REPUBLIQUE ALGERIENNE DEMOCRATIQUE et POPULAIRE.
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU
Faculté des Sciences Département de Mathématiques

⋆⋆⋆

MEMOIRE DE MASTER
en Mathématiques

Spécialité : Analyse Mathématique et Applications

Approximation des estimateurs du modèle GARCH par la méthode du Bootstrap

présenté par

FADLI Ourdia

Devant le jury composé de

⋆⋆⋆

Fazia BEDOUHENE	Professeur	U.M.M.T.O	Présidente
Abdelghani HAMAZ	MCA	U.M.M.T.O	Rapporteur
Farida ACHEMINE	MCA	U.M.M.T.O	Examinatrice

⋆⋆⋆

soutenu le 12 / 10 / 2019

Remerciements

Merci au grand Dieu de m'avoir donné la force, la patience et le courage de mener ce travail jusqu'à la fin.

je tiens, en premier lieu, à remercier mon encadreur Dr. Abdelghani HAMAZ, pour l'orientation, la confiance, la patience qui ont constitué un apport considérable sans lequel ce travail n'aurait pas pu être mené au bon port. Qu'il trouve dans ce travail un hommage vivant à sa haute personnalité.

J'exprime aussi mes précieux sentiments de gratitude et de reconnaissance à tous ceux qui m'ont enseigné, notamment, les cadres enseignants de la faculté de Science, département de Mathématiques, Université Mouloud Mammari.

Je tiens en deuxième lieu à remercier les membres du jury pour avoir accepté d'évaluer ce travail et de participer à la soutenance.

Je tiens encore à dire combien j'apprécie tous ce qui m'a soutenu durant la préparation de ce mémoire, plus particulièrement, ma famille et mes amis.

Mes remerciements vont aussi à tous ceux qui ont collaboré à l'accomplissement de ce travail de prés ou de loin.

Table des matières

Liste des tableaux

Introduction

La très grande fréquence des données financières a favorisé l'étude des impacts de court et de long terme d'une brusque fluctuation sur la dynamique des séries. Elle met en évidence le concept de mémoire longue dans les séries financières et microéconomiques. Pendant de nombreuses années, le modèle de marche aléatoire a été considéré le meilleur prédicteur de la dynamique de la volatilité sur les marchés de change. Ce modèle présume qu'on ne peut pas battre systématiquement le marché : les marchés sont efficients.

L'étude de séries financières révèle une dépendance temporelle du risque qui disparaît souvent progressivement. Les recherches se sont orientées vers l'étude des dynamiques conditionnelles. Toutefois, ces travaux ne différencient pas la dépendance conditionnelle révélant de la moyenne de celle de la variance. Deux volets sont alors issus. Le premier étudie l'équation de la moyenne conditionnelle avec les modèles ARMA, ARIMA, ARFIMA... Le second s'est focalisé sur la variance conditionnelle. Deux classes de modèles non linéaires se sont développées pour caractériser cette variance.

- La première classe se base sur les modèles d'hétéroscédasticité conditionnelle autoregréssive (ARCH) de Engle (1982) et ses extensions,le modèle d'hétéroscédasticité conditionnelle autoregréssif généralisé (GARCH) de Bollerslev (1986) et le modèle GARCH intégré de Engle et de Bollerslev (1986).
- La deuxième classe fait référence aux modèles de volatilité stochastique (SV) introduit par Taylor (1986).

Contrairement aux modèles économétriques standards qui supposent que la variance est toujours constante, les modèles ARCH supposent que la variance conditionnelle est une fonction des carrés des innovations passées. L'application de ces modèles sur les taux d'inflation montre que la volatilité change au cours du temps. Dans tous ces travaux, on suppose que plus l'innovation est lointaine dans le passé plus faible sera son impact sur la variance conditionnelle : la valeur de la volatilité est une fonction décroissante de l'ordre de retard. La spécification linéaire de la variance conditionnelle semble saisir l'effet des erreurs passées ou de mémoire longue trouvé dans les travaux empiriques. Cette spécification ne pose pas de problèmes d'estimation particuliers. Toutefois, pour que la variance conditionnelle soit positive, des contraintes supplémentaires sur les paramètres du modèle sont nécessaires. Pour contourner cette dernière difficulté, Engle (1982, 1983) et Engle et Kraft (1983) proposent de fixer le nombre de retards à prendre en compte au lieu de travailler avec des ARCH infinis. En se basant sur les travaux faits sur l'équation de la moyenne conditionnelle, une extension généralisée des modèles ARCH a été développée : le modèle GARCH. Ce dernier tient compte non seulement de la volatilité courante exprimée par les carrés des résidus passés mais aussi de la volatilité passée, il présente ainsi une spécification plus flexible de la variance conditionnelle.

Toutefois, le modèle GARCH ne permet de modéliser que la dépendance temporelle de court terme : l'effet des chocs décroît exponentiellement dans le temps. Pour étendre la mémoire du modèle GARCH, Engle et Bollerslev (1986) introduisent une extension du modèle GARCH dite intégrée (IGARCH) qui présente une mémoire "explosive" : les effets des chocs sont persistants à l'infini. Compte tenu des propriétés opposées des modèles traditionnels : décroissance exponentielle de l'impact des chocs sur la volatilité des modèles GARCH par opposition leur persistance à l'infinie dans les modèles IGARCH, un troisième modèle plus flexible a été introduit Sur le marché de change, Baillie, Bollerslev et Mikkelsen (1996) ont proposé une version ajustée des deux modèles : c'est le processus GARCH fractionnement intégré (FIGARCH). Ces modèles sont spécifiés par analogie avec les modèles ARFIMA dans l'équation de la moyenne.

Le but de notre travail est de fournir des techniques d'estimation des paramètres du modèle ARCH, GARCH. le plus souvent utilisés dans la modélisation des marchés financiers. A base de l'article "Bootsrap inference for GARCH models by the least absolute deviation estimation" de Qianqian Zhu, Ruochen Zeng and Guodong Li, publié en 2019 dans la revue "Journal of Time series analysis", nous développons en particulier l'estimation des paramètres par la méthode du Bootstrap .

Le travail s'organise comme suit : dans le chapitre 1, nous étudions les modèles ARCH et GARCH en examinant dans le détail les conditions de faible et forte stationnarité. Dans le chapitre 2, nous présentons les méthodes d'estimations des paramètres GARCH en montrant la consistence et la normalité asymptotique. Dans le chapitre 3, nous introduisons l'estimateur qui minimise l'écart absolu moyen (EAM) et nous étudions ces propriétés par les techniques de réechantillonage (Bootstrap). Une application qui concerne le test de Portemanteau sera déduite de cette procédure. Le mémoire sera achevé par une conclusion.

Chapitre 1

Généralités sur les modèles ARCH/GARCH

La modélisation des séries financières se révèle être un problème complexe. Cette complexité tient surtout par l'existence de régularité statistique commune par un très grand nombre de séries financières, de ce fait, celles-ci sont difficiles de reproduire à partir des modèles stochastiques. Dans le but de mettre en pratique les différents fondements théoriques, nous avons élaboré une étude portant sur les modèles ARCH la première pierre de l'édifice de l'économétrie de la volatilité financière fut posée par Robert Engle, (prix Nobel d'économétrie 1982) et GARCH due par Bollerslev (1986).
Ce sont des modèles qui jouent un rôle primordial au niveau de la description des séries financières vu le comportement hétérosédastique de leur variance, la chose qui a toujours été mal considérée par les modèles ARIMA (autorégressif moyenne mobile intégrée) dans lesquels on suppose que la variance est inconditionnelle par rapport au temps. En effet, face aux anomalies des représentations ARMA pour les problèmes monétaires et financiers, Engle (1982)a mis, à la disposition de l'ensemble des acteurs du marché financier, une nouvelle catégorie de modèles autorégressifs conditionnellement hétéroscédastiques aptes à capter le comportement de la volatilité dans le temps. Cette dernière est un paramètre de mesure du risque du rendement et du prix. La volatilité sert également aux calculs pour optimiser la diversification des portefeuilles d'actifs financiers (MEDAF). Les séries monétaires et financières sont caractérisées par le regroupement (clustering) de la volatilité, à savoir les périodes de forte volatilité alternent avec les périodes de faible volatilité. Ce phénomène, que nous appelons aussi l'hétéroscédasticité conditionnelle. Les modèles GARCH standard supposent que les termes d'erreur positifs et négatifs ont un effet symétrique sur la volatilité. Dans la pratique, cette hypothèse est fréquemment violée, en particulier par les

rendements boursiers, dans la mesure où la volatilité augmente davantage après les mauvaises nouvelles que les bonnes nouvelles. Ce soi-disant effet de levier apparaît d'abord dans Black (1976) , qui a noté que : “ Une baisse de la valeur de l'entreprise provoquera un rendement négatif sur son stock et augmentera généralement son effet de levier. [...] Cette augmentation du ratio d'endettement entraînera sûrement une augmentation de la volatilité du titre”. Les rendements négatifs impliquent une plus grande proportion de la dette grâce à une valeur de marché réduite de l'entreprise, ce qui conduit à une volatilité accrue. Le risque, c-à-d. la volatilité, réagit d'abord à des variations plus importantes de la valeur marchande. Néanmoins, il est empiriquement démontré qu'il existe une volatilité élevée après des modifications moins importantes. De son côté, Black n'a rien dit sur l'effet des rendements positifs sur la volatilité. Bien que les rendements positifs entraînent des augmentations moins importantes, ils entraînent une augmentation de la volatilité. D'un point de vue empirique, la volatilité réagit de manière asymétrique au signe des chocs et, par conséquent, un certain nombre d'extensions paramétrées du modèle GARCH standard ont été suggérées récemment (IGARCH, EGARCH, TGARCH..) La famille des modèles ARCH peut se décomposer en deux sous-ensembles :

1. Les modèles ARCH linéaires et les modèles ARCH non linéaires : Reposent sur une spécification quadratique de la variance conditionnelle des perturbations : Modèles ARCH, GARCH et IGARCH (intégrée qui présente une mémoire explosive telle que les effets des chocs sont persistants à l'infini).
2. Les modèles ARCH non linéaires sont caractérisés par des spécifications asymétriques des perturbations. Ce sont les modèles EGARCH(Expontiels), TARCH (Threshold) et TGARCH.

1.1 Modèle ARCH

La nouvelle classe de modèles (ARCH) apte à capter le comportement de la volatilité dans le temps qui est composé de deux équations, la première met en relation le rendement et certaines variables qui l'expliquent et la seconde modélise la variance conditionnelle des résidus.
Ce modèle consiste à introduire une dynamique dans la détermination de la volatilité en supposant que la variance est conditionnelle aux informations dont nous disposons. Il avance une spécification ARCH(q) où le carré des innovations, c'est-à-dire la variance du terme d'erreur au temps t, dépend de l'importance des termes d'erreur au carré des q

périodes passées. Le modèle ARCH(q) permet de générer des épisodes de volatilité importante suivis d'épisodes de volatilité plus faibles.
voilà la première representation telle qu'il à introduit par Engel (1982) :

$$Y_t = \beta X_t + \epsilon_t, \qquad (*)$$

avec

$$\epsilon_t / I_{t-1} \sim \mathcal{N}(0, h_t),$$

où Y_t corespond au variable expliquant les rendements, il peut être un modèle ARMA(p, q) et $I_{t-s} = \sigma(X_{t-s})_{t \leq s}$ est la tribu engendrée par les X_{t-s} $s < t$.
Dans la modélisation ARCH, le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ peut s'écrire sous la forme :

$$\epsilon_t = z_t h_t,$$

avec

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2},$$

où $\{z_t\}$ suite de variables aléatoire $i.i.d$, avec $\alpha_0 > 0$ et $\alpha_i \geq 0$.

1.1.1 Modèle ARCH(1) et ces propiétés

Modèle ARCH(1)

Définition 1.1. Le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ est un ARCH(1) si :

$$\epsilon_t = z_t h_t, \tag{1.1}$$

avec

$$h_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2},$$

$\{z_t\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées $i.i.d.$ d'ésperance nulle et de variance σ^2 avec $\alpha_0 > 0$ et $\alpha_1 \geq 0$.

La composante h_t désigne une variable qui, conditionnellement à l'ensemble d'information des valeurs passées de ϵ_t, $I_{t-1} = \{\epsilon_{t-1}, \epsilon_{t-2}, ..., \epsilon_{t-j}, ...\}$ est déterministe et positive.

La variance conditionnel :

$$\begin{aligned} \mathbb{V}(\epsilon_t/I_{t-1}) &= \mathbb{V}(z_t h_t/I_t) \\ &= h_t^2 \mathbb{V}(z_t/I_t) \\ &= h_t^2 \sigma_z^2. \end{aligned}$$

$\Rightarrow h_t^2$ est la variance conditionnelle de ϵ_t.
Dans ce modèle, le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ est caractérisé par des autocorrélations nulles $\mathbb{E}(\epsilon_t \epsilon_s) = 0$ pour $t \neq s$ ce qui signifie que les ϵ_t sont non corrélées dans le temps. En effet, $\{\epsilon_t; t \in \mathbb{Z}\}$ reste un bruit blanc mais faible.
Des résultats intéressantes sont établis, en considérant le processus autorégressif sur les ϵ_t^2. Pour simplifier, on se limite au cas du ARCH(1). Dans ces conditions :

$$h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 \quad \Leftrightarrow \quad \epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + (\epsilon_t^2 - h_t^2). \tag{1.2}$$

Soit encore :

$$\begin{aligned} \epsilon_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \mu_t, \\ \epsilon_t^2 &= h_t^2 + \mu_t, \end{aligned}$$

où $\mu_t = (\epsilon_t^2 - h_t^2)$ et en ayant les informations disponibles jusqu'au temps $t-1$:

$$\mathbb{E}(\mu_t/I_{t-1}) = 0,$$

avec I_{t-1} c'est l'ensemble de l'information jusqu'à $t-1$, en effet :

$$\begin{aligned} \mathbb{E}(\mu_t/I_{t-1}) &= \mathbb{E}(\epsilon_t^2/I_{t-1}) - \mathbb{E}(h_t^2/I_{t-1}) \\ &= \mathbb{V}(\epsilon_t/I_{t-1}) - \mathbb{E}(h_t^2/I_{t-1}) \\ &= h_t^2 - h_t^2 \\ &= 0. \end{aligned}$$

$\{\mu_t; t \in \mathbb{Z}\}$ est un processus d'innovation pour ϵ_t^2. Ainsi, cette écriture précédente correspond à celle d'un processus $AR(1)$ sur le carré ϵ_t^2.

$$\begin{aligned} \epsilon_t^2 &= h_t^2 + \mu_t \\ \epsilon_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \mu_t. \end{aligned} \tag{1.3}$$

On sait que ce processus $\{\epsilon_t^2; t \in \mathbb{Z}\}$ est stationnaire au second ordre si et seulement si $|\alpha_1| < 1$. Cela s'énonce clairement par le théorème suivant.

Théorème 1.1. *Si un processus $\{\epsilon_t; t \in \mathbb{Z}\}$ est un ARCH(1), alors $\{\epsilon_t^2; t \in \mathbb{Z}\}$ est un AR(1) telle que :*

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \mu_t,$$

où $\mu_t = \epsilon_t^2 - h_t^2$ vérifiant $\mathbb{E}(\mu_t / I_{t-1}) = 0$ est un processus d'innovation pour $\{\epsilon_t^2\}$.

Propiétés de ARCH(1)

Propriétés 1.1.

$$\forall s \geq 1, \quad \text{on a} \quad \mathbb{E}(\epsilon_t / I_{t-s}) = 0. \tag{1.4}$$

Preuve. On utilisant la propriété des espérances itérées, on a :

$$\begin{aligned} \mathbb{E}(\epsilon_t / I_{t-s}) &= \mathbb{E}(\mathbb{E}((\epsilon_t / I_{t-1}) / I_{t-s})) \\ &= \mathbb{E}(0 / I_{t-s}) \\ &= 0. \end{aligned}$$

Car $I_{t-1} \subset t-1\,; \forall s > 1$.

Propriétés 1.2. *La variance conditionnelle du processus $\{\epsilon_t; t \in \mathbb{Z}\}$ satisfaisant une représentation ARCH(1), où $\alpha_0 > 0$, et $0 < \alpha_1 < 1$ définie par l'équation $\epsilon_t = z_t h_t$ est non constante dans le temps et vérifié :*

$$\mathbb{V}(\epsilon_t / I_{t-s}) = \alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \epsilon_{t-s}^2 \quad \forall t \in \mathbb{Z}. \tag{1.5}$$

C'est la propriété centrale des processus ARCH, le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ a une variance conditionnelle qui dépend du temps.

Preuve. On sait que $\mathbb{E}(\epsilon_t / I_{t-s}) = 0$ dès lors, $\mathbb{V}(\epsilon_t / I_{t-s}) = \mathbb{E}(\epsilon_t^2 / I_{t-s})$, en utilisant la la relation où $\{\mu_t; t \in \mathbb{Z}\}$ est un bruit faible.

par itération successive, on a :

$$\begin{aligned}
\epsilon_t^2 &= \alpha_0 + \alpha_1(\alpha_0 + \alpha_1\epsilon_{t-2}^2 + \mu_{t-1}) + \mu_t \\
&= \alpha_0 + \alpha_1\alpha_0 + \alpha_1^2\epsilon_{t-2}^2 + \alpha_1\mu_{t-1} + \mu_t \\
&= \alpha_0 + \alpha_1\alpha_0 + \alpha_1^2(\alpha_0 + \alpha_1\epsilon_{t-3}^2 + \mu_{t-2}) + \alpha_1\mu_{t-1} + \mu_t \\
&= \alpha_0 + \alpha_1\alpha_0 + \alpha_1^2\alpha_0 + \alpha_1^3\epsilon_{t-3}^2 + \alpha_1^2\mu_{t-2} + \alpha_1\mu_{t-1} + \mu_t \\
&= \alpha_0 + \alpha_1\alpha_0 + \alpha_1^2\alpha_0 + \alpha_1^3\alpha_0 + ... + \alpha_1^{s-1}\alpha_0 + \mu_t + \alpha_1\mu_{t-1} + \alpha_1^2\mu_{t-2} + ... + \alpha_1^{s-1}\mu_{t-s+1} + \alpha_1^s\epsilon_{t-s}^2.
\end{aligned}$$

On aura donc :

$$\begin{aligned}
\epsilon_t^2 &= \alpha_0[1 + \alpha_1 + \alpha_1^2 + ... + \alpha_1^{s-1}] + \mu_t + \alpha_1\mu_{t-1} + \alpha_1^2\mu_{t-2} + ... + \alpha_1^{s-1}\mu_{t-s+1} + \alpha_1^s\epsilon_{t-1}^2 \\
&= \alpha_0\left[\frac{1-\alpha_1^s}{1-\alpha_1}\right] + \sum_{j=0}^{s-1}\alpha_1^j\mu_{t-j} + \alpha_1^s\epsilon_{t-s}^2.
\end{aligned}$$

En calculant l'éspérance conditionombelle de chacun de ses nombres :

$$\mathbb{E}(\epsilon_t^2/I_{t-s}) = \alpha_0\left[\frac{1-\alpha_1^s}{1-\alpha_1}\right] + \sum_{j=0}^{s-1}\alpha_1^j\mathbb{E}(\mu_{t-j}/I_{t-s}) + \alpha_1^s\mathbb{E}(\epsilon_{t-s}^2/I_{t-s}).$$

Par définition du bruit blanc, on a :
$\mathbb{E}(\mu_{t-j}/I_{t-s}) = 0, \quad \forall j = 0, 1, ..., s-1$, et par définition $\mathbb{E}(\epsilon_{t-s}^2/I_{t-s}) = \epsilon_{t-s}^2$. On obtient ainsi la formule de la variance :

$$\mathbb{V}(\epsilon_t/I_{t-s}) = \alpha_0\left[\frac{1-\alpha_1^s}{1-\alpha_1}\right] + \alpha_1^s\epsilon_{t-s}^2 \quad \forall t \in \mathbb{Z}.$$

Lorsque s tend vers l'infini, ces variances conditionnelles convergent vers la variance non conditionnelle, et l'on retrouve alors la formule :

$$\begin{aligned}
\mathbb{V}(\epsilon_t) &= \lim_{s\to\infty}\mathbb{V}(\epsilon_t/I_{t-s}) \\
&= \lim_{s\to\infty}(\alpha_0\left[\frac{1-\alpha_1^s}{1-\alpha_1}\right] + \alpha_1^s\epsilon_{t-s}^2) \\
&= \frac{\alpha_0}{1-\alpha_1}.
\end{aligned}$$

car $\alpha_0 > 0$, et $0 < \alpha_1 < 1$. Le résultat selon lequel la variance non conditionnelle est définie par la relation :

$$\mathbb{V}(\epsilon_t) = \frac{\alpha_0}{1-\alpha_1},$$

qui explique les contraintes sur les paramètres de la représentation ARCH, $\alpha_0 > 0$ et $0 < \alpha_1 < 1$.

Propriétés 1.3. *Les autocovariances conditionnelles du processus $\{\epsilon_t; t \in \mathbb{Z}\}$, ARCH(1), définies par l'équation (1.1) sont nulles :*

$$cov(\epsilon_t, \epsilon_{t+k}/I_{t-s}) = 0 \qquad \forall ks \geq 1. \tag{1.6}$$

Le processus est donc un processus sans mémoire conditionnellement à I_{t-s} $\forall k, s \geq 1$.

Preuve. Cette propriété s'obtient de la façon suivante :

$$\begin{aligned} cov(\epsilon_t, \epsilon_{t+k}/I_{t-s}) &= \mathbb{E}[\epsilon_t, \epsilon_{t+k}/I_{t-s}]\mathbb{E}[\epsilon_t/I_{t-s}]\mathbb{E}[\epsilon_{t+k}/I_{t-s}] \\ &= \mathbb{E}[\epsilon_t, \epsilon_{t+k}/I_{t-s}] \\ &= \mathbb{E}[\mathbb{E}(\epsilon_t\epsilon_{t+k}/I_{t+k-1})/I_{t-s}] \\ &= \mathbb{E}[\epsilon_t\mathbb{E}(\epsilon_{t+k}/I_{t+k-1})/I_{t-s}] \\ &= \mathbb{E}[\epsilon_t.0/I_{t-s}] \\ &= 0. \end{aligned}$$

L'absence de corrélations entre les valeurs d'un processus ARCH est une caractéristique très importante de cette famille de modèle, qui les rend utiles pour modéliser certaines séries financières.

Propriétés 1.4. **(i)** *Le moment conditionnel centré d'ordre quatre du processus $\{\epsilon_t; t \in \mathbb{Z}\}$ vérifie*

$$\mathbb{E}(\epsilon_t^4/I_{t-1}) = 3(\alpha_0 + \alpha_1\epsilon_{t-1}^2)^2.$$

(ii) *Sous l'hypothèse $3\alpha_1^2 < 1$, le moment non conditionnel centré d'ordre quatre du processus $\{\epsilon_t; t \in \mathbb{Z}\}$ est égal à :*

$$\begin{aligned} \mathbb{E}(\epsilon_t^4) &= 3\left[\alpha_1^2 + \frac{2\alpha_1\alpha_0^2}{1-\alpha_1} + \alpha_1^2\mathbb{E}(\epsilon_{t-1}^4)\right] \\ &= 3\frac{\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1^2)(1-\alpha_1)}. \end{aligned}$$

(iii) *La kurtosis non conditionnelle associée au processus ARCH(1) est :*

$$k_u = \frac{\mathbb{E}(\epsilon_t^4)}{\mathbb{E}^2(\epsilon_t^2)} = 3\left[\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right] > 3.$$

Preuve. (i) On rappelle que si une variable aléatoire centrée X suit une loi normale centrée, alors :

$$\mathbb{E}(X^4) = 3(\mathbb{V}(X))^2 = (\mathbb{E}(X^2))^2$$

$$\begin{aligned}
\mathbb{E}(\epsilon_t^4/I_{t-1}) &= 3(\mathbb{E}(\epsilon_t^4/I_{t-1})^2 \\
&= 3(\alpha_0 + \alpha_1 \epsilon_{t-1}^2)^2.
\end{aligned}$$

(ii) Sous l'hypothèse $3\alpha_1^2 < 1$, on a :

$$\begin{aligned}
\mathbb{E}(\epsilon_t^4) &= \mathbb{E}[\mathbb{E}(\epsilon_t^4/I_{t-1})] \\
&= \mathbb{E}[3(\alpha_0 + \alpha_0 + \alpha_1 \epsilon_{t-1}^2)^2] \\
&= \mathbb{E}[\alpha_0^2 + 2\alpha_0\alpha_1\alpha_0\epsilon_{t-1}^2 + \alpha_1^2\epsilon_{t-1}^4] \\
&= 3[\alpha_0^2 + 2\alpha_0\alpha_1\mathbb{E}(\epsilon_{t-1}^2) + \alpha_1^2\mathbb{E}(\epsilon_{t-1}^4) \\
&= 3[\alpha_0^2 + 2\alpha_0\alpha_1\frac{\alpha_1}{1-\alpha_1} + \alpha_1^2\mathbb{E}(\epsilon_{t-1}^4)] \\
&= 3\frac{\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1^2)(1-\alpha_1)}.
\end{aligned}$$

(iii) La kurtosis non conditionnelle associée au processus ARCH(1) est définie par :

$$k_u = \frac{\mathbb{E}(\epsilon_t^4)}{\mathbb{E}^2(\epsilon_t^2)}.$$

Et d'après les résultats obtenus précédemment, on obtient :

$$\begin{aligned}
k_u &= \frac{3\alpha_0^2(1-\alpha_1)}{(1-3\alpha_1)(1-\alpha_1)}\frac{(1-\alpha_1)^2}{\alpha_0^2} \\
&= 3\frac{1-\alpha_1^2}{(1-3\alpha_1^2)} > 3.
\end{aligned}$$

Toutes ces propriétés peuvent être généralisées dans le cas d'un processus ARCH(q). En effet, dans la section qui suit, nous montrons la représentation d'un modèle ARCH(q) et ces propriétés.

1.1.2 Modèles ARCH(q) et modèle à erreur ARCH(q)

Modèles ARCH(q)

Définition 1.2. Un processus $\{\epsilon_t; t \in \mathbb{Z}\}$ satisfait une représentation ARCH(q) Si :
$\epsilon_t = z_t h_t, \alpha_0 > 0$, et $\alpha_i \geq 0$ avec

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2},$$

où $\{z_t; t \in \mathbb{Z}\}$ désigne un bruit blanc faible, tel que $\mathbb{E}(z_t) = 0$, et $\mathbb{V}(z_t) = \sigma_z^2$.
Pour ce type de processus, on retrouve les deux propriétés essentielles vues précédement, à savoir la propiété de différence de martingale (ou bruit blanc faible) $\mathbb{E}(\epsilon_t / I_{t-s}) = 0$, et la propriété de variance conditionnelle variable dans le temps puisque :

$$\mathbb{V}(\epsilon_t / I_{t-s}) = h_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2,$$

de plus

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-i}^2 + \mu_t.$$
$$\mathbb{E}(\mu_t / I_{t-1}) = 0.$$

Modèle à erreur ARCH(q)

Définition 1.3. On considère dorénavant le résidu d'un modèle linéaire. Prenons l'exemple d'un modèle linéaire autorégressif d'ordre p avec résidus de type $ARCH(q)$:
$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \epsilon_t$, où $\epsilon_t = z_t h_t, \quad \alpha_0 > 0$, et $\alpha_1 \geq 0$ avec

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_0 \epsilon_{t-i}^2},$$

où $\{z_t; t \in \mathbb{Z}\}$ est une suite de variables aléatoires $i.i.d$(0,1). On a donc un modèle qui décrit à la fois l'évolution de l'espérance conditionnelle et de la variance conditionnelle du processus $\{\epsilon_t; t \in \mathbb{Z}\}$ dans le temps. Envisageons le cas le plus simple d'un processus de type AR(1) à erreur ARCH(1) :

$$X_t = \delta + \alpha X_{t-1} + \epsilon_t, \qquad |\alpha| < 1.$$
$$\epsilon_t = z_t \sqrt{\alpha_0 + \alpha_0 \epsilon_{t-1}^2}.$$

Dans ce cas, les résidus satisfont les propriétés principales étudiées précédemment :

(i) Propriété de différence martingale : $\mathbb{E}(\epsilon_t/I_{t-1}) = 0$, et de façon générale

$$\mathbb{E}(\epsilon_t/I_{t-s}) = 0, \quad \forall s \geq 1. \tag{1.7}$$

(ii) Variance conditionnelle dépendante du temps

$$\mathbb{V}(\epsilon_t/I_{t-s}) = \alpha_0 \left[\frac{1-\alpha_1^s}{1-\alpha_1}\right] + 1 - \alpha_1^s \epsilon_{t-1}.$$

$$\mathbb{V}(\epsilon_t) = \frac{\alpha_0}{1-\alpha_1}. \tag{1.8}$$

(iii) La propriété d'orthogonalité implique que les corrélations conditionnelles sont nulles :

$$cov(\epsilon_t \epsilon_{t+k}/I_{t-s}) = 0, \quad \forall k, s \geq 1.$$

(iiii) Sous l'hypothèse $3\alpha_1^2 < 1$, la distribution des résidus est leptokurtique puisque

$$k_u = 3\frac{1-\alpha_1^2}{(1-3\alpha_1^2)} > 3.$$

Il y'a donc une absence de corrélation entre les valeurs présentes et futures du processus, quels que soient les retards s et k. Mais si la variance conditionnelle de ϵ_t n'est pas constante, la variance non conditionnelle est constante.
On peut, en outre, en déduire un certain nombre de conclusions quant aux processus $\{X_t; t \in \mathbb{Z}\}$ lui même. On peut montrer tout d'abord, que l'espérance conditionnelle de X_t vérifie :

$$\mathbb{E}(\epsilon_t/I_{t-s}) = \delta + \alpha \mathbb{E}(X_{t-1}/I_{t-s}), \quad \forall s \geq 1.$$

Ce qui montre que les prévisions non linéaires de X_t s'obtiennent comme les prévisions linéaires d'un processus $AR(1)$. Plus généralement :

$$X_t = \delta\frac{1-\alpha_1^s}{1-\alpha_0} + \alpha^s X_{t-s} + \alpha\epsilon_{t-1} + ... + \alpha^{s-1}\epsilon_{t-s+1},$$

En effet :

$$\begin{array}{rcl}
X_t & = & \delta + \alpha X_{t-1} + \epsilon_t \\
& = & \delta + \alpha(\delta + \alpha X_{t-2} + \epsilon_{t-1}) + \epsilon_t \\
& = & \delta + (1+\alpha) + \alpha^2 X_{t-2} + \alpha\epsilon_{t-1} + \epsilon_t \\
& = & \delta(1+\alpha+\alpha^2) + \alpha^3 X_{t-3} + \alpha^2\epsilon_{t-2} + \alpha_{t-1} + \epsilon_t \\
& = & \cdot \\
& = & \cdot \\
& = & \cdot \\
& = & \delta(1+\alpha+\alpha^2+...+\alpha^{s-1}) + \alpha^s X_{t-s} + \alpha_{t-1}\epsilon_{t-s+1} + ... + \alpha\epsilon_{t-1} + \epsilon_t \\
& = & \delta\dfrac{1+\alpha^s}{1-\alpha} + \alpha^s X_{t-s} + \alpha^{s-1}\epsilon_{t-s+1} + ... + \alpha\epsilon_{t-1} + \epsilon_t.
\end{array}$$

En prenant l'espérance conditionnelle de deux cotés, on obtient :

$$\mathbb{E}(X_t/I_{t-s}) = \delta\frac{1-\alpha^s}{1-\alpha} + \alpha^s X_{t-s}.$$

De même façon, on peut montrer que la variance conditionnelle de X_t dépend du temps. En effet, on montre quelle dépend du processus $\{\epsilon_{t-s}^2; t \in \mathbb{Z}\}$ de la façon suivante.

Propriétés 1.5. *La variance conditionnelle du processus AR(1) à erreur ARCH(1), X_t, s'écrit :*

$$\mathbb{V}(X_t/I_{t-s}) = \frac{\delta}{1-\alpha_1}\left[\frac{1-\alpha^{2s}}{1-\alpha^2}\right] - \alpha_1\left(\frac{\alpha_1^s-\alpha^{2s}}{\alpha_1-\alpha^2}\right) + \alpha_1\left[\frac{\alpha_1^s-\alpha^{2s}}{\alpha_1-\alpha^2}\right]\epsilon_{t-1}^2. \tag{1.9}$$

Ainsi, la variance d'une erreur de prévision à l'horizon 1, s'écrit :

$$\mathbb{V}(X_t/I_{t-1}) = \delta + \alpha_1\epsilon_{t-1}^2. \tag{1.10}$$

Preuve.

$$\begin{array}{rcl}
\mathbb{V}(X_t/I_{t-1}) & = & \mathbb{V}\left(\dfrac{1-\alpha^s}{1-\alpha} + \alpha^s X_{t-s} + \epsilon_t + \alpha\epsilon_{t-1} + ... + \alpha^{s-1}\epsilon_{t-s+1}/I_{t-s}\right) \\
& = & \mathbb{V}\left(\epsilon_t/I_{t-s} + \alpha^2\mathbb{V}(\epsilon_{t-s}/I_{t-1}) + ... + \alpha^{2(s-1)}\mathbb{V}(\epsilon_{t-s+1}/I_{t-s}\right) \\
& = & \displaystyle\sum_{j=0}^{s-1}\alpha^{2j}\left[\delta\frac{1-\alpha_1^{s-j}}{1-\alpha_1} + \alpha_1^{s-j}\epsilon_{t-s}^2\right] \\
& = & \dfrac{\delta}{1-\alpha_1}\left[\dfrac{1-\alpha^{2s}}{1-\alpha^2} - \alpha_1\dfrac{\alpha_1^s-\alpha^{2s}}{\alpha_1-\alpha^2}\right] + \alpha_1\left[\dfrac{\alpha_1^s-\alpha^{2s}}{\alpha_1-\alpha^2}\right]\epsilon_{t-1}^2.
\end{array}$$

En conclusion, si l'on désire prévoir le processus $\{X_t; t \in \mathbb{Z}\}$ dans le cas d'erreur ARCH(1), l'erreur de prévision a un horizon d'une période admet une variance $\mathbb{V}(X_t/I_{t-1})$ qui varie dans le temps en fonction de la valeur de ϵ^2_{t-s}.

Remarque 1.1. La variance d'une erreur de prévision sur un processus à erreur ARCH dépend du temps.

$$\mathbb{V}(X_t/I_{t-s}) = \mathcal{F}(\epsilon_{t-s}), \quad s \geq 1. \tag{1.11}$$

L'amplitude des intervalles de confiance associée à cette prévision n'est donc pas constante dans le temps.

1.2 Modèle GARCH

Le modèle d'Hétéroscédastique Conditionnelle Autorégressive Généralisée (GARCH), proposé par Bollerslev (1986), a obtenu un énorme succès dans la modélisation de la volatilité.
Sa caractérisation repose essentiellement sur le concept de variance conditionnelle. Dans ces modèles, celle-ci s'écrit comme une fonction affine des valeurs passées du carré de la série. Cette spécification particulière, se révèle très fructueuse car elle permet une étude complète des propriétés des solutions tout en étant assez générale.

1.2.1 Modèle GARCH(p,q) faible

Définition 1.4. On dit que $\{\epsilon_t; t \in \mathbb{Z}\}$ est un processus GARCH(p,q) faible si ses deux premiers moments conditionnels existent et vérifient :

(i) $\mathbb{E}(\epsilon_t/I_{t-s}) = 0, \qquad \forall s \geq 1.$

(ii) il existe des constantes $\alpha_0 > 0, i = 1, ..., q$ telle que

$$\sigma^2 = \mathbb{V}(\epsilon_t/I_{t-s}) \;\; = \;\; \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon^2_{t-1} + \sum_{j=1}^{p} \beta_j h^2_{t-j}. \tag{1.12}$$

L'équation (1.12) peut être écrite de manière symbolique sous la forme plus compacte

$$h^2_t \;\; = \;\; \alpha_0 + \alpha(L)\epsilon^2_t + \beta(L)h^2_t, \qquad t \in \mathbb{Z}, \tag{1.13}$$

où L est l'opérateur retard ($L^i\epsilon^2_t = \epsilon^2_{t-1}$ et $L^i h^2_t = h_{t-1}$) pour tout entier i , α et β sont des polynomes de degrés q et p :

$$\begin{aligned}\alpha(L) &= \alpha_1 L + \alpha_2 L^2 + ... + \alpha_q L^q = \sum_{i=0}^{q} \alpha_1 L^i \\ \beta(L) &= \beta_1 L + \beta_2 L^2 + ... + \beta_p L^p = \sum_{j=1}^{p} \beta_j L^j.\end{aligned}$$

L'innovation du processus $\{\epsilon_t^2; t \in \mathbb{Z}\}$ est par définition la variable $\mu_t = \epsilon_t^2 - h_t^2$. En remplaçant dans l'équation (1.12), les variables h_t^2 par $\epsilon_{t-j}^2 - \mu_{t-j}$, obtient la représentation :

$$\begin{aligned}\epsilon_t^2 - \mu_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j (\epsilon_{t-j}^2 - \mu_{t-j}) \\ \epsilon_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j \epsilon_{t-j}^2 - \sum_{j=1}^{p} \beta_j \mu_{t-1} + \mu_t \\ \epsilon_t^2 &= \sum_{i=1}^{r} (\alpha_t + \beta_t) \epsilon_{t-1}^2 - \sum_{j=1}^{p} \beta_j \mu_{t-j} + \mu_t,\ t \in \mathbb{Z},\end{aligned}$$

où $r = \max(p, q)$ avec la convention $\alpha_t = 0$ (resp $\beta_j = 0$) si $i > q$(resp $j > p$.) La définition 1 ne fournit pas directement de processus la vérifiant. La définition plus restrictive suivante permettra d'obtenir explicitement des processus solutions.

1.2.2 Modèle GARCH(p,q) fort

Définition 1.5. Soit $\{z_t; t \in \mathbb{Z}\}$ une suite de variables aléatoire *i.i.d.*
On dit que $\{\epsilon_t; t \in \mathbb{Z}\}$ est un processus GARCH(p,q) au sens fort (relativement à la suite z_t) s'il vérifie :

$$\begin{aligned}\epsilon_t &= z_t h_t \\ h_t^2 &= \alpha_o + \sum_{t=1}^{q} \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2,\end{aligned} \tag{1.14}$$

où les α_i et α_j sont des constantes positives et α_0 est une constante strictement positive. Il est clair qu'un processus GARCH fort tel que h_t^2 est mesurable par rapport à la tribu I_{t-1} est un processus GARCH faible. La réciproque n'est cependant pas vraie.

1.3 Etude de la stationnarité

Nous allons chercher sous quelles conditions il existe des processus stationnaires (au sens strict et au second-ordre) vérifiant les définitions 1.3 et 1.4. On s'intéresse plus particulièrement aux solutions non anticipatives du modèle (1.14) c'est-à-dire aux processus $\{\epsilon_t; t \in \mathbb{Z}\}$ tel que ϵ_t soit une fonction mesurable des variables z_{t-s}, $s \geq 0$. Nous examinons d'abord le cas du modèle GARCH(1,1) qui peut se traiter avec des techniques élémentaires. On notera, pour $x > 0$, $\log^+ x = \max(logx, 0)$.

Modèle GARCH(1,1)

Définition 1.6. Dans le cas où $p = q = 0$, le modèle (1,14) s'écrit :

$$\begin{aligned} \epsilon_t &= z_t h_t \\ h_t^2 &= \alpha_o + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}^2 < 0, \end{aligned} \tag{1.15}$$

avec $\alpha_0 > 0; \alpha \geq 0$ et $\beta \geq 0$. On pose $\alpha(y) = \alpha y^2 + \beta\alpha_0$.

1.3.1 La stastionarité forte du GARCH(1,1)

Théorème 1.2. *Le modèle GARCH(1,1) est stationnaire strictement :*
Si

$$-\infty \leq \gamma = \mathbb{E}\{\log\{\alpha z_t^2 + \beta\}\} < 0, \tag{1.16}$$

alors la série

$$\mu_t = \{1 + \sum_{t=1}^{\infty} \alpha(z_{t-1}...\alpha(z_{t-i})\}\alpha_0, \tag{1.17}$$

converge presque sûrement (p.s) et le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ défini par $\epsilon_t = \sqrt{\eta_t} z_t$ est l'unique solution strictement stationnaire du modèle (1.15). Cette solution est non anticipative et ergodique.
Si $\gamma \geq 0$ et $\alpha_0 > 0$, il n'existe pas de solution strictement stationnaire.

Preuve. On a : $\epsilon_t = z_t h_t$ et $h_t^2 = \alpha_0 + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}^2$.

Utilisant de manière itérative cette équation, on obtient, pour $N \geq 1$:

$$\begin{aligned}
h_t^2 &= \alpha_0 + \alpha z_{t-1}^2 h_{t-1}^2 + \beta h_{t-1}^2 \\
&= \alpha_0 + (\alpha z_{t-1}^2 + \beta h_{t-1}^2) h_{t-1}^2 \\
&= \alpha_0 + \alpha(z_{t-1}) h_{t-1}^2 \\
&= \alpha_0 + \alpha(z_{t-1})[\alpha_0 + (\alpha z_{t-2}) h_{t-2}^2] \\
&= \alpha_0 + \alpha_0 \alpha(z_{t-1}) + \alpha(z_{t-1})\alpha(z_{t-2}) h_{t-2}^2 \\
&= \alpha_0 + \alpha_0 \alpha(z_{t-1}) + \alpha(z_{t-1})\alpha(z_{t-2})[\alpha_0 + \alpha(z_{t-3}) h_{t-3}^2] \\
&= \alpha_0 + \alpha_0 \alpha(z_{t-1}) + \alpha_0 \alpha(z_{t-1})\alpha(z_{t-2}) + \alpha(z_{t-1})\alpha(z_{t-2})\alpha(z_{t-3}) h_{t-3}^2 \\
&= \vdots \\
&= \alpha_0[1 + \alpha(z_{t-1}) + \alpha(z_{t-1})\alpha(z_{t-2}) + ... + \alpha(z_{t-1})\alpha(z_{t-2})\alpha(z_{t-3})...\alpha(z_{t-N})] \\
&+ \alpha(z_{t-1})\alpha(z_{t-2}) + ... + \alpha(z_{t-N-1}) h_{t-N-1}^2 \\
&= \alpha_0 \left[1 + \sum_{t=1}^{N} \alpha(z_{t-1})...\alpha(z_{t-i}) \right] + \alpha(z_{t-1})...\alpha(z_{t-N-1}) h_{t-N-1}^2 \\
&= \eta_t(N) + \alpha(z_{t-1})...\alpha(z_{t-N-1})^2 h_{t-N-1}. \qquad (1.18)
\end{aligned}$$

Le processus limite $\eta_t = \lim_{N \to +\infty} \eta_t(N)$ existe dans $\overline{\mathbb{R}}_+ = [0, \infty]$ puisque les termes de la somme sont positifs.
De plus, en faisant tendre N vers l'infini dans la relation $\eta_t(N) = \alpha_0 + \alpha(z_{t-1})\eta_{t-1}(N-1)$, on obtient :

$$\eta_t = \alpha_0 + \alpha(z_{t-1})\eta_{t-1}.$$

Nous allons montrer que η_t est presque sûrement finie si et seulement si $\gamma < 0$. Supposons $\gamma < 0$, On utilise la régle de Cauchy pour les séries à termes positifs. On a :

$$\begin{aligned}
[(z_{t-1})...(z_{t-n})]^{\frac{1}{n}} &= \exp[\frac{1}{n} \log\{(z_{t-1})...(z_{t-n})\}] \\
&= \exp[\frac{1}{n} \sum_{i=1}^{n} \log \alpha(z_{t-i})] \to e^{\gamma} \qquad p.s, \qquad (1.19)
\end{aligned}$$

quand $n \to 1$, par application de la loi forte des grands nombres à la suite $i.i.d\{\log \alpha(z_t)\}$. La série définie en (1.17) converge alors presque sûrement dans $\mathbb{R}$. Par suite, le processus $\{\epsilon; t \in \mathbb{Z}\}$ défini par :

$$\epsilon_t = \sqrt{\eta_t} z_t = \{\alpha_0 + \sum_{t=1}^{\infty} \alpha(z_{t-1})...\alpha(z_{t-i})\alpha_0\}^{\frac{1}{2}} z_t, \qquad (1.20)$$

est strictement stationnaire et ergodique et non anticipative (car comme (1.17) fonction mesurable des variables $z_{t-1},\ i \geq 0$. de plus $\epsilon_t = h_t z_t$.
Nous montrons maintenant l'unicité. Soit $\tilde{\epsilon}_t = \sigma_t z_t$ une autre solution strictement stationnaire. D'après (1.18) on a

$$h_t^2 = \eta_t(N) + \alpha(z_{t-1})...(z_{t-N-1})h_{t-N-1}^2,$$

par suite

$$h_t^2 - \eta_t = \{\eta_t(N) - \eta_t\} + \alpha(z_{t-1})...\alpha(z_{t-N-1})h_{t-N-1}^2.$$

Le terme entre accolades à droite de l'égalité tend vers 0 p.s quand $N \to \infty$ Par ailleurs, puisque la série définissant η_t converge p.s, on a $\alpha(z_{t-1})...\alpha(z_{t-n}) \to 0$ avec probabilité 1 quand $n \to \infty$. De plus la loi de h_{t-N-1}^2 est indépendante de N par stationnarité. Par suite $\alpha(z_{t-1})...\alpha(z_{t-N-1})h_{t-N-1} \to 0$ en probabilité lorsque $N \to 1$. De plus, h_{t-N-1}^2 est indépendant de N et donc on a $h_t^2 = \eta_t$ pour tout t, p.s.
Si $\gamma > 0$, par application de la régle de cauchy, $\sum_{n=1}^{N} \alpha(z_{t-1})...\alpha(z_{t-1}) \to 0$ p.s. Lorsque $N \to 1$. Donc si $\alpha_0 > 0, \eta_t = +\infty$, p.s. D'après (1.18), il est clair que $h_t^2 = +\infty$, p.s. Par suite, il n'existe pas de solution finie p.s. pour (1.15). Dans le cas $\gamma = 0$, nous procéderons par l'absurde.
Supposons qu'il existe une solution strictement stationnaire (ϵ_t, h_t^2) de (1.15). Nous avons pour $n > 0$,

$$h_0^2 \geq \alpha_0\{1 + \sum_{i=1}^{n} \alpha(z-1)...\alpha(z-i)\},$$

d'où on déduit que le terme général $\alpha(z-1)...\alpha(z-i)\alpha$ converge vers 0 , p.s quand $n \to \infty$.

Remarque 1.2. **(a)** Le coefficient $\gamma = \mathbb{E}[\log\{\alpha(z_t)\}]$ existe toujours dans $[-\infty, +\infty]$ car $\mathbb{E}[\log^+\{\alpha(z_t)\}] \leq \mathbb{E}[\{\alpha(z_t)\}] = \alpha + \beta$.

(b) Dans le cas où $\alpha_0 = 0$, et $\gamma < 0$, il est clair d'aprés (1.17) que la seul solution stationnaire du modèle est $\epsilon_t = 0$. Il est donc naturel d'imposer $\alpha_0 > 0$ dans la pratique.

(c) On voit que la condition (1.16) dépend de la loi du processus $\{z_t; t \in \mathbb{Z}\}$et qu'elle n'est pas symétrique en α et β.

(d) La condition (1.16) implique $\beta < 1$. Inversement, si

$$\alpha + \beta < 1,$$

(1.16) est vérifiée, car par application de l'inégalité de Jensen

$$\mathbb{E}[\{\alpha(z_t)\}] \leq \log[\mathbb{E}\{\alpha(z_t)\}] = \log(\alpha + \beta) < 0.$$

(e) Si (1.16) est satisfaite, elle l'est également pour tout couple (α_1, β_1) tel que $\alpha_1 = \alpha$ et $\beta_1 = \beta$. En particulier la stationnarité stricte du modèle GARCH implique celle du modèle ARCH obtenu en supprimant β.

(f) Dans le cas $ARCH(1)(\beta = 0)$, la contrainte de stationnarité stricte s'écrit

$$0 \leq \alpha < \exp\{-\mathbb{E}(\log z_t^2)\}. \tag{1.21}$$

1.3.2 La stastionarité faible du GARCH(1,1)

Théorème 1.3. *Stationnarité au second ordre du GARCH(1,1).*
Supposons $\alpha_0 > 0$.

si $\alpha + \beta \geq 1$, *il existe pas de solution GARCH(1,1) non anticipative et stationnaire au second ordre .*

si $\alpha + \beta < 1$, *le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ défini par (1.21), est stationnaire au second ordre. Plus précisément, $\{\epsilon_t; t \in \mathbb{Z}\}$ est un bruit blanc. De plus, il n'existe pas d'autre solution stationnaire au second et non anticipative.*

Preuve. Si $\{\epsilon_t; t \in \mathbb{Z}\}$ est un processus GARCH(1,1), au sens de la définition 1.3, stationnaire au second ordre et non anticipatif, on a

$$\begin{aligned}
\mathbb{E}(\epsilon_t^2) &= \mathbb{E}\{\mathbb{E}(\epsilon_t^2/\epsilon_u, u < t)\} = \mathbb{E}(h_t^2) \\
&= \mathbb{E}\{\alpha_0 + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}^2\} \\
&= \alpha_0 + \alpha\mathbb{E}(\epsilon_{t-1}^2) + \beta\mathbb{E}(\epsilon_{t-1}^2) \\
&= \alpha_0 + (\alpha + \beta)\mathbb{E}(\epsilon_{t-1}^2).
\end{aligned}$$

Soit

$$(1 - \alpha - \beta)\mathbb{E}(\epsilon_t^2) = \alpha_0.$$

Il faut donc $\alpha + \beta < 1$. On obtient de plus : $\mathbb{E}(\epsilon_t^2) > 0$. Inversement, supposons que $\alpha + \beta < 1$. D'aprés la remarque **(d)** précédente, la condition de stationnarité stricte est vérifiée. Il suffit donc de montrer que la solution strictement stationnaire définie en (1.20) admet une variance finie. La variable η_t étant une limite croissante de variables aléatoires positives, d'après le théorème de Beppo Levi, on peut intervertir espérance et somme infinie et écrire :

$$\begin{aligned}
\mathbb{E}(\epsilon_t^2) &= \left(1+\sum_{n=1}^{+\infty}\mathbb{E}\{\alpha(z_{t-1})...\alpha(z_{t-n})\right)\alpha_0\\
&= \left(1+\sum_{n=1}^{+\infty}\{\mathbb{E}(\alpha(z_t))\}^n\right)\alpha_0\\
&= \left(1+\sum_{n=1}^{+\infty}(\alpha+\beta)^n\right)\alpha_0\\
&= \frac{\alpha_0^{n-1}}{1-(\alpha+\beta)}.
\end{aligned}$$

Celà suffit pour affirmer que la solution est faiblement stationnaire. De plus, cette solution est un bruit blanc car :

$$\mathbb{E}(\epsilon_t) = \mathbb{E}\{\mathbb{E}(\epsilon_t/\epsilon_u; u \leq t)\} = 0.$$

et pour tout $j > 0$

$$\begin{aligned}
cov(\epsilon_t, \epsilon_{t-j}) &= \mathbb{E}(\epsilon_t\epsilon_{t-j})\\
&= \mathbb{E}\{\mathbb{E}(\epsilon_t\epsilon_{t-j}/\epsilon_j, j \leq t-1)\}\\
&= \mathbb{E}\{\epsilon_{t-j}\mathbb{E}(\epsilon_t\epsilon_j, j \leq t-j)\} = 0.
\end{aligned}$$

Nous allons montrer l'unicité : Soit $\epsilon_t = \sqrt{\tilde{\eta}_t}z_t$ une autre solution stationnaire au second ordre et non anticipative. On a :

$$|\eta_t - \tilde{\eta}_t| = \alpha_0 + \alpha(z_{t-1}\eta_{t-1} - (\alpha_0 + \alpha(z_{t-1}\tilde{\eta}_{t-1})) = \alpha(z_{t-1})...\alpha(z_{t-n})|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}|$$

et par suite,

$$\begin{aligned}
\mathbb{E}|\eta_t - \tilde{\eta}_t| &= \mathbb{E}\{\alpha(z_{t-1})...\alpha(z_{t-n})\}\mathbb{E}|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}|\\
&= (\alpha+\beta)\mathbb{E}|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}|.
\end{aligned}$$

Notons que la seconde égalité résulte du caractère non anticipatif des solutions, hypothèse qui n'était pas nécessaire pour établir l'unicité de la solution strictement stationnaire. L'espérance de $|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}|$ étant bornée par $\mathbb{E}|\eta_{t-n-1}| + E|\tilde{\eta}_{t-n-1}|$, quantité finie et indépendante de n par stationnarité, et $(\alpha+\beta)^n$ tendant vers 0 quand $n \to 1$, on obtient $\mathbb{E}|\eta_t - \tilde{\eta}_t| = 0$ et donc $\eta_t - \tilde{\eta}_t$ pour tout t, p.s.

1.4 Ecriture vectorielle du GARCH(p,q)

Dans le cas général du GARCH(p,q) fort, l'écriture vectorielle suivante est très utile. On a :

$$Y_t = b_t + A_t Y_{t-1}, \tag{1.22}$$

où $b_t = b_t(z_t) = \begin{bmatrix} \alpha_0 z_t^2 \\ 0 \\ \vdots \\ \alpha_0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{p+q}$, $\qquad Y_t = b_t(z_t) = \begin{bmatrix} \epsilon_t^2 \\ \vdots \\ \epsilon_{t-q+1}^2 \\ h_t^2 \\ \vdots \\ h_{t-p+1}^2 \end{bmatrix} \in \mathbb{R}^{p+q}$,

et

$$A_t = \begin{bmatrix} \alpha_1 z_t^2 & & \cdots & \alpha_q z_t^2 & \beta_1 z_t^2 & & \cdots & \beta_p z_t^2 \\ 1 & 0 & \cdots & 0 & & \cdots & & 0 \\ 0 & 1 & \cdots & 0 & 0 & & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_1 & & \cdots & \alpha_q & \beta_1 & \cdots & & \beta_p \\ 0 & \cdots & 0 & 1 & 0 & \cdots & & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{1.23}$$

A_t une matrice de dimension $(p+q)\times(p+q)$. L'équation(1.22) constitue un modèle vectoriel autorégressif d'ordre 1, avec coefficients positifs et *i.i.d.* La loi de Y_t conditionnelle à son passé infini coincide avec sa loi conditionnelle à Y_{t-1} seulement. En itérant (1.22) on

$$\begin{aligned} Y_t &= b_t + A_t Y_{t-1} \\ &= b_t + A_t(b_{t-1} + A_{t-1} Y_{t-2}) \\ &= \cdots \\ &= b_t + \sum_{k=1}^{\infty} A_t A_{t-1} \ldots A_{t-k+1} b_{t-1}. \end{aligned} \tag{1.24}$$

L'objet de ce qui suit est de trouver des conditions justifiant l'existence de cette série. Une condition suffisante pour que, presque sûrement (1.24) soit strictement superieur à 0, est évidemment :

$$\alpha_0 > 0,\ \alpha_i \geq 0 (i = 1, ..., q),\ \beta_j \geq 0 (j = 1, ..., p). \tag{1.25}$$

Cette condition, très simple à utiliser, n'est cependant pas toujours nécessaire.

1.5 Etude de la stastionnarité

• **Quelques outils d'algebre** : L'outil principale de la stationnarité stricte est le concept de l'exposant de Lyapounov .

Soit A une matrice $(p+q)(p+q)$, son rayon spectrale noté $\rho(A)$ est le plus grand module de ses valeurs propres. Soit$\|.\|$ une norme quelqonque sur l'espace des matrices $(p+q)(p+q)$. On a le résultat d'algebre suivant :

$$\lim_{t\to+\infty}\frac{1}{t}\log\|A^t\| = \log\rho(A). \tag{1.26}$$

Cette propriété s'étend aux matrices aléatoires à travers du résultat suivant :

Théorème 1.4. *Soit $\{A_t, t\in\mathbb{Z}\}$ une suite de matrices aléatoire strictement stationnaire et ergodique telle que $\mathbb{E}\{\log^+(\|A_t\|)\}$ est finie .*
On a

$$\begin{aligned}\lim_{t\to+\infty}\frac{1}{t}\mathbb{E}\{\log\|A_tA_{t-1},...,A_1\|\} &= \gamma \\ &= \inf_{t\in\mathbb{N}}\frac{1}{t}\mathbb{E}\{\log\|A_tA_{t-1},...,A_1\|\},\end{aligned} \tag{1.27}$$

où $\gamma(resp\exp(\gamma))$ s'appele exposant de Lyaponouv (resp.rayon spectral) de la suite de matrice $\{A_t, t\in\mathbb{Z}\}$. De plus

$$\gamma = \lim_{t\to+\infty} p.s.\log\|A_t, A_{t-1},...,A_1\|. \tag{1.28}$$

Le lemme général suivant est très utile pour l'étude des produits des matrices aléatoires.

Lemme 1.1. *Soit $\{A_t, t\in\mathbb{Z}\}$ une suite de matrices i.i.d. telle que $\mathbb{E}\{\log^+\|A_t\|\}$ est finie et de plus grand exposant de Lyapounov γ alors :*

$$\lim_{t\to+\infty} p.s.\|A_t, A_{t-1},...,A_1\| = 0 \quad\Rightarrow\quad \gamma<0. \tag{1.29}$$

1.5.1 La stationnarité forte

Théorème 1.5 (Stationnarité stricte de GARCH(p,q)). *Une condition nécessaire et suffisante d'existance d'un processus GARCH(p,q) strictement stationnaire, solution du modèle (1,14) est que*

$$\gamma<0,$$

où γ est le plus grand exposant de Lyaponouv de la suite $\{A_t, t \in \mathbb{Z}\}$ définie par (1,23). Lorsqu'elle existe, la solution strictement stationnaire est unique, non anticipative et ergodique .

Preuve. Nous utilisons $\|A\| = \sum_{i,j} |a_{ij}|$. Par commodité la norme sera notée d'une manière identique quelle que soit la dimension de la matrice A, avec cette commodité la norme est clairement multiplicative c-à-d : $\|AB\| \leq \|A\|.\|B\|$.
Pour toute matrice A et B telles que AB existe.
Remarquons que les variables μ_t étant des variances finis, tout les termes de la matrice A_t sont intégrable. On a donc

$$\mathbb{E}\{\log^+ \|A_t\|\} \leq \mathbb{E}\{\|A_t\|\} < \infty.$$

Nous supposons que $\gamma < 0$ alors, l'égalité (1,29) implique que la série :

$$Y_t = b_t + \sum_{n=0}^{+\infty} A_t A_{t-1} ... A_{n-1} b_{t-n-1},$$

converge presque sûrement pour toute t. En effet on utilisons la multiplicativité de la norme,

$$\|Y_t\| \leq \|b_t\| + \sum_{n=0}^{+\infty} \|A_t A_{t-1} ... A_{t-n}\| \|b_{t-n-1}\|. \tag{1.30}$$

et

$$\begin{aligned}(\|A_t A_{t-1} ... A_{n-1}\|^{\frac{1}{n}} \|b_{t-n-1}\|)^{\frac{1}{n}} &= \exp\{\frac{1}{n} \log \|A_t A_{t-1} ... A_{t-n}\| + \frac{1}{n} \log \|b_{t-n-1}\|\} \\ &\xrightarrow{p.s} \quad \exp^{\gamma} < 1.\end{aligned}$$

Y_t est bien défini dans $(\mathbb{R}^{*+})^{p+q}$. Soit $Y_{q+1,t}$ la $(q+1)$-ème composante de Y_t.
En posant

$$\epsilon_t = \sqrt{\tilde{Y}_{q+1,t}}\, \mu_t.$$

On définit une solution stricte stationnaire du modèle (1,14). ϵ_t s'éxprime comme fonction mesurable de $\mu_t \mu_{t-1}, ...$, la solution est donc non anticipative et ergodique car $\{\mu_t; t \in \mathbb{Z}\}$ est ergodique.

Supposons qu'il existe une autre solution strictement stationnaire ergodique non anticipative, alors pour toute $n \geq 0$,

$$\begin{aligned} \|Y_t - Y_t^*\| &= \|(A_t Y_{t-1} + b_t) - (A_t Y_{t-1}^* + b_t)\| \\ &= \|A_t (Y_{t-1} - Y_{t-1}^*)\| = ... \\ &= \|A_t A_{t-1} ... A_{t-n} (Y_{t-1} - Y_{t-1}^*)\| \\ &\leq \|A_t A_{t-1} ... A_{t-n}\| \|Y_{t-1} - Y_{t-1}^*\|. \end{aligned}$$

On a $\mathbb{P}\{(Y_t - Y_t^*)\} > 0$. Or on sait que $\|A_t A_{t-1} ... A_{t-n}\| \rightarrow \quad p.s.$ quand $n \rightarrow \infty$ car la série intervenant dans (1.30) converge. Par suite $\mathbb{P}(\|Y_{t-n-1} - Y_{t-n-1^*}\| \rightarrow \infty) > 0$, ce qui implique que $\|Y_{t-n-1}\| \rightarrow \infty$ ou $\|Y_{t-n-1}^*\| \rightarrow \infty$ avec une probabilite positive.
Ceci est impossible car les suites $(Y_t)_t$ et $(Y_t^*)_t$ sont stationnaires. On en conclut que $Y_t = Y_t^*$ pour tout t, p.s.
Nous montrons finalement la partie nécessaire du théorème. D'après le lemme (1,1), il suffit d'établir (1.29). Nous allons montrer que, pour $1 \leq i \leq p + q$

$$lim_{t \to \infty} A_0 ... A_{-t} = 0, \quad p.s. \tag{1.31}$$

On a pour $t > 0$

$$\begin{aligned} Y_0 &= b_0 + A_0 Y_{-1} \\ &= b_0 + \sum_{k=0}^{t-1} A_0 ... A_{-k} b_{-k-1} + A_0 ... A_{-t} Y_{t-1} \\ &= \sum_{k=0}^{t-1} A_0 ... A_{-k} b_{-k-1}. \end{aligned}$$

Car les coefficients des matrices A_t, b_0 et Y_t sont positifs. Par suite la série $\sum_{k=0}^{t-1} A_0 ... A_{-k} b_{-k-1}$ converge et donc $A_0 ... A_{-k} b_{-k-1}$ tend presque sûrement vers 0 quand $k \rightarrow \infty$. Donc $A_0 A_{-1} ... A_{-k} \rightarrow 0 \Rightarrow \gamma < 0$.
Le théorème suivant donne des conditions nécessaire et suffisante de stationnarité du deuxième ordre.

1.5.2 La stationnarité faible

Théorème 1.6. *S'il existe un processus GARCH(p,q), au sens de la définition 1.4, stationnaire au deuxième ordre et non anticipatif , et $\alpha_0 > 0$, alors :*

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1. \tag{1.32}$$

Inversement, si (1,32) est vérifieé l'unique solution strictement stationnaire du modèle (1,14)est un bruit blanc(donc est stationnaire au deuxième ordre).
Il n'existe pas d'autre solution stationnaire au deuxième ordre.

Preuve. Montrons d'abort que (1.32) est une condition nécessaire.
Soit $\{\epsilon_t; t \in \mathbb{Z}\}$ un processus GARCH(p,q) stationnaire au deuxième ordre et non anticipatif, alors :

$$\begin{aligned}
\mathbb{E}\{\epsilon_t^2\} &= \mathbb{E}\{\mathbb{E}\{\epsilon_t^2/\epsilon_j, j \leq t-1\}\} \\
&= \mathbb{E}\{h_t^2\} \\
&= \mathbb{E}\{\alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2\} \\
&= \alpha_0 + \sum_{i=1}^{q} \alpha_i \mathbb{E}\{\epsilon_{t-1}^2\} + \sum_{j=1}^{p} \beta_j \mathbb{E}\{\sigma_{t-j}^2\} \\
&= \alpha_0 + \sum_{i=1}^{q} \alpha_i \mathbb{E}\{\epsilon_t^2\} + \sum_{j=1}^{p} \beta_j \mathbb{E}\{\epsilon_t^2\}.
\end{aligned}$$

Alors, on aura

$$\mathbb{E}\{\epsilon_t^2\}(1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j) = \alpha_0$$
$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1.$$

Chapitre 2

Estimation des paramètres ARCH/GARCH

Dans ce chapitre, nous allons traiter l'estimation des paramètres d'un modèle (G)ARCH, et plus généralement, d'un modèle de régression avec erreur (G)ARCH. Les modèles introduits reposent sur des formulations des moyennes et variances conditionnelles. En pratique celle-ci souvent paramètrées de façon que la moyenne conditionnelle $m_t(\theta)$ et la variance conditionnelle $h_t^2(\theta)$ apparaissent comme des fonctions de paramètres inconus et de valeurs passées du processus. La connaissance de ces moments ne suffit cependant pas sans hypothèse supplémentaire à caractériser la loi conditionnelle du processus.
La vraisemblance est écrite comme si la loi des variables z_t était normale centrée réduite (on parle de pseudo ou quasi-vraisemblance), mais cette hypothèse n'est pas nécessaire pour la convergence forte de l'estimateur. Elle a évidemment un effet sur la variance de la loi normale asymptotique de l'estimateur. Les modèles avec erreurs hétéroscédastiques peuvent être estimés généralement par ces méthodes :

- Estimation par la méthode des moindres carrés ordinaires (MCO).
- Estimateurs par la méthode du Maximum de Vraisemblance (MV).

Nous présenterons ici ces deux méthodes d'estimation. Commençons par étudier la méthode des moindres carrés, ainsi que le cas des estimateurs du (MV) obtenu en cas d'erreur sur la spécification de la loi conditionnelle des résidus.

2.1 Estimation des modèles ARCH par la méthode des MCO

Dans cette partie nous considérons l'estimation par la méthode MCO du modèle ARCH(q) :

$$\begin{aligned} \epsilon_t &= z_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \qquad \alpha_0 > 0, \alpha_i \geq 0, i = 1, ..., q, \end{aligned} \tag{2.1}$$

où $\{z_t; t \in \mathbb{Z}\}$ est une suite de variables $i.i.d$, $\mathbb{E}(z_t) = 0$, $\mathbb{V}(z_t) = 1$. La méthode consiste à tirer partie de la représentation AR sur le carré du processus observé et à appliquer la méthode des moindres carrés (MC). Les estimateurs obtenus sont, au moins pour n grand, moins précis que ceux du maximum de vraisemblance (MV), mais plus faciles à obtenir. Ils peuvent également fournir des valeurs initiales pour la procédure d'optimisation utilisée dans l'obtention d'estimateurs du (MV) plus précis.
On déduit de (2.1) la représentation AR(q) :

$$\epsilon_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \mu_t, \tag{2.2}$$

où $\mu_t = \epsilon_t^2 - h_t^2 = (z_t^2 - 1)h_t^2$. La suite $(\mu_t, \mathcal{F}_{t-1})$ constitue une différence de martingale.
On suppose que l'on dispose d'observations $\epsilon_1, ..., \epsilon_n$ réalisation partielle du processus $\{\epsilon_t; t \in \mathbb{Z}\}$, et de valeurs initiales $\epsilon_0, ..., \epsilon_{1-q}$.

$$\begin{aligned} \epsilon_n^2 &= \alpha_0 + \alpha_1 \epsilon_{n-1}^2 + \alpha_2 \epsilon_{n-2}^2 + ... + \alpha_q \epsilon_{n-q}^2 + \mu_n \\ \epsilon_{n-1}^2 &= \alpha_0 + \alpha_1 \epsilon_{n-2}^2 + \alpha_2 \epsilon_{n-3}^2 + ... + \alpha_q \epsilon_{n-1-q}^2 + \mu_{n-1} \\ &. \\ &. \\ &. \\ \epsilon_1^2 &= \alpha_0 + \alpha_1 \epsilon_0^2 + ... + \alpha_q \epsilon_{1-q}^2 + \mu_1. \end{aligned}$$

Introduisons le vecteur

$$z_{t-1}' = (1, \epsilon_{t-1}^2, ..., \epsilon_{t-q}^2).$$

On déduit le système

$$\epsilon_t^2 = (z_{t-1}'\theta_0 + \mu_t), \ avec \ \theta_0 = (\alpha_0, ..., \alpha_q) \quad t = 1, ..., n. \tag{2.3}$$

Soit $Y = X\theta_0 + U$. En définissant la matrice $n \times q$ et les vecteurs $n \times 1$:

$$X = \begin{bmatrix} z'_{t-1} \\ \vdots \\ z'_0 \end{bmatrix}, \; Y = \begin{bmatrix} \epsilon_n^2 \\ \vdots \\ \epsilon_1^2 \end{bmatrix}, \; U = \begin{bmatrix} \mu_n^2 \\ \vdots \\ \mu_0 \end{bmatrix}.$$

Supposons que la matrice $X'X$ soit inversible. On déduit l'estimateur des MCO de θ :

$$\hat{\theta}_n = (X'X)^{-1}X'Y. \tag{2.4}$$

Pour établir la consistance et la normalité asymtotic de l'estimateur de moindres carrées, les hypothèses suivantes sont nécessaires :

H1 : $\{\epsilon_t\}$ est une solution non anticipative strictement stationnaire .

H2 : $\mathbb{E}(\epsilon_t^4) < +\infty$.

H3 : $\mathbb{P}[z_t^2 = 1] \neq 1$.

Théorème 2.1. ***(Convergence des estimateurs MCO pour un ARCH)***
Soit $\{\hat{\theta}_n\}$ une suite d'estimateurs satisfaisant (2,4). Sous les hypothèses ***H1-H3****, presque sûrement*

$$\hat{\theta}_n \to \theta_0, \; \hat{h}_t^2 \to h_0^2 \, quand \, n \to \infty.$$

Preuve. Etape 1 : Nous avons vu que l'unique solution stationnaire non anticipative de $\{\epsilon_t; t \in \mathbb{Z}\}$ est ergodique. Le processus $\{z_t; t \in \mathbb{Z}\}$ est également ergodique car z_t s'écrit comme fonction mesurable.

$$\frac{1}{n}X'X = \frac{1}{n}\sum_{t=1}^{n} z_{t-1}z'_{t-1} \to \mathbb{E}(z_{t-1}z'_{t-1}) \; p.s. \quad quand \, n \to \infty. \tag{2.5}$$

L'existence de l'espérance est assurée par l'hypothèse **H2**. On a donc

$$\frac{1}{n}X'Y = \frac{1}{n}\sum_{t=1}^{n} z_{t-1}\epsilon_t^2 \to \mathbb{E}(z_{t-1}\epsilon_t^2) \; p.s. \quad quand \, n \to \infty.$$

Etape 2 : Montrons par l'absurde l'inversibilité de la matrice $\mathbb{E}(z_{t-1}z'_{t-1}) = \mathbb{E}(z_t z'_t)$. Supposons qu'il existe un vecteur non nul $c \in \mathbb{R}^{q+1}$, tel que :

$$c'\mathbb{E}(z_t z'_t) = 0$$
$$\mathbb{E}(c'z_t(c'z_t)') = 0.$$

d'où l'on déduit que $c' \, z_t$ est p.s. constant.

Par suite, il existe une combinaison linéaire p.s. égale à une constante des variables $\epsilon_t^2, ..., \epsilon_{t-q+1}$.

Sans perte de généralité, on suppose que le coefficient de ϵ_t^2 égale à 1. Donc $\{z_t\}$ s'exprime p.s comme fonction mesurable des variables $\epsilon_{t-1}, \epsilon_{t-2}, ..., \epsilon_{t-q}$. Comme $\{z_t\}$ est indépendante de ces variables, donc $\{z_t^2\}$ est égale à une constante p.s. Forcément, cette constante ne peut-être que 1, d'où la contradiction avec **H3**.

Etape 3 : Il découle de ce qui précède $\frac{1}{n}X'X$ que est p.s inversible, pour n assez grand et que p.s quand $n \to \infty$,

$$\hat{\theta}_n = \left(\frac{X'X}{n}\right)^{-1} \frac{X'Y}{n} \to \{\mathbb{E}(z_{t-1}z'_{t-1})\}^{-1}\mathbb{E}(z_{t-1}\epsilon_t^2).$$

Etape 4 : Le processus $\{\mu_t; t \in \mathbb{Z}\}$ est l'innovation forte de $\{\epsilon_t^2\}$. On a donc, en particulier, les relations d'orthogonalité

$$\mathbb{E}(\mu_t) = \mathbb{E}(\mu_t\epsilon_t^2 - 1) = ... = \mathbb{E}(\mu_t\epsilon_t^2 - q) = 0,$$

c-à-d $\mathbb{E}(z_{t-1}\mu_t) = 0$. d'où $\mathbb{E}(z_{t-1}\epsilon_t^2) = \mathbb{E}(z_{t-1}z'_{t-1})\theta_0$.

Donc d'après les étapes 2 et 3, $\hat{\theta}_n$ converge p.s vers θ_0. La convergence forte de $\hat{h}_t^2$ vers h_t^2 s'en déduit.

Pour la normalité asymptotique de l'estimateur des MCO, nous devons faire l'hypothèse supplémentaire suivante :

H4 : $\mathbb{E}(\epsilon_t^8) < +\infty$.

Introduisons les matrices carrées symétriques de taille $q + 1$.

$$A = \mathbb{E}(z_{t-1}z'_{t-1})\,, I = \mathbb{E}(h_t^4 z_{t-1}z'_{t-1}).$$

L'inversibilité de la matrice A ce démontre de la même façon comme précédent. celle de I sera montrée dans la preuve du résultat suivant, qui établit la normalité asymptotique de l'estimateur des MCO.

Théorème 2.2. *Sous les hypothèses **H1-H4***

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\mu_4 - 1)A^{-1}IA^{-1}).$$

Preuve. On a :

$$\begin{aligned}
\hat{\theta}_n &= \left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}z'_{t-1}\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}\epsilon_t^2\right) \\
&= \left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}z'_{t-1}\right)^{-1}\left\{\frac{1}{n}\sum_{t=1}^{n} z_{t-1}(z'_{t-1}\theta_0+\mu_t)\right\} \\
&= \theta_0+\left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}z'_{t-1}\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}\mu_t\right).
\end{aligned}$$

Donc

$$\sqrt{n}(\hat{\theta}_n-\theta) = \left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}z'_{t-1}\right)^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} z_{t-1}\mu_t\right\}. \tag{2.6}$$

Soit $\lambda \in \mathbb{R}^{q+1}, \lambda \neq 0$. La suite $(\lambda' z_{t-1}\mu_t, \mathcal{F}_t)$ est une différence de martingale stationnaire, ergodique et de carré intégrable de variance

$$\begin{aligned}
\mathbb{V}(\lambda' z_{t-1}\mu_t) &= \lambda'\mathbb{E}\{z_{t-1}z'_{t-1}\mu_t^2\}\lambda \\
&= \lambda'\mathbb{E}\{z_{t-1}z'_{t-1}(z_t^2-1)h_t^4\}\lambda \\
&= (\mu_t^4-1)\lambda' I\lambda.
\end{aligned}$$

On déduit que, pour tout $\lambda \neq 0$

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\lambda' z_{t-1}\mu_t \xrightarrow{\mathcal{L}} \mathcal{N}(0,(\mu_4-1)\lambda' I\lambda).$$

Par suite, en appliquant la propriété de Cramer-Wold,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} z_{t-1}\mu_t \xrightarrow{\mathcal{L}} \mathcal{N}(0,(\mu_4-1)I). \tag{2.7}$$

On montre que cette loi limite est non dégénérée, c'est-à-dire que I est inversible, par le même raisonnement que celui utilisé pour établir l'inversibilité de A dans la preuve du Théorème (2.1) Par suite, on déduit de (2.5), (2.6) et (2.7), par un raisonnement classique, que $\sqrt{n}(\hat{\theta}_n-\theta)$ est asymptotiquement normal, de moyenne le vecteur nul, et de variance la matrice du théorème.

2.2 Méthode d'estimation du maximum de vraisemblance

Dans cette partie nous étudions la méthode du maximum de vraisemblance conditionnelle (à des valeurs initiales). Nous présentons de façon parallèle la méthode d'estimation du maximum de vraisemblance (MV) sous l'hypothèse de normalité, de la distribution conditionnelle des résidus et la méthode d'estimation du quasi-maximum de vraisemblance (QMV). En effet, l'idée générale des estimateurs du Pseudo Maximum de Vraisemblance (PMV) consiste à démontrer que si l'on commet une erreur sur la distribution conditionnelle des résidus en utilisant à tort une log-vraisemblance fondée sur une loi normale, l'estimateur du (MV) ainsi obtenu peut tout de même être convergent si la vraie loi des résidus appartient à la loi normale (Gourieroux, Montfort, 1989). L'estimateur sera (1) asymptotiquement convergent et (2) asymptotiquement normal. Par conséquent, la fonction de vraisemblance définissant l'estimateur du (MV) sous l'hypothèse de normalité et la fonction de quasi-vraisemblance de l'estimateur du (QMV) sont les mêmes.

2.2.1 Quasi-vraisemblance conditionnelle

On supposera que les observations $\epsilon_1, ..., \epsilon_n$ constituent une réalisation (de longueur n) d'un processus GARCH(p,q), solution strictement stationnaire non anticipative du modèle

$$\begin{aligned} \epsilon_t &= z_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2, \end{aligned} \tag{2.8}$$

où $\{z_t; t \in \mathbb{Z}\}$ une suite de variables aléatoires *i.i.d.* centrées et de variance unité,

$$\alpha_0 > 0,\ \alpha_i \geq 0 (i = 1, ..., q),\ \beta_j \geq 0 (j = 1, ..., p).$$

Les ordres p et q sont supposés connus. Le vecteur des paramètres

$$\theta = (\theta_1, ..., \theta_{p+q+1})' = (\alpha_0, \alpha_1, ..., \alpha_q, \beta_1, ..., \beta_p)', \tag{2.9}$$

appartient à un espace de paramètres $\Theta \subset]0, \infty[\times [0, \infty[^{p+q}$. La vraie valeur du paramètre est inconnue et est notée $\theta_0 = (\alpha_0, \alpha_{01}, ..., \alpha_{0q}, \beta_{01}, ..., \beta_{0p})'$.
Pour écrire la vraisemblance du modèle, il faut spécifier une distribution particulière pour les variables z_t. On considère généralement la quasi-vraisemblance gaussienne, i.e la vraisemblance obtenue à partir d'une loi normale centrée réduite pour les z_t.

La spécification d'une distribution gaussienne pour les variables z_t ne permet pas d'en déduire simplement la loi de l'échantillon. On travaille avec la vraisemblance de $\epsilon_1, ..., \epsilon_n$ conditionnellement à certaines valeurs initiales.

Etant données des valeurs initiales $\epsilon_0, ..., \epsilon_{1-q}, \tilde{h}_0^2, ..., \tilde{h}_{1-p}^2$ que nous allons préciser, la vraisemblance conditionnelle gaussienne $L_n(\theta)$ s'écrit

$$L_n(\theta) = L_n(\theta; \epsilon_1, ..., \epsilon_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{h}_t^2}} \exp\left(\frac{-\epsilon_t^2}{2\tilde{h}_t^2}\right), \tag{2.10}$$

où les $\tilde{h}_t^2$ sont définis récursivement, pour $t \geq 1$, par

$$\tilde{h}_t^2 = \tilde{h}_t^2(\theta) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \tilde{h}_{t-j}^2. \tag{2.11}$$

Pour une valeur donnée de θ, sous l'hypothèse de stationnarité au second ordre, la variance non conditionnelle (correspondant à cette valeur de θ) est un choix raisonnable pour les valeurs initiales inconnues :

$$\epsilon_0^2 = ... = \epsilon_{1-q}^2 = h_0^2 = ... = h_{1-p}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}. \tag{2.12}$$

De telle valeurs initiales ne conviennent pas notamment pour les modèles IGARCH, pour lesquels l'hypothèse de stationnarité au second ordre est relachée, car la constante (2.12) prendrait des valeurs négatives pour certaines valeurs de θ. On peut alors proposer de prendre comme valeurs initiales

$$\epsilon_0^2 = ... = \epsilon_{1-q}^2 = \tilde{h}_0^2 = ... = \tilde{h}_{1-p}^2 = \alpha_0. \tag{2.13}$$

Ou

$$\epsilon_0^2 = ... = \epsilon_{1-q}^2 = \tilde{h}_0^2 = ... = \tilde{h}_{1-p}^2 = \epsilon_1^2. \tag{2.14}$$

Un estimateur du (QMV) de θ est défini comme toute quantité $\hat{\theta}_n$ vérifiant presque sûrement

$$L_n(\hat{\theta}_n) = \sup_{\theta \in \Theta} L_n(\theta). \tag{2.15}$$

On voit, en prenant le logarithme, que maximiser la vraisemblance revient à minimiser par rapport à θ

$$\tilde{I}_n(\theta) \;=\; n^{-1}\sum_{t=1}^{n}\tilde{\ell}_t, \tag{2.16}$$

où $\tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\epsilon_t^2}{\tilde{h}_t^2} + \ln \tilde{h}_t^2$. Et $\tilde{h}_t^2$ est définie en (2.11). Un estimateur du quasi-maximum de vraisemblance est donc une solution mesurable de l'équation

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta}\tilde{I}_n. \tag{2.17}$$

2.2.2 Equations de vraisemblance

On obtient les équations de vraisemblance en annulant la dérivée par rapport à θ du critère $\tilde{I}_n$, ce qui donne

$$\frac{1}{n}\sum_{t=1}^{n}\{\epsilon_t^2 - \tilde{h}_t^2\}\frac{1}{h_t^4}\frac{\partial \tilde{h}_t^2}{\partial\theta} = 0. \tag{2.18}$$

En effet, comme nous le verrons plus précisément dans la partie suivante, le terme de gauche de l'égalité précédente se comporte asymptotiquement comme

$$\frac{1}{n}\sum_{t=1}^{n}\{\epsilon_t^2 - h_t^2\}\frac{1}{h_t^4}\frac{\partial h_t^2}{\partial\theta} = 0. \tag{2.19}$$

L'influence des valeurs initiales étant nulle lorsque $n \to \infty$. Or, pour la vraie valeur du paramètre, l'innovation de ϵ_t^2 est $\mu_t = \epsilon_t^2 - h_t^2$. Donc sous réserve que l'espérance existe, on a

$$\mathbb{E}_{\theta_0}\left(\mu_t \frac{1}{h_t^4(\theta)}\frac{\partial h_t^2(\theta_0)}{\partial\theta}\right).$$

Car $\frac{1}{h_t^4(\theta)}\frac{\partial h_t^2(\theta_0)}{\partial\theta}$ est une fonction mesurable des $\epsilon_{t-i}, i > 0$. Ce résultat n'est autre que la version asymptotique de (2.16) en θ_0, en utilisant le théorème ergodique.

2.2.3 Propriétés asymptotiques de l'estimateur du (QMV)

Dans ce qui suit, nous utilisons comme norme d'une matrice $A = (a_{ij})$ quelconque la norme $\|A\| = \sum |a_{ij}|$. Le rayon spectral d'une matrice A carrée sera noté $\rho(A)$. Le produit de Kronecker sera noté $\bigotimes$.

Convergence forte

Rappelons que le modèle(2.8) possède une solution strictement stationnaire si et seulement si le coefficient de Lyapounov de la suite de matrice

$$A_t = \begin{bmatrix} \alpha_1 z_t^2 & & \cdots & \alpha_q z_t^2 & \beta_1 z_t^2 & & \cdots & \beta_p z_t^2 \\ 1 & 0 & \cdots & 0 & & \cdots & & 0 \\ 0 & 1 & \cdots & 0 & 0 & & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_1 & & \cdots & \alpha_q & \beta_1 & \cdots & & \beta_p \\ 0 & \cdots & 0 & 1 & 0 & \cdots & & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

est strictement négatif. On note $\gamma(\theta)$ ce coefficient de Lyapounov. Notons

$$\mathcal{A}_\theta(z) = \sum_{i=1}^{q} \alpha_i z^i \quad et \quad \mathcal{B}_\theta(z) = 1 - \sum_{j=1}^{q} \beta_j z^j.$$

Par convention $\mathcal{A}_\theta(z) = 0$ si $q = 0$ et $\mathcal{B}_\theta(z) = 1$ si $p = 0$.
Pour la convergence, les hypothèses suivantes sont faites.

A1 : $\theta_0 \in \Theta$ et Θ est compact.

A2 : $\gamma(\theta_0) < 0$ et $\forall \theta \in \Theta$, $\sum_{j=1}^{p} \beta_j < 1$.

A3 : z_t^2 a une loi non dégénérée.

A4 : si $p > 0, \mathcal{A}_{\theta_0}(z)$ et $\mathcal{B}_{\theta_0}(z)$ n'ont pas de racine commune, $\mathcal{A}_{\theta_0}(1) \neq 0$, et $\alpha_{0q} + \beta_{0p} \neq 0$.
Il est pratique d'approximer la suite $\{\tilde{\ell}_t(\theta)\}$ par une suite stationnaire ergodique. La condition de stricte stationnarité **A2** implique que les racines de $\mathcal{B}_\theta(z)$ sont extérieures au disque unité. Notons donc $(h_t^2)_t = \{h_t^2(\theta)\}_t$ la solution strictement stationnaire ergodique et non anticipative de

$$h_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2 \quad \forall t. \tag{2.20}$$

et soit

$$\tilde{I}_n(\theta) = n^{-1} \sum_{t=1}^{n} \tilde{\ell}_t,$$

où $\tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\epsilon_t^2}{h_t^2} + \ln h_t^2$.

Théorème 2.3. *Soit $\{\hat{\theta}_n\}$ une suite d'estimateurs du (QMV) satisfaisant (2.17), avec les conditions initiales (2.13) ou (2.14). Sous les hypothèses **A1-A4**, presque sûrement*

$$\hat{\theta}_n \to \theta_0, \quad quand\, n \to \infty.$$

Remarque 2.1.

1. On ne suppose pas que la vraie valeur θ_0 du paramètre appartient à l'intérieur de Θ. Le théorème permet donc de traiter les cas où certains coefficients, α_i ou β_j , sont nuls.

2. L'hypothèse **A4** disparait dans le cas ARCH. Elle permet de suridentifier l'un des ordres, p ou q, mais pas les deux.

3. L'hypothèse **A4** exclut le cas où tous les θ_{01} sont nuls. Ceci est évidemment nécessaire, sinon le modèle a pour solution un bruit blanc fort qui peut s'écrire de multiples manières. Par exemple, un bruit blanc fort de variance 1 peut s'écrire sous la forme d'un GARCH(1,1) avec
$$h_t^2 = \alpha_0 + 0 \times \epsilon_{t-1}^2 + \beta h_{t-1}^2,$$
pour tous α_0 et β positifs tels que $\alpha_0 = 1 - \beta$.

4. L'hypothèse d'absence de racines communes, dans **A4**, n'est restrictive que si $p > 1$ et $q > 1$. En effet, si $q = 1$ la seule racine de $\mathcal{A}_{\theta_0}(z)$ est 0 et $\mathcal{B}_{\theta_0}(0) \neq 0$.
Si $p = 1$ et $\mathcal{B}_{\theta_0 1} \neq 0$, la seule racine de $\mathcal{B}_{\theta_0}(z)$ est $1/\mathcal{B}_{\theta_0} > 0(si\mathcal{B}_{01} = 0)$, le polynôme n'admet pas de racine). En raison de la positivité des coefficients α_{01}, cette valeur ne peut annuler $\mathcal{A}_{\theta_0}(z)$.

Normalité asymptotique

Pour montrer la normalité asymptotique les hypothèses supplémentaires suivantes sont nécessaires.

A5 : $\theta_0 \in \dot{\Theta}$ où $\dot{\Theta}$ est l'intérieur de Θ.
A6 : $\kappa_z = \mathbb{E}z_t^4 < \infty$.

Théorème 2.4. ***(Normalité asymptotique des estimateurs du (QMV)).*** *Sous les hypothèses* **A1 − A6**, $\sqrt{n}(\tilde{\theta}_n - \theta_0)$ *tend en loi vers une* $\mathcal{N}$, *où*

$$J = \mathbb{E}_{\theta_0}\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta'}\right) = \mathbb{E}_{\theta_0}\left(\frac{1}{h_t^4(\theta_0)}\frac{\partial h_t^2(\theta_0)}{\partial\theta}\frac{\partial h_t^2(\theta_0)}{\partial\theta'}\right). \tag{2.21}$$

Remarque 2.2.

1. L'hypothèse **A5** est classique car elle permet d'utiliser le fait que les conditions du premier ordre sont valides, au moins asymptotiquement. En effet si $\hat{\theta_n}$ est convergent, il appartient également à l'intérieur de Θ pour n grand. En tant que maximum, il doit donc annuler la dérivée de la fonction critère. Cette hypothèse est cependant restrictive car elle exclut par exemple le cas $\alpha_{01} = 0$(il est cependant clair que dans ce cas, $\sqrt{n}(\hat{\alpha}_1 - \alpha_{01})$ est concentrée sur $[0, \infty[$ et ne peut donc être asymptotiquement normale). Ce type de problèmes, dits "de bord", doit faire l'objet d'une étude spécifique.

2. L'hypothèse **A6** ne porte pas sur ϵ_t^2 , et n'exclut bien sûr pas le cas IGARCH. Seule une hypothèse d'existence du moment d'ordre 4 est imposée sur la suite $\{z_t\}$. Cette hypothèse est clairement nécessaire pour l'existence de la variance du vecteur du score $\partial\ell_t(\theta_0)/\partial\theta$.

Chapitre 3

Estimation des modèles GARCH par Bootstrap

Le terme de rééchantillonnage, ou en anglais, "Bootstrap", qui évoque l'action de "se hisser en tirant sur ses propres lacets", désigne un ensemble des méthodes qui consistent à faire de l'inférence statistique sur des "nouveaux" échantillons tirés à partir d'un échantillon initial. Disposant d'un échantillon destiné à donner une certaine information sur une population, on tire au sort, parmi la sous-population réduite à cet échantillon, un nouvel échantillon de même taille n, et on répète cette opération B fois, où B est grand. On analyse ensuite les nouvelles observations ainsi obtenues pour affiner l'inférence faite sur les observations initiales. A priori, on peut avoir des doutes sur l'efficacité d'une telle méthode et penser qu'il n'y a aucune amélioration à espérer en rééchantillonnant à partir du même échantillon. En effet, aucune information supplémentaire ne peut être espérée, toute l'information étant contenue dans l'échantillon initial. Cependant, comme on va le voir, ce rééchantillonnage, s'il ne rajoute aucune information, permet, dans certains cas, d'extraire de l'échantillon de base l'information souhaitée.

Le Bootstrap est une technique de rééchantillonnage introduite par Efron (1979)[15] permettant de simuler la distribution d'un estimateur quelconque pour en apprécier le biais, la variance donc le risque quadratique ou encore pour en estimer un intervalle de confiance même si la loi théorique est inconnue. Les méthodes de Bootstrap sont mises en oeuvre afin d'améliorer la précision des estimations statistiques .
Le Bootstrap généralisée est une combinaison de trois méthodes (le Bootstrap d'Efron, le Bootstrap Bayésien et Bootstrap à pondération aléatoire).

La méthode de vraisemblance quasi-maximum gaussien (QMV) est couramment utilisé pour estimer le modèle GARCH, alors que sa normalité asymptotique nécessite la condition de moment $\mathbb{E}(\eta_t^4) < \infty$; voir Hall et Yao (2003), Berkes et Horváth (2004) et Francq et Zakoian (2004, 2010). Cependant, de nombreuses études empiriques suggérées (Rachev et Mittnik,2000), les séries chronologiques financières et économiques sont généralement à queue lourde.
Dans ce cas, la méthode du Bootsrap généralisée pour l'estimateur qui mimimise l'écart absolu moyen (EAM) est suggérée, et la normalité asymptotique peut être établie dans des conditions de moments clémentes sur les distributions d'innovation. En outre, en tant que dernière étape de la procédure Box-Jenkins, le contrôle diagnostique est d'une importance cruciale. Importance dans la modélisation des séries chronologiques. Li et Li (2005) ont proposé un test de Portemanteau basé sur le résidu absolu auto-corrélations, pour vérifier l'adéquation des modèles GARCH.

Ce test est interéssant dans le sens où les propriétés asymptotiques ne nécessitent que $\mathbb{E}(\eta_t^2) < \infty$, alors que le les tests conventionnels, basés sur des auto-corrélations résiduelles au carré, nécessitent au moins un quatrième moment fini .

Il est à noter que la variance asymptotique de l'estimateur (EAM) et la construction du test de Portemanteau statistiques dépendent touts les deux de la fonction de densité inconnue des innovations. Par conséquent, l'inférence pour les modèles GARCH ajusté par la méthode (EAM) nécessite des méthodes non paramétriques telles que l'estimation de la densité du noyau, et la sélection de paramètres de réglage tels que la bande passante est inévitable dans ces méthodes. Cela impose des difficultés à estimer les distributions asymptotiques. De plus, l'approximation sur échantillon des distributions asymptotiques peut ne pas être assez bon lorsque la taille de l'échantillon est petite, voire modérée. En effet, depuis l'apparition des équipements informatiques puissants, l'approche bootstrap a été largement utilisée pour limiter les distributions de statistiques.
La méthode du Bootstrap, basée sur une estimation pondérée de manière aléatoire, a récemment attiré de plus en plus d'attentions. Par exemple, Rubin (1981) a proposé le Bootstrap Bayésien, également appelé pondérée de manière échangeable. Bootstrap, qui a été étudié par Praestgaard et Wellner (1993) et Cheng (2015), et une pondération aléatoire Bootstrap avec *i.i.d.* poids a été introduit par Zheng (1987). Chatterjee et Bose (2000) ont proposé un Bootstrap généralisé, qui inclut le Bootstrap d'Efron, le Bootstrap Bayésien

et Bootstrap à pondération aléatoire en tant que cas particuliers. De plus, le Bootstrap généralisé a une belle interprétation et des applications utiles dans la pratique.
Bootstrap à base de résidus, tel que la méthode par blocs et le Bootstrap sauvage méthode, joue un rôle important dans la littérature des séries chronologiques, ce qui est dû au fait que les observations sont ordonnées par le temps. Chatterjee et Bose (2005) ont montré par simulation les expériences selon lesquelles le Bootstrap généralisé peut surpasser certaines méthodes de Bootstrap basées sur les résidus pour modèles de séries chronologiques hétéroscédastiques.
De plus, le Bootstrap à pondération aléatoire avec poids *i.i.d.* , cas particulière du Bootstrap généralisé , s'est également avéré performant lors du démarrage des tests de Portemanteau. Motivé par tous ces avantages , cette partie considère l'inférence de Bootstrap généralisée pour les modèles GARCH équipés de la méthode (EAM). Les justifications théoriques sont également établis pour l'estimation (EAM) et le test de Portemanteau.
Nos études de simulation montrent que le Bootstrap généralisé est insensible au choix des poids aléatoires, et nous pouvons simplement utiliser le *i.i.d.* poids dans des applications réelles.

Le long de ce chapitre les notations $\xrightarrow{\mathcal{L}}$ désigne la convergence en loi, $o_p(1)$ désigne une séquence de variables aléatoires convergentes en probabilité à zéro.

3.1 Construction de l'estimateur (EAM)

Rappelons que le modèle d'Hétéroscédastique Conditionnelle Autorégressive Généralisée (GARCH), proposé par Bollerslev (1986),est définis comme suit

$$y_t = \sigma_t\eta_t, \sigma_t^2 = c + \sum_{i=1}^{q}\alpha_i y_{t-i}^2 + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2. \tag{3.1}$$

où $c > 0, \alpha_i \geq 0$ avec $1 \leq i \leq q, \beta_j \geq 0p$, $1 \leq j \leq p$, et les innovations $\{\eta_t\}$ sont *i.i.d.* de moyenne zéro et de variance un.
Notons $\delta > 0$ la médiane de η_t^2, et posons $\epsilon_t = \eta_t/\sqrt{\delta}$ telle que la médiane de ϵ_t^2 est une. Le modèle (3.1) peut ensuite être réécrit en t comme suit :

$$y_t = \epsilon_t\sqrt{h_t},\ h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i y_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j}, \tag{3.2}$$

où $h_t = \delta\sigma_t^2, \alpha_0 = \delta c, \alpha_\sigma a_i = i$, $1 \leq i \leq q, \beta_j = b_j$, $1 \leq j \leq p$ et $\{\epsilon_t\}$ sont *i.i.d.* de moyenne nulle et variance δ^{-1}.

Soit $\theta = (\alpha_0 + \alpha_1, ..., \alpha_q, \beta_1, ..., \beta_q)'$ le vecteur des paramètres du modèle (3.1) et $\theta_0 = (\alpha_{00}, \alpha_{10}, ..., \alpha_{q0}, \beta_{10}, ..., \beta_{p0})'$ sa vraie valeur. Soit $\Theta \subset \mathbb{R}^{p+q+1}$ l'espace des paramètres. Définissons la fonction $h_t(\theta)$ récursivement par

$$h_t(\theta) = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}(\theta). \tag{3.3}$$

et notons $\epsilon_t(\theta) = y_t/\sqrt{h_t(\theta)}$.
Evidemment, $h_t(\theta_0) = h_t$ et $\epsilon_t(\theta_0) = \epsilon_t$.
Notez que $h_t(\theta)$ dépend du passé infini, et les valeurs initiales pour $\{y_0^2, ..., y_{1-q}^2, h_t, ..., h_{1-q}\}$ sont donc nécessaires. Pour plus de simplicité, nous les avons fixé à disons $n^{-1}\sum_{i=1}^{n} y_t^2$. Ces quantités indique les fonctions résultantes de $h_t(\theta)$ et $\epsilon_t(\theta)$ par $\widehat{h}_t(\theta)$ et $\hat{\epsilon}_t(\theta)$ respectivement. Les effets de ces valeurs initiales se révèlent négligeables asymptotiquement. Selon Peng et Yao (2003), nous définissons l'estimateur (EAM) comme étant

$$\widehat{\theta}_n = \arg\min_{\theta\in\Theta} \frac{1}{n} \sum_{t=1}^{n} |\log y_t^2 - \log \widehat{h}_t(\theta)|. \tag{3.4}$$

Malheureusement, la fonction objectif ci-dessus n'est pas convexe par rapport à ce qui rend difficile les dérivations et optimisations numériques. Peng et Yao (2003) ont établi les propriétés asymptotiques de l'estimateurs locaux (EAM) , tandis que ceux des estimateurs globaux ont été dérivés par Chen et Zhu (2015). Cela peut être un coût nécessaire pour analyser des séries chronologiques à queue lourde.

Hypothèse 3.1. **(i)** Θ *est compact et* θ_0 *est un point intérieur dans* Θ *;*

(ii) *Les polynômes* $\sum_{i=1}^{q} \alpha_i x^i$ *et* $1 - \sum_{j=1}^{p} \beta_j x^j$ *n'ont pas de racine commune ;*

(iii) $\{y_t\}$ *est strictement stationnaire et ergodique avec* $\mathbb{E}(y_t^2) < \infty$*.*

Hypothèse 3.2. **(i)** $\mathbb{E}(\epsilon_t^2) < \infty$ *et la médiane de* $z_t = \log(\epsilon_t^2)$ *est zéro ;*

(ii) *La fonction de densité de probabilité* $f(.)$ *de* z_t *est continu à zéro et satisfait* $f(0) > 0$ *;*

(iii) $\sup_{x\subset\mathbb{R}} f(x) < \infty$*.*

Remarque 3.1. L'hypothèse 1 impose des exigences de base aux processus GARCH. Pour le modèle (3.2), la condition nécessaire et suffisante pour l'existence d'une solution unique strictement stationnaire avec $\mathbb{E}(y_t^2) < \infty$ est $\sum_{i=1}^{q} \alpha_i/\delta + \sum_{j=1}^{p} \beta_j < 1$.
l'hypothèse 2 est une configuration générale pour l'estimateur(EAM). En fait, Chen et Zhu

(2015) ont établi la consistance et la normalité asymptotique de $\widehat{\theta}_n$ sous des hypothèses plus légères sur le processus $\{y_t\}$ et l'innovation $\{\epsilon_t\}$, où seul $\mathbb{E}(|y_t|^{2\iota_0}) < \infty$ pour quelque $l_0 \in (0,1)$ et $\mathbb{E}(|\epsilon_t|^{2\imath}) < 0$ pour certains $\imath > 0$ sont requis.
Cependant, nous adoptons des conditions plus fortes dans les hypothèses 1 et 2 en raison du test de Portemanteau dans la section suivante, où au moins $\mathbb{E}(y_t^2) < \infty$ et $\mathbb{E}(\epsilon_t^2) < \infty$ sont nécessaires.
$Sgn(x)$ est le signe de x et Σ désigne la matrice $(p+q+1) \times (p+q+1)$

$$\Sigma = \mathbb{E}[U_t(\theta_0)U_t'(\theta_0)], \tag{3.5}$$

où $U_t(\theta) = h_t^{-1}(\theta)\partial h_t(\theta)/\partial\theta$.
Sous les hypothèses 1 et 2, on peut montrer que

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = \frac{\Sigma^{-1}}{2f(0)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t(\theta_0)sgn(z_t) \xrightarrow{\mathcal{L}} \mathcal{N}(0, [4f^2(0)]^{-1}\Sigma^{-1}). \tag{3.6}$$

Pour estimer la variance asymptotique de $\widehat{\theta}_n$, la matrice Σ peut être remplacée par ses moyennes d'échantillon, tandis que la densité $f(0)$ doit être estimé par la méthode basée sur le noyau, qui est sensible aux largeurs de bande choisies par l'utilisateur.
On utilise alternativement la méthode du Bootstrap généralisé, et nous considérons l'estimateur (EAM) pondérée aléatoirement ci dessous de,

$$\widehat{\theta}_n^* = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{t=1}^{n} \omega_{t,n}|\log y_t^2 - \log\widehat{h}_t(\theta)|, \tag{3.7}$$

où $\{\omega_{t,n}\}$ sont des poids aléatoires non négatifs et indépendants de $\{y_t\}$.

Hypothèse 3.3. *Les conditions suivantes sur les poids aléatoires* $\{\omega_{t,n}\}$ *sont :*

(i) $W_n = (\omega_{1,n}, \ldots, \omega_{n,n})'$ *est échangeable pour chaque n, c'est-à-dire pour toute permutation* $\pi = (\pi_1, \ldots, \pi_n)$ *de* $(1, 2, \ldots, n)$*, la distribution conjointe de* $\pi(P_n) = (\omega_{\pi_1,n}, \ldots, \omega_{\pi_n,n})$ *est la même que celle de* P_n *;*

(ii) $\omega_{t,n} \geq 0$ *et* $\mathbb{E}(\omega_{t,n}) = 1$;

(iii) $\lim_{n\to\infty} \mathbb{V}(\omega_{t,n}) = \tau^2 > 0$ *et* $\mathbb{E}[(\omega_{1,n} - 1)(\omega_{2,n} - 1)] = O(n^{-1})$ *;*

(iv) $\mathbb{E}[(\omega_{t,n} - 1)^4] < \infty$ *et* $\lim_{n\to\infty} \mathbb{E}[(\omega_{1,n} - 1)^2(\omega_{2,n} - 1)^2] = \tau^4$.

Remarque 3.2. D'après l'hypothèse 3(i), il s'ensuit que, pour tout $s \neq t$,

$$\begin{aligned}\mathbb{E}[(\omega_{s,n}-1)(\omega_{t,n}-1)] &= \mathbb{E}[(\omega_{1,n}-1)(\omega_{2,n}-1)],\\ \mathbb{E}[(\omega_{s,n}-1)^2(\omega_{t,n}-1)^2] &= \mathbb{E}[(\omega_{1,n}-1)^2(\omega_{2,n}-1)^2].\end{aligned}$$

De plus, les conditions de l'hypothèse 3 sont les suivantes :
Pour l'hypothèse **3(ii)**, la condition de $\mathbb{E}(\omega_{t,n}) = 1$ est identique à la condition (3.2)dans Chatterjee et Bose (2005), tandis que Cheng (2015) a utilisé $\sum_{t=1}^{n} \omega_{t,n}$.
Les hypothèses **3(iii)** et **(iv)** proviennent des conditions (3.3) - (3.4) et la condition (3.5), également conformes aux conditions imposés par Cheng (2015).
Enfin, l'hypothèse 3 est très général, et inclut de nombreuses méthodes Bootstrap couramment utilisées en tant que cas particuliers.

Exemple 3.1. (Bootstrap à pondération aléatoire avec des poids *i.i.d.* Les poids aléatoires $\{\omega_{t,n}\}$ sont *i.i.d.* de moyenne un et de variance τ^2.
Dans les applications réelles, on peut considérer la distribution exponentielle standard , la distribution de Rademacher, qui prend la valeur de 0 ou 2 avec une probabilité de 0.5, ou la distribution uniforme $U(0.5, 1.5)$.
En particulier, le cas des poids issus de distributions exponentielles standard correspond à le Bootstrap Bayésien.

Exemple 3.2. (Bootstrap d'Efron) Les poids aléatoires $(\omega_{1,n}, ..., \omega_{n,n})$ ont une distribution multinomiale, $Mult_n(n; (n^{-1}, ..., n^{-1}))$ et $\tau^2 = 1$.

Exemple 3.3. (Double Bootstrap). Les poids aléatoires $(\omega_{1,n}, ..., \omega_{n,n})$ ont une distribution multinomiale, $Mult_n(n; (\tilde{w_1}/n, ..., \tilde{w_n}/n))$ et $\tau^2 = 1$, où $(\tilde{w_1}, ..., \tilde{w_n})$ suit $Mult_n(n; (n^{-1}, ..., n^{-1}))$.
Soit $f_t = \sigma(y_t, y_{t-1}, ...)$, et σ -l'algèbre générée par$\{y_s; s \leq t\}$.
Le théorème (2.1) donne la validité théorique du Bootstrap généralisé pour l'estimateur(EAM).

Lemme (L.1). Si l'hypothèse 1 est vérifiée, alors pour tout $k > 0$:

(i) $\mathbb{E}[\sup_{\theta\in\Theta} |\log h_t(\theta)|] < \infty$;

(ii) $\mathbb{E}[\sup_{\theta\in\Theta} \| \frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial\theta} \|^k] < \infty$;

(iii) $\mathbb{E}[\sup_{\theta\in\Theta} \| \frac{1}{h_t(\theta)} \frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta\prime} \|^k] < \infty$.

Lemme (L.2). Soit $\xi_{\rho,t} = 1 + \sum_{j=0}^{\infty} \rho^j |y_{t-j}|$ être une variable aléatoire positive dépendant d'une constante $\varrho \in (0,1)$ Si La l'hypothèse 1 est vérifiée, alors pour toute constante $\iota \in (0,1)$:

(i) $\sup_{\theta\in\Theta} \| \frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \| < C\xi_{\rho,t-1}^{\iota}$;

(ii) $\sup_{\theta\in\Theta} |\widehat{h}_t(\theta) - h_t(\theta)| \leq O(\rho^t)\xi_{\rho,0}^2$,

$\sup_{\theta\in\Theta} |\widehat{h}_t^{-1/2}(\theta) - h_t^{1/2}(\theta)| \leq O(\rho^t)\xi_{\varrho,0}$;

$\sup_{\theta\in\Theta} |\widehat{h}_t^{-1}(\theta) - h_t^{-1}(\theta)| \leq O(\rho^t)\xi_{\rho,0}^{\iota}$.

(iii)$\sup_{\theta\in\Theta} \| \frac{1}{\widehat{h}_t(\theta)} \frac{\partial \widehat{h}_t(\theta)}{\partial \theta} - \| \frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \| < O(\rho^t)\xi_{\rho,0}^{\iota}\xi_{\rho,t-1}^2$;

(iv)$\sup_{\theta\in\Theta} \| \frac{1}{\widehat{h}_t(\theta)} \frac{\partial \widehat{h}_t(\theta)}{\partial \theta \partial \theta\prime} - \| \frac{1}{h_t(\theta)} \frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta\prime} \| < O(\rho^t)\xi_{\rho,0}^{\iota}\xi_{\rho,t-1}^2$.

Lemme (L.3). Pour tout $\theta^\dagger \in \Theta$,soit $B_\eta(\theta) = \{\theta \in \Theta : \|\theta^\dagger - \theta < \eta\}$ est un voisinage ouvert de $\theta^\dagger$ de rayon $\eta > 0$. Si les hypothèses 1 à 3 sont vérifiées, conditionnez f_n, alors :

(i) $\sup_{\theta\in\Theta} |\widehat{L}_n^*(\theta) - L_n^*(\theta)| = o_p^*(1)$;

(ii) $\mathbb{E}[\sup_{\theta\in\Theta} \ell_t^*(\theta)] < \infty$;

(iii)$\mathbb{E}[\ell_t^*(\theta)]$ a un minimum unique à 0 ;

(iv) $\mathbb{E}[\sup_{\theta\in b_\eta(\theta^\dagger)} |\ell_t^*(\theta) - \ell_t^*(\theta^\dagger)|] \to 0$ quand $\eta \to 0$;

où $\widehat{L}_n^*(\theta) = n^{-1}\sum_{t=1}^n \omega_t \widehat{\ell}_t(\theta)$ et $L_n^*(\theta) = n^{-1}\sum_{t=1}^n \omega_t \ell_t(\theta)$,

avec $\widehat{\ell}_t(\theta) = |\log y_t^2 - \log \widehat{h}_t(\theta)|$, $\ell_t(\theta) = |\log y_t^2 - \log h_t(\theta)|$

et $\ell_t^*(\theta) = \omega_t \ell_t(\theta)$.

Lemme (L.4). Si les hypothèse 1 à 3 sont vérifiées, alors $\widehat{\theta}_n^* - \theta = o_p^*(1)$

Théorème 3.1. *Supposons que les propositions 1 à 3 soient vérifiées. Alors,*

$$(\sqrt{n}/\tau)(\widehat{\theta}_n^* - \widehat{\theta}_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, [4f^2(0)]^{-1}\Sigma^{-1})$$

lorsque $n \to \infty$ où Σ est défini comme dans (3.5).
En ce qui concerne l'estimateur global $\widehat{\theta}_n$ de (EAM) en (3.4) dans Chen et Zhu (2015), la clé de la preuve du théorème ci-dessus est établir la $\sqrt{n}$-consistance de $\widehat{\theta}_n^*$; sans confusion, dans ce que suit , nous utiliserons les notations de $\{\omega_t\}$, plutôt que $\{\omega_{t,n}\}$, par souci de simplicité.

Preuve. Pour prouver la validité de l'estimateur Bootstrap, nous utilisons des arguments similaires à ceux de lemme (L.2) dans Chen et Zhu (2015). Définir :

$$\widehat{L}_n^*(\theta) = \frac{1}{n}\sum_{i=1}^{n}\omega_t|\log y_t^2 - \log\widehat{h}_t(\theta)| \; et \; H_n^*(u) = n[\widehat{L}_n^*(\theta_0+u) - \widehat{L}_n^*(\theta_0)];$$

où $u \in \Lambda = \{u : u+\theta_0 \in \Theta\}$.
Notez que $m_t = \log[\widehat{h}_t(\theta_0)/h_t(\theta_0)], z_{1t}(s) = I(z_t < s) - I(z_t > s)$
et $z_{2t}(s) = I(z_t \leq s) - I(z_t \leq 0)$, où $z_t = \log\epsilon_t^2$. On note $\log y_t^2 = z_t + \log h_t(\theta_0)$. De plus, pour $x \neq 0$,

$$|x-y| - |x| = -y[I(x>0) - I(x<0)] + 2\int_0^y [I(x \leq s) - I(x \leq 0)]\mathrm{d}s, \qquad \text{(A.1)}$$

voir Knight (1998). Notons $\widehat{H}_n^*(u) = \sum_{t=1}^n \widehat{A}_t^*(u)$.
Puis en (A.1), il s'ensuit que

$$\widehat{A}_t^*(u) = \omega_t\widehat{q}_t(u)Z_{1t}(m_t) + 2\omega_t\int_0^{\widehat{q}_t(u)}[Z_{2t}(s+m_t) - Z_{2t}(m_t)]\mathrm{d}s, \qquad \text{(A.2)}$$

où $\widehat{q}_t(u) = \log\widehat{h}_t(\theta_0+u) - \log\widehat{h}_t(\theta_0)$ utilisant le développement de Taylor, nous avons $\widehat{q}_t(u) = \widehat{q}_{1t}(u) + \widehat{q}_{2t}(u)$, où $\widehat{q}_{1t}(u) = \frac{u'}{\widehat{h}_t(\theta_0)}\frac{\partial\widehat{h}_t(\theta_0)}{\partial\theta}$ et

$$\widehat{q}_{2t}(u) = \frac{1}{2}u'\left[\frac{1}{\widehat{h}_t(\theta^*)}\frac{\partial^2\widehat{h}_t(\theta^*)}{\partial\theta\partial\theta'} - \frac{1}{\widehat{h}_t^2(\theta^*)}\frac{\partial\widehat{h}_t(\theta^*)}{\partial\theta}\frac{\partial\widehat{h}_t(\theta^*)}{\partial\theta'}\right]u,$$

avec θ^* compris entre θ_0 et $\theta_0 + u$ Notons

$$A_t^0(u) = \omega_t q_t(u)Z_{1t}(0) + 2\omega_t\int_0^{q_t(u)} Z_{2t}(s)\mathrm{d}s \; et$$
$$A_t^*(u) = \omega_t q_t(u)Z_{1t}(m_t) + 2\omega_t\int_0^{q_t(u)}[Z_{2t}(s+m_t) - Z_{2t}(m_t)]\mathrm{d}s,$$

où $q_t(u) = q_{1t}(u) + q_{2t}(u)$ avec $q_{1t}(u) = \frac{u'}{h_t(\theta_0)}\frac{\partial h_t(\theta_0)}{\partial\theta}$ avec $q_{2t}(u)$ telle que :

$$q_{2t}(u) = \frac{1}{2}u'\left[\frac{1}{h_t(\theta^*)}\frac{\partial^2 h_t(\theta^*)}{\partial\theta\partial\theta'} - \frac{1}{h_t^2(\theta^*)}\frac{\partial h_t(\theta^*)}{\partial\theta}\frac{\partial h_t(\theta^*)}{\partial\theta'}\right]u.$$

Dénoter

$$\sum_{t=1}^n \widehat{A}_t^*(u) = \sum_{t=1}^n A_t^*(u) + O_p^*(\|u\|) \qquad \text{(A.3a)}$$
$$= \sum_{t=1}^n A_t^0(u) + O_p^*(\|u\|), \qquad \text{(A.3b)}$$

où les égalités ci-dessus sont identiques dans $u \in \Lambda$.

Nous prouvons d'abord (A.3a). Depuis $|z_{1t}(s)| \leq 1$ et $|z_{2t}(s)| \leq 1$, avec le lemme (L.2) et $\mathbb{E}(\omega_t) = 1$ par l'hypothèse 3, on peut montrer que

$$\begin{aligned}\mathbb{E}^*\left[\sup_{u\in\Lambda}\frac{1}{\|u\|}\left|\sum_{t=1}^{n}\widehat{A}_t^*(u) - \sum_{t=1}^{n}A_t^*(u)\right|\right] &\leq 3\mathbb{E}^*\left[\sum_{t=1}^{n}\omega_t\sup_{u\in\Lambda}\frac{1}{\|u\|}|\widehat{q}_t^*(u) - q_t^*(u)|\right]\\ &\leq \sum_{t=1}^{n}O(\rho^t)\xi_{\rho,0}^{\iota}\xi_{\rho,t-1}^2 = O^*(1).\end{aligned} \quad (A.4)$$

Pour certains $\rho \in (0,1)$ et $\iota \in (0,1)$. Par conséquent, (A.3a) est maintenu par (A.4).

Nous considérons ensuite (A.3b). Notez que $m_t \in F_{t-1}$, et $q_t(\mu) \in F_{t-1}$ où t est le champ généré par $\{y_s, s \leq t\}$.

Puis par (S.12) dans la preuve du lemme (S.3), $|m_t| \leq O(\rho^t)\xi_{\rho,0}^{\iota}\xi_{\rho,t-1}^2$ est valide.

De plus, par la loi des attentes itérées l'inégalité de Hölder, ainsi que les hypothèses à 3 et les lemmes (L.1) et (L.2), on peut vérifier que, là existe un $\iota_0 \in (0,1)$ tel que

$$\begin{aligned}\mathbb{E}\left\{\sup_{u\in\Lambda}\frac{1}{\|u\|}\sum_{t=1}^{n}|q_t(u)[Z_{1t}(m_t) - Z_{1t}(0)]|\right\}^{\iota_0} &\leq C\mathbb{E}\left\{\sum_{t=1}^{n}\{\mathbb{E}[|Z_{1t}(m_t) - Z_{1t}(0)||F_{t-1}]\}^{\iota_0}\right\}\\ &\leq C\mathbb{E}\left\{\sum_{t=1}^{n}[|m_t|\sup_{x\in\mathbb{R}}f(x)]^{\iota_0}\right\}\\ &\leq \sum_{t=1}^{n}O(\rho^{\iota_0 t})\mathbb{E}(\xi_{\rho,0}^{u_o}\,,\,\xi_{\rho,t-1}^{2\iota_0}) < \infty,\end{aligned} \quad (A.5)$$

ce qui implique que :

$$\sup_{u\in\Lambda}\frac{1}{\|u\|}\sum_{t=1}^{n}|q_t(u)[z_{1t}(m_t) - z_{1t}(0)]| = O_p(1). \quad (A.6)$$

Semblable à la preuve de (A.5), avec l'hypothèse 3, nous pouvons montrer que :

$$\begin{aligned}&\mathbb{E}\left\{\sup_{u\in\Lambda}\frac{1}{\|u\|}\sum_{t=1}^{n}|\omega_t - 1||q_t(u)[z_{1t}(m_t) - z_{1t}(0)]|\right\}^{\iota_0}\\ &\leq C\mathbb{E}\left\{\mathbb{E}(|\omega_t - 1|)^{\iota_0}\sum_{t=1}^{n}\mathbb{E}[|z_{1t}(m_t) - z_{1t}(0)|^{\iota_0}|F_{t-1}]\right\} < \infty,\end{aligned}$$

ce qui implique que :

$$\sup_{u\in\Lambda}\frac{1}{\|u\|}\sum_{t=1}^{n}|(\omega_t - 1|)q_t(u)[z_{1t}(m_t) - z_{1t}(0)]| = O_p(1). \quad (A.7)$$

Par conséquent,en peignant (A.6) et (A.7), il s'ensuit que

$$\sum_{t=1}^{n} \omega_t q_t(u)[z_{1t}(m_t) - z_{1t}(0)] = O_p^*(\|u\|). \tag{A.8}$$

Semblable à la preuve de (A.6), nous pouvons montrer que :

$$\sup_{u\in\Lambda} \frac{1}{\|u\|} \sum_{t=1}^{n} \left| \int_0^{q_t(u)} [z_{2t}(s+m_t) - z_{2t}(m_t) - z_{2t}(s)]\mathrm{d}s \right| = O_p(1).$$

En ce qui concerne (A.7), on peut montrer que :

$$\sup_{u\in\Lambda} \frac{1}{\|u\|} \sum_{t=1}^{n} \left| (\omega_t - 1) \int_0^{q_t(u)} [z_{2t}(s+m_t) - z_{2t}(m_t) - z_{2t}(s)]\mathrm{d}s \right| = O_p(1).$$

Par conséquent,

$$\sum_{t=1}^{n} \omega_t \int_0^{q_t(u)} [z_{2t}(s+m_t) - z_{2t}(m_t) - z_{2t}(s)]\mathrm{d}s = O_p^*(\|u\|). \tag{A.9}$$

Notez que

$$\begin{aligned}\sum_{t=1}^{n} [A_t^*(u) - A_t^0(u)] = & \sum_{t=1}^{n} \omega_t q_t(u)[z_{1t}(m_t) - z_{1t}(0)] \\ & + 2\sum_{t=1}^{n} \omega_t \int_0^{q_t(u)} [z_{2t}(s+m_t) - z_{2t}(m_t) - z_{2t}(s)]\mathrm{d}s.\end{aligned}$$

Ceci avec (A.8) et (A.9), alors (A.3b) est valable.
De plus, nous considérons la décomposition de $\sum_{t=1}^{n} A_t^0(u)$ On peut vérifier que :

$$\sum_{t=1}^{n} A_t^0(u) = -\sqrt{n}u' T_{1n}^* + T_{2n}^*(u) + R_n^*(u), \tag{A.10}$$

où

$$T_{1n}^* = \frac{1}{n} \sum_{t=1}^{n} \omega_t \frac{sgn(Z_t)}{h_t(\theta_0)} \frac{\partial h_t(\theta_0)}{\partial \theta}, T_{2n}^*(u) = \sum_{t=1}^{n} \omega_t \xi_t(u).$$

$$R_n^*(u) = -\sum_{t=1}^{n} \omega_t q_{2t}(u) sgn(Z_t) + 2\sum_{t=1}^{n} \omega_t \int_{q_{1t}(u)}^{q_t(u)} Z_{2t}(s)\mathrm{d}s,$$

avec $\xi_t(u) = 2\int_0^{q_{1t}(u)} Z_{2t}(s)\mathrm{d}s$. Pour $R_n^*(u)$, il s'ensuit que

$$R_n^*(u) = R_{1n}(u) + R_{2n}^*(u), \tag{A.11}$$

où

$$R_{1n}(u) = -\sum_{t=1}^{n} q_{2t}(u)sgn(Z_t) + 2\sum_{t=1}^{n}\int_{q_{1t(u)}}^{q_t(u)} Z_{2t}(s)\mathrm{d}s,$$
$$R_{2n}^*(u) = -\sum_{t=1}^{n}(\omega_t - 1)q_{2t}(u)sgn(z_t) + 2\sum_{t=1}^{n}(\omega_t - 1)\int_{q_t(u)}^{q_{1t}(u)} Z_{2t}(s)\mathrm{d}s.$$

D'après la preuve du lemme (A.2) dans Chen et Zhu (2015), pour $u = o_p(1)$,

$$R_{1n}(u) = o_p(n\|u\|^2). \tag{A.12}$$

Ensuite, on considére $R_{2n}^*(u)$. Soit

$$S_{1n}(\theta) = \frac{1}{n}\sum_{t=1}^{n} = \frac{\omega_t - 1}{h_t(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}sgn(Z_t),$$
$$S_{2n}(\theta) = \frac{1}{n}\sum_{t=1}^{n} = \frac{\omega_t - 1}{h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}sgn(Z_t).$$

Par l'hypothèse 3, il en résulte que $\mathbb{E}^*[S_{1n}(\theta)] = \mathbb{E}^*[S_{2n}(\theta)] = 0$.
De plus, par la loi des attentes itérées et l'inégalité de Cauchy-Schwartz, ainsi que les hypothèses 1 à 3 et le lemme (A.1), permettent de vérifier que :

$$\begin{aligned}
\mathbb{E}[S_{1n}(\theta)S_{1n}'(\theta)] &= \frac{1}{n^2}\mathbb{E}\left\{\sum_{t=1}^{n}\mathbb{E}^*[(\omega_t-1)^2]\left[\frac{1}{h_t(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\right]^2\right\} \\
&+ \frac{1}{n^2}\mathbb{E}\left\{\sum_{t\neq s}\mathbb{E}^*[(\omega_t-1)(\omega_s-1)]\left[\frac{1}{h_t(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{1}{h_s(\theta)}\frac{\partial^2 h_s(\theta)}{\partial\theta\partial\theta'}\right]sgn(Z_t)sgn(Z_s)\right\} \\
&\leq \frac{\tau^2}{n^2}\sum_{t=1}^{n}\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{1}{h_t(\theta)}\|\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\right\|^2\right] \\
&+ \frac{C}{n^3}\sum_{t\neq s}\left\{\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{1}{h_t(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\right\|^2\right]\right\}^{1/2}\left\{\mathbb{E}\left[|\sup_{\theta\in\Theta}\left\|\frac{1}{h_s(\theta)}\frac{\partial^2 h_s(\theta)}{\partial\theta\partial\theta'}\right\|^2\right]\right\}^{1/2} \\
&= O(n^{-1}).
\end{aligned}$$

Cela implique que $S_{1n}(\theta) = o_p^*(1)$. De même, nous pouvons montrer que $S_{2n}(\theta) = o_p^*(1)$. ensuite

$$\sum_{t=1}^{n}(\omega_t - 1)q_{2t}(u)sgn(Z_t) = u'[S_{1n}(\theta^*) - S_{2n}(\theta^*)]u = o_p^*(n\|u\|^2). \tag{A.13}$$

De même comme pour (A.13), il peut être prouvé que

$$\sum_{t=1}^{n}(\omega_t-1)\int_{q_t(u)}^{q_{1t}(u)} Z_{2t}\mathrm{d}s = o_p^*(n\|u\|^2).$$

Par conséquent, $R_{2n}(u) = o_p^*(n\|u\|^2)$. Ceci avec (A.11) et (A.12) implique que :

$$R_n^*(u) = o_p^*(n\|u\|^2). \tag{A.14}$$

Pour $T_{2n}^*(u)$, il s'ensuit que :

$$T_{2n}^*(u) = \sum_{t=1}^{n}\xi_t(u) + 2\sum_{t=1}^{n}(\omega_t-1)\int_{0}^{q_{1t}(u)} Z_{2t}\mathrm{d}s. \tag{A.15}$$

Pour $u = o_p(1)$, Chen et Zhu (2015) ont montré que :

$$\begin{aligned}\sum_{t=1}^{n}\xi_t(u) &= \sum_{t=1}^{n}\mathbb{E}[\xi_t(u)]|F_{t-1} + \sum_{t=1}^{n}\{\xi_t(u) - \mathbb{E}[\xi_t(u)|F_{t-1}]\}\\ &= (\sqrt{n}u')[f(0)\Sigma](\sqrt{n}u) + o_p(\sqrt{n}\|u\| + n\|u\|^2).\end{aligned} \tag{A.16}$$

Similaire à la preuve de (A.16), de l'hypothèse3 et du lemme (A.1), nous avons

$$2\sum_{t=1}^{n}(\omega_t-1)\int_{0}^{q_{1t}(u)} Z_{2t}\mathrm{d}s = o_p^*(\sqrt{n}\|u\| + n\|u\|^2),$$

avec (A.15) et (A.16) implique que :

$$T_{2n}^*(u) = (\sqrt{n}u')[f(0)\Sigma](\sqrt{n}u) + o_p^*(\sqrt{n}\|u\| + n\|u\|^2). \tag{A.17}$$

En combinant (A.3a), (A.3b), (A.10), (A.14) et (A.17), nous avons

$$H_n^*(u) = -\sqrt{n}u'T_{1n}^* + (\sqrt{n}u')[f(0)\Sigma](\sqrt{n}u) + o_p^*(\sqrt{n}\|u\| + n\|u\|^2). \tag{A.18}$$

Soit $\widehat{u}_n^* = \widehat{\theta}_n^* - \theta_0$. D'après le lemme (A.4), il est dit que $\widehat{u}_n^* = o_p^*(1)$. Par $(A.18)$, alors nous avons

$$\begin{aligned}H_n^*(\widehat{u}_n^*) &= -\sqrt{n}\widehat{u}_n^{*\prime}T_{1n}^* + (\sqrt{n}\widehat{u}_n^{*\prime}T_{1n}^*)[f(0)\Sigma](\sqrt{n}\widehat{u}_n^*\sqrt{n}\widehat{u}_n^*) + o_p^*(\sqrt{n}\|\widehat{u}_n^*\| + n\|\widehat{u}_n^*\|^2)\\ &\geq -\sqrt{n}\|\widehat{u}_n^*\|[\|T_{1n}^*\| + o_p^*(1)] + n\|\widehat{u}_n^*\|^2[\lambda_{\min} + o_p^*(1)],\end{aligned} \tag{A.19}$$

où $\lambda_{\min}$ est la plus petite valeur propre de $f(0)\Sigma$. Par l'hypothèse 3, il s'ensuit que :

$$\mathbb{E}^*(T_{1n}^*) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{sgn(Z_t)}{h_t(\theta_0)}\frac{\partial h_t(\theta_0)}{\partial\theta} \xrightarrow{\mathcal{L}} \mathcal{N}(0.\Sigma),$$

ce qui implique que $T_{1n}^* = O_p^*(1)$. De plus, depuis $\widehat{\theta}_n^*$ minimise $\widehat{L}_n^*(\theta)$, alors $\widehat{u}_n^*$ est le minimiseur de $H_n^*(u)$ tel que $H_n^*(\widehat{u}_n^*) \leq 0$. Ceci, associé à (A.19), implique que :

$$\sqrt{n}\|\widehat{u}_n^*\| \leq [\lambda_{\min} + o_p^*(1)]^{-1}[\|T_{1n}^*\| + o_p^*(1)] = O_p^*(1). \tag{A.20}$$

Soit $u_n^* = [2f(0)\Sigma]^{-1}T_{1n}^*/\sqrt{n}$ Semblable à la preuve du théorème (2.2) dans Zhu et Ling (2011), par (A.18) et (A.20), on peut vérifier que :

$$\begin{aligned} H_n(\widehat{u}_n^*) &= -2\sqrt{n}\widehat{u}_n^{*\prime}[2f(0)\Sigma]\sqrt{n}u_n^* + \sqrt{n}\widehat{u}_n^{*\prime}[2f(0)\Sigma]\sqrt{n}u_n^* + o_p^*(1), \\ H_n^*(u_n^*) &= -\sqrt{n}u_n^{*\prime}[f(0)\Sigma]\sqrt{n}u_n^* + o_p^*(1). \end{aligned}$$

Puis il s'ensuit que :

$$\begin{aligned} H_n^*(\widehat{u}_n^*) - H_n^*(u_n^*) &= (\sqrt{n}\widehat{u}_n^* - \sqrt{n}u_n^*)'[f(0)\Sigma](\sqrt{n}\widehat{u}_n^* - \sqrt{n}u_n^*) + o_p^*(1) \\ &\geq \lambda_{\min}\|\sqrt{n}\widehat{u}_n^* - \sqrt{n}u_n^*\|^2 + o_p^*(1). \end{aligned}$$

Ceci avec $H_n^*(\widehat{u}_n^*) - H_n^*(u_n^*) \leq 0$, implique que :

$$\sqrt{n}(\widehat{\theta}_n^* - \widehat{\theta}_n) = \frac{\Sigma^{-1}}{2f(0)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\omega_t U_t(\theta_0)sgn(z_t) + o_p^*(1).$$

Donc $\sqrt{n}(\widehat{\theta}_n^* - \theta_0)$ Ceci, associé à (3.6), implique que :

$$\sqrt{n}(\widehat{\theta}_n^* - \widehat{\theta}_n) = \frac{\Sigma^{-1}}{2f(0)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\omega_t - 1)U_t(\theta_0)sgn(z_t) + o_p^*(1). \tag{A.21}$$

Ensuite, par le lemme (4.6) de Praestgaard et Wellner (1993) et la l'hypothèse3, il s'ensuit que

$$(\frac{\sqrt{n}}{\tau}(\widehat{\theta}_n^* - \widehat{\theta}_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, [4f^2(0)]^{-1}\Sigma^{-1}).$$

3.2 Test du Portemanteau par Bootstrap

Le portemanteau est un test qui évalue la corrélation existant entre les résidus.
Soit $\epsilon = (\epsilon_1, ..., \epsilon_n)$ une famille de variable aléatoires de loi à densité. On souhaite tester $H_0 : \{\epsilon\}$ est un bruit blanc fort contre $H_1 : \{\epsilon\}$ n'est pas un bruit blanc fort. On note

$$\hat{\rho}_\epsilon(h) = \frac{\frac{1}{n-h}\sum_{i=1}^{n-h}\epsilon_i\epsilon_{i+h}}{\frac{1}{n}\sum_{i=1}^{n}\epsilon_i^2}.$$

Il s'agit d'un estimateur de la corrélation. En effet, de même que $\frac{1}{n}\sum_{i=1}^{n}\epsilon_i^2$ est un estimateur de la variance $\mathbb{V}(\epsilon_t) = \mathbb{E}(\epsilon_t^2)$, de même $\sum_{i=1}^{n-h}\epsilon_i\epsilon_{i+h}$ est un estimateur de $cov(\epsilon_t, \epsilon_{t+h}) = \mathbb{E}(\epsilon_t\epsilon_{t+h})$. Notons que sous H_0 cette covariance est nulle si $h \neq 0$.
Sous H_0, pour n grand et $k < n/4, (\hat{\rho}_\epsilon(h), h = 1, ..., k)$ est approximativement un bruit blanc gaussien de variance $\frac{1}{n}$. Il s'ensuit que

$$Q_{BP} = n\sum_{h=1}^{k}\hat{\rho}_\epsilon^2(h)$$

suit approximativement une loi du χ^2 à k degrés de liberté. Une trop grande valeur de Q_{BP} indiquerait une certaine corrélation entre les ϵ_i. On en déduit un test au niveau α de H_0 de zone de rejet $Q_{BP} > \chi^2_{1-\alpha}(h)$, avec le quantile d'ordre α de la loi du χ^2 à k degrés de liberté. En fait, on peut améliorer sensiblement ce test en considérant plutôt α la statistique

$$Q_{BP} = n(n+2)\sum_{h=1}^{k}\frac{\hat{\rho}_\epsilon^2(h)}{n-h},$$

qui, sous H_0, suit également, pour n grand, une loi du χ^2 à k degrés de liberté.
Ce test, sous sa première forme, est aussi appelé test de Box-Pierce ; sous sa seconde, il est connu sous le nom de test de Ljung-Box.

Pour une série temporelle observée $\{y_1, ..., y_n\}$, cette section examine les statistiques de test de Portmanteau pour vérifier si le modèle GARCH en (3.1) est correctement spécifié, c'est-à-dire pour vérifier si $\{h_t\}$ en (3.1) avec un certain ordre de (q,p) est suffisant pour interpréter la variance conditionnelle $\mathbb{V}(y_t|F_{t-1})$. Dans la littérature, cela peut être réalisé en vérifiant la signification de la fonction d'autocorrélation de l'échantillon de résidus absolus.
Soit $\mu = \mathbb{E}(|\epsilon_t|)$, et notons $\rho_k = \mathbb{E}[(|\epsilon_t| - \mu)(|\epsilon_{t-k}| - \mu)]/\sigma^2$ la fonction d'autocorrélation de $\{|\epsilon_t|\}$, où $\sigma^2 = \mathbb{V}(|\epsilon_t|) = \mathbb{E}(|\epsilon_t| - \mu)^2$.
Nous allons tester les hypothèses suivantes :
$H_0 : \rho_k = 0$ pour tout $k \geq 1$ vs $\quad H1 : \rho_k \neq 0$ pour au moins un $k \geq 1$
Pour le modèle GARCH ajusté par l'estimateur (EAM) de la section précédente, les résidus correspondants $\{\widehat{\epsilon}_t\}$ peuvent être défini comme étant $\widehat{\epsilon}_t = \widehat{\epsilon}_t(\widehat{\theta}_n)$, et l'autocorrélation

résiduelle absolue de l'échantillon k est de

$$\widehat{\rho}_k = \frac{\sum_{t=k+1}^{n}(|\widehat{\epsilon}_t| - \widehat{\mu})(|\widehat{\epsilon}_{t-k}| - \widehat{\mu})}{\sum_{t=1}^{n}(|\widehat{\epsilon}_t| - \widehat{\mu})^2},\ k \geq 1, \tag{3.8}$$

où $\widehat{\mu} = \sum_{t=1}^{n}(|\widehat{\epsilon}_t|$ et $\widehat{\rho}_k$ est clairement la version exemple de σ_k.
Soit M un entier positif prédéterminé, et notons $\widehat{\rho} = (\widehat{\rho}_1, ..., \widehat{\rho}_M)'$. Sous l'hypothèse que le modèle est correctement spécifiés,

$$\sqrt{n}\widehat{\rho} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Omega), \tag{3.9}$$

où

$$\Omega = I_M + \frac{\mu - 8\mathrm{d}(0)}{4\mu\sigma^4 f^2(0)} H\Sigma^{(-1)}H', \tag{3.10}$$

où I_M est une matrice d'identité de $M \times M$, $d = \mathbb{E}\{|\epsilon_t|[I(|\epsilon_t| > 1) - I(|\epsilon_t| < 1)]\}$, et $H = (H_1, ..., H_M)'$ représente un $M \times (p + q + 1)$ matrice dont le k-ième élément est

$$H_k = \mathbb{E}[0.5\mu h_t^{-1} \partial h_t(\theta_0)/\partial\theta(|\epsilon_{t-k}| - \mu)].$$

Supposons que $\widehat{\Omega}$ soit un estimateur consistant de Ω, et nous pouvons ensuite construire une statistique de test de portemanteau ci-dessous,

$$Q(M) = n\widehat{\rho}'\,\widehat{\Omega}^{-1}\widehat{\rho} \xrightarrow{\mathcal{L}} \chi^2_M, \tag{3.11}$$

où χ^2_M est la distribution chi-carré à M degrés de liberté.

Le résultat en (3.9) peut être utilisé pour vérifier la signification des autocorrélations absolues individuellement, alors que at (3.11) fournit des justifications théoriques pour tester conjointement la signification des premières M autocorrélations absolues. Cependant, la covariance asymptotique Ω dépend de la valeur de $f(0)$, et certaines méthodes non paramètriques ne sont pas évitable de fournir une estimation.

Définir plus précisément l'autocorrélation résiduelle d'amorçage absolue au décalage k inférieur ,

$$\widehat{\rho}^*_k = \frac{\sum_{t=k+1}^{n}\omega_t(|\widehat{\epsilon_t}^*| - \widehat{\mu}^*)(|\widehat{\epsilon}^*_{t-k}| - \widehat{\mu}^*)}{\sum_{t=1}^{n}(|\widehat{\epsilon_t}^*|^2}, \tag{3.12}$$

où $\{\omega_t\}$ sont définis comme dans la section 2, $\widehat{\epsilon^*_t} = \widehat{\epsilon}_t(\theta^*_t)$ et $\widehat{\mu^*} = n_{-1}\sum_{t=1}^{n}(|\widehat{\epsilon^*_t}|$ avec $\widehat{\theta^*_n}$ étant l'estimateur Bootstrap obtenu par (3.7). Soit $\widehat{\rho^*} = (\widehat{\rho}^*_1, ..., \widehat{\rho}^*_M)'$.

Théorème 3.2. *Supposons que les hypothèses 1 à 3 soient vérifiées et que le modèle présenté en (3.2) soit correctement spécifié. Alors*

$$(\sqrt{n}/\tau)(\widehat{\rho}^* - \widehat{\rho}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Omega),$$

lorsque $n \to \infty$, où Ω est défini comme dans (3.10).
Nous pouvons construire la statistique de test du Portemanteau, $O(M) = n\rho'(\widehat{\Omega^*}^{-1}\widehat{\rho})$ où $\widehat{\Omega^*}$ est la covariance de matrice Bootsrap de $(\sqrt{n}/\tau)(\widehat{\rho}^* - \widehat{\rho})$, qui peuvent être utilisés pour tester l'adéquation du modèle ajusté. De plus, le ci-dessus, ainsi que le résultat en (3.9), mèneront aux tests individuels, qui peuvent indiquer comment réviser le modèle GARCH si le test de Portemanteau est rejeté. Plus précisément, la procédure d'amorçage généralisée pour l'estimateur (EAM) et le test de Portemanteau peuvent être résumés ci-dessous :

1. *Générez des pondérations aléatoires $\{\omega_t\}$ à partir d'une distribution non négative satisfaisant l'hypothèse 3. Ensuite, obtenez le estimateur $\widehat{\theta}_n^*$ de (EAM) par (3,4).*
2. *Calculez l'autocorrélation des résidus du Bootstrap $\widehat{\rho}_k^*$ par (3.8) puis $\widehat{\rho}^*$.*
3. *Calculez $\mathbb{E}^{(1)} = (\sqrt{n}/\tau)(\widehat{\theta}_n^* - \widehat{\theta}_n)$ et $T^{(1)} = (\sqrt{n}/\tau)(\widehat{\rho}^* - \widehat{\rho})$. Répétez les étapes 1 et 2 pour B-1 fois et obtenez $\{\mathbb{E}^{(1)}, ..., \mathbb{E}^{(B)}\}$ et $\{T^{(1)}, ..., T^{(B)}\}$, où B est un nombre suffisamment grand. Puis les distributions empiriques de $\{\mathbb{E}^{(b)}\}_{b=1}^B$ et $\{T^{(b)}\}_{b=1}^B$ peut être utilisé pour approximer les distributions asymptotiques de $\sqrt{n}(\widehat{\theta}_n - \theta_n)$ et $\sqrt{n}\widehat{\rho}$ respectivement*

Remarque 3.3. Pour les poids aléatoires, nous avons beaucoup de choix satisfaisant la proposition 3 et, comme mentionné dans section 1, le Bootstrap généralisé inclut même de nombreuses méthodes de Bootstrap couramment utilisées en tant que cas particuliers. Comme le montrent les résultats de la simulation à la section 3, ces différentes pondérations aléatoires ont des performances similaires. En conséquence, dans les applications réelles, nous pouvons simplement utiliser le fichier *i.i.d.* poids aléatoires avec moyenne et variance égales à une.

Preuve. Rappelons que $\mu = \mathbb{E}(|\epsilon_t|), \sigma^2 = \mathbb{V}(|\epsilon_t|)$ et $\widehat{\mu} = n^{-1}\sum_{t=1}^n |\widehat{\epsilon}_t|$.
Tout d'abord, nous allons montrer que :

$$\widehat{\mu}^* = \mu + o_p(1) \quad et \quad \frac{1}{n}\sum_{t=1}^n (|\widehat{\epsilon}_t^*| - \widehat{\mu}^*)^2 = \sigma^2 + o_p(1). \tag{A.22}$$

Notez que $\epsilon_t = \epsilon_t(\theta_0)$ et $\widehat{\epsilon}_t^* = \widehat{\epsilon}_t(\widehat{\theta}_n^*)$. Selon la loi des grands nombres,

$$\begin{aligned}\widehat{\mu}^* - \mu &= \frac{1}{n}\sum_{t=1}^{n}[|\widehat{\epsilon}_t(\widehat{\theta}_n^*)| - |\epsilon_t(\theta_0)|] + o_p(1) \\ &= \frac{1}{n}\sum_{t=1}^{n}[|\widehat{\epsilon}_t(\widehat{\theta}_n^*)| - |\epsilon_t(\widehat{\theta}_n^*)|] + \frac{1}{n}\sum_{t=1}^{n}[|\epsilon_t(\widehat{\theta}_n^*) - |\epsilon_t(\theta_0)|] + o_p(1). \qquad \text{(A.23)}\end{aligned}$$

Par le lemme (L.2), nous avons

$$\sup_{\theta\in\Theta}||\widehat{\epsilon}_t(\theta)| - |\epsilon_t(\theta)|| = |y_t|\sup_{\theta\in\Theta}|\widehat{h}_t^{-1/2}(\theta) - h_t^{-1/2}(\theta)| \leq O(\rho^t)\xi_{\rho,0}|y_t|,$$

ce qui implique que :

$$\frac{1}{n}\sum_{t=1}^{n}[|\widehat{\epsilon}_t(\widehat{\theta}_n^*)| - |\epsilon_t(\widehat{\theta}_n^*)|] \leq \frac{1}{n}\sum_{t=1}^{n} O(\rho^t)\xi_{\rho,0}|y_t| = o_p(1). \qquad \text{(A.24)}$$

De plus, par l'extension de Taylor (S.5) et le fait que $\sup_{\theta\in\Theta} h_t(\theta) \geq \underline{\omega}$, il s'ensuit que :

$$\begin{aligned}||\epsilon_t(\widehat{\theta}_n^*)| - |\epsilon_t(\theta_n)|| &= |y_t|h_t^{1/2}(\widehat{\theta}_n^*)h_t^{-1/2}(\theta_0)|h_t^{1/2}(\widehat{\theta}_n^*) - h_t^{1/2}(\theta_0)| \\ &\leq \frac{|y_t|}{2\omega}\|(\widehat{\theta}_n^*) - \theta_0\|\sup_{\theta\in\Theta}\left\|\frac{1}{h_t(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\right\| \\ &\leq C\|(\widehat{\theta}_n^*) - \theta_0\|\xi_{\rho,t-1}^{\iota}|y_t|.\end{aligned}$$

Ceci, ajouté à l'inégalité de Hölder, au théorème ergodique, la l'hypothèse 1, au lemme (A.2)(i) et au fait que $\widehat{\theta}_n^* - \theta_0 = o_p(1)$ de Lemme (A.4), implique que :

$$\frac{1}{n}\sum_{t=1}^{n}(|\epsilon_t(\widehat{\theta}_n^*)| - |\epsilon_t(\theta_0)|] \leq C\|\widehat{\theta}_n^* - (\theta_0\|\frac{1}{n}\sum_{t=1}^{n}\xi_{\rho,t-1}\iota|y_t| = o_p(1) \qquad \text{(A.25)}$$

En combinant (A.23) - (A.25), nous avons $\widehat{\mu}^* = \mu + o_p(1)$. De même, nous pouvons montrer que

$$\frac{1}{n}\sum_{t=1}^{n}(|\widehat{\epsilon}_t^*| - \widehat{\mu}^*)^2 = \frac{1}{n}\sum_{t=1}^{n}\epsilon_t^2 - \mu^2 + \frac{1}{n}\sum_{t=1}^{n}(\widehat{\epsilon}_t^* + \epsilon_t)(\widehat{\epsilon}_t^* - \epsilon_t) + o_p(1) = \sigma^2 + o_p(1).$$

Par conséquent, (A.22) est vérifié. Ensuite, nous montrons la représentation Bahadur de

$\widehat{\rho}_k^*$. Définir les fonctions suivantes comme suit :

$$\widehat{C}_k(\theta) = \frac{1}{n}\sum_{t=k+1}^{n}[|\widehat{\epsilon}_t(\theta)| - \mu][|\widehat{\epsilon}_{t-k}(\theta)| - \mu],$$

$$C_k(\theta) = \frac{1}{n}\sum_{t=k+1}^{n}[|\epsilon_t(\theta)| - \mu][|\epsilon_{t-k}(\theta)| - \mu],$$

$$\widehat{C}_k^*(\theta) = \frac{1}{n}\sum_{t=k+1}^{n}\omega_t[|\widehat{\epsilon}_t(\theta)| - \mu][|\widehat{\epsilon}_{t-k}(\theta)| - \mu],$$

$$C_k^*(\theta) = \frac{1}{n}\sum_{t=k+1}^{n}\omega_t[|\epsilon_t(\theta)| - |\mu][|\epsilon_{t-k}(\theta)| - \mu].$$

L'hypothèse3 implique que $\mathbb{E}^*(\omega_t) = 1$, puis conditionnée à F_n

$$\mathbb{E}^*\{\sqrt{n}[\widehat{C}_k^*(\widehat{\theta}_n^*) - C_k^*(\widehat{\theta}_n^*)] = \sqrt{n}[\widehat{C}_k^*(\widehat{\theta}_n^*) - C_k(\widehat{\theta}_n^*)]\}. \tag{A.26}$$

De plus, on peut vérifier que :

$$\sqrt{n}[\widehat{C}_k^* - C_k(\theta)] = A_1(\theta) + A_2(\theta) + A_3(\theta), \tag{A.27}$$

où $A_1(\theta)$, $A_2(\theta)$ et $A_3(\theta)$ comme suit :

$$A_1(\theta) = \frac{1}{\sqrt{n}}\sum_{t=k+1}^{n}[|\widehat{\epsilon}_t(\theta)| - |\epsilon(\theta)|][|\widehat{\epsilon}_{t-k}(\theta)| - |\epsilon_{t-k}(\theta)|].$$

$$A_2(\theta) = \frac{1}{\sqrt{n}}\sum_{t=k+1}^{n}[|\epsilon_t(\theta)| - \mu][|\widehat{\epsilon}_{t-k}(\theta)| - |\epsilon_{t-k}(\theta)|].$$

$$A_3(\theta) = \frac{1}{\sqrt{n}}\sum_{t=k+1}^{n}[|\widehat{\epsilon}_t(\theta)| - |\epsilon_t(\theta)|][|\epsilon_{t-k}(\theta)| - \mu)|].$$

Par l'extension de Taylor, Lemmes (A.1) et (A.2) et l'hypothèse 1, nous pouvons montrer que :

$$\sup_{\theta\in\Theta}|A_1(\theta)| \leq \frac{C}{\sqrt{n}}\sum_{t=k+1}^{n}\rho^t|y_t||y_{t-k}|\xi_{\rho,0}^2 = o_p(1).$$

De même, $\sup\limits_{\theta\in\Theta}|A_2(\theta)| = o_p(1)$ et $\sup\limits_{\theta\in\Theta}|A_1(\theta)\ |A_3(\theta)| = o_p(1)$.
Celles-ci avec (A.26) et (A.27) impliquent que $\sqrt{n}\widehat{C}_k^*(\widehat{\theta}_n) = \sqrt{n}C_k^*(\widehat{\theta}_n) + o_p^*(1)$.
De plus, on peut vérifier que :

$$\frac{C_k^*(\theta_0)}{\partial\theta} = \frac{C_k(\theta_0)}{\partial\theta} + J_{1n} + J_{2n}, \tag{A.28}$$

où J_{1n} et J_{1n} comme suit :

$$J_{1n} = -\frac{1}{2n}\sum_{t=k+1}^{n}(\omega_t - 1)\frac{|y_t|}{h_t^{3/2}}\frac{\partial h_t(\theta_0)}{\partial\theta}\left(\frac{|y_{t-k}|}{h_{t-k}^{1/2}} - \mu\right).$$

$$J_{2n} = -\frac{1}{2n}\sum_{t=k+1}^{n}(\omega_t - 1)\left(\frac{|y_t|}{h_t^{1/2}} - \mu\right)\frac{|y_{t-k}|}{h_{t-k}^{3/2}}\frac{\partial h_{t-k}(\theta)}{\partial\theta}.$$

Clairement, $\mathbb{E}^*(J_{1n}) = 0$ et $\mathbb{E}^*(J_{2n}) = 0$. Semblable à la preuve de (A.13), nous pouvons montrer que $\mathbb{E}(J_{1n}J_{1n}' = O(n^{-1})$ et $\mathbb{E}(J_{2n}J_{2n}') = O(n_{-1})$,
ce qui implique que $J_{1n} = o_p^*(1)$ et $J_{2n} = o_p^*(1)$.
Ceux-ci avec (A.28) impliquent que :

$$\partial C_k^*(\theta_0)/\partial\theta = \partial C_k(\theta_0)/\partial\theta + o_p^*(1).$$

Par le théorème ergodique, nous avons

$$\partial C_k(\theta_0)/\partial\theta = -H_k + o_p(1),$$

où $H_k = \mathbb{E}[0.5\mu h_t^{-1}\partial h_t(\theta_0)/\partial\theta(|\epsilon_{t-k}| - \mu)]$. Par conséquent, $\partial\widehat{C}_k^*(\theta_0)/(\partial\theta) = -H_k + o_p^*(1)$. Puis, par l'extension de Taylor, Lemme A.1 et le fait que $\sqrt{n}(\widehat{\theta}_n^* - \theta_0) = O_p(1)$, on peut montrer que :

$$\begin{aligned}\sqrt{n}\widehat{C}_k^*(\widehat{\theta}_n) &= \sqrt{n}\widehat{C}_k^*(\theta_0) + \frac{\partial C_k^*(\theta_0)}{\partial\theta}\sqrt{n}(\widehat{\theta}_n^* - \theta_0)o_p^*(1)\\ &= \sqrt{n}\widehat{C}_k^*(\theta_0) - H_k\sqrt{n}(\widehat{\theta}_n^* - \theta_0) + o_p^*(1).\end{aligned}$$

Ensuite, selon le théorème de Slutsky et (A.22), il s'ensuit que

$$\sqrt{n}\widehat{\rho}_k^* = \frac{\sqrt{n}}{\sigma^2}C_k^*(\theta_0) - \frac{H_k}{\sigma^2}\sqrt{n}(\widehat{\theta}_n^* - \theta_0) + o_p(1). \tag{A.29}$$

De plus, selon Li et Li (2005), nous avons

$$\sqrt{n}\widehat{\rho}_k = \frac{\sqrt{n}}{\sigma^2}C_k(\theta_0) - \frac{H_k}{\sigma^2}\sqrt{n}(\widehat{\theta}_n - \theta_0) + o_p(1).$$

Ceci, conjointement avec (A.29), implique que :

$$\sqrt{n}(\widehat{\rho}_k^* - \widehat{\rho}_k) = \frac{1}{\sigma^2}\frac{1}{\sqrt{n}}\sum_{t=k+1}^{n}(\omega_t - 1)(|\epsilon_t| - \mu)(|\epsilon_{t-k}| - \mu) - \frac{H_k}{\sigma^2}\sqrt{n}(\widehat{\theta}_n^* - \widehat{\theta}_n) + O_p^*(1).$$

Ensuite, par l'hypothèse 3 et (A.21), nous complétons la preuve en appliquant le lemme (4.6) de Praestgaard et Wellner (1993) et le dispositif Cramér-Wold.

3.3 Simulation

Cette section conduit trois expériences de simulation pour évaluer la performance d'un échantillon fini du Bootstrap définis dans les sections 1 et 2.
La première expérience consiste à évaluer la méthode du Bootstrap pour l'estimateur (EAM) de la section 1, et les données les processus de génération sont

$$ARCH(2) : y_t = \epsilon_t \sqrt{h_t}, \qquad h_t = 0.1 + 0.2 y_{t-1}^2 + 0.4 y_{t-2}^2 \tag{3.13}$$
$$GARCH(1,1) : y_t = \epsilon_t \sqrt{h_t}, \qquad h_t = 0.1 + 0.2 y_{t-1}^2 + 0.4 y_{t-1}^2, \tag{3.14}$$

où les $\{\epsilon_t\}$ sont $i.i.d.$ innovations de loi ou Student tel que $\mathbb{E}(\epsilon_t) = 0$ et la médiane de ϵ_t^2 est un.
Nous considérons deux tailles d'échantillons, n = 500 et 1000 , et il y a m = 500 répétitions pour chaque taille d'échantillon. La procédure d'amorçage de la section précédente est conduite avec cinq types de poids aléatoires $\{\omega_t\}$:

(i) $i.i.d.$ Poids exponentiels (1) (P_1) ;

(ii) $i.i.d.$ Poids de Rademacher (P_2) ;

(iii) $i.i.d.$ $U(0.5, 1.5)$ poids (P_3) ;

(iv)$Mult_n(n; (n^{-1}, ..., n^{-1})$ $\quad (P_4)$ poids ;

(v) les poids double Bootstrap (P_5).

En outre, nous utilisons B = 500 échantillons Bootstrapés pour calculer la matrice de covariance asymptotique de l'estimateur (EAM) $\widehat{\theta}_n$.
Notez que $\tau = 1$ pour les poids P_1,P_2 et P_4, $1/\sqrt{12}$ pour P_3et $\tau = \sqrt{2}$ pour P_5. Puisque nous pouvons toujours normaliser les poids $\{\omega_t\}$ de telle sorte que $\lim_{n\to\infty} \mathbb{V}(\omega_t) = 1$, la quantité de τ n'affectera pas le démarrage approximation.
Le tableau 1 présente les biais (Biais), l'écart type empirique (ETE) et l'écart type asymptotique (ETA). de $\widehat{\theta}_n = (\widehat{\theta}_1, ..., \widehat{\theta}_{p+q+1})'$ pour les modèles ARCH et GARCH.
Noter par $\widehat{\theta}_n^{(\ell)} = (\widehat{\theta}_1^{\ell}, ..., \widehat{\theta}_{p+q+1}^{\ell})'$ et $\mathbb{V}_{\mathbb{E}}^{*(\ell)}$ l'estimateur $\widehat{\theta}_n$ et de la matrice de covariance de $\{\mathbb{E}^{(1)}, ..., \mathbb{E}^{(B)}\}$ dans la procédure d'amorçage, respectivement, pour la ℓ réplication.
Les valeurs ETE et ETA de θ_j sont alors définies comme suit :

$$ETE_j \;=\; \sqrt{\frac{1}{1-m}\sum_{\ell=1}^{m}\left[(\widehat{\theta}_j^{\ell}) - (\bar{\theta}_j)\right]^2},$$

$$ETA_j = \frac{1}{m}\sum_{\ell=1}^{m}\frac{(V_E^{*\ell})_{jj}}{n},$$

où $\bar{\theta}_j = m_{-1}\sum_{\ell=1}^{m}\widehat{\theta}_j^{\ell}$, et $(M)_{jj}$ désignent la jième entrée en diagonale de la matrice M. Du tableau 3.1, nous avons les résultats suivants pour les deux modèles :

(1) Lorsque la taille de l'échantillon augmente, les biais, les ETE et les ETA diminuent, et les ETE et les ETA se rapprochent les uns des autres ;

(2) Les ETA calculés par cinq ensembles de pondérations aléatoires sont assez similaires lorsque la taille de l'échantillon est aussi petite que n = 500 et ils deviennent assez proches les uns des autres lorsque n augmente jusqu'à 1000.

La deuxième expérience examine les performances de l'approximation Bootstrap pour les autocorrélations des résidus $\widehat{\rho}_k$ dans la section 2.

Les données et autres paramètres sont conservés comme dans la première expérience. Le tableau 3.2 résume les biais, les ETE et les ETA de l'autocorrélation résiduelle absolue $\widehat{\rho}_k$ aux retards 2, 4 et 6 pour les deux ARCH et les modèles GARCH, où les ETE et les ETA sont calculés comme pour l'estimateur (EAM).

Les conclusions sont similaires comme dans la première expérience. Au fur et à mesure que la taille de l'échantillon augmente, les biais, les ETE et les ETA deviennent plus petits et le processus d'amorçage l'approximation va mieux. De plus, différentes pondérations aléatoires conduisent à des résultats similaires pour la matrice de covariance asymptotique de k.

La troisième expérience consiste à évaluer la procédure d'amorçage pour le test de Portemanteau. Les données sont générées par un modèle ARCH

$$y_t = \epsilon_t\sqrt{h_t}, \qquad h_t = 0.2 + 0.1y_{t-1}^2 + 0.1y_{t-2}^2 + dy_{t-3}^2. \tag{3.15}$$

ou un modèle GARCH

$$y_t = \epsilon_t\sqrt{h_t}, \qquad h_t = 0.2 + 0.1y_{t-1}^2 + dy_{t-2}^2 + 0.1h_{t-1} \tag{3.16}$$

où le départ d = 0, 0.1 ou 0.3, alors que tous les autres paramètres sont conservés des deux expériences précédentes. L'estimateur (EAM) est réalisée en supposant un modèle ARCH(2) pour les données générées par (3.15) et un modèle GARCH(1,1). modèle pour les données de (3.16). Par conséquent, d = 0 correspond à la taille du test de Portemanteau, et $d \neq 0$ correspond à la puissance du test.

Sur la base de la procédure d'amorçage décrite à la section 2, le rejet empirique les taux

		Loi normal							Loi de student						
n		Bias	ETE	ETA_1	ETA_2	ETA_3	ETA_4	ETA_5	Bias	ETE	ETA_1	ETA_2	ETA_3	ETA_4	ETA_5
							ARCH(2)								
500	α_0	0.08	0.36	0.36	0.36	0.37	0.36	0.36	0.48	2.79	1.33	1.81	0.90	1.07	1.27
	α_1	0.04	0.60	0.58	0.58	0.60	0.57	0.57	−0.02	0.49	0.51	0.51	0.53	0.53	0.53
	α_2	−0.17	0.76	0.86	0.86	0.91	0.85	0.84	−0.10	0.78	0.80	0.80	0.83	0.80	0.79
1000	α_0	0.05	0.25	0.25	0.25	0.27	0.25	0.25	0.10	0.41	0.45	0.43	0.42	0.44	0.46
	α_1	0.00	0.41	0.41	0.41	0.44	0.41	0.41	0.03	0.37	0.37	0.37	0.38	0.36	0.37
	α_1	−0.03	0.60	0.60	0.60	0.64	0.61	0.61	−0.09	0.57	0.55	0.55	0.56	0.55	0.55
							GARCH(1.1)								
500	α_0	0.07	0.47	0.48	0.48	0.56	0.46	0.43	0.15	0.40	0.46	0.46	0.47	0.46	0.46
	α_1	0.01	0.65	0.64	0.64	0.71	0.64	0.63	0.13	0.59	0.57	0.57	0.60	0.58	0.58
	α_2	−0.09	1.50	1.42	1.42	0.71	1.41	1.29	−0.32	0.88	0.91	0.91	0.98	0.90	0.86
1000	α_0	0.04	0.32	0.33	0.33	0.35	0.33	0.32	0.09	0.29	0.30	0.30	0.32	0.30	0.30
	α_1	−0.01	0.46	0.45	0.45	0.48	0.45	0.45	0.07	0.42	0.39	0.39	0.41	0.39	0.39
	α_1	−0.10	1.03	1.05	1.06	1.15	1.03	1.00	−0.22	0.68	0.65	0.65	0.69	0.63	0.62

TAB. 3.1 – Biais (×10), ETE (×10) et ETA (×10) de l'estimateur (EAM) pour les innovations de loi normale ou de Student, où ETA_i correspond aux poids aléatoires p_i pour i = 1, 2, 3, 4 et 5, et 0.1 et 2 (ou 1) représentent les valeurs correspondantes

		Loi normal							Loi de Student						
n	k	Bias	ETE	ETA_1	ETA_2	ETA_3	ETA_4	ETA_5	Bias	ETE	ETA_1	ETA_2	ETA_3	ETA_4	ETA_5
							ARCH(2)								
500	2	0.92	4.41	4.62	4.64	5.18	4.67	4.55	0.36	4.39	4.12	4.16	4.29	4.11	4.06
	4	0.10	4.18	4.45	4.50	4.61	4.51	4.45	−0.31	4.28	3.94	3.98	4.02	4.14	4.10
	6	0.14	4.49	4.42	4.47	4.52	4.49	4.42	−0.32	4.37	4.11	4.17	4.19	4.02	4.00
1000	2	0.23	3.00	3.25	3.24	3.55	3.28	3.24	0.14	3.05	2.88	2.88	2.95	2.88	2.86
	4	−0.10	3.14	3.17	3.18	3.26	3.16	3.14	−0.21	3.30	2.93	2.94	2.96	2.93	2.91
	6	0.05	3.20	3.15	3.17	3.19	3.16	3.14	−0.10	2.85	2.87	2.89	2.90	2.88	2.88
							ARCH(1.1)								
500	2	0.71	4.18	4.60	4.64	5.15	4.61	4.49	0.38	4.33	4.16	4.21	4.29	4.17	4.14
	4	0.33	4.58	4.50	4.52	4.89	4.52	4.41	0.11	4.23	4.08	4.12	4.18	4.19	4.14
	6	−0.29	4.64	4.42	4.49	4.71	4.48	4.38	0.35	4.60	4.11	4.15	4.21	4.11	4.07
1000	2	0.56	3.04	3.26	3.26	3.46	3.27	3.22	−0.07	3.14	2.97	2.99	3.03	2.93	2.91
	4	0.04	3.35	3.21	3.23	3.41	3.20	3.14	0.48	3.34	3.06	3.09	3.11	2.96	2.95
	6	0.11	3.11	3.17	3.19	3.29	3.18	3.13	0.32	3.02	2.96	2.98	3.00	3.03	3.01

TAB. 3.2 – Biais (×100), ETE (×100) et ETA (×100) d'autocorrélations des résidus $\widehat{p}_k$ de k=2,4 et 6 , pour les innovations de loi normale ou Loi de Student, où ETA_i correspond aux poids aléatoires p_i pour i = 1, 2, 3, 4 et 5

		Loi normal					Loi de Student				
n	d	P_1	P_2	P_3	P_4	P_5	P_1	P_2	P_3	P_4	P_5
		ARCH(2)									
500	0.0	4.80	4.00	3.20	3.80	4.40	6.00	4.60	4.80	3.80	4.80
	0.1	56.2	54.00	51.8	56.0	60.6	54.8	51.2	50.0	53.4	55.0
	0.3	99.2	99.4	98.4	99.2	99.2	94.8	94.6	94.0	94.2	95.4
1000	0.0	5.2	5.0	4.4	4.6	4.8	5.2	4.4	3.4	5.00	5.20
	0.1	93.2	92.8	94.0	94.6	95	90.4	88.8	88.8	89.8	89.2
	0.3	100	100	100	100	100	98	97.8	98.2	98.4	97.8
		ARCH(1.1)									
500	0.0	6.4	5.2	4.0	5.8	8.0	6.2	5.2	3.4	5.0	5.2
	0.1	22.4	22.0	17.8	22.8	24.0	23.0	20.4	21.6	23.8	25.2
	0.3	95.4	95.6	95.0	96.0	96.6	87.8	88.0	87.0	89.8	89.2
1000	0.0	4.6	5.4	4.2	5.6	5.2	5.6	6.0	5.6	5.8	5.4
	0.1	43.6	42.6	40.0	41.8	42.4	41.4	40.6	41.0	2.8	43.0
	0.3	100	100	100	100	100	99.2	98.8	98.6	98.6	98.6

TAB. 3.3 – Taux de rejet empiriques (%) de la statistique de test Q (M) pour M = 6 au seuil de signification de 5% , pour la loi normale ou Student et d = 0, 0.1 et 0.3, où p_i représente une pondération aléatoire pour i = 1, 2, 3, 4 et 5

sont calculés comme le rapport des rejets sur 500 répétitions.
Le tableau 3.3 donne les taux de rejet empiriques de la statistique de test Q (M) avec M = 6 pour les tests ARCH et GARCH modèles, où Q (M) est calculé par l'approche Bootstrap de la section 2. On voit que le Bootstrap test de Portemanteau donne de bons résultats en termes de taille et de puissance :
La taille se rapproche du niveau nominal 5% lorsque la taille de l'échantillon n augmente à 1 000 et la puissance augmente en fonction de la taille de l'échantillon n ou du départ d augmente.
De plus, les procédures de Bootstrap avec différents poids aléatoires ont des performances similaires, en particulier quand n ou d est grand. En résumé, la méthode de Bootstrap généralisée fonctionne bien dans l'estimateur (EAM) du Bootstrap et les statistiques de test pour des échantillons finis.
En particulier, le choix des poids aléatoires a peu d'influence sur les performances du Bootstrap procédure, et nous pouvons choisir une distribution simple et pratique pour générer des poids aléatoires dans la pratique.

Conclusion

Dans ce travail, nous nous sommes intéressés aux modèles ARCH introduits initialment par Engle en 1982 et généralisé par Bollerslev en 1986. Nous nous sommes concentré sur l'estimation des paramètres de cette classe de modèles. Plus précisément, nous avons traité les questions liées à la consistence et la normalité asymptotique des estimateurs des moindres carré et du maximum de vraissmeblance. Aussi, nous avons introduit une nouvelle classe d'estimateur, si celle qui minimise l'écart absolu moyen.

Aussi, à base des techniques de Bootstrap, nous avons étudier les propriétés de cet estimateur. Une application au test de Portemanteau a été aussi développé.

Comme perspectives futures, il serait intéressant d'élargir notre étude en appliquant le Bootstrap aux autres techniques d'estimation des modèles GARCH, IGARCH, ...

Bibliographie

L'article "Bootsrap inference for GARCH models by the least absolute deviation estimation" de **Qianqian Zhu, Ruochen Zeng and Guodong Li** publié en 2019 dans la revue "Journal of Time series analysis".

Engle, R.F. (1982), AutoRegressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, Econometrica,50, 987-1008.

Bollerslev T.1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31 : 307–327.

Bose A, Chatterjee S. 2002. Comparison of bootstrap and jackknife variance estimators in linear regression : second order results.

Berkes I, Horváth L, Kokoszka P. 2003. GARCH processes : structure and estimation. Bernoulli 9 : 201–227

Berkes I, Horváth L. 2004. The efficiency of the estimators of the parameters in GARCH processes. The Annals of Statistics 32 :633–655.

Francq C, Zakoian J-M. 2010. GARCH Models : Structure, Statistical Inference and Financial Applications. Chichester, UK : John Wileyet Sons. Francq C, Zakoian J-M. 2010.

Zhu K, Ling S. 2011. Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA-GARCH/IGARCH models.

The Annals of Statistics 39 : 2131–2163.

Efron B. 1979. Bootstrap methods : another look at the jackknife. The Annals of Statistics 7 : 1–26.

Efron B, Tibshirani RJ. 1993. An Introduction to the Bootstrap. New York : Chapman and Hall/CRC.

Li G, Li Y, Tsai C-L. 2015. Quantile correlations and quantile autoregressive modeling. Journal of the American StatisticalAssociation 110 : 246–261

Peng L, Yao Q. 2003. Least absolute deviations estimation for ARCH and GARCH models. Biometrika 90 : 967–975.

Praestgaard J, Wellner J. 1993. Exchangeably weighted bootstraps of the general empirical process. Annals of Probability 21 : 2053–2086.

Zhu K. 2016. Bootstrapping the portmanteau tests in weak autoregressive moving average models. Journal of the Royal Statistical Society, Series B 78 : 463–485.

Zheng Y, Zhu Q, Li G, Xiao Z. 2018. Hybrid quantile regression estimation for time series models with conditional heteroscedasticity. Journal of the Royal Statistical Society, Series B 80 : 975–993.