25) et le théorème suivant est pris de ce travail.

Théorème 3.2.3 (Hasza 1977)

Soit le processus $\{X_t, t \geq 0\}$ définie par :

$$X_t = \begin{cases} \varphi_0 + \varphi_1 X_{t-1} + \varepsilon_t & t \geq 1 \\ X_0 & t = 0, \end{cases}$$

et $\{\varepsilon_t\}$ est une suite de variables aléatoires iid $(0, \sigma^2)$. Soit les estimateurs des moindres carrés donnés par :

$$X_t = \begin{cases} \hat{\varphi}_0 = \bar{X}_{(0)} - \hat{\varphi}_1 \bar{X}_{(-1)} \\ \hat{\varphi}_1 = \frac{\sum_{t=1}^n (X_{t-1} - \bar{X}_{(-1)}) (X_t - \bar{X}_{(0)})}{\sum_{t=1}^n (X_{t-1} - \bar{X}_{(-1)})^2} \end{cases}$$

La distribution limite de $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\hat{\varphi}_1 - 1)$ est celle de $W_1 W_2^{-1}$, où W_1 et W_2 sont des variables aléatoires indépendantes :

$$(W_1, W_2) = \left(\sum_{i=1}^{\infty} \varphi_1^{-i} \varepsilon_i, \delta_0 \sum_{j=1}^{\infty} \varphi_1^{-j} \varepsilon_j \right)$$

et

$$\delta_0 = X_0 + \varphi_0 (\varphi_1 - 1)^{-1}.$$

La distribution limite de $\sqrt{n}(\hat{\varphi}_0 - \varphi_0)$ est celle d'une variable aléatoire $N(0, \sigma^2)$ et $\sqrt{n}(\hat{\varphi}_0 - \varphi_0)$ est indépendante de $\varphi_1^n (\hat{\varphi}_1 - \varphi_1)$ en limite. Si $\{\varepsilon_t\} \sim iidN(0, \sigma^2)$, $X_0 = 0$ et $\varphi_0 = 0$, alors $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\hat{\varphi}_1 - 1)$ converge en distribution vers une variable aléatoire de Cauchy.

Démonstration. On a :

$$X_t = \begin{cases} \varphi_0 + \varphi_1 X_{t-1} + \varepsilon_t & t \geq 1 \\ X_0 & t = 0 \end{cases}$$

Donc X_t peut s'écrire sous la forme :

$$\begin{aligned} X_t &= \varphi_0 + \varphi_1^t \sum_{i=1}^t \varphi_1^{-i} \varepsilon_i & t \geq 1 \\ &= \varphi_1^t X_0 + \varphi_1^t \varphi_0 \sum_{j=1}^t \varphi_1^{-j} + \varphi_1^t \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j \\ &= \varphi_1^t \left(\underbrace{X_0 + \varphi_0 \sum_{j=1}^t \varphi_1^{-j} + \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j}_{\delta_{0t}} \right) \\ &= \varphi_1^t Y_t & t \geq 1 \end{aligned}$$

où

$$(Y_t, \delta_{0t}) = \left(\delta_{0t} + \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j, X_0 + \varphi_0 \sum_{j=1}^t \varphi_1^{-j} \right)$$

$$\text{On a } Y_t \xrightarrow{p.s} Y = \delta_0 + \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j \text{ où } \delta_0 = X_0 + \varphi_0 \sum_{j=1}^t \varphi_1^{-j}$$

$$\sum_{j=1}^t \varphi_1^{-j} = \frac{1}{(\varphi - 1)}$$

d'où

$$\delta_0 = X_0 + \varphi_0 (\varphi - 1)^{-1}$$

$$Y_t \xrightarrow{p.s} Y \text{ donc } \varphi_1^{-t} X_t \xrightarrow[t \rightarrow \infty]{p.s} Y.$$

$$\begin{aligned}
 Y_t &= Y_0 + \varphi_0 \sum_{j=1}^t \varphi_1^{-j} + \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j \\
 &= Y_0 + \varphi_0 \sum_{j=1}^t \varphi_1^{-j} + \varphi_0 \sum_{j=t+1}^{\infty} \varphi_1^{-j} + \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j + \sum_{j=t+1}^{\infty} \varphi_1^{-j} \varepsilon_j, \\
 Y - Y_t &= \left(Y_0 + \varphi_0 \sum_{j=1}^t \varphi_1^{-j} + \varphi_0 \sum_{j=t+1}^{\infty} \varphi_1^{-j} + \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j + \sum_{j=t+1}^{\infty} \varphi_1^{-j} \varepsilon_j \right) - \\
 &\quad Y_0 - \varphi_0 \sum_{j=1}^t \varphi_1^{-j} - \sum_{j=1}^t \varphi_1^{-j} \varepsilon_j \\
 &= \varphi_0 \sum_{j=t+1}^{\infty} \varphi_1^{-j} + \sum_{j=t+1}^{\infty} \varphi_1^{-j} \varepsilon_j \\
 &= \sum_{j=t+1}^{\infty} (\varphi_0 \varphi_1^{-j} + \varphi_1^{-j} \varepsilon_j) = O_p(|\varphi_1|^{-1})
 \end{aligned}$$

L'erreur dans l'estimateur des moindres carrés de φ_1 est donnée par :

$$\hat{\varphi}_1 - \varphi_1 = \frac{\sum_{t=1}^n X_{t-1} \varepsilon_t - n \bar{X}_{(-1)} \bar{\varepsilon}}{\sum_{t=1}^n X_{t-1}^2 - n \bar{X}_{(-1)}^2}$$

où

$$\begin{aligned}
 \varphi_1^{-2n} \sum_{t=1}^n X_{t-1}^2 &= \sum_{j=0}^{n-1} \varphi_1^{-2j-2} Y^2 + 2 \sum_{t=1}^n \varphi_1^{2(t-n)-2} Y (Y_{t-1} - Y) + \\
 &\quad \sum_{t=1}^n \varphi_1^{2(t-n)-2} (Y_{t-1} - Y)^2.
 \end{aligned}$$

En utilisant :

$$\begin{aligned}
 \mathbb{E} [(Y_{t-1} - Y)^2] &= (1 - \varphi_1^2)^{-1} \varphi_1^{2t} \sigma^2 + \varphi_0^2 (\varphi_1 - 1)^{-2} \varphi_1^{-2(t-1)} = O_p(|\varphi_1|^{-2t}) \\
 \text{Var} \left[\sum_{t=1}^n \varphi_1^{2(t-n)-2} \sum_{j=t+1}^{\infty} \varphi_1^{-j} \varepsilon_j \right] &= O_p(|\varphi_1|^{-2n}).
 \end{aligned}$$

Et

$$\sum_{t=1}^n \varphi_1^{2(t-n)-2} (Y_{t-1} - Y)^2 = O_p(n |\varphi_1|^{-2n}).$$

On obtient

$$\varphi_1^{-2n} \sum_{t=1}^n X_{t-1}^2 = (\varphi_1^2 - 1)^{-1} X^2 + O_p(|\varphi_1|^{-n}).$$

On a aussi

$$\begin{aligned} \varphi_1^{-n} \sum_{t=1}^n X_{t-1} \varepsilon_t &= Y \sum_{t=1}^n \varphi_1^{t-n-1} \varepsilon_t + O_p(n |\varphi_1|^{-n}) \\ \varphi_1^{-n} \bar{X}_{(-1)} &= n^{-1} (\varphi_1 - 1)^{-1} Y + O_p(|\varphi_1|^{-2}). \end{aligned}$$

Et

$$\left[n^{\frac{1}{2}} \varphi_1^{-n} \bar{X}_{(-1)} \bar{\varepsilon}, n \varphi_1^{-2n} \bar{X}_{(-1)}^2 \right] = O_p(n^{-1}).$$

De plus

$$\varphi_1^{-2n} \sum_{t=1}^n (X_{t-1} - \bar{X}_{(-1)})^2 = (\varphi_1^2 - 1)^{-1} Y^2 + O_p(n^{-1}),$$

et

$$\varphi_1^{-n} \sum_{t=1}^n (X_{t-1} - \bar{X}_{(-1)}) (\varepsilon_t - \bar{\varepsilon}) = Y \sum_{t=1}^n \varphi_1^{t-n-1} \varepsilon_t + O_p(n^{-\frac{1}{2}}).$$

Il s'en suit que

$$(\varphi_1^2 - 1)^{-1} \varphi_1^n (\hat{\varphi}_1 - \varphi_1) = Y^{-1} Z_n + O_p(n^{-\frac{1}{2}}).$$

où

$$Z_n = \sum_{t=1}^n \varphi_1^{t-n-1} \varepsilon_t = \sum_{j=1}^n \varphi_1^{-j} \varepsilon_{n-j+1}.$$

Puisque

$$Z_n = \sum_{j < \frac{n}{2}} \varphi_1^{-j} \varepsilon_{n-j+1} + O_p(|\varphi_1|^{-n}).$$

et

$$Y = \delta_0 + \sum_{j < \frac{n}{2}} \varphi_1^{-j} \varepsilon_j + O_p(|\varphi_1|^{-n}).$$

Z_n et Y sont asymptotiquement indépendantes. Donc $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\widehat{\varphi}_1 - \varphi_1)$ est une distribution limite non dégénérée.

Si $\{\varepsilon_t\}$ est $Ni(0, \sigma^2)$ alors Z_n et Y , étant tous deux une combinaison linéaire des ε_t , ont une distribution normale. La moyenne de Z_n est zéro et la moyenne de Y est δ_0 , donc la distribution asymptotique de $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\widehat{\varphi}_1 - \varphi_1)$ est celle de $W_1 W_2^{-1}$, où (W_1, W_2) est un vecteur normal.

Si $\delta_0 = 0$ et $\{\varepsilon_t\}$ est normal, alors la distribution asymptotique de $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\widehat{\varphi}_1 - \varphi_1)$ est celle de la quotient de deux variables indépendantes normales, qui est la distribution de Cauchy. ■

2.4 Simulation

Dans cette section nous présentons, tout en évoquant des commentaires, les résultats d'une étude de simulation intensive portant, d'une part, sur les comportements de l'estimateur MCO , $\widehat{\varphi}_n$, du paramètre inconnu φ , d'un modèle autorégressif d'ordre un, et l'estimateur de son écart-type $\widehat{\sigma}_{\widehat{\varphi}_n}$, et d'autre part, sur le comportement de ses lois asymptotiques et ceux selon les régions de l'appartenance du paramètre inconnu φ et selon la taille de la série simulée, tout en commentant sur la vitesse de convergence. Pour étudier empiriquement la performance de cet estimateur et le comportement de sa distribution asymptotique, dans les trois régions à savoir de stabilité, de racine unitaire et d'explosibilité (instabilité explosive), nous avons considéré les deux valeurs 0.5 et 0.8 du paramètre φ appartenant à la région de stabilité $|\varphi| < 1$ et la valeur 1.2 pour le cas explosif et nous avons simulé les séries correspondantes, en prenant la valeur initiale égale à 0 ensuite en la prenant égale à 1 et l'écart-type de l'innovation égale à 1.

2.4.1 Cas où $X_0 = 0$

Commençons d'abord, par les résultats de simulation concernant la consistance, au sens de la convergence en probabilité et presque sûrement, de l'estimateur $\widehat{\varphi}_n$ et de l'estimateur de son écart-type $\widehat{\sigma}_{\widehat{\varphi}_n}$ (tout en le comparant avec la valeur théorique). Les valeurs moyennes, basées sur 1000 réplifications, de l'estimateur du MCO , en variant la taille de la

série de $n = 50$ à $n = 500$, avec un pas de 50, ainsi que les estimations de son écart-type $\widehat{\sigma}_{\widehat{\varphi}_n}$ sont portées sur la *Table 2.4.1* ci-dessous.

Pour le cas stationnaire, nous constatons de la *Table 2.4.1* que les résultats des simulations consolident les résultats théoriques concernant la consistance (au sens des modes de convergence) de l'estimateur $\widehat{\varphi}_n$, en particulier le résultat donné par l'expression (2.3.9a) assurant la convergence presque sûre. En effet, la valeur estimée de φ approche la vraie valeur 0.5 (respectivement 0.8) au fur et à mesure que la taille de la série augmente, plus précisément, elle passe de la valeur estimée 0.4821 (avec un écart-type estimé 0.1285), (respectivement 0.7728, avec un écart-type estimé 0.0959) pour $n = 50$ à 0.4973 (avec un écart-type estimé 0.0389), (respectivement 0.7954, avec un écart-type estimé 0.0268) pour $n = 500$.

En conclusion, les simulations montrent empiriquement que l'estimateur du *MCO*, pour les deux modèles correspondant à $\varphi = 0.5$ et $\varphi = 0.8$ appartenant à la région de stabilité, est doté de la propriété de consistance de plus la vitesse de convergence est importante ce qui renforce les résultats théoriques étudiés.

Aussi la *Table 2.4.1* montre clairement que cette propriété de consistance est bien rencontrée aussi dans les cas d'instabilité à savoir l'instabilité à racine unitaire et l'instabilité explosif correspondant, respectivement, à $\varphi = 1$ et $\varphi = 1.2$. De plus, nous pouvons facilement constater que la vitesse de convergence, en particulier pour le cas $\varphi = 1.2$, est plus importante que dans les deux cas stationnaires i.e., $\varphi = 0.5$ et $\varphi = 0.8$.

Itération = 1000 $\sigma^2 = 1$ $X_0 = 0$

Taille de l'échantillon	φ	$\hat{\varphi}$	<i>Std</i>
$n = 50$	0.5	0.4821	0.1285
	0.8	0.7728	0.0959
	1	0.9659	0.0607
	1.2	1.1997	0.0061
$n = 100$	0.5	0.4899	0.0912
	0.8	0.7821	0.0632
	1	0.9806	0.0307
	1.2	1.2000	1.7713×10^{-7}
$n = 150$	0.5	0.4852	0.0713
	0.8	0.7891	0.0524
	1	0.9887	0.0213
	1.2	1.2000	1.4702×10^{-9}
$n = 200$	0.5	0.4948	0.0618
	0.8	0.7932	0.0429
	1	0.9906	0.0159
	1.2	1.2000	2.2705×10^{-14}
$n = 250$	0.5	0.4977	0.0564
	0.8	0.7932	0.0397
	1	0.9929	0.0123
	1.2	1.2000	2.2658×10^{-14}
$n = 300$	0.5	0.4940	0.0489
	0.8	0.7981	0.0347
	1	0.9942	0.0101
	1.2	1.2000	2.2669×10^{-14}
$n = 350$	0.5	0.4964	0.0476
	0.8	0.7952	0.0335
	1	0.9949	0.0087
	1.2	1.2000	2.2660×10^{-14}
$n = 400$	0.5	0.4955	0.0447
	0.8	0.7967	0.0307
	1	0.9955	0.0079
	1.2	1.2000	2.2660×10^{-14}
$n = 450$	0.5	0.4979	0.0400
	0.8	0.7980	0.0287
	1	0.9958	0.0073
	1.2	1.2000	2.2660×10^{-14}
$n = 500$	0.5	0.4973	0.0389
	0.8	0.7954	0.0268
	1	0.9963	0.0065
	1.2	1.2000	2.2665×10^{-14}

Table 2.4.1

Comportement des distributions asymptotiques

Nous avons vu que les résultats de simulation ont bien confirmé la consistance, au sens de la convergence en probabilité et presque sûrement, de l'estimateur *MCO*, $\hat{\varphi}_n$, du paramètre inconnu φ , d'un modèle autorégressif d'ordre un, et ceci dans les trois cas à savoir le cas de stabilité, d'instabilité à racine unitaire et d'instabilité explosif. Dans ce paragraphe, nous verrons que l'étude de simulation confirme empiriquement aussi les résultats théoriques établis concernant le comportement de la distribution asymptotique de cet estimateur dans les trois cas précédents. En effet, nous verrons que la loi empirique de l'estimateur $\hat{\varphi}_n$ converge, quand la taille n tend vers l'infini vers une loi normale pour le cas stable, la loi de Wiener dans le cas d'instabilité à racine unitaire et enfin une loi de Cauchy dans le cas d'instabilité explosif.

a) Distributions asymptotiques dans le cas de stabilité, pour $\varphi = 0.5$ et 0.8

Les graphes des histogrammes et les densités de probabilité ajustées aux valeurs de l'estimateur du *MCO*, $\hat{\varphi}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 50, 100, 150, 200, 250, 300, 350, 400, 450, 500$, générées d'un modèle autorégressif d'ordre un, correspondant aux valeurs appartenant à la région de stabilité $\varphi = 0.5, 0.8$ et où la variance de l'innovation est considérée connue égale à l'unité ($\sigma^2 = 1$) et pour la valeur initiale nulle ($X_0 = 0$) sont donnés dans les deux *Table 2.4.2* et *Table 2.4.2 (suite)* ci-dessous.

De ces différents graphes, nous constatons que les densités de probabilité de la variable aléatoire $\sqrt{n}(\hat{\varphi}_n - \varphi)$ approche, au fur et à mesure que n augmente, d'une densité de probabilité d'une variable aléatoire normalement distribuée. En effet, les tests d'ajustement ont accepté l'hypothèse de normalité de moyenne nulle et de variance 0.75 et 0.36 correspondant, respectivement aux valeurs $\varphi = 0.5, 0.8$, ce qui confirme empiriquement le résultat concernant la convergence en loi de l'estimateur du *MCO*, quand n tend vers l'infini, vers une loi normale et ceci quand le paramètre appartient à la région de stabilité $|\varphi| < 1$.

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 50$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 100$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 150$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 200$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 250$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

Table 2.4.2

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 300$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 350$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 400$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 450$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

$$\sigma^2 = 1, X_0 = 0 \text{ et } n = 500$$

$$\varphi = 0.5$$

$$\varphi = 0.8$$

Table 2.4.2 (suite)

b) Distributions asymptotiques dans le cas d'instabilité (de racine unitaire $\varphi = 1$)

Les graphes des histogrammes des valeurs de l'estimateur du MCO, $\widehat{\varphi}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 50, 100, 150, 200, 250, 300, 350, 400, 450, 500$, générées d'un modèle autorégressif d'ordre un, correspondant à la valeur $\varphi = 1$ appartenant au frontière de stabilité et où la variance de l'innovation est considérée connue égale à l'unité ($\sigma^2 = 1$) et pour la valeur initiale nulle ($X_0 = 0$) sont donnés dans la *Table 2.4.3* sous-dessous. De ces différents graphes, nous constatons que les densités de probabilité de la variable aléatoire $n(\widehat{\varphi}_n - 1)$ sont clairement asymétriques et donc elles ne peuvent être de Gauss, ce qui consolide le résultat théorique consistant à ce que l'estimateur ne converge pas en loi vers une loi normale mais plutôt vers la loi $\frac{1}{2} ((W(1))^2 - 1) / \int_0^1 (W(r))^2 dr$.

$n = 50$	$n = 100$
$n = 150$	$n = 200$
$N = 250$	$N = 300$
$n = 350$	$n = 400$
$n = 450$	$n = 500$

Table 2.4.3

c) Distributions asymptotiques dans le cas d'instabilité explosive $\varphi = 1.2$

Les graphes des histogrammes de la variable aléatoire $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\hat{\varphi}_{1,n} - 1)$, correspondant aux différentes séries temporelles artificielles de taille $n = 50, 100, 150, 200, 250, 300, 350, 400, 450, 500$, générées d'un modèle autorégressif d'ordre un, correspondant à la valeur $\varphi = 1.2$ appartenant à la région explosive et où la variance de l'innovation est considérée connue égale à l'unité ($\sigma^2 = 1$) et pour la valeur initiale nulle ($X_0 = 0$) sont donnés dans la *Table 2.4.4* ci-dessous. De ces différents graphes, nous constatons que les densités de probabilité de la variable aléatoire $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\hat{\varphi}_{1,n} - 1)$ sont de type de Cauchy.

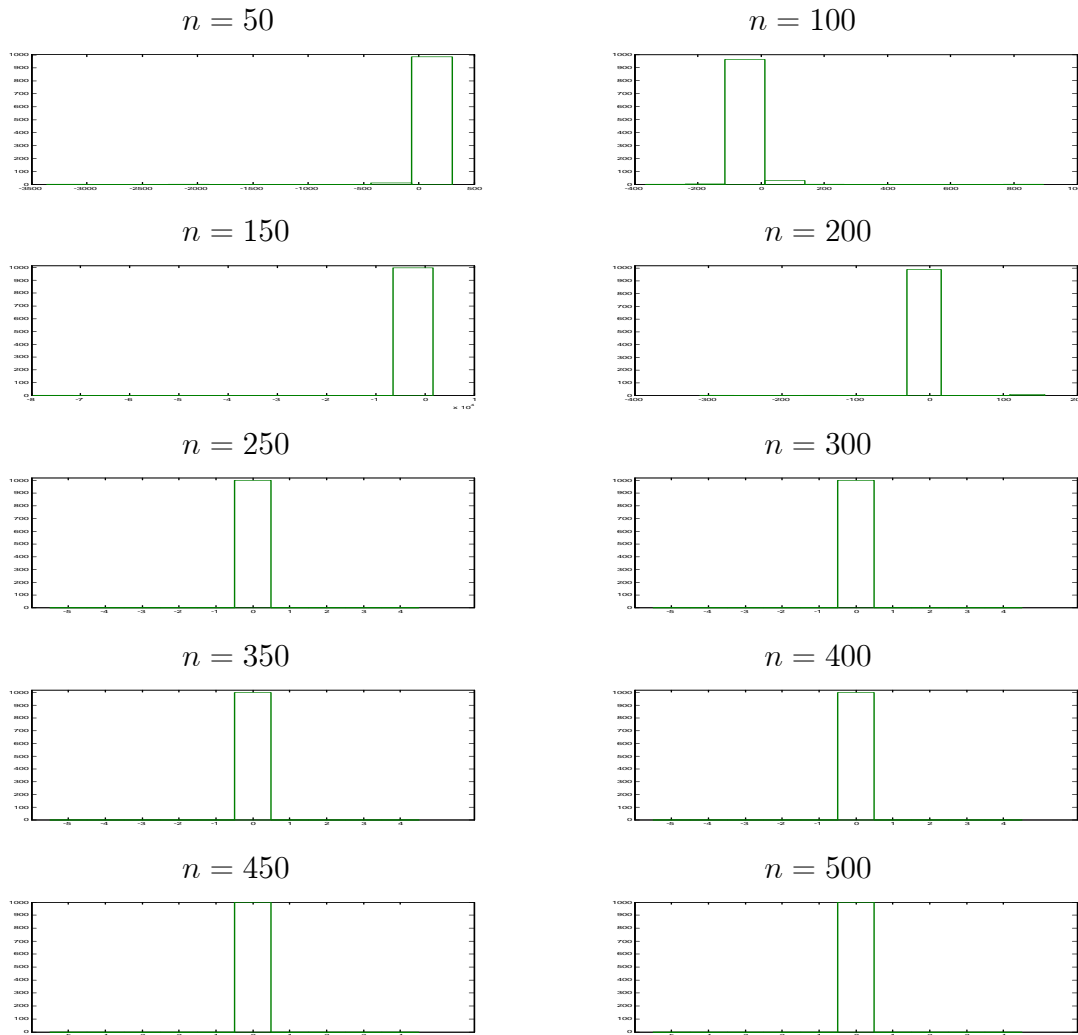
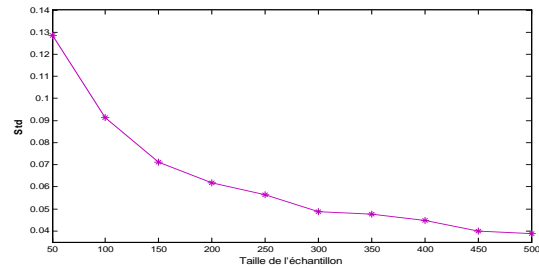
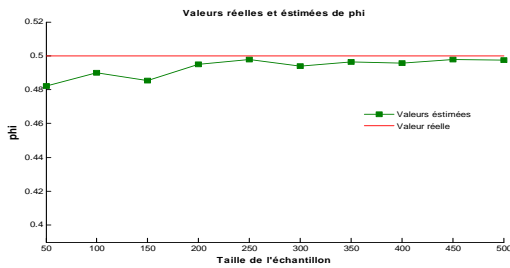


Table 2.4.4

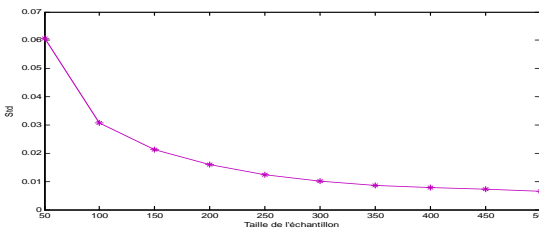
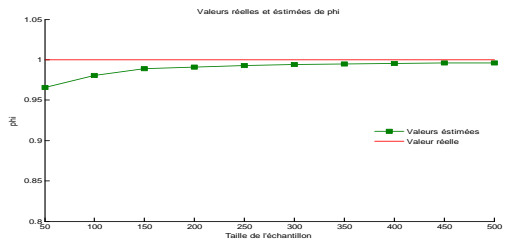
d) Comportement de l'estimateur $\hat{\varphi}_n$ et l'estimateur de son écart-type $\hat{\sigma}_{\hat{\varphi}_n}$ dans les trois régions

d. 1) Dans la région de stabilité ($\varphi = 0.5, 0.8$) et sur la frontière ($\varphi = 1$)

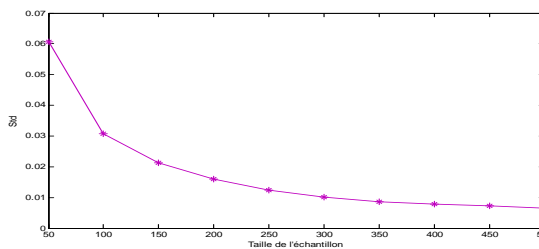
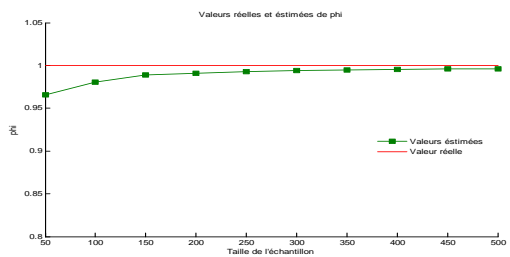
Les graphes de la Table 2.4.5 sous-dessous l'évolution de l'estimateur, $\hat{\varphi}_n$, des moindres carrés du paramètre ainsi que l'évolution de l'estimateur de son écart-type $\hat{\sigma}_{\hat{\varphi}_n}$, pour les deux valeurs $\varphi = 0.5$ et 0.8 , de la région de stabilité ainsi que pour $\varphi = 1$, en fonction des séries temporelles simulées et commentées. A partir des graphes de la gauche de la table, nous constatons clairement que les valeurs estimées du paramètre approche rapidement la vraie valeur du paramètre et ceci dans les deux cas à savoir le cas de stabilité (0.5 et 0.8) et le cas de racine unitaire $\varphi = 1$. Pareillement les estimations de l'écart-type de cet estimateur approche la valeur zéro au fur et à mesure que n augmente. Ces constatations consolident les résultats théoriques et sont étudiées empiriquement dans le paragraphe précédent a).



$\varphi = 0.5$



$\varphi = 0.8$



$\varphi = 1$

Table 2.4.5

d. 2) **A l'extérieur de la région de stabilité** ($\varphi = 1.2$)

Finalement, nous constatons à partir des deux graphes suivants que dans le cas instable explosif l'estimateur converge très rapidement vers la vraie valeur et ceci est reflété encore par le fait que son écart-type devient presque zéro pour des tailles dépassant 100.

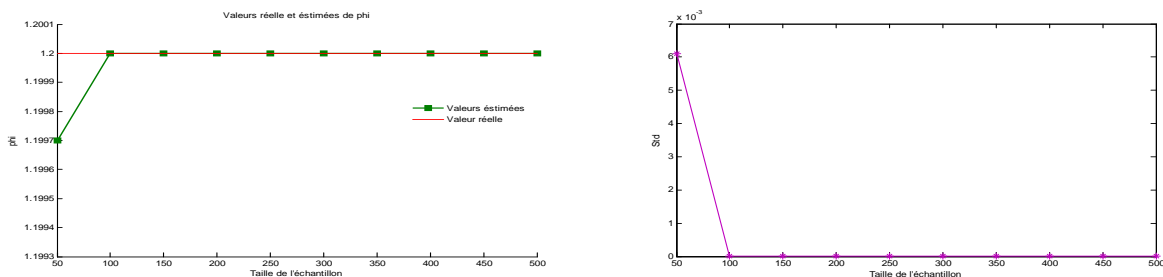


Table 2.4.6

2.4.2 Cas où $X_0 = 1$

Dans ce qui suit, nous présentons les résultats des simulations faites en prenant la valeur initiale égale à 1 et l'écart type de l'innovation égale à 1. Les valeurs moyennes, basées sur 1000 répliquions, de l'estimateur du *MCO*, en variant la taille de la série de $n = 50$ à $n = 500$, avec un pas de 50, ainsi que les estimations de son écart-type $\hat{\sigma}_{\hat{\varphi}_n}$ sont portées sur la *Table 2.4.3* ci-dessous.

Dans le cas stationnaire, nous constatons de la *Table 2.4.7* que la valeur estimée de φ approche de la vraie valeur 0.5 (respectivement 0.8) au fur et à mesure que la taille de la série augmente, plus précisément, elle passe de la valeur estimée 0.4897 (avec un écart-type estimé 0.1178), (respectivement 0.7693, avec un écart-type estimé 0.0970) pour $n = 50$ à 0.4979 (avec un écart-type estimé 0.0386), (respectivement 0.7983, avec un écart-type estimé 0.0270) pour $n = 500$. Ce qui consolide les résultats théoriques concernant la consistance (au sens des modes de convergence) de l'estimateur $\hat{\varphi}_n$, en particulier le résultat donnée par l'expression (2.3.9a) assurant la convergence presque sûre.

Nous concluons donc que l'estimateur du *MCO*, pour les deux modèles correspondant à $\varphi = 0.5$ et $\varphi = 0.8$ appartenant à la région de stabilité, est muni de la propriété de consistance de plus la vitesse de convergence est importante ce qui renforce les résultats théoriques étudié.

Itération = 1000 $\sigma^2 = 1$ $X_0 = 1$

Taille de l'échantillon	φ	$\hat{\varphi}$	<i>Std</i>
$n = 50$	0.5	0.4897	0.1178
	0.8	0.7693	0.0970
	1	0.9627	0.0633
	1.2	1.1999	0.0010
$n = 100$	0.5	0.4976	0.0833
	0.8	0.7875	0.0648
	1	0.9822	0.0308
	1.2	1.2000	1.9217×10^{-7}
$n = 150$	0.5	0.4982	0.0692
	0.8	0.7887	0.0502
	1	0.9885	0.0204
	1.2	1.2000	8.3131×10^{-12}
$n = 200$	0.5	0.4980	0.0601
	0.8	0.7919	0.0436
	1	0.9920	0.0142
	1.2	1.2000	2.3088×10^{-14}
$n = 250$	0.5	0.4967	0.0547
	0.8	0.7920	0.0392
	1	0.9939	0.0115
	1.2	1.2000	2.2663×10^{-14}
$n = 300$	0.5	0.4989	0.0491
	0.8	0.7940	0.0343
	1	0.9946	0.0103
	1.2	1.2000	2.2657×10^{-14}
$n = 350$	0.5	0.4970	0.0454
	0.8	0.7953	0.0316
	1	0.9949	0.0087
	1.2	1.2000	2.2656×10^{-14}
$n = 400$	0.5	0.4973	0.0436
	0.8	0.7962	0.0306
	1	0.9960	0.0073
	1.2	1.2000	2.2663×10^{-14}
$n = 450$	0.5	0.4960	0.0422
	0.8	0.7969	0.0284
	1	0.9962	0.0064
	1.2	1.2000	2.2661×10^{-14}
$n = 500$	0.5	0.4979	0.0386
	0.8	0.7983	0.0270
	1	0.9965	0.0068
	1.2	1.2000	2.2664×10^{-14}

Table 2.4.7

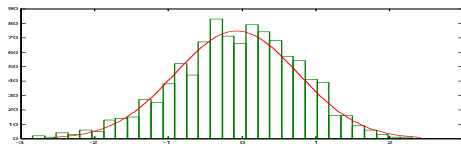
La Table 2.4.7 montre que la propriété de consistance est bien rencontrée aussi dans les cas d'instabilité à savoir l'instabilité à racine unitaire et l'instabilité explosif correspondant, respectivement, à $\varphi = 1$ et $\varphi = 1.2$. De plus, nous pouvons facilement constater que la vitesse de convergence, en particulier pour le cas $\varphi = 1.2$, est plus importante que dans les deux cas stationnaires i.e., $\varphi = 0.5$ et $\varphi = 0.8$.

Comportement des distributions asymptotiques

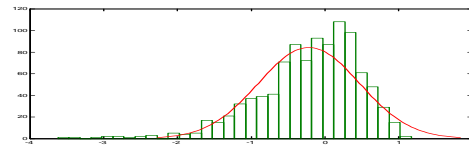
Les densités de probabilité de la variable aléatoire $\sqrt{n}(\widehat{\varphi}_n - \varphi)$ approche, au fur et à mesure que n augmente, d'une densité de probabilité d'une variable aléatoire normalement distribuée. En effet, les tests d'ajustement ont accepté l'hypothèse de normalité de moyenne nulle et de variance 0.75 et 0.36 correspondant, respectivement aux valeurs $\varphi = 0.5, 0.8$. ce qui confirme empiriquement le résultat concernant la convergence en loi de l'estimateur du MCO, quand n tend vers l'infini, vers une loi normale et ceci quand le paramètre appartient à la région de stabilité $|\varphi| < 1$. Dans le cas de racine unitaire ($\varphi = 1$), les densités de probabilité de la variable aléatoire $n(\widehat{\varphi}_n - 1)$ sont asymétriques et donc elles ne peuvent être de Gauss, ce qui consolide le résultat théorique consistant à ce que l'estimateur ne converge pas en loi vers une loi normale mais plutôt vers la loi $\frac{1}{2} ((W(1))^2 - 1) / \int_0^1 (W(r))^2 dr$. Pour le cas explosif ($|\varphi| > 1$), les densités de probabilité de la variable aléatoire $(\varphi_1^2 - 1)^{-1} \varphi_1^n (\widehat{\varphi}_{1,n} - 1)$ sont celle de $W_1 W_2^{-1}$, où W_1 et W_2 sont des variables aléatoires indépendantes telles que :

$$(W_1, W_2) = \left(\sum_{i=1}^{\infty} \varphi_1^{-i} \varepsilon_i, \delta_0 \sum_{j=1}^{\infty} \varphi_1^{-j} \varepsilon_j \right).$$

$\sigma^2 = 1, X_0 = 1$ et $n = 50$

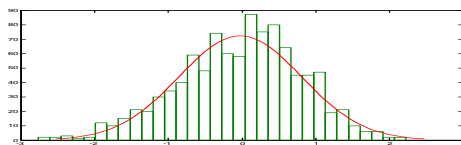


$\varphi = 0.5$

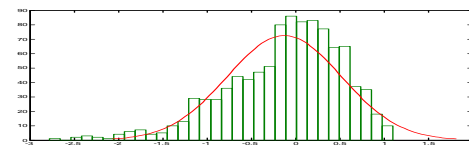


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 100$

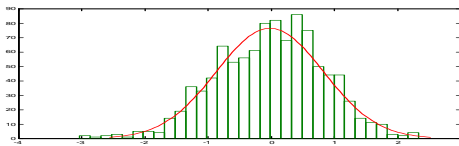


$\varphi = 0.5$

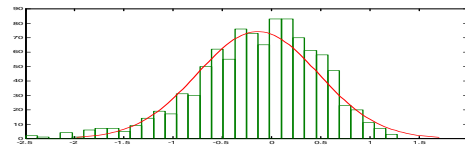


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 150$

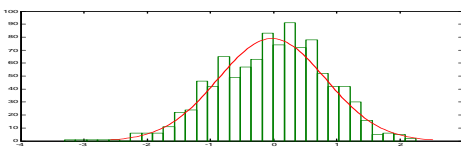


$\varphi = 0.5$

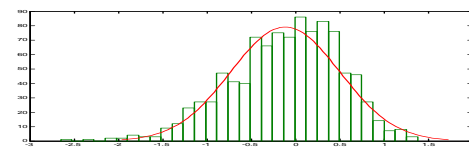


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 200$

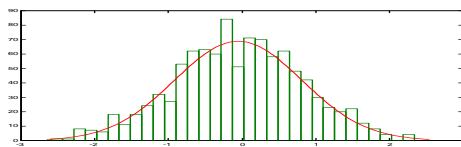


$\varphi = 0.5$

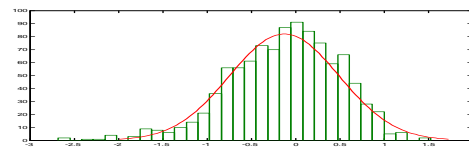


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 250$



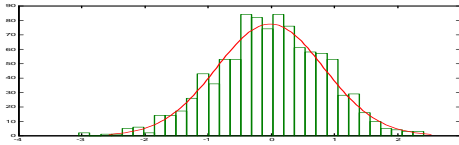
$\varphi = 0.5$



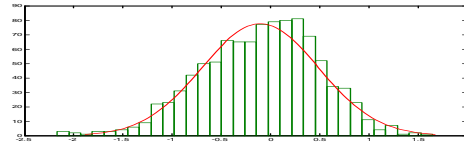
$\varphi = 0.8$

Distributions asymptotiques dans le cas de stabilité, pour $\varphi = 0.5$ et 0.8

$\sigma^2 = 1, X_0 = 1$ et $n = 300$



$\varphi = 0.5$

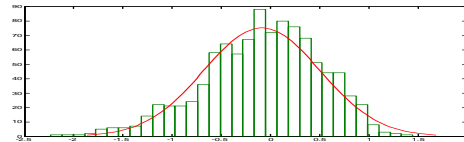


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 350$

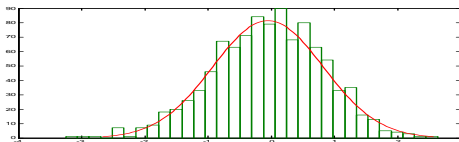


$\varphi = 0.5$

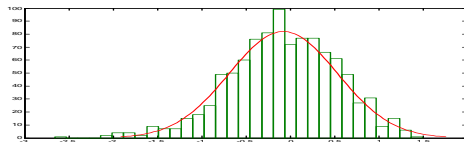


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 400$

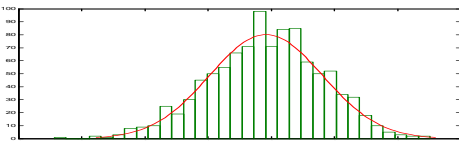


$\varphi = 0.5$

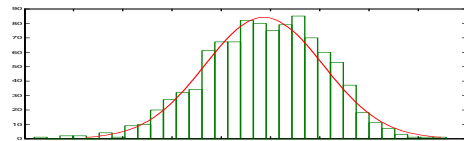


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 450$

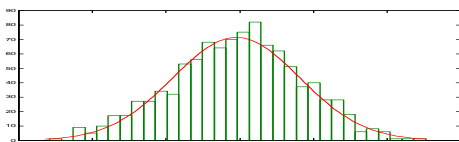


$\varphi = 0.5$

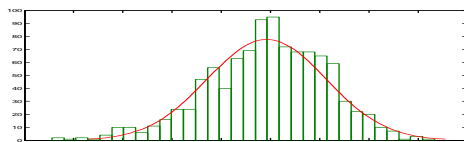


$\varphi = 0.8$

$\sigma^2 = 1, X_0 = 1$ et $n = 500$

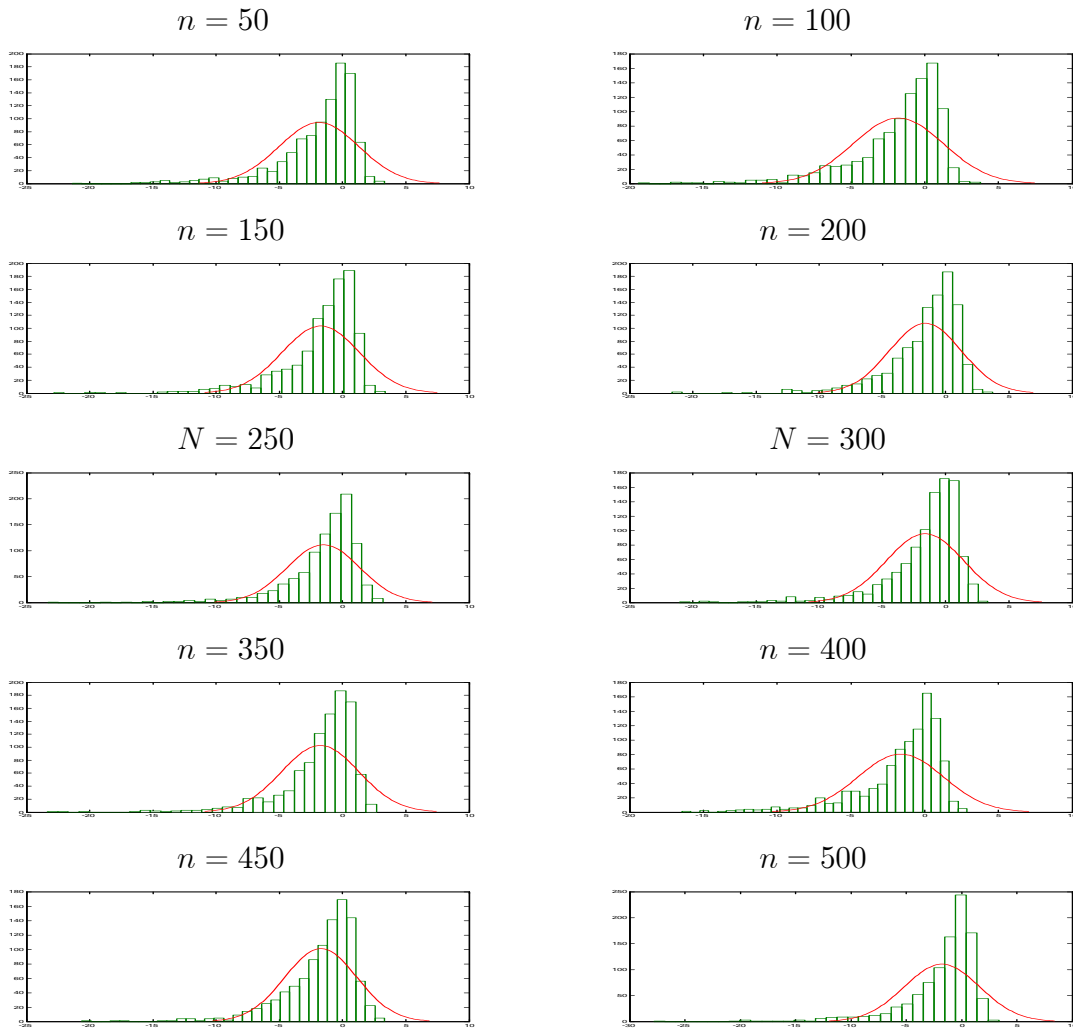


$\varphi = 0.5$

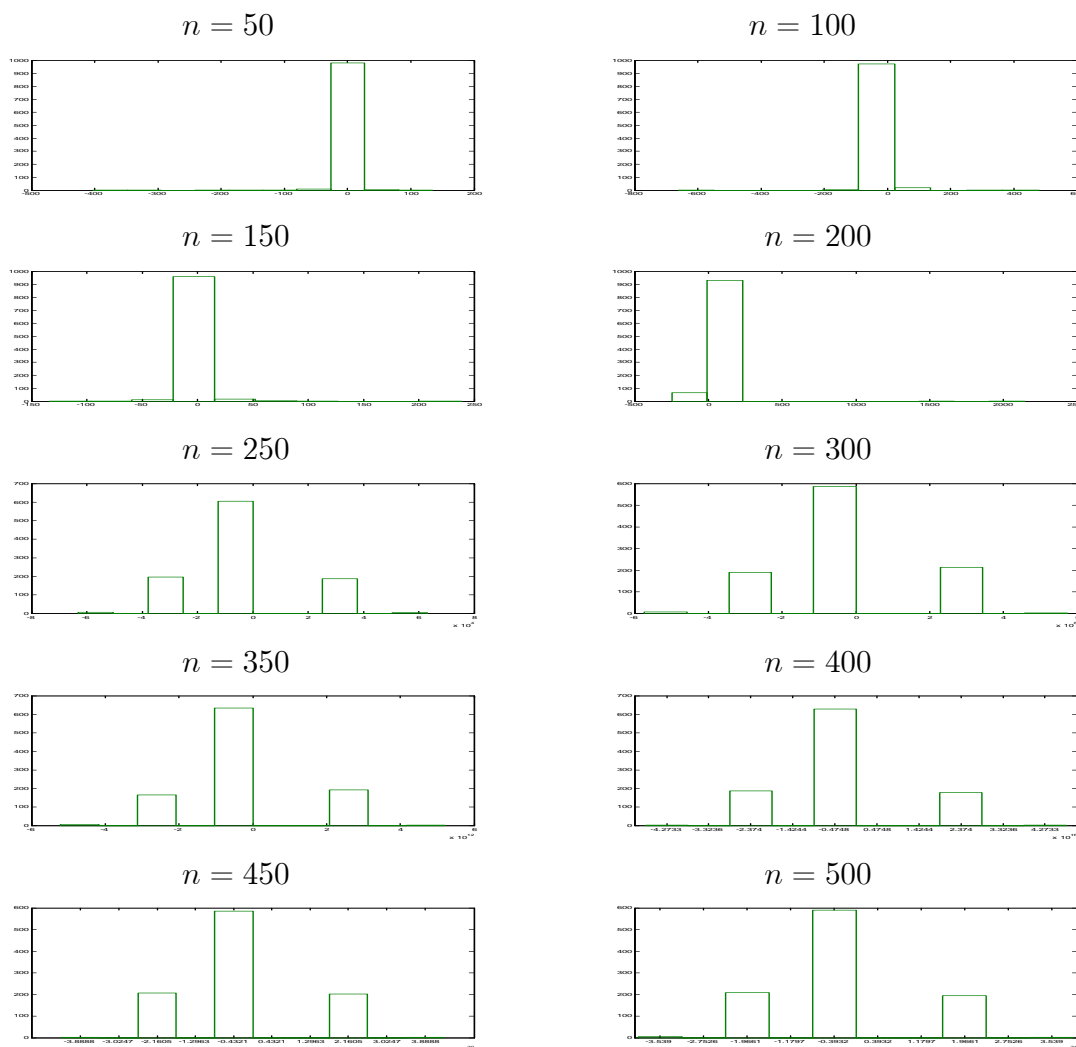


$\varphi = 0.8$

Distributions asymptotiques dans le cas de stabilité, pour $\varphi = 0.5$ et 0.8 (suite)



Distributions asymptotiques dans le cas d'instabilité (de racine unitaire $\varphi = 1$)



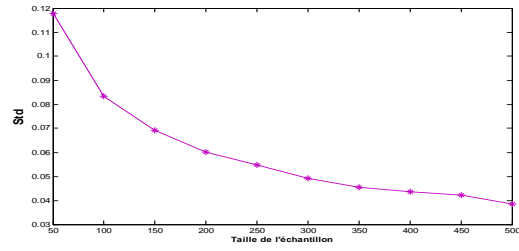
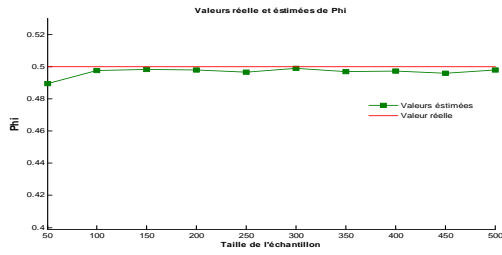
Distributions asymptotiques dans le cas explosif ($\varphi = 1.2$)

Comportement de l'estimateur $\hat{\varphi}_n$ et l'estimateur de son écart-type $\hat{\sigma}_{\hat{\varphi}_n}$ dans les trois régions

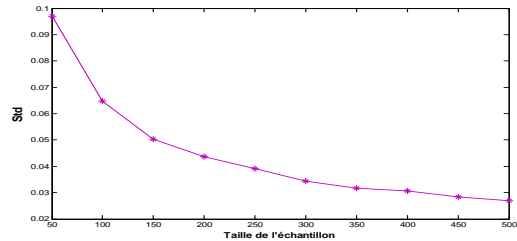
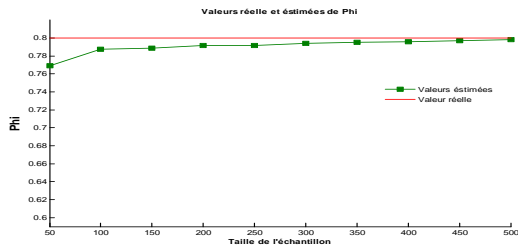
– Dans la région de stabilité ($\varphi = 0.5, 0.8$) et sur la frontière ($\varphi = 1$)

Les graphes de la *Table 2.4.8* ci-dessous l'évolution de l'estimateur, $\hat{\varphi}_n$, des moindres carrés du paramètre ainsi que l'évolution de l'estimateur de son écart-type $\hat{\sigma}_{\hat{\varphi}_n}$, pour les deux valeurs $\varphi = 0.5$ et 0.8 , de la région de stabilité ainsi que pour $\varphi = 1$, en fonction des séries temporelles simulées et commentées. A partir des graphes de la gauche de la table, nous constatons clairement que les valeurs estimées du paramètre approche rapidement la vraie valeur du paramètre et ceci dans les deux cas à savoir le cas de stabilité (0.5 et 0.8) et le

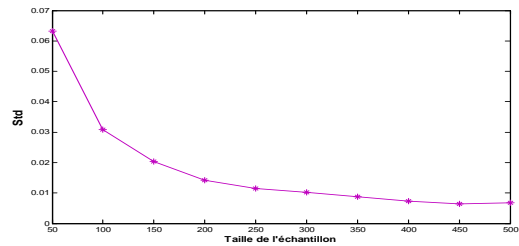
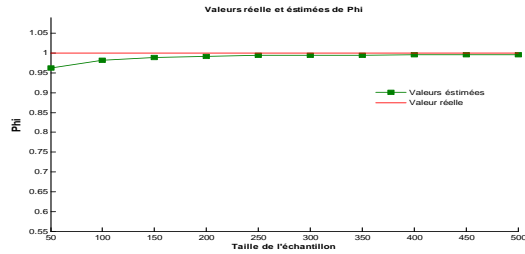
cas de racine unitaire $\varphi = 1$. Pareillement les estimations de l'écart-type de cet estimateur approche la valeur zéro au fur et à mesure que n augmente.



$\varphi = 0.5$



$\varphi = 0.8$



$\varphi = 1$

Table 2.4.8

– A l'extérieur de la région de stabilité ($\varphi = 1.2$)

Finalement, nous constatons à partir des deux graphes suivants que dans le cas instable explosif l'estimateur converge très rapidement vers la vraie valeur et ceci est reflété encore par le fait que son écart-type devient presque zéro pour des tailles dépassant 100.

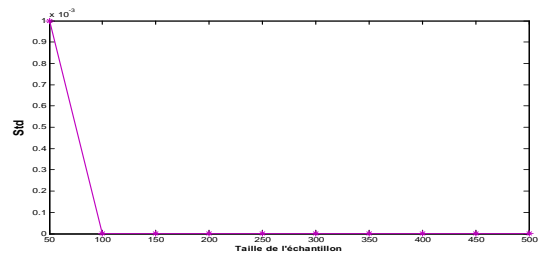
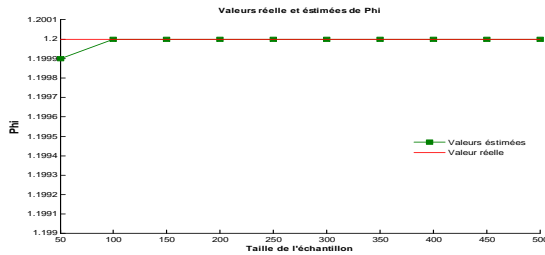


Table 2.4.9

Chapitre 3

Instabilité et Propriétés Asymptotiques des estimateur des MCO d'un $PAR_S(1)$

3.1 Introduction

Nous nous intéressons dans ce chapitre à l'étude des propriétés asymptotiques des estimateurs des MCO des paramètres d'un modèle autorégressif périodique d'ordre un ($PAR_S(1)$).

3.2 Processus périodiquement corrélé

Soit $\{X_t, t \in \mathbb{Z}\}$ un processus stochastique du second ordre, de moyenne et de fonction d'autocovariance définies par :

$$\begin{aligned}\mu_{t+s\tau} &= \mathbb{E}(\mu_t), & \forall t \in \mathbb{Z}, \\ \gamma(t, h) &= cov(X_t, X_h) = \mathbb{E}((X_t - \mu_t)(X_h - \mu_h)), & \forall t, h \in \mathbb{Z}.\end{aligned}$$

Dans le cas où la moyenne μ_t est constante, autrement dit indépendante de t , et la fonction d'autocovariance ne dépend que du retard $|t - h|$, on dit que $\{X_t\}$ est un processus stationnaire du second ordre ou encore faiblement stationnaire. Diverses sont les classes de processus non stationnaires, parmi lesquelles nous citons la classe de processus dont les fonctions d'autocovariance sont des fonctions périodiques dans le temps. La classe des processus périodiquement corrélés a été introduite, dans la littérature de l'étude des processus aléatoires, par *Gladyshev* (1961).

Définition 3.2.1

Un processus du second ordre X_t est dit *périodiquement corrélé* ou *faiblement périodique* s'il existe un entier positif S tel que pour tout $t, h, \tau \in \mathbb{Z}$ on ait :

$$\mu_{t+S\tau} = \mu_t \quad \forall t \in \mathbb{Z},$$

et

$$\gamma(t + S\tau, h + S\tau) = \gamma(t, h).$$

L'entier positif S qui satisfait les deux conditions ci-dessus est appelé *période* du processus périodiquement corrélé. De cette définition nous constatons qu'un processus stationnaire du second ordre est un cas particulier du processus périodiquement corrélé : c'est un processus périodiquement corrélé de période 1 (puisque la moyenne est constante et l'autocovariance dépend du retard).

Propriétés des autocovariances d'un processus S -périodique

Considérons un processus X_t , centré, périodiquement corrélé de période S . Soient, pour tout $t \in \mathbb{Z}$, l'entier positif s , $s = 1, 2, \dots, S$ et l'entier $\tau \in \mathbb{Z}$ tel que $t = s + S\tau$. Notons $\gamma_h^{(s)}$ la fonction d'autocovariance, d'ordre h (h est un entier positif) à l'instant $t = s + S\tau$, du processus S -périodique X_t , c'est-à-dire

$$\gamma_h^{(s)} = \gamma(t, t - h) = \gamma(s, s - h), \quad s = 1, 2, \dots, S, \quad h \in \mathbb{N}.$$

Il est facile de vérifier la propriété importante suivante.

Propriété 3.2.1

La fonction d'autocovariance d'un processus S -périodique X_t , vérifie les deux propriétés suivantes :

$$\gamma_h^{(s+kS)} = \gamma_h^{(s)} \quad \text{et} \quad \gamma_{-h}^{(s)} = \gamma_h^{(s+h)}.$$

Processus périodiquement corrélé et processus multivarié stationnaire

Soit $\underline{Y}(\tau)' = (Y_1(\tau), Y_2(\tau), \dots, Y_S(\tau))$ un processus S -varié (où S est la période du processus S -périodique X_t) dont la i ème composante

$$Y_s(\tau) = X_{s+S\tau}, \quad s = 1, 2, 3, \dots, S \quad \text{et} \quad \tau \in \mathbb{Z}. \quad (3.2.1)$$

Soit $\Gamma(\tau_1, \tau_2)$ la matrice des covariances croisées de $\underline{Y}(\tau_1)$ et $\underline{Y}(\tau_2)$ dont les éléments sont donnés par :

$$\Gamma_{ij}(\tau_1, \tau_2) = Cov\left(Y_i(\tau_1), Y_j(\tau_2)\right)$$

et soient t_1 et t_2 tels que

$$t_1 = i + S\tau_1 \text{ et } t_2 = j + S\tau_2, \forall (\tau_1, \tau_2) \in \mathbb{Z} \times \mathbb{Z}.$$

Les éléments $\Gamma_{ij}(\tau_1, \tau_2)$ peuvent être donnés en fonction des autocovariances $\gamma_h^{(s)}$ du processus S -périodique par la propriété suivante.

Propriété 3.2.2

La éléments, $\Gamma_{ij}(\tau_1, \tau_2)$, $i, j = 1, 2, 3, \dots, S$, de la matrice $\Gamma(\tau_1, \tau_2)$ sont donnés par :

$$\Gamma_{ij}(\tau_1, \tau_2) = \gamma_{i-j+S(\tau_1-\tau_2)}^{(i)}, i, j = 1, 2, 3, \dots, S \text{ et } (\tau_1, \tau_2) \in \mathbb{Z} \times \mathbb{Z}.$$

De plus, si le processus S -varié est stationnaire, alors sa matrice des covariances croisées de délai k est donnée par :

$$\begin{aligned} \text{où } \Gamma(k) &= \left(\Gamma_{ij}(k)\right) = \left(\Gamma_{ij}(\tau, \tau - k)\right), i, j = 1, 2, \dots, S, k \in \mathbb{N}, \\ \Gamma_{ij}(k) &= \gamma_{i-j+kS}^{(i)}, i, j = 1, 2, \dots, S, k \in \mathbb{N}. \end{aligned}$$

Le théorème suivant établit le lien existant entre un processus périodiquement corrélé X_t et le processus multivarié de dimension S lui est associé, $\underline{Y}(\tau)$, donné par (M.1).

Théorème 3.2.1 (Gladyshev (1961))

Un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ est périodiquement corrélé de Période S si, et seulement si, le processus multivarié correspondant $\{\underline{Y}(\tau), \tau \in \mathbb{Z}\}$, défini par :

$$\begin{aligned} \underline{Y}(\tau)' &= (Y_1(\tau), Y_2(\tau), \dots, Y_S(\tau)), \forall \tau \in \mathbb{Z}, \\ \text{ou encore } \underline{Y}(\tau)' &= (X_{1+S\tau}, X_{2+S\tau}, \dots, X_{S+S\tau}) \end{aligned}$$

est **faiblement stationnaire**.

Ce théorème de Gladyshev a été utilisé dès les années soixante dix, en particulier par Tiao et Grupe (1979), pour construire et développer le modèle autorégressif moyenne mobile multivarié correspondant au modèle autorégressif moyenne mobile S -périodique $ARMA_S(p_t, q_t)$. Ces deux auteurs ont montré les conséquences de modéliser un processus périodiquement corrélé par un modèle saisonnier de Box et Jenkins. Troutman (1979) a exploité, de son côté,

cette approche Gladyshev, pour étudier plusieurs propriétés d'un processus périodiquement corrélé, en particulier, les conditions d'indéterminabilité pure (causalité), les propriétés de la structure de la fonction d'autocovariance.

3.2.1 Modèles autorégressifs moyennes mobiles périodiques

Définition 3.2.2

Un processus périodique de période S est dit admettant une représentation autorégressif moyenne mobile d'ordre (p, q) périodique de période S (noté $PARMA_S(p, q)$), s'il est solution de l'équation aux différences stochastiques suivante :

$$X_t - \sum_{i=1}^p \varphi_{t,i} X_{t-i} = \varepsilon_t - \sum_{j=1}^q \theta_{t,j} \varepsilon_{t-j}, \quad t \in \mathbb{Z}, \quad (3.2.2)$$

où $\{\varepsilon_t\}$ est un processus bruit blanc tel que $\mathbb{E}(\varepsilon_t) = 0$ et $\mathbb{E}(\varepsilon_t^2) = \sigma_t^2$

Les paramètres autorégressifs $\varphi_{t,i}$, moyennes mobiles $\theta_{t,j}$ et la variance σ_t^2 sont des fonctions périodiques du temps c'est à dire :

$$\begin{aligned} \varphi_{t+kS,i} &= \varphi_{t,i}, \\ \theta_{t+kS,j} &= \theta_{t,j}, \\ \sigma_{t+kS}^2 &= \sigma_t^2, \end{aligned} \quad \forall t \in \mathbb{Z} \text{ et } \forall k \in \mathbb{N}.$$

En notant L l'opérateur retard tel que $L^i X_t = X_{t-i}$ et $L^j \varepsilon_t = \varepsilon_{t-j}$, l'équation (3.2.2) s'écrira sous la forme suivante :

$$\Phi_t(L) X_t = \Theta_t(L) \varepsilon_t, \quad t \in \mathbb{Z},$$

où $\Phi_t(L) = 1 - \sum_{i=1}^p \varphi_{t,i} L^i$ et $\Theta_t(L) = 1 - \sum_{j=1}^q \theta_{t,j} L^j$.

Cas particuliers

Deux types de modèles importants dérivent de la classe de modèles linéaires périodiques, à savoir les modèles autorégressifs (purs) d'ordre p et à coefficients S -périodiques notés $PAR_S(p)$, et les modèles moyennes mobiles (purs) d'ordre q et à coefficients S -périodiques notés $PMA_S(q)$.

Définition 3.2.3

Un modèle particulier du modèle S -périodique donné par (3.2.2), correspondant au cas où $\varphi_{t,i} = 0$, pour tout $i = 1, \dots, p$ et $t \in \mathbb{Z}$, est dit modèle moyenne mobile S -périodique d'ordre q , noté $PMA_S(q)$, et est donné par l'équation aux différences stochastiques suivante :

$$X_t = \varepsilon_t - \sum_{j=1}^q \theta_{t,j} \varepsilon_{t-j}, \quad t \in \mathbb{Z} \quad (3.2.3)$$

où les paramètres $\theta_{t,j}$, $j = 1, \dots, q$ et $t \in \mathbb{Z}$ et la variance du bruit blanc σ_t^2 sont des fonctions périodiques, en t , de période S .

Définition 3.2.4

Un modèle particulier du modèle S -périodique donné par (3.2.2), correspondant au cas où $\theta_{t,j} = 0$, pour tout $j = 1, \dots, q$ et $t \in \mathbb{Z}$, est dit modèle autorégressif S -périodique d'ordre p , noté $PAR_S(p)$, et est donné par l'équation aux différences stochastiques suivante :

$$X_t - \sum_{i=1}^p \varphi_{t,i} X_{t-i} = \varepsilon_t, \quad t \in \mathbb{Z} \quad (3.2.4)$$

où les paramètres $\varphi_{t,i}$, $i = 1, \dots, p$ et $t \in \mathbb{Z}$ et la variance du bruit blanc σ_t^2 sont des fonctions périodiques, en t , de période S .

Parmi les propriétés importantes du modèle donné par (3.2.2) nous allons énoncer la propriété de causalité et d'inversibilité. Les définitions de ces deux propriétés sont similaires à celles que dans le cas des modèles à coefficients constants.

Définition 3.2.5

Le modèle (3.2.2) est dit causal s'il admet une solution de la forme :

$$X_t = \sum_{i=0}^{\infty} \alpha_{t,i} \varepsilon_{t-i}, \quad t \in \mathbb{Z} \quad (3.2.5)$$

où la série infinie est convergente, en moyenne quadratique. L'expression (3.2.5) est dite représentation de Wold-Cramér du processus $\{X_t, t \in \mathbb{Z}\}$.

Définition 3.2.6

Le modèle (3.2.2) est dit inversible s'il existent des fonctions $\beta_{t,i}$, telles que :

$$\varepsilon_t = \sum_{i=0}^{\infty} \beta_{t,i} X_{t-i}, \quad t \in \mathbb{Z} \quad (3.2.6)$$

où la série infinie est convergente, en moyenne quadratique. L'expression convergente (3.2.6) est dite formule de convolution.

Dans ce qui suit nous allons nous limiter à l'étude du modèle particulier $PAR_S(1)$.

3.2.2 Conditions de causalité et d'inversibilité

Pour l'étude de plusieurs propriétés théoriques des modèles linéaires autorégressifs moyennes mobiles $ARMA$ périodiques, en particulier les propriétés de causalité et d'inversibilité, nous pouvons suivre une des deux approches existant dans la littérature appelées, respectivement, *order-span-lumping* et *period-span-lumping*. La première approche consiste à écrire le modèle périodique scalaire sous la forme d'un modèle multivarié équivalent, de dimension égale à la période du processus périodiquement corrélé S , à matrices de paramètres constantes et par la suite le modèle multivarié classique équivalent sera exploité pour les propriétés sous-jacentes. Cette approche exploite directement le *Théorème 3.2.1 de Gladyshev* (1961) présenté précédemment.

La deuxième approche consiste à ramener un modèle périodique scalaire (ou multivarié), par l'introduction d'un processus multivarié, de dimension égale à l'ordre, dont les composantes sont données en termes du processus périodiquement corrélé sous l'étude à modèle multivarié périodique d'ordre 1. Le modèle équivalent ainsi obtenu par cette approche, introduite pour la première fois par Bentarzi et Hallin (1994), est exploité pour étudier beaucoup de propriétés des modèles périodiques (cf., Bentarzi et Hallin (1994), Bentarzi (1995), Ula et Smadi (1997) et Bentarzi (1998)).

Dans la suite de ce paragraphe, nous présentons brièvement ces deux approches pour l'obtention des conditions nécessaires et suffisantes de la causalité et de l'inversibilité des modèles linéaires périodiques.

Modèle VARMA associé un modèle ARMA périodique

Soit $\{X_t; t \in \mathbb{Z}\}$ un processus S -périodiquement corrélé satisfaisant le modèle autorégressif moyenne mobile périodique $ARMA_S(p_t, q_t)$ suivant :

$$X_t - \sum_{i=1}^{p_t} \varphi_{i,t} X_{t-i} = \varepsilon_t - \sum_{i=1}^{q_t} \theta_{i,t} \varepsilon_{t-i}, \quad t \in \mathbb{Z}, \quad (3.2.7)$$

et considérons les vecteurs aléatoires de dimension S suivant :

$$\underline{Y}(\tau)' = (Y_1(\tau), Y_2(\tau), \dots, Y_S(\tau)), \quad \text{avec } Y_i(\tau) = X_{i+S\tau}, \quad i = 1, 2, \dots, S, \quad \forall \tau \in \mathbb{Z},$$

et
$$\underline{\eta}(\tau)' = (\eta_1(\tau), \eta_2(\tau), \dots, \eta_S(\tau)), \quad \text{avec } \eta_i(\tau) = \varepsilon_{i+S\tau}, \quad i = 1, 2, \dots, S, \quad \forall \tau \in \mathbb{Z},$$

En utilisant les notations et les définitions précédentes on peut construire le modèle autorégressif moyenne mobile de dimension S , VARMA, associé au modèle S -périodique M2.

Théorème 3.2.2

Le modèle S -varié varié autorégressif moyenne mobile, VARMA, associé au modèle univarié autorégressif moyenne mobile d'ordre (p_t, q_t) , S -périodique $ARMA_S(p_t, q_t)$, est donné par :

$$\begin{aligned} \Phi(0) \underline{Y}(\tau) + \Phi(1) \underline{Y}(\tau - 1) + \dots + \Phi(P) \underline{Y}(\tau - P) \\ = \Theta(0) \underline{\eta}(\tau) + \Theta(1) \underline{\eta}(\tau - 1) + \dots + \Theta(Q) \underline{\eta}(\tau - Q), \quad \forall \tau \in \mathbb{Z}, \end{aligned}$$

où $\underline{\eta}(\tau)$ est un processus S -varié bruit blanc dont la matrice variance-covariance est diagonale d'éléments σ_i^2 , $i = 1, 2, \dots, S$. Les ordres (P, Q) sont donnés par :

$$P = \max_{1 \leq i \leq S} \left[\frac{p_i - i}{S} \right] + 1, \quad \text{et} \quad Q = \max_{1 \leq i \leq S} \left[\frac{q_i - i}{S} \right] + 1$$

et les matrices $\Phi(k)$, $k = 1, 2, \dots, P$ et $\Theta(r)$, $r = 1, 2, \dots, Q$ sont données par :

$$\Phi(0) = \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ \varphi_{2,1} & 1 & 0 & \dots & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \varphi_{i,i-1} & \varphi_{i,i-2} & \dots & 1 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \varphi_{S,S-1} & \varphi_{S,S-2} & \varphi_{S,S-3} & \dots & \varphi_{S,1} & 1 \end{pmatrix},$$

et

$$\Phi(k) = \begin{pmatrix} \varphi_{1,kS} & \varphi_{1,kS-1} & \varphi_{1,kS-2} & \cdots & \cdots & \varphi_{1,kS+1-S} \\ \varphi_{2,kS+1} & \varphi_{2,kS} & \varphi_{2,kS-1} & \cdots & \cdots & \varphi_{2,kS+2-d} \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \varphi_{i,kS+i-1} & \varphi_{i,kS+i-2} & \cdots & \cdots & \cdots & \varphi_{i,kS+i-d} \\ \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \varphi_{S,kS+S-1} & \varphi_{S,kS+S-2} & \varphi_{S,kS+S-3} & \cdots & \varphi_{S,kS+1} & \varphi_{S,kS} \end{pmatrix},$$

avec $i = 1, 2, \dots, S$ et $k = 1, 2, \dots, P$ et les matrices $\Theta(r)$, $r = 0, 1, 2, \dots, Q$ sont données par :

$$\Theta(0) = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ \theta_{2,1} & 1 & 0 & \cdots & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \theta_{i,i-1} & \theta_{i,i-2} & \cdots & 1 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \theta_{S,S-1} & \theta_{S,S-2} & \theta_{S,S-3} & \cdots & \theta_{S,1} & 1 \end{pmatrix}$$

et les matrices $\Theta(r)$, $r = 0, 1, 2, \dots, Q$ sont données par :

$$\Theta(r) = \begin{pmatrix} \theta_{1,kS} & \theta_{1,kS-1} & \theta_{1,kS-2} & \cdots & \cdots & \theta_{1,kS+1-S} \\ \theta_{2,kS+1} & \theta_{2,kS} & \theta_{2,kS-1} & \cdots & \cdots & \theta_{2,kS+2-d} \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \theta_{i,kS+i-1} & \theta_{i,kS+i-2} & \cdots & \cdots & \cdots & \theta_{i,kS+i-d} \\ \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \theta_{S,kS+S-1} & \theta_{S,kS+S-2} & \theta_{S,kS+S-3} & \cdots & \theta_{S,kS+1} & \theta_{S,kS} \end{pmatrix},$$

avec $i = 1, 2, \dots, S$ et $r = 1, 2, \dots, Q$.

Pour bien comprendre la démarche de cette construction, il semble très utile de l'appliquer sur un modèle périodique simple. Pour se faire, considérons un processus 4-périodique, satisfaisant un modèle autorégressif moyenne mobile, dont les ordres sont les suivant :

$$p_i = \begin{cases} 2, & i = 1 \\ 1, & i = 2, 4 \\ 3, & i = 3 \end{cases} \quad \text{et } q_i = 1, \text{ pour } i = 1, 2, 3 \text{ et } 4$$

alors, le modèle (4.1) peut s'écrire, dans ce cas particulier, pour $t = i + 4\tau$, sous la forme explicite suivante :

$$\begin{aligned} X_{1+4\tau} - \varphi_{1,1} X_{4+4(\tau-1)} - \varphi_{2,1} X_{3+4(\tau-1)} &= \varepsilon_{1+4\tau} - \theta_{1,1} \varepsilon_{4+4(\tau-1)}, \\ X_{2+4\tau} - \varphi_{1,2} X_{1+4\tau} &= \varepsilon_{2+4\tau} - \theta_{1,1} \varepsilon_{1+4\tau}, \\ X_{3+4\tau} - \varphi_{1,3} X_{2+4\tau} - \varphi_{2,3} X_{1+4\tau} - \varphi_{3,3} X_{4+4(\tau-1)} &= \varepsilon_{3+4\tau} - \theta_{1,1} \varepsilon_{2+4\tau}, \\ X_{4+4\tau} - \varphi_{1,4} X_{3+4\tau} &= \varepsilon_{4+4\tau} - \theta_{1,1} \varepsilon_{3+4\tau}, \end{aligned}$$

En écrivant les équations précédentes sous la forme matricielle, on obtient le modèle multivarié $ARMA(1, 1)$ associé est donc donné par :

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\varphi_{1,2} & 1 & 0 & 0 \\ -\varphi_{1,2} & -\varphi_{1,3} & 1 & 0 \\ 0 & 0 & -\varphi_{1,4} & 1 \end{pmatrix} \underline{Y}(\tau) - \begin{pmatrix} 0 & 0 & \varphi_{2,1} & \varphi_{1,1} \\ 0 & 0 & 0 & 0 \\ 0 & \varphi_{2,3} & \varphi_{1,3} & \varphi_{3,3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \underline{Y}(\tau - 1) \\ & = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\theta_{1,1} & 1 & 0 & 0 \\ 0 & -\theta_{1,1} & 1 & 0 \\ 0 & 0 & -\theta_{1,1} & 1 \end{pmatrix} \underline{\eta}(\tau) - \begin{pmatrix} 0 & 0 & 0 & \theta_{1,1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \underline{\eta}(\tau - 1). \end{aligned}$$

De la manière dont le modèle S -varié $ARMA(P, Q)$ est construit, à partir d'un modèle univarié autorégressif moyenne mobile S -périodique, il découle que beaucoup de propriétés théoriques de ce dernier modèle peuvent être déduites des propriétés liées au modèle S -varié.

2.1) Causalité et inversibilité. Approche de Gladyshev

En particulier, les conditions nécessaires et suffisantes le modèle périodique $ARMA_S(p_t, q_t)$ soit causal (le processus est périodiquement stationnaire) et inversible. Comme les conditions de causalité et d'inversibilité du modèle S -périodique, peuvent être héritiers donc de celles du modèle S -varié stationnaire, il est important de rappeler les conditions de causalité et d'inversibilité de ce dernier modèle.

Rappels

Le modèle autorégressif moyenne mobile multivarié (à coefficients constants) $VARMA(p, q)$, est donné par l'équation aux différences stochastique suivante :

$$\sum_{i=0}^p A_i \underline{X}_{t-i} = \sum_{i=0}^q B_i \underline{\varepsilon}_{t-i} \quad \forall t \in \mathbb{Z},$$

où $\underline{\varepsilon}_t$ est un processus multivarié non corrélé (bruit blanc) et les matrices A et B sont des matrices carrées réelles. Cette classe de modèle, en particulier, la classe particulière des modèles autorégressifs multivariés $VAR(p)$ ont intensivement étudiés. Parmi les résultats existant concernant cette classe de modèles sont liés aux conditions nécessaires et suffisantes pour que le modèle soit causal et inversible.

Causalité (stationnarité)

Théorème 3.2.3 (Hannan 1970)

Une condition nécessaire et suffisante pour que le modèle, $VARMA(p, q)$, soit stationnaire, est que les racines de l'équation caractéristique suivante :

$$\det\left(\sum_{i=0}^p A_i z^{p-i}\right) = 0,$$

sont à l'intérieur du cercle unité.

Inversibilité

Théorème 3.2.4 (Hannan 1970)

Une condition nécessaire et suffisante pour que le modèle, $VARMA(p, q)$, soit inversible, est que les racines de l'équation caractéristique suivante :

$$\det\left(\sum_{i=0}^q B_i z^{q-i}\right) = 0,$$

sont à l'intérieur du cercle unité.

2.2) Causalité et inversibilité deuxième approche

Nous intéressons ici à l'obtention d'une condition nécessaire et suffisante pour la stationnarité périodique (causalité) du modèle autorégressif périodique, $PAR_S(p)$, de période S et d'ordre p , suivant :

$$X_t - \sum_{i=1}^p \varphi_{i,t} X_{t-i} = \varepsilon_t, \quad t \in \mathbb{Z}, \quad (3.2.7)$$

où $\{\varepsilon_t\}$ est un processus bruit blanc tel que $\mathbb{E}(\varepsilon_t) = 0$ et $\mathbb{E}(\varepsilon_t^2) = \sigma_t^2$ et où les paramètres autorégressifs $\varphi_{t,i}$ et la variance σ_t^2 sont périodiques. Dans cette approche, nous construisons le processus multivarié en utilisant l'ordre p au lieu de la période S comme dans l'approche précédent de Gladyshev ce qui permet d'écrire le modèle (3.2.1) sous la forme d'un modèle autorégressif multivarié périodique d'ordre 1. En effet, considérons le processus multivarié, à p dimensions, intérieur :

$$\begin{aligned} \underline{X}_\tau &= (X_{p\tau}, X_{p\tau-1}, X_{p\tau-2}, \dots, X_{p\tau-(p-1)})', \quad \forall \tau \in \mathbb{Z}, \\ \underline{\varepsilon}_\tau &= (\varepsilon_{p\tau}, \varepsilon_{p\tau-1}, \varepsilon_{p\tau-2}, \dots, \varepsilon_{p\tau-(p-1)})', \quad \forall \tau \in \mathbb{Z}. \end{aligned}$$

En écrivant dans (2.1) t sous la forme par $t = s + p\tau$, $s = 1, 2, 3, \dots, S$ et $\tau \in \mathbb{Z}$, nous pouvons réécrire le modèle (2.1) sous la forme d'un modèle autorégressif p -varié suivant :

$$\Phi_{0,\tau} \underline{X}_\tau = \Phi_{1,\tau} \underline{X}_{\tau-1} + \underline{\varepsilon}_\tau, \quad \forall \tau \in \mathbb{Z}, \quad (3.2.8)$$

où les matrices $p \times p$ $\Phi_{0,\tau}$ et $\Phi_{1,\tau}$ sont données comme suit :

$$\Phi_{0,\tau} = \begin{pmatrix} 1 & -\varphi_{1,p\tau} & -\varphi_{2,p\tau} & \cdots & -\varphi_{p-1,p\tau} \\ 0 & 1 & -\varphi_{1,p\tau-1} & \cdots & -\varphi_{p-2,p\tau-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -\varphi_{1,p\tau-(p-2)} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \quad (3.2.9a)$$

$$\Phi_{1,\tau} = \begin{pmatrix} \varphi_{p,p\tau} & 0 & 0 & \cdots & 0 \\ \varphi_{p-1,p\tau-1} & \varphi_{p,p\tau-1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ \varphi_{2,p\tau-p+2} & \varphi_{3,p\tau-p+2} & \cdots & \varphi_{p,p\tau-p+2} & 0 \\ \varphi_{1,p\tau-p+1} & \varphi_{2,p\tau-p+1} & \cdots & \cdots & \varphi_{p,p\tau-p+1} \end{pmatrix}. \quad (3.2.9b)$$

Nous remarquons que les matrices sont périodiques de période \mathbb{S} , où \mathbb{S} est tel que $p\mathbb{S}$ est le plus multiple commun de p et S . Le modèle multivarié périodique (3.2.8) est équivalent au modèle périodique scalaire (3.2.7), alors le modèle (3.2.8) peut être exploité pour déduire beaucoup de propriétés du modèle (3.2.7), en particulier les conditions de causalité. Le théorème suivant établit une condition nécessaire et suffisante pour que le modèle périodique (3.2.8), et par la suite le modèle périodique (3.2.7), soit causal.

Théorème 3.2.5.

Le modèle autorégessif périodique, de période S , donné par (3.2.7) est causal si, et seulement si, les racines de l'équation déterminantale suivante

$$|Iz - \Psi| = 0,$$

sont toutes à l'intérieure du cercle unité, où la matrice Ψ est donnée par :

$$\Psi = \Phi_{0,\mathbb{S}}^{-1} \Phi_{1,\mathbb{S}} \Phi_{0,\mathbb{S}-1}^{-1} \Phi_{1,\mathbb{S}-1} \cdots \Phi_{0,2}^{-1} \Phi_{1,2} \Phi_{0,1}^{-1} \Phi_{1,1}.$$

Démonstration.

Nous remarquons d'abord que la matrice $\Phi_{0,\tau}$ est inversible $\forall \tau \in \mathbb{Z}$. Alors, en multipliant les deux membres du modèle (2.2) par , nous obtenons

$$\Phi_{0,\tau} \underline{X}_\tau = \Phi_{0,\tau}^{-1} \Phi_{1,\tau} \underline{X}_{\tau-1} + \Phi_{0,\tau}^{-1} \underline{\varepsilon}_\tau, \quad \forall \tau \in \mathbb{Z}.$$

En itérant cette équation aux récurrences $m - 1$ ($m > 1$) fois, on obtient

$$\begin{aligned} \underline{X}_\tau &= (\Phi_{0,\tau}^{-1} \Phi_{1,\tau}) (\Phi_{0,\tau-1}^{-1} \Phi_{1,\tau-1}) \dots (\Phi_{0,\tau-m+2}^{-1} \Phi_{1,\tau-m+2}) (\Phi_{0,\tau-m+1}^{-1} \Phi_{1,\tau-m+1}) \underline{X}_{\tau-m} \\ &+ \varepsilon_t + \sum_{j=1}^{m-1} \left\{ (\Phi_{0,\tau}^{-1} \Phi_{1,\tau}) (\Phi_{0,\tau-1}^{-1} \Phi_{1,\tau-1}) \dots (\Phi_{0,\tau-j+2}^{-1} \Phi_{1,\tau-j+2}) (\Phi_{0,\tau-j+1}^{-1} \Phi_{1,\tau-j+1}) \right\} \varepsilon_{t-j}. \end{aligned}$$

Considérons les notations suivantes

$$\begin{aligned} \text{et } A_{j,\tau} &= (\Phi_{0,\tau}^{-1} \Phi_{1,\tau}) (\Phi_{0,\tau-1}^{-1} \Phi_{1,\tau-1}) \dots (\Phi_{0,\tau-j+2}^{-1} \Phi_{1,\tau-j+2}) (\Phi_{0,\tau-j+1}^{-1} \Phi_{1,\tau-j+1}), \\ B_{m,\tau} &= (\Phi_{0,\tau}^{-1} \Phi_{1,\tau}) (\Phi_{0,\tau-1}^{-1} \Phi_{1,\tau-1}) \dots (\Phi_{0,\tau-m+2}^{-1} \Phi_{1,\tau-m+2}) (\Phi_{0,\tau-m+1}^{-1} \Phi_{1,\tau-m+1}). \end{aligned}$$

L'expression précédente s'écrit sous la forme suivante :

$$\underline{X}_\tau = \varepsilon_t + \sum_{j=1}^{m-1} A_{j,\tau} \varepsilon_{t-j} + B_{m,\tau} \underline{X}_{\tau-m}$$

En posant $\tau = s_0 + rS$, $0 \leq s_0 \leq S-1$, nous pouvons réécrire la matrice $B_{m,\tau}$ sous la forme suivante :

$$\begin{aligned} B_{s_0,r} &= \left\{ (\Phi_{0,s_0}^{-1} \Phi_{1,s_0}) (\Phi_{0,s_0-1}^{-1} \Phi_{1,s_0-1}) \dots (\Phi_{0,s_0-(S-1)}^{-1} \Phi_{1,s_0-(S-1)}) \right\}^r \\ &\quad \times (\Phi_{0,s_0}^{-1} \Phi_{1,s_0}) (\Phi_{0,s_0-1}^{-1} \Phi_{1,s_0-1}) \dots (\Phi_{0,1}^{-1} \Phi_{1,1}) \end{aligned}$$

la variance du premier membre de cette dernière expression est donnée par :

$$\text{Var} \left\{ \underline{X}_\tau - \varepsilon_\tau - \sum_{j=1}^{m-1} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \varepsilon_{t-j} \right) \right\} = B_{s_0,r} \text{Var} (\underline{X}_{\tau-m}) B'_{s_0,r}$$

En passant à la limite, quand $r \rightarrow \infty$, nous obtenons

$$\lim_{r \rightarrow \infty} \text{Var} \left\{ \underline{X}_\tau - \varepsilon_\tau - \sum_{j=1}^{m-1} \left(\prod_{i=0}^{j-1} \varphi_{\tau-i} \right) \varepsilon_{\tau-j} \right\} = \lim_{r \rightarrow \infty} \left(B_{s_0,r} \text{Var} (\underline{X}_{\tau-m}) B'_{s_0,r} \right),$$

Pour que $\lim_{r \rightarrow \infty} \left\{ (\Phi_{0,s_0}^{-1} \Phi_{1,s_0}) (\Phi_{0,s_0-1}^{-1} \Phi_{1,s_0-1}) \dots (\Phi_{0,s_0-(S-1)}^{-1} \Phi_{1,s_0-(S-1)}) \right\}^r = 0$, $\forall \tau \in \mathbb{Z}$, il

est suffisante et nécessaire les valeurs propres de la matrice $(\Phi_{0,s_0}^{-1} \Phi_{1,s_0}) (\Phi_{0,s_0-1}^{-1} \Phi_{1,s_0-1}) \dots (\Phi_{0,s_0-(S-1)}^{-1} \Phi_{1,s_0-(S-1)})$, $0 \leq s_0 \leq S-1$, sont toutes à l'intérieur du cercle unité. Comme les valeurs propres de cette dernière matrice sont elles les valeurs propres de la matrice

$$(\Phi_{0,s_0}^{-1} \Phi_{1,s_0}) (\Phi_{0,s_0-1}^{-1} \Phi_{1,s_0-1}) \dots (\Phi_{0,s_0-(S-1)}^{-1} \Phi_{1,s_0-(S-1)}), \quad 0 \leq s_0 \leq S-1.$$

Comme les valeurs propres ne changent pas si on fait une rotation circulaire, alors, en tenant compte de la périodicité, nous avons la matrice

$$\Psi = (\Phi_{0,S}^{-1} \Phi_{1,S}) (\Phi_{0,S-1}^{-1} \Phi_{1,S-1}) \dots (\Phi_{0,1}^{-1} \Phi_{1,1}). \quad \blacksquare$$

Remarque :

Pour l'inversibilité nous suivons la même démarche nous obtenons une condition nécessaire et suffisante pour l'inversibilité d'un modèle moyenne mobile périodique.

3.2.3 Décomposition de Wold-Cramér d'un $PAR_S(1)$

Soit $\{X_t; t \in \mathbb{Z}\}$ un processus du second ordre périodiquement corrélé de période S , satisfaisant un modèle autorégressif périodique d'ordre 1, suivant :

$$X_t = \varphi_t X_{t-1} + \varepsilon_t. \quad (3.2.10)$$

Le théorème suivant établit une condition nécessaire et suffisante pour que ce processus admette une représentation de Wold-Cramér est donc périodiquement stationnaire (causal), tout en donnant son expression.

Théorème 3.2.6.

Un processus, du second ordre, périodiquement corrélé de période S , $\{X_t; t \in \mathbb{Z}\}$, donné par le modèle (3.2.10) est périodiquement stationnaire si, et seulement si,

$$\left| \prod_{s=1}^S \varphi_s \right| < 1,$$

et son décomposition de Wold-Cramér est, sous cette condition, donnée par l'expression convergente, en moyenne quadratique, suivante :

$$X_t = \varepsilon_t - \sum_{j=1}^{\infty} \prod_{i=0}^{j-1} \varphi_{t-i} \varepsilon_{t-j}.$$

La variance de ce processus est donnée par :

$$Var(X_t) = \sigma_t^2 + \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \right)^2 \sigma_{t-j}^2.$$

Démonstration.

Par itération $m - 1$ ($m > 1$) fois, dans le modèle (3.2.10), on obtient

$$X_t = \left(\prod_{i=0}^{m-1} \varphi_{t-i} \right) X_{t-m} + \varepsilon_t + \sum_{j=1}^{m-1} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \right) \varepsilon_{t-j},$$

ou encore

$$X_t - \varepsilon_t - \sum_{j=1}^{m-1} \prod_{i=0}^{j-1} \varphi_{t-i} \varepsilon_{t-j} = \left(\prod_{i=0}^{m-1} \varphi_{t-i} \right) X_{t-m},$$

la variance du premier membre de cette dernière expression est donnée par :

$$Var \left\{ X_t - \varepsilon_t - \sum_{j=1}^{m-1} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \varepsilon_{t-j} \right) \right\} = \left(\prod_{i=0}^{m-1} \varphi_{t-i} \right)^2 Var(X_{t-m}) = \left(\prod_{i=0}^{m-1} \varphi_{t-i} \right)^2 C_1,$$

où C_1 est strictement positif et fini car le processus est du second ordre.

En passant à la limite, quand $m \rightarrow \infty$, nous obtenons

$$\lim_{m \rightarrow \infty} Var \left\{ X_t - \varepsilon_t - \sum_{j=1}^{m-1} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \right) \varepsilon_{t-j} \right\} = M \lim_{m \rightarrow \infty} \prod_{i=0}^{m-1} \varphi_{t-i}^2,$$

L'expression $\prod_{i=0}^{m-1} \varphi_{t-i}^2$ peut être réécrite, pour $m = s_0 + \tau S$, $0 \leq s_0 \leq S - 1$, sous la forme

suivante :

$$\begin{aligned} \prod_{i=0}^{m-1} \varphi_{t-i}^2 &= \left(\varphi_t^2 \varphi_{t-1}^2 \varphi_{t-2}^2 \cdots \varphi_{t-(S-1)}^2 \right)^\tau \prod_{i=0}^{s_0-1} \varphi_{t-i}^2, \\ &= C_2 \left(\varphi_t^2 \varphi_{t-1}^2 \varphi_{t-2}^2 \cdots \varphi_{t-(S-1)}^2 \right)^\tau \prod_{i=0}^{s_0-1} \varphi_{t-i}^2 \text{ où } C_2 \text{ est positif fini.} \end{aligned}$$

Il est clair que $m \rightarrow \infty \implies \tau \rightarrow \infty$ et que l'expression $\prod_{i=0}^{s_0-1} \varphi_{t-i}^2$ est le produit de s_0 termes

finis. Alors, en passant à la limite, quand $m \rightarrow \infty$, on obtient :

$$\lim_{m \rightarrow \infty} \prod_{i=0}^{m-1} \varphi_{t-i}^2 = \left(\prod_{i=0}^{s_0-1} \varphi_{t-i}^2 \right) \lim_{\tau \rightarrow \infty} \left(\varphi_t^2 \varphi_{t-1}^2 \cdots \varphi_{t-(S-1)}^2 \right)^\tau = \lim_{\tau \rightarrow \infty} \left((\varphi_t \varphi_{t-1} \cdots \varphi_{t-(S-1)})^2 \right)^\tau,$$

Cette dernière expression converge, quand $\tau \rightarrow \infty$, vers 0, si, et seulement si, $|\varphi_t \varphi_{t-1} \varphi_{t-2} \cdots \varphi_{t-(S-1)}| < 1$. En écrivant l'instant t sous la forme $t = s + kS$, $1 \leq s \leq S$, on peut réécrire

cette condition nécessaire et suffisante sous la forme équivalente : $\left| \prod_{i=1}^S \varphi_i \right| < 1$. Alors, l'ex-

pression $\lim_{m \rightarrow \infty} Var \left\{ X_t - \varepsilon_t - \sum_{j=1}^{m-1} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \right) \varepsilon_{t-j} \right\}$ converge, quand $m \rightarrow \infty$, vers 0, si, et

seulement si, $\left| \prod_{i=1}^S \varphi_i \right| < 1$. La expression de Wold-Cramér, est alors, donnée comme suit :

$$\left| X_t = \varepsilon_t - \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \right) \varepsilon_{t-j} \right|$$

Comme les ε_{t-j} sont non corrélés, alors la variance du processus est donnée par :

$$Var(X_t) = \sigma_t^2 + \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} \varphi_{t-i} \right)^2 \sigma_{t-j}^2. \blacksquare$$

3.3 Propriétés asymptotiques des estimateurs des paramètres d'un $PAR_S(1)$

Dans cette section, nous obtenons, d'abord, les estimateurs des Moindres Carrés *MCO* des paramètres $\varphi_1, \varphi_2, \dots, \varphi_S$ d'un $PAR_S(1)$ et par la suite, nous étudions les Propriétés Asymptotiques de ces estimateurs selon l'appartenance de ces paramètres à l'une de trois régions dites respectivement : région de stationnarité, région de racine unitaire et région d'explosivité qui seront bien spécifiées ultérieurement.

3.3.1 Estimateurs du MCO des paramètres d'un $PAR_S(1)$

Soit $\{X_t, t \in \mathbb{Z}\}$ un processus stochastique périodiquement corrélé, de période S , satisfaisant le modèle $PAR_S(1)$ causal donné par l'équation aux différences suivante :

$$X_t - \varphi_t X_{t-1} = \varepsilon_t, \quad (3.3.1)$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendante de moyennes et de variances σ_t^2 et où φ_t et σ_t^2 sont périodiques, de période S , c'est-à-dire,

$$\begin{aligned} \varphi_{t+kS} &= \varphi_t, \\ \sigma_{t+kS}^2 &= \sigma_t^2, \end{aligned} \quad \forall t \in \mathbb{Z} \text{ et } \forall k \in \mathbb{N}.$$

Le modèle (3.3.1), peut s'écrire pour $t = s + S\tau$, $s = 1, 2, \dots, S$ et $\tau \in \mathbb{Z}$, sous la forme équivalente suivante :

$$X_{s+S\tau} - \varphi_s X_{s-1+S\tau} = \varepsilon_{s+S\tau}, \quad (3.3.2)$$

En introduisant les vecteurs colonnes suivants :

$$\mathcal{X}_{s,\tau} = \left(\underbrace{0, \dots, 0}_{s-1 \text{ éléments}}, X_{s-1+S\tau}, \underbrace{0, \dots, 0}_{S-s \text{ éléments}} \right)', \quad (3.3.3)$$

et

$$\Phi = (\varphi_1, \varphi_2, \dots, \varphi_S)',$$

on peut réécrire le modèle (3.3.1), sous la forme suivante :

$$X_{s+S\tau} - \Phi' \mathcal{X}_{s,\tau} = \varepsilon_{s+S\tau}. \quad (3.3.4)$$

En considérant ces notations, nous pouvons énoncer le théorème établissant l'estimateur *MCO* d'un $PAR(1)$.

Théorème 3.3.1.

L'estimateur des moindres carrés $\hat{\Phi}_n$ du vecteur $\Phi = (\varphi_1, \varphi_2, \dots, \varphi_S)'$ est donné par :

$$\hat{\Phi}_n = \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau} X_{1+S\tau} / \sum_{\tau=0}^{n-1} X_{0+S\tau}^2 \\ \sum_{\tau=0}^{n-1} X_{1+S\tau} X_{2+S\tau} / \sum_{\tau=0}^{n-1} X_{1+S\tau}^2 \\ \vdots \\ \sum_{\tau=0}^{n-1} X_{S-1+S\tau} X_{S+S\tau} / \sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \end{pmatrix}$$

et nous avons l'expression

$$\sqrt{n} (\hat{\Phi}_n - \Phi) = \sqrt{S} \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau} \varepsilon_{1+S\tau} / \sqrt{n} \right) \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{1+S\tau} \varepsilon_{2+S\tau} / \sqrt{n} \right) \\ \vdots \\ \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau} \varepsilon_{S-1+S\tau} / \sqrt{n} \right) \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau} \varepsilon_{S+S\tau} / \sqrt{n} \right) \end{pmatrix}$$

Démonstration. La somme des carrés des erreurs $Q(\Phi)$ peut s'écrire, pour une taille multiple de S , i.e., $N = nS$, sous la forme :

$$Q(\Phi) = \sum_{\tau=0}^{n-1} \sum_{s=1}^S \varepsilon_{s+S\tau}^2 = \sum_{\tau=0}^{n-1} \sum_{s=1}^S (X_{s+S\tau} - \Phi' \mathcal{X}_{s,\tau}) (X_{s+S\tau} - \Phi' \mathcal{X}_{s,\tau})', \quad (3.3.5)$$

ou encore sous la forme :

$$\begin{aligned} Q &= \sum_{\tau=0}^{n-1} \sum_{s=1}^S (X_{s+S\tau} - \Phi' \mathcal{X}_{s,\tau}) (X_{s+S\tau} - \Phi' \mathcal{X}_{s,\tau})', \\ &= \sum_{\tau=0}^{n-1} \sum_{s=1}^S (X_{s+S\tau}^2 - 2 \Phi' \mathcal{X}_{s,\tau} X_{s+S\tau} + \Phi' \mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau} \Phi), \\ &= \sum_{\tau=0}^{n-1} \sum_{s=1}^S X_{s+S\tau}^2 - 2 \Phi' \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} X_{s+S\tau} \right) + \Phi' \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau} \right) \Phi, \end{aligned}$$

en dérivant cette expression quadratique par rapport au vecteur des paramètres Φ , on obtient l'équation normale suivante :

$$\frac{\partial Q}{\partial \Phi} = -2 \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} X_{s+S\tau} \right) + 2 \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau} \right) \Phi = \underline{0}.$$

Ce qui permet d'obtenir l'estimateur, $\widehat{\Phi}$, des moindres carrés du vecteur, Φ , des paramètres inconnus :

$$\widehat{\Phi} = \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau} \right)^{-1} \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} X_{s+S\tau} \right). \quad (3.3.6)$$

En tenant compte du fait que :

$$\begin{aligned} \sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau} &= \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau}^2 & 0 \cdots 0 \\ 0 & 0 \cdots 0 \\ \vdots & \vdots \\ 0 & 0 \cdots 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \cdots & \cdots 0 \\ 0 & \sum_{\tau=0}^{n-1} X_{1+S\tau}^2 & 0 \cdots 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \cdots 0 \end{pmatrix} \\ &+ \dots + \begin{pmatrix} 0 \cdots & 0 \\ 0 \cdots & 0 \\ \vdots & \vdots \\ 0 \cdots & 0 \sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \end{pmatrix} = \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau}^2 & 0 \cdots & 0 \\ 0 & \sum_{\tau=0}^{n-1} X_{1+S\tau}^2 & 0 \cdots 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots 0 & \sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \end{pmatrix} \\ \sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} X_{s+S\tau} &= \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau} X_{1+S\tau} \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \sum_{\tau=0}^{n-1} X_{1+S\tau} X_{2+S\tau} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots \end{aligned}$$

et

$$\begin{aligned} &+ \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \sum_{\tau=0}^{n-1} X_{S-1+S\tau} X_{S+S\tau} \end{pmatrix} = \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau} X_{1+S\tau} \\ \sum_{\tau=0}^{n-1} X_{1+S\tau} X_{2+S\tau} \\ \vdots \\ \sum_{\tau=0}^{n-1} X_{S-2+S\tau} X_{S-1+S\tau} \\ \sum_{\tau=0}^{n-1} X_{S-1+S\tau} X_{S+S\tau} \end{pmatrix} \end{aligned}$$

Et en remplaçant, dans (3.3.7), $\left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau}\right)^{-1}$ et $\sum_{\tau=0}^{n-1} \sum_{s=1}^S \mathcal{X}_{s,\tau} X_{s+S\tau}$ par leurs expressions explicites, on peut réécrire la formule (3.3.7) sous la forme explicite suivante :

$$\widehat{\Phi}_n = \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2\right)^{-1} & 0 & \cdots & 0 \\ 0 & \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2\right)^{-1} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \\ 0 & \cdots & 0 & \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2\right)^{-1} \end{pmatrix} \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau} X_{1+S\tau} \\ \sum_{\tau=0}^{n-1} X_{1+S\tau} X_{2+S\tau} \\ \vdots \\ \sum_{\tau=0}^{n-1} X_{S-2+S\tau} X_{S-1+S\tau} \\ \sum_{\tau=0}^{n-1} X_{S-1+S\tau} X_{S+S\tau} \end{pmatrix} \quad (3.3.7)$$

$$= \begin{pmatrix} \sum_{\tau=0}^{n-1} X_{0+S\tau} X_{1+S\tau} / \sum_{\tau=0}^{n-1} X_{0+S\tau}^2 \\ \sum_{\tau=0}^{n-1} X_{1+S\tau} X_{2+S\tau} / \sum_{\tau=0}^{n-1} X_{1+S\tau}^2 \\ \vdots \\ \sum_{\tau=0}^{n-1} X_{S-1+S\tau} X_{S+S\tau} / \sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \end{pmatrix}$$

Alors, l'estimateurs, $\widehat{\Phi}_n$, de la méthode des moindres carrés (MCO) du vecteur des paramètres Φ d'un $PAR_S(1)$ est donné par :

$$\widehat{\Phi}_n = \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2\right)^{-1} & \sum_{\tau=0}^{n-1} X_{0+S\tau} X_{1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2\right)^{-1} & \sum_{\tau=0}^{n-1} X_{1+S\tau} X_{2+S\tau} \\ \vdots & \\ \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2\right)^{-1} & \sum_{\tau=0}^{n-1} X_{S-2+S\tau} X_{S-1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2\right)^{-1} & \sum_{\tau=0}^{n-1} X_{S-1+S\tau} X_{S+S\tau} \end{pmatrix} \quad (3.3.8)$$

En remplaçant $X_{s+S\tau}$, donné par le modèle (3.3.4), par $\varphi_s X_{s-1+S\tau} + \varepsilon_{s+S\tau}$ nous obtenons

$$\widehat{\Phi}_n = \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{0+S\tau} (\varphi_1 X_{0+S\tau} + \varepsilon_{1+S\tau}) \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{1+S\tau} (\varphi_2 X_{1+S\tau} + \varepsilon_{2+S\tau}) \\ \vdots \\ \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-2+S\tau} (\varphi_{S-1} X_{S-2+S\tau} + \varepsilon_{S-1+S\tau}) \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-1+S\tau} (\varphi_S X_{S-1+S\tau} + \varepsilon_{S+S\tau}) \end{pmatrix} \quad (3.3.9)$$

$$= \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{S-1} \\ \varphi_S \end{pmatrix} + \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{0+S\tau} \varepsilon_{1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{1+S\tau} \varepsilon_{2+S\tau} \\ \vdots \\ \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-2+S\tau} \varepsilon_{S-1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-1+S\tau} \varepsilon_{S+S\tau} \end{pmatrix}$$

Finalement, nous avons

$$\widehat{\Phi}_n = \Phi + \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{0+S\tau} \varepsilon_{1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{1+S\tau} \varepsilon_{2+S\tau} \\ \vdots \\ \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-2+S\tau} \varepsilon_{S-1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-1+S\tau} \varepsilon_{S+S\tau} \end{pmatrix}$$

où $\widehat{\Phi}_n = (\widehat{\varphi}_{1,n}, \widehat{\varphi}_{2,n}, \dots, \widehat{\varphi}_{S,n})'$ est l'estimateur des moindres carrés ordinaire du vecteur des paramètres $\Phi = (\varphi_1, \varphi_2, \dots, \varphi_S)'$ ou encore

$$\widehat{\Phi}_n - \Phi = \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{0+S\tau} \varepsilon_{1+S\tau} \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{1+S\tau} \varepsilon_{2+S\tau} \\ \vdots \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 \right)^{-1} \sum_{\tau=0}^{n-1} X_{S-1+S\tau} \varepsilon_{S+S\tau} \end{pmatrix}$$

En multipliant, les membres cette dernière égalité, par \sqrt{N} où N est la taille de la série, que nous le prenons, sans perdre de généralité des résultats énoncés, égale un multiple de S , i.e., $N = nS$, nous obtenons l'expression suivante :

$$\sqrt{n} (\widehat{\Phi}_n - \Phi) = \sqrt{S} \begin{pmatrix} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau} \varepsilon_{1+S\tau} / \sqrt{n} \right) \\ \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{1+S\tau} \varepsilon_{2+S\tau} / \sqrt{n} \right) \\ \vdots \\ \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau} \varepsilon_{S-1+S\tau} / \sqrt{n} \right) \\ \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau} \varepsilon_{S+S\tau} / \sqrt{n} \right) \end{pmatrix} \quad (3.3.10)$$

■

3.3.2 Étude des Propriétés Asymptotiques des Estimateurs du MCO d'un $PAR_S(1)$

Nous avons vu que le processus périodiquement corrélé, de période S , satisfaisant le modèle $PAR_S(1)$, (3.3.1), est périodiquement stationnaire si, et seulement si, $\left| \prod_{i=1}^S \varphi_i \right| < 1$.

En notant le produit $\prod_{i=1}^S \varphi_i$ par ϕ , nous concluons que ϕ engendre trois régions différentes à savoir :

- A) Région correspondant aux $(\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_S)$ telle que $\left| \prod_{i=1}^S \varphi_i \right| < 1$ où le processus est stationnaire dite *région de stabilité*
- B) Région correspondant aux $(\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_S)$ telle que $\left| \prod_{i=1}^S \varphi_i \right| = 1$ où le processus est intégré dite *région de racine unitaire*
- C) Région correspondant aux $(\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_S)$ telle que $\left| \prod_{i=1}^S \varphi_i \right| > 1$ où le processus est explosif dite *région explosive*

Nous nous intéressons, dans la présente section, au problème intéressant qui consiste en l'étude de l'influence du type de la région de stabilité sur le comportement de la loi asymptotique du vecteur des estimateurs MCO, $\widehat{\Phi}$, du vecteur des paramètres, Φ . Plus précisément, nous étudions, selon la région, la convergence de la suite des fonctions de répartition empiriques correspondant aux estimateurs $\widehat{\Phi}_n$, $n \in \mathbb{N}^*$ (le vecteur est basé sur un échantillon de taille n), quand $n \rightarrow \infty$, vers une certaine fonction de répartition fixée (ne dépend plus de n). Autrement dit, nous établissons, selon le type de la région spécifiée, la convergence, en loi, quand $n \rightarrow \infty$, de la suite des estimateurs MCO, $\left\{ \widehat{\Phi}_n, n \in \mathbb{N}^* \right\}$.

Estimateurs des moindres carrés pondérés

Nous tenons à remarquer que le vecteur des estimateurs donné par (3.3.10) est obtenu, pour le cas où la variance σ_t est constante. Néanmoins, nous pouvons, facilement, montrer que, dans le cas cadre de périodicité de la variance, cet estimateur est donné par l'estimateur des moindres carrés pondérés suivant :

En effet, comme la variance de ε_t , dans le modèle (3.3.1) est σ_t^2 , alors introduisant un nouveau processus, η_t , dont la variance est constante et est égal à l'unité, définie par

$$\sigma_t \eta_t = \varepsilon_t,$$

et en remplaçant ε_t , dans (3.3.1), par $\sigma_t \eta_t$ on a

$$X_t - \varphi_t X_{t-1} = \sigma_t \eta_t.$$

Comme σ_t est strictement positive ($\sigma_t > 0, \forall t \in \mathbb{Z}$), alors on obtient :

$$\frac{X_t}{\sigma_t} - \varphi_t \frac{X_{t-1}}{\sigma_t} = \eta_t,$$

qui s'écrit pour $t = s + S\tau$ sous la forme :

$$\frac{X_{s+S\tau}}{\sigma_s} - \varphi_s \frac{X_{s-1+S\tau}}{\sigma_s} = \eta_{s+S\tau},$$

Alors, en suivant les mêmes étapes de calcul que dans le cas des estimateurs des moindres carrés ordinaires, nous obtenons :

$$\widehat{\Phi}_n = \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \frac{\mathcal{X}_{s,\tau} \mathcal{X}'_{s,\tau}}{\sigma_s^2} \right)^{-1} \left(\sum_{\tau=0}^{n-1} \sum_{s=1}^S \frac{\mathcal{X}_{s,\tau} X_{s+S\tau}}{\sigma_s^2} \right). \quad (3.3.11)$$

ou plus explicitement

$$\sqrt{n} (\widehat{\Phi}_n - \Phi) = \sqrt{S} \begin{pmatrix} \frac{1}{\sigma_1^2} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{0+S\tau} \varepsilon_{1+S\tau} / \sqrt{n} \right) \\ \frac{1}{\sigma_2^2} \left(\sum_{\tau=0}^{n-1} X_{1+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{1+S\tau} \varepsilon_{2+S\tau} / \sqrt{n} \right) \\ \vdots \\ \frac{1}{\sigma_{S-1}^2} \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{S-2+S\tau} \varepsilon_{S-1+S\tau} / \sqrt{n} \right) \\ \frac{1}{\sigma_S^2} \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau}^2 / n \right)^{-1} \left(\sum_{\tau=0}^{n-1} X_{S-1+S\tau} \varepsilon_{S+S\tau} / \sqrt{n} \right) \end{pmatrix} \quad (3.3.12)$$

A. Loi asymptotique de $\sqrt{n}(\widehat{\Phi} - \Phi)$ dans le cas stationnaire.

Nous avons vu que la région dans laquelle le processus est stationnaire est donnée par $\left| \prod_{s=1}^S \varphi_s \right| < 1$. Alors, sous cette hypothèse, la suite stochastique $\{X_{s+S\tau}, \tau \in \mathbb{N}\}$ correspondant à une valeur quelconque fixé s (la saison s), $s = 1, 2, 3, \dots$, est faiblement stationnaire. Par conséquent, l'obtention de la loi asymptotique de chaque composante (parmi les S composantes) du vecteur $\sqrt{n}(\widehat{\Phi}_n - \Phi)$ donné par (3.3.13) se fait, en appliquant les mêmes arguments utilisés, précédemment, dans le cas où le paramètre φ est constant.

Théorème 3.3.2.

Supposons les paramètres $\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_S$ du modèle (3.3.1) appartient à la région de stabilité donnée par : $\left| \prod_{s=1}^S \varphi_s \right| < 1$, et soit $\widehat{\Phi}_n$ l'estimateur du MCO du vecteur Φ , alors, nous avons la convergence, en loi, quand $n \rightarrow \infty$:

$$\sqrt{nS}(\widehat{\Phi}_n - \Phi) \xrightarrow{\mathcal{L}} N(0, S\Gamma^{-1}).$$

où Γ est la matrice de variance-covariance du vecteur colonne

$$W_n^{(s)} = \left(\frac{1}{\sigma_1^2} \sum_{\tau=0}^{n-1} \frac{X_{0+S\tau} \varepsilon_{1+S\tau}}{\sqrt{n}}, \frac{1}{\sigma_2^2} \sum_{\tau=0}^{n-1} \frac{X_{1+S\tau} \varepsilon_{2+S\tau}}{\sqrt{n}}, \dots, \frac{1}{\sigma_s^2} \sum_{\tau=0}^{n-1} \frac{X_{s-1+S\tau} \varepsilon_{s+S\tau}}{\sqrt{n}}, \dots, \frac{1}{\sigma_S^2} \sum_{\tau=0}^{n-1} \frac{X_{S-1+S\tau} \varepsilon_{S+S\tau}}{\sqrt{n}} \right)',$$

qui est donnée par :

$$\Gamma = \begin{pmatrix} \frac{\mathbb{E}(X_{0+S\tau}^2)}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{\mathbb{E}(X_{1+S\tau}^2)}{\sigma_2^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \frac{\mathbb{E}(X_{S-1+S\tau}^2)}{\sigma_S^2} \end{pmatrix}$$

Démonstration.

Nous donnons une brève démonstration, pour plus de détails voir *Chapitre 2*. (2.3.2). Comme le processus X_t est périodiquement stationnaire, alors pour un s fixé quelconque, $s = 1, 2, \dots, S$, $X_{s-1+S\tau}$ est faiblement stationnaire. alors, la lois asymptotique de la s ième composante, que nous notons $W_n^{(s)}$, du vecteur W_n est la loi normale de moyenne nulle et variance

de variance $\frac{\mathbb{E}(X_{s-1+S\tau}^2)}{\sigma_s^2}$, i.e., $W_n^{(s)} \xrightarrow{\mathcal{L}} N\left(0, \frac{\mathbb{E}(X_{s-1+S\tau}^2)}{\sigma_s^2}\right)$. D'autre part, du fait que $X_{s-1+S\tau}$ est faiblement stationnaire, alors $\frac{1}{\sigma_s^2} \left(\sum_{\tau=0}^{n-1} X_{s-1+S\tau}^2 / n\right)^{-1} \xrightarrow{p} \left(\frac{\mathbb{E}(X_{s-1+S\tau}^2)}{\sigma_s^2}\right)^{-1}$. Alors, en appliquant le *Théorème 1.2.18 (Théorème de Slutsky)*, nous obtenons la convergence, en loi, suivante $Z_n^{(s)} = \sqrt{S} \frac{1}{\sigma_s^2} \left(\sum_{\tau=0}^{n-1} X_{s-1+S\tau}^2 / n\right)^{-1} W_n^{(s)} \xrightarrow{\mathcal{L}} N\left(0, S \left(\frac{\mathbb{E}(X_{s-1+S\tau}^2)}{\sigma_s^2}\right)^{-1}\right)$.

Nous avons montré que la s ième composante du vecteur aléatoire

$$Z = \left(Z_n^{(1)}, Z_n^{(2)}, \dots, Z_n^{(s)}, \dots, Z_n^{(S)}\right)'$$

converge, en loi, vers une variable aléatoire scalaire de loi $N\left(0, S \left(\frac{\mathbb{E}(X_{s-1+S\tau}^2)}{\sigma_s^2}\right)^{-1}\right)$. Pour montrer que le vecteur aléatoire Z converge, en loi, vers une variable aléatoire vectorielle normale nous appliquons le *Théorème 1.2.7 (Cramér-Wold devise)*. En effet, soit

$$\lambda' Z = \sum \lambda_s Z_n^{(s)}.$$

du fait qu'une combinaison linéaire des variables aléatoires non corrélées suivent aussi une loi normale de moyenne $\mathbb{E}(\lambda' Z) = \mathbb{E}\left(\sum \lambda_s Z_n^{(s)}\right) = 0$, et de variance $Var(\lambda' Z) = \lambda' \Gamma \lambda = \Gamma$. Alors, nous avons d'après le *Théorème 1.2.7* $\sqrt{n}S \left(\widehat{\Phi}_n - \Phi\right) \xrightarrow{\mathcal{L}} N(0, S \Gamma^{-1})$. ■

B. Loi asymptotique de $n \left(\widehat{\Phi} - \Phi\right)$ dans le cas de racine unitaire

Boswijk et Franses ont montré que dans le cas où $\left|\prod_{s=1}^S \varphi_s\right| = 1$, on a :

$$n(\widehat{\Phi} - \Phi) \xrightarrow{\mathcal{L}} \left(\int_0^1 W(r)^2 dr\right)^{-1} \int_0^1 W(r) dW(r)$$

3.4 Simulation

Nous présentons dans cette section les résultats d'une étude de simulation intensive portant, d'une part, sur les comportements de l'estimateur MCO , $\widehat{\Phi}_n$, du paramètre inconnu Φ , d'un modèle autorégressif périodique d'ordre 1 et de période 4 $PAR_4(1)$, et l'estimateur de son écart-type $\widehat{\sigma}_{\widehat{\Phi}_n}$, et d'autre part, sur le comportement de ses lois asymptotiques et ce selon les régions de l'appartenance du produit des paramètres inconnus de Φ et selon la taille de la série simulée, tout en commentant sur la vitesse de convergence. Pour étudier empiriquement la performance de cet estimateur et le comportement de sa distribution asymptotique, dans les trois régions à savoir de stabilité, de racine unitaire et d'explosibilité (instabilité explosive), nous avons considéré le vecteur $(1, 0.2, 0.9, 0.5)'$ représentant le vecteur de paramètres $(\varphi_1, \varphi_2, \varphi_3, \varphi_4)'$ pour le cas où $\left| \prod_{s=1}^4 \varphi_s \right| < 1$, les vecteur $(2, 5, 0.1, 1)'$ pour le cas où $\left| \prod_{s=1}^4 \varphi_s \right| = 1$, et le vecteur $(1, 2, 7.5, 0.1)'$ pour le cas explosif où $\left| \prod_{s=1}^4 \varphi_s \right| > 1$.

Commençons d'abord, par les résultats de simulations concernant la consistance, au sens de la convergence en probabilité et presque sûrement, de l'estimateur $\widehat{\Phi}_n$ et de l'estimateur de son écart-type $\widehat{\sigma}_{\widehat{\Phi}_n}$ dans le cas stationnaire $\left(\left| \prod_{s=1}^4 \varphi_s \right| < 1 \right)$. Les valeurs moyennes, basées sur 1000 répliques, de l'estimateur du MCO , en variant la taille de la série de $n = 60$ à $n = 500$, avec un pas de 40, ainsi que les estimations de son écart-type $\widehat{\sigma}_{\widehat{\Phi}_n}$ sont portées sur la *Table 3.4.1* ci-dessous. Nous constatons que les résultats des simulations consolident les résultats théoriques concernant la consistance (au sens des modes de convergence) de l'estimateur $\widehat{\Phi}_n$, en particulier le résultat assurant la convergence presque sûre. En effet, le vecteur estimé $\widehat{\Phi}_n$ approche du vrai vecteurs $(1, 0.2, 0.9, 0.5)'$ au fur et à mesure que la taille de la série augmente, plus précisément, il passe du vecteur estimé $(0.9544, 0.1904, 0.8790, 0.4825)'$ avec un vecteur d'écart-type estimé $(0.2014, 0.2201, 0.1811, 0.1594)'$ pour $n = 60$ au vecteur estimé $(0.9963, 0.2006, 0.8958, 0.5008)'$ avec un écart-type estimé $(0.0672, 0.0762, 0.0636, 0.0564)'$ pour $n = 500$.

En conclusion, les simulations montrent empiriquement que l'estimateur du MCO, pour le modèle correspondant à $\Phi = (1, 0.2, 0.9, 0.5)'$ appartenant à la région de stabilité, est doté de la propriété de consistance.

Les Tables 3.4.2 et 3.4.3 montrent que la propriété de consistance est bien rencontrée aussi dans les deux cas d'instabilité à savoir l'instabilité à racine unitaire et l'instabilité explosif correspondant, respectivement, à $\Phi = (2, 5, 0.1, 1)'$ et $\Phi = (1, 2, 7.5, 0.1)'$.

Itération = 1000

$Var(\varepsilon_t) = (1, 2, 1, 1)'$

Taille de l'échantillon	Φ	$\hat{\Phi}$	$Std(\hat{\Phi})$
$N = 60$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9527 \\ 0.1983 \\ 0.8732 \\ 0.44848 \end{pmatrix}$	$\begin{pmatrix} 0.1965 \\ 0.2253 \\ 0.1857 \\ 0.1603 \end{pmatrix}$
$N = 100$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9777 \\ 0.1910 \\ 0.8785 \\ 0.4860 \end{pmatrix}$	$\begin{pmatrix} 0.1604 \\ 0.1781 \\ 0.1849 \\ 0.1263 \end{pmatrix}$
$N = 140$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9798 \\ 0.1912 \\ 0.8892 \\ 0.4938 \end{pmatrix}$	$\begin{pmatrix} 0.1271 \\ 0.1514 \\ 0.1201 \\ 0.1065 \end{pmatrix}$
$N = 180$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9821 \\ 0.1872 \\ 0.8859 \\ 0.4969 \end{pmatrix}$	$\begin{pmatrix} 0.1177 \\ 0.1286 \\ 0.1053 \\ 0.0927 \end{pmatrix}$
$N = 220$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9925 \\ 0.1928 \\ 0.8875 \\ 0.4971 \end{pmatrix}$	$\begin{pmatrix} 0.1092 \\ 0.1153 \\ 0.0936 \\ 0.0845 \end{pmatrix}$
$N = 260$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9902 \\ 0.1981 \\ 0.8889 \\ 0.4951 \end{pmatrix}$	$\begin{pmatrix} 0.0984 \\ 0.1063 \\ 0.0865 \\ 0.0756 \end{pmatrix}$

Table 3.4.1

Itération = 1000

$Var(\varepsilon_t) = (1, 2, 1, 1)'$

Taille de l'échantillon	Φ	$\hat{\Phi}$	$Std(\hat{\Phi})$
$N = 300$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9888 \\ 0.2041 \\ 0.8909 \\ 0.4979 \end{pmatrix}$	$\begin{pmatrix} 0.0883 \\ 0.0969 \\ 0.0779 \\ 0.0703 \end{pmatrix}$
$N = 340$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9925 \\ 0.1936 \\ 0.8969 \\ 0.4953 \end{pmatrix}$	$\begin{pmatrix} 0.0838 \\ 0.0893 \\ 0.0768 \\ 0.0665 \end{pmatrix}$
$N = 380$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9949 \\ 0.1952 \\ 0.8912 \\ 0.4954 \end{pmatrix}$	$\begin{pmatrix} 0.0785 \\ 0.0898 \\ 0.0709 \\ 0.0661 \end{pmatrix}$
$N = 420$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9947 \\ 0.1974 \\ 0.8919 \\ 0.4988 \end{pmatrix}$	$\begin{pmatrix} 0.0765 \\ 0.0843 \\ 0.0687 \\ 0.0601 \end{pmatrix}$
$N = 460$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9932 \\ 0.1992 \\ 0.8934 \\ 0.4984 \end{pmatrix}$	$\begin{pmatrix} 0.0729 \\ 0.0827 \\ 0.0642 \\ 0.0555 \end{pmatrix}$
$N = 500$	$\begin{pmatrix} 1 \\ 0.2 \\ 0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.9944 \\ 0.1992 \\ 0.8960 \\ 0.5004 \end{pmatrix}$	$\begin{pmatrix} 0.0676 \\ 0.0784 \\ 0.0627 \\ 0.0524 \end{pmatrix}$

Table 3.4.1 (suite)

Itération = 1000

$Var(\varepsilon_t) = (1, 2, 1, 1)'$

Taille de l'échantillon	Φ	$\hat{\Phi}$	$Std(\hat{\Phi})$
$N = 60$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9740 \\ 4.9819 \\ 0.0971 \\ 0.9666 \end{pmatrix}$	$\begin{pmatrix} 0.0787 \\ 0.0559 \\ 0.0082 \\ 0.0857 \end{pmatrix}$
$N = 100$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9853 \\ 4.9904 \\ 0.0976 \\ 0.9764 \end{pmatrix}$	$\begin{pmatrix} 0.0522 \\ 0.0380 \\ 0.0054 \\ 0.0579 \end{pmatrix}$
$N = 140$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9896 \\ 4.9966 \\ 0.0982 \\ 0.9815 \end{pmatrix}$	$\begin{pmatrix} 0.0398 \\ 0.0260 \\ 0.0042 \\ 0.0454 \end{pmatrix}$
$N = 180$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9922 \\ 4.9966 \\ 0.0986 \\ 0.9871 \end{pmatrix}$	$\begin{pmatrix} 0.0286 \\ 0.0213 \\ 0.0032 \\ 0.0313 \end{pmatrix}$
$N = 220$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9934 \\ 4.9970 \\ 0.0987 \\ 0.9876 \end{pmatrix}$	$\begin{pmatrix} 0.0251 \\ 0.0186 \\ 0.0031 \\ 0.0294 \end{pmatrix}$
$N = 260$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9941 \\ 4.9974 \\ 0.0989 \\ 0.9879 \end{pmatrix}$	$\begin{pmatrix} 0.0222 \\ 0.0150 \\ 0.0024 \\ 0.0268 \end{pmatrix}$
$N = 300$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9948 \\ 4.9984 \\ 0.0991 \\ 0.9811 \end{pmatrix}$	$\begin{pmatrix} 0.0196 \\ 0.0136 \\ 0.0022 \\ 0.0218 \end{pmatrix}$
$N = 340$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9954 \\ 4.9985 \\ 0.0992 \\ 0.9912 \end{pmatrix}$	$\begin{pmatrix} 0.0180 \\ 0.0121 \\ 0.0020 \\ 0.0216 \end{pmatrix}$
$N = 380$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9956 \\ 4.9989 \\ 0.0993 \\ 0.9924 \end{pmatrix}$	$\begin{pmatrix} 0.0161 \\ 0.0107 \\ 0.0017 \\ 0.0180 \end{pmatrix}$

Table 3.4.2

Itération = 1000

$$Var(\varepsilon_t) = (1, 2, 1, 1)'$$

Taille de l'échantillon	Φ	$\hat{\Phi}$	$Std(\hat{\Phi})$
$N = 420$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9968 \\ 4.9993 \\ 0.0994 \\ 0.9940 \end{pmatrix}$	$\begin{pmatrix} 0.0134 \\ 0.0094 \\ 0.0015 \\ 0.0154 \end{pmatrix}$
$N = 460$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9966 \\ 4.9987 \\ 0.0994 \\ 0.9942 \end{pmatrix}$	$\begin{pmatrix} 0.0131 \\ 0.0089 \\ 0.0014 \\ 0.0142 \end{pmatrix}$
$N = 500$	$\begin{pmatrix} 2 \\ 5 \\ 0.1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1.9971 \\ 4.9992 \\ 0.0995 \\ 0.9944 \end{pmatrix}$	$\begin{pmatrix} 0.0124 \\ 0.0082 \\ 0.0013 \\ 0.0137 \end{pmatrix}$

Table 3.4.2 (suite)

Itération = 1000

$$Var(\varepsilon_t) = (1, 2, 1, 1)'$$

Taille de l'échantillon	Φ	$\hat{\Phi}$	$Std(\hat{\Phi})$
$N = 60$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 0.9999 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.0017 \\ 0.0001 \\ 0.0010 \\ 0.0001 \end{pmatrix}$
$N = 100$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.0892 \times 10^{-4} \\ 0.1032 \times 10^{-4} \\ 0.1175 \times 10^{-4} \\ 0.0040 \times 10^{-4} \end{pmatrix}$
$N = 140$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.2892 \times 10^{-6} \\ 0.6258 \times 10^{-6} \\ 0.1386 \times 10^{-6} \\ 0.0130 \times 10^{-6} \end{pmatrix}$
$N = 180$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.8476 \times 10^{-9} \\ 0.6218 \times 10^{-9} \\ 0.4337 \times 10^{-9} \\ 0.0403 \times 10^{-9} \end{pmatrix}$

Table 3.4.3

Itération = 1000

$Var(\varepsilon_t) = (1, 2, 1, 1)'$

Taille de l'échantillon	Φ	$\hat{\Phi}$	$Std(\hat{\Phi})$
$N = 220$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.3015 \times 10^{-10} \\ 0.3047 \times 10^{-10} \\ 0.0858 \times 10^{-10} \\ 0.0210 \times 10^{-10} \end{pmatrix}$
$N = 260$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.1542 \times 10^{-12} \\ 0.3791 \times 10^{-12} \\ 0.1094 \times 10^{-12} \\ 0.0309 \times 10^{-12} \end{pmatrix}$
$N = 300$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.1789 \times 10^{-13} \\ 0.2069 \times 10^{-13} \\ 0.0772 \times 10^{-13} \\ 0.0149 \times 10^{-13} \end{pmatrix}$
$N = 340$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0.0204 \times 10^{-14} \\ 0.0117 \times 10^{-14} \\ 0.0126 \times 10^{-14} \\ 0.1416 \times 10^{-14} \end{pmatrix}$
$N = 380$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.0014 \times 10^{-14} \\ 0 \\ 0.1416 \times 10^{-14} \end{pmatrix}$
$N = 420$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.1416 \times 10^{-14} \end{pmatrix}$
$N = 460$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.1416 \times 10^{-14} \end{pmatrix}$
$N = 500$	$\begin{pmatrix} 1 \\ 2 \\ 7.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ 2.0000 \\ 7.5000 \\ 0.1000 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.1416 \times 10^{-14} \end{pmatrix}$

Table 3.4.3 (suite)

Comportement des distributions asymptotiques

Nous avons vu que les résultats de simulation ont bien confirmé la consistance, au sens de la convergence en probabilité et presque sûrement, de l'estimateur MCO , $\widehat{\Phi}_n$, du paramètre inconnu Φ , du modèle $PAR_4(1)$, et ceci dans les trois cas à savoir le cas de stabilité, d'instabilité à racine unitaire et d'instabilité explosif. Dans ce paragraphe, nous verrons que l'étude de simulation confirme empiriquement aussi les résultats théoriques établis concernant le comportement de la distribution asymptotique de cet estimateur dans les trois cas précédents. En effet, nous verrons que la loi empirique de l'estimateur $\widehat{\Phi}_n$ converge, quand la taille n tend vers l'infini vers une loi normale dans le cas stationnaire cas, vers une loi de Wiener dans le cas de racine unitaire et vers une loi de Cauchy dans le cas explosif.

a) Distributions asymptotiques dans le cas de stabilité, pour $\Phi = (1, 0.2, 0.9, 0.5)'$

Nous présentons dans la table (3.4.4) ci dessous les graphes des histogrammes et les densités de probabilité ajustées aux valeurs de l'estimateur du MCO , $\widehat{\Phi}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 60$, et $n = 500$, générées d'un modèle $PAR_4(1)$, correspondant au vecteur appartenant à la région de stabilité $\Phi = (1, 0.2, 0.9, 0.5)'$.

De ces différents graphes, nous constatons que les densités de probabilité de la variable aléatoire $\sqrt{n}(\widehat{\Phi}_n - \Phi)$ approche, au fur et à mesure que n augmente, d'une densité de probabilité d'une variable aléatoire normalement distribuée.

b) Distributions asymptotiques dans le cas de racine unitaire, pour $\Phi = (2, 5, 0.1, 1)'$

De même, des graphes des histogrammes et les densités de probabilité ajustées aux valeurs de l'estimateur du MCO , $\widehat{\Phi}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 60$, et $n = 500$, générées d'un modèle $PAR_4(1)$, correspondant au vecteur appartenant à la région de stabilité $\Phi = (2, 5, 0.1, 1)'$ (table (3.4.5)), nous constatons que les densités de probabilité de la variable aléatoire $n(\widehat{\Phi}_n - \Phi)$ approche, au fur et à mesure que n augmente, d'une densité de probabilité d'une variable aléatoire de loi

$$\left(\int_0^1 W(r)^2 dr \right)^{-1} \int_0^1 W(r) dW(r).$$

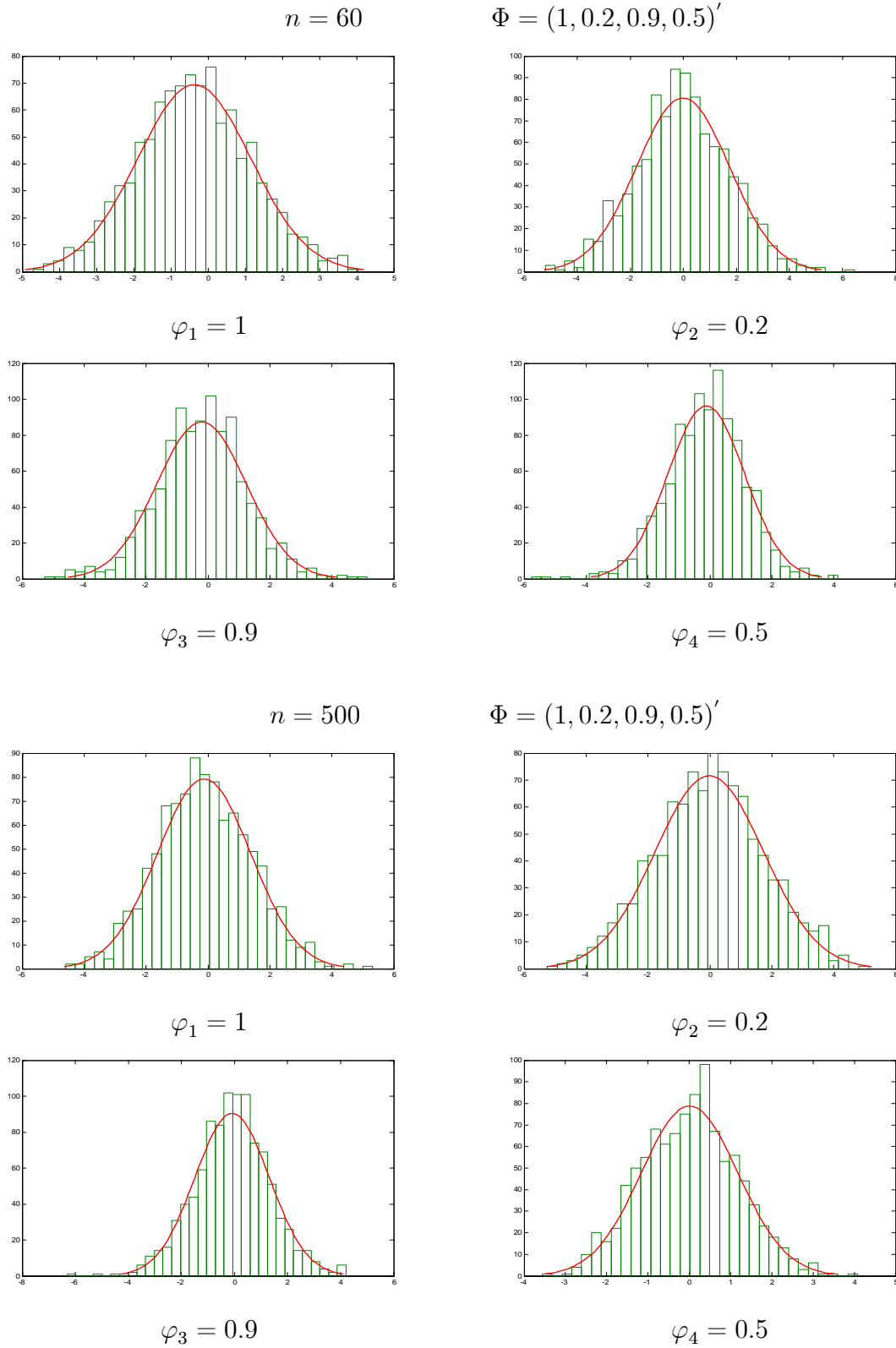
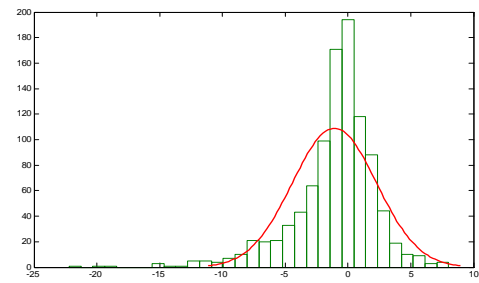
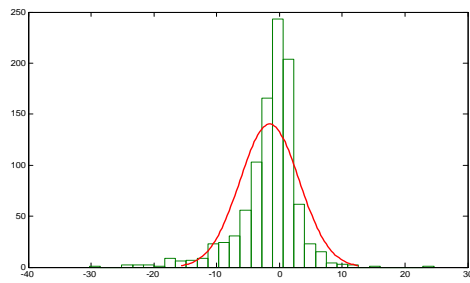


Table 3.4.4

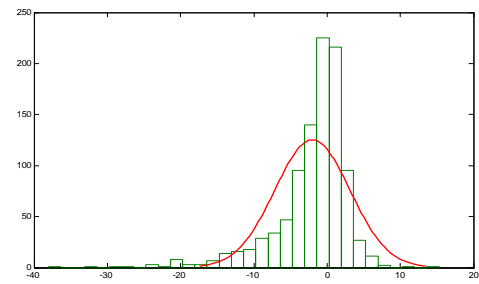
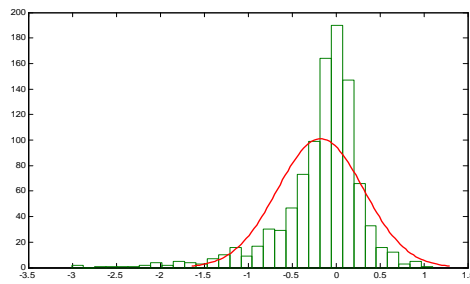
$n = 60$

$\Phi = (2, 5, 0.1, 1)'$



$\varphi_1 = 2$

$\varphi_2 = 5$

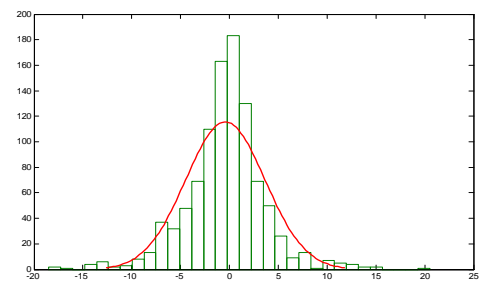
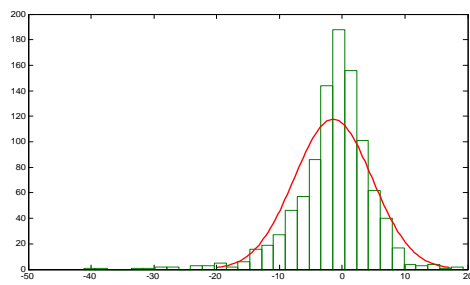


$\varphi_3 = 0.1$

$\varphi_4 = 1$

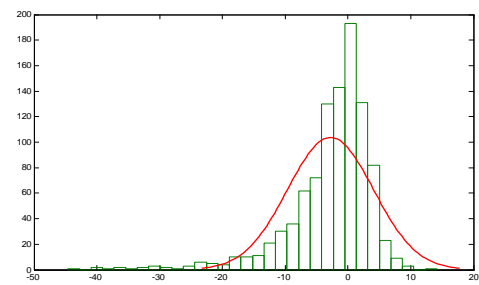
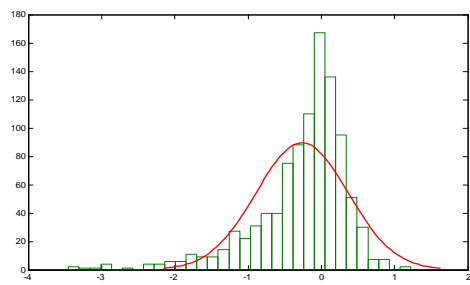
$n = 500$

$\Phi = (2, 5, 0.1, 1)'$



$\varphi_1 = 2$

$\varphi_2 = 5$



$\varphi_3 = 0.1$

$\varphi_4 = 1$

Table 3.4.5

c) **Distributions asymptotiques dans le cas explosif , pour $\Phi = (1, 2, 7.5, 0.1)'$**

Des graphes des histogrammes et les densités de probabilité ajustées aux valeurs de l'estimateur du MCO, $\widehat{\Phi}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 60$, et $n = 300$, générées d'un modèle $PAR_4(1)$, correspondant au vecteur appartenant à la région de stabilité $\Phi = (1, 2, 7.5, 0.1)'$ (table (3.4.6)), nous constatons que les densités de probabilité de la variable aléatoire $\Phi^n (\widehat{\Phi}_n - \Phi)$ approche, au fur et à mesure que n augmente, d'une densité de probabilité d'une variable aléatoire de loi de Cauchy

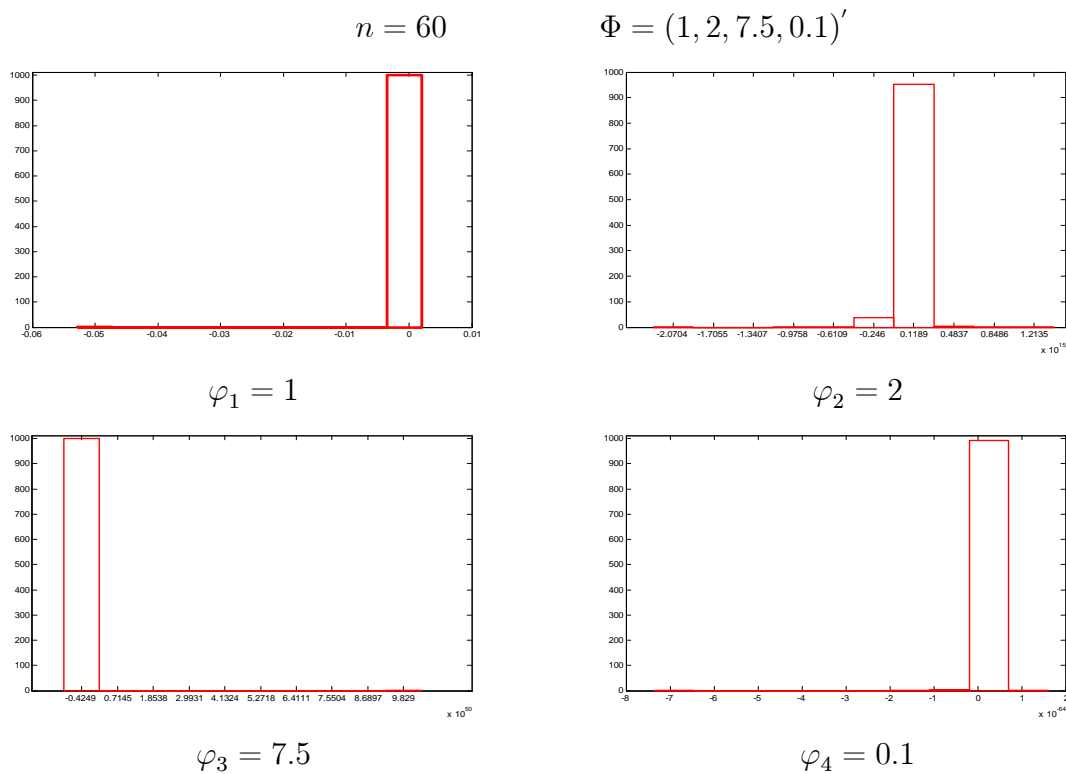


Table 3.4.6

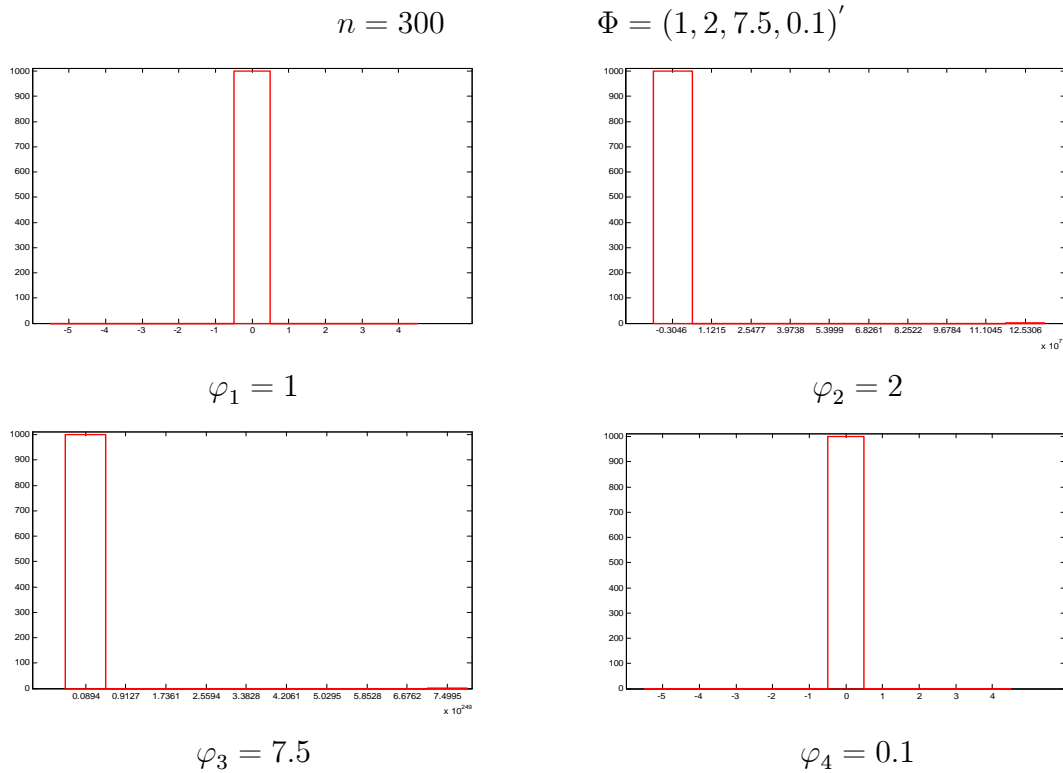


Table 3.4.6 (suite)

d) **Comportement de l'estimateur $\hat{\Phi}_n$ et l'estimateur de son écart-type $\hat{\sigma}_{\hat{\Phi}_n}$ dans les trois régions**

d. 1) **Dans la région de stabilité ($\Phi = (1, 0.2, 0.9, 0.5)'$)**

Les graphes de la *Table 3.4.7* ci-dessous représentent l'évolution de l'estimateur, $\hat{\Phi}_n$, des moindres carrés du paramètres de $\Phi = (1, 0.2, 0.9, 0.5)'$, et ceux ce la *Table 3.4.8* représentent l'évolution de leur écart-type $\hat{\sigma}_{\hat{\Phi}_n}$, en fonction des séries temporelles simulées et commentées. A partir de ces graphes nous constatons clairement que les valeurs estimées des paramètres de Φ approchent rapidement des vraies valeurs des paramètres de Φ . De même, les estimations de l'écart-type de ces estimateurs diminue au fur et à mesure que n augmente.

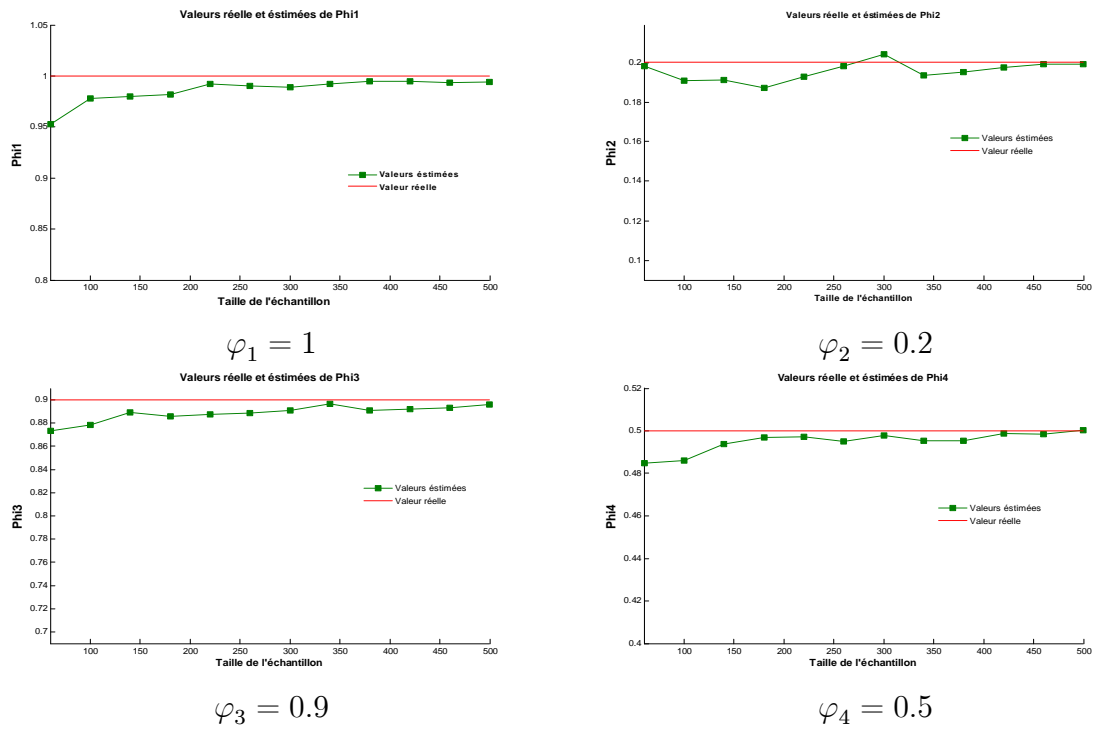


Table 3.4.7

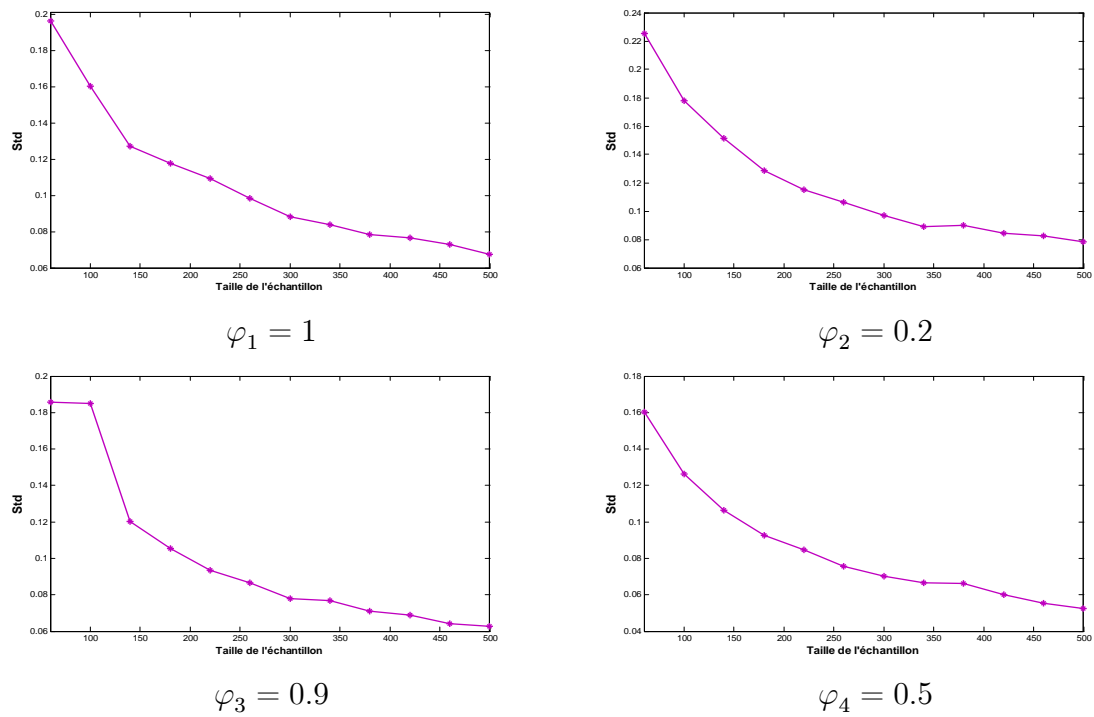


Table 3.4.8

d. 2) Dans la région de racine unitaire ($\Phi = (2, 5, 0.1, 1)'$)

Les graphes de la *Table 3.4.9* ci-dessous représentent l'évolution de l'estimateur, $\hat{\Phi}_n$, des moindres carrés du paramètres de $\Phi = (2, 5, 0.1, 1)'$, et ceux ce la *Table 3.4.10* représentent l'évolution de leur écart-type $\hat{\sigma}_{\hat{\Phi}_n}$, en fonction des séries temporelles simulées et commentées.

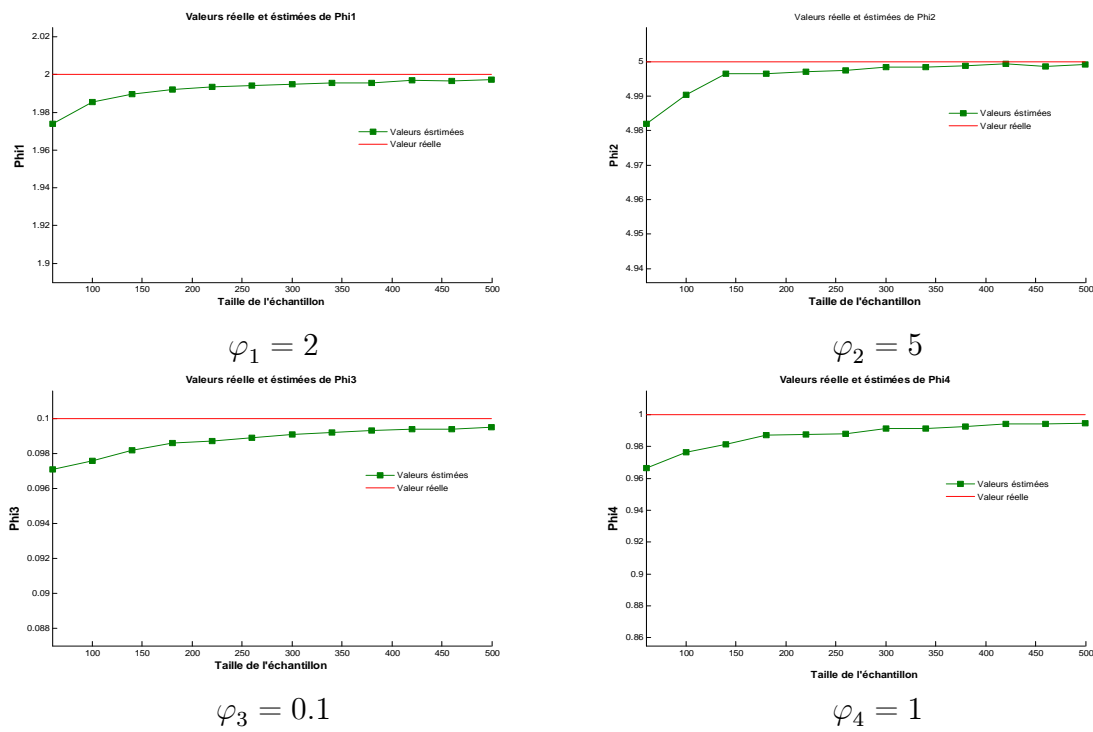


Table 3.4.9

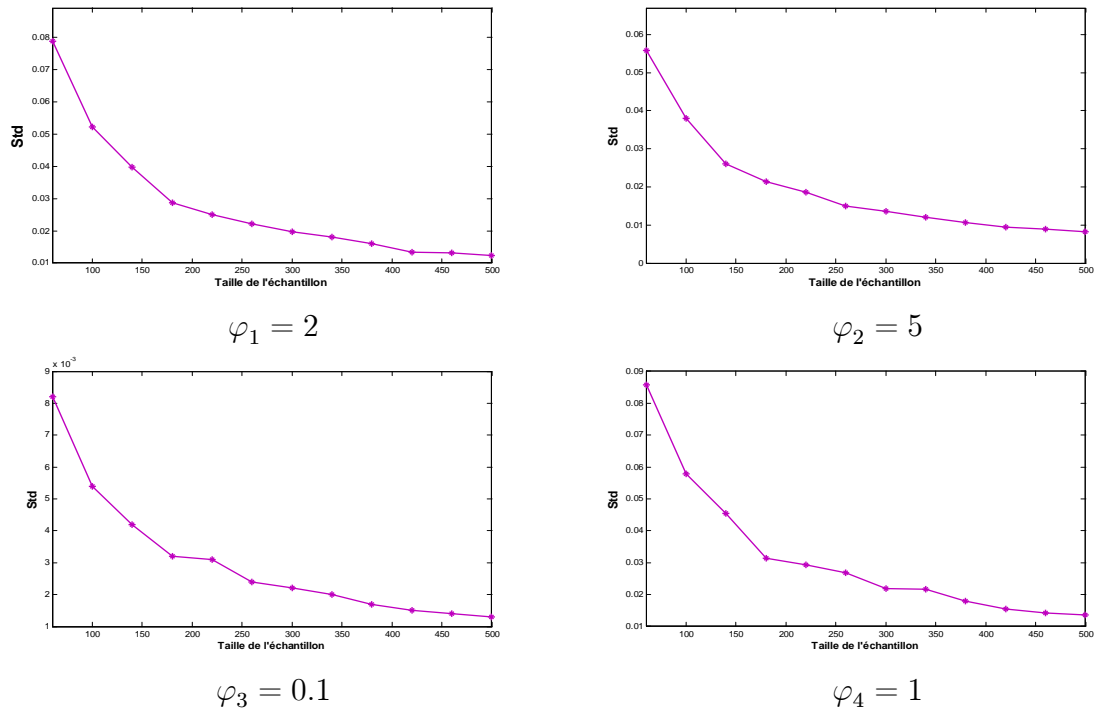


Table 3.4.10

d. 3) Dans la région explosif ($\Phi = (1, 2, 7.5, 0.1)'$)

Les graphes de la Table 3.4.11 ci-dessous représentent l'évolution de l'estimateur, $\hat{\Phi}_n$, des moindres carrés du paramètres de $\Phi = (1, 2, 7.5, 0.1)'$, et ceux ce la Table 3.4.12 représentent l'évolution de leur écart-type $\hat{\sigma}_{\hat{\Phi}_n}$, en fonction des séries temporelles simulées et commentées.

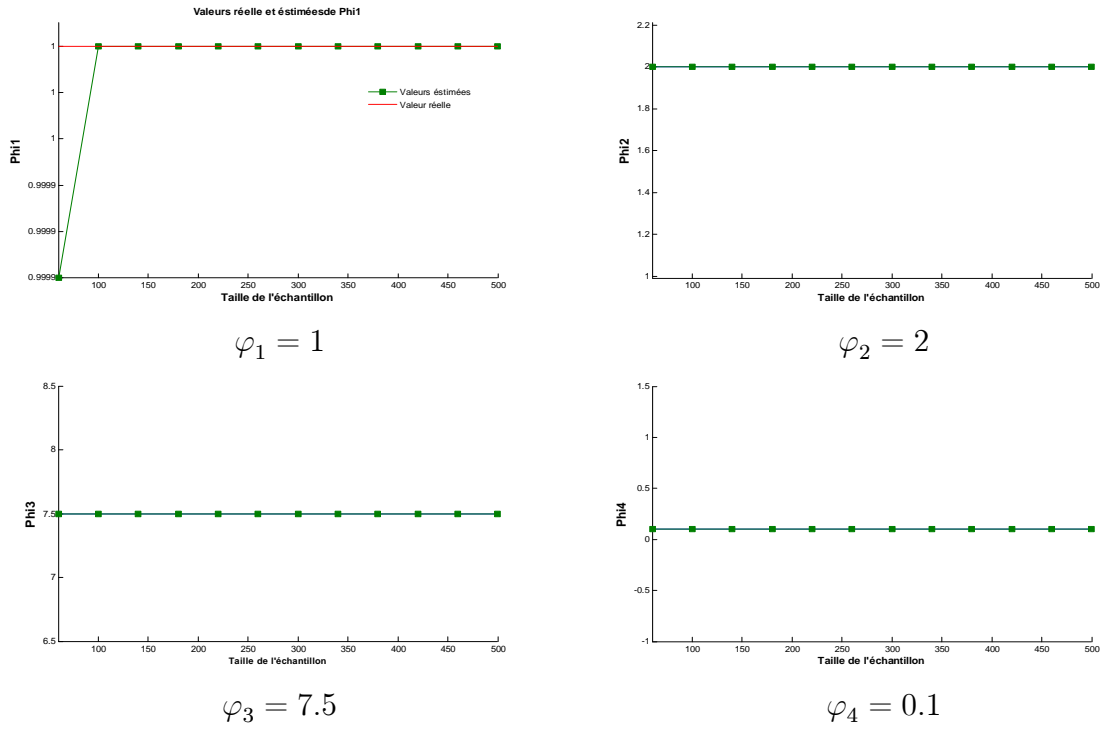


Table 3.4.11

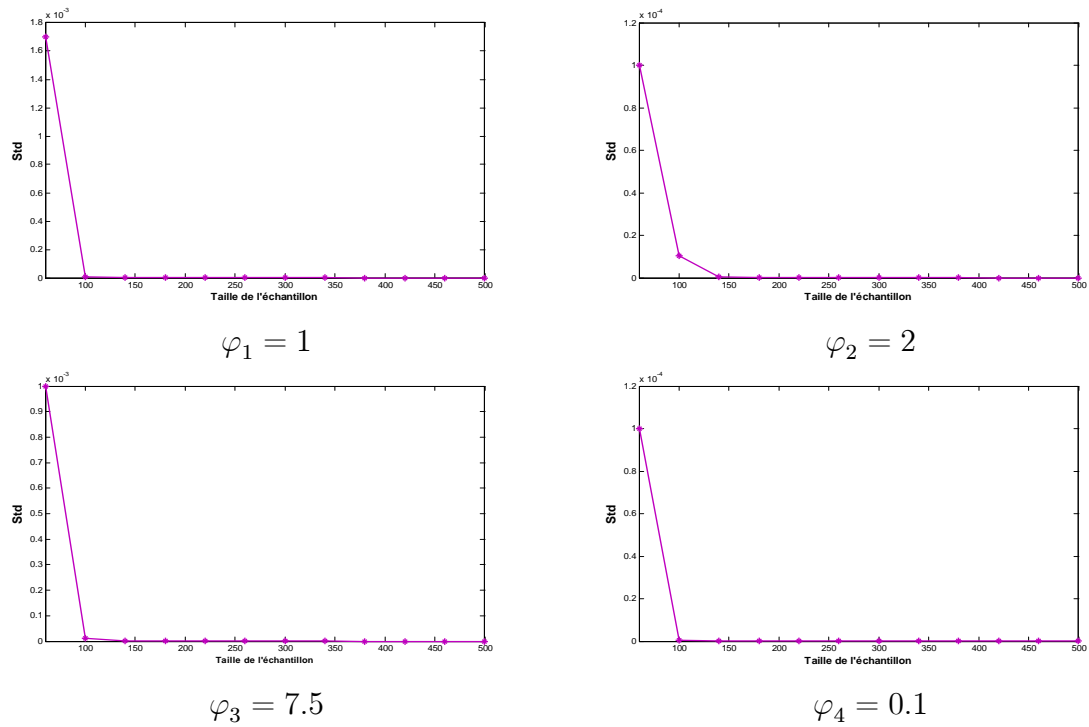


Table 3.4.12

Chapitre 4

Instabilité et Propriétés Asymptotiques des estimateurs des *MCO* d'un *ARCH*(1)

4.1 Introduction

Il est bien connu, de nos jours, que beaucoup de séries chronologiques rencontrées, dans une diversité de domaines en particulier environnemental et économie financière (rendement d'un actif, prix d'action, taux d'intérêt, taux de change de devises,...), révèlent quelques faits stylisés qui ne peuvent pas être capturés et décrits par des modèles linéaires *ARMA* familiarisés par le fameux ouvrage de Box et Jenkins (1970 et 1976).

L'objectif de ce chapitre est l'étude asymptotique des estimateurs des moindres carrés des paramètres du modèle *ARCH*(1). Avant cela, nous allons exposer les caractéristiques d'un modèle *ARCH*(q), par la suite nous passons à l'étude des caractéristiques d'un modèle *ARCH*(1). Et nous terminons par l'étude asymptotique des estimateurs des moindres carrés des paramètres du modèle *ARCH*(1).

4.2 Hétéroscédasticité Conditionnelle

Parmi ces caractéristiques qui ont fait et continu à faire intéresser les analystes des séries chronologiques nous citons le regroupement de la volatilité dont l'évidence empirique remonte aux années soixante (Mandelbort 1963 et Fama 1965). En effet, le phénomène de l'évidence empirique de la volatilité, qui caractérise les séries temporelles financières est clairement constaté et est bien mis en évidence, depuis l'année 1963, par Mandelbrot tout en précisant :

"Large changes tend to be followed by large changes of either sign and small changes tend to be followed by small changes".

Brièvement, la volatilité dans une série chronologique n'est que le phénomène consistant à ce que les grandes valeurs (en valeurs absolues) ont tendance à se regrouper (en formant une étape dite agitée) et les valeurs modestes (toujours en valeurs absolues) ont tendance à se regrouper formant l'étape dite calme.

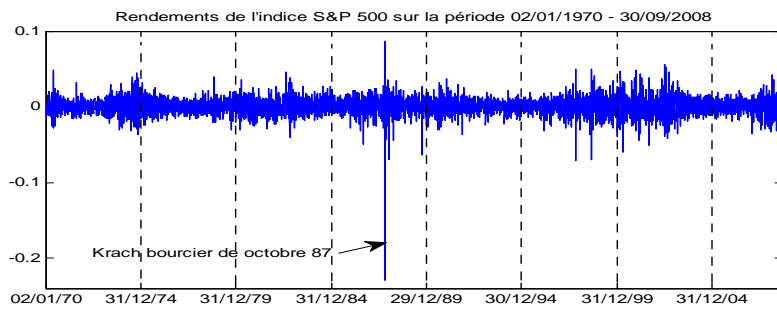
A titre d'exemple, la volatilité dans le domaine économique et financier peut être interprétée comme une mesure de l'instabilité du cours d'un actif financier ; elle mesure l'amplitude des variations d'une action, d'un produit dérivé ou d'un marché. En utilisant la terminologie financière, il s'agit d'un paramètre de quantification du risque de rendement et de prix. La volatilité sert également aux calculs pour optimiser la diversification des portefeuilles d'actifs financiers et l'évaluation des contrats financiers dérivés tels que les options. Les processus générateurs des séries chronologiques où la volatilité révèle, sont généralement non corrélés du départ bien que leurs carrés sont souvent fortement corrélés comme d'ailleurs les processus représentant les rendements en finance. Un exemple typique de ce type de séries chronologiques représentant l'évolution des rendements des actifs financiers qui sont, pour leur importance, fréquemment analysées. Soit p_t le prix d'un actif, à l'instant t , alors le rendement de cet actif est défini comme suit :

Définition 4.2.1.

Le rendement d'un actif financier, à l'instant t , noté r_t , est donné par :

$$r_t = \log \left(\frac{p_t}{p_{t-1}} \right) = \log(p_t) - \log(p_{t-1}). \quad (4.2.1)$$

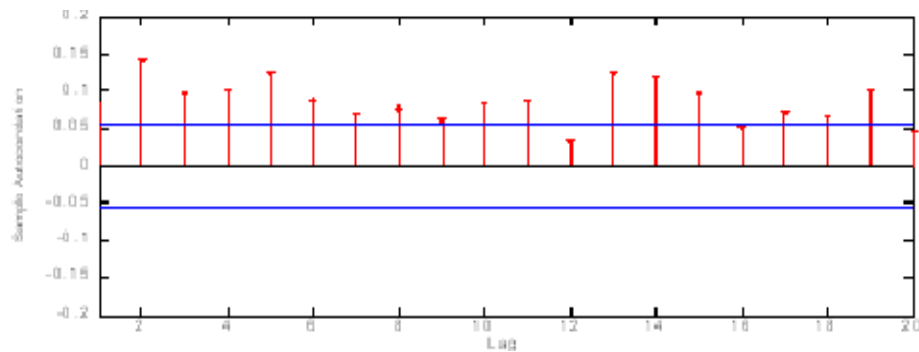
Le *Graph* 4.2.1, ci-dessous, visualise la trajectoire du rendement de l'Indice Boursier *SP500* du 02/01/1970 au 30/09/2008.



Grappe (4.2.1)

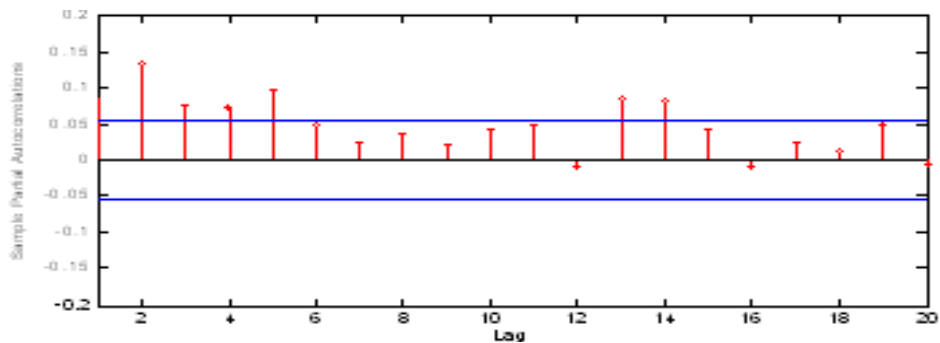
L'application des différents tests de l'hypothèse de la stationnarité, existant dans la littérature de l'analyse des séries chronologiques, indique que les processus des *rendements des actifs financiers*, donnés par (4.2.1), sont de la classe des processus stochastiques, du second ordre, *faiblement stationnaires*.

En dépit, du fait que le processus du rendement r_t est non corrélé, le processus r_t^2 est fortement corrélé (cette constatation est aussi correct pour le processus $|r_t|$).



fonction d'autocorrélation des carrés du rendement

et



fonction d'autocorrélation partielle des carrés du rendement

L'hypothèse de normalité des innovations est fréquemment imposée. Cependant, il est constaté, depuis longtemps, que les résidus des modèles ajustés aux séries temporelles financières ne semblent pas être en accord avec cette hypothèse alors qu'ils sont en conformité totale avec des lois de queues typiquement épaisses (lourdes) et même asymétriques.

Il est montré que la loi des résidus d'un modèle de volatilité stochastique ou déterministe ajusté, même avec innovations normales, est leptokurtique. Ainsi, les modèles de volatilité, outre qu'ils sont adéquats pour capturer la caractéristique de *volatilité clustering*, sont très convenables pour décrire l'excès de *Kurtosis* qui est un fait stylisé observé dans beaucoup de séries chronologiques économiques et financières.

Le fait que les grandes valeurs des carrés des rendements soient généralement précédées de grandes valeurs (indépendamment du signe des rendements) est difficilement compatible avec une variance conditionnelle constante. Ce phénomène est connu sous le nom d'hétéroscédasticité conditionnelle.

Pour ce type de séries temporelles les modèles linéaires Autorégressif Moyenne Mobile (*ARMA*) et même les modèles de la classe plus large *SARIMA* : modèle Autorégressif Moyenne mobile Intégré Saisonnier), se révèlent insuffisants : ils ne permettent pas la prise en compte des phénomènes de variabilités de la volatilité en fonction du temps et des mécanismes d'asymétries car ils sont fondés sur les deux premiers moments non conditionnels. En effet, dans la modélisation *ARMA*, l'espérance et la variance non conditionnelles ne changent pas. dans le temps, elles sont les deux statiques. Par exemple, soit un modèle *AR(1)* :

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t, \text{ où } \varepsilon_t \sim N(0, \sigma^2).$$

La moyenne et la variance inconditionnelles de X_t s'écrivent :

$$\begin{aligned} \mathbb{E}(X_t) &= \mu = \frac{\alpha_0}{1 - \alpha_1} \\ \text{Var}(X_t) &= \frac{\sigma^2}{1 - \alpha_1^2}. \end{aligned}$$

Toute fois l'espérance conditionnelle est dynamique (elle évolue dans le temps) par contre la variance conditionnelle reste encore statique :

$$\begin{aligned} \mathbb{E}(X_t/X_{t-1}) &= \alpha_0 + \alpha_1 X_{t-1}, \\ \text{Var}(X_t) &= \mathbb{E}[(X_t - \mathbb{E}(X_t/X_{t-1}))^2/X_{t-1}], \\ &= \mathbb{E}[\varepsilon_t^2/X_{t-1}] = \mathbb{E}(\varepsilon_t^2) = \sigma^2. \end{aligned}$$

Pour remédier aux insuffisances des modèles classiques, une nouvelle classe de séries chronologiques a été introduite dans la littérature des séries chronologiques. Il s'agit de la classe des modèles Autorégressifs Conditionnels Hétéroscédastiques (*ARCH*) introduite Engle en 1982. Le modèle proposé dans cet article a été employé pour estimer les moyennes et les variances de l'inflation au Royaume-Uni. Engle a élaboré ce modèle après sa visite à l'école de Londres des sciences économiques. Cette recherche s'est vue couronnée par le prix décerné par l'académie Nobel à Engle en 2003. Pour une bonne compréhension du fondement théorique de l'innovation révolutionnaire de Robert Engle (1982), il semble nécessaire de rappeler l'une des plus remarquables innovations en analyse des séries chronologiques, à savoir la décomposition d'un processus faiblement stationnaire due à Wold (1938).

Théorème 4.4.1 (*Décomposition de Wold* 1938)

Tout processus stochastique $\{X_t, t \in \mathbb{Z}\}$, du second ordre, faiblement stationnaire de moyenne μ , possède une décomposition unique donnée par :

$$X_t = \mu + \sum_{s=-\infty}^t \alpha_s \varepsilon_s, \quad t \in \mathbb{Z},$$

où la somme infinie est convergente en moyenne quadratique et par conséquent le processus du second ordre, $\{\alpha_0 \varepsilon_t, t \in \mathbb{Z}\}$, est un processus d'Innovation et le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un processus d'Innovation Normé.

Symboliquement un processus purement indéterminable, faiblement stationnaire de moyenne μ , est donnée par :

$$X_t = \mu + \sum_{s=-\infty}^t \alpha_s \varepsilon_s, \quad t \in \mathbb{Z}, \quad (4.4.1)$$

où le polynôme moyenne mobile, d'ordre infini, $B(L)$, où L désigne l'opérateur de retard, i.e., $L^j Y_t = Y_{t-j}$, est donné par :

$$B(L) = \sum_{j=0}^{\infty} \alpha_j L^j, \quad \text{avec } (\alpha_0 = 1),$$

satisfait la condition de *Wold* (*convergence en quadratique moyenne*).

Un simple calcul montre que la moyenne et la variance non conditionnelles du processus $\{X_t, t \in \mathbb{Z}\}$ sont constantes et données par :

$$\mathbb{E}(X_t) = \mu \quad \text{et} \quad \text{Var}(X_t) = \mathbb{E}(X_t - \mu)^2 = \sigma^2 \sum_{j=0}^{\infty} \alpha_j^2. \quad (4.4.2)$$

Par contre, la moyenne et la variance conditionnelle de X_t (conditionnellement à l'informations acquises jusqu'à l'instant $t - 1$, $I_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$) sont clairement données par

$$\mathbb{E} \left(X_t / I_{t-1} \right) = \mu + \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j},$$

et

$$Var \left(X_t / I_{t-1} \right) = \mathbb{E} \left(X_t - \mathbb{E} \left(X_t / I_{t-1} \right) \right)^2 = \sigma^2. \quad (4.4.3)$$

- Les formules (4.4.2) montrent que le processus stochastique donné par la décomposition de Wold est faiblement stationnaire. De plus ces deux premiers moments sont, symétriquement, invariants dans le temps.
- La première formule de (4.4.3) montre clairement que la moyenne conditionnelle s'adapte aux informations acquises (donc il est dynamique) par contre la formule de la variance conditionnelle (deuxième formule de (4.4.3)) ne se dote pas de cet avantage. L'inconvénient de cette formulation statique c'est qu'elle ne permet pas l'incorporation des informations passées qui permettent de réduire l'incertitude portant sur les conclusions statistiques tirées et par la suite des décisions à prendre.

Pour remédier à ce problème, il a été visé comme objectif le passage d'un modèle statique à un modèle dynamique stochastique. Ce passage permet l'incorporation des observations, qui ne sont que porteuses des informations acquises au cours du cheminement, d'un processus stochastique représentant l'évolution, dans le temps, du phénomène sous-jacent. Cet objectif est, en dépit de sa simplicité et sa clarté, fondamental pour une étude profonde permettant des représentations dynamiques élégantes des séries chronologiques économiques en particulier les financières.

Engle a exploité, d'une façon élégante et très significative, le fait bien connu concernant la *corrélation sérielle* des carrés des Innovations Normées, à savoir le processus *non observable* $\{\varepsilon_t^2, t \in \mathbb{Z}\}$, pour inférer statistiquement le comportement de la variance conditionnelle (conditionnellement à tous le passé, à I_{t-1}) qui n'est que l'objet mathématique représentant la volatilité constatée. Pratiquement, il considère le processus conditionnel suivant :

$$\varepsilon_t / I_{t-1} \sim N(0, h_t), \quad (4.4.4)$$

et, en suivant l'esprit de Wold, il exprime cette variance conditionnelle par une combinaison linéaire, en $\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \varepsilon_{t-3}^2, \dots$, de la forme suivante

$$h_t = \alpha_0 + \sum_{j=1}^{\infty} \alpha_j \varepsilon_{t-j}^2,$$

où $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, 3, \dots$ et enfin $\sum_{j=1}^{\infty} \alpha_j < 1$. (4.4.5)

Le modèle le plus riche ainsi construit est caractérisé par la moyenne et la variance conditionnelle données par :

$$\begin{aligned} \mathbb{E}(X_t/I_{t-1}) &= \mu + \sum_{j=1}^{\infty} \alpha_j \varepsilon_{t-j}, \\ \text{et} \quad \text{Var}(X_t/I_{t-1}) &= h_t = \alpha_0 + \sum_{j=1}^{\infty} \alpha_j \varepsilon_{t-j}^2. \end{aligned} \quad (4.4.6)$$

L'importance remarquable de ce modèle réside dans le fait qu'il reflète à la fois le dynamisme stochastique dans l'espérance conditionnelle ainsi que dans la variance conditionnelle.

Remarques

— **R1** : Il est facile de vérifier que sous ces conditions le processus est faiblement stationnaire et admettant une décomposition de Wold. De plus, l'espérance et la variance non conditionnelles de ce processus sont constantes et données par :

$$\mathbb{E}(X_t) = \mu \quad \text{et} \quad \text{Var}(X_t) = \frac{\alpha_0}{1 - \sum_{j=1}^{\infty} \alpha_j^2}.$$

Par contre l'espérance et la variance conditionnelles sont conjointement dynamique et exprimées par (4.4.6).

— **R2** : On remarque l'expression de la variance conditionnelle ou encore, selon le vocabulaire financier, la volatilité h_t contient une infinité de paramètres inconnus à estimer, ce qui n'est pas pratiquement faisable. Néanmoins, on peut obtenir, sous la condition $\sum_{j=1}^{\infty} \alpha_j < 1$ une approximation d'ordre fini q , pour h_t , donnée par $h_t = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2$.

Le modèle standard ainsi construit, par Engle (1982), pour capturer la volatilité instantanée caractérisant les séries temporelles financières, dit *Modèle Autorégressif Conditionnellement Hétéroscédastique* d'ordre q , (*Autoregressive Conditional Heteroskedastic Model*) noté *ARCH*(q), est donné par :

$$\varepsilon_t = \sqrt{h_t} \eta_t, \quad \eta_t \sim N(0, \sigma^2)$$

où $\varepsilon_t / I_{t-1} \sim N(0, h_t)$ et la variance conditionnelle h_t est donnée par :

$$h_t = \mathbb{E}(\varepsilon_t^2 / I_{t-1}) = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2.$$

4.3 Etude des Modèles ARCH

Durant la dernière décennie, les études appliquées au finance et en économie monétaire utilisent explicitement des modélisations de séries temporelles faisant intervenir des moments du second ordre ou d'ordre supérieur variant dans le temps. En comblant le manquement apparent des structures dynamiques existantes, dans la prise en compte des regroupements de la volatilité, la classe des modèles Autorégressif Conditionnellement Hétéroscédastique (*ARCH*), introduite par Engle (1982), contribue au développement et à la refonte de la théorie financière de façon à inclure la dépendance temporelle de la volatilité. Quoique, cette nouvelle classe de modèle se trouve encore dans une phase expérimentale et en pleine évolution, on assiste, d'ores et déjà, à des extensions et utilisations diverses (modèles *GARCH*, *ARMA* à erreur *ARCH*, ...).

Objectivement, le large développement des prévisions, due au série chronologique, dérive de l'utilisation de la moyenne conditionnée par le passé du processus. Pour de réelle série financière, il est évident d'admettre, qu'à afin d'obtenir un meilleur intervalle de prévision, il suffit d'adjoindre des informations supplémentaires du passé à la prédiction de la variance. Le modèle *ARCH* fut la réponse à cette réflexion, et sa conséquence fut immédiate, puisqu'il donna, ainsi, un nouveau souffle à la macro-économie. L'approche la plus utilisée jusqu'alors, pour modéliser l'hétéroscédasticité était d'introduire une variable exogène, afin de prédire la variance de la variable endogène. Cette solution présente le désavantage de requérir la détermination des causes pouvant induire une variance non constante et, de plus, elle ne prend pas en considération le fait que les moyennes et variances conditionnelles puissent évoluer de manière jointe dans le temps. En supposant la normalité des innovations, un modèle préférable est donné par Engle (1982), permettant à la variance conditionnelle de changer dans le temps, tout en préservant la variance non conditionnelle constante :

$$\varepsilon_t = \eta_t \sqrt{h_t}, \quad (4.5.1a)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2, \quad (4.5.1b)$$

$$\text{et } h_t \text{ i.i.d. } \rightsquigarrow N(0, 1) \quad (4.5.1c)$$

Ce modèle est ce qui a été appelé par Engle (1982) modèle *ARCH* (q). Il apparaît comme un bruit blanc gaussien, multiplié à chaque temps t par une variable aléatoire dont le carré dépend de manière linéaire des q dernières valeurs du processus. Comme on peut l'écrire sous la forme :

$$\varepsilon_t / I_{t-1} \rightsquigarrow N(0, h_t), \quad (4.5.2a)$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2, \quad (4.5.2b)$$

où I_{t-1} désigne l'ensemble d'information disponible à la date $t - 1$.

Ce qui montre que la distribution conditionnelle de ε_t est normale centrée, de variance h_t , laquelle dépend linéairement des q dernières valeurs du processus. Enfin, il vient une dernière formulation de ε_t^2 en posant $u_t = \varepsilon_t^2 - h_t$:

$$\varepsilon_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 + u_t, \quad (4.5.3)$$

Ce qui montre une forme $AR(q)$ pour les carrés du processus d'innovation ε_t^2 , avec toutefois, la remarque que les variables :

$$\varepsilon_t^2 = \eta_t^2 h_t - h_t = (\eta_t^2 - 1) h_t, \quad (4.5.4)$$

ne sont pas de variance constante mais sont de moyenne nulle et non corrélées entre elles.

Note 4.5.1

On note que la dernière formulation donne lieu à des définitions plus générales de modèles $ARCH$ où la normalité n'est pas requise.

Note 4.5.2

Il faut évidemment que $\alpha_0 > 0$ et que $\alpha_1, \alpha_2, \dots, \alpha_q \geq 0$, sans quoi la variance conditionnelle pourrait être négative.

4.3.1 Etude d'un processus ARCH(1)

Considérons le processus $ARCH(1)$, satisfaisant l'équation :

$$\varepsilon_t / I_{t-1} \rightsquigarrow N(0, h_t), \quad (4.5.5a)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad (4.5.5b)$$

Remarquons qu'en vertu du théorème des projections itérées [Cf. Dreesbeke et *al.* (1994)], et compte tenu du fait que l'information contenue dans l'espace engendré par $\{\dots, \varepsilon_{t-h-1}, \varepsilon_{t-h}, \varepsilon_{t-h+1}, \dots, \varepsilon_{t-1}\}$ étant plus consistante que celle contenue dans $\{\dots, \varepsilon_{t-h-1}, \varepsilon_{t-h}\}$ i.e.

$\{\varepsilon_{t-h} \subset \varepsilon_{t-1}\}$, on obtient la propriété suivante $\mathbb{E}(\varepsilon_t / I_{t-h}) = \mathbb{E}(0 / I_{t-h}) = 0, \forall h > 0$.

Moments d'ordre r du processus ARCH(1)

Propriété 4.5.1

Les moments centrés d'ordre impair, s'ils existent, sont nuls, par symétrie.

Propriété 4.5.2 (Engle (1982))

En supposant que le processus démarre infiniment loin dans le passé avec les $2r$ premiers moments finis, le moment d'ordre $2r$ existe, si et seulement si,

$$\alpha_1^r \prod_{j=1}^r (2j - 1) < 1.$$

De plus le moment d'ordre $2k$ ($1 \leq k \leq r$) est donné par le k ième élément du vecteur

$$(I - B_r)^{-1} b_{0,r},$$

c'est-à-dire $\mu_{2k} = \mathbb{E}(\varepsilon^{2k}) = ((I - B_r)^{-1} b_{0,r})_k$, où la matrice, $(r \times r)$, B_r et le vecteur, $(r \times 1)$ sont donnés par :

$$B_r = \begin{pmatrix} \alpha_1^r \theta_r & r \alpha_0 \alpha_1^{r-1} \theta_r & \binom{r}{2} \alpha_0^2 \alpha_1^{r-2} \theta_r & \cdots & r \alpha_0^{r-1} \alpha_1 \theta_r \\ 0 & \alpha_1^{r-1} \theta_{r-1} & (r-1) \alpha_0 \alpha_1^{r-2} \theta_{r-1} & \cdots & (r-1) \alpha_0^{r-2} \alpha_1^2 \theta_{r-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_1^{r-k} \theta_{r-k} & \cdots & (r-k) \alpha_0^{r-k-1} \alpha_1 \theta_{r-k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_1 \theta_1 \end{pmatrix}$$

et

$$b_{0,r} = \left(\alpha_0^r \times \prod_{j=1}^r (2j - 1), \alpha_0^{r-1} \times \prod_{j=1}^{r-1} (2j - 1), \alpha_0^{r-2} \times \prod_{j=1}^{r-2} (2j - 1), \dots, \alpha_0 \times \prod_{j=1}^1 (2j - 1) \right)'$$

Démonstration. Rappelons que pour une variable aléatoire X suivant une loi normale de moyenne nulle, et de variance σ^2 , le moment d'ordre pair est donné par :

$$\mathbb{E}(X^{2m}) = \sigma^{2m} \prod_{j=1}^m (2j - 1).$$

Sachant que la distribution conditionnelle du processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est normale, on a :

$$\begin{aligned} \mathbb{E}\left(\varepsilon_t^{2r} / \varepsilon_{t-1}\right) &= (\sqrt{h_t})^{2r} \prod_{j=1}^r (2j - 1) = (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^r \prod_{j=1}^r (2j - 1), \\ &= \left\{ \sum_{i=0}^{i=r} \binom{r}{i} \alpha_1^{r-i} \varepsilon_{t-1}^{2(r-i)} \alpha_0^i \right\} \prod_{j=1}^r (2j - 1), \end{aligned}$$

$$= \left\{ \binom{r}{0} \alpha_1^{r-0} \varepsilon_{t-1}^{2(r-0)} \alpha_0^0 + \binom{r}{1} \alpha_1^{r-1} \varepsilon_{t-1}^{2(r-1)} \alpha_0^1 + \binom{r}{2} \alpha_1^{r-2} \varepsilon_{t-1}^{2(r-2)} \alpha_0^2 + \dots + \binom{r}{r-1} \alpha_1 \varepsilon_{t-1}^2 \alpha_0^{r-1} + \alpha_0^{r-1} \right\} \prod_{j=1}^r (2j-1).$$

En appliquant l'égalité pour les ordres décroissants $r, r-1, r-2, \dots, 1$, on obtient les r égalités suivantes :

$$\begin{aligned} \mathbb{E} \left(\varepsilon_t^{2r} / \varepsilon_{t-1} \right) &= \left\{ \alpha_1^{r-0} \varepsilon_{t-1}^{2(r-0)} + r \alpha_1^{r-1} \varepsilon_{t-1}^{2(r-1)} \alpha_0^1 + \binom{r}{2} \alpha_1^{r-2} \varepsilon_{t-1}^{2(r-2)} \alpha_0^2 + \dots \right. \\ &\quad \left. + \binom{r}{r-1} \alpha_1 \varepsilon_{t-1}^2 \alpha_0^{r-1} + \alpha_0^r \right\} \prod_{j=1}^r (2j-1) \\ \mathbb{E} \left(\varepsilon_t^{2(r-1)} / \varepsilon_{t-1} \right) &= \left\{ \alpha_1^{r-1} \varepsilon_{t-1}^{2(r-1)} + (r-1) \alpha_1^{r-2} \varepsilon_{t-1}^{2(r-2)} \alpha_0^1 + \dots + \right. \\ &\quad \left. \binom{r-1}{r-2} \alpha_1 \varepsilon_{t-1}^2 \alpha_0^{r-2} + \alpha_0^{r-1} \right\} \prod_{j=1}^{r-1} (2j-1) \\ &\vdots \\ \mathbb{E} \left(\varepsilon_t^{2(r-k)} / \varepsilon_{t-1} \right) &= \left\{ \alpha_1^{r-k} \varepsilon_{t-1}^{2(r-k)} + (r-k) \alpha_1^{r-k-1} \varepsilon_{t-1}^{2(r-k-1)} \alpha_0^1 + \dots + \right. \\ &\quad \left. \binom{r-k}{r-k-1} \alpha_1 \varepsilon_{t-1}^2 \alpha_0^{r-k-1} + \alpha_0^{r-k} \right\} \prod_{j=1}^{r-k} (2j-1) \\ &\vdots \\ \mathbb{E} \left(\varepsilon_t^2 / \varepsilon_{t-1} \right) &= (\alpha_1 \varepsilon_{t-1}^2 + \alpha_0) \prod_{j=1}^1 (2j-1) \end{aligned}$$

En définissant les vecteur colonnes $(r \times 1)$ $W_t = \left(\varepsilon_t^{2r}, \varepsilon_t^{2(r-1)}, \varepsilon_t^{2(r-2)}, \dots, \varepsilon_t^2 \right)'$,

$$b_{0,r} = \left(\alpha_0^r \times \prod_{j=1}^r (2j-1), \alpha_0^{r-1} \times \prod_{j=1}^{r-1} (2j-1), \alpha_0^{r-2} \times \prod_{j=1}^{r-2} (2j-1), \dots, \alpha_0 \times \prod_{j=1}^1 (2j-1) \right)'$$

et en notant $\theta_r = \prod_{j=1}^r (2j-1)$, on peut réécrire les égalités précédentes sous la forme matricielle suivante :

$$\mathbb{E} \left\{ \left(\begin{array}{c} \varepsilon_t^{2r} \\ \varepsilon_t^{2(r-1)} \\ \varepsilon_t^{2(r-2)} \\ \vdots \\ \varepsilon_t^4 \\ \varepsilon_t^2 \end{array} \right) / \varepsilon_{t-1} \right\} = \begin{pmatrix} \alpha_0^r \times \theta_r \\ \alpha_0^{r-1} \times \theta_{r-1} \\ \alpha_0^{r-2} \times \theta_{r-2} \\ \vdots \\ \alpha_0^{r-k} \times \theta_{r-k} \\ \vdots \\ \alpha_0 \times \theta_1 \end{pmatrix} +$$

$$\begin{pmatrix} \alpha_1^r \theta_r & r \alpha_0 \alpha_1^{r-1} \theta_r & \binom{r}{2} \alpha_0^2 \alpha_1^{r-2} \theta_r & \cdots & \binom{r}{r-1} \alpha_0^{r-1} \alpha_1 \theta_r \\ 0 & \alpha_1^{r-1} \theta_{r-1} & (r-1) \alpha_0 \alpha_1^{r-2} \theta_{r-1} & \cdots & \binom{r-1}{r-2} \alpha_0^{r-2} \alpha_1^2 \theta_{r-1} \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & \alpha_1^{r-k} \theta_{r-k} & \cdots & \binom{r-k}{r-k-1} \alpha_0^{r-k-1} \alpha_1 \theta_{r-k} \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & \cdots & \alpha_1 \theta_1 \end{pmatrix} \times \begin{pmatrix} \varepsilon_{t-1}^{2r} \\ \varepsilon_{t-1}^{2(r-1)} \\ \varepsilon_{t-1}^{2(r-2)} \\ \varepsilon_t \\ \vdots \\ \varepsilon_{t-1}^4 \\ \varepsilon_{t-1}^2 \end{pmatrix}.$$

Soit la matrice carrée, $(r \times r)$,

$$B_r = \begin{pmatrix} \alpha_1^r \theta_r & r \alpha_0 \alpha_1^{r-1} \theta_r & \binom{r}{2} \alpha_0^2 \alpha_1^{r-2} \theta_r & \cdots & r \alpha_0^{r-1} \alpha_1 \theta_r \\ 0 & \alpha_1^{r-1} \theta_{r-1} & (r-1) \alpha_0 \alpha_1^{r-2} \theta_{r-1} & \cdots & (r-1) \alpha_0^{r-2} \alpha_1^2 \theta_{r-1} \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & \alpha_1^{r-k} \theta_{r-k} & \cdots & (r-k) \alpha_0^{r-k-1} \alpha_1 \theta_{r-k} \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & \cdots & \alpha_1 \theta_1 \end{pmatrix}$$

alors, on a :

$$\mathbb{E} \left(W_t / \varepsilon_{t-1} \right) = B_r W_{t-1} + b_{0,r},$$

En prenant l'espérance (mathématique) et en notant l'espérance non conditionnelle, $\underline{\mu}_t$, de W_t , c'est-à-dire $\underline{\mu}_t = \mathbb{E}(W_t)$, on obtient l'équation aux différences en $\underline{\mu}_t$ suivante :

$$\underline{\mu}_t = B_r \underline{\mu}_{t-1} + b_{0,r}.$$

On note que la matrice et le vecteur ne dépendent pas du temps, t .

En récurant l'équation précédente m fois, on obtient :

$$\underline{\mu}_t = (B_r)^m \underline{\mu}_{t-m} + \left(I + \sum_{j=1}^{m-1} (B_r)^j \right) b_{0,r}.$$

En prenant $m = t$, on obtient

$$\underline{\mu}_t = (B_r)^t \underline{\mu}_0 + \left(I + \sum_{j=1}^{t-1} (B_r)^j \right) b_{0,r}.$$

Pour assurer la convergence (en moyenne quadratique) la matrice B_r devrait avoir toutes racines à l'intérieur du cercle d'unité. Il est, du fait que B_r est triangulaire, facile de vérifier que les racines (positives) sont :

$$\alpha_1^r \prod_{j=1}^r (2j-1), \alpha_1^{r-1} \prod_{j=1}^{r-1} (2j-1), \dots, \alpha_1.$$

Alors, la condition d'existence du moment d'ordre $2r$ est donnée par :

$$\alpha_1^r \prod_{j=1}^r (2j-1) < 1.$$

Sous cette condition le vecteur $\underline{\mu} = \left(\mathbb{E}(\varepsilon_t^{2r}), \mathbb{E}(\varepsilon_t^{2(r-1)}), \dots, \mathbb{E}(\varepsilon_t^2) \right)'$ des moments peut être donné par l'expression suivante :

$$\underline{\mu} = \left(\sum_{j=1}^{\infty} (B_r)^j \right) b_{0,r} = (I - B_r)^{-1} b_{0,r},$$

où I est une matrice unitaire d'ordre $(r \times r)$. ■

Variance et moment d'ordre 4 d'un ARCH(1)

Propriété 4.5.3

Une condition nécessaire et suffisante pour l'existence du moment d'ordre 4 d'un processus stochastique, $\{\varepsilon_t, t \in \mathbb{Z}\}$, satisfaisant un modèle ARCH(1) est donnée par :

$$3\alpha_1^2 < 1.$$

De plus, sous cette condition, le moment d'ordre 2 (variance non conditionnelle) et le moment d'ordre 4 sont, explicitement, donnés par :

$$\mathbb{E}(\varepsilon_t^2) = \text{Var}(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1} \text{ et } \mathbb{E}(\varepsilon_t^4) = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - 3\alpha_1^2)(1 - \alpha_1)}.$$

Démonstration. On a :

$$\varepsilon_t = \eta_t \sqrt{h_t} \text{ où } \eta_t \rightsquigarrow N(0, 1),$$

$$\text{et } h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2.$$

Alors, en appliquant la *Propriété* 4.5.2, pour $r = 1$, nous obtenons :

$$\mathbb{E}(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1}.$$

Pour calculer $\mathbb{E}(\varepsilon_t^4)$, nous appliquons le même théorème, pour $r = 2$. Alors l'existence du moment $\mathbb{E}(\varepsilon_t^4)$ est assurée par $1 - 3\alpha_1^2 > 0$, et il est donné par :

$$\mathbb{E}(\varepsilon_t^4) = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - 3\alpha_1^2)(1 - \alpha_1)}. \quad \blacksquare$$

La propriété suivante exprime l'expression explicite du coefficient d'aplatissement (kurtosis coefficient) (selon la définition de Fisher) du processus d'innovation d'un processus ARCH(1).

Propriété 4.5.4 (*Kurtosis marginale*)

Le coefficient d'aplatissement de Fisher d'un processus stochastique, possédant le moment d'ordre 4, satisfaisant un modèle ARCH(1) est donné par :

$$K = \frac{\mu_4}{(\mu_2)^2} = \frac{\mathbb{E}(\varepsilon_t^4)}{(\mathbb{E}(\varepsilon_t^2))^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2}.$$

Démonstration. La preuve découle directement des résultats précédents. \blacksquare

Remarque 4.5.1

On montre que le Kurtosis de l'innovation d'un ARCH(1) est strictement supérieur à celui d'une loi normale qui est égal à 3. En effet, on a :

$$K = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2} = \frac{3\left((1 - 3\alpha_1^2) + 2\alpha_1^2\right)}{1 - 3\alpha_1^2} = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2} > 3.$$

Donc la densité de la loi de probabilité du processus d'innovation d'un ARCH(1) est plus aplatie que celle de Gauss. Les queues de cette loi sont donc plus épaisses (plus ouvertes) que celles d'une loi normale. Une loi possédant cette propriété est dite *leptokurtique*.

4.4 Stationnarité d'un modèle ARCH(1)

Nous allons chercher quelles sont les conditions sous lesquelles il existe des processus stationnaires (au sens strict et au second-ordre) vérifiant l'équation (4.6.1b) ci-dessous :

$$\varepsilon_t = \eta_t \sqrt{h_t}, \quad (4.6.1a)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad (4.6.1b)$$

$$\text{et } \eta_t \text{ i.i.d. } \rightsquigarrow N(0, 1) \quad (4.6.1c)$$

On s'intéresse en particulier aux solutions non anticipatives du modèle (4.6.1), c'est à dire aux processus (ε_t) tels que ε_t soit une fonction mesurable des variables η_{t-s} , $s \geq 0$. Pour ces processus, $\sqrt{h_t}$ est indépendante de la tribu engendrée par $\{\eta_{t+i}, i \geq 0\}$ et ε_t est indépendante de la tribu engendrée par $\{\eta_{t+i}, i > 0\}$. De telles solutions sont également ergodiques.

4.4.1 Stationnarité stricte

Théorème 4.6.1 (Stationnarité stricte du ARCH(1))

Si

$$-\infty \leq \gamma := \mathbb{E}[\log(\alpha_1 \eta_t^2)] < 0,$$

la série

$$A_t = \alpha_0 \left(1 + \sum_{i=1}^{\infty} a(\eta_{t-1}) \dots a(\eta_{t-i})\right) \text{ où } a(\eta_t) = \alpha_1 \eta_t^2.$$

converge presque sûrement et le processus (ε_t) défini par $\sqrt{A_t} \eta_t$ est l'unique solution strictement stationnaire du modèle (4.6.1). Cette solution est non anticipative et ergodique. Si $\gamma \geq 0$ et $\alpha_0 > 0$, il n'existe pas de solution strictement stationnaire.

Démonstration. Remarquons que $\gamma = \mathbb{E}[\log(a(\eta_t))]$ existe toujours dans $[-\infty, +\infty[$ car $\mathbb{E}[\log^+(a(\eta_t))] \leq \mathbb{E}[a(\eta_t)] = \alpha_1$

où $\log^+(a(\eta_t)) = \text{Max}[\log(a(\eta_t)), 0]$.

En remplaçant dans (4.6.1b) ε_{t-1}^2 par $\eta_{t-1}^2 h_{t-1}$ on obtient :

$$h_t = \alpha_0 + \alpha_1 \eta_{t-1}^2 h_{t-1}$$

En posant $\alpha_1 \eta_{t-1}^2 = a(\eta_{t-1})$ on obtient :

$$h_t = \alpha_0 + a(\eta_{t-1}) h_{t-1} \quad (4.6.2)$$

Utilisons de manière récursive l'équation (4.6.2), on obtient pour $N \geq 1$:

$$h_t = \underbrace{\alpha_0 \left[1 + \sum_{n=1}^N a(\eta_{t-1}) \dots a(\eta_{t-n}) \right]}_{A_t(N)} + a(\eta_{t-1}) \dots a(\eta_{t-N-1}) h_{t-N-1}$$

$$h_t := A_t(N) + a(\eta_{t-1}) \dots a(\eta_{t-N-1}) h_{t-N-1} \quad (4.6.3)$$

Le processus limite $A_t = \lim_{N \rightarrow \infty} A_t(N)$ existe dans \mathbb{R}^+ puisque les termes de la somme sont positifs. On peut écrire $A_t(N)$ sous la forme :

$$A_t(N) = \alpha_0 + a(\eta_{t-1}) A_{t-1}(N-1)$$

Alors, $\lim_{N \rightarrow \infty} A_t(N) = \alpha_0 + a(\eta_{t-1}) A_{t-1} = A_t$

Nous allons à présent montrer que A_t est presque sûrement finie si et seulement si $\gamma = \mathbb{E}[\log(a(\eta_t))] < 0$.

– Supposons que $\gamma < 0$.

Soit la série de terme positifs $\sum_{n=1}^N a(\eta_{t-1}) \dots a(\eta_{t-n})$, en appliquant la règle de *Cauchy*¹ sur cette série on obtient :

$$\lim_{n \rightarrow \infty} (a(\eta_{t-1}) \dots a(\eta_{t-n}))^{\frac{1}{n}} = \lim_{n \rightarrow \infty} \left[\text{Exp} \left(\frac{1}{n} \sum_{i=1}^n \log(a(\eta_{t-i})) \right) \right].$$

En utilisant le théorème ergodique on a :

$$\frac{1}{n} \sum_{i=1}^n \log(a(\eta_{t-i})) \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E}[\log(a(\eta_t))]$$

Alors

$$\text{Exp} \left[\frac{1}{n} \sum_{i=1}^n \log(a(\eta_{t-i})) \right] \xrightarrow[n \rightarrow \infty]{p.s.} \text{Exp} \{ \mathbb{E}[\log(a(\eta_t))] \} = \text{Exp}(\gamma)$$

Nous avons supposé que $\gamma < 0$, alors $\text{Exp}(\gamma) < 1$. En appliquant le règle de Cauchy on conclut que $\sum_{n=1}^N a(\eta_{t-1}) \dots a(\eta_{t-n})$ converge presque sûrement dans \mathbb{R} .

Donc $A_t = \alpha_0 \left[1 + \sum_{i=1}^{\infty} a(\eta_{t-1}) \dots a(\eta_{t-i}) \right]$ converge presque sûrement dans \mathbb{R} . De plus, le processus

$$\varepsilon_t = \sqrt{A_t} \eta_t = \left\{ \alpha_0 \left[1 + \sum_{i=1}^{\infty} a(\eta_{t-1}) \dots a(\eta_{t-i}) \right] \right\}^{\frac{1}{2}} \eta_t$$

est strictement stationnaire et ergodique puisqu'il est fonction de processus strictement stationnaire et ergodique. Il est aussi non anticipatif comme fonction mesurable des variables η_{t-i} , $i \geq 0$.

¹Soit $(\sum a_n)$ une série à terme positifs et soit $\lambda = \lim \left(a_n^{1/n} \right)$. Alors (i) si $\lambda < 1$ la série $(\sum a_n)$ est convergente, (ii) si $\lambda > 1$ la série $(\sum a_n)$ est divergente.

Montrons maintenant l'unicité de cette solution. Supposons que la solution n'est pas unique et qu'il existe une autre solution strictement stationnaire $\tilde{\varepsilon}_t = \sqrt{h_t}\eta_t$. D'après (4.6.3) on a :

$$h_t - A_t = \{A_t(N) - A_t\} + a(\eta_{t-1}) \dots a(\eta_{t-N-1}) h_{t-N-1}.$$

Il est clair que $A_t(N) - A_t \xrightarrow[N \rightarrow \infty]{p.s.} 0$.

On a vu que la série définissant A_t converge presque sûrement, donc on a :

$$a(\eta_{t-1}) \dots a(\eta_{t-n}) \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

De plus la loi de h_{t-N-1} est indépendante de N par stationnarité. Par suite :

$$a(\eta_{t-1}) \dots a(\eta_{t-N-1}) h_{t-N-1} \xrightarrow[N \rightarrow \infty]{p} 0.$$

Donc on a montré que $h_t - A_t \xrightarrow[N \rightarrow \infty]{p} 0$.

Ce terme étant indépendant de N on a forcément $h_t = A_t, \forall t$. D'où $\varepsilon_t = \sqrt{A_t}\eta_t$ est l'unique solution strictement stationnaire du modèle.

– Supposons que $\gamma > 0$.

D'après la règle de Cauchy, $\sum_{n=1}^N a(\eta_{t-1}) \dots a(\eta_{t-n}) \xrightarrow{p.s.} 0$. Donc si $\alpha_0 > 0$, $A_t = +\infty$ p.s.

D'où $h_t = +\infty$ p.s donc il n'existe pas de solution finie p.s du modèle.

– Supposons que $\gamma = 0$.

Nous procédons par l'absurde. Supposons qu'il existe une solution strictement stationnaire (ε_t, h_t) du modèle ARCH(1). Nous avons pour $n > 0$,

$$h_0 \geq \alpha_0 \left\{ 1 + \sum_{i=1}^n a(\eta_{-1}) \dots a(\eta_{-i}) \right\}$$

d'où on déduit que $a(\eta_{-1}) \dots a(\eta_{-i}) \alpha_0$ converge vers zéro, p.s., quand $n \rightarrow \infty$, ou, de manière équivalente, que

$$\sum_{i=1}^n \log(a(\eta_i)) + \log(\alpha_0) \xrightarrow[n \rightarrow \infty]{p.s.} -\infty \quad (4.6.4)$$

D'après le théorème de Chung-Fuchs² nous avons $\limsup \sum_{i=1}^n \log(a(\eta_i)) = +\infty$ avec probabilité 1, ce qui contredit (4.6.4). ■

²Si X_1, \dots, X_n est une suite iid telle que $\mathbb{E}(X_1) = 0$ et $\mathbb{E}|X_1| > 0$ alors $\lim_{n \rightarrow \infty} \sup \sum_{i=1}^n X_i = +\infty$ et $\lim_{n \rightarrow \infty} \inf \sum_{i=1}^n X_i = -\infty$

4.4.2 Stationnarité au second ordre

Théorème 4.6.1 (*Stationnarité au second ordre du ARCH(1)*)

On suppose $\alpha_0 > 0$, alors

– Si $\alpha_1 \geq 1$, il n'existe pas de solution ARCH(1) non anticipative et stationnaire au second ordre.

– Si $\alpha_1 < 1$, le processus (ε_t) défini par $\varepsilon_t = \sqrt{h_t}\eta_t$ est stationnaire au second ordre. Plus précisément (ε_t) est un bruit blanc faible.

De plus, il n'existe pas d'autre solution stationnaire au second ordre et non anticipative.

4.5 Propriétés asymptotiques des estimateurs des paramètres d'un ARCH(1)

4.5.1 Estimation des paramètres d'un ARCH(1) par la méthode des MCO en deux étapes (2S – LS)

La méthode la plus simple pour estimer les modèles ARCH est la méthode des moindres carrés ordinaires (MCO). Cette procédure d'estimation a pour avantage la simplicité numérique mais présente deux inconvénients :

- Elle n'est pas efficace, par comparaison aux méthodes fondées sur la vraisemblance, ou quasi-vraisemblance.
- Elle nécessite, pour fournir des estimateurs asymptotiquement normaux, des moments d'ordre 8 du processus générateur de la série sous adjacente.

Son extension, la méthode des moindres carrés quasi-généralisée (MCQG) permet de palier le premier inconvénient et de réduire le second : en effet, elle fournit des estimateurs asymptotiquement aussi précis que le quasi-maximum de vraisemblance sous des hypothèses de moments d'ordre 4.

La méthode des MCO a été proposée par Engle (1982) pour les modèles ARCH. Les propriétés théoriques de l'estimateur des MCO ont été établies par Weiss (1984, 1986), dans le cadre ARMA – GARCH, avec des hypothèses de moments d'ordre 8 du processus GARCH.

Pantula (1989) a également étudié les propriétés asymptotiques de la méthode des MCO dans le cas $AR(1) - ARCH(q)$ et donné une forme explicite pour la matrice de variance asymptotique. La méthode des MCQG a été développée, dans le cadre ARCH, par Bose et Mukherjee (2003). Bose et Mukherjee (2001) ont développé une méthode d'estimation des moindres carrés en deux étapes pour estimer les paramètres d'un modèle ARCH (two-stage least squares $2S - LS$). La première étape de cette méthode consiste en l'estimation de ces paramètres en ignorant la variabilité de h_t ainsi que la présence de ces paramètres en h_t . Nous pouvons ainsi obtenir un estimateur de moindres carrés préliminaire que nous notons $\hat{\alpha}_{pr}$, où, comme nous allons voir dans ce qui suit, est le vecteur des paramètres α_0 et α_1 ($\alpha = (\alpha_0, \alpha_1)$). Dans la seconde étape, l'estimateur préliminaire trouvé dans la première étape est utilisé pour construire, tout en prenant en considération $h_t(\alpha)$, l'estimateur amélioré de α noté $\hat{\alpha}_\varphi$.

Cas où $\alpha_1 < 1$

Etape 1 : Estimateur préliminaire du vecteur α et ses propriétés

Rappelons que le modèle ARCH(1) est donnée par :

$$\varepsilon_t = \eta_t \sqrt{h_t} \text{ où } \eta_t \rightsquigarrow N(0, 1), \quad (4.7.1a)$$

$$\text{et } h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2. \text{ où } \alpha_0 > 0 \text{ et } \alpha_1 \geq 0 \quad (4.7.1b)$$

De (4.7.1) on en déduit la représentation AR(1) suivante :

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + u_t, \quad \text{où } u_t = \varepsilon_t^2 - h_t = h_t(\eta_t^2 - 1). \quad (4.7.2)$$

où $u_t = \varepsilon_t^2 - h_t = h_t(\eta_t^2 - 1)$.

En introduisant le vecteur $Z_{t-1} = (1, \varepsilon_{t-1}^2)'$ on déduit de (4.7.2) le système :

$$\varepsilon_t^2 = \alpha Z_{t-1} + u_t, \quad t = 1, \dots, n. \quad (4.7.3)$$

Ainsi, l'estimateur (préliminaire) des moindres carrés de α est donné par

Théorème 4.7.1.

L'estimateur préliminaire du vecteur $\alpha = (\alpha_0, \alpha_1)'$, noté $\hat{\alpha}_{pr}$, de la méthode des moindres carrés est, sous l'hypothèse H1, donné par :

$$\hat{\alpha}_{pr} = \left(\sum_{t=1}^n Z_{t-1} Z_{t-1}' \right)^{-1} \left(\sum_{t=1}^n Z_{t-1} \varepsilon_t^2 \right). \quad (4.7.4)$$

Démonstration. La preuve découle directement en exploitant forme (4.7.3) : ■

En exploitant *Théorème 2.4* (Francq et Zakoïan, (2009)) assurant que la seule solution non anticipative ε_t est ergodique, nous pouvons énoncer la propriété suivante établissant la convergence presque sûre de l'estimateur préliminaire.

Propriété 4.7.1.

L'estimateur préliminaire $\hat{\alpha}_{pr}$ du vecteur α converge presque sûrement, quand $n \rightarrow \infty$, vers la vraie valeur inconnue α , i.e.,

$$\hat{\alpha}_{pr} \xrightarrow{p.s.} \alpha.$$

Démonstration. D'après le *Théorème 2.4* (Francq et Zakoïan, (2009)) l'unique solution non anticipative ε_t est ergodique. Le processus $\{Z_t\}$ est également ergodique puisque Z_t s'écrit comme fonction mesurable des ε_{t-i} , ainsi on peut appliquer le théorème ergodique sur les composantes de la matrice $Z_{t-1}Z'_{t-1}$. D'après le théorème ergodique on a :

$$\frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \xrightarrow{p.s.} \mathbb{E}(Z_{t-1}Z'_{t-1}), \quad (4.7.5a)$$

$$\frac{1}{n} \sum_{t=1}^n Z_{t-1}\varepsilon_t^2 \xrightarrow{p.s.} \mathbb{E}(Z_{t-1}\varepsilon_t^2). \quad (4.7.5b)$$

D'où :

$$\hat{\alpha}_{pr} \xrightarrow{p.s.} (\mathbb{E}(Z_{t-1}Z'_{t-1}))^{-1} \mathbb{E}(Z_{t-1}\varepsilon_t^2),$$

En multipliant (4.7.2) par Z_{t-1} on obtient :

$$Z_{t-1}\varepsilon_t^2 = \alpha Z_{t-1}Z'_{t-1} + Z_{t-1}u_t,$$

On déduit donc l'espérance de $Z_{t-1}\varepsilon_t^2$ qui est donnée par :

$$\mathbb{E}(Z_{t-1}\varepsilon_t^2) = \alpha E(Z_{t-1}Z'_{t-1}) + \mathbb{E}(Z_{t-1}u_t),$$

or, $\mathbb{E}(Z_{t-1}u_t) = 0$, alors $\mathbb{E}(Z_{t-1}\varepsilon_t^2) = \alpha E(Z_{t-1}Z'_{t-1})$.

On conclut donc que $\hat{\alpha}_{pr} \xrightarrow{p.s.} \alpha$. ■

Cet estimateur est aussi doté par la propriété suivante qui établit la convergence, quand $n \rightarrow \infty$, en loi.

Propriété 4.7.2.

L'estimateur préliminaire $\hat{\alpha}_{pr}$ du vecteur α est tel que :

$$\sqrt{n}(\hat{\alpha}_{pr} - \alpha) \xrightarrow{\mathcal{L}} N\left(0, \left(\mathbb{E}(\eta_t^4) - 1\right) A^{-1}BA^{-1}\right),$$

où les matrices carrées A et B sont données comme suit :

$$A = \mathbb{E}\left(Z_{t-1}Z'_{t-1}\right) = \begin{pmatrix} 1 & \mathbb{E}(\varepsilon_{t-1}^2) \\ \mathbb{E}(\varepsilon_{t-1}^2) & \mathbb{E}(\varepsilon_{t-1}^4) \end{pmatrix},$$

$$B = \mathbb{E}\left(h_t^2 Z_{t-1}Z'_{t-1}\right),$$

$$= \begin{pmatrix} \alpha_0 + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^2) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^4) & \alpha_0\mathbb{E}(\varepsilon_{t-1}^2) + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^4) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^6) \\ \alpha_0\mathbb{E}(\varepsilon_{t-1}^2) + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^4) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^6) & \alpha_0\mathbb{E}(\varepsilon_{t-1}^4) + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^6) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^8) \end{pmatrix}.$$

Démonstration. En remplaçant dans (4.7.4) ε_t^2 par sa valeur donnée par (4.7.2) on obtient :

$$\hat{\alpha}_{pr} - \alpha = \left(\sum_{t=1}^n Z_{t-1}Z'_{t-1}\right)^{-1} \left(\sum_{t=1}^n Z_{t-1}u_t\right). \quad (4.7.6)$$

On a vu que :

$$\frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \xrightarrow{p.s} A, \text{ où } A = \mathbb{E}\left(Z_{t-1}Z'_{t-1}\right). \quad (4.7.7)$$

Pour trouver la loi asymptotique de $\sum_{t=1}^n Z_{t-1}u_t$ nous allons utiliser le théorème de Cramér-Wold. Soit $\lambda \in \mathbb{R}^2, \lambda \neq 0$, la suite $\left(\lambda' Z_{t-1}u_t\right)$ est une différence de martingale stationnaire, ergodique.

$$\begin{aligned} \text{Var}(\lambda' Z_{t-1}u_t) &= \lambda' \mathbb{E}(Z_{t-1}Z'_{t-1}u_t^2)\lambda, \\ &= \lambda' \mathbb{E}\left[Z_{t-1}Z'_{t-1}h_t^2(\eta_t^2 - 1)^2\right]\lambda, \\ &= \lambda' \mathbb{E}\left(Z_{t-1}Z'_{t-1}\right) \mathbb{E}(h_t^2) \mathbb{E}\left[(\eta_t^2 - 1)^2\right]\lambda, \\ &= \underbrace{\left(\mathbb{E}(\eta_t^4) - 1\right)}_B \lambda' \mathbb{E}\left(h_t^2 Z_{t-1}Z'_{t-1}\right)\lambda. \end{aligned}$$

Donc

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1}u_t \xrightarrow{\mathcal{L}} N\left(0, \left(\mathbb{E}(\eta_t^4) - 1\right) \lambda' B \lambda\right)$$

Et en appliquant le théorème de Cramér-Wold on obtient :

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} u_t \xrightarrow{\mathcal{L}} N(0, (\mathbb{E}(\eta_t^4) - 1) B), \quad (4.7.8)$$

de (4.7.7) et (4.7.8) on déduit que :

$$\sqrt{n}(\hat{\alpha}_{pr} - \alpha) \xrightarrow{\mathcal{L}} N(0, A^{-1} (\mathbb{E}(\eta_t^4) - 1) B), \quad (4.7.9)$$

avec

$$A = \mathbb{E} (Z_{t-1} Z'_{t-1}) = \begin{pmatrix} 1 & \mathbb{E}(\varepsilon_{t-1}^2) \\ \mathbb{E}(\varepsilon_{t-1}^2) & \mathbb{E}(\varepsilon_{t-1}^4) \end{pmatrix},$$

$$B = \mathbb{E} (h_t^2 Z_{t-1} Z'_{t-1}),$$

$$= \begin{pmatrix} \alpha_0 + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^2) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^4) & \alpha_0\mathbb{E}(\varepsilon_{t-1}^2) + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^4) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^6) \\ \alpha_0\mathbb{E}(\varepsilon_{t-1}^2) + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^4) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^6) & \alpha_0\mathbb{E}(\varepsilon_{t-1}^4) + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^6) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^8) \end{pmatrix},$$

où

$$\mathbb{E}(\varepsilon_{t-1}^2) = \frac{\alpha_0}{1 - \alpha_1},$$

$$\mathbb{E}(\varepsilon_{t-1}^4) = \frac{\alpha_0^2\mathbb{E}(\eta_t^4)(1 + \alpha_1)}{(1 - \alpha_1)(1 - \alpha_1^2\mathbb{E}(\eta_t^4))} = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}, \text{ car } \mathbb{E}(\eta_t^4) = 3,$$

$$\begin{aligned} \mathbb{E}(\varepsilon_{t-1}^6) &= \frac{\alpha_0^3(1 + 2\alpha_1)\mathbb{E}(\eta_t^4) + \alpha_0^3\alpha_1^2(2 + \alpha_1)\mathbb{E}(\eta_t^4)\mathbb{E}(\eta_t^6)}{(1 - \alpha_1)(1 - \alpha_1^2\mathbb{E}(\eta_t^4))(1 - \alpha_1^3\mathbb{E}(\eta_t^6))}, \\ &= \frac{3\alpha_0^3(1 + 2\alpha_1) + 45\alpha_0^3\alpha_1^2(2 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)(1 - 15\alpha_1^3)}, \text{ car } \mathbb{E}(\eta_t^4) = 3 \text{ et } \mathbb{E}(\eta_t^6) = 15, \end{aligned}$$

$$\mathbb{E}(\varepsilon_{t-1}^8) =$$

$$\frac{4\alpha_0^4\alpha_1 + 2\alpha_0^4\alpha_1^2(3 - 2\alpha_1)\mathbb{E}(\eta_t^4) + 4\alpha_0^4\alpha_1^3(1 + \alpha_1)\mathbb{E}(\eta_t^6) + 4\alpha_0^4\alpha_1^5(2 + \alpha_1)\mathbb{E}(\eta_t^4)\mathbb{E}(\eta_t^6)}{(1 - \alpha_1)(1 - \alpha_1^2\mathbb{E}(\eta_t^4))(1 - \alpha_1^3\mathbb{E}(\eta_t^6))(1 - \alpha_1^4\mathbb{E}(\eta_t^8))} \mathbb{E}(\eta_t^8).$$

$$\mathbb{E}(\varepsilon_{t-1}^8) = \frac{4\alpha_0^4\alpha_1 + 6\alpha_0^4\alpha_1^2(3 - 2\alpha_1) + 60\alpha_0^4\alpha_1^3(1 + \alpha_1) + 180\alpha_0^4\alpha_1^5(2 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)(1 - 15\alpha_1^3)(1 - 105\alpha_1^4)} 105,$$

$$\text{car } \mathbb{E}(\eta_t^8) = 105. \quad \blacksquare$$

Étape 2 : Estimation fonctionnelle du vecteur α et ses propriétés

Bien que l'estimateur préliminaire $\hat{\alpha}_{pr}$, obtenu dans la première étape converge presque sûrement vers le vrai vecteur inconnu α et converge en loi vers une loi gaussienne, il souffre de la nécessité de l'existence des moments d'ordre 8 qui est une condition forte. Pour face à cet inconvénient Schick (1996) a proposé une classe d'estimateurs de α paramétrée par

une fonctionnelle, dont l'estimateur $\hat{\alpha}_{pr}$ se déduit pour le cas particulier d'une fonction identité. Soit φ_p l'ensemble des fonctions bornées et mesurables de \mathbb{R}^p dans \mathbb{R}^p . Pour $\varphi \in \varphi_1$, l'estimateur fonctionnel $\hat{\alpha}_\varphi$ est établi par le théorème suivant.

Théorème 4.7.2.

L'estimateur fonctionnel $\hat{\alpha}_\varphi$ du vecteur $\alpha = (\alpha_0, \alpha_1)'$ est donné par :

$$\tilde{\alpha}_\varphi = \left(\sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \right)^{-1} \left(\sum_{t=1}^n \varphi(Z_{t-1}) \varepsilon_t^2 \right)$$

La convergence presque sûre de $\tilde{\alpha}_\varphi$, quand $n \rightarrow \infty$, vers α est assurée par la propriété suivante.

Propriété 4.7.3.

L'estimateur fonctionnel $\hat{\alpha}_\varphi$ du vecteur α est tel que :

$$\tilde{\alpha}_\varphi \xrightarrow{p.s.} \alpha.$$

Démonstration. Cet estimateur est fortement convergent sous l'hypothèse que $\mathbb{E} \left(\varphi(Z_{t-1}) Z_{t-1} \right)$ soit finie. En effet,

$$\begin{aligned} \tilde{\alpha}_\varphi &= \left(\sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \right)^{-1} \sum_{t=1}^n \varphi(Z_{t-1}) (\alpha Z_{t-1} + u_t), \\ &= \left(\sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \right)^{-1} \left(\sum_{t=1}^n \varphi(Z_{t-1}) Z_{t-1} \alpha + \sum_{t=1}^n \varphi(Z_{t-1}) u_t \right), \\ &= \left(\sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \right)^{-1} \left(\sum_{t=1}^n \varphi(Z_{t-1}) Z_{t-1} \alpha + \sum_{t=1}^n \varphi(Z_{t-1}) u_t \right), \\ &= \alpha + \left(\frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \right)^{-1} \frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) u_t. \end{aligned} \tag{4.7.10}$$

Le théorème ergodique stipule d'une part que

$$\frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E} \left(\varphi(Z_{t-1}) Z'_{t-1} \right), \tag{4.7.11}$$

sous peine que $\mathbb{E} \left(\varphi(Z_{t-1}) Z'_{t-1} \right) < \infty$. Or une condition nécessaire et suffisante pour que $\mathbb{E} \left(\varphi(Z_{t-1}) Z'_{t-1} \right) < \infty$ sera donnée plus tard en fonction d'un choix particulier $\varphi(\cdot)$. D'autre part comme le processus (W_t) défini par

$$\begin{aligned} W_t &= \varphi(Z_{t-1}) u_t, \\ &= \varphi(Z_{t-1}) (\varepsilon_t^2 - \alpha Z_{t-1}), \quad t \in \mathbb{Z}, \end{aligned}$$

est strictement stationnaire et ergodique, donc toujours par le théorème ergodique

$$\begin{aligned}
 \frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) u_t &= \frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) (\varepsilon_t^2 - \alpha Z_{t-1}), \\
 &\xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E}(\varphi(Z_{t-1}) (\varepsilon_t^2 - \alpha Z_{t-1})), \\
 &= \mathbb{E}(\mathbb{E}(\varphi(Z_{t-1}) (\varepsilon_t^2 - \alpha Z_{t-1}) / \mathcal{F}_{t-1})), \\
 &= \mathbb{E}(\varphi(Z_{t-1}) \mathbb{E}((\varepsilon_t^2 - \alpha Z_{t-1}) / \mathcal{F}_{t-1})), \\
 &= 0.
 \end{aligned} \tag{4.7.12}$$

Finalemt par (4.7.9), (4.7.11) et le lemme de Slutsky (si $(X_n) \xrightarrow[p.s.]{} X$ et $(Y_n) \xrightarrow[p.s.]{} Y$ alors $(X_n Y_n) \xrightarrow[p.s.]{} XY$), il s'ensuit que

$$\tilde{\alpha}_\varphi \xrightarrow[n \rightarrow \infty]{p.s.} \alpha. \quad \blacksquare$$

Cet estimateur est aussi doté par la propriété suivante qui établit la convergence, quand $n \rightarrow \infty$, en loi. En effet, nous la propriété suivante.

Propriété 4.7.4.

L'estimateur fonctionnel $\hat{\alpha}_\varphi$ du vecteur α est tel que :

$$\sqrt{n} (\hat{\alpha}_\varphi - \alpha) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, \left(\mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right) \right)^{-1} \right).$$

Démonstration. L'estimateur préliminaire $\hat{\alpha}_\varphi$ est asymptotiquement gaussien à condition que

$$\mathbb{E} \left(\varphi(Z_{t-1}) \varphi'(Z_{t-1}) (\varepsilon_t^2 - \alpha Z_{t-1})^2 \right),$$

soit finie. En effet, en réécrivant (4.7.10) comme suit

$$\sqrt{n} (\hat{\alpha}_\varphi - \alpha) = \left(\frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \right)^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n \varphi(Z_{t-1}) u_t, \tag{4.7.13}$$

il est facile de voir que

$$\sum_{t=1}^n \frac{1}{\sqrt{n}} \varphi(Z_{t-1}) u_t := \sum_{t=1}^n \kappa_{n,t},$$

vérifie les conditions du théorème central limite des martingales (cf, Corollary 3.1 de Hall et Heyde, 1980) ($S_n = \sum_{t=1}^n \kappa_{n,t}$ étant un tableau de martingales, Martingale-array). En particulier,

$$\begin{aligned}
 \sum_{t=1}^n \kappa_{n,t}^2 &= \frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) \varphi'(Z_{t-1}) u_t^2, \\
 &\xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E} \left(\varphi(Z_{t-1}) \varphi'(Z_{t-1}) (\varepsilon_t^2 - h_t(\hat{\alpha}_{pr})) \right), \\
 &= \mathbb{E} \left(\varphi(Z_{t-1}) \varphi'(Z_{t-1}) \right) \mathbb{E} \left(h_t^2(\hat{\alpha}) (\eta_t^2 - 1)^2 \right).
 \end{aligned}$$

D'où,

$$\frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) \varphi'(Z_{t-1}) u_t^2 \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, \mathbb{E} \left(\varphi(Z_{t-1}) \varphi'(Z_{t-1}) \right) \mathbb{E} \left(h_t^2(\hat{\alpha}) \right) \right). \quad (4.7.14)$$

D'autre part, (4.7.11) implique que

$$\frac{1}{n} \sum_{t=1}^n \varphi(Z_{t-1}) Z'_{t-1} \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E} \left(\varphi(Z_{t-1}) Z'_{t-1} \right), \quad (4.7.15)$$

et par le lemme de Slutsky et en utilisant (4.7.13)-(4.7.11), on trouve

$$\sqrt{n} (\tilde{\alpha}_\varphi - \alpha) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, \Omega_\varphi),$$

où

$$\Omega_\varphi = \left(\mathbb{E} \left(\varphi(Z_{t-1}) Z'_{t-1} \right) \right)^{-2} \mathbb{E} \left(\varphi(Z_{t-1}) \varphi'(Z_{t-1}) \right) \mathbb{E} \left(h_t^2(\hat{\alpha}) \right). \quad (4.7.16)$$

Ainsi, il revient de trouver la fonction $\varphi^*(\cdot) \in \varphi_1$ pour laquelle Ω_{φ^*} est la plus petite possible dans le sens où

$$\Omega_{\varphi^*} \leq \Omega_\varphi \quad \text{pour tout } \varphi \in \varphi_1,$$

et qui ne nécessite pas l'existence des moments pour le processus. Or, voyant la structure de Ω_φ on peut penser que la fonction optimale φ^* est telle que

$$\varphi^*(Z_{t-1}) Z'_{t-1} = \varphi(Z_{t-1}) \varphi'(Z_{t-1}) h_t^2(\alpha),$$

ou encore

$$\varphi^*(Z_{t-1}) = \frac{Z_{t-1}}{h_t^2(\alpha)}, \quad (4.7.17)$$

$$\begin{aligned}
 \Omega_{\varphi^*} &= \left(\mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right) \right)^{-2} \mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^4(\alpha)} \right) \mathbb{E} \left(h_t^2(\alpha) \right), \\
 &= \left(\mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right) \right)^{-1} \left(\mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right) \right)^{-1} \mathbb{E} \left(\frac{h_t^2 Z_{t-1} Z'_{t-1}}{h_t^4(\alpha)} \right).
 \end{aligned}$$

ce qui donne

$$\Omega_{\varphi^*} = \left(\mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right) \right)^{-1}. \quad (4.7.18)$$

Effectivement, (4.7.18) est vraie puisque par l'inégalité de Cauchy-Schwarz

$$\begin{aligned} & \left(\mathbb{E} (\varphi(Z_{t-1}) Z'_{t-1}) \right)^{-1} \mathbb{E} (\varphi(Z_{t-1}) \varphi'(Z_{t-1})) \mathbb{E} (h_t^2(\alpha)) \\ &= \left(\mathbb{E} (\varphi^*(Z_{t-1}) Z'_{t-1}) \right)^{-1} \mathbb{E} (\varphi^*(Z_{t-1}) \varphi^{*'}(Z_{t-1})) \mathbb{E} (h_t^2(\alpha)), \\ &\geq \frac{1}{\mathbb{E} \left(\frac{Z_{t-1}}{h_t^2(\alpha)} \right)}. \end{aligned}$$

On note que $\mathbb{E} \left(\frac{Z_{t-1}}{h_t^2(\alpha)} \right)$ (et donc Ω_{φ^*}) existe pourvu que $\alpha_1 < \exp(-\mathbb{E}(\log \eta_t^2)) \simeq 3.5$ sans aucune condition supplémentaire sur les moments. Donc tenant compte de (4.7.16)-(4.7.17), on peut proposer l'estimateur fonctionnel $\tilde{\alpha}_{\tilde{\varphi}}$ pour le cas particulier

$$\tilde{\varphi}(Z_{t-1}) = \frac{Z_{t-1}}{h_t^2(\alpha)},$$

Avec une telle fonction $\tilde{\varphi}(Z_{t-1})$, l'estimateur fonctionnel $\tilde{\phi}_{\tilde{h}}$ et sa distribution limite sont donnés par

$$\tilde{\alpha}_{\tilde{\varphi}} = \left(\sum_{t=1}^n \frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right)^{-1} \left(\sum_{t=1}^n \frac{Z_{t-1} \varepsilon_t^2}{h_t^2(\alpha)} \right), \quad (4.7.19)$$

$$\sqrt{n}(\tilde{\alpha}_{\tilde{\varphi}} - \alpha) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, \left(\mathbb{E} \left(\frac{Z_{t-1} Z'_{t-1}}{h_t^2(\alpha)} \right) \right)^{-1} \right). \blacksquare$$

4.5.2 Estimation des paramètres d'un ARCH(1) par la méthode des MCO en deux étapes (2S – WLS)

2S – WLS dans le cas stable

La méthode d'estimation 2S – LS ne converge pas dans le cas d'instabilité, pour remédier à ce problème Aknouche (2010) a proposé une méthode d'estimation en deux étapes améliorée

"two-stage weighted least squares" ($2S - WLS$) qui permet d'estimer les paramètres d'un modèle ARCH(1) et dans le cas d'instabilité et dans le cas d'instabilité. Cette méthode nécessite aucune condition sur les moments du processus.

Soit le modèle ARCH(1) donné par (4.7.1). Soient $\{u_t, t \in \mathbb{N}\}$ une suite de variables aléatoires iid de moyenne 0 et de variance 1, et le paramètre $\alpha = (\alpha_0, \alpha_1)'$ tel que $\alpha_0 > 0$ et $\alpha_1 > 0$. Posons $u_t = \eta_t^2 - 1$ et $v_t = \varepsilon_t^2 - h_t(\alpha) = h_t(\alpha)u_t$, l'équation (4.7.1) peut être réécrite comme suit :

$$\varepsilon_t^2 = \mathcal{Z}'_{t-1} \alpha + h_t(\alpha)u_t, \quad t \in \mathbb{Z} \quad (4.7.20)$$

où $\mathcal{Z}_{t-1} = (1, \varepsilon_{t-1}^2)'$ et $h_t(\alpha) = \alpha' \mathcal{Z}_{t-1}$.

La condition nécessaire et suffisante sur α_1 pour que (4.7.1) ait une solution strictement stationnaire et ergodique non anticipative est donnée par :

$$\mathbf{A}_0: 0 \leq \alpha_1 < \exp(-E \log(\eta_t^2)).$$

de plus on suppose que :

$$\mathbf{A}_1: E(\eta_t^4) < \infty.$$

Soit \mathcal{G} l'ensemble des fonctions bornées et continues de \mathbb{R}^2 dans \mathbb{R}^2 . Considérons l'estimateur fonctionnel $\hat{\alpha}(g)$ indexé par $g \in \mathcal{G}$ et donnée par

$$\hat{\alpha}(g) = \left(\sum_{t=1}^n g(\mathcal{Z}_{t-1}) \mathcal{Y}'_{t-1} \right)^{-1} \sum_{t=1}^n g(\mathcal{Z}_{t-1}) \varepsilon_t. \quad (4.7.21)$$

Posons comme hypothèse :

$$\mathbf{A}_2: E(g(\mathcal{Z}_{t-1}) \mathcal{Z}'_{t-1}) < \infty.$$

$$\mathbf{A}_3: E(g(\mathcal{Y}_{t-1}) g(\mathcal{Y}_{t-1})' h_t^2(\theta)) < \infty.$$

Théorème 4.7.3. (Aknouche, 2010)

i) Sous \mathbf{A}_0 et \mathbf{A}_2 $\hat{\alpha}(g)$ converge presque sûrement vers α quand $n \rightarrow \infty$.

ii) Si de plus \mathbf{A}_1 et \mathbf{A}_3 sont vérifiées alors

$$\sqrt{n}(\hat{\alpha}(g) - \alpha) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, \Delta(g)),$$

où

$$\Delta(g; \alpha) = \kappa E(g(\mathcal{Z}_0) \mathcal{Z}'_0)^{-1} E\left(g(\mathcal{Z}_0) g(\mathcal{Z}_0)' (\alpha' \mathcal{Z}_0)^2\right) E(g(\mathcal{Z}_0) \mathcal{Z}'_0)^{-1}, \quad (4.7.22)$$

$$\kappa = E(\eta_t^4) - 1.$$

L'objectif est de trouver la fonction g^* pour laquelle $\Delta(g^*; \alpha)$ est minimale dans le sens où

$$\Delta(g^*; \alpha) \leq \Delta(g; \alpha), \quad \text{pour tout } g \in \mathcal{G}.$$

Puisque la fonction optimale g^* est telle que

$$g^*(\mathcal{Z}_0) \mathcal{Z}'_0 = g^*(\mathcal{Z}_0) g^*(\mathcal{Z}_0)' h_1^2(\alpha),$$

donnant

$$\begin{aligned} g^*(\mathcal{Z}_0) &= g^*(\mathcal{Z}_0; \alpha) \\ &= \frac{1}{(\alpha' \mathcal{Z}_0)^2} \mathcal{Z}_0. \end{aligned}$$

Ainsi l'estimateur optimal $\hat{\alpha}(g^*) = \hat{\alpha}(g^*; \alpha)$ dans la classe des estimateurs fonctionnels de la forme (2.4.21) est donné par :

$$\hat{\alpha}(g^*; \alpha) = \left(\sum_{t=1}^n \frac{1}{(\alpha' \mathcal{Z}_{t-1})^2} \mathcal{Z}_{t-1} \mathcal{Z}'_{t-1} \right)^{-1} \sum_{t=1}^n \mathcal{Z}_{t-1} \frac{\varepsilon_t^2}{(\alpha' \mathcal{Z}_{t-1})^2}, \quad (4.7.23)$$

et sa variance asymptotique est donnée par :

$$\begin{aligned} \Delta(g^*) &= \Delta(g^*, \alpha), \\ &= \kappa \left(E \left(\frac{1}{(\alpha' \mathcal{Z}_0)^2} \mathcal{Z}_0 \mathcal{Z}'_0 \right) \right)^{-1}. \end{aligned} \quad (4.7.24)$$

Au lieu d'estimer le α préliminaire par la méthode des *MCO* comme l'ont fait Bose et Mukherjee (2003), on tire profit de la forme *WLS* donnée par (4.7.19). Ainsi, pour tout $\alpha^* = (\alpha_0^*, \alpha_1^*)'$ tels que $\alpha_0^* > 0$ et $\alpha_1^* > 0$, on considère l'estimateur donné par (4.7.19), mais amélioré par le carré de la variance conditionnelle $h_t^2(\alpha^*)$ évalué au point α^*

$$\begin{aligned} \hat{\alpha}_{WLS} &= \hat{\alpha}(g^*; \alpha_1) \\ &= \left(\sum_{t=1}^n \frac{1}{(\alpha^{*'} \mathcal{Z}_{t-1})^2} \mathcal{Z}_{t-1} \mathcal{Z}'_{t-1} \right)^{-1} \sum_{t=1}^n \mathcal{Z}_{t-1} \frac{\varepsilon_t^2}{(\alpha^{*'} \mathcal{Z}_{t-1})^2}. \end{aligned} \quad (4.7.25)$$

Sous les conditions des moments suivantes :

$$E \left(\frac{1}{(\alpha^* \mathcal{Z}_0)^4} \mathcal{Z}_0 \mathcal{Z}'_0 (\alpha' \mathcal{Z}_0)^2 \right) < \infty, \quad (2.4.26a)$$

$$E \left(\frac{1}{(\alpha^* \mathcal{Z}_0)^2} \mathcal{Z}_0 \mathcal{Z}'_0 \right) < \infty, \quad (2.4.26b)$$

$\hat{\alpha}_{WLS}$ est consistant et asymptotiquement Gaussien avec variance asymptotique

$$\Delta(g^*; \alpha, \alpha^*) = \kappa E \left(\frac{1}{(\alpha^* \mathcal{Z}_0)^2} \mathcal{Z}_0 \mathcal{Z}'_0 \right)^{-1} E \left(\frac{1}{(\alpha^* \mathcal{Z}_0)^4} \mathcal{Z}_0 \mathcal{Z}'_0 (\alpha' \mathcal{Z}_0)^2 \right) E \left(\frac{1}{(\alpha^* \mathcal{Z}_0)^2} \mathcal{Z}_0 \mathcal{Z}'_0 \right)^{-1}. \quad (4.7.27)$$

Cependant, les conditions des moments (4.7.26) sont toujours satisfaites pour tout $\alpha^* = (\alpha_0^*, \alpha_1^*)'$ tel que $\alpha_0^* > 0$ et $\alpha_1^* > 0$; ainsi $\hat{\alpha}_{WLS}$ est consistant et asymptotiquement Normal seulement sous l'hypothèse de stationnarité stricte **A0** et l'hypothèse **A1**.

Corollaire 4.7.1. (Aknouche, 2010)

Sous **A0** et **A1** et pour tout $\alpha^* = (\alpha_0^*, \alpha_1^*)'$ tel que $\alpha_0^* > 0$ et $\alpha_1^* > 0$, on a :

i) $\hat{\alpha}_{WLS}$ converge p.s vers α quand $n \rightarrow \infty$;

ii)

$$\sqrt{n}(\hat{\alpha}_{WLS} - \alpha) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, \Delta(g^*; \alpha, \alpha^*)).$$

où $\Delta(g^*; \alpha, \alpha^*)$ est donnée par (4.7.27).

Dans la seconde étape on propose le second estimateur $\hat{\alpha}_{2S-WLS}$ qui a encore la forme *WLS* (4.7.23), mais amélioré par le carré de la variance conditionnelle estimée

$$\hat{\alpha}_{2S-WLS} = \left(\sum_{t=1}^n \frac{1}{(\hat{\alpha}'_{WLS} \mathcal{Z}_{t-1})^2} \mathcal{Z}_{t-1} \mathcal{Z}'_{t-1} \right)^{-1} \sum_{t=1}^n \mathcal{Z}_{t-1} \frac{\varepsilon_t^2}{(\hat{\alpha}'_{WLS} \mathcal{Z}_{t-1})^2}. \quad (4.7.28)$$

Théorème 4.7.4. (Aknouche, 2010)

Sous **A0** et **A1**

$$\sqrt{n}(\hat{\alpha}_{2S-WLS} - \alpha) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, \Delta(g^*; \alpha)),$$

où $\Delta(g^*; \alpha)$ est donnée par (4.7.24).

2S – WLS dans le cas instable

Nous avons vu que les propriétés asymptotiques de *WLS* et *2S-WLS* ont été obtenues en se basant sur l'hypothèse de stationnarité **A1**. Lorsque cette hypothèse est non vérifiée, les résultats limites usuels, tel que le théorème ergodique, ne sont pas applicables. Dans ce cas instable, Jensen et Rahbek (2004) ont montré que le, α_0 étant supposé connu, le *QMLE* du paramètre α_1 est asymptotiquement normal. leur résultat est basé sur le Théorème 2 de Nelson (1990) qui montre que sous l'hypothèse $\alpha_1 \geq \exp(-E \log(\eta_t^2))$

$$\varepsilon_t \xrightarrow[t \rightarrow \infty]{a.s.} \infty. \quad (4.7.29)$$

Notons que pour (4.7.29) les termes de la matrice $\frac{1}{n} \sum_{t=1}^n \frac{1}{(\theta^* \mathcal{Z}_{t-1})^2} \mathcal{Z}_{t-1} \mathcal{Z}'_{t-1}$ sont tels que

$$\frac{1}{(\theta^* \mathcal{Z}_{t-1})^2} \mathcal{Z}_{t-1} \mathcal{Z}'_{t-1} = \begin{pmatrix} \frac{1}{(\alpha_0^* + \alpha_1^* \varepsilon_{t-1}^2)^2} & \frac{\varepsilon_{t-1}^2}{(\alpha_0^* + \alpha_1^* \varepsilon_{t-1}^2)^2} \\ \frac{\varepsilon_{t-1}^2}{(\alpha_0^* + \alpha_1^* \varepsilon_{t-1}^2)^2} & \frac{\varepsilon_{t-1}^4}{(\alpha_0^* + \alpha_1^* \varepsilon_{t-1}^2)^2} \end{pmatrix}$$

$$\xrightarrow[t \rightarrow \infty]{a.s.} \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{(\alpha_1^*)^2} \end{pmatrix},$$

alors, par le lemme de Cesaro on a :

$$\frac{1}{n} \sum_{t=1}^n \frac{1}{(\theta_1^* \mathcal{Y}_{t-1})^2} \mathcal{Y}_{t-1} \mathcal{Y}'_{t-1} \xrightarrow[n \rightarrow \infty]{a.s.} \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\alpha_1^2} \end{pmatrix},$$

où la matrice limite est singulière. Puisque l'estimateur *WLS* ne converge pas donc on considère, pour α_0 fixé et connu, l'estimation de α_1 seulement parce que la matrice correspondante est non singulière.

Considérons le modèle *ARCH* donné par (4.7.1) avec α_0 connu, e.g. $\alpha_0 = 1$. pour ce modèle l'estimation *WLS* (4.7.25) pour α_1 se réduit à

$$\hat{\alpha}_{WLS}(\alpha_1) = \left(\sum_{t=1}^n \frac{\varepsilon_{t-1}^4}{(1 + \alpha_1^* \varepsilon_{t-1}^2)^2} \right)^{-1} \sum_{t=1}^n \frac{\varepsilon_{t-1}^2 (\varepsilon_t^2 - 1)}{(1 + \alpha_1^* \varepsilon_{t-1}^2)^2}, \quad (4.7.30)$$

donc en utilisant (4.7.20) avec $\alpha_0 = 1$ et (4.7.30) on obtient

$$\sqrt{n} (\hat{\alpha}_{WLS}(\alpha_1^*) - \alpha_1) = \left(\frac{1}{n} \sum_{t=1}^n \frac{\varepsilon_{t-1}^4}{(1 + \alpha_1^* \varepsilon_{t-1}^2)^2} \right)^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\varepsilon_{t-1}^2 (1 + \alpha_1^* \varepsilon_{t-1}^2) u_t}{(1 + \alpha_1^* \varepsilon_{t-1}^2)^2}.$$

Le résultat suivant montre que pour tout α_1^* positif fixé, le $\hat{\alpha}_{WLS}(\alpha_1^*)$ est asymptotiquement Gaussien avec variance asymptotique indépendante de α_1^* et est la même que celle du *QMLE* (Jensen et Rahbek, 2004).

Théorème 4.7.5. (Aknouche, 2010)

Sous **A1** et $\alpha_1 \geq \exp(-E \log(\eta_t^2))$ on a pour tout α_1^*

$$\sqrt{n}(\widehat{\alpha}_{WLS}(\alpha_1^*) - \alpha_1) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, \kappa \alpha_1^2).$$

4.6 Simulation

Dans cette section nous présentons, tout en évoquant des commentaires, les résultats d'une étude de simulation intensive portant, d'une part, sur les comportements de l'estimateur *MCO*, $\widehat{\alpha}_n$, du paramètre inconnu α , d'un modèle autorégressif hétéroscédastique d'ordre un, et l'estimateur de son écart-type $\widehat{\sigma}_{\widehat{\alpha}_n}$, et d'autre part, sur le comportement de ses lois asymptotiques et ce selon les régions de l'appartenance du paramètre inconnu α et selon la taille de la série simulée, tout en commentant sur la vitesse de convergence. Pour étudier empiriquement la performance de cet estimateur et le comportement de sa distribution asymptotique, dans les trois régions à savoir de stabilité, de racine unitaire et d'explosibilité (instabilité explosive), nous avons considéré les deux vecteurs $(1, 0.2)$ et $(0.3, 0.6)$ représentant le vecteur de paramètres (α_0, α_1) pour le cas où le paramètre α_1 appartient à la région de stabilité $|\alpha_1| < 1$ et les vecteurs $(0.5, 1)$ et $(0.7, 1.3)$ respectivement pour le cas de racine unitaire et le cas explosif. Nous avons pris pour ces deux derniers cas le vecteur des valeurs initiales des paramètres $\alpha^* = (1, 0.5)'$.

Commençons d'abord, par les résultats de simulations concernant la consistance, au sens de la convergence en probabilité et presque sûrement, de l'estimateur $\widehat{\alpha}_n$ et de l'estimateur de son écart-type $\widehat{\sigma}_{\widehat{\alpha}_n}$ dans le cas stationnaire. Les valeurs moyennes, basées sur 1000 répliques, de l'estimateur du *MCO*, en variant la taille de la série de $n = 50$ à $n = 500$, avec un pas de 50, ainsi que les estimations de son écart-type $\widehat{\sigma}_{\widehat{\alpha}_n}$ sont portées sur la *Table 4.8.1* ci-dessous. Nous constatons que les résultats des simulations consolident les résultats théoriques concernant la consistance (au sens des modes de convergence) de l'estimateur $\widehat{\alpha}_n$, en particulier le résultat assurant la convergence presque sûre. En effet, le vecteur estimé de α approche vers les vrais vecteurs $(1, 0.2)$ et $(0.3, 0.6)$ au fur et à mesure que la taille de la série augmente, plus précisément, il passe du vecteur estimé $(0.9980, 0.2093)$ avec un vecteur d'écart-type estimé $(0.2595, 0.1514)$, (respectivement $(0.3436, 0.4845)$), avec

un écart-type estimé $(0.1188, 0.2665)$ pour $n = 50$ au vecteur estimé $(1.006, 0.1891)$ avec un écart-type estimé $(0.0878, 0.0672)$, (respectivement $(0.3043, 0.5842)$), avec un écart-type estimé $(0.0342, 0.1061)$ pour $n = 500$.

En conclusion, les simulations montrent empiriquement que l'estimateur du MCO, pour les deux modèles correspondant à $\alpha = (1, 0.2)$ et $\alpha = (0.3, 0.6)$ appartenant à la région de stabilité, est doté de la propriété de consistance.

La *Table 4.8.2* montre clairement que cette propriété de consistance est bien rencontrée aussi dans les cas d'instabilité à savoir l'instabilité à racine unitaire et l'instabilité explosif correspondant, respectivement, à $\alpha = (0.5, 1)$ et $\alpha = (0.7, 1.3)$.

Itération = 1000

Taille de l'échantillon	(α_0, α_1)	$(\hat{\alpha}_0, \hat{\alpha}_1)$	$Std(\hat{\alpha}_0, \hat{\alpha}_1)$
$N = 50$	(1, 0.2)	(0.9980, 0.2093)	(0.2595, 0.1514)
	(0.3, 0.6)	(0.3436, 0.4845)	(0.1188, 0.2665)
$N = 100$	(1, 0.2)	(1.0108, 1.1920)	(0.1860, 0.1195)
	(0.3, 0.6)	(0.3214, 0.5263)	(0.0840, 0.2276)
$N = 150$	(1, 0.2)	(1.0181, 1.1826)	(0.1603, 0.1076)
	(0.3, 0.6)	(0.3129, 0.5571)	(0.0683, 0.1957)
$N = 200$	(1, 0.2)	(1.0118, 1.1856)	(0.1409, 0.0978)
	(0.3, 0.6)	(0.3124, 0.5600)	(0.0574, 0.1695)
$N = 250$	(1, 0.2)	(1.0142, 1.1828)	(0.1308, 0.0890)
	(0.3, 0.6)	(0.3091, 0.5713)	(0.0520, 0.1493)
$N = 300$	(1, 0.2)	(1.0143, 1.1836)	(0.1177, 0.0843)
	(0.3, 0.6)	(0.3089, 0.5743)	(0.0436, 0.1328)
$N = 350$	(1, 0.2)	(1.0127, 0.1873)	(0.1005, 0.0791)
	(0.3, 0.6)	(0.3038, 0.5796)	(0.0407, 0.1243)
$N = 400$	(1, 0.2)	(1.0129, 0.1870)	(0.0980, 0.0770)
	(0.3, 0.6)	(0.3038, 0.5796)	(0.0375, 0.1156)
$N = 450$	(1, 0.2)	(1.0105, 0.1905)	(0.0964, 0.0682)
	(0.3, 0.6)	(0.3039, 0.5871)	(0.0377, 0.1084)
$N = 500$	(1, 0.2)	(1.006, 0.1891)	(0.0878, 0.0672)
	(0.3, 0.6)	(0.3043, 0.5842)	(0.0342, 0.1061)

Table 4.8.1

Itération = 1000

 $\alpha^* = (1, 0.5)'$

Taille de l'échantillon	(α_0, α_1)	$(\hat{\alpha}_0, \hat{\alpha}_1)$	$Std(\hat{\alpha}_0, \hat{\alpha}_1)$
$N = 50$	(0.5, 1)	(0.5629, 0.8792)	(0.2023, 0.4601)
	(0.7, 1.3)	(0.8298, 1.1549)	(0.3575, 0.4960)
$N = 100$	(0.5, 1)	(0.5328, 0.9276)	(0.1402, 0.2615)
	(0.7, 1.3)	(0.7523, 1.2249)	(0.2160, 0.3079)
$N = 150$	(0.5, 1)	(0.5197, 0.9532)	(0.1127, 0.2099)
	(0.7, 1.3)	(0.7355, 1.2366)	(0.1748, 0.2425)
$N = 200$	(0.5, 1)	(0.5144, 0.9630)	(0.0952, 0.1886)
	(0.7, 1.3)	(0.7274, 1.2643)	(0.1542, 0.2103)
$N = 250$	(0.5, 1)	(0.5116, 0.9708)	(0.0857, 0.1633)
	(0.7, 1.3)	(0.7197, 1.2770)	(0.1401, 0.1821)
$N = 300$	(0.5, 1)	(0.5103, 0.9757)	(0.0767, 0.1536)
	(0.7, 1.3)	(0.7143, 1.2770)	(0.1204, 0.1664)
$N = 350$	(0.5, 1)	(0.5093, 0.9880)	(0.0737, 0.1367)
	(0.7, 1.3)	(0.7169, 1.2820)	(0.1093, 0.1594)
$N = 400$	(1, 0.2)	(0.5083, 0.9784)	(0.0688, 0.1301)
	(0.3, 0.6)	(0.7119, 1.2812)	(0.1050, 0.1489)
$N = 450$	(0.5, 1)	(0.5075, 0.9888)	(0.0624, 0.1208)
	(0.7, 1.3)	(0.7130, 1.2801)	(0.0976, 0.1348)
$N = 500$	(0.5, 1)	(0.5110, 0.9813)	(0.0625, 0.1163)
	(0.7, 1.3)	(0.7122, 1.2865)	(0.0920, 0.1284)

Table 4.8.2

Comportement des distributions asymptotiques

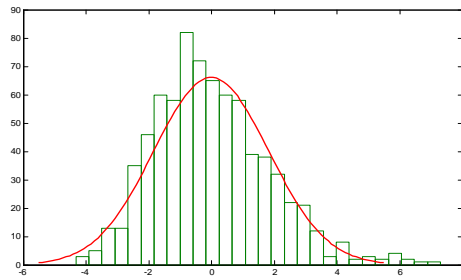
Nous avons vu que les résultats de simulation ont bien confirmé la consistance, au sens de la convergence en probabilité et presque sûrement, de l'estimateur MCO , $\hat{\alpha}_n$, du paramètre inconnu α , d'un $ARCH(1)$, et ceci dans les trois cas à savoir le cas de stabilité, d'instabilité à racine unitaire et d'instabilité explosif. Dans ce paragraphe, nous verrons que l'étude de simulation confirme empiriquement aussi les résultats théoriques établis concernant le comportement de la distribution asymptotique de cet estimateur dans les trois cas précédents. En effet, nous verrons que la loi empirique de l'estimateur $\hat{\alpha}_n$ converge, quand la taille n tend vers l'infini vers une loi normale pour les trois cas.

a) Distributions asymptotiques dans le cas de stabilité, pour $\alpha = (1, 0.2)$ et $\alpha = (0.3, 0.6)$

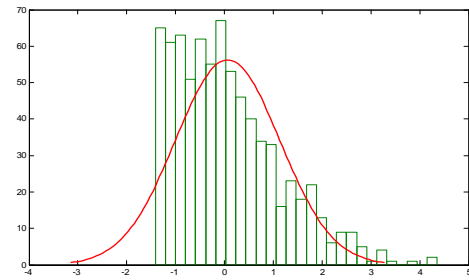
Les graphes des histogrammes et les densités de probabilité ajustées aux valeurs de l'estimateur du MCO , $\hat{\alpha}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 50, 100, 200, 250, 300, 350, 400, 450, 500$, générées d'un modèle $ARCH(1)$, correspondant aux valeurs appartenant à la région de stabilité $\alpha = (1, 0.2)$ et $\alpha = (0.3, 0.6)$ sont donnés dans la *Table 4.8.3* ci-dessous.

De ces différents graphes, nous constatons que les densités de probabilité de la variable aléatoire $\sqrt{n}(\hat{\alpha}_n - \alpha)$ approche, au fur et à mesure que n augmente, d'une densité de probabilité d'une variable aléatoire normalement distribuée.

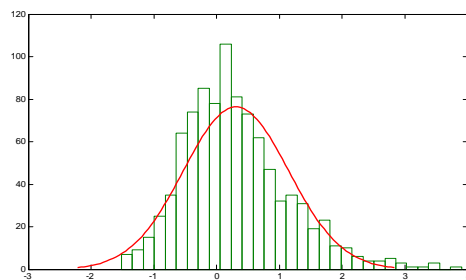
$N = 50$



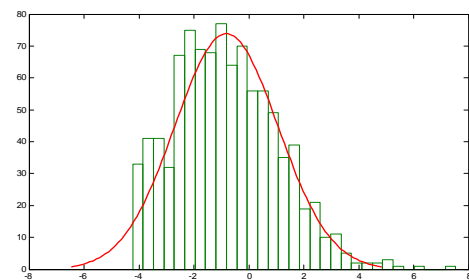
$\alpha_0 = 1$



$\alpha_1 = 0.2$

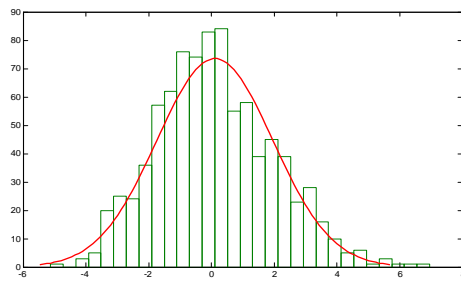


$\alpha_0 = 0.3$

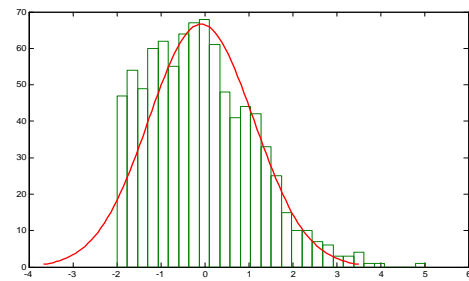


$\alpha_1 = 0.6$

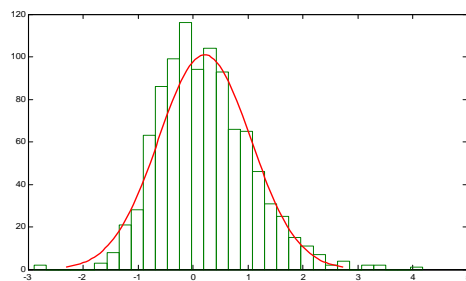
$N = 100$



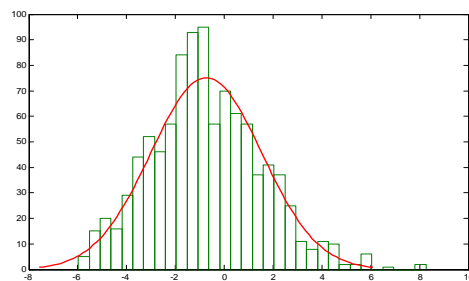
$\alpha_0 = 1$



$\alpha_1 = 0.2$



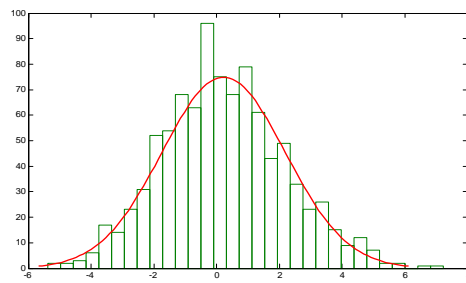
$\alpha_0 = 0.3$



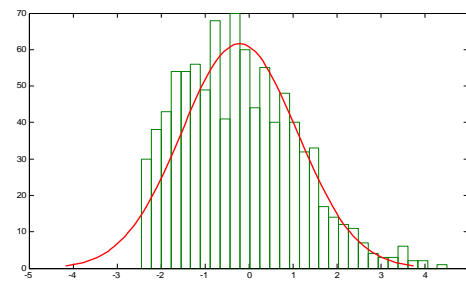
$\alpha_1 = 0.6$

Table 4.8.3

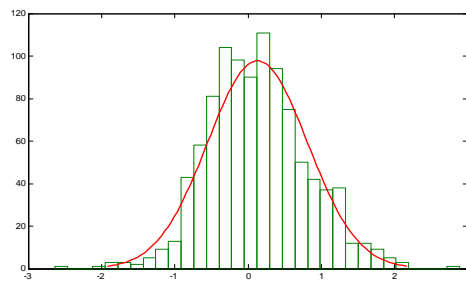
$N = 150$



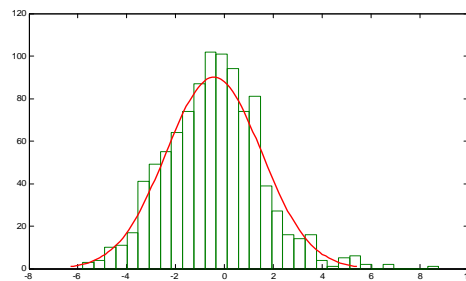
$\alpha_0 = 1$



$\alpha_1 = 0.2$

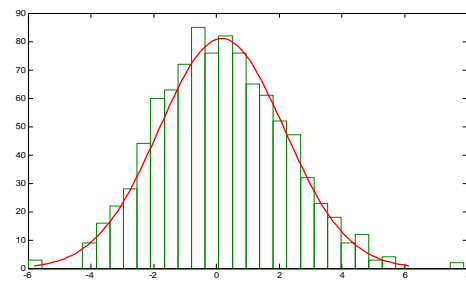


$\alpha_0 = 0.3$

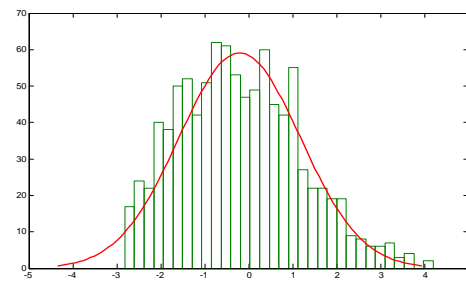


$\alpha_1 = 0.6$

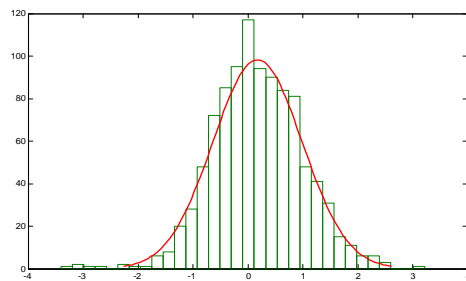
$N = 200$



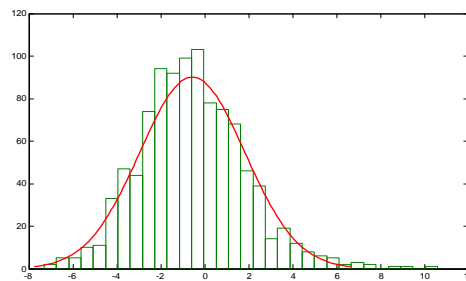
$\alpha_0 = 1$



$\alpha_1 = 0.2$



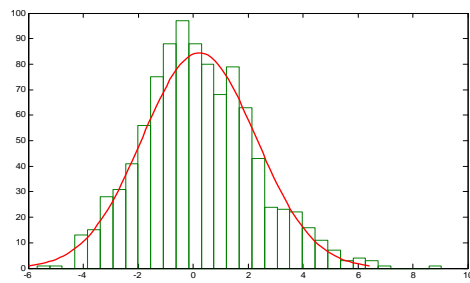
$\alpha_0 = 0.3$



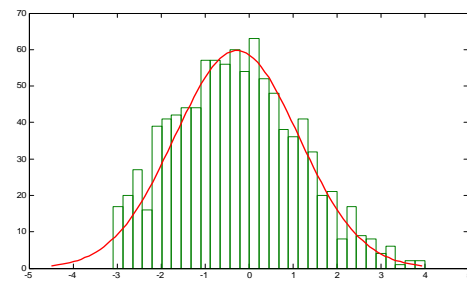
$\alpha_1 = 0.6$

Table 4.8.3 (suite)

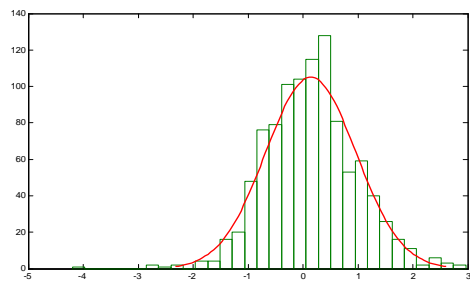
$N = 250$



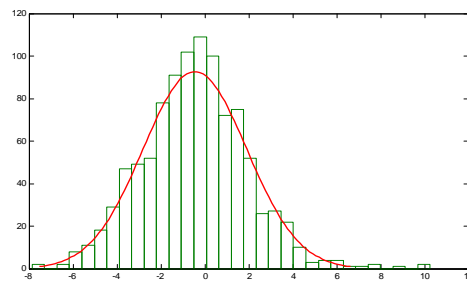
$\alpha_0 = 1$



$\alpha_1 = 0.2$

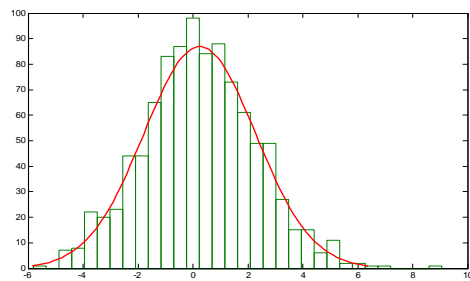


$\alpha_0 = 0.3$

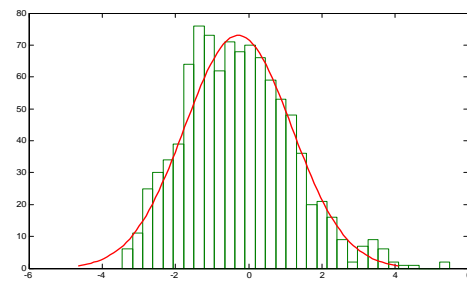


$\alpha_1 = 0.6$

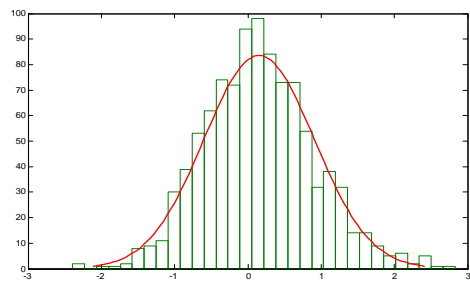
$N = 300$



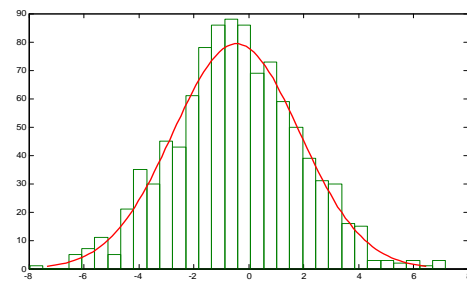
$\alpha_0 = 1$



$\alpha_1 = 0.2$



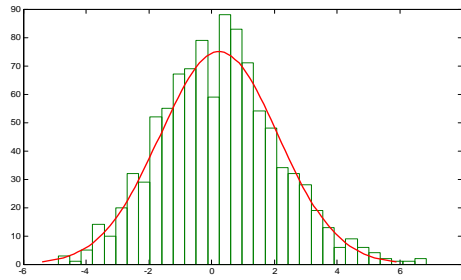
$\alpha_0 = 0.3$



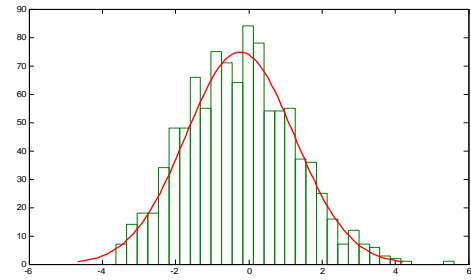
$\alpha_1 = 0.6$

Table 4.8.3 (suite)

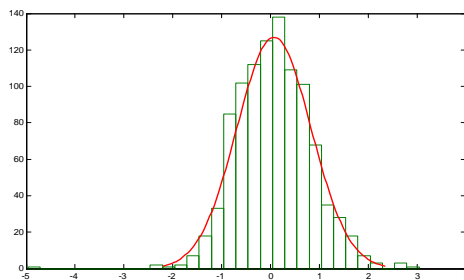
$N = 350$



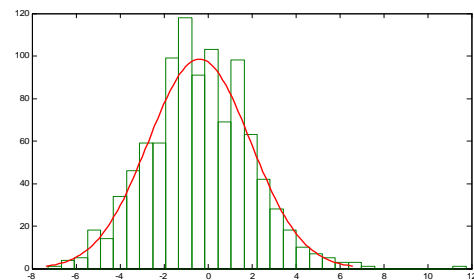
$\alpha_0 = 1$



$\alpha_1 = 0.2$

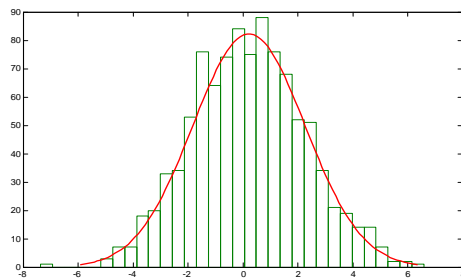


$\alpha_0 = 0.3$

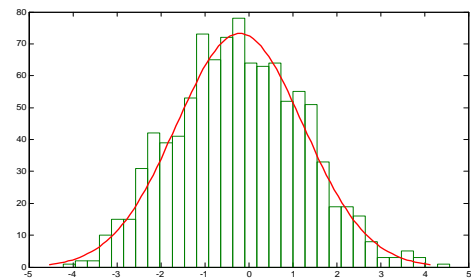


$\alpha_1 = 0.6$

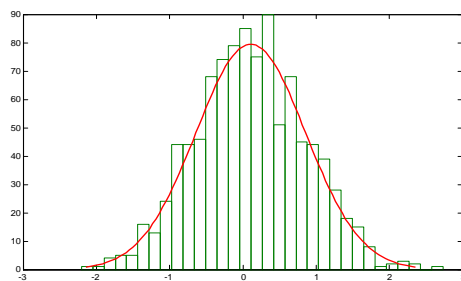
$N = 400$



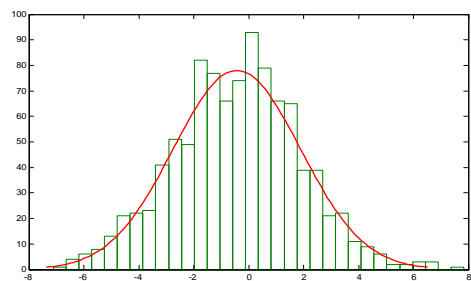
$\alpha_0 = 1$



$\alpha_1 = 0.2$



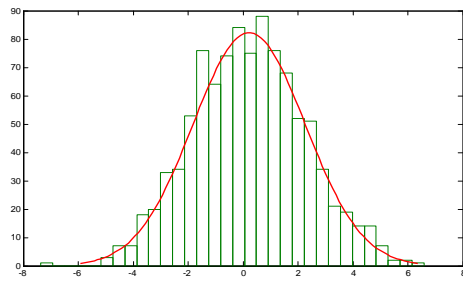
$\alpha_0 = 0.3$



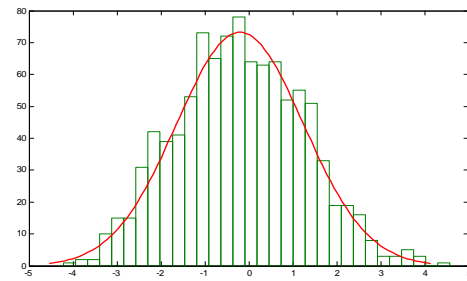
$\alpha_1 = 0.6$

Table 4.8.3 (suite)

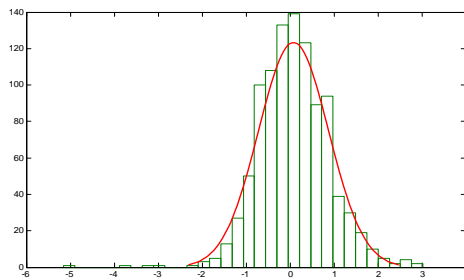
$N = 450$



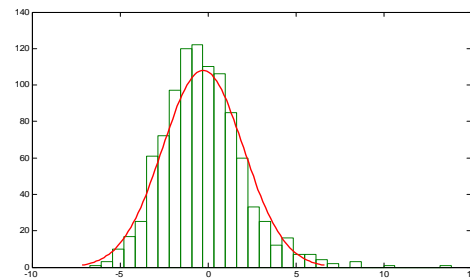
$\alpha_0 = 1$



$\alpha_1 = 0.2$

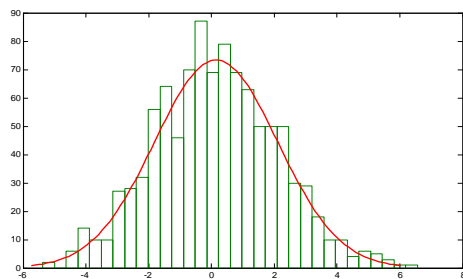


$\alpha_0 = 0.3$

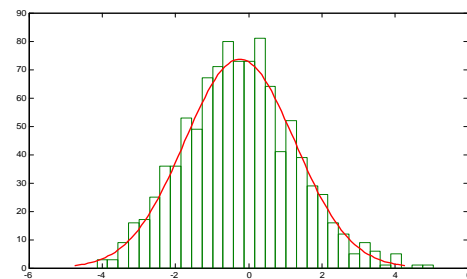


$\alpha_1 = 0.6$

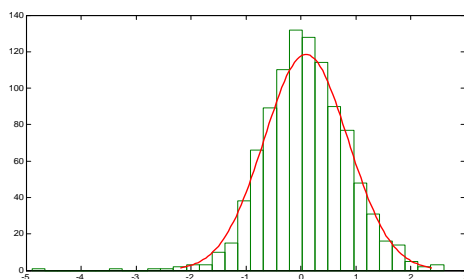
$N = 500$



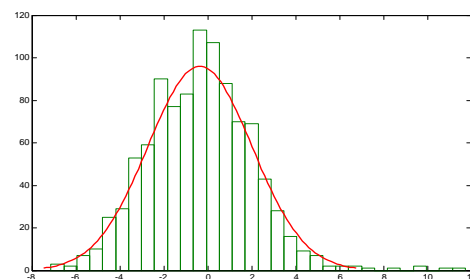
$\alpha_0 = 1$



$\alpha_1 = 0.2$



$\alpha_0 = 0.3$



$\alpha_1 = 0.6$

Table 4.8.3 (suite)

b) Distributions asymptotiques dans le cas d'instabilité (de racine unitaire $\alpha = (0.5, 1)$ et explosif $\alpha = (0.7, 1.3)$)

Les graphes des histogrammes des valeurs de l'estimateur du MCO, $\hat{\alpha}_n$, correspondant aux différentes séries temporelles artificielles de taille $n = 50, 100, 200, 250, 300, 350, 400, 450, 500$, générées d'un modèle ARCH(1), correspondant aux vecteurs $\alpha = (0.5, 1)$ et $\alpha = (0.7, 1.3)$ sont respectivement donnés dans les Table 4.8.4 ci-dessous. De ces différents graphes, nous constatons que les densités de probabilité de la variable aléatoire $\sqrt{n}(\hat{\alpha}_n - \alpha)$ sont de type Normal.

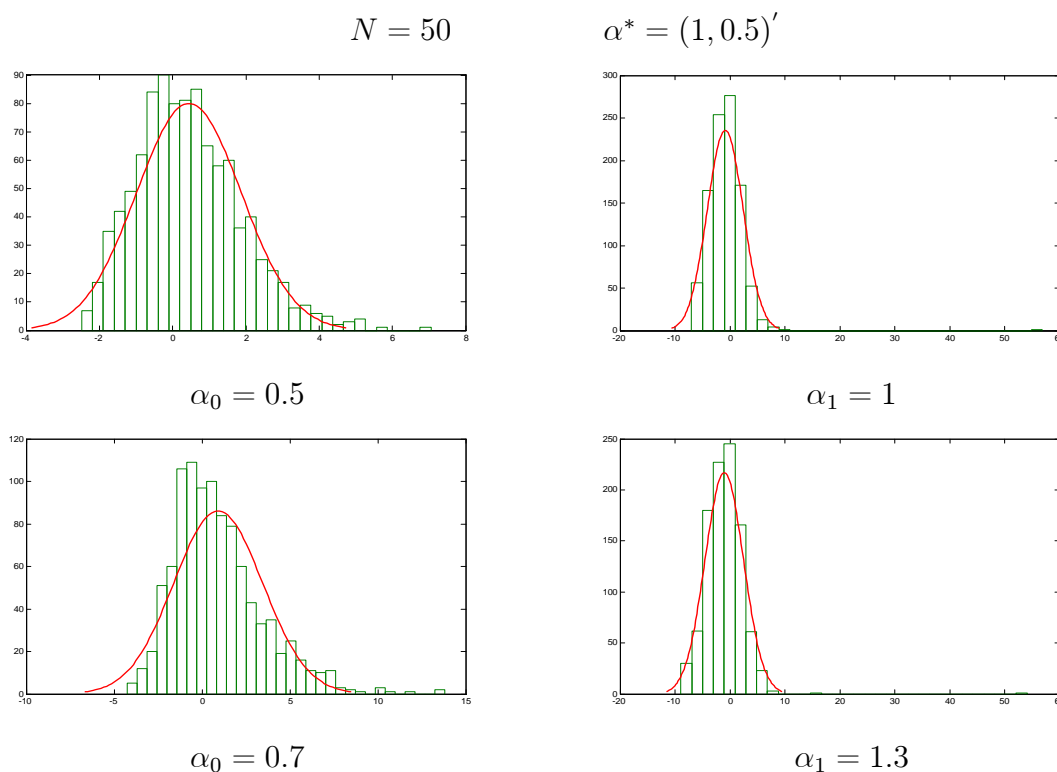
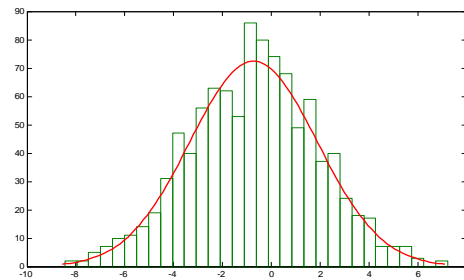
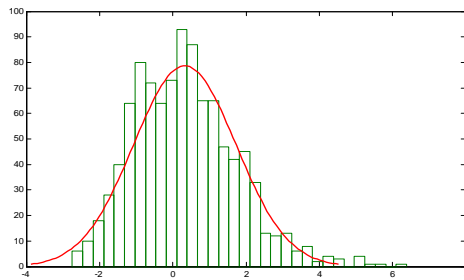


Table 4.8.4

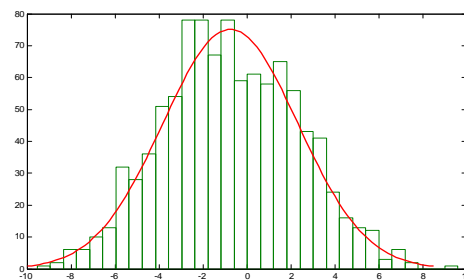
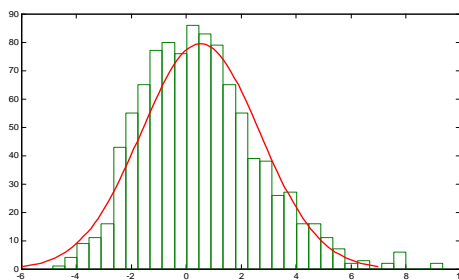
$N = 100$

$\alpha^* = (1, 0.5)'$



$\alpha_0 = 0.5$

$\alpha_1 = 1$

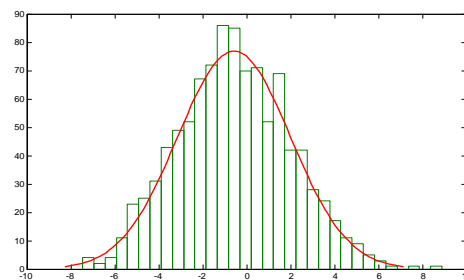
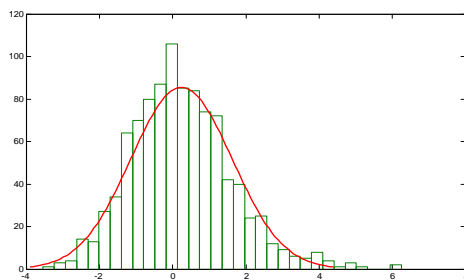


$\alpha_0 = 0.7$

$\alpha_1 = 1.3$

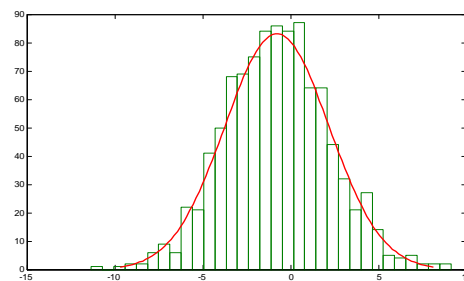
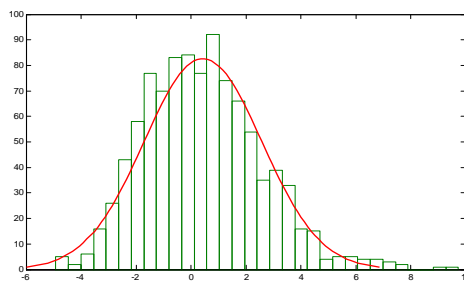
$N = 150$

$\alpha^* = (1, 0.5)'$



$\alpha_0 = 0.5$

$\alpha_1 = 1$



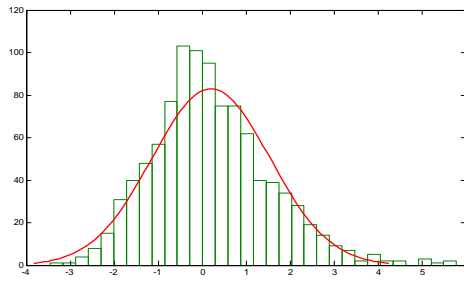
$\alpha_0 = 0.7$

$\alpha_1 = 1.3$

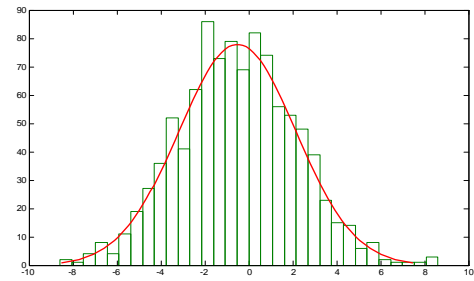
Table 4.8.4 (suite)

$N = 200$

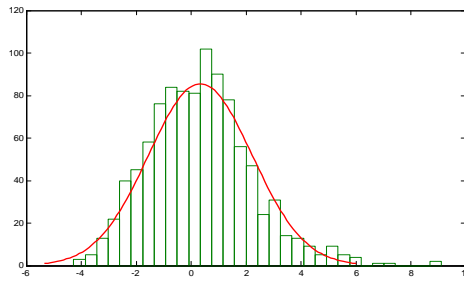
$\alpha^* = (1, 0.5)'$



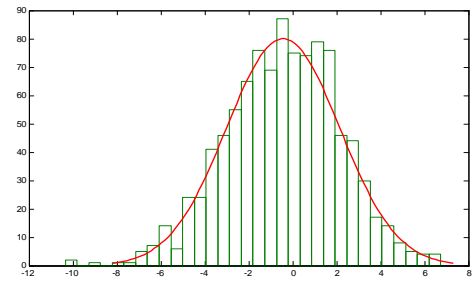
$\alpha_0 = 0.5$



$\alpha_1 = 1$



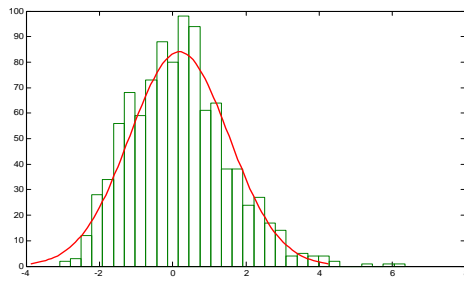
$\alpha_0 = 0.7$



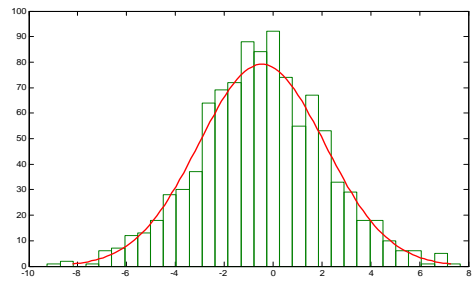
$\alpha_1 = 1.3$

$N = 250$

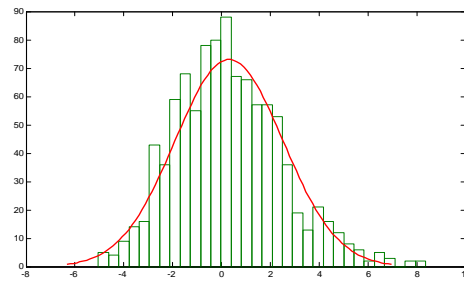
$\alpha^* = (1, 0.5)'$



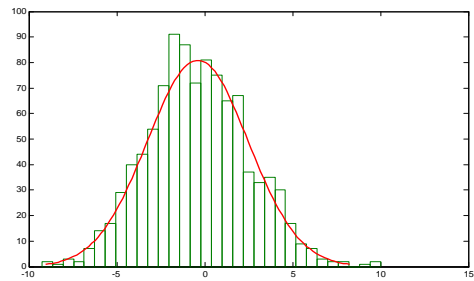
$\alpha_0 = 0.5$



$\alpha_1 = 1$



$\alpha_0 = 0.7$

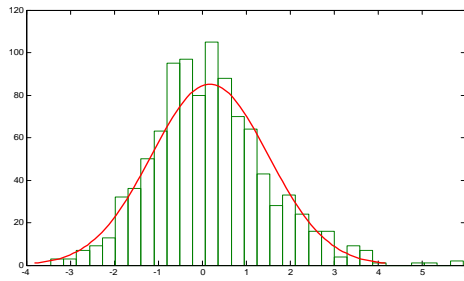


$\alpha_1 = 1.3$

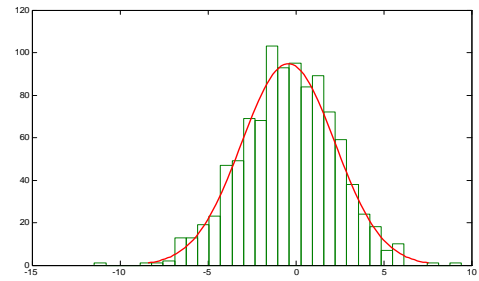
Table 4.8.4 (suite)

$N = 300$

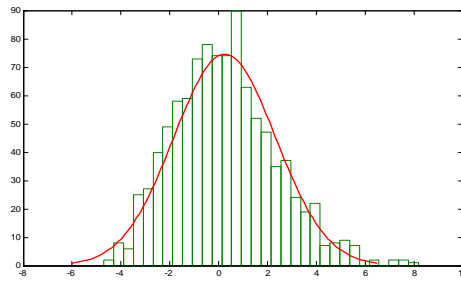
$\alpha^* = (1, 0.5)'$



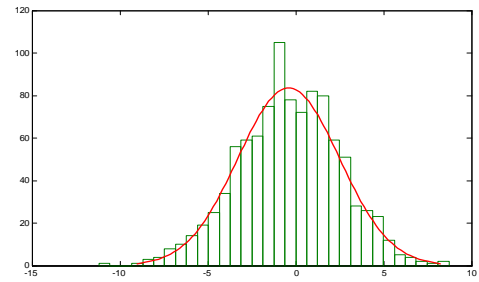
$\alpha_0 = 0.5$



$\alpha_1 = 1$



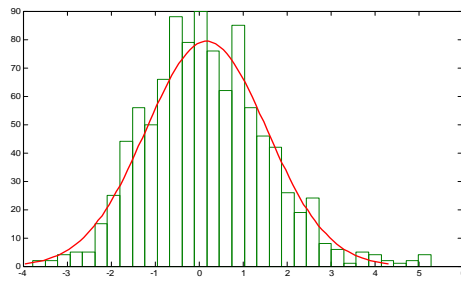
$\alpha_0 = 0.7$



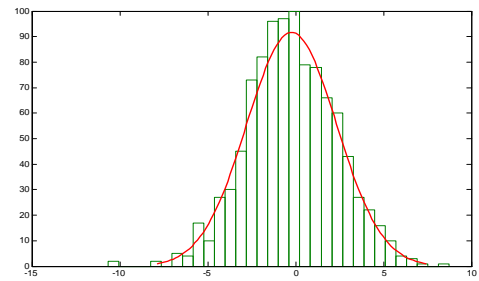
$\alpha_1 = 1.3$

$N = 350$

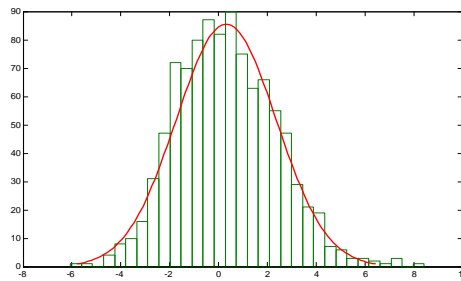
$\alpha^* = (1, 0.5)'$



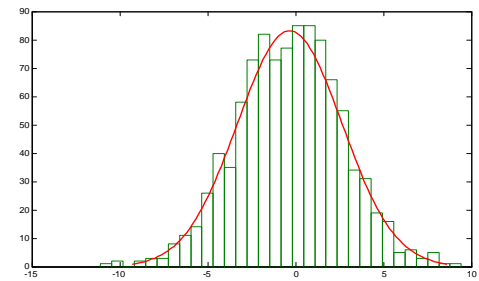
$\alpha_0 = 0.5$



$\alpha_1 = 1$



$\alpha_0 = 0.7$

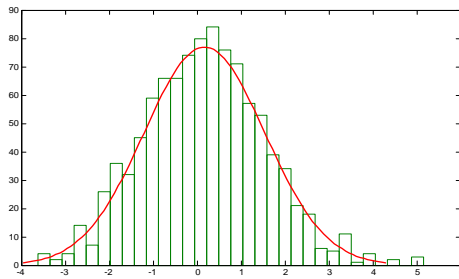


$\alpha_1 = 1.3$

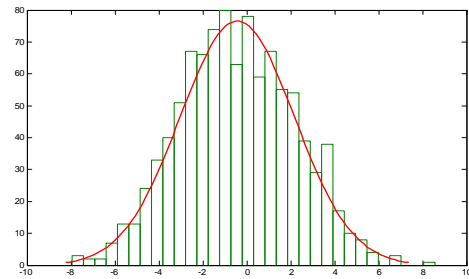
Table 4.8.4 (suite)

$N = 400$

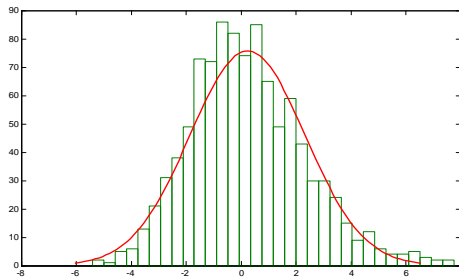
$\alpha^* = (1, 0.5)'$



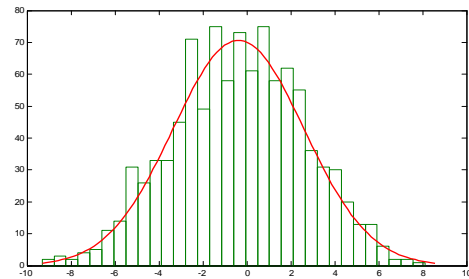
$\alpha_0 = 0.5$



$\alpha_1 = 1$



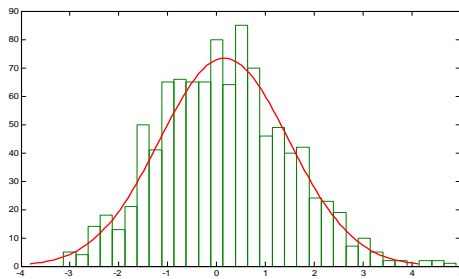
$\alpha_0 = 0.7$



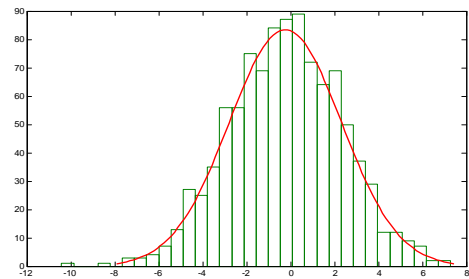
$\alpha_1 = 1.3$

$N = 450$

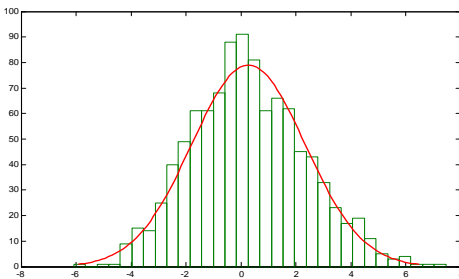
$\alpha^* = (1, 0.5)'$



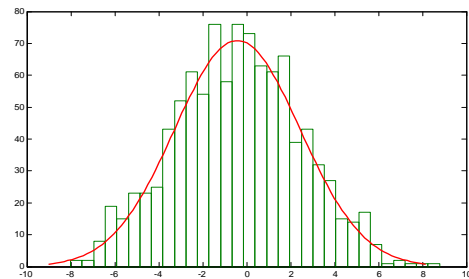
$\alpha_0 = 0.5$



$\alpha_1 = 1$



$\alpha_0 = 0.7$



$\alpha_1 = 1.3$

Table 4.8.4 (suite)

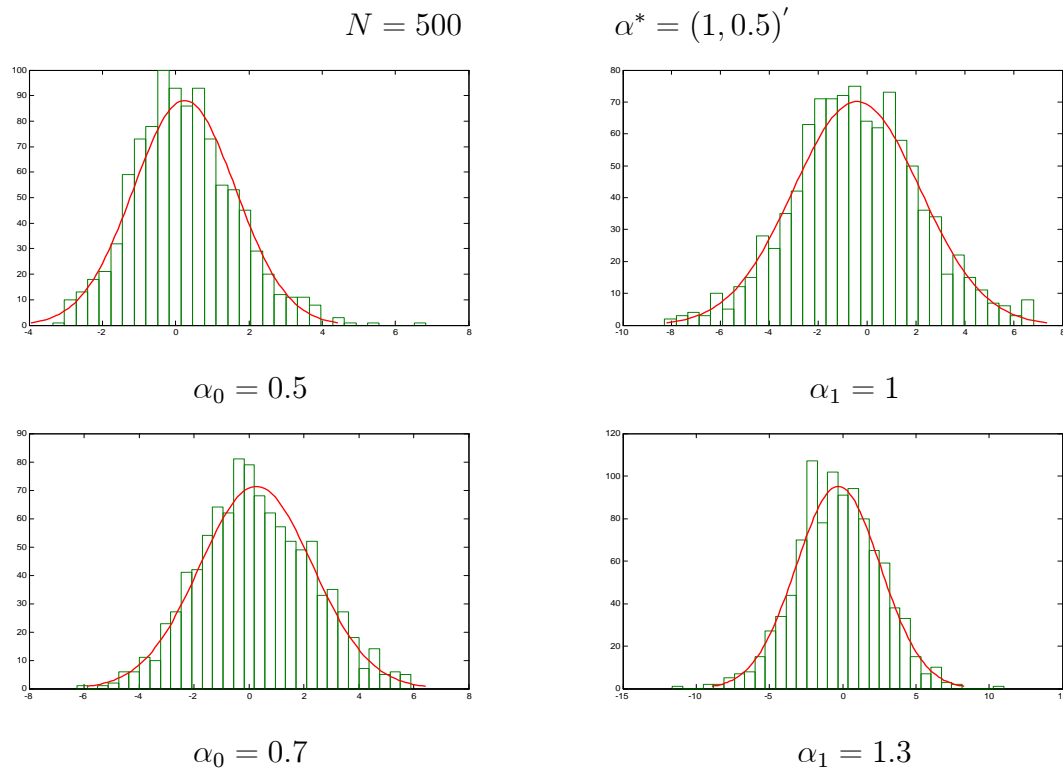
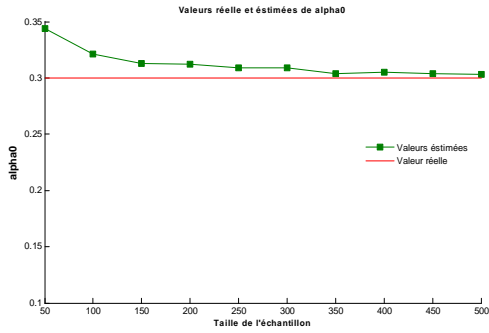


Table 4.8.4 (suite)

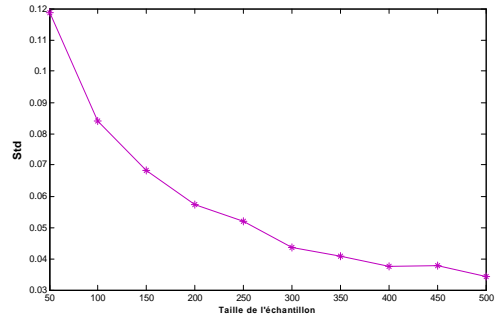
c) Comportement de l'estimateur $\hat{\alpha}_n$ et l'estimateur de son écart-type $\hat{\sigma}_{\hat{\alpha}_n}$ dans les trois régions

c. 1) Dans la région de stabilité $\alpha = (1, 0.2)$ et $\alpha = (0.3, 0.6)$

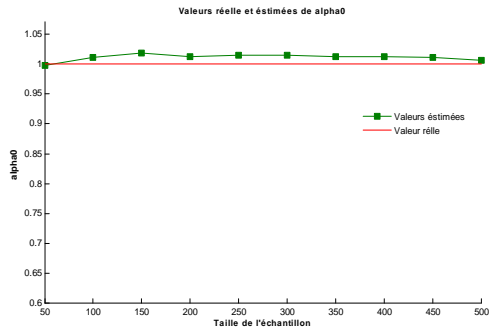
Les graphes de la Table 4.8.5 ci-dessous représente l'évolution de l'estimateur, $\hat{\alpha}_n$, des moindres carrés du paramètre α ainsi que l'évolution de l'estimateur de son écart-type $\hat{\sigma}_{\hat{\alpha}_n}$, pour $\alpha = (1, 0.2)$ et $\alpha = (0.3, 0.6)$, en fonction des séries temporelles simulées et commentées. A partir des graphes de la gauche de la table, nous constatons clairement que les valeurs estimées du paramètre approche rapidement de la vraie valeur du paramètre. Pareillement les estimations de l'écart-type de cet estimateur approche de la valeur zéro au fur et à mesure que n augmente.



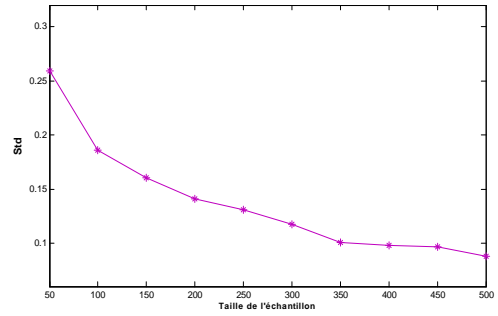
$\alpha_0 = 0.3$



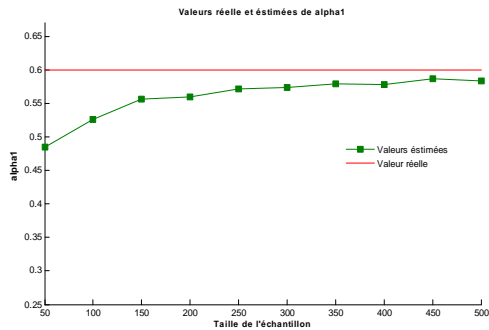
$Std(\hat{\alpha}_0)$



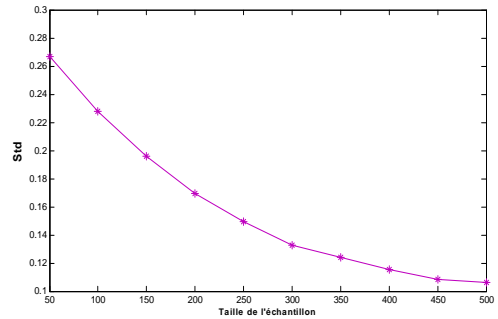
$\alpha_0 = 1$



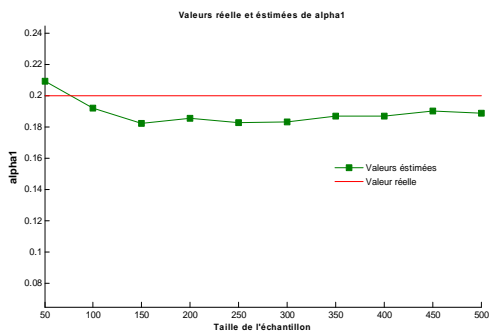
$Std(\hat{\alpha}_0)$



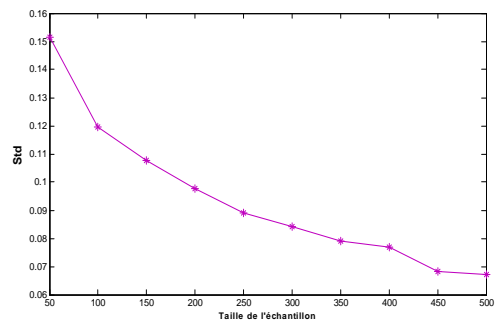
$\alpha_1 = 0.6$



$Std(\hat{\alpha}_1)$



$\alpha_1 = 0.2$



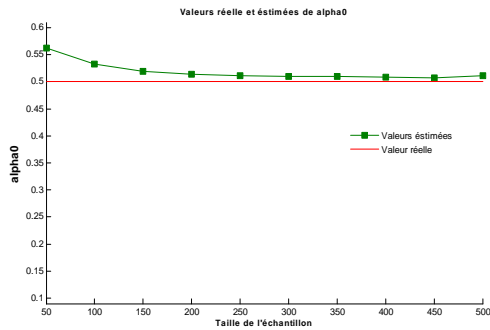
$Std(\hat{\alpha}_1)$

Table 4.8.5

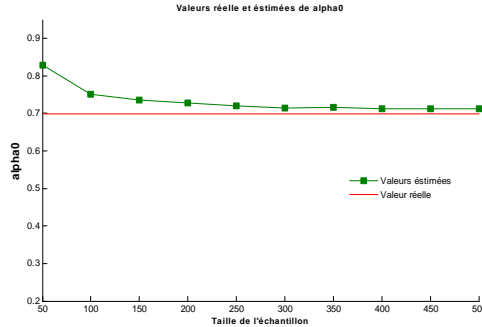
c. 2) Dans la région de racine unitaire $\alpha = (0.5, 1)$ et explosif $\alpha = (0.7, 1.3)$

Les graphes de la Table 4.8.6 ci-dessous représente l'évolution de l'estimateur, $\hat{\alpha}_n$, des moindres carrés du paramètre α ainsi que l'évolution de l'estimateur de son écart-type $\hat{\sigma}_{\hat{\alpha}_n}$, pour $\alpha = (0.5, 1)$ et $\alpha = (0.7, 1.3)$, en fonction des séries temporelles simulées et commentées.

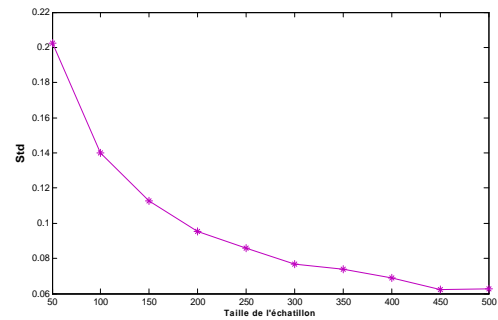
A partir des graphes de la gauche de la table, nous constatons clairement que les valeurs estimées du paramètre approche rapidement de la vraie valeur du paramètre. De même les estimations de l'écart-type de ces estimateurs approchent de la valeur zéro au fur et à mesure que n augmente.



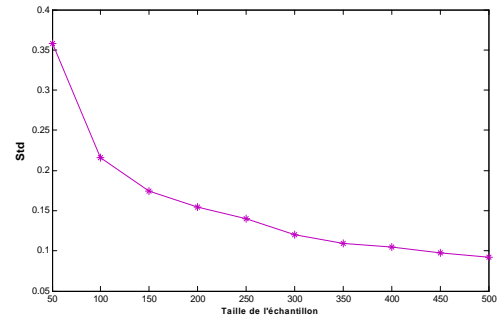
$\alpha_0 = 0.5$



$\alpha_0 = 0.7$

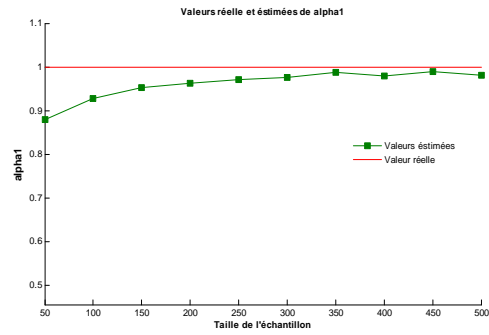


$Std(\hat{\alpha}_0)$

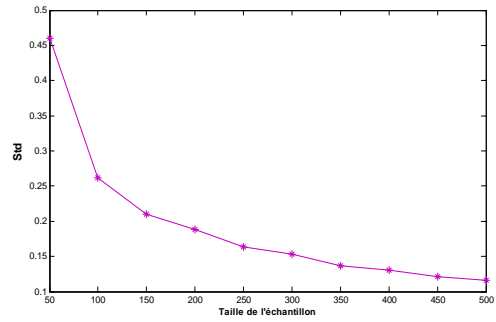


$Std(\hat{\alpha}_0)$

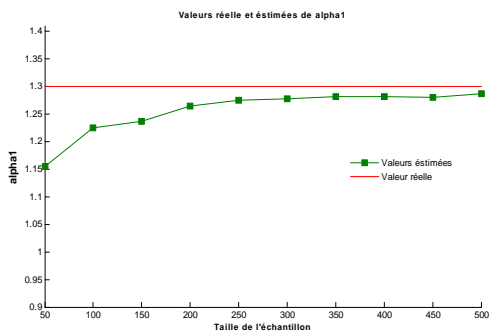
Table 4.8.6



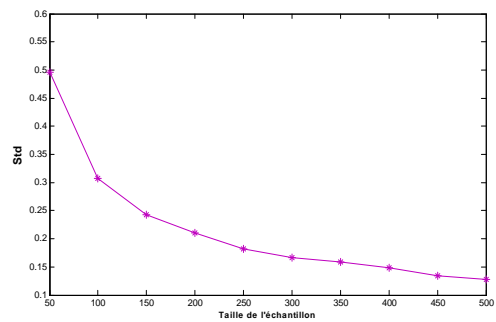
$$\alpha_1 = 1$$



$$Std(\hat{\alpha}_1)$$



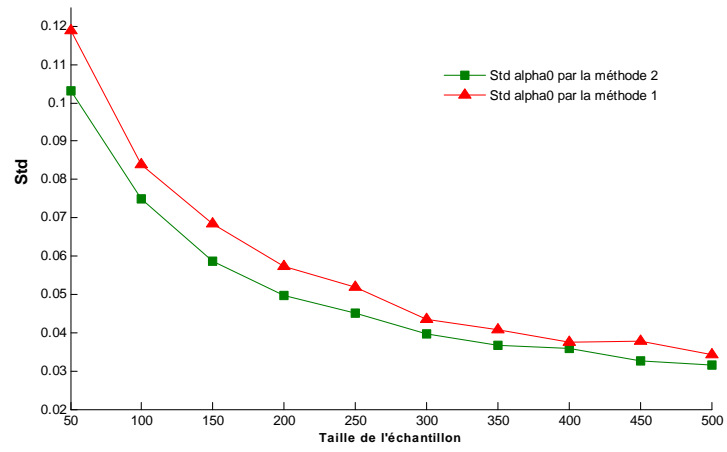
$$\alpha_1 = 1.3$$



$$Std(\hat{\alpha}_1)$$

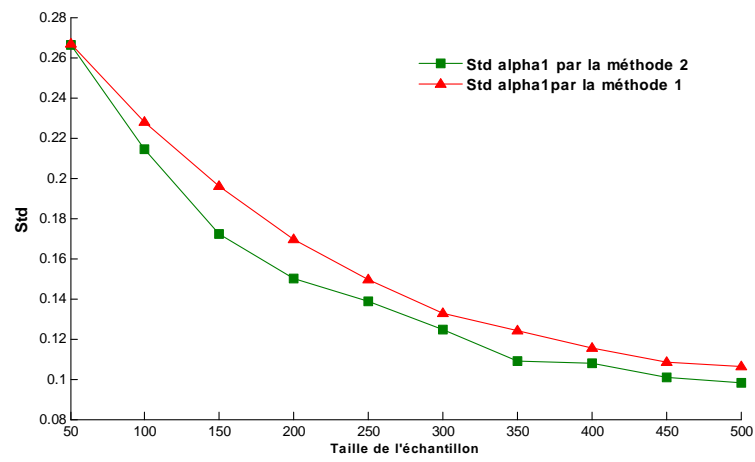
Table 4.8.6 (suite)

Les graphes ci-dessous montrent que la méthode $2S - WLS$ a amélioré les résultats d'estimation dans le cas stationnaire.



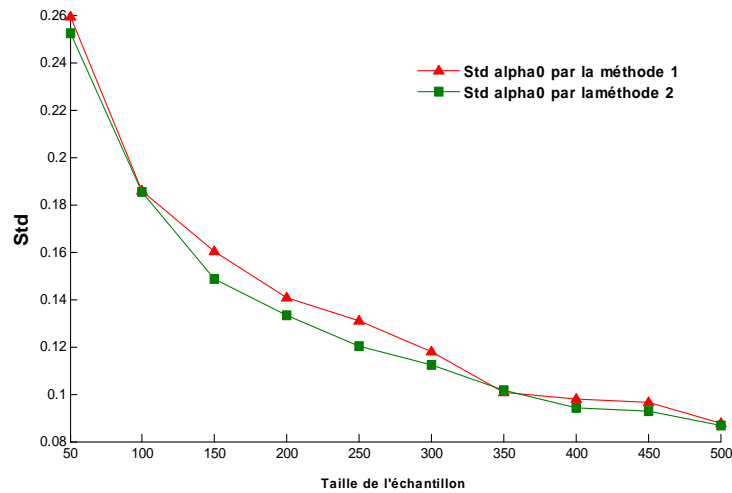
$\alpha_0 = 0.3$

$\alpha_0^* = 1$



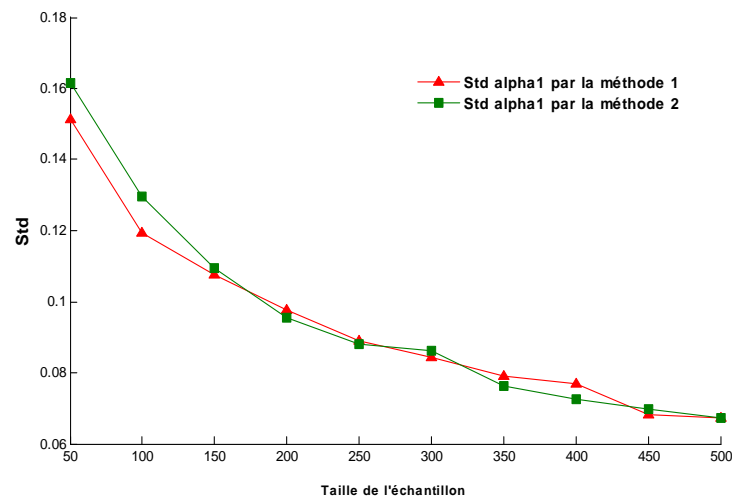
$\alpha_1 = 0.6$

$\alpha_1^* = 0.5$



$$\alpha_0 = 1$$

$$\alpha_0^* = 1$$



$$\alpha_1 = 0.2$$

$$\alpha_1^* = 0.2$$

La table ci-dessous présente les résultats d'estimation du paramètre α_1 par le $2S - WLS$ dans le cas où $\alpha_0 = 1$ et $\alpha_1 = 4$, c'est à dire lorsque α_1 est en dehors de la région de stabilité. De ces résultats nous concluons que l'estimateur $2S - WLS$ converge même dans le cas d'instabilité.

n	100	200	300	400	500	600	700
$\hat{\alpha}_1(WLS)$	4.0266	4.0043	3.9919	4.0012	4.0015	4.0002	4.0000
Std	0.5417	0.3850	0.3273	0.2817	0.2175	0.2273	0.2153

Conclusion

Dans ce mémoire nous avons étudié l'instabilité, ses différentes régions et son influence sur les propriétés probabilistes et statistiques de trois classes de modèles (équations) de séries chronologiques dont les deux premières sont linéaires et la troisième non linéaire de type conditionnellement hétéroscédastique. Nous pouvons conclure que pour les modèles linéaires, la condition de stabilité est la même par rapport à toutes les propriétés probabilistes mais elle est différente d'une propriété à une autre pour le cas de modèles non linéaires. En revanche, les propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance et des moindres carrés généralisés dépendent des régions de stabilité et instabilité dans le cas linéaire mais dans le cas non linéaire, elles sont insensibles à ces différentes régions.

Bibliographie

- [1] Aknouche, A. (2010). Efficient two-stage weighted least squares estimation of *ARCH* processes in the stable and instable cases. Preprint.
- [2] Aknouche A. and Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic *GARCH* and periodic *ARMA – GARCH* processes. *Journal of Time Series Analysis*, 28, 19 – 46.
- [3] Aknouche, A. and Bentarzi, M. (2008). On the existence of higher-order moments for periodic *GARCH* models. *Statistics and Probability Letters*, 78, 3262-3268.
- [4] Aknouche, A. and Guerbyenne, H. (2009). On some probabilistic properties of double periodic AR models. *Statistics & Probability Letters*, 79, 407-413.
- [5] Aknouche, A., (2006). *Etude théorique et algorithmique de la modélisation des séries chronologiques périodiques*. Thèse de Doctorat. Faculté de Mathématiques, U.S.T.H.B. Alger.
- [6] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. New York : John Wiley& Sons.
- [7] Babillot, M., Bougerol, P. and Elie, L. (1997). The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case. *Annals of Probability*, 25, 478-493.
- [8] Bentarzi, M. and Hallin, M. (1994). On the Invertibility of Periodic Moving Average Models. *Journal of Time Series Analysis*, **15**, 263-268.
- [9] Bentarzi, M. (1995). "*Modèles de séries chronologiques à coefficients périodiques*". Thèse de Doctorat es sciences. Institut de mathématiques, U.S.T.H.B, Alger, Algérie.
- [10] Bentarzi, M. and Hallin, M. (1994). On the Invertibility of Periodic Moving Average Models. *Journal of Time Series Analysis*. Vol 15, N°3, 263-268.

- [11] Bentarzi, M.(1998). Model-Building Problem of Periodically Correlated m-Variate Moving Average Processes. *Journal of Multivariate Analysis*. 66, 1-21.
- [12] Bera, A. K., et Higgins, M. L. (1993). ARCH Models : Properties, Estimation and Testing. *Journal of Economic Surveys*, 7 (4), 307-366.
- [13] Berkes, I., Horváth, L. and Kokoska, P. S. (2003). *Asymptotics in Statistics*. Springer-Verlag, New York.
- [14] Bibi, A. and Aknouche, A. (2009). On the probabilistic properties of periodic GARCH processes. *Mathematical Methods of Statistics*,17, 305-316.
- [15] Bloomfield, P., Hurd, H. L. and Lund, R. B.(1994). Periodic correlation in stratospheric ozone data. *J. Time. Ser. Anal.* 15,127-150.
- [16] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- [17] Brockwell, P. J. and Davis, R. A.(1991). *Time series : Theory and methods*, 2nd ed. New York : Springer.
- [18] Brockwell, P. J. and Davis, R. A.(2002). *Introduction to Time Series and Forecasting*, 2nd ed. New York : Springer.
- [19] Bose, A. and Mukherjee, K. (2003). Estimating the ARCH parameters by solving linear equations. *Journal of Time Series Analysis*, **24**, 127-136.
- [20] Boswijk, H. P. and Franses, P. H. (1996). Unit roots in periodic autoregressions. *Journal of time series analysis*, 17, 221 – 45.
- [21] Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115 – 27.
- [22] Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis, Forecasting and Control*. San Francisco : Holden-Day.
- [23] Cramér, H. (1961). On some classes of nonstationary stochastic processes. *Pro. 4th Berkeley Symp. Math. Statist. Prob*, Vol. 2, 57 – 78.
- [24] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with estimates of variance of U.K. Inflation. *Econometrica*, 50, 987 – 1008.

- [25] Engle R.F. (2001). The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, 15 (4), 157-168.
- [26] Feigin, P. D. and Tweedie, R. L. (1985) Random coefficient autoregressive processes :a Markov chain analysis of stationarity and finiteness of moments. *Journal of Time Series Analysis* 6, 1 – 14.
- [27] Fuller, W. A. (1996). *Introduction to Statistical Time Series*, 2nd ed. New York : John Wiley & Sons.
- [28] Francq, C. et Zakoïan, J. M. (2009). *Modèles GARCH : Structure, Inférence Statistique et Applications Financières*. Economica.
- [29] Ghysels, E., Hall, A. and Lee, H. S. (1996). On Periodic Structures and Testing for Seasonal Unit Roots. *Journal of the American Statistical Association*. Vol. 91, N.436, 1551 – 1559.
- [30] Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet. Math.*, 2, 385 – 88.
- [31] Gouriéroux, C. (1992). *Modèles ARCH et Applications Financières*. Collection ENSAE, Economica.
- [32] Guerbyenne, H. (2007). *Modèles de séries chronologiques conditionnellement hétéroscédastiques*. Thèse de Doctorat. Faculté de Mathématiques, U.S.T.H.B. Alger.
- [33] Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its application*. Academic press.
- [34] Hamilton, James. D. (1994). *Time series analysis*. Princeton University Press.
- [35] Hannan, E. J. (1970). *Multiple Time Series*. New York : John Wiley & Sons.
- [36] Hasza, D. P. (1977). *Estimation in nonstationary time series*. Ph.D. thesis, Iowa State University, Ames, Iowa.
- [37] Mann, H. B. and Wald, A (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica*, Vol. 11, pp. 173 – 200.
- [38] Osborn, D. R. and Smith, J. P. (1989). The performance of periodic autoregressive models in forecasting seasonal U. K. consumption. *Journal of Business and Economic Statistics*, 7, 117 – 127.

- [39] Phillip Hans Franses and Richard Paap. (2004). *Periodic time series models*. Oxford University Press Inc, New York.
- [40] Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica* 55, 277 – 301.
- [41] Schick, A. (1996). \sqrt{n} -consistent estimation in random coefficient autoregressive model. *The Australian Journal of statistics*, 38, 155 – 52.
- [42] Slutsky, E. (1927). "The summation of random causes as the source of cyclic processes" (in Russian). In *Problems of Economic Conditions, English translation* (1937) in *Econometrica*, 5, 105 – 146.
- [43] Soren Tolver Jensen and Anders Rahbek. (2004). Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica*, Vol. 72, 641 – 646
- [44] Ula, T. A. and Smadi, A. A. (1977). periodic stationary conditions for autoregressive moving average processes as eigenvalue problems. *Water Resources. Reserach*. 33.
- [45] Weiss, A. A. (1986). Asymptotic theory for ARCH models : estimation and testing. *Econometric Theory*, 2, 107 – 131.
- [46] White, J. S (1958). *The limiting distribution of the serial correlation coefficient in explosive case*. Vol. 29, pp 1188 – 1197.
- [47] White, H. (1984). *Asymptotic Theory for Econometricians*. New York Academic Press.
- [48] Wold, H. O. A. (1938). *A Study of the Analysis of Stationary Time Serie*. (2nd ed. 1954). Uppsala, Almqvist and Wiksell.
- [49] Wold, H. O. A. (1948). "On prediction in stationary time series." *The Annals of Mathematical Statistics*, 19, 558–567.
- [50] Yule, G. U. (1926). "Why do we sometimes get nonsense-correlations between time-series? — A study in sampling and the nature of time-series." *Journal of the Royal Statistical Society*, 89, 1, 1–63.487