



N° Réf :

Centre Universitaire
Abd Elhafid Boussouf Mila

Institut des Sciences et Technologie

Département de Mathématiques et Informatique

Mémoire préparé en vue de l'obtention du diplôme de Master

En : Mathématiques
Spécialité : Mathématiques Appliquées

Estimateur de Quasi Maximum de vraisemblance pour le model GJR-GARCH avec bruit $N(1,1)$

Préparé par :

Boulaiche Amina
Djamaa Waim

Soutenue devant le jury :

MOHAMED SALAH Abdelouaheb (Pr)	C.U.Abd Elhafid Boussouf	Président
BOULAROUK Yakoub (MCA)	C.U.Abd Elhafid Boussouf	Rapporteur
BOUOUDEN Rabeh (MCA)	C.U.Abd Elhafid Boussouf	Examineur

Année Universitaire : 2020/2021

Remerciement

Nous remercions Allah le tout puissant de nous avoir donné le courage et la santé afin de pouvoir élaborer ce modeste travail et nous souhaitons adresser nos remerciements les plus sincères aux personnes qui nous ont apportés leur aide et qui ont contribué à l'élaboration de ce mémoire.

Nous voudrions aussi exprimer nos remerciements sincères au nos encadrement Mr.Boularouk Yakoub qui nous a fait l'honneur d'être rapporteur de la mémoire, ainsi que pour son encadrement, ses nombreux conseils et son soutien constant tout au long de ce sujet.

Nous tenons à exprimer nos sincères remerciements a tous les professeurs de départements mathématique et informatique spécialité mathématique.

Nos sincères remerciements s'adressent également aux membres de jury pour avoir accepté d'évaluer ce travail.

En fin un grand merci a nos familles qui jouent le rôle le plus important soit dans ce travail soit dans nos vies personnelles, et en remercie tous nos collègues qui nous ont soutenues tout au long de nos études et tous ceux qui de près ou lion, ont contribué à la réalisation de ce mémoire.

Dédicace

Avant tout, je remercie Allah le tout puissant qui m'a donné la santé, le courage et la volenté de réaliser ce travail.

Je dédie ce travail à :

Mes chers parents qui à vrai dire aucune dédicace ne pourrait vous montrer ma gratitude, et mon respect, et mon amour envers vous,

A mes chers frères, ainsi que toute la famille Djamaa,

A notre encadrement pour son soutien, son encouragement et son support.

A mes amies, et mes collègues de la promotion 2021.

Djamaa Waim.

Dédicace

Avant tout, je remercie Allah le tout puissant qui m'a donné la santé, le courage et la volenté de réaliser ce travail.

Je dédie ce travail à :

Mes chers parents qui à vrai dire aucune dédicace ne pourrait vous montrer ma gratitude, et mon respect, et mon amour envers vous,

A mes chers frères, ainsi que toute la famille Boulaiche,

A notre encadrement pour son soutien, son encouragement et son support.

A mes amies, et mes collègues de la promotion 2021.

Boulaiche Amina.

Résumé

Dans ce mémoire,
Nous étudions l'estimateur de quasi-maximum vraisemblance d'un processus GJR-GARCH avec bruit $N(1,1)$. Pour cela, nous commençons par un aperçu général sur les notions de base et les définitions (processus stochastique, la loi Normale, processus GJR-GARCH ,...). Par suite, nous étudions la consistance et la normalité asymptotiques de l'estimateur de quasi-maximum de vraisemblance. Enfin, utilisant le programme R, nous confirmons les résultats théorique obtenu par des simulation numériques de type Monte-Carlo.

Mots clés :

Modèle GJR-GARCH, Fonction vraisemblance, Estimateur Quasi-Maximum de vraisemblance Normale, Normale(1,1), Consistance forte, Normalité asymptotique.

Abstact

In this dissertation,

We study the quasi-maximum likelihood estimator of a GJR-GARCH process with $N(1,1)$ noise. For this, we start with a general overview on the basics and definitions (stochastic process, Normal law, GJR-GARCH process,...). Therefore, we study the asymptotic consistency and normality of the quasi-maximum likelihood estimator. Finally, using the R program, we confirm the theoretical results obtained by numerical silmulation of Monte Carlo type.

Keywords :

GJR-GARCH model, Likelihood function, Normal Quasi-Maximum Likelihood Estimator, Strong consistency, Asymptotic normality.

Table des matières

1	Définitions	6
1.1	Processus stochastique	7
1.1.1	Stationnarité	7
1.1.2	Le processus bruit blanc (BB)	9
1.1.3	Causalité	10
1.1.4	Les séries chronologiques	10
1.2	Loi Normale	12
1.3	Fonction Gamma	13
1.4	Le maximum de vraisemblance	13
1.4.1	Estimateur	13
1.4.2	La vraisemblance	14
1.4.3	L'estimateur de quasi maximum de vraisemblance . . .	15
1.5	Inégalité de Cauchy-Schwarz	15
1.6	Convergence presque sûrement des variables aléatoires	15
2	Processus puissance asymétrique GARCH modèle GJR-GARCH	16
2.1	Modèle GJR-GARCH(p,q)	17
2.2	Définition et hypothèse	22
2.2.1	Définition d'estimateur	22
2.2.2	Existence et stationnarité	25
2.2.3	Hypothèses nécessaires pour la convergence d'EQMV- Normale	29
2.3	comportement asymptotique d'EQMV	30
2.3.1	Propriétaires asymptotiques de la quasi vraisemblance .	30
2.3.2	Consistance forte	30
2.3.3	La normalité asymptotique	36

3	Étude numériques	42
3.1	La consistance	43
3.2	La normalité asymptotique	46

Table des figures

3.1	Représentation en boîte de α_0	44
3.2	Représentation en boîte de α	45
3.3	Représentation en boîte de β	45
3.4	Représentation en boîte de γ	46
3.5	Représentation en histogramme pour différents paramètres des modèle.	47

Liste des tableaux

3.1	Racine-erreur quadratique moyenne des $\hat{\theta}_n^{(0)}$ et $\hat{\theta}_n$ pour les processus GJR-GARCH considérés.	43
-----	---	----

Notations générales

Notations

$(\Omega, A, \mathbb{P})$ espace de probabilité.

$\mathcal{F}$ filtration (suite de tribu).

Ensemble et espace

$\mathbb{N}, \mathbb{Z}, \mathbb{R}$ les entiers positives, les entiers, les nombres réels.

Fonction

γ_X la fonction d'autocovariance de (X_t) .

$\mathbb{I}$ fonction indicatrice.

Processus

i.i.d indépendant et identiquement distribué.

$N(1, 1)$ la loi normale de moyenne 1 et variance 1.

p.s presque sûrement.

Estimation

arg argument.

Θ l'ensemble de stationnarité.

θ l'ensemble des paramètres.

$\hat{\theta}$ estimateur de θ .

θ_0 la vraie valeur de paramètre.

Probabilité

$\xrightarrow{D}$ La convergence en distribution.

$\xrightarrow{p.s}$ LA convergence presque sûrement.

Introduction

La théorie des séries chronologiques est d'un usage constant quand on traite des données échantillonnées dans le temps. Les chronologiques (ou temporelles) sont appliquées de nos jours dans des domaines aussi variés que l'économétrie, la médecine ou la démographie, pour n'en citer qu'une petite partie. La modélisation représentant une étape cruciale dans l'étude des séries chronologiques. Elle a connu une grande évolution durant les dernières années et plusieurs modèles de représentation ont été proposés.

Depuis les travaux de Wold (1938), l'intérêt pour le développement de modèles de séries chronologiques, pouvant répondre aux besoins de l'utilisateur, a augmenté. Les modèles de séries chronologique linéaires à coefficients constants ont connu une ère de prospérité grâce, en particulier, aux travaux de Box et Jenkins (1970) et leur fameux ouvrage qui les a popularisés avec, en particulier, leur méthodologie : identification, estimation, validation. Ces modèles, qui supposent une variance des erreurs constante, ont vite montré leurs limites, en particulier, dans la modélisation des séries chronologiques macroéconomiques et financières on la focalisation sur les premiers moments conditionnels-les moments d'ordre supérieurs étant traités comme des paramètres de nuisance-supposés constants par rapport au temps, s'est révélée limitative. De plus, l'importance croissante motivée par les considérations sur le risque et sur l'incertitude dans la théorie économique moderne ont nécessité le développement de nouvelles techniques pour les séries chronologiques économétriques permettant à la variance et à la covariance de dépendre du temps. Ainsi est née, sous l'impulsion du génie d'Engle (1982), la classe des modèles ARCH (auto régressifs conditionnellement hétéroscédastiques) suggérés afin de saisir les caractéristiques particulières des séries de données d'observations financières. Les modèles ARCH font la distinction entre les moments du second ordre conditionnels et inconditionnels (marginaux). Le principe général proposé par Engle (1982) consiste à supposer que la variance dépend de

l'ensemble informationnel dont on dispose. Il propose une spécification ARCH (q) où le carré des perturbations suit un processus auto régressif d'ordre q. Les modèles ARCH sont donc des modèles auto régressifs conditionnellement hétéroscédastiques. Engle (1982) a donc proposer ces processus pour palier aux insuffisances de la classe des représentations ARMA, notamment en ce qui concerne les séries financières qui présentent une volatilité (ou variabilité instantanée mesurée par la variance conditionnelle) fonction du temps et par des ajustements asymétriques. Ainsi, les modèles ARCH sont basés sur une paramétrisation endogène de la variance conditionnelle. La famille des modèles ARCH peut se décomposer en deux sous-ensembles : les modèles ARCH linéaires et les modèles ARCH non linéaires.

Les premiers reposent sur une spécification quadratique de la variance conditionnelle des perturbations : modèles ARCH(q), GARCH(p,q) et IGARCH(p, q). Les modèles ARCH non linéaires sont caractérisés par des spécifications asymétriques des perturbations. Ce sont les modèles EGARCH(p,q), TARCH(q) et TGARCH(p,q). (Bresson et Pirotte, Séries temporelles).

le modèle $ARCH(q)$ introduit Engle (1982). On considère un processus X_t tel que :

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2.$$

ou ε_t est une série non corrélée de moyenne nulle, et de variance conditionnelle h_t^2 qui varie avec le temps.

le modèle ARCH a été généralisé par Bollerslev dans l'article (étudiant de Engle) en 1986, est appelé GARCH(p,q) qui s'écrit sous la forme :

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i h_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \varepsilon_t^2,$$

ou h_t est la variance conditionnelle à la date t. $\alpha_0 > 0$, $\alpha_i \geq 0$ pour $i = 1, \dots, q$. $\beta_j \geq 0$ pour $j = 1, \dots, p$. Avec p, q deux entiers positifs non nuls et ε_t obéit à une $N(0,1)$ les modèles GARCH les plus populaires, qui se divisent en deux catégories : symétriques et asymétriques.

- GARCH asymétriques

Constatation Empirique : L'accroissement de volatilité dû à une baisse de prix est généralement supérieur à celui résultant d'une hausse de même ampleur. L'idée derrière les modèles asymétriques est toute simple : l'effet hétéroscédastique n'est sans doute pas le même suivant que l'erreur précédente

est positive ou négative. On note parmi les modèles asymétriques, ceux qui sont le plus utilisés, les modèles EGARCH, AGARCH, NGARCH et GJR-GARCH.

Afin de prendre en compte la modification d'un coefficient selon la survenue d'un évènement, il est courant d'introduire une nouvelle explicative construite comme produit d'une indicatrice de l'évènement en question et de la variable initiale. C'est l'idée adoptée par Glosten, Jagannathan et Runkle (1993) qui, partant de l'écriture GARCH(p,q) de base, arrivent à introduire le modèle suivant :

$$h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i X_{t-i}^2 + \gamma_i \mathbb{I}_{X_{t-i} < 0} X_t^2) + \sum_{i=1}^p \beta_i h_{t-i}.$$

Le modèle GJR-GARCH capte l'effet d'asymétrie des perturbations sur la variance conditionnelle. Dans leurs études sur l'indice de la valeur pondérée par la capitalisation boursière de CRSP, Glosten, Jagannathan et Runkle [Glosten et al., 1993] trouvent que les chocs négatifs provoquent une augmentation de la variance conditionnelle plus forte que les chocs positifs. De plus, au sujet de l'impact de la variance conditionnelle sur l'espérance conditionnelle du taux de rendements excédentaires, le coefficient estimé est négatif, comme dans l'article de Nelson (1991). Cette relation négative entre les rendements conditionnels et la variance conditionnelle peut être expliquée par l'effet de levier, ou bien par la volatilité rétroactive qui suggère qu'une hausse anticipée de la volatilité accroisse le rendement exigé par les investisseurs, puis-que le titre deviendra plus risqué, et ceci implique que la valeur du titre diminue immédiatement, toutes choses étant égales par ailleurs.

Ce mémoire étudie principalement le QMLE pour le modèle GJR-GARCH basé sur la distribution de Normale(1,1) organisé comme suit :

- Dans le première chapitre nous avons présenté des définitions principaux , propriétés des processus stochastiques des séries chronologiques et quelques notions de base dans le premier chapitre.
- Dans le deuxième chapitre nous avons défini l'estimateur de quasi-maximum vraisemblance Normale. En suit, nous étudions l'existence et la stationnarité de processus GJR-GARCH. En fin, nous prouvons la consistance forte et la normalité asymptotique de cet l'estimateur.

- Dans le troisième chapitre, nous avons effectué une étude numérique à l'aide de langage R pour vérifier les résultats théorique obtenus dans le deuxième chapitre.

Chapitre 1

Définitions

Dans ce chapitre nous définissons tout d'abord certain concepts et propriétés fondamentales (processus stochastiques, séries chronologiques, loi normale (Gauss), Estimateur, la vraisemblance, EQMV, ...) dans un contexte d'estimation de quasi maximum vraisemblance d'un processus GJR-GARCH avec erreur normale (1,1).

1.1 Processus stochastique

La théorie des processus stochastiques tente de construire des modèles de probabilité pour des phénomènes qui évoluent au fil du temps.

Définition 1.1.1 *Soit une famille des variables aléatoires $X(t)$ obéissant à la même loi de probabilité et caractérisées par un paramètre réel t . En générale, t est une date temporelle et $t \in [0, +\infty[$. Une telle famille s'appelle un "processus stochastique".*

On peut classer les processus stochastique en deux classes :

- **Les processus à temps discrète** : t ne prend que des valeurs discrètes.
- **Les processus à temps continu** : t varie continûment de 0 à $+\infty$. La famille des $X(t)$ comprend donc une infinité de variable aléatoires.

Chaque variable $X(t)$ peut prendre plusieurs valeurs possibles avec une certaine probabilité. Ces valeurs représentent des états. Ces derniers peuvent être discrets ou continus.

Définition 1.1.2 (L'opérateur de retard)

L'opérateur de retard L appliqué à une série chronologique génère la valeur précédente :

$$LY_t = Y_{t-1},$$

pour tout $t > 1$.

Si l'opérateur est appliqué k fois, on obtient :

$$L^k Y_t = Y_{t-k}.$$

1.1.1 Stationnarité

Espérance mathématique

Définition 1.1.3 *Soit X une variable aléatoire réelle continue admet une densité de probabilité f_X si :*

$$E\{|X|\} = \int_{\mathbb{R}} |x|f_X(x)dx < +\infty.$$

L'espérance mathématique de X est défini par :

$$\begin{aligned} E\{X\} &= \int_{\Omega} X dp, \\ &= \int_{\mathbb{R}} xf_X(x)dx, \end{aligned}$$

avec Ω l'ensemble des variables aléatoires, P la loi de probabilité.

Définition 1.1.4 (Moment d'ordre r)

Soit X un v.a.r continue de densité de probabilité f_X et $r \in \mathbb{N}^*$ si :

$$\int_{\mathbb{R}} |x|^r f_X(x)dx < +\infty.$$

On dit que X a un moment d'ordre r :

$$\begin{aligned} E\{X^r\} &= \int_{\Omega} X^r dp, \\ &= \int_{\mathbb{R}} x^r f_X(x)dx. \end{aligned}$$

Et le moment d'ordre r centré est donné par :

$$\begin{aligned} E\{X^r\} &= \int_{\Omega} (X - \mu)^r dp, \\ &= \int_{\mathbb{R}} (x - \mu)^r f_X(x)dx, \end{aligned}$$

où $\mu = E\{X\}$.

La stationnarité au second ordre (au sens faible)

Définition 1.1.5 Soit X_t un processus stochastique

- La moyenne de X_t est définie par :

$$\mu_X(t) = E\{X_t\}.$$

- La fonction d'autocovariance X_t est donnée par :

$$\begin{aligned}\gamma_X(r, s) &= \text{cov}(X_r, X_s), \\ &= \{(X_r - \mu_X(r))\} E\{X_s - \mu_X(s)\}.\end{aligned}$$

Définition 1.1.6 Si processus $(X_t)_{t \in \mathbb{Z}}$ admet un moment d'ordre deux, on dit que X_t est stationnaire au sens faible (ou stationnaire au second ordre) lorsque les deux propriétés suivantes sont simultanément vérifiées :

1. $\forall t \in \mathbb{Z}, \quad \mu_X(t)$ indépendante du temps.
2. $\forall t \in \mathbb{Z}, h \in \mathbb{Z}, \quad \gamma_X(t, t+h)$ indépendante du temps.

La stationnarité au sens stricte(fortement stationnaire)

Définition 1.1.7 Un processus stochastique $(X_t)_{t \in \mathbb{Z}}$ est dit strictement stationnaire (fortement stationnaire) si $X_{t_1}, \dots, X_{t_n}$ et $X_{t_1+h}, \dots, X_{t_n+h}$ ont la même loi de probabilité.

Remarque 1.1.1 Si X_t est strictement stationnaire alors X_t est également faiblement stationnaire.

1.1.2 Le processus bruit blanc (BB)

Définition 1.1.8 Un processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un Un bruit blanc faible s'il satisfait les trois conditions suivantes :

1. $\forall t \in \mathbb{Z}, \quad E\{\varepsilon_t\} = 0.$
 2. $\forall t \in \mathbb{Z}, h \in \mathbb{Z}, \quad \gamma_\varepsilon(h) = E\{\varepsilon_t \varepsilon_{t-h}\} = \begin{cases} \sigma^2 & \text{si } h = 0 \\ 0 & \text{sinon.} \end{cases}$
- Si (ε_t) est i.i.d alors (ε_t) est un BB fort.

Remarque 1.1.2

- Un BB faible est faiblement stationnaire .
- Un BB fort est fortement stationnaire.

1.1.3 Causalité

Définition 1.1.9 *Un processus stochastique X_t est dit causal s'il existe une suite des constantes $(\alpha_k)_{k \geq 0}$ tel que $\sum_{k \in \mathbb{Z}} |\alpha_k| < \infty$, avec*

$$X_t = \sum_{k=0}^{\infty} \alpha_k \eta_{t-k}.$$

1.1.4 Les séries chronologiques

On appelle série chronologique (temporelle), une suite finie de données quantitatives indexée par le temps. L'indice temps peut être selon les cas, la seconde, la minute, l'heure, etc. Dans des nombreux domaines variés que l'économie, finance, démographie, biologie, médecine, météorologie, pollution....

Définition 1.1.10 *Une série temporelle (ou série chronologique) est une réalisation d'un processus aléatoire indexé par le temps, noté X_t . Pour chaque t , X_t est une variable aléatoire dont on a une réalisation x_t .*

Classification des séries chronologiques

On peut classer les séries chronologiques selon des critères variés : domaines d'application, séries réelles / complexes, séries stationnaires ou non stationnaires (avec tendances, avec facteurs saisonniers, processus intégrés), séries représentées de façon temporelle ou spectrale.

Objectifs principaux dans l'étude des séries chronologiques

Les principaux objectifs de la modélisation des séries temporelles sont les suivants :

- Décrire. Par exemple :
 - En économétrie, détecter puis analyser les périodes de crises et croissances ;
 - En reconnaissance vocale, reconnaître les mots dans des signaux ;

- Dans le séquençage du génome, détecter les parties de l'ADN qui contiennent de l'information.
- Comparer deux séries temporelles. Par exemple, l'évolution démographique de deux régions ou deux séquences d'ADN.
- Prédire l'évolution future de la série temporelle à partir de celles qui ont été observées. Par exemple, la température ou le cours d'une action du lendemain ou l'évolution de la population mondiale au cours du siècle prochain.

Modélisation d'une séries chronologiques

Un modèle est une image simplifiée de la réalité qui vise à traduire les mécanismes de fonctionnement du phénomène étudié et permet de mieux les comprendre.

On distingue principalement deux types de modèles :

1. **les modèles déterministes** : Ces modèles relèvent de la Statistique Descriptive. Ils ne font intervenir que de manière sous-jacente le calcul des probabilités et consistent à supposer que l'observation de la série à la date t est une fonction du temps t et d'une variable ϵ_t centrée faisant office d'erreur au modèle, représentant la différence entre la réalité et le modèle proposé :

$$X_t = f(t, \epsilon_t).$$

On suppose de plus que les ϵ_t sont décorréliées.

Les deux modèles de ce type les plus usités sont les suivants :

- **Le modèle additif** : C'est le "modèle classique de décomposition" dans le traitement des modèles d'ajustement. La variable X_t s'écrit comme le somme de trois termes :

$$X_t = Z_t + S_t + \epsilon_t,$$

où Z_t représente la tendance (déterministe), S_t la saisonnalité (déterministe aussi) et ϵ_t les composantes ("erreurs au modèle") aléatoires i.i.d.

- **Le modèle multiplicatif** : La variable X_t s'écrit au terme d'erreur près comme le produit de la tendance et d'une composante de saisonnalité :

$$X_t = Z_t(1 + S_t)(1 + \epsilon_t).$$

L'ajustement est ici multiplicatif et intervient dans les modèles GARCH.

2. **Les modèles stochastiques** : Ils sont du même type que les modèles déterministes à ceci près que les variables de bruit ϵ_t ne sont pas i.i.d mais possèdent une structure de corrélation non nulle : ϵ_t est une fonction des valeurs passées ($\pm$ lointaines suivant le modèle) et d'un terme d'erreur η_t

$$\epsilon_t = g(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \eta_t).$$

1.2 Loi Normale

Les lois normales sont parmi les lois de probabilité les plus utilisées pour modéliser des phénomènes naturels issus de plusieurs événements aléatoires. Elles sont en lien avec de nombreux objets mathématiques dont le mouvement brownien, le bruit blanc gaussien ou d'autres lois de probabilité. Elles sont également appelées lois **gaussiennes**, **lois de Gauss** ou **lois de Laplace-Gauss** des noms de Laplace (1749-1827) et Gauss (1777-1855), deux mathématiciens, astronomes et physiciens qui l'ont étudiée.

Définition 1.2.1 *La loi normale est une loi dont nous ferons grand usage. Elle est utilisée pour une variable aléatoire continue.*

- *Elle est définie à partir de densité de probabilité :*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

où μ représente la moyenne et σ l'écart-type. Cette loi dépend donc deux paramètres, μ et σ .

et on note :

$$X \sim N(\mu, \sigma^2).$$

1.3 Fonction Gamma

Fonction gamma, généralisation de la fonction factorielle aux valeurs non intégrales, introduite par le mathématicien suisse Leonhard Euler au 18^{ème} siècle.

Pour un nombre entier positif n , la factorielle (écrire comme $n!$) est définie par $n! = 1 * 2 * 3 * \dots * (n - 1) * n$. Mais cette formule n'a pas de sens si n n'est pas un entier.

Pour étendre la factorielle à tout nombre réel $x > 0$ (que x soit un nombre entier ou non), la fonction gamma est définie comme :

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} \exp(-t) dt.$$

1.4 Le maximum de vraisemblance

1.4.1 Estimateur

En mathématiques, un estimateur est une statistique permettant d'évaluer un paramètre inconnu relatif à une loi de probabilité (comme son espérance ou sa variance). Il peut par exemple servir à estimer certaines caractéristiques d'une population totale à partir de données obtenues sur un échantillon comme lors d'un sondage. La définition et l'utilisation de tels estimateurs constitue la statistique différentielle.

Définition 1.4.1 • *Estimer la paramètre θ consiste à donner une valeur approchée à ce paramètre.*

- *Une statistique de l'échantillon est un v.a $\varphi(X_1, \dots, X_n)$ où φ est une application de $\mathbb{R}^n \rightarrow \mathbb{R}^P$, $P \geq 1$.*
- *Un estimateur $\hat{\theta}_n$ de θ est une statistique de l'échantillon i.e :*

$$\hat{\theta}_n = h(X_1, \dots, X_n),$$

telle que pour chaque réalisation $(x_1, \dots, x_n)$ de l'échantillon la valeur

$$\hat{\theta}_n = h(X_1, \dots, X_n).$$

S'appelle une estimation de θ .

- $\hat{\theta}_n$ est une variable aléatoire de loi de probabilité qui dépend du paramètre inconnu θ .
- $\hat{\theta}_n$ est un estimateur sans biais (ou non biaisé) du paramètre θ si :

$$E\{\hat{\theta}_n\} = \theta.$$

- $\hat{\theta}_n$ est un estimateur asymptotiquement sans biais (ou non biaisé) du paramètre θ si :

$$\lim_{n \rightarrow +\infty} E\{\hat{\theta}_n\} = \theta.$$

1.4.2 La vraisemblance

On appelle vraisemblance de l'échantillon $(X_1, \dots, X_n)$ la loi de probabilité de ce n-uple, note $L(x_1, \dots, x_n, \theta)$ et définie par :

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n P_\theta(X_i = x_i),$$

si X un v.a discrète, et donne par :

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f_{X_i, \theta}(x_i),$$

si X un v.a continu.

Fonction de log-vraisemblance

Définition 1.4.2 On appelle fonction de log-vraisemblance pour $(x_1, x_2, \dots, x_n)$ toute fonction de θ définie par :

$$L_n(x_1, \dots, x_n; \theta) = \log(L(x_1, \dots, x_n; \theta)).$$

Elle n'a de sens que si θ vérifie $L_n(x_1, \dots, x_n; \theta) > 0$.

La fonction logarithme népérien étant croissante, l'estimateur maximum de

vraisemblance $\hat{\theta}$ de θ pour $(x_1, \dots, x_n)$ vérifie :

$$\hat{\theta}_n = \arg \max_{\theta} (L_n(\epsilon_1, \dots, \epsilon_n, \theta)).$$

1.4.3 L'estimateur de quasi maximum de vraisemblance

En général, la fonction de vraisemblance pour un processus X n'est pas calculable puisque le passé du phénomènes $(X_0, X_{-1}, \dots)$ sont généralement inconnus. C'est pourquoi on utilise la quasi-vraisemblance qu'on obtient en remplaçant le passé du processus par des zéros. L'estimateur de quasi vraisemblance maximale (EQMV) est défini par :

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} \log(QL_{\theta}(X_1, \dots, X_n)).$$

1.5 Inégalité de Cauchy-Schwarz

Théorème 1.5.1 Soient $(U_1, \dots, U_n)$ et $(V_1, \dots, V_n)$ des réels(ou des complexes). Alors :

$$\sum_{k=1}^n |U_k V_k| \leq \left(\sum_{k=1}^n |U_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^n |V_k|^2 \right)^{\frac{1}{2}}.$$

1.6 Convergence presque sûrement des variables aléatoires

Définition 1.6.1 Nous disons que la suite X_n converge presque sûrement vers X si :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(\sup_{k \geq n} |X_k - X| > \varepsilon) = 0.$$

Ceci est noté $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$.

Chapitre 2

Processus puissance asymétrique GARCH modèle GJR-GARCH

Dans ce chapitre, nous allons traiter l'estimation des paramètres d'un modèle GJR-GARCH avec erreur normale $(1, 1)$, la méthode de l'estimation la plus utilisée est la méthode du quasi maximum de vraisemblance. Au début nous commençons par l'étude de l'existence et la stationnarité pour ce modèle. Par suite on va étudier en détail la méthode de quasi maximum de vraisemblance de loi normale obtenu sous l'hypothèse que $(\eta_t)_{t \in \mathbb{Z}}$ suit une loi normale $(1, 1)$. Enfin, nous prouvons les propriétés stationnaires de QMV c'est à dire la consistance forte et la normalité asymptotique de cet estimateur paramétrique.

2.1 Modèle GJR-GARCH(p,q)

Définition 2.1.1 *Un processus X_t satisfait une représentation GJR-GARCH(p,q) si et seulement si :*

$$\begin{cases} X_t = \eta_t \sqrt{h_t}, \\ h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i X_{t-i}^2 + \gamma_i \mathbb{I}_{X_{t-i} < 0} X_{t-i}^2) + \sum_{i=1}^p \beta_i h_{t-i}. \end{cases} \quad (2.1.1)$$

Où $\alpha_0 > 0$, $\alpha_i \geq 0$, $-1 < \gamma_i < 1$, pour $i = 1, \dots, q$, $\beta_i \geq 0$ pour $i = 1, \dots, p$, avec $\alpha_q > 0$, $\beta_p > 0$ et $\sum_{i=1}^p \beta_i < 1$.

• On note le vecteur des paramètres par : $\theta = (\alpha_0, \alpha_1, \dots, \alpha_q, \gamma_1, \dots, \gamma_q, \beta_1, \dots, \beta_p)$.

• Le résidu normalisé η_t est un bruit faible et $\mathbb{I}_{X_{t-i} < 0}$ désigne la fonction indicatrice telle que :

$$\begin{cases} \mathbb{I}_{X_{t-i} < 0} = 1 & \text{si } X_{t-i} < 0 \\ \mathbb{I}_{X_{t-i} < 0} = 0 & \text{sinon.} \end{cases}$$

Nous ne considérerons que le cas simple d'un processus GJR-GARCH(1,1) tel que :

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \gamma_1 \mathbb{I}_{X_{t-1} < 0} X_{t-1}^2 + \beta_1 h_{t-1},$$

ce que l'on peut encore écrire comme :

$$h_t = \alpha_0 + \alpha_{pos} \mathbb{I}_{X_{t-1} \geq 0} X_{t-1}^2 + \alpha_{neg} \mathbb{I}_{X_{t-1} < 0} X_{t-1}^2 + \beta_1 h_{t-1}.$$

permettant de lire directement les coefficients spécifiques afférents aux résidus positifs $\alpha_{pos} = \alpha_1$ ou négatifs $\alpha_{neg} = \alpha_1 + \gamma_1$.

Proposition 2.1.1 *Soit X_t le processus GJR-GARCH définie par (2.1.1), si $(1 - \sum_{i=1}^p \beta_i L^i)^{-1}$ existe, alors pour tout $t \in \mathbb{Z}$ on peut écrire la variance conditionnelle h_t sous la forme suivante :*

$$h_t = b_0 + \sum_{i \geq 1} b_i^+ L^i (\max(X_t, 0))^2 + \sum_{i \geq 1} b_i^- L^i (\max(-X_t, 0))^2, \quad (2.1.2)$$

où $b_0 = (1 - \sum_{i=1}^p \beta_i)^{-1} \alpha_0$ et les coefficients $(b_i^+, b_i^-)_{i \geq 1}$ sont définis par les relations récurrent suivant :

$$\begin{cases} b_i^+ = \sum_{k=1}^p \beta_k b_{i-k}^+ + \alpha_i & \text{avec } \alpha_i = 0 \quad \text{pour } i > q; \\ b_i^- = \sum_{k=1}^p \beta_k b_{i-k}^- + (\alpha_i + \gamma_i) & \text{avec } (\alpha_i + \gamma_i) = 0 \quad \text{pour } i > q, \end{cases}$$

avec $b_i^+ = b_i^- = 0$ pour $i < 0$.

Preuve 1 (preuve de la Proposition 2.1.1)

On a pour tout $t \in \mathbb{Z}$ la variance conditionnelle est définie par :

$$h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i X_{t-i}^2 + \gamma_i \mathbb{I}_{X_{t-i} < 0} X_t^2) + \sum_{i=1}^p \beta_i h_{t-i}. \quad (2.1.3)$$

En remplaçant $X_t = \max(X_t, 0) + \min(X_t, 0)$ dans (2.1.3) on obtient :

$$h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i (\max(X_{t-i}, 0))^2) + \sum_{i=1}^q (\alpha_i + \gamma_i) (\min(X_t, 0))^2 + \sum_{i=1}^p \beta_i h_{t-i}. \quad (2.1.4)$$

par la multiplication de $(1 - \sum_{i=1}^p \beta_i L^i)$ et (2.1.4) on obtient :

$$(1 - \sum_{i=1}^p \beta_i L^i) h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i (\max(X_{t-i}, 0))^2) + \sum_{i=1}^q (\alpha_i + \gamma_i) (\min(X_t, 0))^2, \quad (2.1.5)$$

donc on peut écrire (2.1.5) de la manière suivant :

$$B(L) h_t = \alpha_0 + \Theta^+(L) (\max(X_t, 0))^2 + \Theta^-(L) (\max(-X_t, 0))^2, \quad (2.1.6)$$

avec

$$\begin{aligned} B(L) &= (1 - \sum_{i=1}^p \beta_i L^i), \\ \Theta^+(L) &= \sum_{i=1}^q \alpha_i L^i = \sum_{i=1}^q \theta_i^+ L^i, \\ \text{et } \Theta^-(L) &= \sum_{i=1}^q (\alpha_i + \gamma_i) L^i = \sum_{i=1}^q \theta_i^- L^i. \end{aligned}$$

Comme l'inverse de l'opérateur $B(L)$ existe alors on a :

$$h_t = b_0 + \sum_{i \geq 1} b_i^+ L^i (\max(X_t, 0))^2 + \sum_{i \geq 1} b_i^- L^i (\max(-X_t, 0))^2. \quad (2.1.7)$$

De (2.1.6) et (2.1.7) on peut déterminer les coefficients b_0 et $(b_i^+, b_i^-)_{i \geq 1}$:

$$\begin{cases} B(L) \sum_{i \geq 0} b_i^+ L^i = \alpha_0 + \Theta^+(L); \\ B(L) \sum_{i \geq 0} b_i^- L^i = \alpha_0 + \Theta^-(L), \end{cases} \quad (2.1.8)$$

implique que

$$b_0 = b_0^+ = b_0^- = \frac{\alpha_0}{B(L)} = (1 - \sum_{i=1}^p \beta_i)^{-1} \alpha_0.$$

Le système (2.1.8) est équivalent au système

$$\begin{cases} (1 - \sum_{i=1}^p \beta_i L^i) \sum_{i \geq 0} b_i^+ L^i = \sum_{i=0}^q \theta_i^+ L^i \\ (1 - \sum_{i=1}^p \beta_i L^i) \sum_{i \geq 0} b_i^- L^i = \sum_{i=0}^q \theta_i^- L^i \end{cases}, \text{ avec } \theta_0 = \alpha_0,$$

d'où

$$\begin{cases} \sum_{i \geq 0} b_i^+ L^i - \sum_{i \geq 0} \sum_{j=1}^p \beta_j b_i^+ L^{i+j} = \sum_{i=0}^p \theta_i^+ L^i \\ \sum_{i \geq 0} b_i^- L^i - \sum_{i \geq 0} \sum_{j=1}^p \beta_j b_i^- L^{i+j} = \sum_{i=0}^p \theta_i^- L^i \end{cases}, \text{ avec } \theta_0 = \alpha_0.$$

De la première équation on a :

$$\begin{aligned} b_1^+ &= \beta_1 b_0^+ + \alpha_1, \\ b_2^+ &= \beta_1 b_1^+ + \beta_2 b_0^+ + \alpha_2, \\ b_3^+ &= \beta_1 b_2^+ + \beta_2 b_1^+ + \beta_3 b_0^+ + \alpha_3, \\ &\dots \end{aligned}$$

De la deuxième équation on a :

$$\begin{aligned} b_1^- &= \beta_1 b_0^- + (\alpha_1 + \gamma_1), \\ b_2^- &= \beta_1 b_1^- + \beta_2 b_0^- + (\alpha_2 + \gamma_2), \\ b_3^- &= \beta_1 b_2^- + \beta_2 b_1^- + \beta_3 b_0^- + (\alpha_3 + \gamma_3), \\ &\dots \end{aligned}$$

Alors les coefficients $(b_i^+, b_i^-)_{i \geq 1}$ sont définis par les relations de récurrence :

$$\begin{cases} b_i^+ = \sum_{k=1}^p \beta_k b_{i-k}^+ + \alpha_i & \text{avec } \alpha_i = 0 \quad \text{pour } i > q, \\ b_i^- = \sum_{k=1}^p \beta_k b_{i-k}^- + (\alpha_i + \gamma_i) & \text{avec } (\alpha_i + \gamma_i) = 0 \quad \text{pour } i > q. \end{cases}$$

Lemme 2.1.1 Soit $0 < \rho_0 < 1$, en définir $U = \{\theta \mid \beta_1 + \beta_2 + \beta_3 + \dots + \beta_p \leq \rho_0\}$. pour tout $\theta \in U$ on a :

$$b_i^+ \leq C_1 \rho_0^{\frac{i}{p}} \quad 0 \leq i < \infty, \quad (2.1.9)$$

$$b_i^- \leq C_2 \rho_0^{\frac{i}{p}} \quad 0 \leq i < \infty, \quad (2.1.10)$$

avec $C_1, C_2 > 0$ constant plus grand.

Preuve 2 (preuve du Lemme 2.1.1)

Nous donnons la même démarche de Berkes 2003

On va montrer ce lemme par récurrence.

• Pour $i=0$ on a :

$$b_0^+ = b_0^- = \frac{\alpha_0}{1 - \sum_{i=1}^p \beta_i} \leq \frac{\alpha_0}{1 - \rho_0} \rho_0^{\frac{q}{p}}. \quad (2.1.11)$$

relation vérifié.

• Pour $i \geq 1$:

Supposons que le lemme est vérifié pour tout $i > R = \max(p, q)$ et on va montrer qu'il est resté vrai pour i :

De (2.1.9) on trouve,

$$\begin{aligned} b_i^+ &\leq (\beta_0 + \beta_1 + \dots + \beta_p) \max_{1 \leq k \leq p} b_{i-k}, \\ &\leq \rho_0 C_1 \rho_0^{\frac{(i-p)}{p}}, \\ &\leq C_1 \rho_0^{\frac{i}{p}}. \end{aligned}$$

Et de (2.1.10) :

$$\begin{aligned} b_i^- &\leq (\beta_0 + \beta_1 + \dots + \beta_p) \max_{1 \leq k \leq p} b_{i-k}, \\ &\leq \rho_0 C_2 \rho_0^{\frac{(i-p)}{p}}, \\ &\leq C_2 \rho_0^{\frac{i}{p}}. \end{aligned}$$

Ce qui achève la preuve.

2.2 Définition et hypothèse

2.2.1 Définition d'estimateur

Soit $(\eta_1, \eta_2, \dots, \eta_n)$ une trajectoire observée de η où $\theta \subset \mathbb{R}^{1+2q+p}$ est l'ensemble des paramètres inconnus du modèle, pour estimer θ nous considérons la log-vraisemblance de $(X_1, X_2, \dots, X_n)$. Si f est la densité de probabilité de η définit par :

$$f_\eta(\eta_t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\eta_t-1)^2}. \quad (2.2.1)$$

- Le modèle GJR-GARCH(p,q) s'écrit sous la forme :

$$\begin{aligned} x_t &= \eta_t h_t^{1/2}, \\ h(t) &= \alpha_0 + \sum_{i=1}^q (\alpha_i X_t^2 + \gamma_i \mathbb{I}_{X_{t-i} < 0} X_t^2) + \sum_{i=1}^p \beta_i h_{t-i}. \end{aligned}$$

Et comme on a :

$$f_{(\eta_1, \eta_2, \dots, \eta_n)}(\eta_1, \eta_2, \dots, \eta_n) = \prod_{i=1}^n f_\eta(\eta_i),$$

de (2.2.1) on obtient :

$$\begin{aligned} f_{(\eta_1, \eta_2, \dots, \eta_n)}(\eta_1, \eta_2, \dots, \eta_n) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\eta_i-1)^2}, \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \sum_{i=1}^n (\eta_i-1)^2}. \end{aligned}$$

Nous utilisons l'intégrale pour obtenir la fonction $f(X_1, \dots, X_n)$

$$\begin{aligned} \oint f_{(\eta_1, \eta_2, \dots, \eta_n)}(\eta_1, \eta_2, \dots, \eta_n) d\eta_1 \dots d\eta_n &= \oint \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \sum_{i=1}^n (\eta_i-1)^2} d\eta_1 \dots d\eta_n, \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \oint e^{-\frac{1}{2} \sum_{i=1}^n (\eta_i-1)^2} d\eta_1 \dots d\eta_n. \end{aligned} \quad (2.2.2)$$

On fait un changement de variable $(\eta_1, \dots, \eta_n) \longrightarrow (X_1, \dots, X_n)$

on a $X_t = \eta_t h_t^{1/2}$ tel que $\eta_t \sim \mathcal{N}(1, 1)$,

$$\Rightarrow \eta_t = X_t h_t^{-1/2},$$

alors $d\eta_t = h_t^{-1/2} dX_t, \forall t = 1 \dots n$.

Donc (2.2.2) devient :

$$\begin{aligned} \oint f_{(\eta_1, \eta_2, \dots, \eta_n)}(\eta_1, \eta_2, \dots, \eta_n) d\eta_1 \dots d\eta_n &= \frac{1}{(2\pi)^{\frac{n}{2}}} \oint e^{-\frac{1}{2} \sum_{t=1}^n (X_t (h_\theta^t)^{-\frac{1}{2}} - 1)^2} (h_\theta)^{-\frac{1}{2}} dX_1 \dots (h_\theta)^{-\frac{n}{2}} dX_n, \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \oint e^{-\frac{1}{2} \sum_{t=1}^n (X_t (h_\theta^t)^{-\frac{1}{2}} - 1)^2} \left(\prod_{t=1}^n (h_\theta^t)^{-\frac{1}{2}} \right) dX_1 \dots dX_n, \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \oint e^{-\frac{1}{2} \sum_{t=1}^n (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2} \left(\prod_{t=1}^n (h_\theta^t)^{-\frac{1}{2}} \right) dX_1 \dots dX_n. \end{aligned}$$

D'où

$$\begin{aligned} f(X_1, \dots, X_n) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \left(\prod_{t=1}^n (h_\theta^t)^{-\frac{1}{2}} \right) e^{-\frac{1}{2} \sum_{t=1}^n (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2}, \\ &= \left(\prod_{t=1}^n \frac{1}{2\pi} (h_\theta^t)^{-\frac{1}{2}} \right) e^{-\frac{1}{2} \sum_{t=1}^n (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2}. \end{aligned}$$

• La fonction vraisemblance :

$$l(\theta) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \sum_{t=1}^n (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2} \left(\prod_{t=1}^n (h_\theta^t)^{-\frac{1}{2}} \right). \quad (2.2.3)$$

- La fonction log-vraisemblance :

$$\begin{aligned}
L(\theta) &= \log l(\theta), \\
&= \log \left(\frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \sum_{t=1}^n (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2} \left(\prod_{t=1}^n (h_\theta^t)^{-\frac{1}{2}} \right) \right), \\
&= -n \log \sqrt{2\pi} - \frac{1}{2} \sum_{t=1}^n \log(h_\theta^t) - \frac{1}{2} \sum_{t=1}^n (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2, \\
&= -n \log \sqrt{2\pi} - \frac{1}{2} \sum_{t=1}^n \left(\log(h_\theta^t) + (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2 \right).
\end{aligned}$$

- Nous définissons respectivement la vraisemblance de loi Normale et le quasi-vraisemblance de loi Normale par :

$$\begin{aligned}
L_n(\theta) &= - \sum_{t=1}^n q_t(\theta), & q_t(\theta) &= \log(h_\theta^t) + (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2, \\
\hat{L}_n(\theta) &= - \sum_{t=1}^n \hat{q}_t(\theta), & \hat{q}_t(\theta) &= \log(\hat{h}_\theta^t) + (\hat{h}_\theta^t)^{-1} (X_t - (\hat{h}_\theta^t)^{\frac{1}{2}})^2.
\end{aligned}$$

Comme $(X_0, X_{-1}, \dots)$ sont inconnus, généralement h_θ^t ne sont pas calculables pour cela, au lieu d'utiliser la log-vraisemblance on utilise la quasi log-vraisemblance .

Un EQMV de loi Normale $\hat{\theta}_n$ est le maximisateur de $\hat{L}_n$:

$$\begin{aligned}
\hat{\theta}_n &= \arg \max_{\theta \in \Theta} \hat{L}_n, \\
&= \arg \min_{\theta \in \Theta} \sum_{i=1}^n \hat{q}_i(\theta).
\end{aligned}$$

2.2.2 Existence et stationnarité

Proposition 2.2.1 *Le GJR-GARCH définie dans (2.1.1) est un processus causal faiblement stationnaire satisfaisant $E\{X^2\} < \infty$. Si et seulement si*

$$\sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i < 1. \quad (2.2.4)$$

Pour assurer la stationnarité de (2.1.1), nous définissons l'ensemble

$$\Theta = \{\theta \in \mathbb{R}^{1+2q+p} \setminus \sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i < 1\}.$$

Dans la proposition suivante, nous établissons une représentation pour h_t^2 en termes de $\{X_{t-i}^2\}_{i \geq 1}$.

Proposition 2.2.2 *Si $\theta_0 \in \Theta$, alors il existe une représentation causale unique de h_t^2 donnée par*

$$h_t^2 = c_0 + \sum_{i \geq 1} c_i X_{t-i}^2, \quad (2.2.5)$$

où, il existe $0 < \rho, \rho_ < 1$ tel que,*

$$c_i \leq C_1 \rho^i, \quad (2.2.6)$$

$$\frac{\partial c_i(\theta)}{\partial \theta_k} \leq C_2 i \rho_*^i, \quad \forall 1 \leq j \leq 1 + 2q + p. \quad (2.2.7)$$

Preuve 3 (preuve de la Proposition 2.2.1)

L'idée de la preuve suit celle du théorème 1 de Bollerslev (1986)([4]).

Soit X_t est le processus GJR-GARCH(p, q) définit par :

$$X_t = \eta_t \sqrt{h(t)} \quad \eta \sim \mathcal{N}(1, 1), \quad (2.2.8)$$

$$h(t) = \alpha_0 + \sum_{i=1}^q (\alpha_i X_t^2 + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} X_t^2) + \sum_{i=1}^p \beta_i h_{t-i}. \quad (2.2.9)$$

Substituent (2.2.8) dans (2.2.9) on obtient :

$$\begin{aligned}
h(t) &= \alpha_0 + \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 h_{t-i} + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2 h_{t-i}) + \sum_{i=1}^p \beta_i h_{t-i}, \\
&= \alpha_0 + \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 (\alpha_0 + \sum_{j=1}^q (\alpha_j \eta_{t-i-j}^2 h_{t-i-j} + \gamma_j \mathbb{I}_{\eta_{t-i-j} < 0} \eta_{t-i-j}^2 h_{t-i-j}) + \sum_{j=1}^p \beta_j h_{t-i-j})) \\
&\quad + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2 (\alpha_0 + \sum_{j=1}^q (\alpha_j \eta_{t-i-j}^2 h_{t-i-j} + \gamma_j \mathbb{I}_{\eta_{t-i-j} < 0} \eta_{t-i-j}^2 h_{t-i-j}) + \sum_{j=1}^p \beta_j h_{t-i-j})) \\
&\quad + \sum_{i=1}^p \beta_i (\alpha_0 + \sum_{j=1}^q (\alpha_j \eta_{t-i-j}^2 h_{t-i-j} + \gamma_j \mathbb{I}_{\eta_{t-i-j} < 0} \eta_{t-i-j}^2 h_{t-i-j}) + \sum_{j=1}^p \beta_j h_{t-i-j}), \\
&\quad \cdot \\
&\quad \cdot \\
&\quad \cdot \\
&= \alpha_0 \sum_{k=1}^{+\infty} M(t, k).
\end{aligned} \tag{2.2.10}$$

Alors

$$\begin{aligned}
M(t, 0) &= 1, \\
M(t, 1) &= \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2) + \sum_{i=1}^p \beta_i, \\
M(t, 2) &= \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2) M(t, 1) + \sum_{i=1}^p \beta_i M(t, 1), \\
M(t, 3) &= \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 M(t, 2) + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2 M(t, 2)) + \sum_{i=1}^p \beta_i M(t, 2).
\end{aligned}$$

Et en générale

$$M(t, k+1) = \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 M(t-i, k) + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2 M(t-i, k)) + \sum_{i=1}^p \beta_i M(t-i, k). \tag{2.2.11}$$

On a η_t^2 est i.i.d, le moment de $M(t, k)$ est indépendant de t alors :

$$E\{M(t, k)\} = E\{M(s, k)\} \quad \text{pour tout } k, t, s. \quad (2.2.12)$$

De (2.2.8) on déduire que

$$\begin{aligned} E\{\eta_t^2\} &= \text{Var}(\eta_t) + E\{\eta_t\}^2, \\ &= 2. \end{aligned} \quad (2.2.13)$$

$$\begin{aligned} E\{\mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2\} &= \int_{-\infty}^0 x^2 f(x), \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} x^2 f(x), \\ &= \frac{1}{2} E\{\eta_t^2\}, \\ &= 1. \end{aligned} \quad (2.2.14)$$

De (2.2.11), (2.2.12), (2.2.13) et (2.2.14) on obtient :

$$\begin{aligned} E\{M(t, k+1)\} &= E\left\{ \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 M(t-i, k) + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2 M(t-i, k) + \sum_{i=1}^p \beta_i M(t-i, k) \right\}, \\ &= E\left\{ \sum_{i=1}^q (\alpha_i \eta_{t-i}^2 + \gamma_i \mathbb{I}_{\eta_{t-i} < 0} \eta_{t-i}^2) + \sum_{i=1}^p \beta_i \right\} E\{M(t, k)\}, \\ &= \left(\sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i \right) E\{M(t, k)\}, \\ &\quad \cdot \\ &\quad \cdot \\ &\quad \cdot \\ &= \left(\sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i \right)^{k+1} E\{M(t, 0)\}, \\ &= \left(\sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i \right)^{k+1}. \end{aligned} \quad (2.2.15)$$

En fin de (2.2.8), (2.2.10) et (2.2.15) on obtient :

$$\begin{aligned}
E\{X_t^2\} &= E\{\eta_t^2 h_t\}, \\
&= E\left\{\alpha_0 \sum_{k=1}^{\infty} M(t, k) \eta_t^2\right\}, \\
&= \alpha_0 E\{\eta_t^2\} E\left\{\sum_{k=1}^{\infty} M(t, k)\right\}, \\
&= 2\alpha_0 \sum_{k=1}^{\infty} E\{M(t, k)\}, \\
&= 2\alpha_0 \left(1 - \sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i\right)^{-1}.
\end{aligned}$$

Donc $E\{X_t^2\} < \infty$ si et seulement si :

$$\sum_{i=1}^q (2\alpha_i + \gamma_i) + \sum_{i=1}^p \beta_i < 1.$$

2.2.3 Hypothèses nécessaires pour la convergence d'EQMV-Normale

L'EQMV-Normale pourrait converger et être asymptotiquement Gaussien mais cela nécessiter quelques hypothèses supplémentaires sur Θ et la fonction h_θ :

- **Hypothèses H1**(*compacité*) : Θ est un ensemble compacte .
- **Hypothèses H2**(*Borne inférieure de la variance conditionnelle*) : il existe une constante $\alpha_0 > 0$ tel que $\forall \theta \in \Theta$ alors $h_\theta(x) > \alpha_0$ pour tout $x \in \mathbb{R}^n$.
- **Hypothèses H3** : $E\{(\eta_t - 1)^2\} = 1$.
- **Hypothèses H4** (*Identifiabilité*) : La fonction h_θ est telle que pour tout $\theta_1, \theta_2 \in \Theta$, alors $h_{\theta_1} = h_{\theta_2}$ implique que $\theta_1 = \theta_2$.

2.3 comportement asymptotique d'EQMV

2.3.1 Propriétaires asymptotiques de la quasi vraisemblance

Lemme 2.3.1 *Supposons que $\theta_0 \in \Theta$ et que X_t est la solution causal et stationnaire de $(X_t)_{t \in \mathbb{Z}}$, alors pour tout $t \in \mathbb{N}^\infty$:*

$$E | \hat{h}_\theta^t - h_\theta^t |_\Theta^r \leq CE \{ | X_0 |^r \} (\sum_{i \geq t} \alpha_i^{(0)}(h_\theta^t, \Theta))^r.$$

Avec

$$\alpha_i^{(0)}(h_\theta^t, \Theta) = \sup_{\theta \in \Theta} \max(|b_i^+(\theta)|, |b_j^-(\theta)|).$$

2.3.2 Consistance forte

Théorème 2.3.1 *Supposons que les hypothèses sont satisfaites, alors la suite d'EQMV-Normale $(\hat{\theta}_n)$ converge fortement, c'est-à-dire*

$$\hat{\theta}_n \longrightarrow \theta_0 \quad p.s.$$

Preuve 4 (preuve du Théorème 2.3.1)

Nous prouvons ce théorème en deux étapes dans la première étape une forte loi uniforme des grandes nombres sur θ satisfaite par $\frac{1}{n} \hat{L}_n(\theta)$ qui converge vers $L(\theta) = -E\{q_t(\theta)\}$ est établi. Par suite, dans la deuxième étape nous prouvons que $L(\theta)$ admet un maximum unique en θ_0 .

On prouve **l'étape(1)** par le même façon de la preuve du théorème 1 de Bardet, Boularouk et Djaballah (2017)([6]), la loi uniforme forte des grands nombre satisfaite pour la moyenne d'échantillon $(\hat{q}_t)_{t \in \mathbb{N}^*}$ est obtenu en prouvant que $E\{|q_t(\theta)|\} < \infty$. Pour tout $t \in \mathbb{Z}$ et sous l'hypothèse H2, on a :

$$\begin{aligned}
|q_t(\theta)| &= |\log(h_\theta^t) + (h_\theta^t)^{-1}(X_t - (h_\theta^t)^{\frac{1}{2}})^2| \\
&\leq |\log(h_\theta^t)| + (X_t(h_\theta^t)^{-\frac{1}{2}} - 1)^2 \\
&\leq |h_\theta^t| + (X_t(h_\theta^t)^{-\frac{1}{2}} - 1)^2 \\
&\leq |h_\theta^t| + \left(\frac{X_t}{(\alpha_0)^{\frac{1}{2}}} - 1\right)^2.
\end{aligned}$$

Donc :

$$\sup |q_t(\theta)| \leq |h_\theta^t| + \left(\frac{X_t}{(\alpha_0)^{\frac{1}{2}}} - 1\right)^2, \quad (2.3.1)$$

où, $EX_t^2 < \infty$ d'après la proposition 2.2.1 et $E\{|h_\theta^t|\} < \infty$ d'après le Lemme 1 de Bardet et Wintenberger (2009) ([1]) ce qui implique que $E\{|q_t(\theta)|\} < \infty$, d'où on conclut que :

$$\left|\frac{1}{n}L_n(\theta) - L(\theta)\right| \xrightarrow[n \rightarrow +\infty]{p.s} 0. \quad (2.3.2)$$

Maintenant nous allons montrer que :

$$\frac{1}{n}|\widehat{L}_n(\theta) - L_n(\theta)| \xrightarrow[n \rightarrow +\infty]{p.s} 0. \quad (2.3.3)$$

En effet, pour tout $\theta \in \Theta$ et $t \in \mathbb{N}^$*

$$\begin{aligned}
\hat{q}_t(\theta) - q_t(\theta) &= \log(\hat{h}_\theta^t) + (\hat{h}_\theta^t)^{-1}(X_t - (\hat{h}_\theta^t)^{\frac{1}{2}})^2 - \log(h_\theta^t) - (h_\theta^t)^{-1}(X_t - (h_\theta^t)^{\frac{1}{2}})^2 \\
&= \log\left(\frac{\hat{h}_\theta^t}{h_\theta^t}\right) + \left(X_t(\hat{h}_\theta^t)^{-\frac{1}{2}} - 1\right)^2 - \left(X_t(h_\theta^t)^{-\frac{1}{2}} - 1\right)^2 \\
&= \log\left(\frac{\hat{h}_\theta^t}{h_\theta^t}\right) + \left(X_t(\hat{h}_\theta^t)^{-\frac{1}{2}} - X_t(h_\theta^t)^{-\frac{1}{2}}\right)\left(X_t(\hat{h}_\theta^t)^{-\frac{1}{2}} + X_t(h_\theta^t)^{-\frac{1}{2}} - 2\right) \\
&\leq \log\left(\frac{\hat{h}_\theta^t}{h_\theta^t}\right) + \left(X_t(\hat{h}_\theta^t)^{-\frac{1}{2}} - X_t(h_\theta^t)^{-\frac{1}{2}}\right)\left(X_t(\hat{h}_\theta^t)^{-\frac{1}{2}} + X_t(h_\theta^t)^{-\frac{1}{2}}\right) \\
&\leq \frac{\hat{h}_\theta^t}{h_\theta^t} - 1 + \left((\hat{h}_\theta^t)^{-1} - (h_\theta^t)^{-1}\right)X_t^2 \\
&= \frac{\hat{h}_\theta^t - h_\theta^t}{h_\theta^t} + \left((\hat{h}_\theta^t)^{-1} - (h_\theta^t)^{-1}\right)X_t^2
\end{aligned}$$

Alors :

$$\begin{aligned}
|\hat{q}_t(\theta) - q_t(\theta)| &\leq |\hat{h}_\theta^t - h_\theta^t|\alpha_0^{-1} + |\hat{h}_\theta^t - h_\theta^t|\alpha_0^{-2}X_t^2 \\
&\leq C|\hat{h}_\theta^t - h_\theta^t|(1 + X_t^2),
\end{aligned}$$

avec $C > 0$.

D'après le corollaire 1 de Kounias et Weng(1969)([16]), il suffit de montrer que il existe un $s \in [0, 1]$ telle que :

$$\sum_{t \geq 1} \frac{1}{t^s} E\{|\hat{q}_t(\theta) - q_t(\theta)|^s\} < \infty$$

De l'inégalité de Cauchy-Schwarz avec $s = 1$

$$E\{|\hat{q}_t(\theta) - q_t(\theta)|\} \leq CE\{(1 + X_t^2)^2\}^{\frac{1}{2}} \times E\{(|\hat{h}_\theta^t - h_\theta^t|)^2\}^{\frac{1}{2}} \quad (2.3.4)$$

D'après la proposition(2.2.1),le lemme(2.3.1), et la relation (2.3.4) on a $E\{|X_t|^2\} < \infty$ et $E\{|\hat{\sigma}_t|^2\} < \infty$ ce qui implique :

$$\begin{aligned} E\{|\hat{q}_t(\theta) - q_t(\theta)|\} &\leq CE\{|X_0|^2\}^{\frac{1}{2}} (\sum_{j \geq t} \alpha_j^{(0)}(\sigma_\theta^t, \Theta)) \\ &\leq C(\sum_{j \geq t} \alpha_j^{(0)}(\sigma_\theta^t, \Theta)) \\ &\leq C \sum_{j \geq t} \rho^j \\ &\leq C' \rho^t \end{aligned}$$

tel que $C' = C \times \max(C_1, C_2)$ et $\rho = \rho_0^{\frac{1}{p}}$.

Nous avons donc :

$$\sum_{t \geq 1} \frac{1}{t} E\{|\hat{q}_t(\theta) - q_t(\theta)|\} \leq A \sum_{t \geq 1} \frac{\rho^t}{t} \quad (\text{avec } A = \frac{C'}{1 - \rho})$$

Cette série est fini car $0 < \rho < 1$, par conséquent, nous obtenons :

$$\frac{1}{n} \sum_{t=1}^n |\hat{q}_t(\theta) - q_t(\theta)| \xrightarrow[n \rightarrow +\infty]{p.s} 0 \quad \Leftrightarrow \quad \frac{1}{n} |\hat{L}_n(\theta) - L_n(\theta)| \xrightarrow[n \rightarrow +\infty]{p.s} 0$$

Comme on a :

$$\begin{aligned} \left| \frac{1}{n} L_n(\theta) - L(\theta) \right| &= \left| \frac{1}{n} \widehat{L}_n(\theta) - \frac{1}{n} L_n(\theta) + \frac{1}{n} L_n(\theta) - L(\theta) \right| \\ &\leq \frac{1}{n} |\widehat{L}_n(\theta) - L_n(\theta)| + \left| \frac{1}{n} L_n(\theta) - L(\theta) \right| \end{aligned}$$

Donc :

$$\frac{1}{n} \widehat{L}_n(\theta) \xrightarrow[n \rightarrow +\infty]{p.s} L(\theta) \quad (2.3.5)$$

Pour prouver l'**étape(2)** on suppose que $L(\theta)$ admet deux maximum θ et θ_0 telle que $\theta \neq \theta_0$, pour $\theta \in \Theta$ on a :

$$\begin{aligned} L(\theta) &= -E\{q_t(\theta)\} \\ &= -E\{\log(h_\theta^t) + (X_t(h_\theta^t)^{-\frac{1}{2}} - 1)^2\} \\ &= -E\{\log(h_\theta^t) + (h_\theta^t)^{-1} (X_t - (h_\theta^t)^{\frac{1}{2}})^2\} \\ &= -E\{\log(h_\theta^t) + (h_\theta^t)^{-1} (\eta_\theta^t (h_{\theta_0}^t)^{\frac{1}{2}} - (h_\theta^t)^{\frac{1}{2}})^2\} \\ &= -E\left\{ \log(h_\theta^t) + \left(\frac{h_{\theta_0}^t}{h_\theta^t} \right) \left(\eta_t - \left(\frac{h_\theta^t}{h_{\theta_0}^t} \right)^{\frac{1}{2}} \right)^2 \right\} \end{aligned}$$

D'autre par on a :

$$\begin{aligned} L(\theta_0) &= -E\{q_t(\theta_0)\} \\ &= -E\{\log(h_{\theta_0}^t) + (X_t(h_{\theta_0}^t)^{-\frac{1}{2}} - 1)^2\} \\ &= -E\{\log(h_{\theta_0}^t) + (h_{\theta_0}^t)^{-1} (X_t - (h_{\theta_0}^t)^{\frac{1}{2}})^2\} \\ &= -E\{\log(h_{\theta_0}^t) + (\eta_t - 1)^2\} \\ &= -E\{\log(h_{\theta_0}^t) - 1 \end{aligned}$$

Par suite en utilisant l'Hypothèse H3, nous obtenons :

$$\begin{aligned}
L(\theta_0) - L(\theta) &= -E\left\{\log(h_\theta^t) + (\eta_t - 1)\right\} + E\left\{\log(h_\theta^t) + \left(\frac{h_{\theta_0}^t}{h_\theta^t}\right)\left(\eta_t - \left(\frac{h_\theta^t}{h_{\theta_0}^t}\right)^{\frac{1}{2}}\right)^2\right\} \\
&= E\left\{\log\left(\frac{h_\theta^t}{h_{\theta_0}^t}\right) - (\eta_t - 1)^2 + \left(\frac{h_{\theta_0}^t}{h_\theta^t}\right)\left(\eta_t - \left(\frac{h_\theta^t}{h_{\theta_0}^t}\right)\right)\right\} \\
&= E\left\{\log\left(\frac{h_\theta^t}{h_{\theta_0}^t}\right)\right\} - 1 + E\left\{\eta_t - \left(\frac{h_\theta^t}{h_{\theta_0}^t}\right)^{\frac{1}{2}} + 1 - 1\right\}^2 \\
&= E\left\{\log\left(\frac{h_\theta^t}{h_{\theta_0}^t}\right)\right\} - 1 + E\left\{(\eta_t - 1) + \left(1 - \left(\frac{h_\theta^t}{h_{\theta_0}^t}\right)^{\frac{1}{2}}\right)^2\right\}
\end{aligned}$$

mais pour un η_t suivant une distribution de probabilité symétrique, nous avons pour tout $m \in \mathbb{R}^*$, $E\{(\eta_t - 1 + m)^2\} > E\{(\eta_t - 1)^2\} = 1$, d'où :

$$\begin{aligned}
L(\theta_0) - L(\theta) &> E\left\{\log \frac{h_\theta^t}{h_{\theta_0}^t}\right\} + \frac{h_{\theta_0}^t}{h_\theta^t} - 1 \\
&> H\left(\frac{h_{\theta_0}^t}{h_\theta^t}\right)
\end{aligned}$$

avec $H(x) = -\log(x) - 1 + x$. Mais pour tout $x \in [0, 1[\cup]1, +\infty[$, $H(x) > 0$ et $H(1) = 0$. Par conséquent si $\theta \neq \theta_0$, on a $H\left(\frac{h_{\theta_0}^t}{h_\theta^t}\right) > 0$ ($= 0$ si $\theta = \theta_0$). Cela implique, de l'hypothèse **H4** (identifiabilité), que $L(\theta_0) - L(\theta) > 0$ presque sûrement pour tout $\theta \in \Theta$, $\theta \neq \theta_0$. D'où une borne supérieure de $L(\theta)$ est atteinte seulement pour $\theta = \theta_0$ qui est un maximum unique.

2.3.3 La normalité asymptotique

Maintenant, nous donnons le théorème prouvant la normalité asymptotique pour l'estimateur.

Lemme 2.3.2 *Soit $\theta_0 = (\alpha_0, \alpha_1, \dots, \alpha_q, \gamma_1, \dots, \gamma_q, \beta_1, \dots, \beta_p)$ avec X_t une solution stationnaire de (2.1.1), alors on a :*

$$\frac{1}{\sqrt{n}} \frac{\partial L_n(\theta_0)}{\partial \theta} \xrightarrow[n \rightarrow +\infty]{D} N_d(0, G(\theta_0)). \quad (2.3.6)$$

avec $G(\theta_0)$ la matrice définie dans par

$$G(\theta_0) = E \left\{ \frac{\partial q_t(\theta_0)}{\partial \theta_i} \frac{\partial q_t(\theta_0)}{\partial \theta_j} \right\} \text{ pour } 1 \leq i, j \leq p + q + 1. \quad (2.3.7)$$

Preuve 5 (preuve du Lemme 2.3.2)

Le preuve est divisée en deux parties, premièrement nous prouvons que $(\partial q_t(\theta)/\partial \theta_i; F_t) = h(X_{t-1}, \dots)$ est une martingale de différences et deuxièmement nous avons appliqué le théorème central limite pour les martingales de différences

Pour prouver que $(\partial q_t(\theta)/\partial \theta_i; F_t)$ est un martingale-différences, si ses espérances mathématiques par rapport au passé sont nulles :

$$\begin{aligned} E \left\{ \frac{\partial q_t}{\partial \theta_i} \right\} &= E \left\{ h_t^{-1} \frac{\partial h_t}{\partial \theta_i} - h_t^{-2} \frac{\partial h_t}{\partial \theta_i} (X_t - h_t^{\frac{1}{2}})^2 + h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (\eta_t - 1) \right\}, \\ &= E \left\{ h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (1 - (\eta_t - 1)^2 + (\eta_t - 1)) \right\}, \\ &= 0. \end{aligned}$$

$$\begin{aligned} E \left\{ \left| \frac{\partial q_t}{\partial \theta_i} \right| \right\} &= E \left\{ \left| h_t^{-1} \frac{\partial h_t}{\partial \theta_i} - h_t^{-2} \frac{\partial h_t}{\partial \theta_i} (X_t - h_t^{\frac{1}{2}})^2 + h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (\eta_t - 1) \right| \right\}, \\ &= E \left\{ \left| h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (\eta_t - 1)^2 + h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (1 + \eta_t - 1) \right| \right\}, \\ &= E \left\{ \left| h_t^{-1} \frac{\partial h_t}{\partial \theta_i} \right| |\eta_t - (\eta_t - 1)^2| \right\}, \\ &\leq E \left\{ \left| h_t^{-1} \frac{\partial h_t}{\partial \theta_i} \right| (\eta_t + (\eta_t - 1)^2) \right\}, \\ &\leq CE \left\{ \left| \frac{\partial h_t}{\partial \theta_i} \right| \right\}, \\ &< \infty. \end{aligned}$$

Donc $(\partial q_t(\theta)/\partial \theta_i; F_t)$ est un martingale-differences.
Maintenant, pour appliquer le théorème centrale limite d'un martingale-difference,
nous prouvons que $E\left\{\frac{\partial q_t(\theta)}{\partial \theta_i}\right\} < \infty$.

$$\begin{aligned}
E\left\{\frac{\partial q_t}{\partial \theta_i}\right\}^2 &= E\left\{\left(h_t^{-1}\frac{\partial h_t}{\partial \theta_i}(\eta_t - 1)^2 + h_t^{-1}\frac{\partial h_t}{\partial \theta_i}(1 + \eta_t - 1)\right)^2\right\}, \\
&= E\left\{\left(h_t^{-1}\frac{\partial h_t}{\partial \theta_i}\right)^2\left(1 - (\eta_t - 1)^2 + (\eta_t - 1)\right)^2\right\}, \\
&= E\left\{\left(h_t^{-1}\frac{\partial h_t}{\partial \theta_i}\right)^2\left(\eta_t - (\eta_t - 1)^2\right)^2\right\}, \\
&\leq C'E\left\{\frac{\partial h_t^2}{\partial \theta_i}\right\}, \\
&< \infty.
\end{aligned}$$

Par suite

$$E\left\{\frac{\partial q_t(\theta)}{\partial \theta}\right\}^2 = \sum_{t=n}^n E\left\{\frac{\partial q_t(\theta)}{\partial \theta}\right\}^2 < \infty.$$

$$\begin{aligned}
E\left\{\frac{\partial q_t}{\partial \theta_i} \frac{\partial q_t}{\partial \theta_j}\right\}^2 &= E\left\{\left(h_t^{-2}\frac{\partial h_t}{\partial \theta_i} \frac{\partial h_t}{\partial \theta_j}\right)(-(\eta_t - 1)^2 + 1 + \eta_t - 1)^2\right\}, \\
&= E\left\{\left(h_t^{-2}\frac{\partial h_t}{\partial \theta_i} \frac{\partial h_t}{\partial \theta_j}\right)\right\}E\{(-(\eta_t - 1)^2 + \eta_t)^2\}, \\
&= E\left\{\left(h_t^{-2}\frac{\partial h_t}{\partial \theta_i} \frac{\partial h_t}{\partial \theta_j}\right)\right\}.
\end{aligned}$$

La théorème centrale limite pour les différences de martingale implique
que :

$$n^{-\frac{1}{2}} \frac{\partial L_n(\theta_0)}{\partial \theta} \xrightarrow[n \rightarrow +\infty]{D} N_d(0, G(\theta_0)), \quad (2.3.8)$$

tell que $G(\theta_0)$ définie dans 2.3.7

Théorème 2.3.2 (la normalité asymptotique)

Soit X_t une solution stationnaire de (2.1.1), avec $\theta_0 \in \Theta$, et que les conditions du théorème 2.3.1 vérifiées, si la fonction de probabilité cumulée de ξ_0 est continûment différentiable alors

$$\sqrt{n}(\theta_n - \theta_0) \xrightarrow[n \rightarrow +\infty]{D} N_d(0, F^{-1}GF^{-1}) \quad (2.3.9)$$

où les matrices $G(\theta_0)$ définit dans 2.3.7 et $F(\theta_0)$ est définit par :

$$F(\theta_0) = E\left\{ \frac{\partial^2 q_t(\theta_0)}{\partial \theta_i \partial \theta_j} \right\} \text{ pour } 1 \leq i, j \leq p + q + 1. \quad (2.3.10)$$

Preuve 6 (preuve du Théorème 2.3.2)

La première dérivée de L_n est

$$\frac{\partial L_n}{\partial \theta_i} = - \sum_{t=1}^n \frac{\partial q_t}{\partial \theta_i},$$

avec

$$\frac{\partial q_t}{\partial \theta_i} = h_t^{-1} \frac{\partial h_t}{\partial \theta_i} - h_t^{-2} \frac{\partial h_t}{\partial \theta_i} (X_t - h_t^{\frac{1}{2}})^2 + h_t^{-1} h_t^{-\frac{1}{2}} \frac{\partial h_t}{\partial \theta_i} (X_t - (h_t)^{\frac{1}{2}}).$$

De puis $\theta_0 \in \Theta$ et $\widehat{\theta}_n \rightarrow \theta_0$, extension de Taylor de $\frac{1}{\sqrt{n}} \frac{\partial L(\widehat{\theta}_n)}{\partial \theta_i}$ implique

$$\frac{1}{\sqrt{n}} \frac{\partial L_n(\widehat{\theta}_n)}{\partial \theta_i} = \frac{1}{\sqrt{n}} \frac{\partial L_n(\theta_0)}{\partial \theta_i} + \frac{1}{n} \frac{\partial^2 L_n(\bar{\theta}_{n,i})}{\partial \theta \partial \theta_i} \sqrt{n}(\widehat{\theta}_n - \theta_0)$$

Avec n suffisamment grande telle que $(\bar{\theta}_{n,i}) \in \Theta$, qui est entre $\widehat{\theta}_n$ et θ_0
En utilise la consistence de $\widehat{\theta}_n$ et résultat de lemme(2.3.3), on déduire que

$$F_n = -\frac{1}{n} \left(\frac{\partial^2 L_n(\bar{\theta}_{n,i})}{\partial \theta \partial \theta_i} \right)_{1 \leq i \leq d} \rightarrow F(\theta_0).$$

tel que $F(\theta_0)$ définie dans (2.3.10)

et

$$\begin{aligned} \frac{\partial^2 q_t(\theta)}{\partial \theta_j \partial \theta_i} &= (h_\theta^t)^{-1} \frac{\partial^2 h_\theta^t}{\partial \theta_j \partial \theta_i} - (h_\theta^t)^{-2} \frac{\partial h_\theta^t}{\partial \theta_j} \frac{\partial h_\theta^t}{\partial \theta_i} + \left[\left(2(h_\theta^t)^{-3} \frac{\partial h_\theta^t}{\partial \theta_j} \frac{\partial h_\theta^t}{\partial \theta_i} - (h_\theta^t)^{-2} \frac{\partial^2 h_\theta^t}{\partial \theta_j \partial \theta_i} \right) (X_t - (h_\theta^t)^{\frac{1}{2}})^2 \right. \\ &\quad \left. + \left((h_\theta^t)^{-2} \frac{\partial h_\theta^t}{\partial \theta_j} \frac{\partial h_\theta^t}{\partial \theta_i} \right) \left((h_\theta^t)^{-\frac{1}{2}} (X_t - (h_\theta^t)^{\frac{1}{2}}) \right) \right] + \left[\left(+ \frac{3}{2} (h_\theta^t)^{-\frac{5}{2}} \frac{\partial h_\theta^t}{\partial \theta_j} \frac{\partial h_\theta^t}{\partial \theta_i} + (h_\theta^t)^{-1} \frac{\partial^2 h_\theta^t}{\partial \theta_j \partial \theta_i} \right) \right. \\ &\quad \left. \left((h_\theta^t)^{-\frac{1}{2}} (X_t - (h_\theta^t)^{\frac{1}{2}}) \right) + \frac{1}{2} (h_\theta^t)^{-2} \frac{\partial h_\theta^t}{\partial \theta_j} \frac{\partial h_\theta^t}{\partial \theta_i} \right], \end{aligned}$$

et comme on a $E\{(h_\theta^t)^{-1}(X_t - (h_\theta^t)^{\frac{1}{2}})^2\} = E\{(\eta_t - 1)^2\} = 1$, et $E\{(h_\theta^t)^{-\frac{1}{2}}(X_t - (h_\theta^t)^{\frac{1}{2}})\} = E\{\eta_t - 1\} = 0$, nous obtenons

$$E\left\{\frac{\partial^2 q_t(\theta)}{\partial \theta_j \partial \theta_i}\right\} = \frac{3}{2} E\left\{(h_\theta^t)^{-2} \frac{\partial h_\theta^t}{\partial \theta_j} \frac{\partial h_\theta^t}{\partial \theta_i}\right\}.$$

$F(\theta_0)$ est une matrice inversible, et il existe n grand qui met F_n est une matrice inversible. en plus, puisque $E\left\{\left\|\frac{\partial^2 q_t(\theta)}{\partial \theta_j \partial \theta_i}\right\|\right\} < \infty$, la loi uniforme des grands nombres pour $\frac{\partial^2 q_t(\theta)}{\partial \theta_j \partial \theta_i}$. implique que

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -nF_n^{-1} \frac{1}{\sqrt{n}} \left(\frac{\partial L_n(\hat{\theta}_n)}{\partial \theta} - \frac{\partial L_n(\theta_0)}{\partial \theta} \right). \quad (2.3.11)$$

En utilisant le lemme 2.3.2 on obtient théorème 2.3.2 puisque $\frac{\partial \hat{L}_n(\hat{\theta}_n)}{\partial \theta} = 0$ ($\hat{\theta}_n$ est un extremum local de $\hat{L}_n$),

$$E\left\{\frac{1}{\sqrt{n}} \left| \frac{\partial L_n}{\partial \theta} - \frac{\partial \hat{L}_n}{\partial \theta} \right| \right\} \rightarrow 0.$$

$$\begin{aligned}
\left| \frac{\partial \hat{q}_t(\theta)}{\partial \theta_i} - \frac{\partial q_t(\theta)}{\partial \theta_i} \right| &= \left| \hat{h}_t^{-1} \frac{\partial \hat{h}_t}{\partial \theta_i} - \hat{h}_t^{-1} \frac{\partial \hat{h}_t}{\partial \theta_i} (X_t - \hat{h}_t^{\frac{1}{2}})^2 + \hat{h}_t^{-1} \frac{\partial \hat{h}_t}{\partial \theta_i} (\eta_t - 1) - \left(h_t^{-1} \frac{\partial h_t}{\partial \theta_i} - h_t^{-2} \frac{\partial h_t}{\partial \theta_i} \right. \right. \\
&\quad \left. \left. (X_t - h_t^{\frac{1}{2}})^2 + h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (\eta_t - 1) \right) \right|, \\
&= \left| \hat{h}_t^{-1} \frac{\partial \hat{h}_t}{\partial \theta_i} - \hat{h}_t^{-1} \frac{\partial \hat{h}_t}{\partial \theta_i} (\eta_t - 1)^2 + \hat{h}_t^{-1} \frac{\partial \hat{h}_t}{\partial \theta_i} (\eta_t - 1) - \left(h_t^{-1} \frac{\partial h_t}{\partial \theta_i} - h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (\eta_t - 1)^2 \right. \right. \\
&\quad \left. \left. + h_t^{-1} \frac{\partial h_t}{\partial \theta_i} (\eta_t - 1) \right) \right|, \\
&= \frac{1}{2} \left| \left(\hat{h}_t^{-2} \frac{\partial \hat{h}_t^2}{\partial \theta_i} - h_t^{-2} \frac{\partial h_t^2}{\partial \theta_i} \right) \eta_t - \left(\hat{h}_t^{-2} \frac{\partial \hat{h}_t^2}{\partial \theta_i} - h_t^{-2} \frac{\partial h_t^2}{\partial \theta_i} \right) (\eta_t - 1)^2 \right|, \\
&\leq \frac{1}{2\alpha^2} \left| \left(\frac{\partial \hat{h}_t^2}{\partial \theta_i} - \frac{\partial h_t^2}{\partial \theta_i} \right) \eta_t - \left(\frac{\partial \hat{h}_t^2}{\partial \theta_i} - \frac{\partial h_t^2}{\partial \theta_i} \right) (\eta_t - 1)^2 \right|, \\
&\leq \frac{1}{2\alpha^2} \left(\left| \frac{\partial \hat{h}_t^2}{\partial \theta_i} - \frac{\partial h_t^2}{\partial \theta_i} \right| |\eta_t| + \left| \frac{\partial \hat{h}_t^2}{\partial \theta_i} - \frac{\partial h_t^2}{\partial \theta_i} \right| (\eta_t - 1)^2 \right).
\end{aligned}$$

de la proposition 2.2.2, on peut écrire h_t^2 ce la forme causal $h_t^2 = \sum_{i \geq 1} c_i X_{t-i}^2$, où

$$\frac{\partial c_i(\theta)}{\partial \theta_k} \leq C_2 i p_*^i,$$

pour tout $\theta \in \Theta$ et $i \leq 0 < \infty$. Donc

$$\left| \frac{\partial \hat{q}_t(\theta)}{\partial \theta_i} - \frac{\partial q_t(\theta)}{\partial \theta_i} \right| \leq \frac{1}{2\alpha^2} (|\eta_t| + (1 + \eta_t)^2) \sum_{t \leq i < \infty} i p_*^i X_{t-i}^2.$$

$$\text{Donc, } E \left\{ \left| \frac{\partial \hat{q}_t(\theta)}{\partial \theta_i} - \frac{\partial q_t(\theta)}{\partial \theta_i} \right| \right\} < \infty.$$

Lemme 2.3.3 Soit $\theta_0 = (\alpha_0, \alpha_1, \dots, \alpha_q, \gamma_1, \dots, \gamma_q, \beta_1, \dots, \beta_p)$ avec X_t une solution stationnaire de (2.1.1), alors on a :

$$\left\| \frac{1}{n} \frac{\partial^2 L_n(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right\|_{\Theta} \xrightarrow[n \rightarrow +\infty]{p.s} 0 \text{ avec } \frac{\partial^2 L_n(\theta)}{\partial \theta \partial \theta'} = -E \left\{ \frac{\partial^2 q_0(\theta)}{\partial \theta \partial \theta'} \right\}. \quad (2.3.12)$$

Preuve 7 (preuve du lemme 2.3.3)

Le deuxième processus dérivé $(\frac{\partial^2 q_t(\theta)}{\partial \theta^2})_{t \in \mathbb{Z}}$ est stationnaire ergodique (c'est

une fonction mesurable de $X_t, X_{t-1}, \dots$) par conséquent, il satisfait à une loi uniforme des grands nombres, si son premier moment uniforme est borné. Par conséquent, en utilisant la borne $h_\theta^{-1} \leq \alpha_0^{-1}$, il existe $c > 0$ tel que

$$\left| \frac{\partial^2 q_t(\theta)}{\partial \theta_i \partial \theta_j} \right| \leq C \left[\left(\left| \frac{\partial^2 h_t}{\partial \theta_i \partial \theta_j} \right| + \left| \frac{\partial h_t}{\partial \theta_j} \right| \left| \frac{\partial h_t}{\partial \theta_i} \right| \right) + \left(\left| \frac{\partial^2 h_t}{\partial \theta_i \partial \theta_j} \right| + \left| \frac{\partial h_t}{\partial \theta_j} \right| \left| \frac{\partial h_t}{\partial \theta_i} \right| \right) |X_t - h_t| \right].$$

Nous concluons que $E \left\{ \left| \frac{\partial^2 q_t(\theta)}{\partial \theta_i \partial \theta_j} \right| \right\} < \infty$ depuis, tout $t \in \mathbb{Z}, 1 \leq i, j \leq d$ tq

$$d = i + j + 1$$

$$E\{X_t^2\} < +\infty, E\{|h_t|\} < +\infty, E\left\{ \left| \frac{\partial h_t}{\partial \theta_i} \right| \right\} < +\infty, E\left\{ \left| \frac{\partial^2 h_t}{\partial \theta_i \partial \theta_j} \right| \right\} < +\infty.$$

En conséquence, le LUGN pour $\frac{\partial^2 q_t(\theta)}{\partial \theta^2}$

Chapitre 3

Étude numériques

Dans ce chapitre nous réalisons les expériences de Mont-Carlo sur le comportement du QMLE-Normale (1,1), dans le cas du processus GJR-GARCH(1,1) définie par $X_t = \eta_t \sqrt{h_t}$, à l'aide du programme R pour différentes tailles d'échantillons. Nous illustrons la consistance du paramètre $\hat{\theta}_n$ obtenu à travers le tableau. Ensuite, en réalisant une représentation en boîte des paramètres de processus qui explique la différence entre le RMSE ($\hat{\theta}_n^{(0)}$) et le RMSE ($\hat{\theta}_n$) pour. En fin nous utilisons une représentation de l'histogramme et la densité des paramètres pour confirmer la normalité asymptotique des paramètres de modèle.

3.1 La consistance

Dans cette section, nous illustrons le RMSE ($\hat{\theta}_n$) pour plusieurs tailles d'échantillons ($n = 100, n = 1000, n = 5000$) avec des valeurs différentes de la moyenne du bruit blanc ($m = 0.3, m = 0.7, m = 1, m = 1.4$), nous exposons les résultats dans le tableau 3.1.

Considérons le GJR-GARCH(1,1) définie par :

$$\begin{cases} X_t = \eta_t h_t^{\frac{1}{2}} \\ h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \gamma_1 \mathbb{I}_{X_{t-1} < 0} X_{t-1}^2 + \beta_1 h_{t-1} \end{cases}$$

Avec $\alpha_0 = 0.7, \alpha = 0.1, \beta = 0.1, \gamma = 0.2$

		$m = 0.3$		$m = 0.7$		$m = 1$		$m = 1.4$	
		$\hat{\theta}_n^{(0)}$	$\hat{\theta}_n$	$\hat{\theta}_n^{(0)}$	$\hat{\theta}_n$	$\hat{\theta}_n^{(0)}$	$\hat{\theta}_n$	$\hat{\theta}_n^{(0)}$	$\hat{\theta}_n$
n=100	α_0	0.305	0.305	0.433	0.318	0.777	0.254	1.401	0.341
	α	0.122	0.102	0.123	0.067	0.154	0.064	0.234	0.070
	β	0.333	0.309	0.323	0.342	0.247	0.259	0.187	0.174
	γ	0.332	0.258	0.744	0.319	1.285	0.466	4.412	1.649
	Somme	1.092	0.974	1.623	1.046	2.463	1.043	6.234	2.234
n=1000	α_0	0.147	0.200	0.375	0.153	0.720	0.118	1.447	0.282
	α	0.043	0.040	0.062	0.031	0.104	0.022	0.198	0.037
	β	0.142	0.178	0.144	0.121	0.130	0.128	0.079	0.078
	γ	0.112	0.100	0.206	0.106	0.388	0.152	3.701	0.257
	Somme	0.444	0.518	0.787	0.417	1.342	0.42	2.425	0.654
n=5000	α_0	0.090	0.146	0.353	0.130	0.704	0.052	1.384	0.249
	α	0.023	0.026	0.051	0.019	0.101	0.008	0.194	0.034
	β	0.068	0.077	0.086	0.081	0.060	0.055	0.042	0.040
	γ	0.048	0.054	0.120	0.053	0.241	0.063	0.537	0.165
	Somme	0.292	0.303	0.61	0.283	1.106	0.178	2.157	0.488

TABLE 3.1 – Racine-erreur quadratique moyenne des $\hat{\theta}_n^{(0)}$ et $\hat{\theta}_n$ pour les processus GJR-GARCH considérés.

• Le tableau 3.1 montre une relation inverse entre la taille de l'échantillon et chacun des deux estimateurs ($\hat{\theta}_n^{(0)}, \hat{\theta}_n$), ce qui valide la consistance de ces

estimateurs. Nous notons que le $RMSE(\hat{\theta}_n)_{n=5000}$ est le plus proche de zéro pour les différentes valeurs de la moyenne. ces arguments valident l'utilité et la performance de la procédure d'estimation.

La représentation en boîte

Nous représentons les paramétriser $(\alpha_0, \alpha, \beta, \gamma)$ en boîte pour différents tailles d'échantillons avec $m=0$, $m=1$, dans les figures 3.1, 3.2, 3.3, 3.4 respectivement.

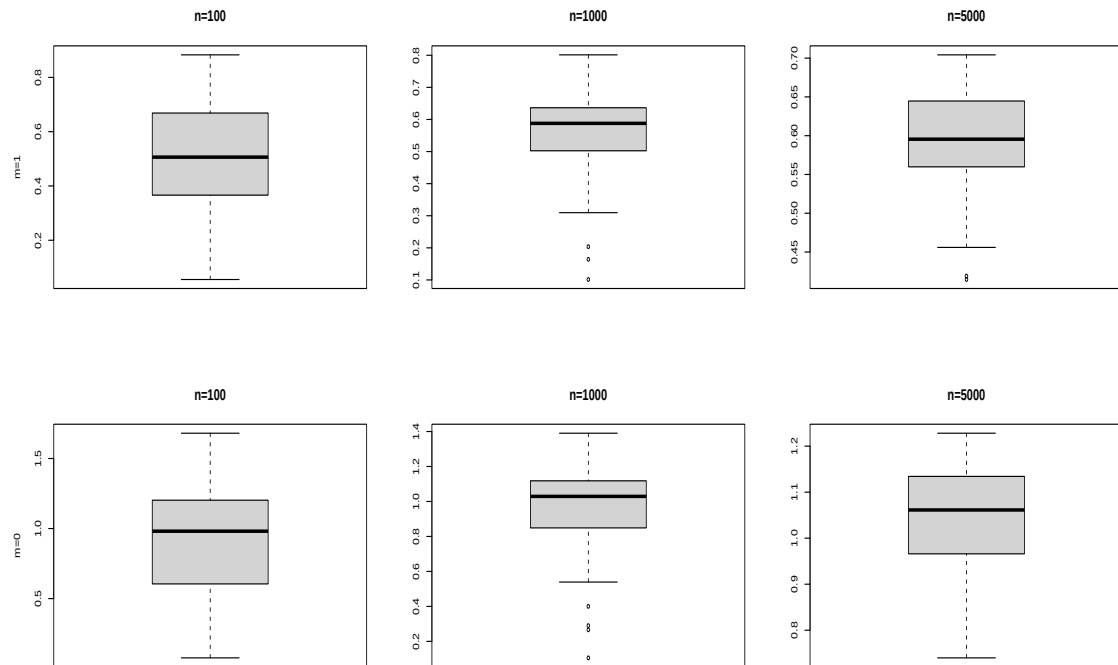


FIGURE 3.1 – Représentation en boîte de α_0 .

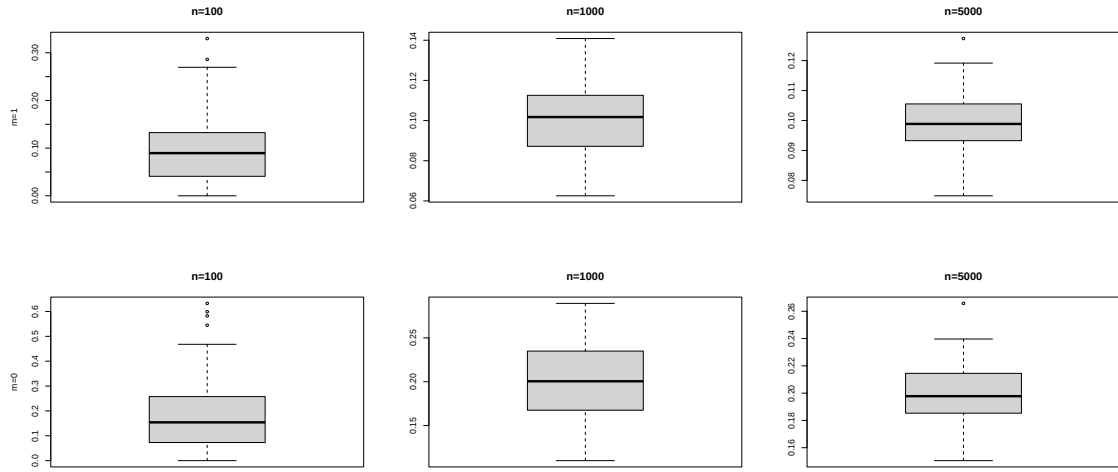


FIGURE 3.2 – Représentation en boîte de α .

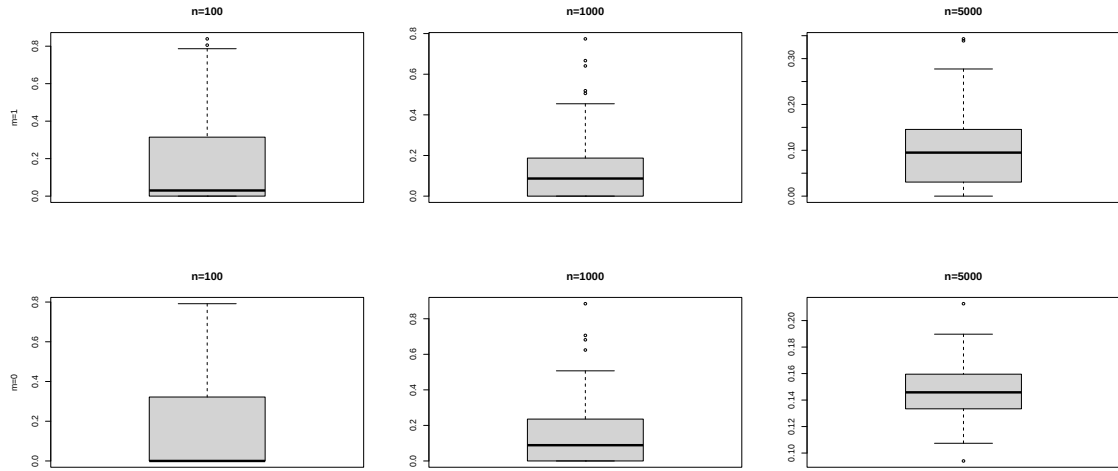


FIGURE 3.3 – Représentation en boîte de β .

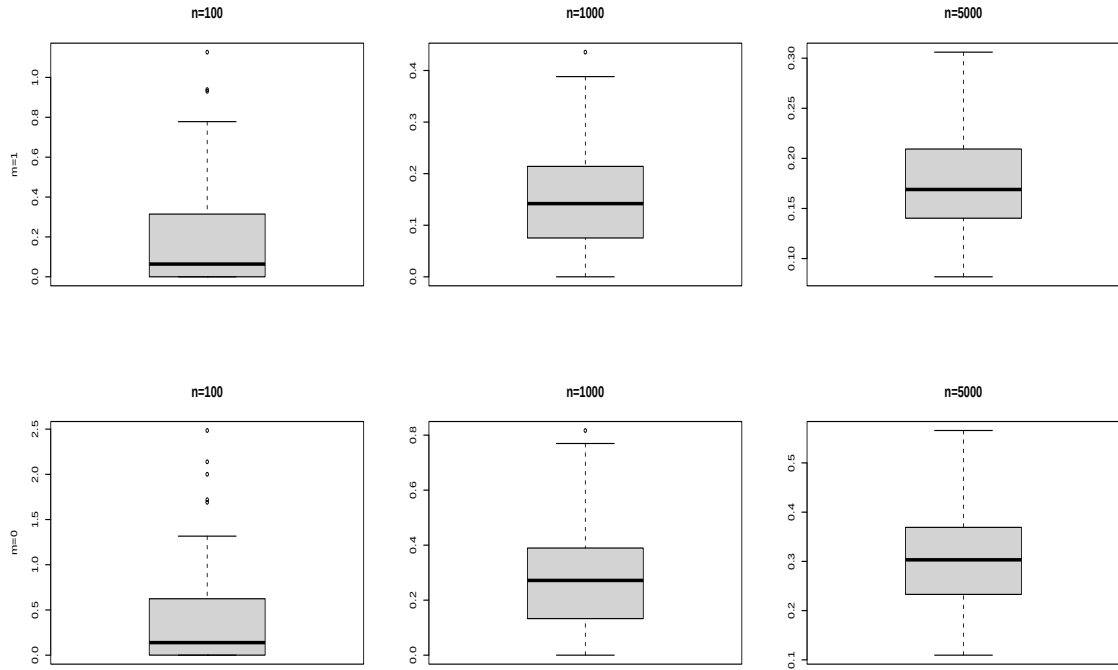


FIGURE 3.4 – Représentation en boîte de γ .

Il est clair que le RMSE diminue à mesure que la taille de l'échantillon augmente et les champs des paramètres se rapprochent du champ réel pour la base d'erreur moyenne ($m=1$) mieux que ($m=0$). ce qui valide les résultats théoriques (cohérence des estimateurs).

3.2 La normalité asymptotique

Nous avons fait une représentation de l'histogramme avec une courbe représentant la densité des chaque paramètres de processus GJR-GARCH, nous trouvons :

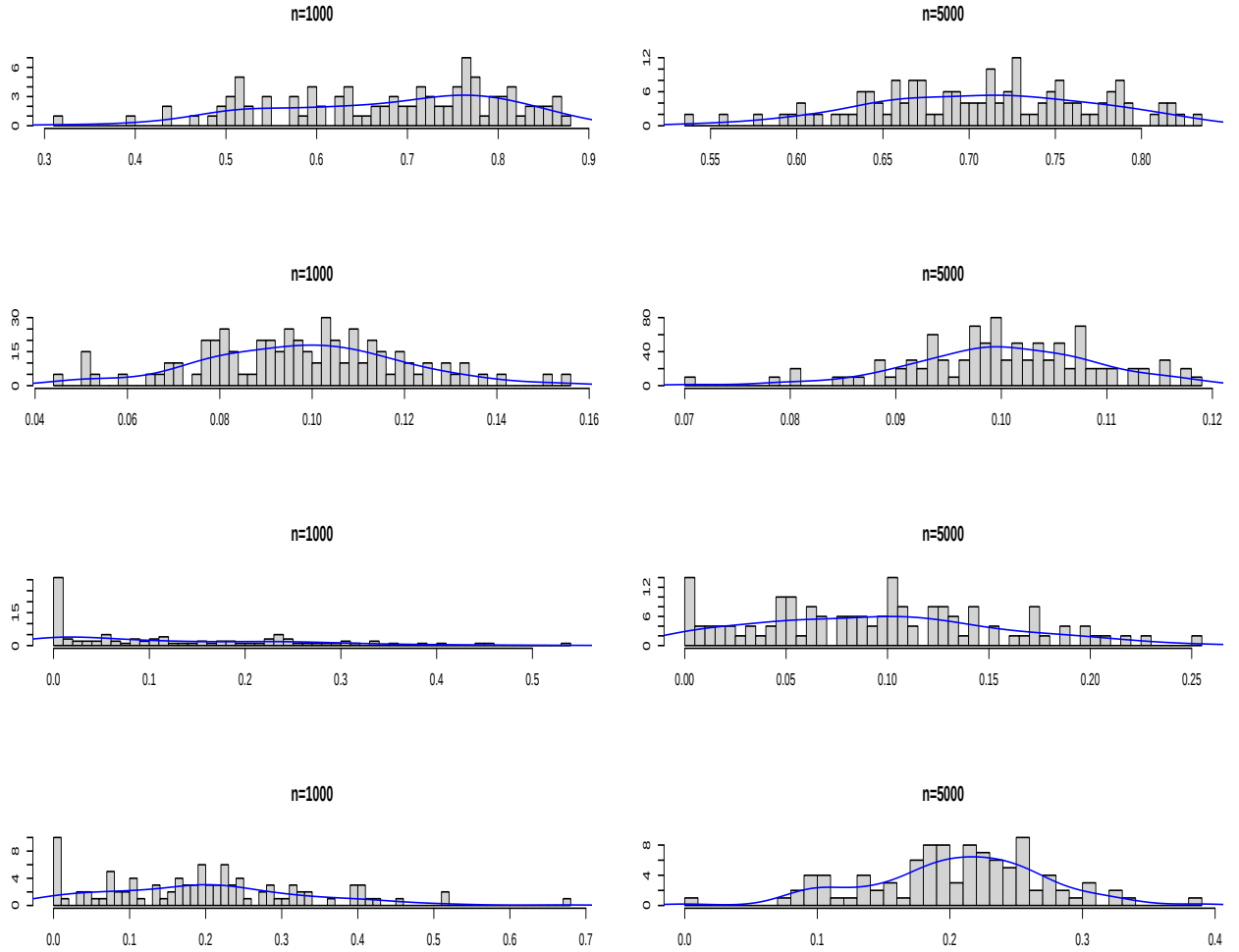


FIGURE 3.5 – Représentation en histogramme pour différents paramètres des modèle.

On remarque que plus la taille est grande, plus la densité correspond à la fréquence de l'histogramme, et les intervalles sur l'axe x convergent vers les exactes valeurs des paramètres.

Conclusion

Dans cette mémoire, notre travail s'est basé sur l'étude de modèle GJR-GARCH dans le domaine du modélisation des série chronologique. Tous d'abord nous avons mentionne des définitions fondamentales, propriétés des processus stochastique et quelques notions de bases.

Par suit nous avons défini le modèle GJR-GARCH qui introduit par Glosten, Jagannathan et Runkle et d'estimateur de quasi maximum vraisemblance Normale avec des conditions assurant l'existence et la stationnarité. Nous avons proposé des Hypothèses pour montrer les propriétés de la consistance forte. Nous avons terminé cette par la prouve de la normalité asymptotique de l'EQMV-Normale(1,1).

En fin, nous présenté un application simulée de type Monte-Carlo pour confirmer les résultats théorique obtenus. Nous avons fait une comparaison entre les deux estimateurs $(\hat{\theta}_n^{(0)}, \hat{\theta}_n)$. Et nous avons remarquer que le GJR-GARCH avec erreurs Normale(1,1) est mieux dans certains cas que GJR-GARCH avec erreurs Normale(1,0). dans le dernière partie nous avons fais une représentation en boit et de l'histogramme avec courbe densité des paramètres de processus.

Bibliographie

- [1] Bardet, J.M. and Wintenberger, O. (2009), Asymptotic normality of the Quasi-Maximum likelihood estimator for multidimensional causal process, *The Annals of Statistics*, 37, 2730-2759.
- [2] Barnarda, W. Trindade, A.A. Wickramasingheb, R.I.P (2014), Autoregressive Moving Average Models Under Exponential Power Distributions, *ProbStat Forum*, 7, 65-77.
- [3] Berkes, I., Horváth, L. and Kokoszka, P. (2003), GARCH processes : structure and estimation, *Bernoulli*, 9, 201-227.
- [4] Bollerslev, T. (1986), Generalised autoregressive conditional heteroscedasticity, *Journal of Econometrics*, 31, 307-327.
- [5] Boularouk, Y. (2021), Standard Laplace quasi-maximum likelihood estimator for GARCH processes, *Communications in Statistics - Simulation and Computation*.
- [6] Bardet, J. M, Y. Boularouk, K. Djaballah. (2017). Asymptotic behaviour of the Laplacian quasi-maximum likelihood estimator of affine causal processes. *Electronic Journal of Statistics* 11 (1) :452–79.
- [7] Boularouk, Y.(2016) Estimation des paramètres de processus causaux affines par la méthode du maximum de vraisemblance, thèse, université des sciences et technologies Houari Boumediene.
- [8] Christophe Hurlin. (2006-2007), Master Econometrie et statistique appliquée, université d'Orléans.
- [9] DeFrutos, R.F. Serrano, G.R.(1997), A generalized least squares estimation method for invertible vector moving average models, *Econom Lett* 57, 149-156.
- [10] Dugre, J.P.(1980), Generating covariance sequences and the calculation of quantization and rounding error variances in digital filters, *Acoustics, Speech and n.a. Processing, IEEE Transactions*, Vol. 28, 102 - 104.

- [11] Engle, R.F. (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987-1008.
- [12] Francq, C. and Zakoian, J.-M. (2004), Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, *Bernoulli*, 10, 605-637.
- [13] Francq, C. Lepage, G. and Zakoian, J.-M. (2011), Two-stage non Gaussian QML estimation of GARCH models and testing the efficiency of the Gaussian EQMV. *Journal of Econometrics*, 165, 246-257.
- [14] CHOUEIB, H. (2016), Evaluation d'options européennes sous un modele Bad Environment Good Environment (BEGE), HEC Montréal.
- [15] Gérard-Michel Cochard (2019), Introduction aux processus stochastiques et à la simulation, Great Britain by ISTE Editions Ltd.
- [16] Kounias, E.G. and Weng, T.-S. (1969), An inequality and almost sure convergence. *Annals of Mathematical Statistics*, 40, 1091-1093.
- [17] Xuan, H, L.Song, M. Amin, and Y.Shi.2017.Qusi-maximum likelihood estimator of Laplace(1,1)for GARCH models.*Open Mathematics* 15, 1573-48.