

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



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 fondamentales et appliquées

L'Estimateur de Quasi Maximum de Vraisemblance Laplace pour le processus GARCh

Préparé par : Bouchouk Hadda
Bouhannache Houda

Devant le jury

Hadj Ali Thoria (MAA)

Boularouk Yakoub (MCB)

Labeled Boujama (MAA)

C.U.Abd Elhafid Boussouf

C.U.Abd Elhafid Boussouf

C.U.Abd Elhafid Boussouf

Président

Rapporteur

Examinateur

Année Universitaire : 2016/2017

Remerciement

Le travail présenté dans ce mémoire a été effectué dans le cadre de la préparation du diplôme de Master spécialité «mathématique fondamentale et appliquer» à l'institut science et technologie.

Nous remercions Dieu le tout puissant qui ma donné courage, patience et force jusqu'à la fin de la réalisation de ce modeste travail.

Ce mémoire n'aurait pas pu avoir lieu sans un encadrement de très grande qualité, celui de docteur **Yakoub Boularouk** enseignant au département de mathématique et informatique pour avoir proposé et encadré ce sujet de mémoire et pour son appui scientifique et sa disponibilité durant la préparation de notre travail de fin d'étude. Ainsi que pour ses aptitudes pédagogiques dont il a fait preuve, et ses remarques constructives, la qualité des discussions qui ont ponctuées ce travail. On le remercie également pour son aide précieuse et continuelle pendant la rédaction de ce mémoire.

Nous tenons également à remercier toute nos famille pour la confiance, le soutien moral et financier qu'elles nous ont vouée jusqu'à ce stade de nos études.

Nous terminons cette page en remerciant tous nos amis et ceux dont les noms sont cités, qu'ils sachent qu'il ne sont pas oubliés.

DÉDICACE

Avec énorme plaisir, un cœur ouvert et une immense joie, que je dédie notre travail à mes très chers, respectueux et magnifiques parents qui m'ont soutenue tout au long de ma vie.

A Mon père (Boudjamaa), ma Mère (Atika), mon frère (laid "Khalil") et mes sœurs (Fouziya, Fadila, Warda, Aziza) et les petits (Rima, Youssef, zinou, Mohammed).

A mes amis (khadidja, Fatima, Souad, Amina, Sara, Imen, Loubi, hano, mankousha (Hadjer), sara, moufida, Imen, jiji, hanane, ines, maryouma, amira, racha, habiba, nihad, naima, et groupe Minouchette et ma très chers Houda).

A mes Collègues d'étude Ridha Lala Bouali qui m'ont aidé moralement et réellement a mes étude durent tout les années.

A notre encadreur Yakoub Boularouk.

Hadda

DÉDICACE

Avec énorme plaisir, un cœur ouvert et une immense joie, que je dédie notre travail à mes très chers, respectueux et magnifiques parents qui m'ont soutenue tout au long de ma vie.

A Mon père (Zidane), ma Mère (badia), mon frère (Mizou) et mes sœurs (Nadjah, Amal, Randa, Asma) et les petits (bitou, fougou, sanouma, ayouna et aloua).

A mes amis (maryouma, Sabrina, Zey dalnim, Béatrice, MiMa, Gleaming Eyes, DiDa Mayassine, Siham SouSou, Kila, Loubi, hano, Hadjer, hasna, Amina, Bouchra et ma très chers Hadouch).

A mes Collègues d'étude qui m'ont aidé moralement et réellement a mes étude tout les années.

A notre encadreur Yakoub Boularouk.

Houda

Table des matières

Introduction Générale	v
1 Définitions et Notation	1
1.1 Processus stochastique	2
1.1.1 Stationnarité	2
1.1.2 Ergodicité	5
1.1.3 Causalité	5
1.1.4 Classification des processus stochastique	6
1.1.5 Les série chronologique	7
1.2 La loi Laplace	10
1.2.1 Densité de probabilité	10
1.2.2 Fonction de répartition	10
1.3 Le Maximum vraisemblance	11
1.3.1 Estimateur	11
1.3.2 La vraisemblance	11
1.3.3 L'estimateur de maximum de vraisemblance	12
1.4 Processus Garch(p,q)	13

2	L'EQMV-Laplace de processus GARCH	18
2.1	L'étude EQMV-Laplace	18
2.1.1	Existence et stationnarité	21
2.1.2	Hypothèse nécessaire pour la convergence d'EQMV- Laplace	24
2.2	Consistance forte	25
2.2.1	Propriétés asymptotique de la quasi vraisemblance . .	25
2.3	la Normalité asymptotique	32
3	Application numérique	41
3.1	Estimation des paramètres d'un modèle GARCH	41
3.2	Etude comparative	45
	Bibliographie	48
	Conclusion générale	50

Table des figures

1.1	Évolution du cours du Dow Jones entre 1928 et 2004, données mensuelles	14
1.2	Production de poissons, en milliers de francs, en Finistère nord (Brest, Morlaix, Paimpol)	14
1.3	Indice mensuel des prix à la consommation I_t	15
1.4	Trafic voyageur de la SNCF en 2ième classe	15
1.5	Accroissement relatif mensuel de l'indice des prix	16
1.6	Modèle additif. Amplitude constante autour de la tendance . .	16
1.7	Modèle multiplicatif. Amplitude proportionnelle à la tendance	17
1.8	loi de Laplace $\theta = 0, \phi=1$	17

Introduction Générale

Une série chronologique (ou série temporelle) est une série d'observations ordonnées chronologiquement. Elles se rencontrent naturellement dans une grande variété de domaines. On peut citer : l'économie (taux de chômage, ...), la finance (cours d'action, taux d'intérêt, ...), l'écologie (pollution à l'ozone,...), le transport (avec l'exemple célèbre du trafic aérien international), la démographie. La modélisation représente une étape cruciale dans l'étude des séries chronologiques, elle a connu une grande évolution durant les cinquante dernières années et plusieurs modèles de représentation ont été proposés.

Les modèles linéaires de séries chronologiques ont contribué de manière significative dans la modélisation d'un grand nombre des phénomènes, et donné plusieurs théories mathématiques qui ont permis de prévoir les valeurs futures, mais il est pris sur ces formules linéaires qu'ils ne pouvant pas traduire le caractère cinétique de ces phénomènes, cela a conduit à entraver le développement de plusieurs aspects de la modélisation des séries chronologiques, l'hypothèse de linéarité qui caractérise ces modèles exige que les composants de temps est caractérisé par un temps fixe.

L'analyse de l'évolution du phénomène non linéaire a commencé dans les

mathématiques, la physique et d'autres sciences où les modèles non Linéaire dépend du temps sont utilisés dans l'analyse des problèmes financiers et monétaires. Les statisticiens ont également développé de nouveaux modèles, tels que les modèles d'auto-régression des moyennes mobiles et les modèles non-linéaire, ce qui a donné de même comme un nouveau monde pour la modélisation des séries chronologiques.

En 1970, Box et Jenkins ont publié un livre "*Time series : forecasting and control*", dans lequel ils ont présenté et étudié les modèles Autorégressifs moyenne mobile (ARMA) et les modèles Auto-régressifs moyenne mobile intégré (ARIMA) de bruits Gaussiens et de variance constante, ces deux processus proposés dépendent de leurs passés et celui des bruits.

En 1982, Engel a proposé le modèle Autorégressif conditionnellement hétéroscédastique (ARCH) avec bruits Gaussiens, qui a permis à la variance conditionnelle de changer en fonction des erreurs passées au fil du temps, tout en laissant la variance inconditionnelle constante. Ce modèle s'est déjà révélé utile dans la modélisation de plusieurs phénomènes.

Dans Engle (1982), Engle (1983) et Engle et Kraft (1983), des modèles pour le taux d'inflation sont construits en admettant l'hypothèse que l'inflation tend à changer au fil du temps. En 1984, Weiss a considéré les modèles ARMA avec erreurs ARCH de bruits Gaussiens. Il a utilisé ces modèles pour la modélisation de seize séries chronologiques micro-économiques américaines.

En 1986, Bollerslev a proposé une généralisation de ces modèles ARCH aux modèles Autorégressifs conditionnellement hétéroscédastique généralisé (GARCH), dont la variance conditionnelle au présent dépend de son passé

et du passé de processus.

En 2007, Doukhan et Wintenberger ont écrit sous forme générale la classe des processus causaux affines, avec $f_{\theta}^t = 0$, et $M_{\theta}^t = \sigma_{\theta}^t$.

$$X_t = M_{\theta}^t \varepsilon_t + f_{\theta}^t,$$

qui est un cas plus générale des processus GARCH.

Il y a plusieurs méthodes d'estimation paramétriques qu'on peut utiliser, par exemples : La méthode des moindres carrés (MC), la méthode d'écart absolu minimal (EAM), la méthode du maximum de vraisemblance et d'autres.

Dans notre travail, nous intéressons à la méthode du maximum de vraisemblance, cette méthode est la plus utilisée dans l'estimation des paramètres de modèles de séries chronologiques, Box et Jenkins l'ont utilisée pour l'estimation des modèles ARMA, Engel l'a utilisée pour les modèles ARCH, Bollerselv pour les modèles GARCH. La méthode du maximum de vraisemblance exige la connaissance de la loi de probabilité du bruit, pour cette raison les travaux cités ci-dessus et d'autres, supposent que la loi de probabilité est Gaussienne, elle fournit une bonne représentation pour plusieurs séries chronologiques c'est la loi la plus utilisée en statistique.

En 1988, Li et McLeod ont considéré le problème de la modélisation des séries chronologiques ARMA entraînée par les bruits non Gaussiens, comme ils ont prouvé sous des conditions générales la normalité asymptotique de l'estimateur du maximum de vraisemblance. Depuis, plusieurs travaux ont étudié les séries chronologiques avec bruits non Gaussien. En 1997, Shephard et Michael ont changé de direction en fournissant des méthodes pour l'estima-

tion des modèles de séries chronologiques avec bruits non Gaussiens, Fancq et al. (2011) ont écrit un article dans lequel ils ont proposé une estimation en deux étapes pour les paramètres des modèles GARCH avec bruits non Gaussiens.

Bardet et Boularouk (2017) ont montré la normalité et la consistance de l'estimateur de maximum de vraisemblance pour la classe des processus caux affines avec erreur de type Laplace.

Dans le premier chapitre, nous présenterons un rappel des principales définitions, propriétés des processus stochastiques des séries chronologiques et quelques notions utiles dans notre travail.

Dans le deuxième chapitre, nous avons étudié les propriétés asymptotiques de l'estimateur de quasi-vraisemblance pour le processus GARCH avec bruits Laplace au lieu des bruits Gaussiens étudié dans la littérature. Nous définissons la fonction de vraisemblance et l'estimateur de quasi-vraisemblance pour le processus GARCH avec bruits Laplace, puis nous définissons l'espace de stationnarité.

Nous prouvons la consistance d'estimateur de quasi vraisemblance pour le processus GARCH, sous condition d'existence des moments de premier ordre qui est une condition plus faible que celle exigée pour le cas des processus avec bruits Gaussien. De plus, nous prouvons la normalité asymptotique de cet estimateur, sous condition d'existence des moments d'ordre deux, un résultat qui se trouve plus performant que celui des processus avec bruits Gaussiens.

Dans le troisième chapitre, utilisant le logiciel R, nous exposons à travers un exemple de série chronologique l'estimation des paramètres de modèle

GARCH. Dans le but de confirmer les résultats asymptotiques obtenus et de comparer l'EQMV-Laplace avec l'EQMV-Gauss, nous effectuons des expériences types monte-Carlo sur plusieurs modèles de séries chronologiques de différentes tailles et de lois de probabilité du bruit.

Finalement, nous avons effectués une application numérique sur des séries chronologiques simulées.

Chapitre 1

Définitions et Notation

Ce premier chapitre donne les définitions et les propriétés des principales notions utiles pour comprendre et traiter l'estimateur de quasi maximum de vraisemblance Laplace d'un processus GARCH, nous allons discuter dans un premier temps, sur un niveau descriptif de cadre général le processus stochastique.

Le processus stochastique ou processus aléatoire représente une évolution à temps discrète ou continu (processus stochastique continu) d'une variable aléatoire. Mais dans ce chapitre nous étudions le processus stochastique dans le cas discret et les séries chronologiques.

Dans un deuxième temps, nous présentons la loi Laplace et quelque propriété de l'estimateur du quasi-maximum de vraisemblance sont établies en fin de chapitre.

1.1 Processus stochastique

Définition 1. *Un processus stochastique (aléatoire) notée $X = (X_t)_{t \in T}$ $T = t_1, t_2, \dots, t_n$ est une suite de variables aléatoires indexées dans le temps et définies sur un même espace de probabilité $(\Omega, \mathbb{P}, \mathcal{F})$. Ainsi, pour chaque instant du temps, la valeur de la quantité étudiée X_t est appelée variable aléatoire et l'ensemble des valeurs X_t quand t varie est appelée processus aléatoire.*

Définition 2. (*L'opérateur de retard*) : *L'opérateur L est dit opérateur de retard s'il décale le processus d'une unité de temps vers le passé.*

$$LX_t = X_{t-1} \quad \text{pour tout } t > 1$$

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

$$L^k X_t = X_{t-k}$$

1.1.1 Stationnarité

Lorsque chacune des variables X_t vérifie $E(X_t) < \infty$, la loi du processus est partiellement résumée par l'espérance des différentes variables et par leurs covariances. Ces moments dépendent en général du temps, ce qui est gênant quand de l'observation de réalisations du processus on veut tirer de l'information sur la loi sous-jacente de ce processus. Pour pouvoir obtenir une accumulation d'information on est amené à considérer des processus dits stationnaires

Définition 3. (*Moments d'ordre r*) : *Soit X une variable aléatoire*

admettant une densité f . Pour tout entier r , on dit que X admet un moment d'ordre r si la variable aléatoire X^r admet une espérance, c'est le cas si seulement si $x \mapsto |x|^r f(x)$ est intégrable. Dans ce cas le moment d'ordre r de X est :

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx \quad (1.1)$$

Remarque 1.

1. Le moment d'ordre 0 d'une variable aléatoire existe toujours et vaut $E(X^0) = E(1) = 1$.
2. Le moment d'ordre 1 est l'espérance de X .

Remarque 2. Soit X une variable aléatoire symétrique, c'est-à-dire admettant une densité f paire. Si X admet un moment d'ordre $2r + 1$ impair, alors ce moment, intégrale d'une fonction impaire et intégrable, est nul : $E(X^{2r+1}) = 0$. Bien évidemment, si X n'admet pas de moment d'ordre $2r+1$, on ne peut rien conclure.

Définition 4. (Moment centrés) : Soit X une variable aléatoire à densité, admettant une espérance. On dit que X admet un moment centré d'ordre r si la variable aléatoire centrée $X' = X - E(X)$ admet un moment d'ordre r . Le moment centré d'ordre r de X est :

$$E(X'^r) = E[(X - E(X))^r] \quad (1.2)$$

Définition 5. (La stationnarité au sens strict) : On dit que le processus X_t est stationnaire au sens strict (ou fortement stationnaire) si la

loi de $(X_{t_1}, \dots, X_{t_n})$ est la même que la loi de $X_{t_1+\tau}, \dots, X_{t_n+\tau}$, pour tout (t_1, t_2, \dots, t_n) avec $t_i \in T$, pour $i = 1, \dots, n$ et pour tout $\tau \in T$ avec $t_i+\tau \in T$. Ainsi, un processus aléatoire est strictement stationnaire si toutes ces caractéristiques, c'est-à-dire tous ces moments sont invariants pour tout changement de l'origine du temps.

Définition 6. (La stationnarité du second ordre) : Un processus X_t est dit stationnaire du second ordre (ou faiblement stationnaire) si $X_t, t \in T$ est du 2nd ordre et si les deux premiers moments sont invariants dans le temps :

- $E(X_t) = m = Cste \forall t \in T$
- $Var(X_t) = \sigma^2 = \gamma(0) < \infty$
- $cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = \gamma(h) \forall t \in T, \forall h \in T$

En résumé, un processus X_t est dit stationnaire du second ordre si sa moyenne, sa variance et sa covariance sont indépendantes du temps et si sa variance est finie.

Définition 7. (Bruit blanc) : Est un processus aléatoire stationnaire du second ordre centré et dont la densité spectrale de puissance est constante sur tout l'axe des fréquences.

Remarque 3. On a $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible si :

- $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une suite des variables aléatoire réelles identiquement distribuées
- $\forall (t, t') \in \mathbb{Z}^2, t \neq t' : Cov(\varepsilon_t, \varepsilon_{t'}) = 0$
- $\forall t \in \mathbb{Z} : E(\varepsilon_t) = 0 \quad \text{et} \quad E(\varepsilon_t^2) = \sigma_\varepsilon^2$

1.1.2 Ergodicité

Pour estimer la loi d'un processus, on cherche à accumuler de l'information en faisant tendre le nombre d'observations vers l'infini. Pour que ce mécanisme d'accumulation fonctionne, il faut que le processus ait une mémoire finie. C'est à dire qu'à partir d'un certain nombre d'observation, il n'y ait plus d'information nouvelle, mais simplement confirmation des informations passées.

Définition 8. (*La moyenne temporelle*) : La moyenne temporelle du processus stochastique $X_t, t \in T$ dans le cas discret est défini par :

$$\bar{X} = \frac{1}{2N + 1} \sum_{k=-N}^N X_k$$

où N est un entier naturel.

Remarque 4. On appelle réalisation ou trajectoire du processus $X_t, t \in T$ le graphique de $X(t)$ en fonction de t .

Définition 9. On dit que le processus stochastique X_t est ergodique si toute caractéristique du processus peut être obtenue, à partir d'une seule trajectoire du processus.

1.1.3 Causalité

La notion de causalité introduite par Wiener en 1956, Granger en 1969 et Christopher A. Sims dans les années 1980, apparaît comme le soubassement de l'analyse de relations dynamiques entre les séries chronologiques. Bien que

formellement définie, elle demeure sujette à plusieurs controverses parmi les économistes.

Définition 10. *Un processus stochastique X_t est dit causal s'il existe une suite de constantes $(\alpha_k)_{k \in \mathbb{Z}}$ tel que : $\sum_{k \in \mathbb{Z}} |\alpha_k| < \infty$, avec :*

$$X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \quad (1.3)$$

1.1.4 Classification des processus stochastique

Une variable aléatoire X associée à chaque $\omega \in \Omega$ une réalisation $X(\omega)$, un processus stochastique $(X_t)_{t \in T}$ associée à chaque ω une fonction (ou trajectoire) $(X_t(\omega))_{t \in T}$:

$$\begin{aligned} T &\longrightarrow E \\ t &\longmapsto X_t(\omega) \end{aligned}$$

où E est l'espace d'arrivée des variables aléatoires X_t . Passer des variables aléatoires aux processus stochastiques revient à passer en analyse des points aux fonctions.

- Lorsque l'ensemble des temps T est au plus dénombrable (par exemple $T = \mathbb{N}$), on parle de processus stochastiques à temps discret.
- Lorsqu'il est continu (i.e. $T = [0; t_0]$ ou $T = \mathbb{R}_+$), on parle de processus stochastiques à temps continu.

1.1.5 Les série chronologique

La théorie des séries chronologiques (ou temporelles) est appliquée de nos jours dans des domaines variés que l'économétrie, la médecine ou la démographie.

Définition 11. *Une série chronologique $(x_t, t \in T)$ est une suite d'observations indexées par un ensemble ordonné $T = t_1, t_2, \dots, t_n$.*

Remarque 5.

- *Une série chronologique est encore appelée série temporelle.*
- *Une série chronologique est une observation d'un processus stochastique à temps discret.*

1.1.5.1 Objectifs de l'étude d'une série chronologique

L'étude d'une série chronologique permet d'analyser, d'écrire et d'expliquer un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décision. Cette étude permet aussi de faire un contrôle, par exemple pour la gestion des stocks, le contrôle d'un processus chimique.

1.1.5.2 Description d'une série chronologique

On peut décomposer la série $(x_t, t \in T)$ en plusieurs parties :

1. Tendence représente l'évolution à long terme.

Par exemple, la série de la Figure 1.3. a tendance à augmenter de façon linéaire

2. La composante saisonnière (ou saisonnalité) correspond à un phénomène qui se répète à intervalles de temps réguliers (périodiques). En général, c'est un phénomène saisonnier d'où le terme de variations saisonnières.

Par exemple, la série de la Figure 1.4. présente des cycles réguliers au cours du temps et de même amplitude.

3. Phénomène cyclique : Il s'agit d'un phénomène se répétant mais contrairement à la saisonnalité sur des durées qui ne sont pas fixes et généralement plus longues.

4. La composante résiduelle (ε_t) correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire. On parle aussi d'aléas.

Par exemple, la série de Figure 1.5. a un comportement assez irrégulier : il y a comme une sorte de bruit de faible amplitude qui perturbe les données.

1.1.5.3 Modélisation de la série temporelle

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 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'er-

reur au modèle, représentant la différence entre la réalité et le modèle proposé :

$$X_t = f(t, \varepsilon_t)$$

On suppose de plus que les ε_t sont décorrélées. Les deux modèles de ce type les plus utilisés sont les suivants :

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

$$Y_t = f_t + S_t + X_t$$

où f_t représente la tendance (déterministe), S_t la saisonnalité (déterministe aussi) et X_t processus stochastique.

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

$$Y_t = f_t(1 + S_t)(1 + X_t)$$

2. Les modèles stochastiques : Une partie importante de l'analyse des séries chronologiques est consacrée aux modèles linéaires.

1.2 La loi Laplace

La loi (distribution) de Laplace est une variante de la distribution exponentielle. Une telle distribution est symétrique par rapport à l'origine et la valeur absolue d'une variable laplacienne est exponentiellement distribuée (pour cette raison, une variable laplacienne est parfois dite exponentielle double). Si le paramètre de cette exponentielle est $\lambda > 0$, λ sera également le paramètre de la variable laplacienne.

Remarque 6. *La loi Laplace est une loi de probabilité continue et symétrique autour de son espérance mathématique θ qui correspond également au mode et la médiane de la distribution, ou ϕ est un paramètre de dispersion.*

1.2.1 Densité de probabilité

Une variable aléatoire possède une distribution Laplace, si sa densité de probabilité est :

$$f(x) = \frac{1}{2} \lambda \exp(-\lambda|x|) \quad -\infty < x < \infty$$

1.2.2 Fonction de répartition

La fonction de répartition de la loi Laplace est donnée par :

$$\begin{aligned} F(x) &= \begin{cases} \frac{1}{2} \int_{-\infty}^0 \lambda \exp(\lambda x) dx & x < 0 \\ \frac{1}{2} \int_{-\infty}^0 \lambda \exp(\lambda x) dx + \frac{1}{2} \int_0^x \lambda \exp(-\lambda x) dx & x > 0 \end{cases} \\ &= \begin{cases} \frac{1}{2} \exp(\lambda x) & x < 0 \\ 1 - \frac{1}{2} \exp(-\lambda x) & x > 0 \end{cases} \end{aligned}$$

1.3 Le Maximum vraisemblance

1.3.1 Estimateur

Pour estimer θ on ne dispose que des données x_1, \dots, x_n , donc une estimation de θ sera une fonction de ces observations.

Définition 12. Une statistique t est une fonction des observations x_1, \dots, x_n :

$$\begin{aligned} t : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x_1, \dots, x_n &\longmapsto t(x_1, \dots, x_n) \end{aligned}$$

puisque les observations x_1, \dots, x_n sont des réalisations des variables aléatoires X_1, \dots, X_n , la quantité calculable à partir des observations $t(x_1, \dots, x_n)$ est une réalisation de la variable aléatoire $t(X_1, \dots, X_n)$.

on note souvent $t_n = t(x_1, \dots, x_n)$ et $T_n = t(X_1, \dots, X_n)$.

Définition 13. Un estimateur d'une grandeur θ est une statistique T_n à valeurs dans l'ensemble des valeurs possibles de θ . Une estimation de θ est une réalisation t_n de l'estimateur T_n .

1.3.2 La vraisemblance

Définition 14. On appelle fonction de vraisemblance (ou plus simplement vraisemblance) pour l'échantillon x_1, \dots, x_n , la fonction du paramètre θ :

$$L(\theta, x_1, \dots, x_n) = \begin{cases} P(X_1 = x_1, \dots, X_n = x_n, \theta) & \text{si les } X_i \text{ sont discrètes} \\ f_{X_1, \dots, X_n}(x_1, \dots, x_n, \theta) & \text{si les } X_i \text{ sont continues} \end{cases}$$

Si les X_i sont indépendantes et de même loi. Dans ce cas, la fonction de vraisemblance s'écrit :

$$L(\theta, x_1, \dots, x_n) = \begin{cases} \prod_{i=1}^n P(X_i = x_i, \theta) = \prod_{i=1}^n P(X = x_i, \theta) & \text{si les } X_i \text{ sont discrètes} \\ \prod_{i=1}^n f_{X_i}(x_i, \theta) = \prod_{i=1}^n f(x_i, \theta) & \text{si les } X_i \text{ sont continues} \end{cases}$$

Remarque 7. La probabilité et la densité utilisées dans cette définition sont des fonctions des observations x_1, \dots, x_n , dépendant du paramètre θ . A l'inverse, la fonction de vraisemblance est considérée comme une fonction de θ dépendant des observations x_1, \dots, x_n , ce qui permet par exemple, de dériver cette fonction par rapport à θ .

1.3.3 L'estimateur de maximum de vraisemblance

Pour n quelconque, il est logique de dire que la valeur la plus vraisemblable de θ est la valeur pour laquelle la probabilité d'observer x_1, \dots, x_n est la plus forte possible.

Définition 15. L'estimation de maximum de vraisemblance de θ est la valeur $\hat{\theta}_n$ de θ qui rend maximale la fonction de vraisemblance $L(\theta, x_1, \dots, x_n)$. L'estimateur de maximum de vraisemblance (EMV) de θ est la variable aléatoire correspondante.

Donc $\hat{\theta}_n$ sera en général calculé en maximisant la log-vraisemblance :

$$\hat{\theta}_n = \arg \max_{\theta} \ln L(\theta, x_1, \dots, x_n)$$

avec la log-vraisemblance est défini par :

$$\ln[L(\theta)] = \begin{cases} \sum_{i=1}^n \ln[P(x_i, \theta)] & \text{si } x_i \text{ sont discrète} \\ \sum_{i=1}^n \ln[f(x_i, \theta)] & \text{si } x_i \text{ sont continue} \end{cases}$$

1.4 Processus Garch(p,q)

Définition 16. *Le processus GARCH(p,q) a été introduit (par Bollerslev (1986) [10]) comme solution du système d'équation :*

$$\begin{cases} X_t = \sigma_t \cdot \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{i=1}^q d_i \sigma_{t-i}^2 \end{cases} \quad (1.4)$$

avec $\alpha_0 > 0$, $\alpha_i \geq 0$ pour $(i = 1, \dots, p)$, $d_i \geq 0$ pour $(i = 1, \dots, q)$, α_p, d_q strictement positifs et $\sum_{i=1}^q d_i < 1$, notons $\theta = (\alpha_1, \dots, \alpha_p, d_1, \dots, d_q)$.

On peut représenter σ_t sous la forme causale $\sigma_t^2 = \sum_{j=0}^{\infty} b_j X_{t-j}^2$ d'où le système d'équations (1.4) admet la représentation causal affine $X_t = M_{\theta}^t + f_{\theta}^t$ avec $M_{\theta}^t = \sigma_t^2 = \sum_{j=0}^{\infty} b_j X_{t-j}^2$, $f_{\theta}^t = 0$.

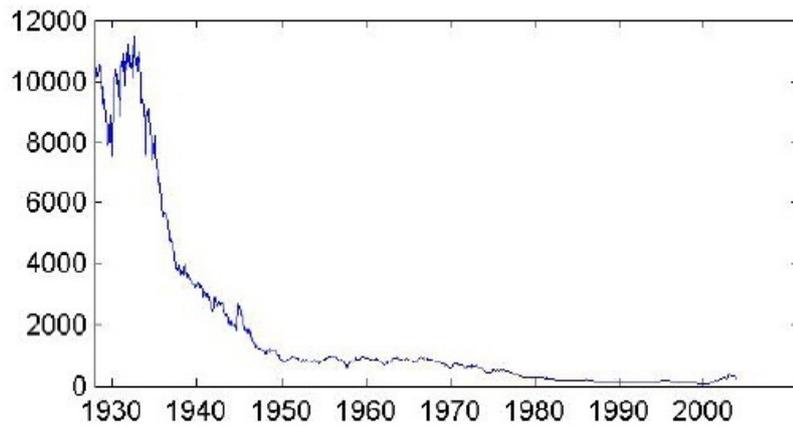


FIGURE 1.1 – Évolution du cours du Dow Jones entre 1928 et 2004, données mensuelles

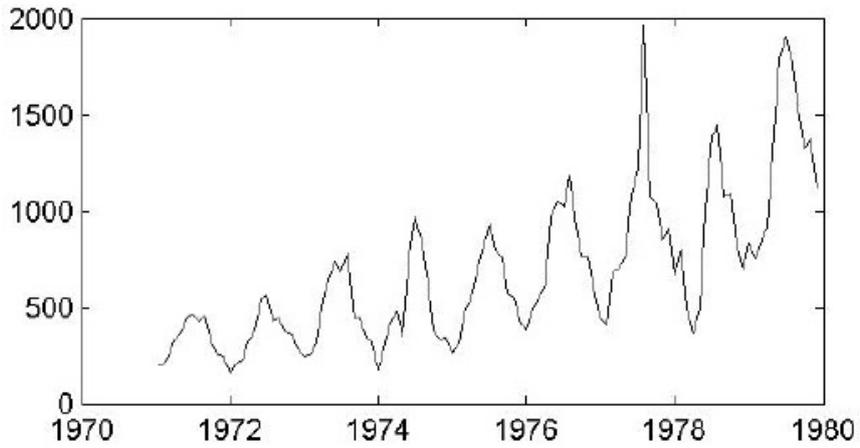


FIGURE 1.2 – Production de poissons, en milliers de francs, en Finistère nord (Brest, Morlaix, Paimpol)

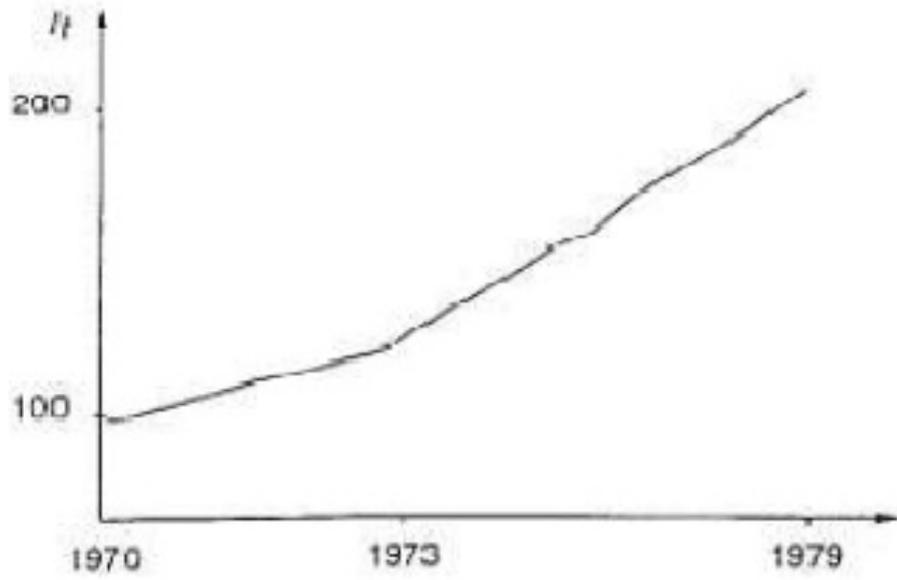


FIGURE 1.3 – Indice mensuel des prix à la consommation I_t



FIGURE 1.4 – Trafic voyageur de la SNCF en 2ième classe

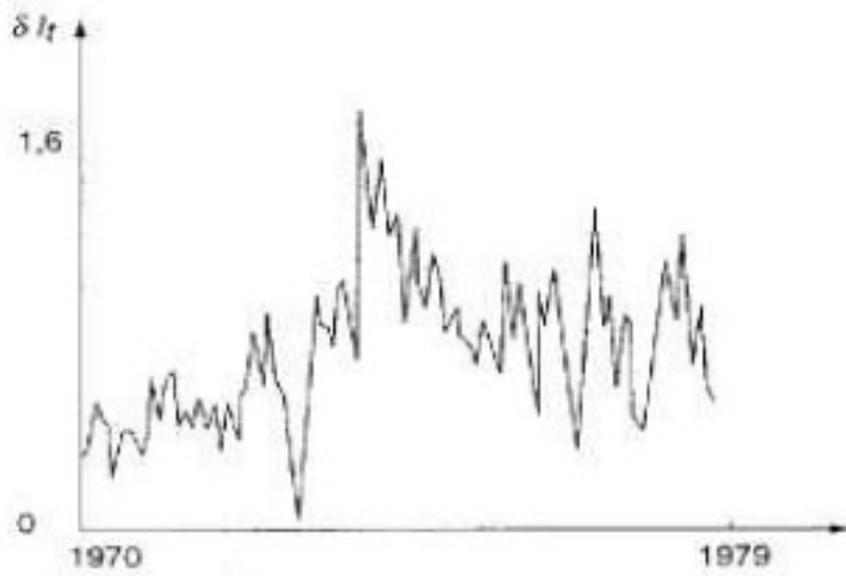


FIGURE 1.5 – Accroissement relatif mensuel de l'indice des prix

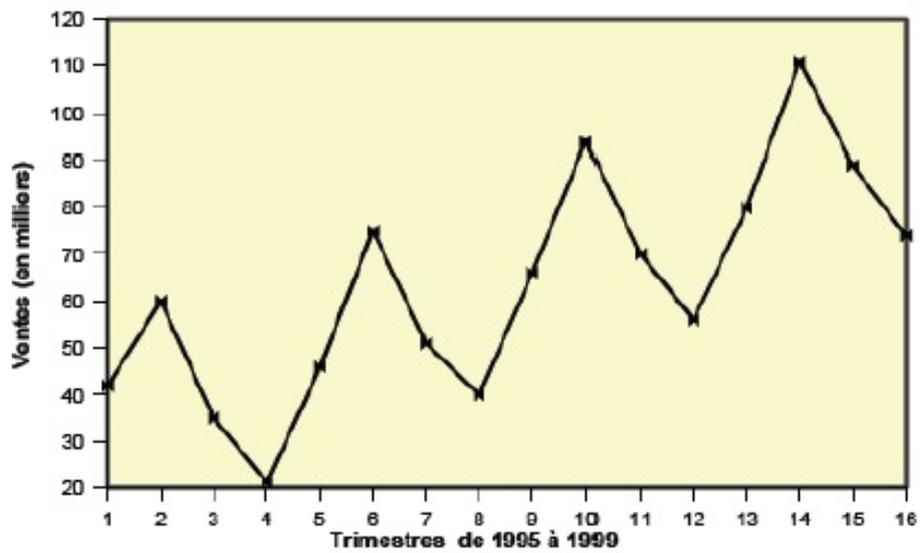


FIGURE 1.6 – Modèle additif. Amplitude constante autour de la tendance

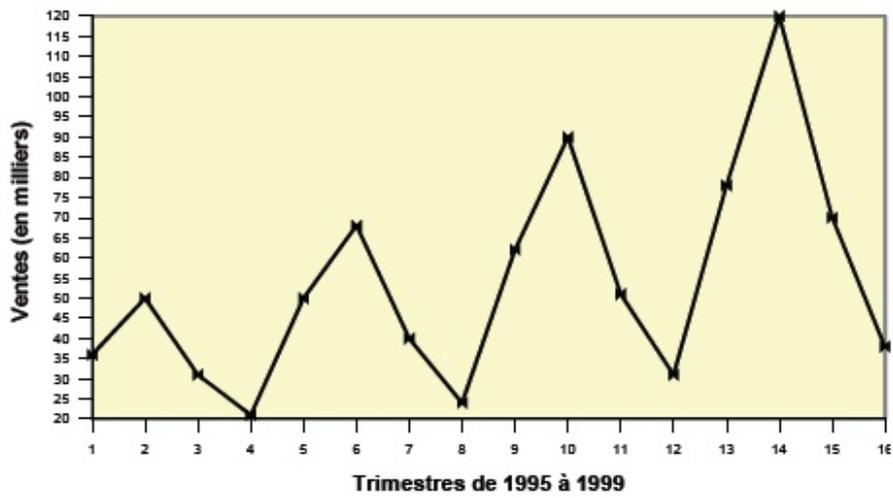


FIGURE 1.7 – Modèle multiplicatif. Amplitude proportionnelle à la tendance

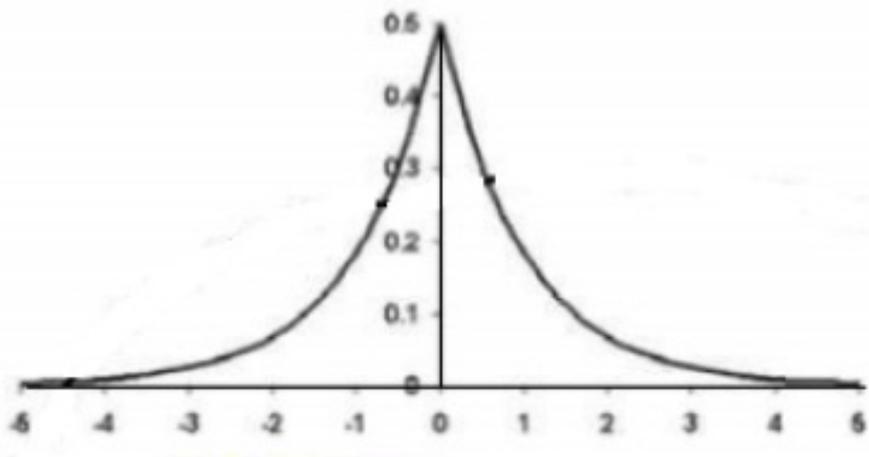


FIGURE 1.8 – loi de Laplace $\theta = 0, \phi = 1$

Chapitre 2

L'EQMV-Laplace de processus GARCH

Dans ce chapitre, nous traitons l'estimation des paramètres d'un modèle GARCH, la méthode d'estimation envisagée est la méthode du quasi-maximum de vraisemblance. Dans un premier temps, nous étudions en détail la méthode du quasi-maximum de vraisemblance, et calculons de la log-vraisemblance Laplace, qu'est écrite comme si la loi des variables ε_t était Laplace. Par suite, nous prouvons la consistance et la normalité asymptotique de cet estimateur paramétrique.

2.1 L'étude EQMV-Laplace

Soit (X_1, \dots, X_n) une trajectoire observée de X solution ou $\theta \in \mathbb{R}^n$ est inconnu. Pour estimer θ nous considérons la log-vraisemblance de (X_1, \dots, X_n) conditionnellement à (X_0, X_{-1}, \dots) .

Nous définissons respectivement la vraisemblance de Laplace et le quasi-vraisemblance de Laplace par les étapes suivant :

1. Hypothèse sur le bruit(ε_t). Dans le présent travail, nous supposons que le bruit suit une loi Laplace centré et réduite, c'est à dire :

$$g(\varepsilon_t) = \frac{1}{2}e^{-|\varepsilon_t|} \quad (2.1)$$

2. Le changement des variables :

on a :

$$X_t = \sigma_t \cdot \varepsilon_t \Rightarrow \varepsilon_t = X_t \cdot \sigma_t^{-1}$$

faire des calcules à utiliser l'intégrale pour obtenir $f(X_1 \cdots X_n)$ alors :

$$\begin{aligned} \varepsilon_t &= X_t \cdot \sigma_t^{-1} \\ \Rightarrow \frac{d\varepsilon_t}{dX_t} &= \sigma_t^{-1}, \forall t = 1, \dots, n \end{aligned}$$

$$\begin{aligned} \int f(\varepsilon_1, \dots, \varepsilon_n) d\varepsilon_1, \dots, \varepsilon_n &= \frac{1}{2^n} \int e^{-\sum_{t=1}^n |\varepsilon_t|} d\varepsilon_1 \cdots d\varepsilon_n \\ &= \frac{1}{2^n} \int e^{-\sum_{t=1}^n |X_t| \sigma_t^{-1}} \sigma_1^{-1} \cdots \sigma_n^{-1} dX_1 \cdots dX_n \end{aligned}$$

$$\begin{aligned} f(X_1 \cdots X_n) &= \frac{1}{2^n} \sigma_t^{-1} e^{-\sum_{t=1}^n |X_t| \sigma_t^{-1}} \\ &= \prod_{t=1}^n \frac{1}{2} \sigma_t^{-1} e^{-|X_t| \sigma_t^{-1}} \end{aligned}$$

3. Maximum log-vraisemblance

$$\begin{aligned}\log f(X_t) &= \log\left(\prod_{t=1}^n \frac{1}{2} \sigma_t^{-1} e^{-|X_t| \sigma_t^{-1}}\right) \\ &= -\sum_{t=1}^n \log(2) - \log(\sigma_t) - |X_t| \sigma_t^{-1}\end{aligned}$$

$$L_n(\theta) = \log f = -\sum_{t=1}^n q_t(\theta)$$

tel que :

$$\theta = \{(\alpha_i)_{i=1,p}, (d_i)_{i=1,q}\}$$

$$q_t(\theta) = \log(\sigma_t) + |X_t| \sigma_t^{-1}$$

4. Comme X_0, X_{-1}, \dots sont inconnus, en général σ_t ne sont pas calculables. Pour cela, au lieu d'utiliser directement la log-vraisemblance, un quasi-log-vraisemblance est considéré. D'où, la vraisemblance de Laplace et le quasi-vraisemblance définis par :

$$\hat{L}_n(\theta) = -\sum_{t=1}^n \hat{q}_t(\theta) \tag{2.2}$$

où

$$\hat{q}_t(\theta) = \log(\hat{\sigma}_t^{-1}) + |X_t| \hat{\sigma}_t^{-1}$$

2.1.1 Existence et stationnarité

Comme dans Doukhan et Wintenberger [9] et Bardet et Wintenberger [2], plusieurs inégalités de type Lipschitz sur σ_θ est nécessaire pour l'obtention de l'existence et la r-stationnarité (définie dans [9]) d'une solution causale et ergodique de

$$X_t = \sigma_{\theta_0}(X_{t-1}, X_{t-2}, \dots)\varepsilon_t \quad (2.3)$$

Pour $k = 0, 1, 2$ et un certain espace Θ de \mathbb{R}^d , définissons la condition Lipschitzienne sur la fonction σ_θ :

Hypothèse ($A_k(\sigma_t, \Theta)$) : $\forall x \in \mathbb{R}^\infty, \theta \in \Theta \mapsto \sigma_\theta(x) \in \mathcal{C}^k(\Theta)$, $\partial_\theta^k \sigma_\theta$ satisfait $|\partial_\theta^k \sigma_\theta(0)|_\Theta < \infty$ et il existe une suite $(\alpha_j^{(k)}(\sigma_t, \Theta))_j$ de nombres non négatifs telle que $\forall x, y \in \mathbb{R}^\mathbb{N}$

$$\|\partial_\theta^k \sigma_\theta(x) - \partial_\theta^k \sigma_\theta(y)\|_\Theta \leq \sum_{j=1}^{\infty} \alpha_j^{(k)}(\sigma_t, \Theta) |x_j - y_j| \quad \text{avec,} \quad \sum_{j=1}^{\infty} \alpha_j^{(k)}(\sigma_t, \Theta) < \infty$$

Afin d'assurer l'existence d'une solution stationnaire d'ordre r ($r \geq 1$) pour (2.3), l'ensemble $\Theta(r)$ définit par :

$$\Theta(r) = \{\theta \in \mathbb{R}^d, (A_0(\sigma_t, \{\theta\})) \text{ est vérifiée, } (E[|\varepsilon_0|^r])^{\frac{1}{r}} \sum_{j=1}^{\infty} \alpha_j^{(0)}(\sigma_t, \{\theta\}) < 1\}$$

Proposition 1. *Le processus Garch(p, q) est stationnaire au second ordre avec $E(X_t) = 0$, $Var(X_t) = \alpha_0 E(\varepsilon_t^2)(1 - A(1) - B(1))^{-1}$ et $Cov(X_t, X_s) = 0$ pour $t \neq s$ si seulement si*

$$E(\varepsilon_t^2)(A(1) + B(1)) < 1$$

avec : $A(1) = \sum_{i=1}^p \alpha_i$, $B(1) = \sum_{i=1}^q \beta_i$ et $E(\varepsilon_t^2) = 1$

Démonstration 1. Soit X_t est le processus Garch(p, q) défini par :

$$X_t = \varepsilon_t \sigma_t^{\frac{1}{2}} \quad \varepsilon_t \sim \mathcal{L}(0, 1) \quad (2.4)$$

$$\sigma_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i} \quad (2.5)$$

substituent (2.4) dans (2.5) on obtient :

$$\begin{aligned} \sigma_t &= \alpha_0 + \sum_{i=1}^p \alpha_i (\varepsilon_{t-i} \sigma_{t-i}^{\frac{1}{2}})^2 + \sum_{i=1}^q \beta_i \sigma_{t-i} \\ &= \alpha_0 + \sum_{i=1}^p \alpha_i \sigma_{t-i} \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i} \\ &= \alpha_0 + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 (\alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i-j}^2 \sigma_{t-i-j} + \sum_{i=1}^q \beta_i \sigma_{t-i-j}) \\ &\quad + \sum_{j=1}^q \beta_j (\alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i-j}^2 \sigma_{t-i-j} + \sum_{i=1}^q \beta_i \sigma_{t-i-j}) \\ &\quad \vdots \\ &= \alpha_0 \sum_{k=0}^{\infty} M(t, k) \end{aligned} \quad (2.6)$$

avec

$$M(t, k) = \prod_{i=1}^p \alpha_i^{a_i} \prod_{j=1}^q \beta_j^{b_j} \prod_{l=1}^n \varepsilon_{t-s_l}^2$$

pour

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j = k, \quad \sum_{i=1}^p a_i = n$$

et

$$1 \leq s_1 < s_2 < \dots < s_n \leq \max\{kp, (k-1)p + q\}$$

alors

$$\begin{aligned}
M(t, 0) &= 1 \\
M(t, 1) &= \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \\
M(t, 2) &= \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 \left(\sum_{i=1}^p \alpha_i \varepsilon_{t-i-j}^2 + \sum_{i=1}^q \beta_i \right) + \sum_{j=1}^q \beta_j \left(\sum_{i=1}^p \alpha_i \varepsilon_{t-i-j}^2 + \sum_{i=1}^q \beta_i \right)
\end{aligned}$$

et en générale

$$M(t, k+1) = \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 M(t-i, k) + \sum_{i=1}^q \beta_i M(t-i, k) \quad (2.7)$$

On a ε_t^2 est iid, le moment de $M(t, k)$ est indépendant de t alors :

$$E(M(t, k)) = E(M(s, k)) \text{ pour tout } k, t, s \quad (2.8)$$

de (2.7) et (2.8) on obtient :

$$\begin{aligned}
E(M(t, k+1)) &= \left(\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i \right) E(M(t, k)) \\
&\vdots \\
&= \left(\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i \right)^{k+1} E(M(t, 0)) \\
&= \left(\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i \right)^{k+1} \quad (2.9)
\end{aligned}$$

de (2.4), (2.6) et (2.9) on obtient :

$$\begin{aligned}
E(X_t^2) &= E(\sigma_t \varepsilon_t^2) \\
&= E(\alpha_0 \sum_{k=0}^{\infty} M(t, k) \varepsilon_t^2) \\
&= \alpha_0 E(\varepsilon_t^2) E(\sum_{k=0}^{\infty} M(t, k)) \\
&= \alpha_0 E(\varepsilon_t^2) \sum_{k=0}^{\infty} E(M(t, k)) \\
&= \alpha_0 E(\varepsilon_t^2) (1 - \sum_{i=1}^p \alpha_i - \sum_{i=1}^q \beta_i)^{-1}
\end{aligned}$$

si seulement si :

$$\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i < 1$$

Proposition 2. Si $\theta_0 \in \Theta(r)$ pour un certain $r \geq 1$, alors il existe une unique solution causal de (2.3) (X_t indépendant de $(\varepsilon_i)_{i>t}$ pour $t \in \mathbb{Z}$), qui est stationnaire, ergodique et satisfait $E(|X_0|^r) < \infty$.

2.1.2 Hypothèse nécessaire pour la convergence d'EQMV-Laplace

L'EQMV Laplace pourrait converger et être asymptotiquement Gaussien mais cela nécessite quelques hypothèses supplémentaires sur Θ et σ_θ :

- **Hypothèse H1** (Compacité) : Θ est un ensemble compact.
- **Hypothèse H2** (Borne inférieure de la variance conditionnelle) : il existe une constante $\alpha_0 > 0$ tel que $\forall \theta \in \Theta$, alors $\sigma_\theta(x) > \alpha_0$ pour tout $x \in \mathbb{R}^N$.
- **Hypothèse H3** (Identifiabilité) : La fonction σ_θ est tel que : pour

tout $\theta_1, \theta_2 \in \Theta$ alors $\sigma_{\theta_1} = \sigma_{\theta_2}$ implique que $\theta_1 = \theta_2$.

2.2 Consistance forte

2.2.1 Propriétés asymptotique de la quasi vraisemblance

La quasi vraisemblance est une approximation de la vraisemblance obtenu en remplaçant $\sigma_\theta^t = \sigma_\theta(X_{t-1}, X_{t-2}, \dots)$ par $\hat{\sigma}_\theta^t = \sigma_\theta(X_{t-1}, X_{t-2}, \dots, \mu)$ où $\mu = (u_n)_{n \geq 0}$ est égale à zéro.

Le lemme suivant nous permet d'avoir une estimation de l'erreur d'approximation, il représente une étape cruciale dans la preuve de la consistance d'EQMV.

lemme 2.2.1. *Supposons que $\theta_0 \in \Theta(r)$ pour $r \geq 1$ et que X est la solution causal et stationnaire de (2.3). Si $(A_0(\sigma_t, \Theta))$ est satisfaite, alors $\sigma_\theta^t \in L^r(\mathcal{C}(\Theta, \mathbb{R}^m))$ et il existe une constante $C > 0$ indépendante de t tel que :*

$$E[\|\hat{\sigma}_\theta^t - \sigma_\theta^t\|_\Theta^r] \leq CE[|X_0|^r] \left(\sum_{j \geq t} \alpha_j(\sigma_t) \right)^r \text{ pour tout } t \in \mathbb{N}^*$$

Démonstration 2. *On a $H_\theta^{t,p} = H_\theta(X_{t-1}, \dots, X_{t-p}, 0, \dots)$, tel que $H = \sigma^2$ pour tout $P \in \mathbb{N}$ et $t \in \mathbb{N}$, d'après proposition 2, $\theta_0 \in \Theta(r)$ et les inégalités communes satisfaites par les normes de la matrice*

$H_\theta^{t,p} \in \mathbb{L}^{r/2}(\mathcal{C}(\Theta, \mathcal{M}_m))$ depuis, on note $\sigma_\theta^{t,p} = \sigma_\theta(X_{t-1}, \dots, X_{t-p}, 0 \dots)$

$$\|H_\theta^{t,p}\|_\Theta^{r/2} \leq \|\sigma_\theta^{t,p}\|_\Theta^r \leq \left(\|\sigma_\theta(0)\|_\Theta + \sum_{j=1}^{\infty} \|X_{t-j}\| \alpha_j^{(0)}(\sigma_t, \Theta) \right)^r$$

Nous concluons comme ci-dessus que $H_\theta^t \in \mathbb{L}^{r/2}(\mathcal{C}(\Theta, \mathcal{M}_m))$ pour $p < q$

$$\|H_\theta^{t,p} - H_\theta^{t,q}\|_\Theta^{r/2} \leq \|\sigma_\theta^{t,p} - \sigma_\theta^{t,q}\|_\Theta^{r/2} (\|\sigma_\theta^{t,p}\|_\Theta^{r/2} + \|\sigma_\theta^{t,q}\|_\Theta^{r/2})$$

L'inégalité de cauchy schwarz implique que :

$$\begin{aligned} E[\|H_\theta^{t,p} - H_\theta^{t,q}\|_\Theta^{r/2}] &\leq (E[\|\sigma_\theta^{t,p} - \sigma_\theta^{t,q}\|_\Theta^r])^{1/2} \\ &\times [(E[\|\sigma_\theta^{t,p}\|_\Theta^r])^{1/2} + (E[\|\sigma_\theta^{t,q}\|_\Theta^r])^{1/2}] \\ &\leq B \left(E \left[\left(\sum_{p < j \leq q} \alpha_j^{(0)}(\sigma_t, \Theta) \|X_{t-j}\| \right)^r \right] \right)^{1/2} \\ &\leq B (E[\|X_0\|^r])^{1/2} \left(\sum_{p < j \leq q} \alpha_j^{(0)}(\sigma_t, \Theta) \right)^{r/2} \end{aligned}$$

pour une constante $B > 0$

1. On remarque que $\|X_0 X_0'\| \leq \|X_0\|^2$.
2. On a $(H_\theta^{t,p})_{p \in \mathbb{N}^*}$ converge vers H_θ^t on $\mathbb{L}^{r/2}(\mathcal{C}(\Theta, \mathcal{M}_m))$ ainsi, il existe une sous-suite $(p_k)_{k \in \mathbb{N}}$ tel que $\|H_\theta^{t,p_k} - H_\theta^t\|_\Theta \xrightarrow[k \rightarrow \infty]{a.s.} 0$.
grâce à la continuité du déterminant $(\det H_\theta^{t,p_k})_{k \in \mathbb{N}}$ converge aussi a.s vers $\det H_\theta^t$ alors $\det H_\theta^t \geq \underline{H}$, H_θ^t est une matrice inversible. à partir de la relation élémentaire entre la norme matricielle et le déterminant

$$\|(\widehat{H}_\theta^t)^{-1}\|_{\Theta} \leq \underline{H}^{-1/m}.$$

Nous allons d'abord prouver la forte consistance d'une suite de l'EQMV Laplace pour une solution de (2.3).

Théorème 1. *Supposons que les hypothèses **H1**, **H2** et **H3** sont satisfaites et que $\theta_0 \in \Theta(r) \cap \Theta$ avec $r \geq 1$. Soit X une solution stationnaire de (2.3). Si $(A_0(\sigma_t, \Theta))$ demeure avec :*

$$\alpha_j^0(\sigma_t, \Theta) = \mathcal{O}(j^{-\ell}) \text{ pour certain } \ell > \frac{2}{\min(r, 2)} \quad (2.10)$$

alors, la suite d'EQMV-Laplace $(\widehat{\theta}_n)_n$ converge fortement, c'est-à-dire

$$\widehat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$$

Les conditions requises pour cette forte consistance d'EQMV-Laplace sont presque les mêmes que celles requises pour la forte consistance d'une suite d'EQMV-Gaussien sauf que $r \in [1, 2)$ est avérée possible dans le théorème 1 et non dans le cas d'EQMV-Gaussien (voir Bardet et Wintenberger, 2009)[2]. En outre, si $r = 2$ la condition (2.10) sur les coefficients Lipschitzienne est plus faible pour l'EQMV-Laplace que pour l'EQMV-Gaussien. Comme nous le verrons ci-dessous de nombreuses séries chronologiques satisfaisant les hypothèses du Théorème 1.

Preuve de Théorème 1

La preuve de théorème se fait en deux étapes et suit la même procédure que dans Bardet et Wintenberger (2009)[2] :

- Dans la première étape(i) : Une forte loi uniforme des grands nombres (sur θ) satisfaite par $\frac{1}{n}\widehat{L}_n(\theta)$ qui converge vers $L(\theta) = -E[Q_0(\theta)]$ est établi.
- Dans la deuxième étape(ii) : On prouve que $L(\theta)$ admet un maximum unique en θ_0 . Ces deux conditions impliquent la forte consistance de $\widehat{\theta}_n$ (Jeantheau, 1998)[14].

(i)De la même façon et pour les mêmes raisons de la preuve du théorème 1 de Bardet et Wintenberger (2009)[2], la loi uniforme forte des grands nombres satisfaite par la moyenne d'échantillon $(\widehat{q}_t)_{t \in \mathbb{N}^*}$ (défini dans (2.2))est obtenu en prouvant que $E[|q_t(\theta)|_{\Theta}] < \infty$.

Du lemme 2.2.1 pour tout $t \in \mathbb{Z}$

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

donc :

$$\sup_{\theta \in \Theta} |q_t(\theta)| \leq \frac{1}{\alpha_0}|X_t| + |\log(\alpha_0)| + \|\sigma_\theta^t\|_{\Theta}$$

avec $r \geq 1$, on a :

- De la proposition 2 pour tout $t \in \mathbb{Z}$, $E[|X_t|] < \infty$
- Du Lemme 2.2.1 $E[\|\sigma_\theta^t\|_{\Theta}^r] < \infty$, ce qui implique que $E[\|\sigma_\theta^t\|_{\Theta}] < \infty$

Par conséquent, pour tout $t \in \mathbb{Z}$

$$E [\|q_t(\theta)\|_{\Theta}] < \infty$$

ainsi, de la loi forte uniforme des grands nombres pour $(q_t(\theta))$ on déduit que :

$$\left\| \frac{L_n(\theta)}{n} - L(\theta) \right\|_{\Theta} \xrightarrow[n \rightarrow \infty]{p.s.} 0 \quad (2.11)$$

Nous allons montrer maintenant que $\frac{1}{n} \|\widehat{L}_n(\theta) - L_n(\theta)\|_{\Theta} \xrightarrow[n \rightarrow \infty]{p.s.} 0$. En effet, pour tout $\theta \in \Theta$ et $t \in N^*$

$$\begin{aligned} |\widehat{q}_t(\theta) - q_t(\theta)| &\leq |\log(\widehat{\sigma}_\theta^t) - \log(\sigma_\theta^t) + (\widehat{\sigma}_\theta^t)^{-1}|X_t| - (\sigma_\theta^t)^{-1}|X_t|| \\ &\leq |\log(\widehat{\sigma}_\theta^t) - \log(\sigma_\theta^t)| + |((\widehat{\sigma}_\theta^t)^{-1} - (\sigma_\theta^t)^{-1})|X_t|| \\ &\leq |\widehat{\sigma}_\theta^t - \sigma_\theta^t| \alpha_0^{-1} + |\widehat{\sigma}_\theta^t - \sigma_\theta^t| \alpha_0^{-2} |X_t| \\ &\leq C(1 + |X_t|)(|\widehat{\sigma}_\theta^t - \sigma_\theta^t|) \end{aligned}$$

avec $C > 0$ Par conséquent, on a :

$$\|\widehat{q}_t(\theta) - q_t(\theta)\|_{\Theta} \leq C(1 + |X_t|)(\|\widehat{\sigma}_\theta^t - \sigma_\theta^t\|_{\Theta})$$

par le corollaire 1 de Kounias et Weng (1969)[15], la preuve est obtenue s'il existe un $s \in (0, 1]$ tel que :

$$\sum_{t \geq 1} \frac{1}{t^s} E [\|q_t(\theta) - \widehat{q}_t(\theta)\|_{\Theta}^s] < \infty \quad (2.12)$$

prouvons (2.12) avec $s = \frac{r}{2}$ quand $r \in [1, 2]$

De l'inégalité de Cauchy-Schwarz et l'hypothèse $A_0(\sigma_t, \Theta)$

$$\begin{aligned} E \left[\|\hat{q}_t(\theta) - q_t(\theta)\|_{\Theta}^{\frac{r}{2}} \right] &\leq C(E[(1 + |X_t|)^r])^{\frac{1}{2}} \\ &\times (E[(\|\hat{\sigma}_\theta^t - \sigma_\theta^t\|_{\Theta})^r])^{\frac{1}{2}} \end{aligned}$$

En utilisant le lemme 2.2.1 et les précédents résultats prouvés impliquant $E[|X_t|^r] < \infty$ et $E[\|\sigma_\theta^t\|_{\Theta}^r] < \infty$ nous obtenons :

$$\begin{aligned} E \left[\|\hat{q}_t(\theta) - q_t(\theta)\|_{\Theta}^{\frac{r}{2}} \right] &\leq C(\sum_{j \geq t} \alpha_j^0(\sigma_t, \Theta))^{\frac{r}{2}} \\ &\leq Ct^{-\frac{(\ell-1)r}{2}} \end{aligned}$$

où la dernière inégalité est obtenu de la condition (2.10) du Théorème 1 Nous avons donc :

$$\sum_{t \geq 1} \frac{1}{t^{\frac{r}{2}}} E \left[\|\hat{q}_t(\theta) - q_t(\theta)\|_{\Theta}^{\frac{r}{2}} \right] \leq A \sum_{t \geq 1} t^{-\frac{r\ell}{2}}$$

qui est fini quand $r\ell > 2$ quand $r \geq 2$, il suffit de considérer le cas $r = 2$.

En conséquence, nous obtenons :

$$\frac{1}{n} \sum_{t=1}^n \|\hat{q}_t(\theta) - q_t(\theta)\|_{\Theta} \xrightarrow[n \rightarrow \infty]{p.s} 0 \quad \frac{1}{n} \|\hat{L}_n(\theta) - L_n(\theta)\|_{\Theta} \xrightarrow[n \rightarrow \infty]{p.s} 0 \quad (2.13)$$

en utilisant (2.11) nous aurons :

$$\frac{1}{n} \|\hat{L}_n(\theta) - L(\theta)\|_{\Theta} \xrightarrow[n \rightarrow \infty]{p.s} 0 \quad (2.14)$$

(ii) Pour $\theta \in \Theta$, nous étudions :

$$L(\theta) = -E[q_0(\theta)]$$

qui peut être considéré comme une divergence de Kullback-Leibler. Nous avons

$$\begin{aligned} L(\theta) &= -E[\log(\sigma_\theta^t) + (\sigma_\theta^t)^{-1}|X_t|] \\ &= -E[\log(\sigma_\theta^t) + \frac{\sigma_{\theta_0}^t}{\sigma_\theta^t}|\varepsilon_t|] \end{aligned}$$

Par conséquent, utilisant $E[|\varepsilon_t|] = 1$, nous obtenons :

$$\begin{aligned} L(\theta_0) - L(\theta) &= E[\log(\frac{\sigma_\theta^t}{\sigma_{\theta_0}^t}) + \frac{\sigma_{\theta_0}^t}{\sigma_\theta^t}|\varepsilon_t| - 1] \\ &= E[\log(\frac{\sigma_\theta^t}{\sigma_{\theta_0}^t}) - 1 + \frac{\sigma_{\theta_0}^t}{\sigma_\theta^t}E[|\varepsilon_t|(X_{t-k})_{k \geq 1}]] \end{aligned}$$

pour $\theta \neq \theta_0$

$$\begin{aligned} L(\theta_0) - L(\theta) &> E[\log(\frac{\sigma_\theta^t}{\sigma_{\theta_0}^t}) - 1 + \frac{\sigma_{\theta_0}^t}{\sigma_\theta^t}] \\ &> h(\frac{\sigma_\theta^t}{\sigma_{\theta_0}^t}) \end{aligned}$$

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 $\sigma_\theta \neq \sigma_{\theta_0}$, on a $h(\frac{\sigma_\theta^t}{\sigma_{\theta_0}^t}) > 0$ (> 0 est remplacé par $= 0$ si $\sigma_\theta = \sigma_{\theta_0}$). Cela implique, de l'hypothèse **H3**(identifiabilité), que $L(\theta_0) - L(\theta) > 0$ presque sûrement pour tout $\theta \in \Theta, \theta \neq \theta_0$. D'où une borne

supérieur de $L(\theta)$ est atteinte seulement pour $\theta = \theta_0$ qui est un maximum unique.

2.3 la Normalité asymptotique

Une extension du Théorème 1 prouvé dans Davis et Dunsmuir (1997)[8] sera donnée ci-après et consistera une étape essentielle pour la preuve de la normalité asymptotique de l'estimateur.

Théorème 2. *Soit $(Z_t)_{t \in \mathbb{Z}}$ une suite de (variable aléatoire iid) telle que $\text{Var}(Z_0) = \sigma_Z^2 < \infty$, avec la fonction de distribution commune qui est symétrique ($F(-x) = 1 - F(x)$ pour $x \in \mathbb{R}$) et continûment différentiable dans un voisinage de 0 avec une dérivée $f(0)$ en 0. Notons $\mathcal{F}_t = \sigma(Z_t, Z_{t-1}, \dots)$ tel que σ est un tribu borélienne pour $t \in \mathbb{Z}$ et soit $(Y_t)_{t \in \mathbb{Z}}$ et $(V_t)_{t \in \mathbb{Z}}$ deux processus stationnaire adaptés à $(\mathcal{F}_t)_t$ et comme $E[Y_0^2 V_0^2] < \infty$ alors :*

$$\sum_{t=1}^n V_{t-1} (|Z_t - n^{-1/2} Y_{t-1}| - |Z_t|) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(f(0)E[V_0 Y_0^2], E[V_0^2 Y_0^2]) \quad (2.15)$$

La normalité asymptotique de l'EQMV-Laplace peut être établie utilisant des hypothèses supplémentaires.

Preuve de Théorème 2

Nous suivons le même schéma de preuve que Davis et Dunsmuir (1997)[8]

Donc, désignons par :

$$\begin{aligned}
S_n &= \sum_{t=1}^n V_{t-1} (|Z_t - n^{-1/2}Y_{t-1}| - |Z_t|) \\
&= -n^{-1/2} \sum_{t=1}^n V_{t-1} Y_{t-1} \operatorname{sgn}(Z_t) \\
&\quad + 2 \sum_{t=1}^n V_{t-1} (n^{-1/2}Y_{t-1} - Z_t) (\mathbf{1}_{0 < Z_t < n^{-1/2}Y_{t-1}} - \mathbf{1}_{n^{-1/2}Y_{t-1} < Z_t < 0}) \\
&= A_n + B_n
\end{aligned}$$

Dés que $E[V_{t-1}Y_{t-1}\operatorname{sgn}(Z_t)/\mathcal{F}_{t-1}] = E[\operatorname{sgn}(Z_t)]E[V_{t-1}Y_{t-1}] = 0$ et $E[V_0^2Y_0^2] < \infty$, on peut appliquer le théorème central limite pour la suite de différences de martingales stationnaire (voir Billingsley, 1968)[3] et :

$$A_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, E[V_0^2Y_0^2]) \quad (2.16)$$

Considérant B_n et définissant $W_{nt} = V_{t-1}(n^{-1/2}Y_{t-1} - Z_t)\mathbf{1}_{0 < Z_t < n^{-1/2}Y_{t-1}}$.

Utilisant les mêmes arguments que Davis et Dunsmuir (1997)[8], nous obtenons également :

- $\limsup_{n \rightarrow \infty} n E[W_{nt}^2] = 0$;
- $E[W_{nt}/\mathcal{F}_{t-1}] \simeq \frac{1}{2n} f(0) V_{t-1} Y_{t-1}^2$ pour $|n^{-1/2}Y_{t-1}| < \varepsilon$;
- $\sum_{t=1}^n W_{nt} \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \frac{1}{2} f(0) E[V_0 Y_0^2 \mathbf{1}_{Y_0 > 0}]$.

Donc, nous déduisons que :

$$B_n \xrightarrow[n \rightarrow \infty]{\mathcal{P}} f(0) E[V_0 Y_0^2] \quad (2.17)$$

La preuve est achevée de (2.16) et (2.17) .

Soit $\mathcal{D}^{(2)}(\Theta)$ l'espace de Banach de fonctions deux fois continuellement différentiable sur Θ muni de la norme uniforme :

$$\|h\|_{2,\Theta} = \|h\|_{\Theta} + \left\| \frac{\partial h}{\partial \theta} \right\|_{\Theta} + \left\| \frac{\partial^2 h}{\partial \theta \partial \theta'} \right\|_{\Theta}$$

lemme 2.3.1. *Soit $\theta_0 \in \Theta(r)$ ($r \geq 1$) et supposons que pour $i = 0, 1, 2$, $(A_i(\sigma_t, \Theta))$ demeurent vérifier, alors : $\sigma_{\theta}^t \in \mathbb{L}^r(\mathcal{D}^{(2)}(\Theta))$.*

Théorème 3. *Supposons que $\theta_0 \in \overset{\circ}{\Theta} \cap \Theta(r)$ où $r \geq 2$ et $\overset{\circ}{\Theta}$ désigne l'intérieur de Θ . Soit X la solution stationnaire de l'équation (2.3). Supposons que les hypothèses de théorème 1 demeurent vérifier et pour $i = 1, 2$, supposons que $(A_i(\sigma_t, \Theta))$ est vérifiée. Alors, si la fonction de probabilité cumulative de ε_0 est continument différentiable dans un voisinage de 0 avec une dérivée $g(0)$ en 0 et si la matrice Γ_{σ_t} , défini dans (2.25), est une matrice symétrique définie positive, alors :*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}_d(0, \Gamma_{\sigma_t}^{-1}((\sigma_{\varepsilon}^2 - 1)\Gamma_{\sigma_t})\Gamma_{\sigma_t}^{-1}) \quad (2.18)$$

Comme il a déjà été prouvé pour l'estimateur de la médiane (voir Van der Vaart (1996) [17]) ou pour l'estimateur des moindres écarts absolus des processus ARMA (voir Davis et Dunsmuir (1997)) [8], c'est peu surprenant que la fonction de densité de probabilité g du bruit blanc $(\varepsilon_i)_i$ influe la covariance asymptotique de (2.18). Cependant, quand $f_{\theta} = 0$, ce n'est pas le cas et c'est ce qui se passe pour les processus GARCH (voir Francq et al. (2011)[11], ou Francq et Zakoian (2015)[10]) où g n'apparaît pas dans la covariance

asymptotique.

Preuve de Théorème 3

Soit $v = \sqrt{n}(\theta - \theta_0) \in \mathbb{R}^d$. Nous allons prouver en 2/ que la maximisation de $\hat{L}_n(\theta)$ est équivalente à la maximisation de $L_n(\theta)$ qui est équivalente à la maximisation de :

$$W_n(v) = - \sum_{t=1}^n (q_t(\theta_0 + n^{-1/2}v) - q_t(\theta_0)) \quad (2.19)$$

$$\begin{aligned} &= \sum_{t=1}^n \log \left(\frac{(\sigma_{\theta_0 + n^{-1/2}v}^t)^{-1}}{(\sigma_{\theta_0}^t)^{-1}} \right) + (\sigma_{\theta_0}^t)^{-1} |X_t| \\ &- \sum_{t=1}^n (\sigma_{\theta_0 + n^{-1/2}v}^t)^{-1} |X_t| \end{aligned} \quad (2.20)$$

par rapport à v . Par conséquence, il existe une suite $(\hat{v}_n)_n$ où \hat{v}_n est un maximisateur de $W_n(v)$ tel que $\hat{v}_n = \sqrt{n}(\hat{\theta}_n - \theta_0)$. En 1/ nous allons fournir un théorème limite satisfait par $W_n(v)$. Ensuite, nous allons prouver en 3/ que $(W_n(\cdot))_n$ converge comme étant un processus de $\mathcal{C}(\mathbb{R}^d)$ (espace des fonctions continues sur \mathbb{R}^d) vers un processus limite W . Ainsi $(\hat{v}_n)_n$ converge vers le maximiseur de W .

1/ Tout d'abord, nous allons étudier le comportement asymptotique de $W_n(v)$.

Nous avons :

$$W_n(v) = \sum_{t=1}^n \log \left(\frac{(\sigma_{\theta_0 + n^{-1/2}v}^t)^{-1}}{(\sigma_{\theta_0}^t)^{-1}} \right) + |X_t| ((\sigma_{\theta_0}^t)^{-1} - (\sigma_{\theta_0 + n^{-1/2}v}^t)^{-1})$$

Nous avons :

$$W_n(v) = - \sum_{t=1}^n \log \left(\frac{\sigma_{\theta_0+n^{-1/2}v}^t}{\sigma_{\theta_0}^t} \right) + |\varepsilon_t| \left(1 - \frac{\sigma_{\theta_0}^t}{\sigma_{\theta_0+n^{-1/2}v}^t} \right)$$

Utilisation du développement de Taylor, on déduit par que pour chaque $t \in \{1, \dots, n\}$, il existe un $\bar{\theta}_1^t$ et un $\bar{\theta}_2^t$ dans le segment $[\theta_0, \theta_0 + n^{-1/2}v]$ tel que :

$$\begin{aligned} \log \left(\frac{\sigma_{\theta_0+n^{-1/2}v}^t}{\sigma_{\theta_0}^t} \right) &= n^{-1/2} (\sigma_{\theta_0}^t)^{-1} v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} \\ &\quad + \frac{1}{2} n^{-1} \left\{ v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\bar{\theta}_1^t} - (\sigma_{\theta_0}^t)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\bar{\theta}_1^t} \right)^2 \right\} \\ \frac{\sigma_{\theta_0}^t}{\sigma_{\theta_0+n^{-1/2}v}^t} &= 1 - n^{-1/2} (\sigma_{\theta_0}^t)^{-1} v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} \\ &\quad + \frac{1}{2} n^{-1} \left\{ 2 (\sigma_{\theta_0}^t)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\bar{\theta}_2^t} \right)^2 - v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\bar{\theta}_2^t} \right\} \end{aligned}$$

Alors :

$$\begin{aligned} W_n(v) &= n^{-1/2} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-1} v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} (|\varepsilon_t| - 1) \\ &\quad + \frac{1}{2n} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-1} \left\{ v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\bar{\theta}_2^t} v |\varepsilon_t| - v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\bar{\theta}_1^t} v \right\} \\ &\quad + \frac{1}{2n} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-2} \left\{ \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\bar{\theta}_1^t} \right)^2 - 2 \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\bar{\theta}_2^t} \right)^2 |\varepsilon_t| \right\} \\ &= W_n^{(1)}(v) + W_n^{(2)}(v) + W_n^{(3)}(v) \end{aligned}$$

Utilisant le théorème central limite pour les différences de martingale (voir [3]), comme :

$E[(\sigma_{\theta_0}^t)^{-1}v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} (|\varepsilon_t| - 1)/\mathcal{F}_{t-1}] = 0$ et $E[\|(\sigma_{\theta_0}^t)^{-1}v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0}\|_{\Theta}^2] < \infty$
de lemme 2.3.1) on a :

$$W_n^{(1)}(v) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, E[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2] (\sigma_{\varepsilon}^2 - 1)) \quad (2.21)$$

Sachant que $\theta \in \Theta \mapsto \frac{\partial \sigma_{\theta}^t}{\partial \theta}$ et $\theta \mapsto \frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2}$ sont des fonctions continues, $\bar{\theta}_1^t \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \theta_0$ et $\bar{\theta}_2^t \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \theta_0$ nous affirmons que $W_n^{(2)}(v)$ a la même distribution limite que $\frac{1}{2n} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-1} v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\theta_0} v(|\varepsilon_t| - 1)$
Notons que $E[(\sigma_{\theta_0}^t)^{-1}v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\theta_0} v(|\varepsilon_t| - 1)/\mathcal{F}_{t-1}] = 0$ et que $E[\|(\sigma_{\theta_0}^t)^{-1}v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\theta_0} v\|] < \infty$ (du Lemme 2.3.1). D'où, par la loi forte des grands nombres pour les différences de martingales (voir [3]), nous obtenons :

$$\frac{1}{2n} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-1} v' \left(\frac{\partial^2 \sigma_{\theta}^t}{\partial \theta^2} \right)_{\theta_0} v(|\varepsilon_t| - 1) \xrightarrow[n \rightarrow \infty]{p.s.} 0$$

et cela implique que :

$$W_n^{(2)}(v) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} 0 \quad (2.22)$$

Des arguments similaires aux précédentes induisent que $W_n^{(3)}(v)$ à la même distribution limite que $\frac{1}{2n} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} \right)^2 (1 - 2|\varepsilon_t|)$. De la loi forte des grands nombres pour les différences de martingales (voir [3]), nous obtenons :

nons :

$$\begin{aligned} \frac{1}{2n} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} \right)^2 (1 - 2|\varepsilon_t|) &\xrightarrow[n \rightarrow \infty]{p.s.} \frac{1}{2} E \left[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 (1 - 2|\varepsilon_0|) \right] \\ &\xrightarrow[n \rightarrow \infty]{p.s.} -\frac{1}{2} E \left[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 \right] \end{aligned}$$

et cela implique que :

$$W_n^{(3)}(v) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} -\frac{1}{2} E \left[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 \right] \quad (2.23)$$

Enfin, de (2.21) , (2.22) et (2.23), nous avons :

$$W_n(v) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N} \left(-\frac{1}{2} E \left[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 \right], E \left[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 \right] (\sigma_{\varepsilon}^2 - 1) \right)$$

Enfin, nous obtenons le comportement de $W_n(v)$ défini dans (2.20). Cependant, nous devons préciser la relation de dépendance asymptotique de $W_n^{(1)}$. En effet, ce terme converge vers une loi Gaussienne. Cela implique la considération de comportement asymptotique qui pourraient être réduits au comportement asymptotique de :

$$n^{-1/2} \sum_{t=1}^n (\sigma_{\theta_0}^t)^{-1} v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} (|\varepsilon_t| - 1)$$

à partir de la preuve de théorème 2 et en utilisant à nouveau le théorème central limite pour les différences de martingales, nous obtenons comme variance

asymptotique :

$$\begin{aligned} E \left[(\sigma_{\theta_0}^t)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^t}{\partial \theta} \right)_{\theta_0} (|\varepsilon_t| - 1) \right)^2 \right] &= E \left[(\sigma_{\theta_0}^0)^{-2} \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 \right] (\sigma_{\varepsilon}^2 - 1) \\ &= E \left[(\sigma_{\theta_0}^0)^{-2} \left\{ (\sigma_{\varepsilon}^2 - 1) \left(v' \left(\frac{\partial \sigma_{\theta}^0}{\partial \theta} \right)_{\theta_0} \right)^2 \right\} \right] \end{aligned}$$

comme ε_t admet une distribution de probabilité symétrique avec une médiane et espérance nulle. Alors, il n'y a pas de terme de covariance et finalement nous obtenons :

$$W_n(v) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} W(v) = v' \left(-\frac{1}{2} \Gamma_{\sigma_t} \right) v + v' N \quad (2.24)$$

avec :

$$\begin{cases} N \stackrel{\mathcal{L}}{=} \mathcal{N}(0, (\sigma_{\varepsilon}^2 - 1) \Gamma_{\sigma_t}) \\ \Gamma_{\sigma_t} = \left(E \left[\left(\frac{\partial \log(\sigma_{\theta}^0)}{\partial \theta_i} \right)_{\theta_0} \left(\frac{\partial \log(\sigma_{\theta}^0)}{\partial \theta_j} \right)_{\theta_0} \right] \right)_{1 \leq i, j \leq d} \end{cases} \quad (2.25)$$

2/ Nous considérons maintenant l'approximation $\widehat{W}_n(v)$ de $W_n(v)$ définie par :

$$\widehat{W}_n(v) = - \sum_{t=1}^n (\widehat{q}_t(\theta_0 + n^{-1/2}v) - \widehat{q}_t(\theta_0)) \quad \text{pour tout } v \in \mathbb{R}^d$$

partir des hypothèses des Théorèmes 1 et (2.13) on a :

$$\frac{1}{n} \sum_{t=1}^n \|\widehat{q}_t(\theta) - q_t(\theta)\|_{\Theta} \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

Alors on a :

$$\widehat{W}_n(v) = W_n(v) + R_n(v)$$

avec

$$\left[\sup_{v \in \mathbb{R}^d} |R_n(v)| \right] \leq 2 \sum_{t=1}^n [\|\widehat{q}_t(\theta) - q_t(\theta)\|_{\Theta}] \xrightarrow[n \rightarrow \infty]{p.s.} 0$$

et donc :

$$\widehat{W}_n(v) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} W(v) \tag{2.26}$$

avec W définie dans (2.24)

3/ De (2.26), la preuve du théorème 2 et les mêmes arguments que ceux de la précédente preuve du théorème 2 de Davis et Dunsmuir (1997) [8], nous déduisons que les distributions finies $(\widehat{W}_n(v_1), \dots, \widehat{W}_n(v_k))$ converge vers $(W(v_1), \dots, W(v_k))$ pour tout $(v_1, \dots, v_k) \in (\mathbb{R}^d)^k$. De plus, toujours en suivant la preuve de théorème 2, $(W_n(v))_v$ converge vers $(W(v))_v$ comme étant un processus dans l'espace des fonctions continues \mathcal{C}^0 . Par conséquent, le maximum \widehat{v} de $\widehat{W}_n(v)$ satisfait :

$$\widehat{v} = (\Gamma_{\sigma_t})^{-1}N$$

et cela implique 2.18

Chapitre 3

Application numérique

Dans ce chapitre, nous allons étudier numériquement, la modélisation des processus GARCH à base d'erreur Laplace. Pour cela

- Dans la première section, nous modélisons deux séries chronologiques simulées ARCH(1) de paramètres $\omega = 1.2, \alpha = 0.5$ et GARCH(1,1) de paramètres $\omega = 1.2, \alpha = 0.5, \beta = 0.3$.
- Dans la deuxième section, nous présentons une étude comparative entre l'EQMV-Laplace et l'EQMV-Gauss en exposons les résultats des expériences de Monte-Carlo sur le modèle GARCH(1,1).

3.1 Estimation des paramètres d'un modèle GARCH

L'estimation d'un ARCH ou GARCH se fait par la méthode du maximum de vraisemblance. Cette méthode nécessite de préciser la loi du bruit

blanc ε_t . Nous avons supposé que $\varepsilon_t \sim \mathcal{L}(0, 1)$ un bruit blanc Laplace. Nous illustrons la méthode en estimant les paramètres de deux modèles de séries chronologiques simulée.

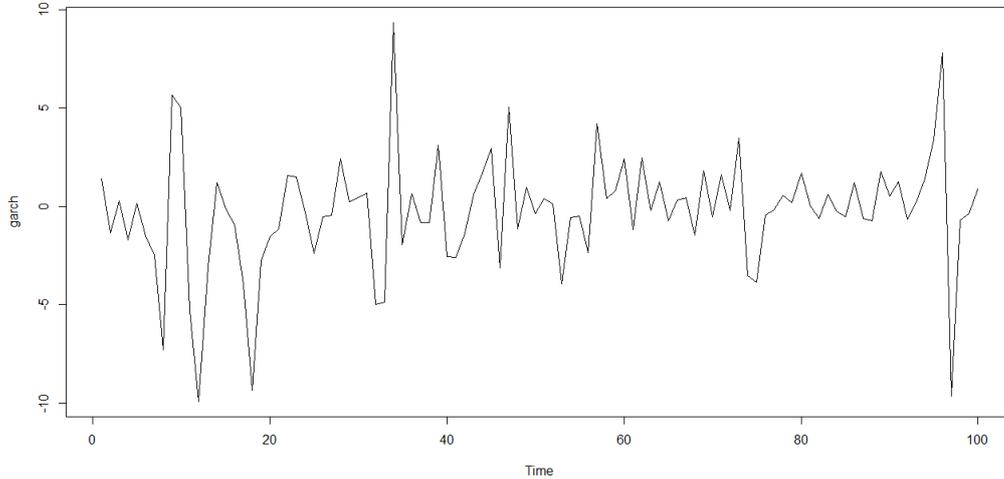
Exemple 3.1.1. (*Modèle ARCH(1)*)

Soit x_t la série chronologique ARCH(1) simulée de taille $n = 100$, présentée dans le tableau 3.1

Time t	1	2	3	4	5	6	7	8	9	10
x_t	1.412	-1.334	0.291	-1.722	0.149	-1.521	-2.487	-7.296	5.683	5.019
Time t	11	12	13	14	15	16	17	18	19	20
x_t	-5.286	-9.913	-3.028	1.227	-0.111	-0.95	-3.975	-9.35	-2.766	-1.547
Time t	21	22	23	24	25	26	27	28	29	30
x_t	-1.15	1.559	1.504	-0.282	-2.417	-0.532	-0.437	2.432	0.252	0.445
Time t	31	32	33	34	35	36	37	38	39	40
x_t	0.667	-5.004	-4.879	9.34	-1.961	0.637	-0.828	-0.805	3.112	-2.553
Time t	41	42	43	44	45	46	47	48	49	50
x_t	-2.585	-1.382	0.593	1.702	2.938	-3.113	5.054	-1.147	0.978	-0.379
Time t	51	52	53	54	55	56	57	58	59	60
x_t	0.416	0.1	-3.929	-0.572	-0.508	-2.376	4.202	0.392	0.821	2.431
Time t	61	62	63	64	65	66	67	68	69	70
x_t	-1.195	2.461	-0.218	1.26	-0.751	0.336	0.425	-1.475	1.803	-0.542
Time t	71	72	73	74	75	76	77	78	79	80
x_t	1.593	-0.195	3.487	-3.543	-3.85	-0.472	-0.177	0.564	0.207	1.693
Time t	81	82	83	84	85	86	87	88	89	90
x_t	0.031	-0.603	0.599	-0.257	-0.515	1.222	-0.63	-0.725	1.765	0.502
Time t	91	92	93	94	95	96	97	98	99	100
x_t	1.247	-0.641	0.181	1.393	3.396	7.805	-9.638	-0.681	-0.382	0.888

TABLE 3.1 – Série ARCH(1) de paramètres $\omega = 1.2, \alpha = 0.5$.

Nous représentons graphiquement la série comme fonction du temps



La série chronologique ARCH(1).

La fonction de vraisemblance pour le modèle ARCH(1) est donnée par

$$L_n(\theta) = -\frac{1}{2} \sum_{t=1}^n \left[\log(\omega + \alpha_1 X_{t-1}^2) + (\omega + \alpha_1 X_{t-1}^2)^{-1} |X_t| \right]. \quad (3.1)$$

Pour obtenir \hat{L}_n , nous remplaçons $(X_{-i}), i \geq 0$ par zéro.

Les valeurs de ω, α_1 qui muniment (3.1) sont :

$$(\hat{\omega}, \hat{\alpha}) = \arg \max_{\theta} \ln L(\theta, x_1, \dots, x_n) = (1.17, 0.47)$$

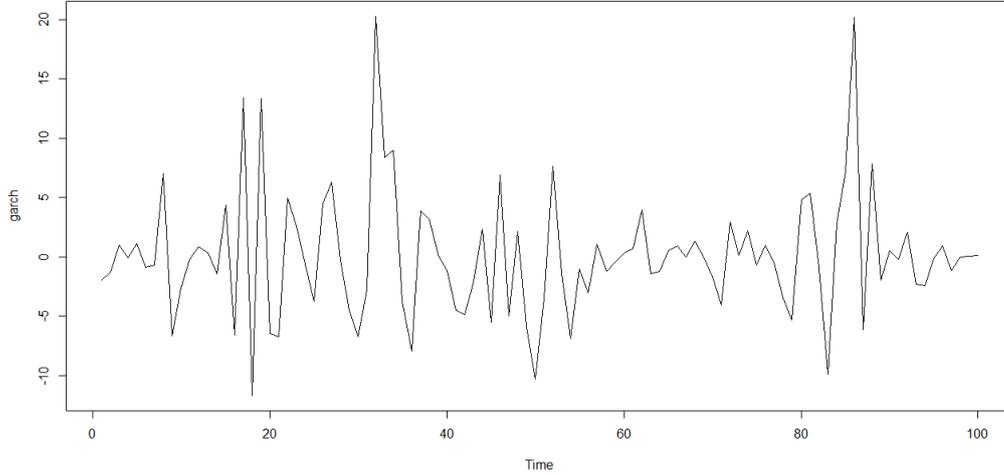
Exemple 3.1.2. (Modèle GARCH(1,1))

Soit y_t la série chronologique GARCH(1,1) simulée de taille $n = 100$, présentée dans le tableau 3.2

Time t	1	2	3	4	5	6	7	8	9	10
x_t	-1.961	-1.253	1.012	-0.051	1.11	-0.854	-0.662	7.01	-6.647	-2.611
Time t	11	12	13	14	15	16	17	18	19	20
x_t	-0.184	0.843	0.306	-1.412	4.357	-6.615	13.4	-11.656	13.317	-6.417
Time t	21	22	23	24	25	26	27	28	29	30
x_t	-6.708	4.926	2.589	-0.605	-3.761	4.503	6.325	-0.292	-4.529	-6.694
Time t	31	32	33	34	35	36	37	38	39	40
x_t	-2.752	20.244	8.353	8.999	-3.756	-7.928	3.889	3.159	0.108	-1.164
Time t	41	42	43	44	45	46	47	48	49	50
x_t	-4.514	-4.847	-2.064	2.345	-5.542	6.913	-4.963	2.121	-5.822	-10.285
Time t	51	52	53	54	55	56	57	58	59	60
x_t	-3.529	7.675	-1.44	-6.836	-1.001	-3.055	1.044	-1.233	-0.38	0.311
Time t	61	62	63	64	65	66	67	68	69	70
x_t	0.764	3.98	-1.411	-1.21	0.509	0.939	-0.009	1.304	-0.032	-1.685
Time t	71	72	73	74	75	76	77	78	79	80
x_t	-4.055	2.976	0.139	2.197	-0.695	0.97	-0.455	-3.427	-5.329	4.797
Time t	81	82	83	84	85	86	87	88	89	90
x_t	5.382	-0.928	-9.903	2.839	7.191	20.19	-6.121	7.821	-1.972	0.53
Time t	91	92	93	94	95	96	97	98	99	100
x_t	-0.185	2.048	-2.261	-2.443	-0.153	0.912	-1.163	-0.008	0.082	0.13

TABLE 3.2 – Série GARCH(1,1) de paramètres $\omega = 1.2, \alpha_1 = 0.5, \beta_1 = 0.3$.

Nous représentons graphiquement la série come fonction du temps



La série chronologique $GARCH(1,1)$.

La fonction de vraisemblance pour le modèle $GARCH(1,1)$ est donnée par

$$L_n(\theta) = -\frac{1}{2} \sum_{t=1}^n \left[\log(\omega + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2) + (\omega + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^{-1} |X_t| \right]. \quad (3.2)$$

Pour obtenir \hat{L}_n , nous remplaçons $(X_{-i}), i \geq 0$ par zéro.

Les valeurs de ω, α_1 qui munimisent (3.2) sont :

$$(\hat{\omega}, \hat{\alpha}_1, \beta_1) = \arg \max_{\theta} \ln L(\theta, x_1, \dots, x_n) = (1.28, 0.50, 0.26)$$

3.2 Etude comparative

Pour illustrer les résultats asymptotiques énoncés, nous réalisons des expériences de Monte-Carlo sur le comportement d'EQMV-Laplace (noté $\hat{\theta}_n^{LQL}$)

une séries chronologiques type GARCH(1,1), pour plusieurs tailles d'échantillons et distributions de probabilités (après la normalisation de sorte que $E|\varepsilon_t| = 1$). Nous proposons également une comparaison avec les résultats obtenus par l'EQMV Gaussien noté $\hat{\theta}_n^{GQL}$ (après une normalisation de sorte que $E|\varepsilon_t|^2 = 1$).

Plus précisément, les distributions de probabilités considérées de (ε_t) sont :

- Distribution Gaussienne centrée notée \mathcal{N} ;
- Distribution de Laplace Centrée notée \mathcal{L} ;
- Distribution Uniforme Centrée notée \mathcal{U} ;
- Distribution de Student Centrée avec degré de liberté 3, notée t_3 ;
- Mélange Gaussien centré et normalisé avec une distribution de probabilité $0.05 * \mathcal{N}(-2, 0.16) + \mathcal{N}(0, 1) + 0.05 * \mathcal{N}(2, 0.16)$ et notée \mathcal{M} .

Toutes ces distributions de probabilité sont normalisées de sorte que $E[|\varepsilon_0|] = 1$, requis pour l'EQMV- Laplace. Pour utiliser l'EQMV Gaussien nécessitant $E[|\varepsilon_0|] = 1$, il est nécessaire de considérer le modèle avec $\sigma_\theta^t = \frac{E[|\varepsilon_0|]}{\sqrt{Var(\varepsilon_0)}}\sigma_\theta$ au lieu de σ_θ .

Dans le tableau 3.3, nous avons les expériences de Monte-Carlo sur le processus GARCH(1,1) définie par $X_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 + \beta \varepsilon_{t-1}^2} \varepsilon_t$ avec $\alpha_0 = 0.2$, $\alpha_1 = 0.4$ et $\beta = 0.2$.

n	\mathcal{L}		\mathcal{N}		t_3		\mathcal{U}		\mathcal{M}	
	$\hat{\theta}_n^{GQL}$	$\hat{\theta}_n^{LQL}$								
α_0	100	0.112	0.105	0.095	0.100	0.126	0.081	0.047	0.134	0.114
	1000	0.036	0.032	0.028	0.028	0.058	0.023	0.018	0.066	0.051
	5000	0.016	0.014	0.012	0.012	0.043	0.010	0.015	0.040	0.023
α_1	100	0.162	0.157	0.149	0.150	0.364	0.115	0.070	0.507	0.429
	1000	0.061	0.056	0.449	0.449	0.333	0.030	0.033	0.160	0.136
	5000	0.029	0.026	0.020	0.020	0.193	0.013	0.030	0.086	0.058
β	100	0.225	0.209	0.188	0.190	0.429	0.163	0.105	0.483	0.390
	1000	0.060	0.055	0.051	0.051	0.174	0.044	0.022	0.170	0.169
	5000	0.027	0.024	0.022	0.022	0.075	0.019	0.009	0.072	0.075

TABLE 3.3 – Racine d'erreur moyenne quadratique $\hat{\theta}_n^{LQL}$ et $\hat{\theta}_n^{GQL}$ pour les processus GARCH(1, 1).

Bibliographie

- [1] Agnes,L.(2011), Séries Chronologiques. Université de Toulouse le Mirail,4,5,7, 14-16.
- [2] 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.
- [3] Billingsley, P. (1968), *Convergence of Probability Measures*. Wiley, New York.
- [4] Bollerslev, Tim.(1986),Generalized autoregressive conditional heteroskedasticity, University of California at San Diego, La Jolla, CA 92093, USA Institute of Economics, University of Aarhus, Denmark,310,323-325.
- [5] Boularouk Yakoub.(2010),Modeles stochastique et leur prévision par des processus autoregressifs-mémoire présenté pour l'obtention du Diplôme de Magister en Mathématiques,9.
- [6] Boularouk Yakoub.(2016), Estimation des paramètres de processus causaux affines par la méthode du maximum de vraisemblance-Thèse Présentée pour l'obtention du diplôme de doctorat en : mathématiques spécialité : statistiques,13-28.

- [7] Corinne Perraudin (2004-2005) Séries Chronologiques. Université Paris, 4-6.
- [8] Davis, R. and Dunsmuir, W. (1997), Least Absolute Deviation Estimation for Regression with ARMA Errors. *Journal of Theoretical Probability*, 10, 481-497.
- [9] Doukhan, P. and Wintenberger, O. (2007), Weakly dependent chains with infinite memory. *Stochastic Process. Appl*, 118, 1997-2013.
- [10] Francq, C. and Zakoian, J-M. (2015), Risk-parameter estimation in volatility models. *Journal of Econometrics*, 184, 158-173.
- [11] 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.
- [12] Gaudoin, O et BÉGUIN, M. (2009). *Principes et méthodes statistiques. Ensimag-2ème Année, INP Grenoble*. 36, 38, 39.
- [13] Jean-marc Bardet and Olivier Wintenberge. (2009), Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes.
- [14] Jeantheau, T. (1998), Strong consistency of estimators for multivariate arch models. *Econometric Theory*, 14, 70-86.
- [15] Kounias, E.G. and Weng, T.S. (1969), An inequality and almost sure convergence. *Annals of Mathematical Statistics*, 40, 1091-1093.
- [16] Lefebvre, Mario *Processus stochastique appliqués*. (2005) École Polytechnique de Montréal.

- [17] Van der Vaart, A.W. (2000), *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Conclusion générale

Dans cette mémoire, notre travail s'est porté sur la modélisation des séries chronologiques. Nous avons montré la consistance des estimateurs de quasi vraisemblance pour le processus GARCH, sous l'hypothèse que le bruit est non Gaussien, précisément un bruit Laplace. Nous avons réussi à prouver la consistance seulement sous condition d'existence des moments d'ordre un, une condition plus faible que celle exige dans la modélisation de processus avec bruits Gaussiens.

Comme nous avons montré la normalité asymptotique sous condition d'existence des moments d'ordre deux au lieu des moments d'ordre quatre nécessaire pour le processus de bruit Gaussien. Á travers des expériences numériques de type Monte-Carlo, nous avons pu vérifier les résultats théoriques obtenus. De même, nous avons montrer que l'EQMV-Laplace fournit une estimation plus exact que l'EQMV-Gauss, sauf bien sur dans le cas d'une distribution Gaussienne.

Résumé

La modélisation stochastique est une étape crucial dans l'étude des séries chronologiques, la prévision est encore et toujours vitales pour le monde des entreprises désireuses d'anticiper leurs activités, leurs besoins et les moyens à mettre en œuvre pour satisfaire la demande des clients. L'approche ARCH-GARCH est introduit pour éviter la lacune majeur des modèles ARMA. Cette approche constitue un outil puissant pour traiter les phénomènes non linéaires et les mouvements imprévisibles ou volatils du marché financier.

A cet effet, notre travail est divisé en trois chapitre :

Le premier chapitre est un aperçu général sur des définitions et des notations pour éclairer et facilité les informations qui on ensuite. Dans le deuxième chapitre, nous prouvons les propriétés asymptotiques de l'estimateur quasi-maximum de vraisemblance Laplace (EQMV-Laplace) pour le modèle GARCH, à savoir la consistance forte et la normalité asymptotique.

Finalement, dans le troisième chapitre, nous présentons une application numérique sur une série chronologique GARCH simulée, comme nous exposons une étude de comparaison numérique entre l'estimateur quasi-maximum de vraisemblance Laplace (EQMV-Laplace) et l'estimateur quasi-maximum de vraisemblance Gaussien (EQMV-Gaussien).

Mots clés :

Processus stochastique, bruit blanc, stationnarité, série chronologique, la loi Laplace, estimateur de quasi-maximum vraisemblance, processus GARCH,

consistance forte, la normalité asymptotique.

Abstract

Stochastic modeling is a crucial step in time series study, and forecasts are still vital for the world of companies that want to anticipate their activities, their needs and the means to be implemented to meet the demand client. The ARCH-GARCH approach is introduced to avoid the major flaw in ARMA models. This approach is a powerful tool for dealing with non-linear phenomena and unpredictable or volatile financial market movements. To this end, our work is divided into three chapters :

The first chapter is a general overview on definitions and notations to illuminate and facilitated the information that one then. In the second chapter, we prove the asymptotic properties of the quasi-maximum Laplace likelihood estimator (EQMV-Laplace) for the GARCH model, namely the strong consistency and the asymptotic normality. Finally, in the third chapter, We present a numerical application on a simulated GARCH time series, as we show a numerical comparison study between the quasi-maximum likelihood estimator Laplace (EQMV-Laplace) and the quasi-maximum Gaussian likelihood estimator (Gaussian EQMV).

Keywords : Stochastic process, white noise, stationarity, time series, the law laplace, quasi-maximum likelihood estimator, GARCH process, strong consistency, asymptotic normality.

تلخيص:

النمذجة العشوائية هو خطوة حاسمة في دراسة السلاسل الزمنية, التوقعات لا تزال و دائما حيوية لرجال الاعمال الراغبين في توقع انشطتهم و احتياجاتهم و ذلك العملاء. يتمحور هذا العمل حول سيرورة تقبل تمثيل انحذار ذاتي تغاير شرطي لتلبية طلب غير ثابت. هذا النهج هو أداة قوية للتعامل مع الظواهر غير الخطية و الحركات المتقلبة و القلقة في السوق المالية, و بشكل اكثر تحديدا, فإنها يمكن أن تقدم نماذج و حلول للسلاسل الزمنية المالية.

ينقسم هذا العمل إلى ثلاثة أقسام :

الفصل الأول هو لمحة عامة عن التعاريف و الرموز من اجل توضيح و تسهيل المعلومات.
في الفصل الثاني نثبت خصائص(EQMV-Laplace)من اجل النموذج GARCH

,و نظرية الاتساق القوي و التقارب الطبيعي

و أخيرا, في الفصل الثالث نقدم تطبيق على السلاسل الزمنية و مقارنة عددية بين (EQMV-Laplace) و (EQMV-Gaussien).

كلمات مفتاحية :

عملية عشوائية, الضوضاء البيضاء, السكون, السلسلة الزمنية, قانون لابلاس, مقدر الحد الأقصى التقريبي الاحتمالي, نموذج GARCH, الاتساق القوي

, التقارب الطبيعي