

N° D'ORDRE : 19/2012/M/MT

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie Houari Boumediene
Faculté des Mathématiques



MÉMOIRE
PRÉSENTÉ POUR L'OBTENTION DU DIPLOME DE MAGISTER

EN RECHERCHE OPÉRATIONNELLE
OPTION : MÉTHODES STOCHASTIQUES

Par : MESSAHLI Leila

**Estimation d'un modèle GARCH en présence de
données de haute fréquence**

Soutenu publiquement, le 11/01/2012, devant le jury composé de :

M ^r A. BERRACHEDI	Professeur,	à l'USTHB	Président.
M ^{me} H. GUERBYENNE	Maître de Conférences/ A,	à l'USTHB	Directrice de mémoire.
M ^r H. BELBACHIR	Maître de Conférences/ A,	à l'USTHB	Examinateur.
M ^{me} N. EL SAADI	Maître de Conférences/ A,	à l'ENSSEA	Examinatrice.
M ^r F. HAMDI	Maître de Conférences/ A,	à l'USTHB	Invité.



Louange À Dieu, le miséricordieux, sans Lui rien de tout cela n'aurait pu être.

Je tiens en premier lieu à remercier M^{me} H. GUERBYENNE pour l'honneur qu'elle ma fait en acceptant de m'encadrer. Son exigence et son ouverture scientifique, sa rigueur, son enthousiasme très communicatif, ses critiques et ses encouragements m'ont beaucoup aidés. J'apprécie sa vaste connaissance qui m'a été très utile et je lui suis reconnaissante de m'avoir donnée l'opportunité de mener à bien cette recherche.

Je remercie vivement M^r A. BERRACHEDI pour l'honneur qu'il m'a fait en acceptant de présider le jury de ce mémoire.

Je remercie M^r M. BELBACHIR, M^{me} N. ELSAADI et M^r F. HAMDI pour l'honneur qu'ils m'ont fait en acceptant d'examiner ce travail et pour le temps consacré à la lecture et à la correction de ce mémoire.

Je voudrais aussi remercier ma famille pour le soutien qu'elle m'a procuré et particulièrement pour leurs encouragements, car sans eux, je n'aurais pas fini ce mémoire.

Enfin, merci à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.



Je dédie ce modeste travail :

À la mémoire de mes deux très chers grands pères .

À ma mère et mon père.

À mes frères et soeurs.

À ma promotion Magister MS.

À tous mes amis

L. MESSAHLI

Estimation d'un modèle GARCH en présence de données de haute fréquence

Leila MESSAHLI

Département de Recherche Opérationnelle

Faculté de Mathématiques,

Université des Sciences et de la Technologie Houari Boumediene,

U. S. T. H. B.

Table des matières

Table des matières	i
<i>Introduction générale</i>	<i>1</i>
1 Quelques outils mathématiques	8
1.1 Introduction	8
1.2 Processus stochastiques	8
1.2.1 Introduction	8
1.2.2 Notions élémentaires sur les processus aléatoires	9
1.2.3 Stationnarité et ergodicité	10
1.2.4 Équations aux récurrences stochastiques	14
1.3 Processus d'Itô et Quelques processus stochastiques en temps continu	18
1.3.1 Le processus d'Itô unidimensionnel	22
1.3.2 Formule d'Itô	22
1.4 Martingales et Semimartingales	26
1.4.1 Introduction	26
1.4.2 Martingales	27
1.4.3 Semimartingales et martingales locales	28
2 Structure probabiliste et estimation du modèle GARCH (p, q)	33
2.1 Introduction	33

2.2	Faits stylisés des séries financières	34
2.2.1	Non stationnarité de processus des prix	35
2.2.2	Regroupement des volatilités (volatility clustering)	38
2.2.3	Absence d'autocorrélation des rendements logarithmiques	39
2.2.4	Distribution des rendements	40
2.2.5	Présence de mémoire longue	44
2.2.6	Effet de levier	45
2.2.7	Saisonnalité	45
2.3	Structure probabiliste du modèle GARCH	46
2.3.1	Présentation et définition du modèle GARCH	46
2.3.2	Étude de la stationnarité des processus GARCH	47
2.3.3	Existence des moments d'ordres supérieurs et structure d'autocovariance du modèle GARCH	60
2.4	Estimation d'un modèle GARCH	63
2.4.1	Estimation du modèle par la méthode du quasi-maximum de vraisemblance	64
2.4.2	Convergence forte de l'estimateur du QMV	67
2.4.3	Normalité asymptotique des estimateurs du QMV	68
2.4.4	Simulation	69
2.5	Conclusion	72
3	<i>Proxies de volatilité et volatilité réalisée</i>	74
3.1	Introduction	74
3.2	Modèle d'échelle	75
3.3	Proxies de volatilité	78
3.3.1	Définitions et notations	78
3.3.2	Ordre sur les proxies	80

3.4	Variance réalisée	86
3.4.1	Modèle en temps discret	88
3.4.2	Propriétés statistiques de la variance réalisée en absence des effets de microstructure	89
3.4.3	Propriétés statistiques de la variance réalisée en présence des effets de microstructure	93
3.4.4	Quelques propriétés des volatilités réalisées et de leur logarithme de l'indice S&P 500 <i>Futures</i>	96
3.5	Test de l'hypothèse de semimartingale pour l'indice S&P 500 <i>Futures</i>	101
3.5.1	Cadre théorique de l'approche utilisée	102
3.5.2	Mise en oeuvre du test	104
3.6	Conclusion	122
4	<i>Estimation de modèles GARCH en présence de données de haute fréquence.</i>	123
4.1	Présentation du modèle	123
4.2	Propriétés asymptotiques des estimateurs du quasi maximum de vraisemblance basés sur des proxies de volatilité	126
4.2.1	Consistance et normalité asymptotique des estimateurs du QMV gaussien	126
4.3	Estimation du modèle log-GARCH	129
4.3.1	Consistance et normalité asymptotique des estimateurs du QMV log-gaussien	129
4.3.2	Gain en efficacité du QMV log-gaussien relativement au QMV gaussien	131
4.4	Simulation et application	131
4.4.1	Simulation	131
4.4.2	Description des données	134
4.4.3	Mise en oeuvre du modèle	135

4.5	Semimartingales, variation quadratique et volatilité réalisée	137
4.6	Conclusion	142
	Conclusion générale	143
	Annexe	145
	Bibliographie	154

Alà barakati Allah

Introduction générale

"L'économie est une science consistant à penser en termes de modèles."

Keynes

Volatilité est sans doute l'un des concepts les plus intrigants et mystérieux en finance. Intrigant dans le sens qu'il est le paramètre crucial dans la plupart des processus financiers et mystérieux vue l'ambiguïté de sa définition. En finance et en économétrie financière, l'analyse des variations du prix d'un actif et de son instabilité est fondamentale. Il est bien connu que les rendements journaliers des actifs financiers sont difficiles, voire impossible à prédire, bien que leur volatilité semble, relativement facile à prédire. De ce fait, la volatilité joue un rôle clé dans l'évaluation des produits dérivés, dans l'allocation des actifs ou dans la gestion du risque. La volatilité mesure le degré de fluctuation d'un actif financier, d'un produit dérivé, d'un marché,....etc. Elle est associée à l'imprévu, l'incertain et le risque.

Le concept de volatilité est probablement l'un des sujets qui suscitent le plus de recherches dans le domaine de la finance mathématique. Louis Bachelier (1900) a été le premier à quantifier le concept du mouvement brownien, introduit par le botaniste écossais Robert Brown (1827). Il a montré que c'est le processus stochastique adéquat pour modéliser l'évolution des prix des actions. Pour cet auteur le degré des fluctuations des prix dépend d'un paramètre inconnu σ , dit, coefficient d'instabilité, appelé aujourd'hui *volatilité*. Osborne (1959) et le lauréat du prix Nobel d'économie en 1970, Paul Samuelson (1965), ont redécouvert la thèse de Bachelier et étendu le modèle à un mouvement brownien géométrique, excluant ainsi les prix négatifs. L'étude la plus reconnue dans le monde de la finance mathématique est certainement le modèle d'évaluation des options de Black-Scholes et Merton (1973). Ce modèle permet d'évaluer le prix d'une option dans le cas de *non arbitrage*, en se basant sur l'hypothèse que l'évolution d'un titre suit un mouvement brownien géométrique et que la

volatilité est constante. Cependant, l'hypothèse de la constance de volatilité dans la formule de Black-Scholes était rejetée depuis le début comme cela est démontré dans les travaux de Black (1976), Schmalensee et Trippi (1978) et Christie (1982), et surtout après le crash (1987), qui a mis en évidence l'effet *smile* qui pourtant ne devrait pas apparaître sous les hypothèses de Black-Scholes.

Tout commence en 1963 et 1965, lorsque les deux auteurs Mandelbrot et Fama ont constaté deux faits sur les données financières. Premièrement, elles ont des queues plus épaisses que celles de la loi normale, *i.e.*, les valeurs loin de la moyenne apparaissent avec des probabilités plus élevées que pour une loi normale ; le coefficient d'aplatissement étant élevé. Deuxièmement des regroupements de volatilité «volatility clustering», sont remarqués dans les faits empiriques, tout en précisant "*Large changes tend to be followed by large changes of either sign and small changes tend to be followed by small changes*". Par suite, d'autres propriétés sont signalées. Il est connu de nos jours, que les séries financières des prix exhibent certaines régularités statistiques dites faits stylisés, à savoir : regroupement de volatilité (volatility clustering), excès de kurtosis, asymétrie, effet de levier, périodicité dans la structure des covariances, persistance dans la structure dynamique, et bien d'autres. Ces propriétés illustrent la difficulté de la modélisation des séries financières. Les modèles linéaires classiques de type *ARMA* se révèlent incompatibles pour capturer ces faits stylisés car ils sont centrés sur la fonction d'autocovariance.

Les modèles économétriques introduits dans la littérature, afin de prendre en compte les propriétés particulières des séries financières se présentent généralement sous la forme multiplicative $\varepsilon_n = \sigma_n Z_n$, où la variable aléatoire σ_n est appelée *volatilité*. Le fait que les grandes valeurs des carrés des rendements soient généralement précédées par de grandes valeurs est difficilement compatible avec une variance conditionnelle constante. Ce phénomène est connu sous le nom d'hétéroscédasticité conditionnelle. Les processus *ARCH* introduits par Engle (1982) forment une classe de processus stochastiques largement utilisée dans la finance et l'économie pour modéliser l'hétéroscédasticité conditionnelle. Les modèles *GARCH*, ou modèles autorégressifs conditionnellement hétéroscédastiques généralisés (généralisés par Bollerslev en 1986), viennent apporter une réponse aux faits empiriques répertoriés ci-dessus, en autorisant une dépendance linéaire de la volatilité avec ses valeurs historiques, et en ajustant cette volatilité avec le carré des rendements observés.

Une gamme d'articles est consacrée à l'étude de la structure probabiliste des modèles *GARCH*. La stationnarité stricte du modèle *GARCH* (1, 1) a été étudiée par Nelson (1990a). Pour les modèles *GARCH* (p, q), les conditions de stationnarité stricte sont établies par Bougerol et Picard (1992b). La condition de stationnarité au second ordre du modèle *GARCH* (p, q) est établie dans Bollerslev (1986), de même que l'existence d'une représentation *ARMA* pour le carré d'un *GARCH* (Bollerslev, 1988). La condition nécessaire et suffisante d'existence des moments d'ordre pair a été établie par Ling et McAleer (2002).

L'utilisation de la méthode du quasi maximum de vraisemblance (QMV) dans l'estimation des paramètres du modèle *GARCH* est particulièrement intéressante (sous certaines conditions de régularités), contrairement aux méthodes des moindres carrés ordinaires et moindres carrés quasi-généralisées. Les propriétés asymptotiques des estimateurs du QMV pour les modèles *ARCH* ont été établies par Weiss (1986) sous des conditions d'existence des moments d'ordre 4. Le premier article donnant une preuve rigoureuse, sous des hypothèses très faibles, des propriétés asymptotiques de l'estimateur du QMV dans le modèle *GARCH* (p, q) est celui de Berkes, Horváth et Kokoszka (2003a, b). Ces hypothèses ont été légèrement affaiblies par Francq et Zakoïan (2004). De son côté Straumann et Mikosch (2006), ont étudié le modèle conditionnellement hétéroscédastique généralisant le modèle *GARCH* (1, 1), par l'approche d'une représentation sous forme d'une équation aux récurrences stochastiques.

La mise en place de systèmes électroniques de cotation, d'information et de gestion de marchés conduit maintenant à des bases de données incluant des informations nombreuses et variées. Ces données de cotations (tick by tick data) correspondent à des durées très courtes entre enregistrements successifs, de l'ordre 3 à 10 secondes pour les marchés très liquides (données haute-fréquence). Des bases de données intra-journalières ont commencé à apparaître.

La plupart des séries financières, sont analysées à différentes fréquences (mois, semaine, jour, minute, seconde,...etc.). Le choix de la fréquence observée a souvent une importance cruciale quant aux propriétés de la série observée. Les modèles en temps continu semblent un bon cadre pour tenir en compte les données de haute fréquence.

Par essence, la volatilité est un phénomène non observable ; elle est inhérente à la plupart des modèles financiers. Comme elle ne s'observe pas naturellement, alors, il est nécessaire de

construire un certain nombre de statistiques H_n dites *proxies* et l'on se base sur ces dernières lors de spécification, d'estimation et d'évaluation des modèles de volatilité.

En statistique le terme proxy est utilisé pour une variable qui n'est pas de premier intérêt, en elle-même, mais qui est étroitement liée à un objet d'intérêt. Elle est reliée au concept d'estimateur, mais les proxies tendent à remplacer les variables d'intérêt non observables, tandis que un estimateur est une statistique désignée pour l'estimation des paramètres (inconnus). Un bon estimateur des paramètres est celui qui possède une petite racine carrée de l'erreur moyenne quadratique, en parallèle, une bonne proxy d'une variable est celle qui possède une forte corrélation avec cette variable.

Mesurer la volatilité à une fréquence très élevée, de façon presque continue en utilisant des données intra-journalières est possible. L'idée de la volatilité réalisée est initialement proposée par French, Schwert et Stambaugh (1987). Taylor et Xu (1997) ainsi qu'Andersen et Bollerslev (1997) montrent que la volatilité journalière réalisée peut être construite en additionnant les rendements intrajournaliers au carré, en prenant la racine carrée de cette dernière. Une motivation théorique de la construction de la variance réalisée réside dans le fait que la variation quadratique d'une grande classe de semimartingales est un estimateur sans biais de la variance conditionnelle des rendements (voir, par exemple, Barndorff-Nielsen et Shephard (2002), et Andersen, Bollerslev, Diebold, et Labys (2003), McAleer et Medeiros (2008), Andersen, Bollerslev, et Diebold (2009)). En effet, la disponibilité des données de haute fréquence crée un intérêt pour cette approche. Une volatilité observable donne lieu à de nouvelles opportunités : on peut l'analyser, l'optimiser et la prévoir en utilisant des techniques plus simples que les modèles économétriques complexes utilisés lorsque la volatilité est latente. Effectivement, les proxies de volatilité construites sur la base des données intra-journalières ont un pouvoir de prévision supérieur à celles des proxies construites à partir des rendements journaliers (voir, par exemple, Ghysels, Santa-Clara, et Valkanov (2006), Engle et Gallo (2006)). Généralement, la recherche d'une bonne proxy souffre de l'absence d'une méthode de comparaison de ces proxies.

Depuis le début de ce siècle, l'attention des chercheurs est consacré vers l'application des modèles de semimartingales pour les données de haute fréquence. Deux résultats sont à signaler

- (1) les données de haute fréquence (financières) sont utilisées pour estimer la volatilité, en utilisant la volatilité réalisée,
- (2) les modèles de type *ARFIMA* pour la volatilité réalisée offre un cadre impressionnant de prédiction de volatilité.

Intuitivement, une semimartingale est un processus qui se compose du signal entaché d'un bruit. Le processus des prix logarithmiques (ou des rendements) forme une semimartingale s'il est la somme de deux processus, le processus A_t , le signal et M_t , le bruit, où t est un nombre réel positif. Le processus A_t est caractérisé par des trajectoires relativement régulières, il est supposé être à variations bornées, c'est la partie modélisable. Le processus M_t est la partie non prédictible et erratique, c'est la martingale locale.

La classe des semimartingales est une grande classe, en particulier la plupart des modèles financiers en temps continu se trouve dans cette classe (mouvement brownien, processus d'Itô, et les processus Lévy, par exemple). L'application sur la série de données réelles de l'indice S&P 500 *futures* montrent que le processus qui lui est associé forme une semimartingale (Peters et de Vilder (2006)).

Datant des années quatre-vingt dix, des recherches sur les modèles *GARCH* et les données de haute fréquence existent, mais ne conservent pas la spécification du modèle *GARCH*. Les résultats concernant l'aggrégation temporelle des modèles *GARCH* établis par Drost et Nijman (1993), montrent que le modèle *GARCH* n'est pas atteint sous cette aggrégation. Faisant tendre le pas du temps vers zéro, les limites de processus de type *GARCH* sont des processus de diffusion (Klüppelberg, Lindner, et Maller, 2004, Kallsen et Vesenmayer, 2009, Nelson, 1990b).

de Vilder et Visser (2008) ont introduit un modèle pour tenir compte du mouvement intra-journalier des rendements des actifs financiers, dit modèle d'échelle. Visser (2008) a incorporé les données de haute fréquence dans le modèle (journalier) en temps discret, *GARCH*, et par suite il a développé une théorie sur l'estimation des paramètres du modèle *GARCH*, fondée sur la méthode du QMV. L'idée est de construire la fonction de vraisemblance en se basant sur des proxies H_n , qui exploitent l'information contenue dans le processus des rendements intra-journaliers (le modèle d'échelle). Néanmoins, d'autres auteurs comme Lildholdt (2002),

Brandt et Jones (2006) ont utilisé la proxy basée sur les valeurs extrêmes max-min, dans l'estimation et prédiction du modèle *GARCH*, en supposant un mouvement brownien pour les rendements intra-journaliers, qui est un cas particulier du modèle d'échelle.

L'objectif principal de ce mémoire est l'estimation des paramètres d'un modèle *GARCH* en présence de données de haute fréquence. Après avoir passé en revue les propriétés statistiques les plus saillantes caractérisant les séries financières, nous étudions la structure probabiliste (stationnarité, ergodicité, conditions d'existence des moments d'ordres supérieurs, etc.) ainsi que les propriétés asymptotiques de l'estimateur du QMV, à savoir la consistance et la normalité asymptotique du modèle *GARCH*.

Étant donné les propriétés asymptotiques intéressantes de la variance réalisée sous l'hypothèse de semimartingalité du processus sous-jacent, nous testons cette hypothèse sur les données de l'indice S&P 500 *futures*, où nous acceptons l'hypothèse que le processus associé à l'indice S&P 500 *futures* est une semimartingale. Par la suite, nous étudions les propriétés statistiques de l'estimateur du quasi maximum de vraisemblance (QMV) selon la théorie développée par Visser. Nous terminons ce mémoire par une étude de simulation et une application sur la série de l'indice S&P 500 *futures*, où nos résultats montrent qu'on a un gain en efficacité dans l'estimation des paramètres du modèle *GARCH*, en utilisant la volatilité réalisée comme proxy.

Ce travail intitulé "*estimation d'un modèle GARCH en présence de données de haute fréquence*", comporte quatre chapitres.

Dans le premier chapitre intitulé "**Quelques outils mathématiques**", nous introduisons les notions et les concepts de base, ainsi que les outils principaux utilisés dans les chapitres suivants, qui ont permis par suite de faire ce travail.

Le chapitre 2 intitulé "**Structure probabiliste et estimation du modèle *GARCH* (p, q)**", aborde l'aspect probabiliste et estimation du modèle *GARCH* classique, après avoir rappelé les faits stylisés, généralement observés dans les marchés financiers (nous les illustrons sur les données journalières de l'indice S&P 500 *futures*). Dans un premier temps, nous spécifions une écriture autorégressive généralisée du modèle qui nous permettra d'établir, à partir de l'exposant de Lyapounov, les propriétés probabilistes du modèle *GARCH*, à savoir, la stationnarité stricte et au second ordre, l'ergodicité, l'existence des moments d'ordres supé-

rieurs et la structure d'autocovariance. Ensuite, nous passons à l'estimation des paramètres de ce modèle en utilisant la méthode du quasi maximum de vraisemblance (QMV). Les propriétés asymptotiques de ces estimateurs (la consistance et la normalité asymptotique) sont étudiées. Enfin, une étude de simulation intensive a été faite. Elle nous a permis de confirmer les résultats théoriques obtenus.

Le chapitre 3 intitulé "**Proxies de volatilité et volatilité réalisée**", est essentiellement consacré à l'étude des proxies de volatilité en présence des données intra-journalières de haute fréquence dans les modèles de volatilité, en temps discret (la volatilité est un facteur d'échelle). Après avoir introduit le processus des rendements intra-journalier $R_n(\cdot)$ dit modèle d'échelle (*scaling*), nous abordons le problème d'optimisation de ces proxies (le critère à utiliser, l'existence de proxies optimales, ...) d'une part. D'autre part, nous examinons les propriétés statistiques (convergence et distribution asymptotique) de l'une des proxies de volatilité les plus utilisées en économétrie financière, à savoir la volatilité journalière réalisée (elle se construit en additionnant les carrés des rendements intra-journaliers et en prenant la racine carrée), cela sera fait en absence / présence des effets de microstructure. Nous terminons ce chapitre par le test de semimartingale sur le processus des prix associé à l'indice S&P 500 *futures*.

Dans le dernier chapitre dont l'intitulé est "**Estimation de modèles GARCH en présence de données de haute fréquence**", nous montrons à quel point l'estimation des paramètres du modèle GARCH par la méthode du QMV sera améliorée, si une proxy adéquate de volatilité est utilisée en présence de données de haute fréquence. Les propriétés asymptotiques des estimateurs, à savoir la consistance et la normalité asymptotique sont conservées *i.e.*, comme dans le cas classique pour les rendements *close-to-close*. Un gain en efficacité des estimateurs est obtenu si l'on travaille sur le logarithme des proxies H_n au lieu des proxies H_n . Lors de l'étude des propriétés statistiques, nous utilisons la théorie développée par Visser (2008) fondée sur la méthode du QMV. L'intuition sous-jacente est de construire la fonction de vraisemblance en se basant sur des proxies H_n , qui exploitent l'information contenue dans le processus des rendements intra-journaliers $R_n(\cdot)$. Une étude de simulation et une application sur la série de l'indice S&P 500 *futures* ont été faites. Elles confortent les résultats théoriques.

1

Quelques outils mathématiques

1.1 Introduction

Les processus stochastiques offrent un cadre mathématique à de nombreux phénomènes évolutifs dans le temps. Nous consacrons le premier chapitre à la présentation des principaux outils nécessaires à l'étude des modèles de séries chronologiques financières. En effet, les propriétés probabilistes, à savoir : la stationnarité (stricte et au second ordre), l'ergodicité (forte et faible) caractérisant la structure de dépendance des processus stochastiques solutions d'équations aux différences stochastiques sont étudiées.

Du fait que notre intérêt est centré autour de la modélisation de processus en présence de données de haute fréquence, les concepts de martingale en temps continu, de martingale locale, de semimartingale, de variation quadratique, sont introduits.

1.2 Processus stochastiques

1.2.1 Introduction

Il est bien connu, que la modélisation d'une série chronologique repose sur la construction d'un modèle pour le processus stochastique sous-jacent, *i.e.*, le processus générateur

de cette série temporelle. La plupart de ces modèles admettent une représentation markovienne linéaire de la forme $Y_t = A_t Y_{t-1} + B_t$, $t \in \mathbb{Z}$, où la suite de matrices $\{(A_t, B_t), t \in \mathbb{Z}\}$ peut être indépendante et identiquement distribuée (*i.i.d.*). Cette dernière représentation est d'une importance majeure lors de l'étude de l'aspect probabiliste de ces modèles. De ce fait, il est nécessaire d'introduire les concepts de base de la théorie des processus aléatoires et d'évoquer quelques résultats existants dans la littérature concernant la stationnarité (stricte et second ordre) l'ergodicité (faible et forte) et les équations aux récurrences stochastiques.

1.2.2 Notions élémentaires sur les processus aléatoires

Dans des domaines assez variés : l'économie, les finances, la physique, la biologie, etc., on rencontre une multitude de phénomènes évolutifs dans le temps, dans l'espace, où dans les deux en même temps. Afin de pouvoir étudier leur structure probabiliste et statistique, nous utilisons les processus stochastiques (aléatoires) comme modèle mathématique.

Ainsi, dans ce qui suit, nous donnons la définition des processus stochastiques évoluant dans une seule dimension (scalaire).

Définition 1.2.1

On appelle processus stochastique de domaine d'évolution \mathbb{T} , défini sur un espace probabilisé (Ω, \mathcal{F}, P) , une famille de variables aléatoires $\{X_t, t \in \mathbb{T}\}$, à valeurs dans un espace des états E muni d'une tribu \mathcal{E} .

Remarques :

- 1) Si l'espace des états $(E, \mathcal{E}) \subset (\mathbb{N} \text{ ou } \mathbb{Z}, \mathcal{P}(\mathbb{N}) \text{ ou } \mathcal{P}(\mathbb{Z}))$ (dans le sens où l'ensemble des états possibles E est dénombrable), alors, on est en présence d'un processus aléatoire à *espace des états discret*. Par contre, si l'ensemble des états possibles est non dénombrable ($E \subset \mathbb{R}$), on dit que le processus est à *espace des états continu*.
- 2) Si le domaine d'évolution $\mathbb{T} = \mathbb{R}$, on parle de processus en temps *continu*.
- 3) Si le domaine d'évolution $\mathbb{T} = \mathbb{Z}$ ou \mathbb{N} , on parle de processus en temps *discret*.

Dans notre travail, nous considérons que le domaine d'évolution \mathbb{T} est le temps et nous

nous intéressons à la fois aux processus en temps discret et aux processus en temps continu mais à espace des états continu.

1.2.3 Stationnarité et ergodicité

Ces deux notions jouent un rôle important en analyse des séries chronologiques, en particulier lors de la recherche des conditions d'existence et d'unicité de solutions des modèles décrivant les phénomènes sous-jacent d'une part, et l'étude des propriétés statistiques (convergence, normalité asymptotique, etc.) d'autre part.

Dans la suite de cette section, nous considérons un processus aléatoire $\{X_t, t \in \mathbb{Z}\}$, défini sur un espace de probabilité (Ω, \mathcal{F}, P) à valeurs dans un espace des états $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. (sauf mention contraire).

Stationnarité

Dans la théorie des processus stochastiques, une place particulièrement importante est tenue par les processus dont les lois de probabilité présentent une invariance pour toute translation dans le temps (nous considérons ici que le domaine d'évolution \mathbb{T} est l'ensemble des temps). Cette propriété d'invariance temporelle est couramment utilisée, en analyse des séries chronologiques. Elle nous permet de travailler avec n'importe quelle portion de processus. Ce concept est formulé comme suit :

Stationnarité stricte des processus aléatoires

Définition 1.2.2

Un processus aléatoire $\{X_t, t \in \mathbb{Z}\}$, est dit strictement stationnaire si, et seulement si,

$$P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n) = P(X_{t_1+h} \leq x_1, \dots, X_{t_n+h} \leq x_n) \\ \forall n \in \mathbb{N}^*, \forall h, t_i \in \mathbb{Z}, \text{ et } \forall x_i \in \mathbb{R} \quad i = 1, \dots, n.$$

Autrement dit, les distributions fini-dimensionnelles de toutes sous suites finies $(X_{t_1}, \dots, X_{t_n})$, sont invariantes par translation sur le domaine d'évolution.

Le processus strictement stationnaire le plus simple est la suite *i.i.d.*

Stationnarité faible des processus aléatoires

Le concept précédent nécessite la connaissance de la loi sous-jacente qui n'est pas aisée à vérifier dans la pratique. Il est souvent plus utile de définir la stationnarité de manière moins stricte que celle donnée ci-dessus, à savoir la stationnarité au second ordre, ou stationnarité faible qui nécessite la connaissance des deux premiers moments du processus pour que l'on puisse étudier sa structure probabiliste et les propriétés statistiques des paramètres sous-jacents.

Définition 1.2.3

Un processus aléatoire $\{X_t, t \in \mathbb{Z}\}$, est dit faiblement stationnaire (stationnaire au second ordre, ou stationnaire en covariance) si, et seulement si les trois conditions suivantes sont vérifiées :

- (i) $\mathbb{E}(X_t^2) < \infty$ existe, $\forall t \in \mathbb{Z}$,
- (ii) $\mathbb{E}(X_t) = m$, $\forall t \in \mathbb{Z}$, (la constance de la fonction moyenne),
- (iii) $Cov(X_t, X_{t+h}) = \gamma_X(h)$, $\forall t, h \in \mathbb{Z}$, (dépend seulement de l'ordre du retard).

Le processus faiblement stationnaire le plus simple est le processus bruit blanc.

Ergodicité

La notion d'ergodicité d'un processus stationnaire traduit le fait qu'il peut prendre n'importe quelle valeur dans l'espace des états indépendamment de sa valeur initiale, et qu'il ne peut être absorbé par un sous ensemble strictement inclu dans cet espace des états.

Grâce à cette notion d'ergodicité, le théorème ergodique peut être établi pour les processus stationnaires (strictement, faiblement) qui est une extension des théorèmes des lois des grands nombres (loi forte et loi faible) dans le cas *i.i.d.*, stipulant que les moyennes temporelles convergent vers les moyennes théoriques.

Dans un premier temps, nous présentons les définitions formelles de la notion d'ergodicité dans un cadre général et nous les projetons sur les processus aléatoires, dans un deuxième temps.

Définition 1.2.4

Soit l'espace probabilisé (Ω, \mathcal{F}, P) . Une application $\mathcal{T} : \Omega \rightarrow \Omega$ est dite une transformation préservant la mesure P si elle est \mathcal{F} mesurable et $P(\mathcal{T}^{-1}A) = P(A)$, pour chaque ensemble $A \in \mathcal{F}$. De plus, si \mathcal{T} est bijective et \mathcal{T}^{-1} est \mathcal{F} -mesurable, il s'ensuit que \mathcal{T} est inversible ; dans ce cas \mathcal{T}^{-1} préserve la mesure : $P(A) = P(\mathcal{T}^{-1}\mathcal{T}A) = P(\mathcal{T}A)$.

Remarques :

- 1) Si l'image par \mathcal{T} de l'ensemble A est égale à cet ensemble lui-même à savoir, $\mathcal{T}(A) = A$, dans ce cas, A est dit \mathcal{T} -invariant. Ainsi, tout ensemble de probabilité 0 ou 1 est invariant.
- 2) Toute transformation bijective préservant la mesure, peut être utilisée pour engendrer des processus stochastiques strictement stationnaires.
- 3) Chaque processus strictement stationnaire a une distribution fini-dimensionnelle identique à celle d'un processus donnée par une transformation bijective préservant la mesure.

Après avoir exposé la définition d'une transformation préservant la mesure et celle des ensembles invariants, nous lançons la définition suivante de l'ergodicité.

Définition 1.2.5

Une transformation bijective préservant la mesure \mathcal{T} est dite ergodique si chaque ensemble A \mathcal{T} -invariant est de probabilité 1 ou 0.

Maintenant, nous allons considérer le cas des processus aléatoires. Commençons d'abord par définir le concept d'invariance d'un espace par rapport à une transformation bijective donnée \mathcal{T} .

Définition 1.2.6

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

(i) opérateur de translation dans le temps (shift) si

$$\mathcal{T}(\dots, X_{-1}, X_0, X_1, \dots) = (\dots, X_0, X_1, X_2, \dots),$$

(ii) invariante si $\forall A \subset \mathbb{R}^{\mathbb{Z}}, \mathcal{T}(A) = A$, où

$$\mathcal{T}(A) = \{ \mathcal{T}(\dots, X_{-1}, X_0, X_1, \dots), (\dots, X_0, X_1, X_2, \dots) \in A \subset \mathbb{R}^{\mathbb{Z}} \}.$$

Lorsque le processus $\{X_t, t \in \mathbb{Z}\}$, est strictement stationnaire et de moyenne finie m , la suite des moyennes temporelles (empiriques) $\{\bar{X}_t, t \in \mathbb{Z}\}$ où $\bar{X}_t = \frac{1}{n} \sum_{t=1}^n X_t$ converge presque sûrement vers une certaine variable aléatoire. Mais, si en outre le processus remplit la propriété d'ergodicité, cette variable ne peut être que la variable aléatoire dégénérée (une constante) *i.e.*, la moyenne m .

Théorème 1.2.1 (Karlin et Taylor, 1975)

Soit $\{X_t, t \in \mathbb{Z}\}$, un processus strictement stationnaire ayant une moyenne finie m . Alors, on a :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n X_t = m,$$

avec une probabilité égale à 1 (c'est la convergence presque sûre).

D'un point de vue pratique, l'intérêt de ce théorème réside dans le fait que les propriétés statistiques des processus strictement stationnaires et ergodiques peuvent être obtenues par l'observation d'une seule réalisation, mais sur une période suffisamment grande.

Certaines transformations de suites stationnaires et ergodiques restent stationnaires et ergodiques. C'est la stabilité de la classe des processus stationnaires et ergodiques par translation mesurable.

Théorème 1.2.2 (Breiman, 1968)

Si $\{X_t, t \in \mathbb{Z}\}$, est une suite strictement stationnaire ergodique et si $\{Y_t, t \in \mathbb{Z}\}$, est défini par

$$Y_t = f(\dots, X_{t-1}, X_t, X_{t-2}, \dots),$$

où f est une fonction mesurable de $\mathbb{R}^{\mathbb{Z}}$ dans \mathbb{R} , alors $\{Y_t, t \in \mathbb{Z}\}$, est également une suite strictement stationnaire et ergodique.

Le deuxième type d'ergodicité est l'ergodicité faible, dont la définition est la suivante.

Définition 1.2.7

Un processus faiblement stationnaire $\{X_t, t \in \mathbb{Z}\}$, dont les deux premiers moments existent et sont finis i.e, $\mathbb{E}(X_t) = \mu < \infty$ et $\mathbb{E}(X_t^2) < \infty \quad \forall t \in \mathbb{Z}$, est dit ergodique pour la moyenne, si la moyenne temporelle $\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t$ converge en probabilité vers μ lorsque n tend vers l'infini.

1.2.4 Équations aux récurrences stochastiques

En analyse des séries chronologiques, de nombreux modèles peuvent être exprimés sous forme d'une équation aux récurrences stochastiques

$$X_t = A_t X_{t-1} + B_t, \quad t \in \mathbb{Z}, \quad A_t \in \mathcal{M}(d) \text{ et } X_t, B_t \in \mathbb{R}^d, \quad (1.2.1)$$

où B_t est un vecteur aléatoire et $\mathcal{M}(d)$, désigne l'espace des matrices carrées aléatoires de dimension d , dont la solution, lorsqu'elle existe, représente le processus modélisant le phénomène sous-jacent. Un modèle satisfaisant l'écriture (1.2.1) est dit modèle autorégressif généralisé.

Le coefficient de plus grand exposant de Lyapounov γ

L'outil principal pour l'étude de la stricte stationnarité est le concept du coefficient de plus grand exposant de Lyapounov. Dans ce qui suit, nous introduisons quelques résultats concernant ce coefficient. Considérons une matrice réelle A , de dimension d . Son rayon spectral, noté, $\rho(A)$ est défini comme étant la plus grande valeur propre, en module. Nous avons la propriété suivante

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|A^t\| = \log \rho(A) \quad (1.2.2)$$

où $\|\cdot\|$ désigne une norme quelconque sur l'espace des matrices réelles $M(d)$.

Cette dernière propriété peut être étendue au cas des matrices aléatoires comme le montre le théorème suivant.

Théorème 1.2.3 (in Francq et Zakoïan, 2009)

Soit $\{A_t, t \in \mathbb{Z}\}$, une suite de matrices aléatoires, strictement stationnaires et ergodiques, telle que $\mathbb{E} \{\log^+ \|A_t\|\}^1$ est finie. On a

$$\lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \{\log \|A_t \dots A_1\|\} = \gamma = \inf_{t \in \mathbb{N}^*} \frac{1}{t} \mathbb{E} \{\log \|A_t \dots A_1\|\}, \quad (1.2.3)$$

et γ (resp. $\exp(\gamma)$) s'appelle plus grand exposant de Lyapounov (resp. rayon spectral) de la suite des matrices $\{A_t, t \in \mathbb{Z}\}$. De plus

$$\gamma = \lim_{t \rightarrow \infty} \frac{1}{t} \log \|A_t A_{t-1} \dots A_1\| \text{ p.s.} \quad (1.2.4)$$

Si la matrice carrée A_t est constante pour tout $t \in \mathbb{Z}$, en utilisant (1.2.2), nous retrouvons que $\gamma = \log \rho(A)$.

Comme les propriétés probabilistes (stationnarité, ergodicité, etc.) des modèles de séries chronologiques sont liées à celles des processus solutions de l'équation (1.2.1), il semble utile que l'on s'intéresse à l'une de ces propriétés, à savoir, par exemple, la stationnarité. Dans cette section, nous nous intéressons aux conditions d'existence et d'unicité d'une solution stationnaire de cette équation. Brandt (1986), a établi un théorème dans lequel il a donné les conditions suffisantes d'existence d'une unique solution strictement stationnaire pour (1.2.1). Il l'a fait dans le cadre unidimensionnel ($d = 1$), où la séquence (A_t, B_t) est strictement stationnaire et ergodique, mais la démonstration reste valable pour le cas multidimensionnel.

Théorème 1.2.4 (Brandt, 1986)

Soit $\{(A_t, B_t), t \in \mathbb{Z}\}$, un processus strictement stationnaire et ergodique tel que $\mathbb{E}(\log^+ \|A_0\|)$ et $\mathbb{E}(\log^+ \|B_0\|)$ sont toutes les deux finies. Supposons que le plus grand exposant de Lyapounov γ défini par

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

¹Tout au long de ce mémoire, nous utilisons les notations x^+ , x^- pour désigner $\max(x, 0)$ et $\max(-x, 0)$ resp.

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

$$X_t = \sum_{k=0}^{\infty} A_t A_{t-1} \dots A_{t-k+1} B_{t-k},$$

converge presque sûrement, et le processus $\{X_t, t \in \mathbb{Z}\}$, est l'unique solution strictement stationnaire de (1.2.1).

De plus, Bougerol et Picard (1992a) ont établi les conditions nécessaires et suffisantes pour l'existence d'une solution strictement stationnaire indépendante du futur (non anticipative), et cela dans le cas où (A_t, B_t) est une séquence de variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*). Ils stipulent que, sous une condition d'irréductibilité, s'il y a une solution non anticipative strictement stationnaire de l'équation (1.2.1), alors, le plus grand exposant de Lyapounov γ est strictement négatif.

Pour introduire le théorème de Bougerol et Picard (1992a) (théorème 1.2.6), nous commençons par présenter quelques définitions.

Définition 1.2.8

Un modèle autorégressif généralisé avec coefficients *i.i.d.* est un modèle qui satisfait l'écriture d'une équation aux récurrences stochastiques

$$X_t = A_t X_{t-1} + B_t, t \in \mathbb{Z}, \quad (1.2.5)$$

où $\{(A_t, B_t), t \in \mathbb{Z}\}$, est une séquence donnée de variables aléatoires *i.i.d.* définie sur l'espace probabilisé (Ω, \mathcal{F}, P) à valeurs dans $\mathcal{M}(d) \times \mathbb{R}^d$. Une solution de cette équation est une séquence $\{X_t, t \in \mathbb{Z}\}$, de variables aléatoires \mathbb{R}^d dimensionnelle vérifiant (1.2.5).

Après avoir défini un modèle autorégressif généralisé, nous passons à la définition d'une solution non anticipative.

Définition 1.2.9

Une solution strictement stationnaire non anticipative de (1.2.5) est un processus strictement stationnaire $\{X_t, t \in \mathbb{Z}\}$, qui est une solution de (1.2.5) tel que, pour chaque entier p , X_p est indépendante des variables aléatoires $\{(A_t, B_t), t > p\}$.

Enfin, nous présentons la définition d'un modèle irréductible.

Définition 1.2.10

Un sous espace affine H de \mathbb{R}^d est dit invariant sous le modèle (1.2.5) si $\{A_0x + B_0, x \in H\}$, est contenue dans H presque sûrement. Le modèle (1.2.5) est dit irréductible si \mathbb{R}^d est le seul sous espace affine invariant.

Théorème 1.2.5 (Bougerol et Picard, 1992a)

Considérons un modèle autorégressif généralisé donné par (1.2.5) avec des coefficients *i.i.d.* Supposons que ce modèle est irréductible et qu'il a une solution strictement stationnaire non anticipative $\{X_t, t \in \mathbb{Z}\}$. Alors les trois assertions suivantes sont vérifiées

- (i) $A_0A_{-1}\dots A_{-k}$ convergent vers 0 presque sûrement quand $k \rightarrow +\infty$,
- (ii) Pour chaque entier t ,

$$X_t = \sum_{k=0}^{\infty} A_t A_{t-1} \dots A_{t-k+1} B_{t-k},$$

où la série converge presque sûrement,

- (iii) Cette solution est l'unique solution strictement stationnaire de (1.2.5).

Le résultat principal de Bougerol et Picard (1992a) est traduit par le théorème suivant.

Théorème 1.2.6 (Bougerol et Picard, 1992a)

Supposons que le modèle autoregressif généralisé (1.2.5) avec des coefficients *i.i.d.* est irréductible et $\mathbb{E}(\log^+ \|A_0\|)$ et $\mathbb{E}(\log^+ \|B_0\|)$ sont tous les deux finis. Alors, le modèle (1.2.5) a une solution strictement stationnaire non anticipative si et seulement si le plus grand exposant de Lyapounov γ est strictement négatif.

Proposition 1.2.1 (Bougerol et Picard, 1992a)

Supposons que le modèle autoregressif généralisé (1.2.5) a une solution strictement stationnaire non anticipative $\{X_t, t \in \mathbb{Z}\}$. Soit H un sous espace affine minimal de \mathbb{R}^d tel que $P(X_0 \in H) = 1$. Alors, H est invariant sous le modèle (1.2.5) et chaque sous espace invariant de H porte une solution strictement stationnaire non anticipative.

Une conséquence immédiate du théorème (1.2.4), théorème (1.2.6) et de la proposition citée ci-dessus (1.2.1) est le corollaire suivant.

Corollaire 1.2.1 (Bougerol et Picard, 1992a)

Considérons un modèle où $\mathbb{E}(\log^+ \|A_0\|)$ et $\mathbb{E}(\log^+ \|B_0\|)$ sont tous les deux finis. Supposons qu'il existe une solution strictement stationnaire non anticipative qui n'est pas portée par un hyperplan affine. Alors, les trois conditions suivantes sont équivalentes.

- (i) *Le plus grand exposant de Lyapounov est strictement négatif.*
- (ii) *Le modèle est irréductible.*
- (iii) *Il existe une unique solution stationnaire.*

Lemme 1.2.1 (Bougerol et Picard, 1992a)

Soit $\{A_t, t \in \mathbb{N}\}$, une suite ergodique et strictement stationnaire de matrices aléatoires à valeurs dans $\mathcal{M}(d)$, telle que $\mathbb{E}(\log^+ \|A_0\|)$ est finie, alors

$$\lim_{t \rightarrow \infty} \|A_t \dots A_1\| = 0, p.s \Rightarrow \gamma < 0, \quad (1.2.6)$$

où γ est le plus le grand exposant de Lyapounov associé à cette suite de matrices aléatoires.

1.3 Processus d'Itô et Quelques processus stochastiques en temps continu

Le mouvement brownien standard unidimensionnel

Un exemple particulièrement important de processus stochastique à temps continu est le mouvement brownien. Il servira de base pour la construction de la plupart des modèles d'actifs financiers et de taux d'intérêt. Ce processus trouve son origine dans la biologie, où le botaniste Robert Brown (1827), avait observé au microscope le mouvement irrégulier des particules de pollen en suspension dans l'eau. Ces particules étaient animées de mouvements erratiques incessants. En 1900, Louis Bachelier introduit le mouvement brownien pour modéliser la dynamique des prix des actions boursiers. Par suite, en 1905, Smoluchowski décrit ce processus comme une limite de promenades aléatoires ; à la même année Albert Einstein

construit un modèle probabiliste de mouvement brownien. La première étude mathématique rigoureuse d'un tel processus est faite par Robert Wiener (1923). Depuis le mouvement brownien continue de passionner les probabilistes, aussi bien pour l'étude de ses trajectoires que pour la théorie de l'intégration stochastique.

Pour modéliser le fait que l'incertitude des évènements de Ω devient de moins en moins incertaine lorsque le temps s'écoule, autrement dit, on possède de plus en plus d'information, la notion de filtration a été introduite. Une filtration sur (Ω, \mathcal{F}, P) est une famille croissante de sous tribus de \mathcal{F} : $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ pour tout $0 \leq s \leq t$, où \mathcal{F}_t représente l'information disponible jusqu'à l'instant t et elle augmente au fur et à mesure que le temps augmente, notée $\mathbb{F} = (\mathcal{F}_t, 0 \leq t \leq \infty)$. Le quadruplet $(\Omega, \mathcal{F}, \mathbb{F}, P)$ est appelé espace probabilisé et *filtré*. Un exemple d'une filtration du processus stochastique $X = \{X_t, 0 \leq t < \infty\}$, est la *filtration naturelle* (ou *canonique*) $\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$.

Cette filtration est la plus petite σ -algèbre par rapport à laquelle X_s est mesurable pour tout $0 \leq s \leq t$. \mathcal{F}_t^X s'interprète comme toute l'information qu'on peut extraire de l'observation des trajectoires de X entre 0 et t .

On dit que le processus $\{X_t\}$, est adapté à la filtration \mathbb{F} ou $\{X_t\}$ est \mathbb{F} -adapté si X_t est \mathcal{F}_t -mesurable, $\forall t \geq 0$.

Définition 1.3.1

On appelle *mouvement brownien standard unidimensionnel* (ou *processus de Wiener*), un processus stochastique $W = \{W_t, t \geq 0\}$, adapté construit sur l'espace probabilisé et filtré $(\Omega, \mathcal{F}, \mathbb{F}, P)$ à valeurs réelles, tels que :

- (i) $\forall \omega \in \Omega, W_0(\omega) = 0$ p.s.,
- (ii) $\forall 0 \leq t_0 \leq t_1 \leq \dots \leq t_m$, les variables aléatoires $W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}}$, sont indépendantes,
- (iii) $\forall s, t \geq 0$ tels que $s < t$, la variable aléatoire $W_t - W_s$ est de loi normale de moyenne nulle et de variance $t - s$,
- (iv) $\forall \omega \in \Omega$, la trajectoire $t \rightarrow W_t(\omega)$ est continue.

Notons que cette définition suffit à garantir que les accroissements d'un mouvement

brownien sont stationnaires.

Afin de prendre en compte le comportement très erratique des cours des actifs financiers, Louis Bachelier les modélise à l'aide d'un mouvement brownien avec tendance (*drift*), mouvement brownien arithmétique

$$S_t = S_0 + \mu t + \sigma W_t,$$

où μ est la tendance et σ , la volatilité de l'actif. Dans une telle modélisation la variable S_t peut prendre des valeurs négatives, ce qui n'a pas de sens financier.

Équation différentielle stochastique unidimensionnelle

Une équation différentielle stochastique (ÉDS) est une généralisation de la notion d'équation différentielle, prenant en compte un terme de bruit blanc.

Les ÉDS permettent de modéliser des trajectoires aléatoires, tels que les cours boursiers, ou les mouvements de particules soumises à des phénomènes de diffusion. Les domaines d'application des ÉDS sont vastes : physique, biologie, écologie, finance mathématique, traitement du signal, dynamique des populations, etc.

Définition 1.3.2

Une équation différentielle stochastique est une équation de la forme

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t, \quad X_0 = x_0, \quad (1.3.1)$$

où l'équation intégrale stochastique suivante est dite solution de (1.3.1)

$$X_t = x_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

avec μ (*drift*) et σ (*coefficient du diffusion*) sont des fonctions mesurables définies sur $[0, \infty[\times \mathbb{R}$, $x_0 \in \mathbb{R}$ et le processus $(W_t, t \geq 0)$ est un mouvement brownien standard dans \mathbb{R} .

La première intégrale $\int_0^t \mu(s, X_s) ds$ est de Riemann ou de Lebesgue, quant à la deuxième *i.e.*, $\int_0^t \sigma(s, X_s) dW_s$, est l'intégrale stochastique au sens d'Itô. Cette dernière est plus dé-

licate à définir. De façon générale, il s'agit de définir l'intégrale stochastique $I \equiv \int_0^t Y_s dW_s$ où $\{Y_t, 0 \leq t \leq T\}$ est un processus stochastique, et de spécifier les conditions à imposer à Y_t pour que cette définition ait sens. On peut penser à définir I , dans l'esprit d'une intégrale de Stieltjes, en suivant la méthode standard qui consiste, à partir d'une partition $[t_0 = 0, t_1, \dots, t_k, \dots, t_m = t]$ de l'intervalle $[0, t]$, à considérer les sommes de Riemann $\mathbf{S}_m = \sum_{k=1}^m Y_{t_k^*} (W_k - W_{k-1})$ où $t_k^* \in (t_{k-1}, t_k)$ et à définir I comme la limite de \mathbf{S}_m (presque sûre ou en moyenne quadratique) quand la partition devient de plus en plus fine. On peut montrer que les \mathbf{S}_m convergent si le processus Y_t satisfait certaines conditions, mais que la limite dépend de la position assignée à t_k^* dans l'intervalle $[t_{k-1}, t_k]$. Il est d'usage de choisir la spécification de l'intégrale correspondant à $t_k^* = t_{k-1}$, donc de définir I comme la limite des sommes $\mathbf{S}_m = \sum_{k=1}^m Y_{t_{k-1}} (W_k - W_{k-1})$. Ainsi définie, il s'agit de l'intégrale d'Itô qui la seule utilisée en finance².

Le problème est de montrer que, sous certaines conditions sur μ et σ , l'ÉDS donnée en (1.3.1) admet une unique solution, comme le montre le théorème suivant.

Théorème 1.3.1 (existence et unicité de la solution des ÉDS)

Soient $T > 0$ et $\mu(.,.) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$; $\sigma(.,.) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ des fonctions mesurables satisfaisant

$$|\mu(t, x)| + |\sigma(t, x)| \leq C(1 + |x|), \quad t \in [0, T], x \in \mathbb{R},$$

où C est une constante, tel que

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq D|x - y|; \quad t \in [0, T], x, y \in \mathbb{R},$$

où D est une constante. Soit Z une variable aléatoire indépendante de la σ -algèbre engendrée par $W_s, s \geq 0$, tel que $\mathbb{E}|Z| < \infty$.

Alors, ÉDS

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t, \quad 0 \leq t \leq T, X_0 = Z,$$

admet une solution. Cette solution est unique, continue et adaptée à la filtration générée par

²Pour plus de détails, voir R. Portait & P. Poncet (Finance de Marché), B. Oksendal (Stochastic Differential Equations) et I. Karatzas & S.E Shreve (Brownian Motion and Stochastic Calculus).

Z et $\{W_s, s \leq t\}$, avec

$$\mathbb{E} \left(\int_0^T |X_t|^2 dt \right) < \infty.$$

Nous nous limitons ici à des ÉDS temporellement homogènes, dans lesquelles les fonctions μ et σ ne dépendent pas de t , autrement dit

$$dX_t = \mu(X_t) dt + \sigma(X_t) dW_t, \quad X_0 = x_0,$$

Néanmoins, la formule (ou lemme) d'Itô sera donné dans sa version généralisée.

1.3.1 Le processus d'Itô unidimensionnel

Soit $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t, t \geq 0\}, P)$ un espace probabilisé et filtré et $\{W_t\}$, un mouvement brownien standard dans \mathbb{R} .

Définition 1.3.3

On appelle processus d'Itô unidimensionnel, un processus $\{X_t\}$, à valeurs dans \mathbb{R} tel que :

$$X_t = X_0 + \int_0^t u(s, \omega) ds + \int_0^t v(s, \omega) dW_s,$$

tels que :

- u et v sont des processus adaptés à la filtration \mathbb{F} .
- $P \left(\int_0^t |u(\omega, s)| dt < \infty \right) = 1$
- $P \left(\int_0^t v(\omega, s)^2 dt < \infty \right) = 1.$

Écrit sous sa forme différentielle, le processus d'Itô devient

$$dX_t = udt + vdW_t.$$

1.3.2 Formule d'Itô

Le lemme d'Itô établi par Itô est l'équivalent stochastique du théorème fondamental du calcul de dérivée (il faut noter qu'un terme s'ajoute). Il nous permettra de déterminer

l'équation différentielle stochastique satisfaite par certains processus stochastiques donnés. Le théorème fondamental du calcul de dérivée stipule que si $\frac{df}{dx} : \mathbb{R} \rightarrow \mathbb{R}$ représente la dérivée de la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, alors $f(b) - f(a) = \int_a^b \frac{df}{dx}(x) dx$

Théorème 1.3.2 (formule d'Itô)

Soit $\{X_t\}$, un processus d'Itô i.e.,

$$dX_t = udt + vdW_t.$$

Soit $f(t, x) \in C^2([0, \infty[\times \mathbb{R})$ (i.e., f est deux fois continûment différentiable sur $[0, \infty[\times \mathbb{R}$).

Alors, $Y_t = f(t, X_t)$ est aussi un processus d'Itô et

$$dY_t = \frac{\partial f}{\partial t}(t, X_t) dt + \frac{\partial f}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t) (dX_t)^2, \quad (1.3.2)$$

où

$$\begin{cases} (dX_t)^2 = (dX_t)(dX_t), \\ (dt)(dt) = (dt)(dW_t) = (dW_t)(dt) = 0, \\ (dW_t)(dW_t) = dt. \end{cases}$$

Dans ce qui suit, nous donnons deux exemples de processus d'Itô, à savoir le mouvement brownien géométrique et le processus d'Orstein-Uhlenbeck.

Le mouvement brownien géométrique

Définition 1.3.4

Un processus stochastique de la forme $\{S_t, t \geq 0\}$, est dit mouvement brownien géométrique s'il est solution de l'ÉDS suivante

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \quad (1.3.3)$$

où μ et σ sont des constantes et $\{W_t\}$, représente un mouvement brownien standard unidimensionnel.

À l'aide du lemme d'Itô (équation (1.3.2)), il est possible de montrer que le processus stochastique $S = \{S_t, t \geq 0\}$, défini par

$$S_t = S_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right). \quad (1.3.4)$$

est solution à l'équation différentielle stochastique (1.3.3).

Comme W_t est de loi $N(0, t)$, alors, $\ln \left(\frac{S_t}{S_0} \right) = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t$ est de loi $N \left(\left(\mu - \frac{\sigma^2}{2} \right) t, \sigma^2 t \right)$, et S_t est de loi lognormale. Il est évident que dans ce cas, les cours des actifs financiers S_t restent toujours positifs. À noter que le rendement de détention R_t d'un titre risqué, lorsque la capitalisation se fait de façon continue, est défini par $R_t = \ln \left(\frac{S_t}{S_0} \right)$.

Processus d'Orstein-Uhlenbeck (mean-reverting process)

Certains processus stochastiques n'incorporent pas une dérive mais ont plutôt tendance à retourner vers un niveau moyen à plus ou moins long terme. C'est le cas des taux d'intérêt, des prix de certaines matières premières comme le pétrole, d'électricité et des taux de change de certaines devises. Nous sommes alors en présence d'un processus Ornstein-Uhlenbeck (O-U), encore appelé processus de retour vers la moyenne. Ce dernier a été proposé par les deux auteurs Orstein et Uhlenbeck (1930) dans un contexte physicien comme alternative aux mouvements browniens. Ce processus a connu une variété de domaines d'application ; en finance on le trouve sous l'appellation "modèle de Vasicek" mais la plupart des probabilistes utilisent l'appellation O-U de retour à la moyenne pour faire référence à ce modèle.

Définition 1.3.5

Un processus $\{X_t, t \geq 0\}$, est dit processus d'Orstein-Uhlenbeck (O-U) s'il est solution de l'équation différentielle stochastique (ÉDS) suivante :

$$dX_t = -\delta X_t dt + \sigma dW_t, \quad (1.3.5)$$

où δ : la vitesse de retour de $\{X_t, t \geq 0\}$ vers sa moyenne de long terme (vélocité).

σ : la volatilité de processus sous-jacent.

$(W_t)_{t \geq 0}$: mouvement brownien standard.

Les deux paramètres δ et σ sont strictement positifs

En appliquant la formule d'Itô (équation (1.3.2)) qui est l'une des pierres angulaires du calcul stochastique à la fonction aléatoire $X_t e^{\delta t}$, et par suite en intégrant ses deux membres, on obtient la forme explicite de la solution exacte de (1.3.5), à savoir

$$X_t = X_0 e^{-\delta t} + \sigma e^{-\delta t} \int_0^t e^{\delta s} dW_s, \quad X_0 = x_0. \quad (1.3.6)$$

Si nous notons Y_t , le taux d'intérêt à l'instant t et μ le niveau moyen d'équilibre de long terme de Y_t (l'équilibre à long terme), nous parlerons de processus de Vasicek (1977) (appelé encore processus d'O-U de retour à la moyenne), qui est une variante du processus d'O-U, où un coefficient de dérive constant a été ajouté. Alors, l'équation (1.3.5) devient

$$dY_t = -\delta (Y_t - \mu) dt + \sigma dW_t, \quad (1.3.7)$$

Remplaçons X_t dans l'équation (1.3.6) par $(Y_t - \mu)$, nous obtenons $Y_t - \mu = (Y_0 - \mu) e^{-\delta t} + \sigma e^{-\delta t} \int_0^t e^{\delta s} dW_s$, $Y_0 = y_0$, d'où la solution exacte de l'ÉDS donnée en (1.3.7) est

$$Y_t = Y_0 e^{-\delta t} + \mu (1 - e^{-\delta t}) + \sigma e^{-\delta t} \int_0^t e^{\delta s} dW_s.$$

Le processus $\{Y_t, t \geq 0\}$, est un processus gaussien de fonction d'espérance $\mathbb{E}(Y_t) = Y_0 e^{-\delta t} + \mu (1 - e^{-\delta t})$ et de fonction de variance $V(Y_t) = \frac{\sigma^2}{2\delta} (1 - e^{-2\delta t})$.

Autrement dit, $Y_t \sim N\left(Y_0 e^{-\delta t} + \mu (1 - e^{-\delta t}), \frac{\sigma^2}{2\delta} (1 - e^{-2\delta t})\right)$, où \sim désigne suit une loi de probabilité.

Remarques :

- 1) La distribution du processus $\{Y_t, t \geq 0\}$, converge (quand t tend vers l'infini) vers une distribution gaussienne d'espérance μ et de variance $\frac{\sigma^2}{2\delta}$.
- 2) L'un des inconvénients du modèle de Vasicek est qu'il suppose que la volatilité des taux d'intérêts est constante et indépendante du niveau de processus $\{Y_t, t \geq 0\}$.
- 3) La solution d'une équation différentielle stochastique (ÉDS) consiste à trouver le(s) processus stochastiques satisfaisant l'équation donnée. Or, le lemme d'Itô permet l'inverse,

c'est-à-dire permet de trouver l'ÉDS satisfaite par un processus stochastique donné. S'il advient que le processus stochastique choisi satisfasse l'ÉDS donnée, alors nous aurons trouvons une solution à cette équation. En ce sens, le lemme d'Itô nous fournit des solutions à des ÉDS. Cependant, il ne s'agit pas d'une méthode systématique permettant l'obtention de solutions. Son utilisation demande du flair.

1.4 Martingales et Semimartingales

1.4.1 Introduction

Dans cette section, nous considérons des processus stochastiques réels, définis sur un espace probabilisé, ayant pour domaine d'évolution, un intervalle de temps continu fini ou infini (sauf mention contraire).

Étant donné un processus stochastique $X = \{X_t, 0 \leq t < \infty\}$, ce dernier est dit *càd-làg* si pour chaque $\omega \in \Omega$, la trajectoire $X(\omega)$, est continue à droite et admet une limite à gauche presque sûrement.

Un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq \infty\}, P)$ est dit satisfaire les *conditions habituelles* si

- (i) La tribu \mathcal{F}_0 contient tous les ensembles P négligeables de \mathcal{F} .
- (ii) La filtration \mathbb{F} est continue à droite *i.e.*, $\mathcal{F}_{t+} = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t$, pour tout $0 \leq t < \infty$.

Dans ce cas, la filtration \mathbb{F} est dite satisfaire les hypothèses habituelles.

Intuitivement, la continuité à droite de \mathcal{F}_t signifie qu'ayant observé toute l'information disponible jusqu'en t inclus, on n'apprend rien de plus par une observation infinitésimale dans le futur. On dit que le processus est *non anticipatif*. Il est clair que tout processus X est adapté par rapport à sa filtration naturelle $\mathbb{F}^X = \{\mathcal{F}_t^X, 0 \leq t < \infty\}$.

Dans le reste de ce qui suit, nous supposons toujours que les hypothèses habituelles sont vérifiées.

1.4.2 Martingales

Martingale

Les processus stochastiques sont caractérisés par une structure de dépendance entre les membres de la famille des variables aléatoires qui le constituent. La propriété de martingale exprime une relation qui peut avoir lieu dans de nombreux contextes (la théorie des jeux, etc.), elle est devenue un outil indispensable dans la théorie des probabilités et les probabilités appliquées.

En calcul stochastique, une martingale désigne un type de processus tel que sa valeur espérée connaissant l'information disponible à une certaine date s est la valeur à cette même date.

Pour motiver l'idée de martingale, on considère une suite de variables aléatoires réelles intégrables (X_1, X_2, \dots) , définie sur (Ω, \mathcal{F}, P) . Supposons que X_t est le montant de la fortune que gagne le joueur au temps t . Dans le cas où le jeu est équitable, la fortune espérée à l'instant $t + 1$ connaissant les t résultats précédents du jeu *i.e.*, la tribu engendrée par (X_1, X_2, \dots, X_t) , est exactement le niveau du fortune obtenu au $t^{\text{ème}}$ instant. Une description mathématique de ce modèle est donnée par la définition suivante.

Définition 1.4.1

Une suite de variables aléatoires réelles $\{X_t, t \in \mathbb{N}\}$, définie sur (Ω, \mathcal{F}, P) associée d'une filtration $\{\mathcal{F}_t, t \in \mathbb{N}\}$, (*i.e.*, une suite croissante de tribus $\{\mathcal{F}_t, t \in \mathbb{N}\}$) est dite \mathcal{F}_t -martingale et on écrit $\{X_t, \mathcal{F}_t, t \in \mathbb{N}\}$, si et seulement si

- (i) $\{X_t, t \in \mathbb{N}\}$ est adapté à la filtration $\{\mathcal{F}_t, t \in \mathbb{N}\}$, autrement dit X_t est \mathcal{F}_t -mesurable, (l'information portée par le processus $\{X_t, t \in \mathbb{N}\}$ ne peut être meilleure que celle donnée par la filtration $\{\mathcal{F}_t, t \in \mathbb{N}\}$).
- (ii) $\mathbb{E}(|X_t|) < \infty, \forall t \in \mathbb{N}, i.e., X_t$ est intégrable pour tout entier t ,
- (iii) $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ presque sûrement (*p.s.*), pour tout $s \leq t, s, t \in \mathbb{N}$.

Remarques :

- 1) Si $\{X_t, t \in \mathbb{N}\}$, respecte les deux premières conditions de la définition (1.4.1), et

$\mathbb{E}(X_t | \mathcal{F}_s, \forall s \leq t) \geq X_s$, presque sûrement, alors, on l'appelle sous-martingale et si $\mathbb{E}(X_t | \mathcal{F}_s, \forall s \leq t) \leq X_s$, presque sûrement, on l'appelle sur-martingale.

2) Lorsque le domaine d'évolution est \mathbb{R}^+ , nous considérons la filtration $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$, et nous disons que nous sommes en présence d'une (sous/sur)martingale à temps continu.

Différence de martingale

Nous restons toujours dans le même contexte, mais cette fois-ci, on s'intéresse au gain net du joueur, et on cherche une description mathématique de ce phénomène. Pour ce faire, on considère une suite de martingales $\{X_t, t \in \mathbb{N}\}$, et on définit une nouvelle variable aléatoire $D_t = X_t - X_{t-1}$, $t \in \mathbb{N}^*$ (on prend $X_0 = 0$), en utilisant les propriétés de l'espérance conditionnelle et le fait que $\{X_t, t \in \mathbb{N}\}$, est une martingale, on démontre

$$\mathbb{E}(D_t | X_1, \dots, X_{t-1}) = X_{t-1} - X_{t-1} = 0.$$

D'où la définition suivante d'une suite de différence de martingale.

Définition 1.4.2

Une suite de variables aléatoires réelles $\{D_t, t \in \mathbb{N}^*\}$, définie sur (Ω, \mathcal{F}, P) associée d'une filtration $\{\mathcal{F}_t, t \in \mathbb{N}\}$, est dite une différence de martingale si et seulement si

- (i) $\{D_t, t \in \mathbb{N}^*\}$ est adapté à la filtration $\{\mathcal{F}_t, t \in \mathbb{N}\}$,
- (ii) $\mathbb{E}(|D_t|) < \infty, \forall t \in \mathbb{N}^*, i.e., X_t$ est intégrable pour tout entier t ,
- (iii) $\mathbb{E}(D_t | \mathcal{F}_s) = 0$ presque sûrement (p.s), pour tout $s \leq t, s, t \in \mathbb{N}$.

Dans le reste de ce qui suit nous ne considérons que les martingales en temps continu.

1.4.3 Semimartingales et martingales locales

Nous nous intéressons à savoir si un évènement donné, caractérisé par sa première date $T(\omega)$ d'apparition, a eu lieu ou non avant la date t sachant l'observation de l'information \mathcal{F}_t . Ceci conduit à la notion de temps d'arrêt³.

³Nous travaillons toujours sur l'espace probabilisé complet et filtré $(\Omega, \mathcal{F}, \mathbb{F}, P)$, satisfaisant les hypothèses habituelles.

Définition 1.4.3

Une variable aléatoire $T : \Omega \rightarrow [0, \infty]$, i.e., est appelé temps d'arrêt (par rapport à la filtration $\mathbb{F} = (\mathcal{F}_t, 0 \leq t \leq \infty)$) si :

$$\{T \leq t\} = \{\omega \in \Omega : T(\omega) \leq t\} \in \mathcal{F}_t, \forall t, 0 \leq t \leq \infty.$$

Si T est un temps d'arrêt, nous posons $t \wedge T = \min(t, T)$, $\forall t \geq 0$.

Définition 1.4.4

Étant donné un processus $\{X_t, t \geq 0\}$, et un temps d'arrêt T . Le processus X^T défini par $X_t^T = X_{T \wedge t}$ est dit processus arrêté en T .

À noter que si le processus $\{X_t, t \geq 0\}$ est adapté et càdlàg et si T est un temps d'arrêt, alors

$$X_t^T = X_{T \wedge t} = X_t \mathbb{1}_{[t < T]} + X_T \mathbb{1}_{[t \geq T]}.$$

Définition 1.4.5

Soit $M = \{M_t, 0 \leq t < \infty\}$ un processus stochastique. On dit que M est une \mathbb{F} -martingale locale s'il existe une suite croissante de \mathbb{F} -temps d'arrêt, $(T_n)_{n \geq 1}$ avec $\lim_{n \rightarrow \infty} T_n = \infty$ p.s. telle que

- (i) M_0 est \mathcal{F}_0 mesurable,
- (ii) pour tout $n \geq 0$, $(M_{t \wedge T_n} - M_0, t \geq 0)$ est une \mathbb{F} -martingale.

La suite de temps d'arrêt est dite réductrice pour la martingale locale M .

Remarques :

- 1) Nous n'avons pas besoin d'hypothèse d'intégrabilité sur M_0 .
- 2) Toute martingale est une martingale locale mais la réciproque n'est pas vraie : les martingales locales sont beaucoup plus générales que les martingales.

Il est intéressant d'avoir des conditions assurant qu'une martingale locale est une martingale. Le critère suivant est très utile en pratique.

Proposition 1.4.1

Soit $M = (M_t)_{0 \leq t < \infty}$ une martingale locale. Supposons que

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |M_s| \right] < \infty, \quad \forall t, 0 \leq t < \infty.$$

Alors, M est une martingale.

Nous introduisons quelques définitions qui seront utilisées par suite.

Définition 1.4.6

Un processus stochastique $A = \{A_t, t \geq 0\}$, est dit à variation finie si toutes ses trajectoires sont càd-làg et à variation finie sur chaque intervalle compact de \mathbb{R}^+ , i.e., pour tout $\omega \in \Omega$, $\sup \sum_{i=1}^m |A_{t_i}(\omega) - A_{t_{i-1}}(\omega)| < +\infty$, où le supremum porte sur toutes les subdivisions $0 = t_1 \leq t_2 \leq \dots \leq t_m = t$ sur intervalle $[0, t]$.

Définition 1.4.7

Un processus $X = \{X_t, t \geq 0\}$ est dit à variation quadratique finie, si pour tout $t \geq 0$, la limite en probabilité de $\sum_{i=1}^m (X_{t_i} - X_{t_{i-1}})^2$ sur la subdivision $\Delta = \{0 = t_1 < t_2 \dots < t_m = t\}$ de $[0, t]$ quand le pas $|\Delta| = \sup_{1 \leq k \leq m} |t_k - t_{k-1}|$ tend vers 0, est finie. On note

$$\langle X, X \rangle_t = p \lim_{|\Delta| \rightarrow 0} \sum_{i=1}^m (X_{t_{i+1}} - X_{t_i})^2,$$

où $\langle X, X \rangle_t$ est appelée variation quadratique de X et $p \lim$ désigne la limite en probabilité.

Théorème 1.4.1 (Dambis (1965), Dubins-Schwartz (1965))

Soit $M = \{M_t, t \geq 0\}$, une martingale locale continue telle que $\lim_{t \rightarrow \infty} \langle M, M \rangle_t = \infty$. Définissons le temps d'arrêt

$$T_\tau = \inf \{t \geq 0; \langle M, M \rangle_t > \tau\}, \forall \tau \geq 0$$

Alors, le processus $W_\tau = M_{T_\tau}$ est un mouvement brownien unidimensionnel. En particulier, la filtration $\{\mathcal{L}_\tau = \mathcal{F}_{T_\tau}, 0 \leq \tau < \infty\}$, satisfait les conditions usuelles et nous avons :

$$M_t = W_{\langle M, M \rangle_t}, 0 \leq t < \infty.$$

Nous introduisons finalement une classe fondamentale de processus à variation quadratique finie, étendant les sur (sous) martingales (locales), cette classe est largement utilisé dans la modélisation en finance. C'est la classe des processus de semimartingales.

Définition 1.4.8

Une semimartingale est un processus càd-làg adapté $X = \{X_t, t \geq 0\}$ admettant une décomposition de la forme :

$$X = X_0 + M + A \tag{1.4.1}$$

où M est une martingale locale continue càd-làg nulle en 0, et A est un processus adapté à variation finie et nul en 0. Une semimartingale continue est une semimartingale telle que dans la décomposition (1.4.1), M et A sont continus. Une telle décomposition où M et A sont continus, est unique.

Quelques exemples de processus de semimartingales sont donnés à travers les théorèmes et les corollaires suivants.

Théorème 1.4.2 (in Protter, 2004)

Un processus adapté de trajectoires càd-làg et à variation finie sur des ensembles compacts est une semimartingale.

Théorème 1.4.3 (in Protter, 2004)

Une martingale de carré intégrable, de trajectoires càd-làg est une semimartingale

Corollaire 1.4.1 (in Protter, 2004)

Une martingale locale avec des trajectoires continues forme une semimartingale.

Corollaire 1.4.2 (in Protter, 2004)

Le processus de Wiener est une semimartingale.

Corollaire 1.4.3 (in Protter, 2004)

Chaque martingale locale, càd-làg et localement de carré intégrable est une semimartingale.

et nous avons la propriété que pour tout $t \geq 0$, si $0 = t_1 \leq \dots \leq t_m = t$ est une subdivision de $[0, t]$ de pas tendant vers 0 tel que $\sup_{1 \leq k \leq m} |t_k - t_{k-1}|$, nous avons la convergence en probabilité :

Nous définissons la variation quadratique d'une semimartingale continue $X = M + A$ par

$$\langle X, X \rangle_t = p \lim_{|\Delta| \rightarrow 0} \sum_{i=1}^m (X_{t_i} - X_{t_{i-1}})^2.$$

où $\Delta \{0 = t_1 < t_2 \dots < t_m = t\}$ subdivision de $[0, t]$. et par suite la propriété suivante :

$$\langle X, X \rangle = \langle M, M \rangle$$

2

Structure probabiliste et estimation du modèle GARCH (p, q)

2.1 Introduction

La modélisation des séries financières est un problème complexe. Cette complexité n'est pas uniquement due à la grande variété des séries utilisées (prix d'action, taux d'intérêt, taux de change,...), à l'importance de la fréquence d'observation (seconde, minute, heure, jour,...) ou à la disponibilité d'échantillons de très grande taille. Elle tient surtout à l'existence de régularités statistiques communes à de très grands nombres de séries financières (stylized facts) et difficiles à reproduire artificiellement à partir de modèles stochastiques. Les modèles linéaires *ARMA* décrits dans le fameux ouvrage de Box et Jenckis (1970) se révèlent incapables de capturer ces faits stylisés.

Tout commence en 1963, lorsque Mandelbrot constate deux faits sur les données financières. Premièrement, elles ont des queues plus épaisses que celles de la loi normale, *i.e.*, les valeurs loin de la moyenne apparaissent avec des probabilités plus élevées que pour une loi normale; le coefficient d'aplatissement étant élevé. Deuxièmement des regroupements de volatilité «*volatility clustering*», sont remarqués dans les faits empiriques. Pour répondre à ce besoin, les modèles *ARCH* ont été intrduits par Engle en 1982 et généralisés par la suite par

Bollerslev en 1986, en modèles $GARCH$.

Le présent chapitre consiste essentiellement à donner une représentation générale des modèles $GARCH$, autrement dit, leur structure probabiliste, ainsi que leur estimation. Dans un premier lieu, nous spécifions une écriture autorégressive généralisée du modèle qui nous permettra d'établir, à partir du plus grand exposant de Lyapounov, les propriétés probabilistes du modèle $GARCH$, à savoir, la stationnarité stricte et au second ordre, l'ergodicité, l'existence des moments supérieurs et la structure d'autocovariance. Ensuite, nous passons à l'estimation d'un tel modèle, en utilisant la méthode du quasi maximum de vraisemblance (QMV) où nous étudions les propriétés asymptotiques de ses estimateurs (la consistance et la normalité asymptotique). Enfin, une étude de simulation a été faite afin de consolider les résultats théoriques concernant la consistance de l'estimateur du QMV .

2.2 Faits stylisés des séries financières

Dans cette section, nous passons en revue les principaux faits stylisés des séries financières de prix et de rendements.

Sur les marchés financiers, on observe les prix des actifs financiers mais l'attention de l'analyse des séries financières se focalise plus sur les rendements que sur les prix. Cela est dû aux deux raisons suivantes.

- Raison théorique et empirique : les processus des rendements ont des propriétés probabilistes plus attractives que celles des processus des prix telles que la stationnarité et l'ergodicité.

- Raison économique : les rendements ont une signification et une interprétation économique ; ils représentent un gain d'investissement lors de la positivité des rendements, une perte s'il sont négatifs.

Notons $S_{n,\Delta}$ le cours intrajournalier en fonction du temps de la $n^{\text{ème}}$ journée ouvrable (normalisée à $[0, 1]$) d'un actif financier qui pourra être une action, un indice financier, un taux de change ou un contrat à terme. Alors le rendement logarithmique de cet actif financier est donné par la définition suivante.

Définition 2.2.1

Le rendement logarithmique (également appelé rendement) du cours d'un actif financier pour une échelle de temps Δ , de la $n^{\text{ème}}$ journée, noté $r_{n,\Delta}$, est donné par

$$r_{n,\Delta} = \log \left(\frac{S_{n,\Delta}}{S_{n,\Delta-1}} \right),$$

À noter que r_n correspond au rendement logarithmique journalier ($r_{n,1} = r_n$).

Maintenant, nous introduisons ces caractéristiques saillantes – que sont les faits stylisés – des séries financières, et nous les illustrons sur les données journalières de l'indice boursier S&P 500, sur la période 02/01/1990 au 31/12/2007 (4538 observations).

2.2.1 Non stationnarité de processus des prix

Le processus stochastique $S = \{S_n, n \geq 0\}$, associé aux prix (journaliers) des actifs, est généralement non stationnaire au second ordre et ses trajectoires sont proches de celles d'une marche aléatoire sans terme constant. Contrairement aux séries des prix, celles des rendements exhibent des trajectoires compatibles avec la stationnarité faible, voir le tracé de la série S&P 500 sur la figure 2.2.1a et 2.2.1b respectivement.

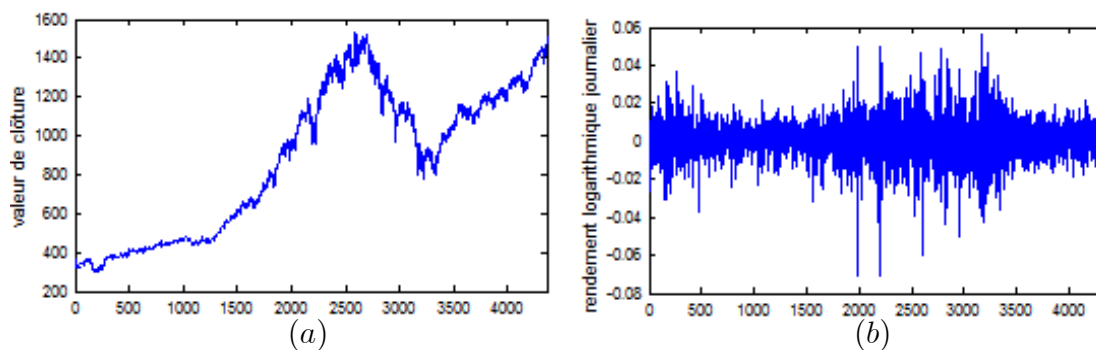


Figure 2.2.1 : (a) Valeurs de clôture de l'indice S&P 500, du 02/01/1990 au 31/12/2007 et (b) rendements logarithmiques journaliers correspondants.

Afin de détecter la non stationnarité, nous utilisons le test de racine unitaire de Dickey-Fuller (*Dickey-Fuller* (DF)) à partir des trois modèles suivants :

$$\text{modèle 1 : } \Delta S_n = \phi S_{n-1} + \varepsilon_n,$$

$$\text{modèle 2 : } \Delta S_n = \phi S_{n-1} + c + \varepsilon_n,$$

$$\text{modèle 3 : } \Delta S_n = \phi S_{n-1} + c + \beta n + \varepsilon_n.$$

Nous commençons par estimer le modèle 3, $\Delta S_n = \phi S_{n-1} + c + \beta n + \varepsilon_n$, incluant une constante et une tendance. la valeur estimée de la statistique DF est égale à -1.895 . Cette dernière est supérieure à la valeur critique -3.41 , au seuil 5%. Donc, nous acceptons l'hypothèse nulle de racine unitaire ($\phi = 0$).

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.895751	0.6565
Test critical values:		
1% level	-3.960061	
5% level	-3.410795	
10% level	-3.127192	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(S)
 Method: Least Squares
 Date: 11/21/11 Time: 10:45
 Sample (adjusted): 2 4538
 Included observations: 4537 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
S(-1)	-0.001578	0.000833	-1.895751	0.0581
C	0.720155	0.404217	1.781603	0.0749
@TREND(1)	0.000417	0.000246	1.693566	0.0904

Figure 2.2.2 : Test de racine unitaire sur la série S_n : modèle 3.

Il faut à présent évaluer la validité du modèle 3. Pour ce faire, nous testons la nullité du coefficient de la tendance conditionnellement à la présence d'une racine unitaire. Ce qui revient à effectuer le test suivant : $H_0^{(3)} : (c, \beta, \phi) = (c, 0, 0)$. La statistique de Fisher F_3^1

$^1 F_3 = \frac{(SCR_{3,c} - SCR_3)/2}{SCR_3/(N-3)}$, où $SCR_{3,c}$, c'est la somme des carrés des résidus du modèle 3 contraint sous $H_0^{(3)}$ et SCR_3 est la somme des carrés des résidus du modèle 3 non contraint.

nous donne une valeur égale à 1.79. Cette valeur est inférieure à la valeur critique 6.27 (lue à partir de la table de Dickey-Fuller), au seuil 5%. Ainsi, nous acceptons l'hypothèse $H_0^{(3)}$, et par suite le modèle 3 est remis en cause. (Figure 2.2.2)

Passons à l'estimation du modèle 2, $\Delta S_n = \phi S_{n-1} + c + \varepsilon_n$, où nous testons la présence d'une racine unitaire. Comme la valeur de la statistique DF est égale à -0.854 est supérieure à la valeur d' DF tabulée, au seuil 5%, nous acceptons l'hypothèse nulle de racine unitaire ($(\phi = 0)$) dans le modèle 2. De la même manière que précédemment (modèle 3), nous devons valider notre diagnostic en vérifiant que le modèle 2 est le bon modèle. Autrement dit, nous testons la nullité de la constante conditionnellement à la présence d'une racine unitaire. Ce qui revient à effectuer le test suivant : $H_0^{(2)} : (c, \phi) = (0, 0)$. En utilisant le test de Fisher de la nullité conjointe des coefficients du modèle. La statistique de Fisher F_2 égale à 0.729, lue directement à partir de figure (2.2.3) ci-dessous. Cette valeur est inférieure à la valeur critique 4.589 (lue à partir de la table de Dickey-Fuller), au seuil 5%. Alors, l'hypothèse $H_0^{(2)}$ est acceptée et nous devons recommencer ce test à partir du modèle 1 sans constante, ni tendance. (Figure 2.2.3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.854032	0.8029
Test critical values:		
1% level	-3.431608	
5% level	-2.861981	
10% level	-2.567048	

*MacKinnon (1996) one-sided p-values.
F-statistic = 0.729370

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(S)
Method: Least Squares
Date: 11/21/11 Time: 10:48
Sample (adjusted): 2 4538
Included observations: 4537 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
S(-1)	-0.000341	0.000400	-0.854032	0.3931
C	0.551827	0.391888	1.408125	0.1592

Figure 2.2.3 : Test de racine unitaire sur la série S_n : modèle 2.

Au seuil de 5%, $DF_{calculée} = 1.112 > DF_{tabulée} = -1.94$, nous acceptons l'hypothèse nulle de racine unitaire ($\phi = 0$) dans le modèle 1. (Voir la Figure (2.2.4) ci-dessous).

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	1.112654	0.9315
Test critical values:		
1% level	-2.565472	
5% level	-1.940894	
10% level	-1.616653	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(S)
 Method: Least Squares
 Date: 11/21/11 Time: 10:52
 Sample (adjusted): 2 4538
 Included observations: 4537 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
S(-1)	0.000176	0.000158	1.112654	0.2659

Figure 2.2.4 : Test de racine unitaire sur la série S_n : modèle 1.

Les résultats du test DF , donnés dans les figures 2.2.2, 2.2.3 et 2.2.4, nous permettent de conclure que la série des prix de l'indice S&P 500 est issue d'un processus non stationnaire, possédant une tendance stochastique qui peut être représentée par une pure marche aléatoire : $S_n = S_{n-1} + \varepsilon_n$ où ε_n est un bruit blanc.

2.2.2 Regroupement des volatilités (volatility clustering)

La volatilité fait référence au degré de fluctuation des rendements des actifs; empiriquement, de fortes (faibles) valeurs ou de fortes (faibles) variations, tendent à être suivies par d'autres grandes (petites) variations *i.e.*, de même ampleur, mais pas nécessairement dans le même sens. C'est le phénomène de regroupement (accumulation) des volatilités. Ce regroupement des volatilités est entièrement dû aux corrélations des séries financières.

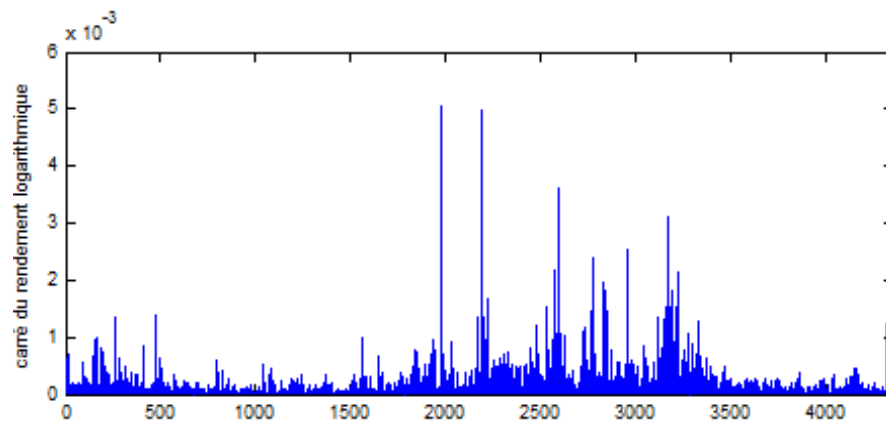


Figure 2.2.5 : Carrés des rendements logarithmiques journaliers de l'indice S&P 500

Sur la figure (2.2.5), nous avons présenté les carrés des rendements logarithmiques journaliers de l'indice S&P 500. Le comportement intermittent des rendements logarithmiques journaliers est très visible. On voit clairement le regroupement des volatilités par paquets (la même observation peut être tirée de la visualisation de la figure 2.2.1 (b)).

2.2.3 Absence d'autocorrélation des rendements logarithmiques

La série des variations du prix (rendements logarithmiques) présente, généralement, de très faibles autocorrélations, la rendant proche d'un bruit blanc (figure 2.2.6a). C'est la traduction immédiate de l'hypothèse des marchés efficients.

La justification la plus fréquente de cette absence de corrélation invoque le principe d'absence d'arbitrage des marchés liquides. Si de telles corrélations, faciles à détecter, existaient, il serait aisé pour un intervenant de concevoir une stratégie d'arbitrage qui en moyenne lui permettrait de gagner de l'argent. Or, la mise en œuvre de ces stratégies, devrait éliminer mécaniquement ces corrélations, au moins pour des temps supérieurs au temps de réaction du marché, qui typiquement est de l'ordre de quelques minutes pour revenir à l'équilibre. Néanmoins, pour des séries intra-journalières avec intervalles de temps très courts ($\Delta < 5$ minutes), des autocorrélations significatives peuvent apparaître en raison des effets de microstructure (voir le chapitre 3). En revanche, les séries des carrés ou des valeurs absolues des rendements sont souvent fortement autocorrélées. (figures 2.2.6b, 2.2.9 ci-dessous).

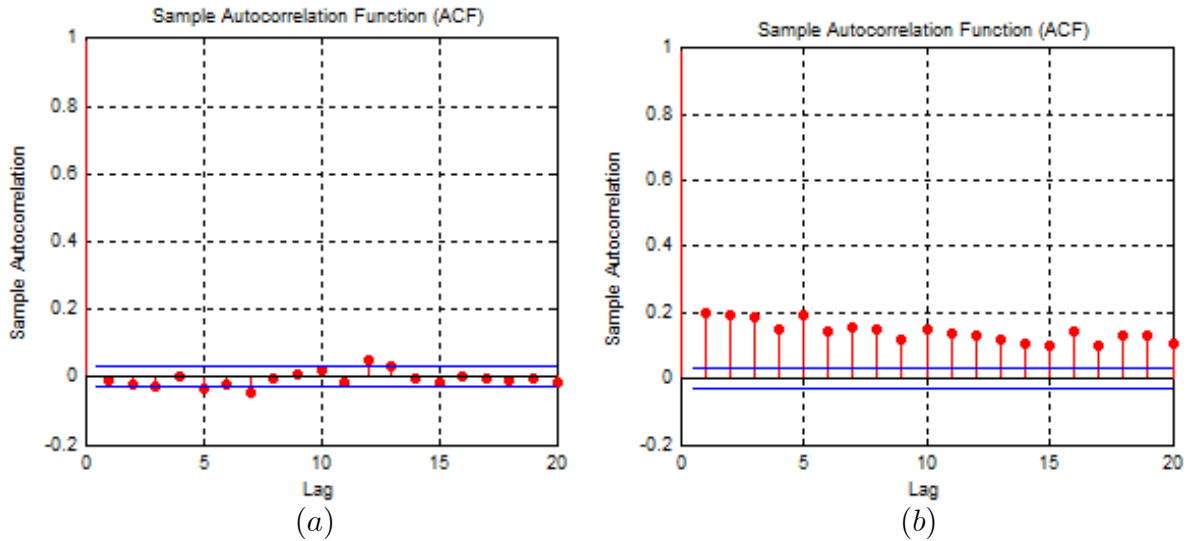


Figure 2.2.6 : (a) Autocorrélations empiriques des rendements logarithmiques journaliers et (b) leur carré pour l'indice S&P 500.

2.2.4 Distribution des rendements

La première modélisation des cours boursiers par un mouvement brownien, repose sur l'hypothèse que les rendements (logarithmiques) suivent une loi gaussienne. Cette modélisation reste encore aujourd'hui au cœur de la théorie financière.

Néanmoins, la pratique nous montre que les distributions empiriques des rendements logarithmiques ont des queues plus épaisses que la distribution normale. Ce constat a été signalé pour la première fois par Mandelbrot (1963) pour les prix du coton et a été, depuis lors, observé pour différents marchés. Ce sera encore plus flagrant si l'on considère les rendements sur de petits intervalles de temps, ce qui n'est possible que grâce aux données à haute fréquence.

Une manière de quantifier plus précisément les déviations par rapport à une loi normale est de calculer la *kurtosis* et la *skewness* des rendements logarithmiques.

Queues de distribution épaisses (excès de kurtosis)

L'hypothèse de normalité des rendements est généralement rejetée. Les queues des distributions de probabilités empiriques des rendements sont généralement plus épaisses que

celles d'une loi gaussienne. On parle alors de distribution *leptokurtique*.

Une mesure de cette effet est obtenue à partir du coefficient d'aplatissement (la kurtosis) dont on rappelle la définition.

Définition 2.2.2

On appelle *kurtosis* d'une variable aléatoire X , de moyenne μ , de variance σ_X^2 et dont le moment d'ordre 4 existe, le nombre réel

$$Kr = \frac{\mathbb{E}(X - \mu_X)^4}{\sigma_X^4}.$$

Cette kurtosis est comparée à celle d'une loi normale qui est égale à 3.

- Lorsque $Kr = 3$, la distribution correspondante est dite *mésokurtique*, c'est le cas de la loi normale.
- Si $Kr > 3$, une telle variable est caractérisée par une distribution pointue, par conséquent des queues plus épaisses et la distribution est dite *leptokurtique*.
- Si $Kr < 3$, la distribution de X est relativement plus aplatie que celle de la loi normale ; ces queues sont moins épaisses et la distribution est dite *platikurtique*.

Asymétrie (perte/gain)

La distribution des cours des actifs financiers est généralement asymétrique ; il y a plus de mouvements à la baisse qu'à la hausse. Ceci renforce encore les raisons de mise en cause de l'hypothèse de normalité en analyse des séries chronologiques financières.

Rappelons qu'un test simple de l'hypothèse d'asymétrie consiste à tester la nullité du moment centré d'ordre 3 de la distribution, *i.e.*, la *skewness*.

Définition 2.2.3

La *skewness* (ou coefficient d'asymétrie) d'une variable aléatoire X , telle que son moment d'ordre 3 existe, est définie comme le troisième moment centré normalisé par le cube de son

écart-type :

$$Sw = \frac{\mathbb{E}(X - \mu_X)^3}{(\sigma_X)^3}.$$

Le coefficient d'asymétrie est un indicateur de l'asymétrie de la loi de la variable aléatoire. Une valeur absolue statistiquement différente de zéro indique l'asymétrie de cette loi.

- Si $Sw < 0$, dès lors que la distribution de X est non symétrique ; plus précisément, la probabilité d'obtenir des valeurs inférieures à sa moyenne est supérieure à celle d'obtenir des valeurs plus grandes que la moyenne, autrement dit, une queue de distribution étalée vers la gauche.
- Si $Sw > 0$, la distribution est asymétrique et la variable aléatoire a tendance à prendre des valeurs supérieures à sa valeur moyenne avec une probabilité plus grande que celle associée aux valeurs inférieures à la moyenne, , autrement dit, une queue de distribution étalée vers la droite.
- Il est à noter que $Sw = 0$, n'implique pas que la distribution est symétrique

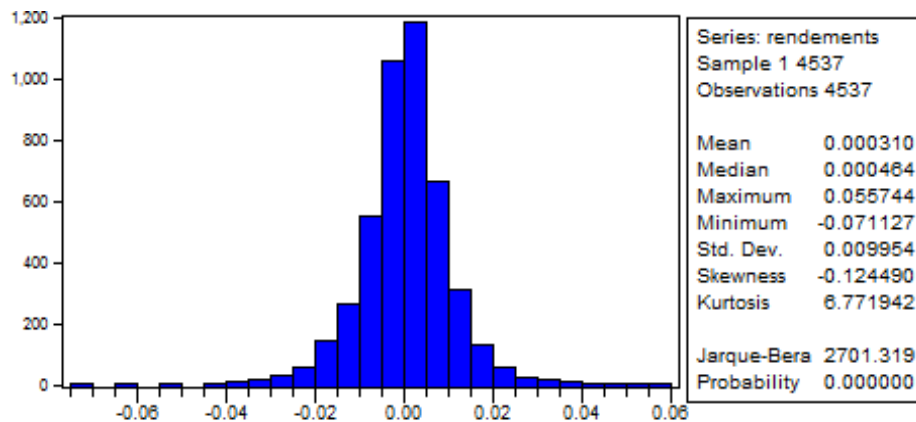


Figure 2.2.7 : Histogramme accompagné d'un petit tableau donnant les statistiques : moyenne, écart-type, skewness, kurtosis, la statistique du test de Jarque et bera et sa p-value (etc.) des rendements logarithmiques journaliers de l'indice S&P 500.

Le test d'asymétrie : $\left| \frac{\widehat{sw}}{\sqrt{6/n}} \right| = |-3.4233| > 1.96$, au seuil $\alpha = 0.05$. Donc, on accepte l'hypothèse d'asymétrie. Où \widehat{sw} est la valeur estimée du coefficient d'asymétrie Sw .

Le test d'aplatissement : $\left| \frac{\widehat{kr}-3}{\sqrt{24/n}} \right| = |51.8614| > 1.96$, au seuil $\alpha = 0.05$. Alors, l'aplatissement normal est rejeté. Où \widehat{kr} est la valeur estimée de la kurtosis Kr .

Le test de Jarque-Bera est un test de normalité fondé sur les propriétés de symétrie et d'aplatissement de la loi normale. La statistique du test est donnée par

$$JB = \frac{n}{6} \left(Sw^2 + \frac{(Kr - 3)^2}{4} \right). \quad (2.2.1)$$

Le test de normalité de Jarque-Bera confirme la non normalité de la distribution inconditionnelle des rendements logarithmiques journaliers puisque la p-value est presque nulle.

Une autre façon de tester, visuellement, l'adéquation ou non des rendements avec une distribution donnée, par exemple gaussienne, est de faire une analyse des quantiles, en traçant les quantiles empiriques contre les quantiles théoriques d'une loi donnée. Si les distributions sont identiques, le tracé doit coïncider approximativement avec la bissectrice $y = x$. Si au contraire elles sont différentes, on doit observer des déviations.

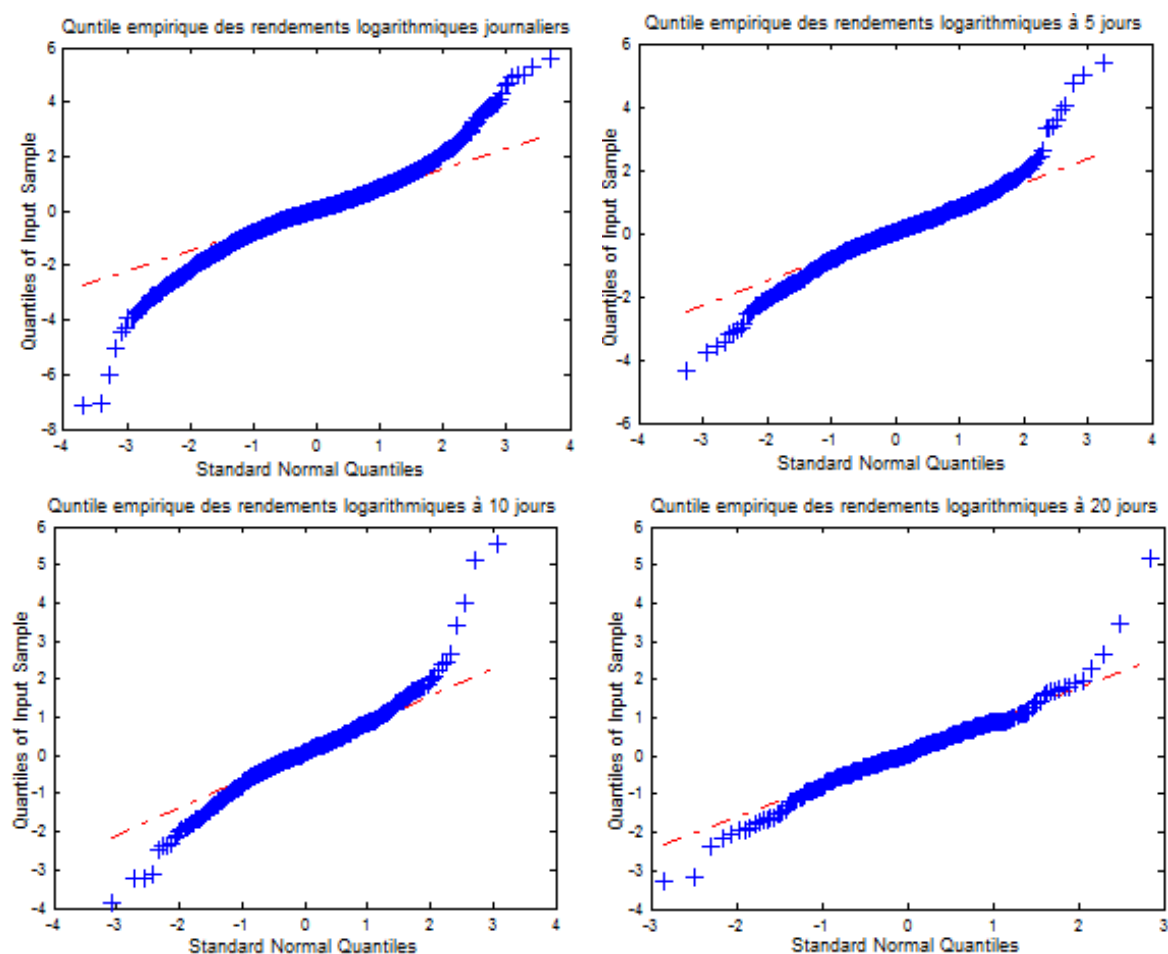


Figure 2.2.8 : Quantiles empiriques des rendements logarithmiques normalisés par leur écart-type de l'indice S&P 500, aux échelles journalières, à 5 jours, à 10 jours et à 20 jours.

Les tracés des figures 2.2.8 montrent clairement la convergence de la distribution vers la gaussienne, présentée par le trait pointillé, au fur et à mesure que l'échelle augmente. À noter que les déviations diminuent lorsque l'échelle croît mais elles le font lentement et restent considérables même au bout du 20 jours.

2.2.5 Présence de mémoire longue

Une autre caractéristique importante des séries financières est celle de la mémoire longue. Cette notion est apparue dans les années 1950, à partir des travaux de Hurst en hydrologie et, depuis, cette notion a été appliquée dans divers autres domaines tels que : l'économie, la météorologie, la finance, etc.

Si le processus des rendements logarithmiques est non autocorrélé, il ne l'est pas pour celui des rendements logarithmiques absolus ou carrés dont la fonction d'autocorrélation décroît lentement dans le temps.

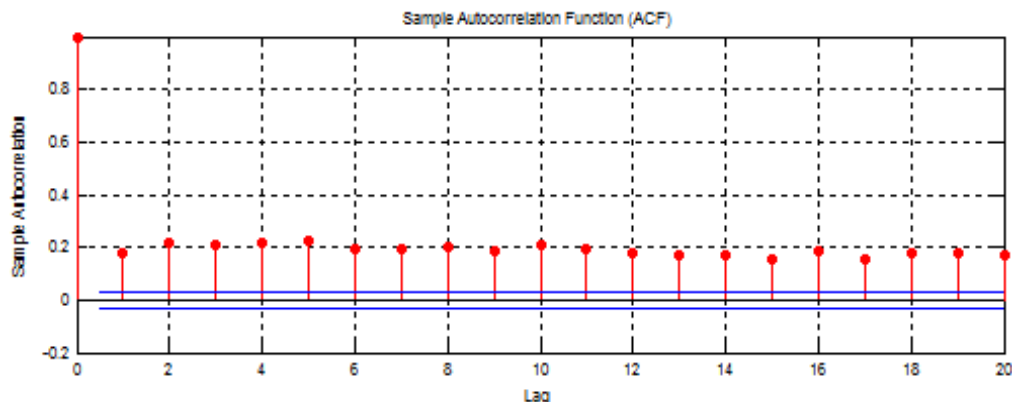


Figure 2.2.9 : Autocorrélations empiriques des rendements journaliers absolus de l'indice S&P 500.

Contrairement à la fonction d'autocorrélation des rendements logarithmiques (reflétée par le corrélogramme tracé sur la figure (2.2.6a)) celui des rendements logarithmiques absolus est caractérisé par une longue signification positive, indiquant ainsi l'existence d'une dépendance entre les différents rendements logarithmiques.

2.2.6 Effet de levier

Cette propriété notée par Black en 1976 repose sur l'observation qu'il existe une asymétrie entre l'impact des valeurs passées positives et des valeurs passées négatives sur la volatilité des cours ou des rendements. Ainsi, les valeurs négatives (baisse des cours) tendent à engendrer un accroissement de la volatilité supérieur à celui induit par des valeurs positives (hausse des cours). Cette propriété ne doit pas être confondue avec celle de l'asymétrie de la distribution des cours ou des rendements. Il s'agit ici d'une asymétrie de la relation liant des valeurs passées des cours ou des rendements à la volatilité de ces derniers.

2.2.7 Saisonnalité

La volatilité tend à augmenter lorsque les marchés ne fonctionnent pas (week-end, fêtes, jours fériés, etc.), reflétant ainsi l'information accumulée pendant cet arrêt. Notons que l'effet

saisonnier est également très présent pour les séries intra-journalières.

Nous avons décrit les propriétés qui nous semblent plus importantes, mais d'autres existent. Nous citons, à titre d'exemple, la périodicité dans la structure des covariances, la multimodalité et bien d'autres.

Toutes ces observations ont conduit à des modèles d'évolution des cours plus généraux que ceux comportant un mouvement brownien ou une semimartingale, par exemple, les modèles $GARCH$. Ce genre de modèle semble alors mieux tenir de quelques propriétés statistiques précitées. Le reste de ce chapitre est consacré à l'étude de cette classe de modèle.

2.3 Structure probabiliste du modèle $GARCH$

2.3.1 Présentation et définition du modèle $GARCH$

Les modèles linéaires classiques $ARMA$, popularisés par Box et Jenckis (1970), s'avèrent incapables de capturer les faits stylisés des séries financières évoqués ci-dessus. Afin de remédier à ces insuffisances, deux grandes classes de modèles ont été introduites, à savoir les modèles de type $GARCH$, et les modèles à *volatilité stochastiques*. Notre étude tout au long de ce mémoire sera portée seulement sur le premier type *i.e.*, les modèles $GARCH$.

Les modèles $ARCH$ ont été introduits par Engle en 1982, puis généralisés en modèles $GARCH$ en 1986 par Bollerslev. La modélisation $GARCH$ est devenue un outil incontournable en finance, particulièrement utile pour analyser et prévoir la volatilité. Ces modèles rendent mieux compte des faits stylisés des séries financières que ne le faisaient les modèles $ARMA$.

Définition 2.3.1

On dit que le processus $\{\varepsilon_n, n \in \mathbb{Z}\}$ suit un modèle $GARCH(p, q)$ fort s'il est défini par une équation de la forme

$$\varepsilon_n = Z_n \sqrt{\sigma_n^2}, \quad (Z_n) \text{ i.i.d}(0, 1), \quad (2.3.1a)$$

$$V(\varepsilon_n | \mathcal{F}_{n-1}) = \sigma_n^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{n-i}^2 + \sum_{j=1}^p \beta_j \sigma_{n-j}^2, \quad \forall n \in \mathbb{Z}, \quad (2.3.1b)$$

où les coefficients α_i, β_j et ω sont des constantes positives i.e., $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0$, $i = \overline{1, q}, j = \overline{1, p}$ et $(\alpha_q, \beta_p) \neq (0, 0)$, et \mathcal{F}_{n-1} est la tribu engendée par le passé du processus jusqu'à l'instant $n - 1$.

L'écriture abrégée de l'équation donnée par (2.3.1) est la suivante :

$$\sigma_n^2 = \omega + a(B)\varepsilon_n^2 + b(B)\sigma_n^2, \quad n \in \mathbb{Z},$$

avec

$$\begin{aligned} a(B) &= \alpha_1 B + \dots + \alpha_q B^q, \quad \alpha_i \in \mathbb{R}^+, \quad \forall i = 1, 2, \dots, q \text{ et } \alpha_q \neq 0, \\ b(B) &= \beta_1 B + \dots + \beta_p B^p, \quad \beta_j \in \mathbb{R}^+, \quad \forall j = 1, 2, \dots, p \text{ et } \beta_p \neq 0, \end{aligned}$$

où B est l'opérateur de retard tel que : $B^i \varepsilon_n = \varepsilon_{n-i}, B^j \sigma_n^2 = \sigma_{n-j}^2$.

Si $b(B) = 0$, l'équation (2.3.1) devient $\sigma_n^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{n-i}^2$, et le processus est appelé $ARCH(q)$.

Le carré d'un processus $\{\varepsilon_n, n \in \mathbb{Z}\}$, satisfaisant la formalisation d'un modèle $GARCH(p, q)$, donnée par (2.3.1), peut s'écrire sous forme d'une structure linéaire d'un modèle $ARMA(r, p)$. Par définition, l'innovation du processus (ε_n^2) est $v_n = \varepsilon_n^2 - \sigma_n^2$ (qui forme un bruit blanc). Remplaçons les variables σ_{n-j}^2 par $\varepsilon_{n-j}^2 - v_{n-j}$ dans l'équation (2.3.1), nous obtenons la représentation

$$\varepsilon_n^2 = \omega + \sum_{i=1}^r (\alpha_i + \beta_j) \varepsilon_{n-i}^2 + v_n - \sum_{j=1}^p \beta_j v_{n-j},$$

où $r = \max(p, q)$, avec la convention $\alpha_i = 0$ si $i > q$ et $\beta_j = 0$ si $j > p$.

À partir de cette représentation, nous pouvons facilement calculer les moments conditionnels et non conditionnels du processus $(\varepsilon_n, n \in \mathbb{Z})$ en outre, cette représentation sera utile lors de l'identification et de l'estimation des processus $GARCH$.

2.3.2 Étude de la stationnarité des processus $GARCH$

Comme nous l'avons déjà signalé au chapitre 1, la stationnarité stricte d'un processus admettant une écriture d'une équation aux différences stochastiques, découle de l'existence

d'une solution strictement stationnaire de cette équation. Ainsi, nous allons nous intéresser aux conditions d'existence de telles solutions (au sens strict et au second ordre) pour les processus vérifiant (2.3.1). Nous nous intéressons plus particulièrement aux solutions non anticipatives où les processus (ε_n) s'écrivent comme fonction mesurable des variables $Z_{n-s}, s \geq 0$. Pour ces processus, σ_n^2 est indépendant de la tribu engendrée par $\{Z_{n+h}, h \geq 0\}$, et (ε_n) est indépendant de la tribu engendrée par $\{Z_{n+h}, h > 0\}$. De telles solutions sont également ergodiques.

Un processus $\{\varepsilon_n, n \in \mathbb{Z}\}$, satisfaisant l'écriture d'un GARCH (p,q) donnée en (2.3.1) admet la représentation autoregressive généralisée suivante :

$$X_n = A_n X_{n-1} + B_n, n \in \mathbb{Z}, \tag{2.3.2}$$

où

$$\begin{aligned} X_n &= \left(\varepsilon_n^2 \quad \varepsilon_{n-1}^2 \quad \cdots \quad \varepsilon_{n-q+1}^2 \quad \sigma_n^2 \quad \cdots \quad \sigma_{n-p+1}^2 \right)' \in \mathbb{R}^{p+q} \\ B_n &= \left(\omega Z_n^2 \quad 0 \quad \cdots \quad \omega \quad 0 \quad \cdots \quad 0 \right)' \in \mathbb{R}^{p+q} \end{aligned}$$

où ω dans le vecteur B_n se trouve à la $(q+1)^{\text{ème}}$ ligne.

$$A_n = \left(\begin{array}{ccc|ccc} \alpha_1 Z_n^2 & \cdots & \alpha_q Z_n^2 & \beta_1 Z_n^2 & \cdots & \beta_p Z_n^2 \\ & I_{(q-1) \times (q-1)} & O_{(q-1) \times 1} & & O_{(q-1) \times p} & \\ \hline \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ & O_{(p-1) \times q} & & & I_{(p-1) \times (p-1)} & O_{(p-1) \times 1} \end{array} \right) \tag{2.3.3}$$

Existence d'une solution strictement stationnaire

L'équation (2.3.2) admet une solution unique, strictement stationnaire et ergodique de la forme :

$$X_n = \sum_{i=0}^{\infty} \prod_{j=0}^{i-1} A_{n-j} B_{n-i}, \tag{2.3.4}$$

où la série (2.3.4) converge presque sûrement, pourvu que le plus grand exposant de Lyapounov soit strictement négatif, avec la convention $\prod_{j=0}^{-1} A_{n-j} B_n = B_n$.

Théorème 2.3.1 (Bougerol et Picard, 1992b)

Une condition nécessaire et suffisante d'existence d'un processus $GARCH(p, q)$ strictement stationnaire, solution du modèle (2.3.1) avec $\omega > 0$, est que le plus grand exposant de Lyapounov de la suite $\{A_n, n \in \mathbb{Z}\}$, définie en (2.3.3) soit strictement négatif ($\gamma < 0$). Lorsqu'elle existe, la solution strictement stationnaire est unique, non anticipative et ergodique.

Remarque :

La condition nécessaire et suffisante d'existence d'une solution strictement stationnaire du modèle $GARCH(p, q)$ est établie par Bougerol et Picard (1992b). Néanmoins, nous considérons l'écriture selon une équation aux récurrences stochastiques donnée en (2.3.2), par Francq et Zakoïan (2004).

Démonstration.(du théorème 2.3.1)

Nous notons x^+ (resp. x^-), $\max(x, 0)$ (resp. $\max(-x, 0)$) et nous utilisons la norme matricielle définie par $\|A\| = \sum_{i,j} |a_{i,j}|$.

Utilisons de manière récursive l'équation (2.3.3), nous obtenons pour tout $N \geq 0$:

$$\begin{aligned} X_n &= \underbrace{B_n + \sum_{t=1}^N A_n \dots A_{n-t+1} B_{n-t}}_{\widetilde{X}_n(N)} + A_n \dots A_{n-N} X_{n-N-1}, \\ &= \widetilde{X}_n(N) + A_n \dots A_{n-N} X_{n-N-1}. \end{aligned} \quad (2.3.5)$$

Nous allons à présent montrer que la suite $\widetilde{X}_n(N)$ converge presque sûrement, vers une limite \widetilde{X}_n , quand N tend vers l'infini.

Supposons que $\gamma < 0$.

En utilisant la multiplicativité de la norme, nous avons

$$\begin{aligned} \left\| \widetilde{X}_n(N) \right\| &\leq \|B_n\| + \sum_{t=1}^N \|A_n \dots A_{n-t+1}\| \|B_{n-t}\|, \\ &\leq \|B_n\| + \sum_{t=1}^{\infty} \|A_n \dots A_{n-t+1}\| \|B_{n-t}\|. \end{aligned}$$

Par suite, en appliquant le critère de la racine $n^{\text{ème}}$ de Cauchy, nous obtenons

$$\begin{aligned} \lim_{t \rightarrow \infty} \|A_n \dots A_{n-t+1}\|^{1/t} \|B_{n-t}\|^{1/t} &= \lim_{t \rightarrow \infty} \exp \left\{ \frac{1}{t} \log \|A_n \dots A_{n-t+1}\| + \frac{1}{t} \log \|B_{n-t}\| \right\}, \\ &= \exp \{ \gamma \} \text{ p.s, si } \frac{1}{t} \log \|B_{n-t}\| \xrightarrow[t \rightarrow \infty]{p.s} 0 \end{aligned}$$

Pour montrer que $\frac{1}{t} \log \|B_{n-t}\| \xrightarrow[t \rightarrow \infty]{p.s} 0$, nous procédons comme suit :

$$\mathbb{E} |\log \|B_{n-t}\|| = \mathbb{E} (\log^+ \|B_{n-t}\| + \log^- \|B_{n-t}\|),$$

or,

$$\begin{aligned} \log^- \|B_{n-t}\| &= \max(-\log \|B_{n-t}\|, 0), \\ &= \max(-\log \omega - \log(Z_n^2 + 1), 0), \\ &\leq \max(-\log \omega, 0) = \log^- \omega, \end{aligned}$$

d'où

$$\mathbb{E} |\log \|B_{n-t}\|| \leq \mathbb{E} (\log^+ \|B_{n-t}\|) + \log^- \omega < \infty. \quad (2.3.6)$$

Il reste à montrer que si (2.3.6) est vérifiée, $\frac{\log \|B_{n-t}\|}{t} \xrightarrow[t \rightarrow \infty]{p.s} 0$, ce qui revient à montrer que

$$P \left(\lim_{t \rightarrow \infty} \left| \frac{\log \|B_{n-t}\|}{t} \right| > e \right) = 0, \forall e > 0.$$

Nous avons

$$\begin{aligned} \sum_{t=1}^{\infty} P \left(\frac{|\log \|B_{n-t}\||}{t} > e \right) &= \sum_{t=1}^{\infty} P \left(\frac{|\log \|B_n\||}{t} > e \right), \text{ car la suite } (B_n)_n \text{ est i.i.d.} \\ &\leq \int_0^{\infty} P \left(\frac{|\log \|B_n\||}{x} > e \right) dx, \\ &= \int_0^{\infty} P \left(\frac{|\log \|B_n\||}{e} > x \right) dx = e^{-1} \mathbb{E} (|\log \|B_n\||) < \infty. \end{aligned}$$

D'après le théorème de Borel-Cantelli, $P \left(\limsup_{t \rightarrow \infty} \left(\frac{|\log \|B_{n-t}\||}{t} > e \right) \right) = 0$, et comme

$0 \leq P \left(\lim_{t \rightarrow \infty} \left(\frac{|\log \|B_{n-t}\||}{t} > e \right) \right) \leq P \left(\limsup_{t \rightarrow \infty} \left(\frac{|\log \|B_{n-t}\||}{t} > e \right) \right)$, alors $\log \|B_{n-t}\|$ converge presque sûrement vers zéro, quand t tend à l'infini.

Nous avons supposé que $\gamma < 0$, alors, $\exp \{ \gamma \} < 1$. En appliquant le critère de la racine $n^{\text{ème}}$ de Cauchy, nous concluons que $\widetilde{X}_n(N)$ converge presque sûrement.

Soit $\widetilde{X}_{q+1,n}$, la $q+1^{\text{ème}}$ composante de \widetilde{X}_n . Le processus $\varepsilon_n = \sqrt{\widetilde{X}_{q+1,n}} Z_n$, est une solution

du modèle (2.3.1), elle est non anticipative car d'après (2.3.4), (ε_n) s'exprime comme fonction mesurable des variables Z_{n-i} , $i \geq 0$. Cette solution est également strictement stationnaire et ergodique puisque (Z_n) est ergodique.

Montrons maintenant l'unicité de cette solution. Supposons que la solution n'est pas unique et qu'il existe une autre solution strictement positive (X_n) de (2.3.2). D'après (2.3.5) on a :

$$\|X_n - \tilde{X}_n\| \leq \|\tilde{X}_n(N) - \tilde{X}_n\| + \|A_n \dots A_{n-N}\| \|X_{n-N-1}\|.$$

Il est clair que $\tilde{X}_n(N) - \tilde{X}_n \xrightarrow[N \rightarrow \infty]{p.s} 0$.

Nous avons vu que la série définissant \tilde{X}_n converge presque sûrement, donc nous avons :

$$\|A_n \dots A_{n-N}\| \xrightarrow[N \rightarrow \infty]{p.s} 0.$$

De plus la loi de X_{n-N-1} ne dépend pas de N par stationnarité. Par suite :

$$\|A_n \dots A_{n-N}\| \|X_{n-N-1}\| \xrightarrow[N \rightarrow \infty]{p} 0.$$

Donc, nous avons montré que $X_n - \tilde{X}_n \xrightarrow[N \rightarrow \infty]{p} 0$.

Ce terme étant indépendant de N , on a nécessairement $X_n = \tilde{X}_n, \forall n$. D'où $\varepsilon_n = \sqrt{\tilde{X}_{q+1,n}} Z_n$, est l'unique solution strictement stationnaire du modèle.

Enfin, nous montrons la partie nécessaire du théorème, *i.e.*, si nous avons une solution strictement stationnaire du modèle (2.3.1), alors le plus grand exposant de Lyapounov (γ) est strictement négatif. D'après le lemme (1.2.1) (Bougerol et Picard, 1992a), il suffit de montrer

$$\text{que } \lim_{n \rightarrow \infty} \|A_0 \dots A_{-n}\| \xrightarrow{p.s} 0.$$

Nous utilisons la notation $x \geq y$, qui signifie que toutes les composantes du vecteur x sont supérieures ou égales à celles du vecteur y qui est de même dimension.

Nous allons montrer que, pour $1 \leq i \leq q + p$

$$\lim_{n \rightarrow \infty} A_0 \dots A_{-n} e_i = 0, \quad p.s \quad (2.3.7)$$

où e_i est le $i^{\text{ème}}$ élément de la base canonique de \mathbb{R}^{p+q} .

Soit (ϵ_n) une solution strictement stationnaire de (2.3.1), et d'après (2.3.2) nous avons pour tout n strictement positif

$$\begin{aligned} X_0 &= B_0 + A_0 X_{-1}, \\ &= B_0 + \sum_{k=0}^{n-1} A_0 \dots A_{-k} B_{-k-1} + A_0 \dots A_{-n} X_{-n-1}, \\ &\geq \sum_{k=0}^{n-1} A_0 \dots A_{-k} B_{-k-1}, \end{aligned}$$

car les coefficients des matrices A_n , B_0 et X_n sont positifs.

Par suite, la série $\sum_{k=0}^{n-1} A_0 \dots A_{-k} B_{-k-1}$ converge et donc $A_0 \dots A_{-k} B_{-k-1} \xrightarrow[k \rightarrow \infty]{p.s} 0$.

Or, $B_{-k-1} = \omega Z_{-k-1}^2 e_1 + \omega e_{q+1}$. Donc $A_0 \dots A_{-k} B_{-k-1}$ se décompose en deux termes positifs et nous obtenons

$$\lim_{k \rightarrow \infty} A_0 \dots A_{-k} \omega Z_{-k-1}^2 e_1 = 0, \quad \lim_{k \rightarrow \infty} A_0 \dots A_{-k} \omega e_{q+1} = 0, \quad p.s \quad (2.3.8a)$$

Le fait que ω est différent de zéro, (2.3.8a) devient

$$\lim_{k \rightarrow \infty} A_0 \dots A_{-k} Z_{-k-1}^2 e_1 = 0, \quad \lim_{k \rightarrow \infty} A_0 \dots A_{-k} e_{q+1} = 0, \quad p.s \quad (2.3.8b)$$

1) $i = 1, \dots, q$.

Si $i = q$, on a : $A_{-k} e_q = \alpha_q Z_{-k}^2 e_1 + \alpha_q e_{q+1}$, d'après (2.3.8b), on conclut que (2.3.7) est vérifiée, pour $i = q$, pour montrer que (2.3.7) est vraie pour le reste des valeurs de i , on utilise la relation suivante

$$A_{-k} e_i = \alpha_i Z_{-k}^2 e_1 + \alpha_i e_{q+1} + e_{i+1}, \quad i = 1, \dots, q-1. \quad (2.3.9)$$

Pour $i = q-1$, la relation (2.3.7) donne

$$\begin{aligned} \lim_{k \rightarrow \infty} A_0 \dots A_{-k} e_{q-1} &= \alpha_{q-1} \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} Z_{-k+1}^2 e_1 + \\ &\alpha_{q-1} \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} Z_{-k+1}^2 e_{q+1} + \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} e_q = \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} e_q = 0. \end{aligned}$$

Donc, pour $i = q-1$ (2.3.7) est vérifiée, et par récurrence pour le reste des valeurs de i en utilisant (2.3.9)

2) $i = q + 1, \dots, q + p$.

D'après (2.3.8b), (2.3.7) est vérifiée pour $i = q + 1$. On utilise la relation

$$A_{-k}e_{q+j} = \beta_j Z_{-k}^2 e_1 + \beta_j e_{q+1} + e_{q+1+j}, \quad j = 1, \dots, p, \quad (2.3.10)$$

avec par convention $e_{q+1+j} = 0$, on obtient, pour $j = 1$,

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} A_0 \dots A_{-k} e_{q+1} = \beta_1 \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} Z_{-k+1}^2 e_1 + \\ &\beta_1 \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} Z_{-k+1}^2 e_{q+1} + \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} e_{q+2} \geq \lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} e_{q+2} \geq 0. \end{aligned}$$

D'où $\lim_{k \rightarrow \infty} A_0 \dots A_{-k+1} e_{q+2} = 0$. Anisi, (2.3.7) est vérifiée pour $i = q + 2$. Nous procédons de la même manière *i.e.*, par récurrence pour le reste des valeurs de i , pour montrer que (2.3.7) est vraie.

Remarques :

- 1) Il est remarquable que la solution strictement stationnaire d'un modèle $GARCH$, si elle existe, est non anticipative, contrairement aux modèles $ARMA$ pour lesquels la condition de stationnarité implique l'existence de solutions stationnaires pouvant dépendre du futur comme du passé du bruit.
- 2) La condition $\gamma < 0$, est également nécessaire pour l'obtention d'une solution strictement stationnaire, en plus, elle garantit l'existence de certains moments du processus solution.
- 3) Empiriquement, la condition que le plus grand exposant de Lyapounov soit strictement négatif, peut être vérifiée en calculant des approximations de ce coefficient par des méthodes de simulations. De plus ce coefficient dépend de la loi du processus Z_n , ce qui présente un inconvénient du point de vue pratique.
- 4) Si un modèle $GARCH$ est tel qu'il admet une solution strictement stationnaire, tout modèle obtenu en remplaçant les α_i, β_j par des coefficients plus petits l'est également. Effectivement, le coefficient γ du modèle ainsi défini sera nécessairement inférieur à celui du modèle initial, car avec la norme utilisée, $0 \leq A \leq B$ implique $\|A\| \leq \|B\|$. En particulier, la stationnarité du modèle $GARCH$ implique celle du modèle $ARCH$ obtenu en supprimant les coefficients β_j .

Cas particulier : $GARCH(1, 1)$

Afin d'illustrer la condition de stationnarité stricte des processus $GARCH$, nous nous plaçons dans le cas d'un processus $GARCH(1, 1)$, qui admet la représentation

$$\begin{aligned} \varepsilon_n &= Z_n \sqrt{\sigma_n^2}, & (Z_n) \text{ i.i.d}(0, 1), \\ \sigma_n^2 &= \omega + \alpha \varepsilon_{n-1}^2 + \beta \sigma_{n-1}^2, & \omega > 0, (\alpha, \beta) > (0, 0). \end{aligned} \quad (2.3.11)$$

Le modèle (2.3.11) est une solution de l'équation aux récurrence stochastique $X_n^* = A_n^* X_{n-1}^* + B_n^*$, où

$$X_n^* = \begin{pmatrix} \varepsilon_n^2 \\ \sigma_n^2 \end{pmatrix}, \quad A_n^* = \begin{pmatrix} \alpha Z_n^2 & \beta Z_n^2 \\ \alpha & \beta \end{pmatrix} \text{ et } B_n^* = \begin{pmatrix} \omega Z_n^2 \\ \omega \end{pmatrix}.$$

On a donc

$$\prod_{k=0}^{n-1} A_{n-k}^* = \prod_{k=1}^{n-1} (\alpha Z_{n-k}^2 + \beta) A_n^*, \quad \forall n \geq 1.$$

Par suite

$$\log \left\| \prod_{k=0}^{n-1} A_{n-k}^* \right\| = \sum_{k=1}^{n-1} \log (\alpha Z_{n-k}^2 + \beta) + \log \|A_n^*\|,$$

et d'après le théorème ergodique, nous obtenons

$$\gamma^* = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left\| \prod_{k=0}^{n-1} A_{n-k}^* \right\| = \mathbb{E} \{ \log (\alpha Z_n^2 + \beta) \}.$$

Théorème 2.3.2 (Stationnarité stricte du modèle $GARCH(1, 1)$)

Si

$$-\infty \leq \gamma^* := \mathbb{E} \{ \log (\alpha Z_n^2 + \beta) \} < 0,$$

avec $a(Z_n) = \alpha Z_n^2 + \beta$, la série

$$h_n^* = \omega \left\{ 1 + \sum_{k=1}^{\infty} a(Z_{n-1}) \dots a(Z_{n-k}) \right\},$$

converge presque sûrement et le processus défini (ε_n) par $\varepsilon_n = \sqrt{h_n^*} Z_n$ est l'unique solution strictement stationnaire, non anticipative et ergodique du modèle (2.3.11).

Si $\gamma^* \geq 0$ et $\omega > 0$, il n'existe pas de solution strictement stationnaire.

Remarques :

- 1) Lorsque $\omega = 0$ et $\gamma^* < 0$, la seule solution strictement stationnaire du modèle (2.3.11) est nulle ($\varepsilon_n = 0$). Alors, de point de vue application statistique, il est naturel d'imposer à ω d'être strictement positif.
- 2) La condition de stationnarité stricte $\gamma^* = \mathbb{E} \{ \log(\alpha Z_n^2 + \beta) \} < 0$, implique $\beta < 1$. Inversement, si $\alpha + \beta < 1$, la condition $\gamma^* = \mathbb{E} \{ \log(\alpha Z_n^2 + \beta) \} < 0$, est vérifiée, par l'application de l'inégalité de Jensen, $\mathbb{E} \{ \log(\alpha Z_n^2 + \beta) \} \leq \log \mathbb{E}(\alpha Z_n^2 + \beta)$ et comme $\log \mathbb{E}(\alpha Z_n^2 + \beta) = \log(\alpha + \beta) < 0$, il s'ensuit que $\mathbb{E} \{ \log(\alpha Z_n^2 + \beta) \} < 0$.

Le corollaire suivant donne une condition nécessaire de stationnarité stricte des processus $GARCH$, sous trois formes différentes.

Corollaire 2.3.1 (conséquence de la stationnarité stricte)

Si le plus grand exposant de Lyapounov $\gamma = \inf \{ \mathbb{E}(\frac{1}{n} \log \|A_0 A_{-1} \dots A_{-n+1}\|), n \in \mathbb{N} \}$, est strictement négatif, les conditions suivantes sont équivalentes :

- (i) $\sum_{j=1}^p \beta_j < 1$,
- (ii) $1 - \beta_1 z - \dots - \beta_p z^p = 0 \Rightarrow |z| > 1$,
- (iii) $\rho(K) < 1$, où K est la sous-matrice de A_n définie par :

$$K = \left(\begin{array}{ccc|c} \beta_1 & \dots & & \beta_p \\ \hline & I_{(p-1) \times (p-1)} & & O_{(p-1) \times (1)} \end{array} \right).$$

De plus, si $\gamma < 0$, alors il existe s strictement positif tel que

$$\mathbb{E}(\sigma_n^{2s}) < \infty \text{ et } \mathbb{E}(|\varepsilon_n|^{2s}) < \infty.$$

Existence d'une solution stationnaire au second ordre

Le théorème suivant donne les conditions nécessaires et suffisantes de stationnarité au second ordre du modèle $GARCH$.

Théorème 2.3.3 (Bollerslev, 1986)

S'il existe un processus $GARCH(p, q)$, tel que son premier moment conditionnel est une différence de martingale et son deuxième moment conditionnel satisfait l'écriture (2.3.1b), de plus ces deux premiers moments conditionnels existent, si ce processus est stationnaire au second ordre et non anticipatif et si $\omega > 0$, alors

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (2.3.12)$$

Inversement, si (2.3.12) est vérifiée, l'unique solution strictement stationnaire est un bruit blanc faible (donc est stationnaire au second ordre). Il n'existe pas d'autre solution non anticipative et stationnaire au second ordre.

Démonstration.(du théorème 2.3.3)

Commençons par démontrer la partie nécessaire de ce théorème *i.e.*, s'il existe une solution stationnaire au second ordre, alors la condition (2.3.12) est vérifiée.

Soit (ε_n) , un processus $GARCH(p, q)$, stationnaire au second ordre et non anticipatif. Alors, $\mathbb{E}(\varepsilon_n^2) = \mathbb{E}(\mathbb{E}(\varepsilon_n^2/\mathcal{F}_{n-1})) = \mathbb{E}(\sigma_n^2) > 0$ et ne dépend pas du temps n . En prenant l'espérance des deux membres de (2.3.1b), on obtient

$$\mathbb{E}(\varepsilon_n^2) = \omega + \sum_{i=1}^q \alpha_i \mathbb{E}(\varepsilon_{n-i}^2) + \sum_{j=1}^p \beta_j \mathbb{E}(\sigma_{n-j}^2) = \omega + \sum_{i=1}^q \alpha_i \mathbb{E}(\varepsilon_n^2) + \sum_{j=1}^p \beta_j \mathbb{E}(\varepsilon_n^2).$$

Ainsi,

$$\left(1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j\right) \mathbb{E}(\varepsilon_n^2) = \omega. \quad (2.3.13)$$

Puisque ω est strictement positif, on doit avoir (2.3.12), pour que (2.3.13) soit vérifiée.

Passons maintenant à la démonstration de la partie suffisante *i.e.*, si la condition (2.3.12) est vérifiée, alors l'unique solution strictement stationnaire forme un bruit blanc, *i.e.*, une solution stationnaire au second ordre et non anticipatif; de plus cette dernière solution est unique.

Supposons que (2.3.12) soit vraie et cherchons une solution $GARCH$ qui soit stationnaire. Pour $n, k \in \mathbb{Z}$, définissons les vecteurs à valeurs dans \mathbb{R}^d suivants :

$$X_k(n) = \begin{cases} 0 & \text{si } k < 0 \\ B_n + A_n X_{k-1}(n-1) & \text{si } k \geq 0. \end{cases}$$

Nous avons

$$X_k(n) - X_{k-1}(n) = \begin{cases} 0 & \text{si } k < 0 \\ B_n & \text{si } k = 0 \\ A_n \{X_{k-1}(n-1) - X_{k-2}(n-1)\} & \text{si } k > 0. \end{cases}$$

Utilisons ces relations de façon récursive nous obtenons pour, $k > 0$

$$\begin{aligned} X_k(n) - X_{k-1}(n) &= A_n \dots A_{n-k+1} \{X_1(n-k) - X_0(n-k)\}, \\ &= A_n \dots A_{n-k+1} B_{n-k}, \text{ où } B_{n-k} = X_1(n-k) - X_0(n-k) \end{aligned}$$

Puisque la matrice $A_n \dots A_{n-k+1} B_{n-k}$ est positive et ses termes sont indépendants (car chaque terme s'exprime comme fonction d'une variable Z_{n-i} , où le processus (Z_n) est *i.i.d.*).

Alors, nous avons, pour $k > 0$

$$\begin{aligned} \mathbb{E} \|X_k(n) - X_{k-1}(n)\| &= \|\mathbb{E}(A_n \dots A_{n-k+1} B_{n-k})\|, \\ &= \|\mathbb{E}(A_n) \dots \mathbb{E}(A_{n-k+1}) \mathbb{E}(B_{n-k})\|, \\ \text{avec } A &= \mathbb{E}(A_n) \text{ et } B = \mathbb{E}(B_n) \\ \mathbb{E} \|X_k(n) - X_{k-1}(n)\| &= \|A \dots AB\| = \|A^k B\|, \\ &= (1, \dots, 1) A^k B, \end{aligned}$$

car tous les termes du vecteur $A^k B$ sont positifs.

Calculons le déterminant de $(\lambda I_{p+q} - A)$ (nous notons \det le déterminant d'une matrice).

Pour ce faire, retranchons la $q + 1^{\text{ème}}$ ligne de $(\lambda I_{p+q} - A)$ à la première, puis développons le déterminant par rapport à la première ligne, nous obtenons : $\det(\lambda I_{p+q} - A) =$

$$\begin{aligned}
 & \det \left(\begin{array}{cccc|cccc}
 \lambda - \alpha_1 & -\alpha_2 & \cdots & -\alpha_q & -\beta_1 & \cdots & & -\beta_p \\
 -1 & \lambda & & 0 & & & & \\
 & \ddots & \ddots & & & & & \\
 & & -1 & \lambda & & & & \\
 \hline
 -\alpha_1 & \cdots & & -\alpha_q & \lambda - \beta_1 & -\beta_2 & & -\beta_p \\
 & & & & -1 & \lambda & 0 \cdots & 0 \\
 & & & & 0 & \ddots & \ddots & \vdots \\
 & & & & \vdots & & & \\
 & & & & 0 & \cdots & -1 & \lambda
 \end{array} \right)_{p+q} \\
 & = \det \left(\begin{array}{cccc|cccc}
 \lambda & 0 & \cdots & 0 & -\lambda & 0 & \cdots & 0 \\
 -1 & \lambda & & 0 & & & & \\
 0 & \ddots & \ddots & & & & & \\
 \vdots & & & -1 & \lambda & & & \\
 \hline
 -\alpha_1 & \cdots & & -\alpha_q & \lambda - \beta_1 & -\beta_2 & & -\beta_p \\
 & & & & -1 & \lambda & 0 \cdots & 0 \\
 & & & & 0 & \ddots & \ddots & \vdots \\
 & & & & \vdots & & & \\
 & & & & 0 & \cdots & -1 & \lambda
 \end{array} \right)_{p+q} \\
 & = \lambda \det \left(\begin{array}{cccc|cccc}
 \lambda & 0 & \cdots & 0 & & & & \\
 -1 & \lambda & & \vdots & & & & \\
 0 & \ddots & \ddots & & & & & \\
 \vdots & & & & & & & \\
 0 & \cdots & -1 & \lambda & & & & \\
 \hline
 -\alpha_2 & \cdots & & -\alpha_q & \lambda - \beta_1 & -\beta_2 & \cdots & -\beta_p \\
 & & & & -1 & \lambda & 0 \cdots & 0 \\
 & & & & 0 & \ddots & \ddots & \vdots \\
 & & & & \vdots & & & \\
 & & & & 0 & \cdots & -1 & \lambda
 \end{array} \right)_{p+q-1} +
 \end{aligned}$$

$$\begin{aligned}
 & (-\lambda)(-1)^{q+2} \det \left(\begin{array}{cccc|cccc}
 -1 & \lambda & 0 \dots & 0 & & & & \\
 0 & \ddots & \ddots & \vdots & & & & \\
 \vdots & & & -1 & \lambda & & & \\
 -\alpha_1 & \dots & & & -\alpha_q & -\beta_2 & \dots & -\beta_p \\
 \hline
 & & & & & \lambda & 0 & \dots & 0 \\
 & & & & & -1 & & & \vdots \\
 & & & & & 0 & \ddots & \ddots & \\
 & & & & & \vdots & & & \\
 & & & & & 0 & \dots & -1 & \lambda
 \end{array} \right)_{p+q-1} \\
 & = \det \left(\begin{array}{cccc}
 \lambda - \beta_1 & -\beta_2 & 0 \dots & -\beta_p \\
 -1 & \lambda & 0 & 0 \\
 & \ddots & \ddots & \vdots \\
 0 & & -1 & \lambda
 \end{array} \right)_p (\lambda)^q + \\
 & \det \left(\begin{array}{cccc}
 -1 & \lambda & 0 & \dots & 0 \\
 & \ddots & \ddots & & \vdots \\
 0 & & -1 & & \lambda \\
 -\alpha_1 & \dots & & & -\alpha_q
 \end{array} \right)_q \times \det \left(\begin{array}{cccc}
 \lambda & 0 & \dots & 0 \\
 -1 & \lambda & & \vdots \\
 & \ddots & \ddots & \\
 0 & & -1 & \lambda
 \end{array} \right)_{p-1} - \lambda (-1)^q \\
 & = \det \left(\begin{array}{cccc}
 \lambda - \beta_1 & -\beta_2 & \dots & -\beta_p \\
 -1 & \lambda & & 0 \\
 & \ddots & \ddots & \vdots \\
 0 & & -1 & \lambda
 \end{array} \right)_p \lambda^q + \det \left(\begin{array}{cccc}
 -\alpha_1 & -\alpha_2 & \dots & -\alpha_q \\
 -1 & \lambda & & 0 \\
 & \ddots & \ddots & \vdots \\
 0 & & -1 & \lambda
 \end{array} \right)_q \lambda^p (-1)^{q+1} (-1)^{q-1}
 \end{aligned}$$

Pour calculer le déterminant de ces deux dernières matrices, on développe par rapport à la dernière colonne et par récurrence, nous obtenons

$$\det(\lambda I_{p+q} - A) = \lambda^q (\lambda^p - \lambda^{p-1} \beta_1 - \dots - \beta_p) + \lambda^p (-\lambda^{q-1} \alpha_1 - \dots - \alpha_q).$$

Alors,

$$\det(\lambda I_{p+q} - A) = \lambda^{p+q} \left(1 - \sum_{i=1}^q \alpha_i \lambda^{-i} - \sum_{j=1}^p \beta_j \lambda^{-j} \right)$$

Si $|\lambda| \geq 1$ et en utilisant l'inégalité triangulaire $|a - b| \geq |a| - |b|$, on tire

$$|\det(\lambda I_{p+q} - A)| \geq \left| 1 - \sum_{i=1}^q \alpha_i \lambda^{-i} - \sum_{j=1}^p \beta_j \lambda^{-j} \right| \geq 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j > 0.$$

Remarques :

- 1) La condition de stationnarité au second ordre $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$ implique celle de stationnarité stricte $\gamma < 0$ *i.e.*, la solution stationnaire au second ordre l'est également strictement.
- 2) Sous les conditions du théorème 2.3.3, l'unique solution stationnaire de modèle (2.3.1) forme un bruit blanc de variance

$$\text{var}(\varepsilon_n) = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}.$$

2.3.3 Existence des moments d'ordres supérieurs et structure d'autocovariance du modèle GARCH

Existence des moments d'ordre supérieurs pour le modèle GARCH

Examinons maintenant le problème de l'existence des moments d'ordres supérieurs du modèle $GARCH(p, q)$, nous ne considérons que les moments d'ordres pair (sous l'hypothèse que la loi de Z_n est symétrique), sinon ces moments semblent difficiles à calculer.

Théorème 2.3.4 (Ling et McAlee, 2002)

Soit $A^{(m)} = \mathbb{E}(A_n^{\otimes m})$ où A_n est définie par (2.3.3). Supposons que $\mathbb{E}(Z_n^{2m}) < \infty$. Une condition nécessaire et suffisante d'existence des moments jusqu'à l'ordre m du processus (ε_n^2) , défini comme la première composante de (X_n) est que

$$\rho(A^{(m)}) < 1.$$

Alors, la série (X_n) définie en (2.3.4) est convergente en moyenne d'ordre m , pour tout $n \in \mathbb{Z}$, et le processus (ε_n^2) , est strictement stationnaire.

Démonstration.(Théorème 2.3.4)

Commençons par démontrer la partie suffisante du théorème, *i.e.*, si le $\rho(A^{(m)}) < 1$, alors, il existe une solution strictement stationnaire à l'ordre $2m$.

Posons, pour $k > 0$,

$$X_{n,k} = \underbrace{A_n A_{n-1} \dots A_{n-k+1}}_{A_{n,k}} B_{n-k}, \text{ avec convention } A_{n,0} = I_{p+q} \text{ et } X_{n,0} = B_n.$$

En utilisant les propriétés du produit de Kronecker et celles de la norme, nous obtenons

$$\begin{aligned} \mathbb{E} \|X_{n,k}\|^m &= \mathbb{E} \|A_{n,k} B_{n-k}\|^m, \\ &= \mathbb{E} (\|A_{n,k} B_{n-k}\| \dots \|A_{n,k} B_{n-k}\|), \\ &= \mathbb{E} \|A_{n,k} B_{n-k} \otimes \dots \otimes A_{n,k} B_{n-k}\|, \\ &= \mathbb{E} \|(A_{n,k} B_{n-k})^{\otimes m}\| = \mathbb{E} \|(A_{n,k})^{\otimes m} (B_{n-k})^{\otimes m}\|, \\ &= \|\mathbb{E} (A_n^{\otimes m} \dots A_{n-k+1}^{\otimes m} B_{n-k}^{\otimes m})\|, \end{aligned}$$

Étant données l'indépendance et l'équidistributivité qui caractérisent les termes du produit

$$A_n A_{n-1} \dots A_{n-k+1} B_{n-k}, \text{ nous aurons}$$

$$\mathbb{E} \|X_{n,k}\|^m = \|(A^{\otimes m})^k B^{\otimes m}\|.$$

Notons $A^{(m)}$ (resp. $B^{(m)}$), $\mathbb{E}(A^{\otimes m})$ (resp. $\mathbb{E}(B^{\otimes m})$)

$$\text{Alors, } \mathbb{E} \|X_{n,k}\|^m = \|(A^{(m)})^k B^{(m)}\|.$$

L'objectif est de montrer que X_n est finie en moyenne d'ordre m .

Nous avons

$$\begin{aligned} \|X_n\|_m &= \{\mathbb{E} \|X_n\|^m\}^{\frac{1}{m}} = \left\{ \mathbb{E} \left\| \sum_{k=0}^{\infty} X_{n,k} \right\|^m \right\}^{\frac{1}{m}}, \\ &\leq \sum_{k=0}^{\infty} \{\mathbb{E} \|X_{n,k}\|^m\}^{\frac{1}{m}} = \sum_{k=0}^{\infty} \|X_{n,k}\|_m. \end{aligned}$$

Or, d'après la sous-multiplicativité de la norme

$$\sum_{k=0}^{\infty} \|X_{n,k}\|_m = \sum_{k=0}^{\infty} \|(A^{(m)})^k B^{(m)}\|_m^{\frac{1}{m}} \leq \left\{ \sum_{k=0}^{\infty} \|(A^{(m)})^k\|_m^{\frac{1}{m}} \right\} \|B^{(m)}\|_m^{\frac{1}{m}}.$$

Si le rayon spectral de $A^{(m)}$ est strictement inférieur à l'unité, il est clair que $\|(A^{(m)})^k\|$ converge à vitesse exponentielle, lorsque k tend vers l'infini. Ainsi, la norme L^m de X_n est finie et par conséquent (X_n) est presque sûrement fini. C'est la solution strictement stationnaire de (2.3.2) et cette solution converge en moyenne d'ordre m , donc elle appartient

à L^m .

Comme $\|\varepsilon_n^2\|_m \leq \|X_n\|_m$, car la norme de X_n est supérieure à celle de chacune de ses composantes, alors une condition suffisante d'existence de $\mathbb{E}(\varepsilon_n^{2m})$ est que $\rho(A^{(m)}) < 1$.

Supposons que (ε_n^2) appartient à L^m et montrons qu'une condition nécessaire pour que cette dernière soit vraie est que $\rho(A^{(m)}) < 1$ doit être vérifiée

Nous avons, pour tout $t \geq 0$

$$\begin{aligned} \mathbb{E}(X_n^{\otimes m}) &= \mathbb{E}\left\{\sum_{k=0}^t X_{n,k} + A_n A_{n-1} \dots A_{n-t} X_{n-t-1}\right\}^{\otimes m}, \\ &\geq \mathbb{E}\left\{\sum_{k=0}^t X_{n,k}\right\}^{\otimes m}, \\ &\geq \sum_{k=0}^t \mathbb{E}(X_{n,k})^{\otimes m} = \sum_{k=0}^t \left((A^{(m)})^k B^{(m)}\right). \end{aligned}$$

Puisque toutes les composantes de $\mathbb{E}(X_n^{\otimes m})$ sont finies, nous obtenons

$$\lim_{t \rightarrow \infty} (A^{(m)})^t B^{(m)} = 0. \quad (2.3.14)$$

Pour que (2.3.14) soit vérifiée, il suffit de montrer que

$$\lim_{t \rightarrow \infty} (A^{(m)})^t = 0. \quad (2.3.15)$$

et ceci est équivalent à $\rho(A^{(m)}) < 1$. Pour déduire (2.3.14) de (2.3.15), nous devons montrer que pour un entier k fixé, les composantes de $\left((A^{(m)})^t B^{(m)}\right)$ sont toutes strictement positives.

Considérons la forme des matrices A_n donnée en (2.3.3).

- α_q et β_p sont tous les deux non nuls.

Dans ce cas la première composante de $X_{n,k}$ ne sera jamais nulle presque sûrement (resp. la $(q+1)^{\text{ème}}$ composante ne sera jamais nulle), pour tout $k \geq q+1$ (resp. pour tout $k \geq p+1$). De plus, nous avons $X_{n,k} = A_n X_{n-1,k-1}$, $\forall k > 0$, qui se traduit par la $i^{\text{ème}}$ composante de $X_{n,k}$ est la $(i-1)^{\text{ème}}$ composante de $X_{n-1,k-1}$ pour $i = 2, \dots, q$, il l'est aussi pour $j = 2, \dots, p$.

Ainsi, on conclut qu'aucune des q premières composantes de $X_{n,2q}$ (resp. les p dernières composantes de $X_{n,2p}$), n'est nulle presque sûrement, il en est de même pour $X_{n,k}$, $\forall k \geq 2q$ (resp. $X_{n,k}$, $\forall k \geq 2p$). Tenant compte de la positivité des variables $X_{n,k}$, ceci montre que toutes les composantes de $\left((A^{(m)})^t B^{(m)} \right)$ sont strictement positives, $\forall k \geq \max(2q, 2p)$.

- $\alpha_q \beta_p = 0$

Dans ce cas, nous pouvons remplacer X_n par un vecteur de dimension plus petite, obtenu en supprimant la composante ε_{n-q+1}^2 si α_q est nul, et la composante σ_{n-p+1}^2 , si β_p est nul. La matrice A_n est alors remplacée par une matrice de dimension plus petite, mais ayant les mêmes valeurs propres non nulles que celle de départ. Il en est de même pour la matrice $(A^{(m)})$. Si $\alpha_q \beta_p \neq 0$, nous retournons au cas précédent, sinon nous continuons le processus de réduction des dimensions.

Structure d'autocovariance des carrés du modèle $GARCH(p, q)$

Comme nous l'avons déjà vu, un processus (ε_n) satisfaisant l'écriture d'un $GARCH(p, q)$ peut s'écrire sous forme d'un modèle $ARMA$, pour le carré de ce processus s'il est stationnaire à l'ordre 4. Ainsi, nous pouvons exploiter cette dernière forme pour le calcul de la fonction d'autocovariance, néanmoins, ce calcul est difficile à réaliser. Francq et Zakoïan, 2009, utilisent la représentation markovienne (2.3.2) et proposent un algorithme qui permet de le faire en utilisant la représentation markovienne (2.3.2). Les autocovariance des carrés du modèle $GARCH$ sont positives.

2.4 Estimation d'un modèle $GARCH$

L'idée fondamentale de la méthode d'estimation par maximum de vraisemblance, comme son nom l'indique, est de trouver un ensemble d'estimations de paramètre, appelé $\hat{\theta}$, telle que la vraisemblance d'avoir obtenu l'échantillon que nous utilisons soit maximisée. Dans certains cas, les modèles économétriques conduisent à des expressions complexes de la distribution conditionnelle des variables endogènes. Les fonctions de vraisemblance qui leur sont associées sont difficilement manipulables, elles reposent sur l'hypothèse que le modèle

est complètement spécifié, ce qui rend difficile l'estimation des paramètres d'intérêt. Pour pallier à ces inconvénients, on peut approximer la fonction de vraisemblance, ou bien ne pas entièrement spécifier la structure paramétrique de ce modèle. Cette dernière solution est adoptée par la méthode du quasi de maximum de vraisemblance (QMV), qui est appliquée lorsque le modèle est seulement spécifié au premier et /ou au deuxième ordre.

L'estimateur du maximum de vraisemblance est non consistant si le modèle sous-jacent est mal spécifié, autrement dit, la fonction de vraisemblance utilisée n'est pas fondée sur la vraie distribution. Par contre, la méthode d'estimation QMV donne des estimateurs consistants asymptotiquement normaux, et cela dans le cas où la fonction de vraisemblance est fondée sur une loi appartenant à une famille exponentielle linéaire, lorsque le premier moment conditionnel est spécifié et à une famille exponentielle quadratique, lorsque le deuxième moment conditionnel est spécifié, même si la vraie distribution n'appartient pas à cette famille (voir Gouriéroux, Monfort et Trognon (1984)).

Cette méthode d'inférence s'est généralisée au cas dynamique, son utilisation pour les modèles $GARCH$ s'avère très intéressante, car les estimateurs qu'elle fournit ont de bonnes propriétés asymptotiques. Celles-ci, ont été établies par Weiss (1986) sous des conditions d'existence des moments d'ordre 4. Dans le cadre d'un $GARCH(1, 1)$ ces mêmes propriétés ont été démontrées sous des hypothèses de stationnarité stricte pour l'estimateur du QMV local, par Lumsdaine (1996). Dans Lee et Hansen (1994), la convergence de l'estimateur global est obtenue sous l'hypothèse de stationnarité au second-ordre.

2.4.1 Estimation du modèle par la méthode du quasi-maximum de vraisemblance

Nous considérons une réalisation $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)$ de processus $GARCH(p, q)$, solution strictement stationnaire, non anticipative du modèle

$$\varepsilon_n = Z_n \sqrt{h_n}, \quad (Z_n) \text{ i.i.d } (0, 1), \quad (2.4.1)$$

$$V(\varepsilon_n | \mathcal{F}_{n-1}) = h_n = \omega^0 + \sum_{i=1}^q \alpha_i^0 \varepsilon_{n-i}^2 + \sum_{j=1}^p \beta_j^0 h_{n-j}, \quad \forall n \in \mathbb{Z},$$

tels que : $\omega^0 > 0$, $\alpha_i^0 \geq 0$ et $\beta_j^0 \geq 0$ $i = 1, \dots, q$, et $j = 1, \dots, p$.

Ce modèle vérifie $\mathbb{E}(\varepsilon_n | \mathcal{F}_{n-1}) = 0$, et $Var(\varepsilon_n / \mathcal{F}_{n-1}) = h_n$, où \mathcal{F}_{n-1} est l'information dont on dispose jusqu'à la date $n - 1$.

Soit $\theta = (\theta_1, \dots, \theta_{p+q+1})' = (\omega, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$, le vecteur des paramètres appartenant à un espace compact des paramètres $\Theta \subset]0, +\infty[\times [0, +\infty[^{p+q}$. Nous notons la vraie valeur inconnue du vecteur des paramètres $\theta^0 = (\omega^0, \alpha_1^0, \dots, \alpha_q^0, \beta_1^0, \dots, \beta_p^0)'$.

Le fait que la distribution de (h_1, \dots, h_N) est inconnue, il n'y a pas une expression explicite de densité de probabilité de $(\varepsilon_1, \dots, \varepsilon_N)$, même en faisant l'hypothèse de la distribution normale de Z_n . Pour surmonter ces difficultés, nous considérons la fonction du quasi-maximum de vraisemblance conditionnelle gaussienne, où la vraisemblance conditionnelle est obtenue à partir d'une loi normale centrée réduite des variables Z_n .

Étant données des valeurs initiales $\varepsilon_0, \dots, \varepsilon_{1-q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2$, la fonction de vraisemblance conditionnelle L de θ est donnée par

$$\begin{aligned} L_N(\theta) &= f_\theta(\varepsilon_1, \dots, \varepsilon_N / \varepsilon_0, \dots, \varepsilon_{1-q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2), \\ &= \prod_{n=1}^N \frac{1}{\sqrt{2\pi\tilde{\sigma}_n^2}} \exp\left(-\frac{\varepsilon_n^2}{2\tilde{\sigma}_n^2}\right), \end{aligned}$$

où $\tilde{\sigma}_n^2$ s'obtient à partir de l'équation récursive suivante

$$\tilde{\sigma}_n^2 = \tilde{\sigma}_n^2(\theta) = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{n-i}^2 + \sum_{j=1}^p \beta_j \tilde{\sigma}_{n-j}^2, \forall n \geq 1.$$

Par exemple, les valeurs initiales peuvent être choisies comme $\epsilon_0^2 = \dots = \epsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}$, sous l'hypothèse de stationnarité au second ordre. Nous pouvons prendre comme valeurs initiales

$$\varepsilon_0^2 = \dots = \varepsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \omega. \quad (2.4.2a)$$

Un autre choix pour $\epsilon_0^2, \dots, \epsilon_{1-q}^2, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2$ est

$$\varepsilon_0^2 = \dots = \varepsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \varepsilon_1^2 \quad (2.4.2b)$$

Alors, la pseudo log-vraisemblance est de la forme

$$\log(L_N(\theta)) = \frac{1}{N} \sum_{n=1}^N \left(-\log(2\pi\tilde{\sigma}_n^2)^{\frac{1}{2}} - \frac{\varepsilon_n^2}{2\tilde{\sigma}_n^2} \right) \propto -\frac{1}{N} \sum_{n=1}^N \underbrace{\left(\log\tilde{\sigma}^2 + \frac{\varepsilon_n^2}{\tilde{\sigma}_n^2} \right)}_{\tilde{l}_n(\theta)}$$

Nous posons $\tilde{I}_N(\theta) = N^{-1} \sum_{n=1}^N \tilde{l}_n(\theta)$. Maximiser le logarithme de la vraisemblance revient à minimiser $\tilde{I}_N(\theta)$ par rapport au paramètre θ . Un estimateur du quasi-maximum de vraisemblance est une solution mesurable de l'équation

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L_N(\theta) = \arg \min_{\theta \in \Theta} \tilde{I}_N(\theta). \quad (2.4.3)$$

Le choix des valeurs initiales n'influent pas sur les propriétés asymptotiques de l'estimateur du QMV, néanmoins ce choix révèle important en pratique.

En annulant la première dérivée par rapport à θ , on tire

$$\frac{1}{N} \sum_{n=1}^N (\varepsilon_n^2 - \tilde{\sigma}_n^2) \frac{1}{\tilde{\sigma}_n^4} \frac{\partial \tilde{\sigma}_n^2}{\partial \theta} = 0. \quad (2.4.4)$$

Il sera pratique d'approximer la suite $(\tilde{l}_n(\theta))$ par une suite stationnaire ergodique $I_N(\theta) = N^{-1} \sum_{t=1}^N l_n(\theta)$, où $l_n(\theta) = \left(\log \sigma_n^2 + \frac{\varepsilon_n^2}{\sigma_n^2} \right)$. On note par, $(\sigma_n)_n = \{\sigma_n(\theta)\}$ la solution strictement stationnaire, ergodique et non anticipative de

$$\sigma_n^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{n-i}^2 + \sum_{j=1}^p \beta_j \sigma_{n-j}^2, \forall n \in \mathbb{Z}. \quad (2.4.5)$$

Alors, (2.4.4) devient

$$\frac{1}{N} \sum_{n=1}^N (\varepsilon_n^2 - \sigma_n^2) \frac{1}{\sigma_n^4} \frac{\partial \sigma_n^2}{\partial \theta} = 0.$$

Notons que $\sigma_n^2(\theta^0) = h_n$.

2.4.2 Convergence forte de l'estimateur du QMV

Soit la suite de matrices aléatoires $\mathbf{A}^0 = (A_n^0, n \in \mathbb{Z})$, où A_n^0 juste A_n donnée en (2.3.3), en remplaçant θ par θ^0 , *i.e.*,

$$A_n^0 = \left(\begin{array}{ccc|ccc} \alpha_1^0 Z_n^2 & \cdots & \alpha_q^0 Z_n^2 & \beta_1^0 Z_n^2 & \cdots & \beta_p^0 Z_n^2 \\ & I_{(q-1)} & O_{(q-1) \times 1} & & O_{(q-1) \times p} & \\ \hline \alpha_1^0 & \cdots & \alpha_q^0 & \beta_1^0 & \cdots & \beta_p^0 \\ & O_{(p-1) \times q} & & I_{(p-1)} & & O_{(p-1) \times 1} \end{array} \right)$$

Notons $\mathcal{A}_\theta(z) = \sum_{i=1}^q \alpha_i z^i$ et $\mathcal{B}_\theta(z) = 1 - \sum_{j=1}^p \beta_j z^j$, avec convention $\mathcal{A}_\theta(z) = 0$ si $q = 0$ et $\mathcal{B}_\theta(z) = 1$ si $p = 0$.

Pour étudier la convergence forte des estimateur du QMV $\hat{\theta}_N$, nous considérons les hypothèses suivantes établies par Francq et Zakoïan (2004).

A1 : $\theta^0 \in \Theta$ et Θ est compact.

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

$$\begin{aligned} \gamma(\mathbf{A}^0) &= \inf_{n \in \mathbb{N}^*} \frac{1}{n} \mathbb{E} (\log \|A_n^0 A_{n-1}^0 \cdots A_1^0\|), \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n^0 A_{n-1}^0 \cdots A_1^0\| \text{ p.s.} \end{aligned}$$

A3 : Z_n^2 a une loi non dégénérée et $\mathbb{E}(Z_n^2) = 1$.

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 $\alpha_q^0 + \beta_p^0 \neq 0$.

Le résultat suivant établit la convergence forte de $\hat{\theta}_N$.

Théorème 2.4.1 (Francq et Zakoïan, 2004)

Soit $(\hat{\theta}_N)_N$ une suite d'estimateurs du QMV satisfaisant (2.4.3), avec les conditions initiales (2.4.2a) ou (2.4.2b). Sous les hypothèses **A1-A4**,

$$\hat{\theta}_N \rightarrow \theta^0 \text{ p.s. } \text{ quand } N \rightarrow \infty.$$

Le symbole \rightarrow désigne la convergence presque sûre.

Remarques :

- 1) La première partie de l'hypothèse d'identifiabilité (**A4**) concernant les racines communes a été déjà faite par Berkes et *al* (2003a,b). Il est remarquable que la vraie valeur θ^0 du paramètre n'appartient pas à l'intérieur de Θ , ainsi, le théorème permet de traiter les cas où certains coefficients α_i ou β_j , sont nuls. Cela est essentiel, en particulier, pour atteindre (maintenir) les situations de suridentification de l'un des ordres, p ou q , mais pas les deux en même temps. Généralement, on estime de manière consistante les paramètres d'un $GARCH(p-1, q)$ (ou d'un $GARCH(p, q-1)$, si nous utilisons un $GARCH(p, q)$).
- 2) Quand $p \neq 0$, l'hypothèse (**A4**) exclut le cas où tous les coefficients α_i^0 sont nuls. Ceci est évidemment nécessaire, sinon le modèle a pour solution un bruit blanc *i.i.d.*.
- 3) L'hypothèse d'absence de racines communes, dans (**A4**), est toujours vérifiée lorsque $p > 1$ et $q > 1$. Si $q = 1$, la seule racine de $\mathcal{A}_{\theta^0}(z)$ est zéro et $\mathcal{B}_{\theta^0}(0) \neq 0$. Si $p = 1$ et $\beta_1^0 \neq 0$, la seule racine de $\mathcal{B}_{\theta^0}(z)$ est $1/\beta_1^0 > 0$ (si $\beta_1^0 \neq 0$, le polynôme n'admet pas de racine). En raison de la positivité des coefficients α_i^0 , cette valeur ne peut annuler $\mathcal{A}_{\theta^0}(z)$.
- 4) L'hypothèse que $\mathbb{E}(Z_n) = 0$ n'est pas nécessaire pour l'établissement des propriétés asymptotiques de l'estimateur du QMV. La variance conditionnelle de (ε_n) est seulement proportionnelle à h_n dans ce modèle *i.e.*, $var(\varepsilon_n/\mathcal{F}_{n-i}, i > 0) = \{1 - (\mathbb{E}(Z_n))^2\} h_n$. L'hypothèse $\mathbb{E}(Z_n^2) = 1$ est faite pour des raisons d'identifiabilité et n'est pas restrictive dès que $\mathbb{E}(Z_n^2) < \infty$.

2.4.3 Normalité asymptotique des estimateurs du QMV

Revenons maintenant à la normalité asymptotique de l'estimateur $\hat{\theta}_N$. Pour ce faire, nous considérons les conditions supplémentaires suivantes (faites toujours par Francq et Zakoian (2004).

A5 : $\theta^0 \in \overset{\circ}{\Theta}$, où $\overset{\circ}{\Theta}$ est l'intérieur de Θ .

A6 : $\kappa_Z = \mathbb{E}(Z_n^4) < \infty$.

La loi limite de $\widehat{\theta}_N$ est donné par le théorème suivant.

Théorème 2.4.2 (Francq et Zakoian, 2004)

Sous les hypothèses **A1-A6**,

$$\sqrt{N} \left(\widehat{\theta}_N - \theta^0 \right) \xrightarrow{\mathcal{L}} N \left(0, (\kappa_Z - 1) J^{-1} \right),$$

où

$$J := \mathbb{E}_{\theta^0} \left(\frac{\partial^2 l_n(\theta^0)}{\partial \theta \partial \theta'} \right) = \mathbb{E}_{\theta^0} \left(\frac{1}{\sigma_n^4(\theta^0)} \frac{\partial \sigma_n^2(\theta^0)}{\partial \theta} \frac{\partial \sigma_n^2(\theta^0)}{\partial \theta'} \right)$$

est une matrice définie positive.

Remarques :

- 1) L'hypothèse **A5** est standard car elle permet d'utiliser le fait que les conditions du premier ordre sont valides, au moins asymptotiquement. Cette hypothèse est nécessaire pour l'obtention de normalité asymptotique, en effet, si certaines composantes de θ^0 sont nulles, **A5** n'est pas vérifiée et le théorème ne peut pas conclure. Par exemple, lorsque $\alpha_1^0 = 0$, la loi de $\sqrt{N}(\widehat{\alpha}_1 - \alpha_1^0)$ est, pour tout N , concentrée sur $[0, \infty[$ et ne peut pas être asymptotiquement normale. Ce type de problèmes, dits "effet de bord", ont été étudiés par Andrew (1999) dans le cadre d'un $GARCH(1, q)$.
- 2) Ce théorème ne requiert aucune hypothèse sur la distribution de (Z_n) , apart l'existence du moment d'ordre 4. Cette hypothèse est clairement nécessaire pour l'existence de la variance asymptotique du vecteur du score $\partial l_n(\theta^0) / \partial \theta$. On note aussi que cette hypothèse n'implique pas l'existence du moment d'ordre deux de processus observé (ε_n) .

2.4.4 Simulation

Dans cette section, nous présentons des résultats numériques d'une étude de simulation illustrant la consistance au sens de la convergence presque sûre (500 réplifications du modèle

$GARCH(1, 1)$ ont été effectuées) de l'estimateur du QMV, $\hat{\theta}_N$, en échantillon fini, dans le cas stationnaire, pour les tailles suivantes : $N = 250, 500, 1000, 5000$ et 10000 . Les estimations moyennes du paramètre $\theta^0 = (\omega^0, \alpha_1^0, \beta_1^0)$, ainsi que leurs écarts-type empiriques (Root Mean-Square Error ($RMSE$)) correspondants sont portés dans les tables (2.4.1 – 2.4.2 et 2.4.3) ci-dessous pour les différentes valeurs de v où ce dernier désigne le degré de liberté (pour les simulations, nous avons utilisé une loi de Student pour $v = \infty, v = 9$ et $v = 5$ et on s'est basé sur la méthode de vraisemblance fondée sur la loi gaussienne.

v	Taille de l'échantillon	ω^0	$\hat{\omega}$	$RMSE(\hat{\omega})$
∞	$n = 250$	(0.014)	$\begin{pmatrix} 0.0962 \\ 0.033 \\ 0.0199 \\ 0.0149 \\ 0.0145 \end{pmatrix}$	$\begin{pmatrix} 0.1944 \\ 0.0618 \\ 0.0117 \\ 0.0034 \\ 0.0025 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			
9	$n = 250$	(0.014)	$\begin{pmatrix} 0.0887 \\ 0.0296 \\ 0.0206 \\ 0.0148 \\ 0.0146 \end{pmatrix}$	$\begin{pmatrix} 0.1563 \\ 0.0297 \\ 0.012 \\ 0.0037 \\ 0.0029 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			
5	$n = 250$	(0.014)	$\begin{pmatrix} 0.1094 \\ 0.0327 \\ 0.0203 \\ 0.0151 \\ 0.0145 \end{pmatrix}$	$\begin{pmatrix} 0.2377 \\ 0.0447 \\ 0.0153 \\ 0.0043 \\ 0.0029 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			

Table 2.4.1

v	Taille de l'échantillon	α_1^0	$\hat{\alpha}_1$	$RMSE(\hat{\alpha}_1)$
∞	$n = 250$	(0.084)	$\begin{pmatrix} 0.0871 \\ 0.0849 \\ 0.0843 \\ 0.0839 \\ 0.0841 \end{pmatrix}$	$\begin{pmatrix} 0.0459 \\ 0.0276 \\ 0.0188 \\ 0.0084 \\ 0.0059 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			
9	$n = 250$	(0.084)	$\begin{pmatrix} 0.0934 \\ 0.0854 \\ 0.0848 \\ 0.0841 \\ 0.0840 \end{pmatrix}$	$\begin{pmatrix} 0.0536 \\ 0.0353 \\ 0.0202 \\ 0.0092 \\ 0.0074 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			
5	$n = 250$	(0.084)	$\begin{pmatrix} 0.0992 \\ 0.0906 \\ 0.0858 \\ 0.0843 \\ 0.0844 \end{pmatrix}$	$\begin{pmatrix} 0.0822 \\ 0.0442 \\ 0.0293 \\ 0.0131 \\ 0.0094 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			

Table 2.4.2

v	Taille de l'échantillon	β_1^0	$\hat{\beta}_1$	$RMSE(\hat{\beta}_1)$
∞	$n = 250$	(0.905)	$\begin{pmatrix} 0.8205 \\ 0.8831 \\ 0.8984 \\ 0.9040 \\ 0.9043 \end{pmatrix}$	$\begin{pmatrix} 0.1928 \\ 0.0701 \\ 0.0233 \\ 0.0093 \\ 0.0064 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			
9	$n = 250$	(0.905)	$\begin{pmatrix} 0.8054 \\ 0.8858 \\ 0.8959 \\ 0.9038 \\ 0.9042 \end{pmatrix}$	$\begin{pmatrix} 0.2141 \\ 0.0524 \\ 0.0248 \\ 0.0104 \\ 0.0082 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			
5	$n = 250$	(0.905)	$\begin{pmatrix} 0.7607 \\ 0.8744 \\ 0.8950 \\ 0.903 \\ 0.9038 \end{pmatrix}$	$\begin{pmatrix} 0.2595 \\ 0.0774 \\ 0.0347 \\ 0.0137 \\ 0.0095 \end{pmatrix}$
	$n = 500$			
	$n = 1000$			
	$n = 5000$			
	$n = 10000$			

Table 2.4.3

Il est remarquable que pour des échantillons de taille moyenne ($N = 250, 500, 1000$) la méthode du QMV surestime la constante ω^0 et α_1^0 mais d'un degré plus faible pour ce dernier, contrairement au paramètre β_1^0 qui est sous-estimé par cette méthode. De plus, nous constatons que les valeurs des paramètres estimés par la méthode de QMV se rapprochent de celles de la vraie valeur du paramètre $\theta^0 = (0.014, 0.084, 0.95)$ au fur à mesure que la taille de l'échantillon augmente et leurs $RMSE$ associés diminuent. Néanmoins, la convergence de l'estimateur du QMV se manifeste clairement à partir d'une taille d'échantillon égale à 5000, plus précisément pour $N = 5000$ le vecteur des paramètres estimés $\widehat{\theta}_N = (\widehat{\omega}, \widehat{\alpha}_1, \widehat{\beta}_1)$ passe d'une valeur de $(0.0148, 0.0841, 0.9038)$ lorsque $v = 9$ (respectivement $\widehat{\theta}_N = (0.0151, 0.0843, 0.903)$ pour $v = 5$) à une valeur de $(0.0146, 0.0840, 0.9042)$ pour $N = 10000$ et $v = 9$ (respectivement $\widehat{\theta}_N = (0.0145, 0.0844, 0.9038)$ pour $v = 5$). Ainsi, ces résultats de simulations confirment les résultats théoriques concernant la consistance de l'estimateur du QMV.

2.5 Conclusion

D'un point de vue statistique, les séries financières sont des objets complexes qui possèdent un certain nombre de propriétés connues et résumées sous le nom de "faits stylisés".

La modélisation classique $ARMA$, ne rend pas compte de ces propriétés que nous venons de décrire. Dans ce chapitre, nous avons présenté une classe de processus qui permet de combler partiellement ces lacunes, les processus $GARCH$. Nous avons étudié la structure probabiliste dans un premier temps et l'estimation des paramètres, dans un deuxième, temps du modèle associé.

La modélisation de la volatilité des rendements par un modèle $GARCH$, est d'une importance majeure dans la quantification des risques de management. Dans ce modèle, c'est le second moment conditionnel qui varie en fonction du temps suivant un certain modèle de régression déterministe. Par exemple, la volatilité (au carré) d'aujourd'hui décrite par un $GARCH(1, 1)$ est expliquée par la volatilité d'hier et ce qui s'est passé hier, *i.e.* les rendements.

La volatilité est inobservable. De nombreux auteurs ont exploité des données intrajourna-

lières, c'est-à-dire, des données de haute fréquence pour proposer des mesures non paramétriques pour la volatilité. Comment choisir ces proxies de volatilité ? Peuvent-elles améliorer l'estimation des paramètres du modèle sous considération ? C'est l'objet des deux chapitres suivants.

3

Proxies de volatilité et volatilité réalisée

3.1 Introduction

Le concept de volatilité est probablement l'un des sujets qui suscite le plus de recherches dans le domaine de la finance mathématique. Cet intérêt pour la volatilité est motivé par deux raisons importantes : le nombre de plus en plus de compagnies utilisant les outils de gestion des risques et le grand nombre de produits dérivés transigés dans les marchés financiers mondiaux. De même, quand une compagnie veut étudier son degré d'exposition à un risque financier, elle doit être capable d'évaluer la volatilité de chacun des biens qu'elle possède.

La volatilité peut être définie comme la variabilité des prix des actifs financiers, des cours des actions, par exemple. Autrement dit, c'est la variabilité de la variable sous considération. Plus, la variable fluctue durant une période, plus elle est censée être volatile. La volatilité est associée à l'imprévu, l'incertain et le risque. Pour le public général, le terme est synonyme de risque. Ainsi, une forte volatilité est considérée comme un symptôme de perturbations. Les transactions des valeurs boursières ne sont pas à leurs justes valeurs et le marché de capital ne fonctionne pas bien comme il le faut.

Par essence, la volatilité est un phénomène non observable ; elle est inhérente à la plupart des modèles financiers. Comme elle ne s'observe pas naturellement, alors, il est nécessaire de

construire un certain nombre de statistiques dites *proxies*, et l'on se base sur ces dernières lors de spécification, d'estimation et d'évaluation des modèles de volatilité.

Mesurer la volatilité, en utilisant des données intra-journalières, à une fréquence très élevée, de façon presque continue est possible. L'idée de la volatilité réalisée est initialement proposée par French, Schwert et Stambaugh (1987). Taylor et Xu (1997) ainsi qu'Andersen et Bollerslev (1997) montrent que la volatilité journalière réalisée peut être construite en additionnant les rendements intra-journaliers au carré, puis en prenant la racine carrée de cette quantité.

Une motivation théorique de la construction de la volatilité réalisée réside dans le fait que la variation quadratique d'une grande classe de semimartingales est un estimateur sans biais de la variance conditionnelle des rendements, (voir, par exemple, Barndorff-Nielsen et Shephard (2002), et Andersen, Bollerslev, Diebold, et Labys (2003)). En effet, la disponibilité de données de haute fréquence crée un intérêt pour cette approche.

Une volatilité observable donne lieu à de nouvelles opportunités : on peut l'analyser, l'optimiser et la prévoir en utilisant des techniques plus simples que les modèles économétriques complexes utilisés lorsque la volatilité est latente. Effectivement, les *proxies* de volatilité construites sur la base de données intra-journalières ont un pouvoir de prévision supérieur à celles construites à partir des rendements journaliers, (voir, par exemple, Ghysels, Santa-Clara, et Valkanov (2006), Engle et Gallo (2006)). Généralement, la recherche d'une bonne proxy souffre de l'absence d'une méthode de comparaison de ces proxies.

3.2 Modèle d'échelle

Les modèles à temps discret, comme le modèle *GARCH* et les modèles de volatilité stochastique, souvent admettent la forme multiplicative suivante

$$r_n = \sigma_n Z_n, n \in \mathbb{Z}, \quad (3.2.1)$$

où (Z_n) sont des variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*) de moyenne nulle et de variance égale à l'unité *i.e.*, $\mathbb{E}(Z_n) = 0$, $\mathbb{E}(Z_n^2) = 1$, et (σ_n) est une suite de variables aléatoires strictement positives. De plus Z_n et σ_n sont indépendantes. Du

fait que, le facteur d'échelle σ_n est non observable, il est naturel de penser à son estimation en se basant sur des données intra-journalières (haute fréquence). Pour que cet objectif soit atteint, une extension au temps continu du modèle (3.2.1) a été faite par de Vilder et Visser (2008), en gardant la même structure multiplicative donnée en (3.2.1) où la variable aléatoire Z_n a été remplacée par un processus stochastique $\Psi_n(\cdot)$ càdlàg (continu à droite et possédant une limite à gauche). Pour chaque jour ouvrable, $n = 1, 2, \dots, N$, on observe le processus des rendements logarithmiques intra-journaliers, écrit sous forme du produit d'un processus Ψ_n (standard day process) et une variable d'échelle σ_n .

Nous avons la définition suivante.

Définition 3.2.1

Les processus des rendements logarithmiques intra-journaliers $R_n(\cdot)$ sont dits satisfaire le modèle d'échelle s'ils admettent la représentation

$$R_n(u) = \sigma_n \Psi_n(u), \quad 0 \leq u \leq 1, \quad (3.2.2)$$

tels que :

- (i) Les facteurs d'échelle σ_n sont strictement positifs,
- (ii) Les processus càdlàg $\Psi_n(\cdot)$ ont la même distribution de probabilité pour chaque trading day, n ,
- (iii) Le processus $\Psi_n(\cdot)$ est indépendant de \mathcal{G}_{n-1} , pour tout n , où

$$\mathcal{G}_{n-1} = \sigma \left\{ (\Psi_i(u))_{i \leq n-1}, 0 \leq u \leq 1, (\sigma_i)_{i \leq n} \right\},$$

la tribu engendrée par le passé des processus jusqu'à l'instant $n - 1$ (la clôture de la $(n - 1)^{\text{ème}}$ jour ouvrable).

Remarques :

- 1) Les deux dernières conditions dans la définition (3.2.2), impliquent que la suite des processus $(\Psi_n(\cdot))_n$ est *i.i.d.* Le processus $\Psi_n(\cdot)$ peut être n'importe quel modèle représentant la trajectoire des rendements intra-journaliers à une échelle près. À noter que le facteur d'échelle σ_n (volatilité intra-journalière) reste invariant au cours de la journée ouvrable concernée n .

- 2) $R_n(1)$ est égale au rendement *close-to-close*. Si nous posons $Z_n = \Psi_n(1)$, nous retrouvons $r_n = R_n(1) = \sigma_n Z_n$.
- 3) À la fois σ_n et $\Psi_n(\cdot)$ ne sont pas supposés avoir une structure particulière.

Nous définissons les rendements logarithmiques *close-to-close* par $r_n = R_n^C - R_{n-1}^C$, pour tout $n \in \mathbb{Z}$, où R_n^C est la valeur de clôture pour la $n^{\text{ème}}$ journée ouvrable, avec la convention $R_0^C \equiv 0$. Nous notons R_n la valeur d'ouverture de la journée suivante ($n+1$), et nous adoptons l'hypothèse standard de continuité à droite et d'existence de la limite à gauche des trajectoires de ce processus (*càdlàg*).

Considérons la $n^{\text{ème}}$ journée $[n-1, n)$. Soit $\vartheta \in [0, 1[$, désignant le temps au cours de la journée ouvrable. Nous posons $t = n-1 + \vartheta$. Le processus des rendements cumulatifs $\{R_t, t \in \mathbb{R}\}$ satisfait la représentation

$$R_{n-1+\vartheta} = R_{n-1}^C + \sigma_n \Psi_n(\vartheta), \quad 0 \leq \vartheta < 1,$$

Remarques :

- 1) Pour chaque modèle à temps discret de la forme (3.2.1) des rendements logarithmiques *close-to-close* (r_n), il existe une extension en temps continu $R_n(\cdot)$, satisfaisant la représentation (3.2.2).
- 2) Le rendement *overnight*, en partant de la $(n-1)^{\text{ème}}$ journée ouvrable à la $n^{\text{ème}}$ journée ouvrable est donné par $R_{n-1} - R_{n-1}^C$.

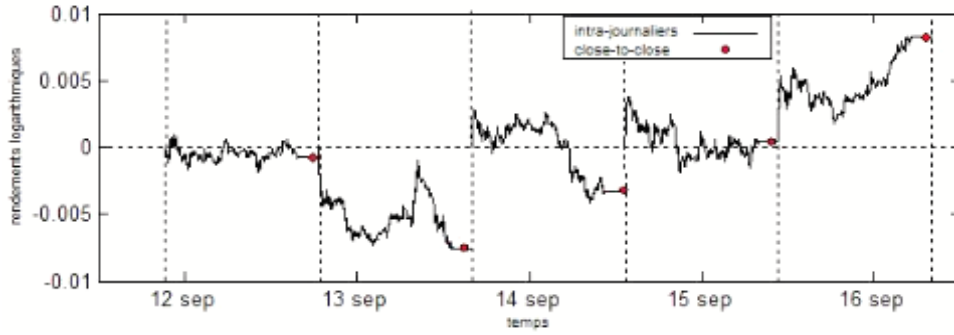


Figure 3.2.1 : Cinq trajectoires de processus des rendements intra-journaliers $R_n(\cdot)$, ainsi que les rendements *close-to-close* r_n , de l'indice boursier *S&P500* spot, commençant de 12-09-2005.

La figure (3.2.1) ci-dessus montre cinq trajectoires de processus des rendements logarithmiques intra-journaliers $R_n(\cdot)$. Pour chaque jour ouvrable n , nous observons le processus $R_n(\cdot)$, où les rendements sont calculés par rapport à la valeur de clôture (close) du jour précédent. Ce processus $R_n(\cdot)$, commence par la valeur de rendement *overnight* à l'instant $u = 0$, et se termine par le rendement *close-to-close* $R_n(1) = r_n$ (à l'instant $u = 1$). Il est remarquable que les trajectoires de processus des rendements intra-journaliers n'ont pas une forme unique, ils diffèrent d'une journée à une autre.

3.3 Proxies de volatilité

En statistique, le terme proxy est utilisé pour une variable qui n'est pas d'intérêt premier, en elle-même, mais qui est étroitement liée à un objet d'intérêt. Elle est reliée au concept d'estimateur. Mais les proxies tendent à remplacer les variables d'intérêt non observables, tandis qu'un estimateur est une statistique dont le rôle est l'estimation de paramètres. Un bon estimateur des paramètres est celui qui possède, par exemple, la plus petite erreur quadratique moyenne. Par analogie, une bonne proxy est celle qui possède une forte corrélation avec la variable.

3.3.1 Définitions et notations

Nous considérons l'espace de Skorohod $\mathbb{D}[0, 1]$ des fonctions càdlàg dans $[0, 1]$. Nous supposons que les éléments de l'espace de $\mathbb{D}[0, 1]$ sont continus en 1, de plus l'espace $\mathbb{D}[0, 1]$

est séparable, muni d'une métrique complète est doté de la topologie de Skorohod. L'espace $C[0, 1]$ des fonctions continues sur l'intervalle unité $[0, 1]$ est un sous espace vectoriel du $\mathbb{D}[0, 1]$.

Dans le reste de cette partie nous considérons le modèle d'échelle (3.2.2) et les proxies qu'on définira par la suite sont dites proxy de volatilité. Rappelons qu'une fonctionnelle H définie sur l'espace vectoriel D à valeurs dans $[0, \infty[$, est dite positivement homogène si $H(\alpha f) = \alpha H(f)$, $\alpha \in [0, \infty[$, $f \in D$. À noter que dans notre cas les fonctionnelles proxy sont des fonctionnelles proxy pour Ψ .

Définition 3.3.1

Soit H une fonctionnelle mesurable, positivement homogène, définie sur un sous espace vectoriel D de l'espace de Skorohod $\mathbb{D}[0, 1]$. Supposons que $\Psi \in D$ p.s, et $H(\Psi)$ est strictement positive p.s. Alors, H est une proxy fonctionnelle et la variable aléatoire $H_n \equiv H(R_n(\cdot))$ est dite proxy¹.

À titre d'exemple de proxies de volatilité, nous citons : les rendements absolus journaliers *close-to-close*, les rendements absolus *overnight*, l'étendue des rendements (max-min) et la volatilité réalisée qui sera définie ci-dessous.

Remarques :

- 1) Pour une réalisation ψ_n de Ψ , qui n'appartient pas au sous-espace vectoriel D , on peut définir $H_n = 0$. Une fonctionnelle proxy n'est pas linéaire, en effet $H(\Psi) + H(-\Psi) > 0$, mais $H(\Psi - \Psi) = H(0) = 0$ et de plus n'est pas symétrique : $H(\Psi) \neq H(-\Psi)$.
- 2) Les proxies sont linéaires en facteur d'échelle σ_n . En effet, $H_n = H(\sigma_n \Psi_n)$, et par homogénéité de la fonctionnelle H , on obtient

$$H_n = \sigma_n H(\Psi_n) \quad (3.3.1)$$

où $(H(\Psi_n))_n$ est une suite d'innovations *i.i.d.* strictement positives.

De plus σ_n et $H(\Psi_n)$ sont indépendantes dès que σ_n et Ψ_n le sont. Ainsi la proxy H_n donne des informations concernant le facteur d'échelle σ_n . L'équation (3.3.1) a la même

¹Les deux termes H_n, H désignent la notion de proxy

forme que celle des rendements des actifs en temps discret donnée en (3.2.1).

À partir d'une fonctionnelle de proxy, nous pouvons construire d'autres fonctionnelles de proxy, comme le montre la proposition suivante.

Proposition 3.3.1 (de Vilder et Visser, 2008)

Soient $H^{(i)}$, $i = 1, \dots, d$, des fonctionnelles de proxy. Définissons une fonction mesurable, positivement homogène : $G : [0, \infty[^d \rightarrow [0, \infty[$. De plus, on considère que $G(\underline{x}) > 0$ si $\underline{x} \neq 0_{\mathbb{R}^d}$. Alors, la fonctionnelle de proxy $H : f \mapsto G(H^{(1)}(f), \dots, H^{(d)}(f)) \forall f \in D$ est une fonctionnelle de proxy.

Exemple

On considère deux fonctionnelles de proxy $H^{(1)}$ et $H^{(2)}$. Les fonctionnelles $H : f \mapsto H(f)$ données ci-dessous définissent à leur tour des fonctionnelles de proxy.

1. $H(f) \equiv aH^{(1)}(f)$, pour $a > 0$.
2. $H(f) \equiv \varpi_1 H^{(1)}(f) + \varpi_2 H^{(2)}(f)$, pour $\varpi_1, \varpi_2 \in]0, \infty[$.
3. $H(f) \equiv (H^{(1)}(f))^{\varpi_1} (H^{(2)}(f))^{\varpi_2}$ pour $\varpi_1, \varpi_2 \in \mathbb{R}$, tel que $\varpi_1 + \varpi_2 = 1$.
4. $H(f) \equiv \max \{H^{(1)}(f), H^{(2)}(f)\}$ et $H(f) \equiv \min \{H^{(1)}(f), H^{(2)}(f)\}$.

3.3.2 Ordre sur les proxies

Proxy optimale

La restriction aux proxies de volatilité positivement homogènes, nous permet de les comparer et de les optimiser sans recours à l'identification du modèle de facteur d'échelle σ_n .

En introduisant l'opérateur logarithmique sur la proxy H_n donnée en (3.3.1), nous obtenons

$$\log(H_n) = \log(\sigma_n) + U_n, \quad (3.3.2)$$

où $U_n = \log(H(\Psi_n))$ est dit l'erreur (bruit). Pour juger de la qualité d'une proxy, nous utilisons le critère de l'erreur quadratique moyenne (EQM). Ce critère nous permet de com-

parer et d'optimiser les proxies :

$$\mathbb{E}(\log(H_n) - \log(\sigma_n))^2 = \mathbb{E}(U_n^2) = \text{var}(U_n) + (\mathbb{E}(U_n))^2.$$

Les proxies H_n ne sont pas nécessairement sans biais $\mathbb{E}(U_n) \neq 0$. Pour contourner le problème des proxies biaisées, de Vilder et Visser (2008), définissent une nouvelle proxy à une constante près strictement positive, $\tilde{H}_n = c H_n$, cela conduit à une EQM :

$$\begin{aligned} \mathbb{E}(\log(\tilde{H}_n) - \log(\sigma_n))^2 &= \text{var}(\tilde{U}_n) + (\mathbb{E}(\tilde{U}_n))^2, \\ &= \text{var}(U_n) + (\mathbb{E}(U_n + \log c))^2, \end{aligned}$$

où $\tilde{U}_n = U_n + \log c$.

La constante $c = c^*$ qui minimise EQM est obtenu en posant $\mathbb{E}(\tilde{U}_n) = 0$, ce qui donne

$$\mathbb{E}(\log(\tilde{H}_n) - \log(\sigma_n))^2 = \text{var}(U_n).$$

Minimiser l'EQM revient à minimiser la quantité $\text{var}(U_n)$.

Définition 3.3.2

Soient les deux proxies suivantes $H^{(1)}$ et $H^{(2)}$ dont la variance de leurs logarithmes est $(\lambda^{(i)})^2 = \text{var}(\log(H^{(i)}(\Psi)))$, $i = 1, 2$ respectivement. Alors, la proxy $H^{(1)}$ est dite meilleure que $H^{(2)}$ si sa variance $(\lambda^{(1)})^2$ est inférieure ou égale à celle de $H^{(2)}$, autrement dit, $(\lambda^{(1)})^2 \leq (\lambda^{(2)})^2$. Une proxy H^* possédant une variance finie est dite optimale si elle est la meilleure parmi toutes les autres proxies i.e., $\text{var}(\log(H^*(\Psi))) = \inf_H \text{var}(\log(H(\Psi)))$.

La proposition ci-dessous montre qu'une bonne proxy est caractérisée par une forte corrélation avec le facteur d'échelle σ_n .

Proposition 3.3.2 (de Vilder et Visser (2008))

Si $0 < \text{var}(\log(\sigma_n))$, alors

$$\text{corr}(\log(H_n), \log(\sigma_n)) = \left(1 + \frac{\lambda^2}{\text{var}(\log(\sigma_n))}\right)^{-\frac{1}{2}}, \quad (3.3.3)$$

où $\lambda^2 = \text{var}(\log(H(\Psi)))$.

De plus, supposons que $(\lambda^{(1)})^2 = \text{var}(\log(H^{(1)}(\Psi))) \leq (\lambda^{(2)})^2 = \text{var}(\log(H^{(2)}(\Psi))) < \infty$,

alors, (3.3.3) implique

$$\text{corr}(\log(H_n^{(1)}), \log(\sigma_n)) \geq \text{corr}(\log(H_n^{(2)}), \log(\sigma_n)).$$

Dans le cas où $\lambda^2 = 0$, on aura une corrélation parfaite entre la proxy H_n et le facteur d'échelle σ_n qui se traduit par : $\text{corr}(\log(H_n), \log(\sigma_n)) = 1$.

Preuve (de la proposition 3.3.2)

Rappelons que : $\text{var}(U_n) = \text{var}(\log(H(\Psi_n)))$. Dès que $\log(H_n) = \log(\sigma_n) + U_n$, on a

$$\begin{aligned} \text{corr}(\log(H_n), \log(\sigma_n)) &= \frac{\text{cov}(\log(H_n), \log(\sigma_n))}{[\text{var}(\log(H_n)) \text{var}(\log(\sigma_n))]^{\frac{1}{2}}}, \\ &= \frac{\text{var}(\log(\sigma_n)) + \text{cov}(\log(\sigma_n), U_n)}{[(\text{var}(\log(\sigma_n)) + \text{var}(U_n)) \text{var}(\log(\sigma_n))]^{\frac{1}{2}}}. \end{aligned}$$

Étant donné l'indépendance entre σ_n et Ψ_n , alors, $\log(\sigma_n)$ et U_n le sont aussi. Après simplifications on obtient (3.3.3).

Si $\lambda^2 = 0$, cette corrélation nous donne une valeur qui est égale à l'unité.

L'identification des deux processus Ψ_n et σ_n n'est pas nécessaire lors de l'étude des proxies, comme le montre la proposition suivante.

Proposition 3.3.3 (de Vilder et Visser (2008))

Considérons les deux fonctionnelle proxies $H^{(1)}$ et $H^{(2)}$. De plus, supposons que toutes les deux représentations (σ_n, Ψ_n) , (σ'_n, Ψ'_n) satisfont le modèle d'échelle donné en (3.2.2). Si $H^{(1)}$ est meilleure que $H^{(2)}$ pour Ψ , alors, $H^{(1)}$ l'est aussi pour Ψ' .

Preuve (de Proposition 3.3.3)

L'objectif est de montrer que : $var (\log (H^{(1)} (\Psi'_n))) \leq var (\log (H^{(2)} (\Psi'_n)))$ si $var (\log (H^{(1)} (\Psi_n))) \leq var (\log (H^{(2)} (\Psi_n)))$. On a par hypothèse $\sigma_n \Psi_n = \sigma'_n \Psi'_n$ et de plus σ_n et Ψ_n sont indépendants, alors

$$\begin{aligned} var (\log (\sigma_n)) + var (\log (H^{(1)} (\Psi_n))) &\leq var (\log (\sigma_n)) + var (\log (H^{(2)} (\Psi_n))), \\ &= var (\log (\sigma'_n)) + var (\log (H^{(2)} (\Psi'_n))). \end{aligned}$$

Comme : $var (\log (\sigma_n)) + var (\log (H^{(1)} (\Psi_n))) = var (\log (\sigma'_n)) + var (\log (H^{(1)} (\Psi'_n)))$.
On aura $var (\log (\sigma'_n)) + var (\log (H^{(1)} (\Psi'_n))) \leq var (\log (\sigma'_n)) + var (\log (H^{(2)} (\Psi'_n)))$ et par conséquence : $var (\log (H^{(1)} (\Psi'_n))) \leq var (\log (H^{(2)} (\Psi'_n)))$.

Le théorème suivant garantit l'existence d'une proxy optimale sous certaine condition.

Remarque :

- Considérons une bonne proxy H et un paramètre α strictement positif. Alors, la fonctionnelle définie par $G(f) = \alpha H(f)$ constitue à son tour une bonne proxy, vue que $var (\log (G(f))) = var (\log (\alpha H(f))) = var (\log (\alpha)) + var (\log (H(f))) = var (\log (H(f)))$.

Théorème 3.3.1. (de Vilder et Visser, 2008)

S'il existe une fonctionnelle proxy avec une variation de nuisance finie, alors il existe une fonctionnelle proxy optimale.

Proposition 3.3.4 (de Vilder et Visser 2008)

Supposons que $H^{(1)}$ et $H^{(2)}$ sont deux fonctionnelle proxy optimales. Alors, il existe une constante α strictement positive telle que $H^{(1)} (\Psi) \stackrel{p.s}{=} \alpha H^{(2)} (\Psi)$.

D'un point de vue théorique, le théorème précédent (3.3.1), garantit l'existence de proxies optimales. Néanmoins, il présente des limites en pratique. Cela est dû à la forme inconnue du processus $\Psi_n(\cdot)$ et de sa non observabilité. Selon les deux auteurs de Vilder et Visser (2008), la comparaison entre les proxies, en pratique, se fait en utilisant la variance estimée du logarithme de leurs proxies. Ce qui nous permet d'établir la relation d'équivalence suivante

$$\left(\lambda^{(1)}\right)^2 \leq \left(\lambda^{(2)}\right)^2 \iff \text{var} \left(\log \left(H_n^{(1)}\right)\right) \leq \text{var} \left(\log \left(H_n^{(2)}\right)\right). \quad (3.3.4)$$

Ainsi, pour comparer les proxies, il suffit de comparer la variance de leurs proxies logarithmiques.

En pratique, on utilise les variances estimées.

Dans le cadre de recherche d'une bonne proxy, de Vilder et Visser (2008) ont proposé une méthode pour optimiser cette dernière. Ils définissent une nouvelle proxy $H_n^{(\varpi)}$ dite proxy géométrique issue d'une combinaison d'un nombre fini de proxy $H_n^{(i)}, i = 1, \dots, d$, tel que $H_n^{(\varpi)} = \prod_{i=1}^d \left(H_n^{(i)}\right)^{\varpi_i}$ où $\varpi = (\varpi_1, \dots, \varpi_d)'$, sous la contrainte $\sum_{i=1}^d \varpi_i = 1, \varpi_i \in \mathbb{R}$.

Nous notons Λ la matrice de variance covariance des erreurs $U^{(i)} = \log \left(H^{(i)}(\Psi)\right)$, *i.e.*,

$$\Lambda = \text{var} \left((U^{(1)}, \dots, U^{(d)})' \right). \quad (3.3.5)$$

L'erreur $U^{(\varpi)}$ associée à la proxy $H_n^{(\varpi)}$ est égale à $\sum_{i=1}^d \varpi_i U^{(i)}$, dont la variance est $\lambda_{\varpi}^2 = \varpi' \Lambda \varpi$ qui atteint son minimum global pour

$$\varpi^* = \frac{\Lambda^{-1} \iota}{\iota' \Lambda^{-1} \iota}, \quad \iota = (1, \dots, 1)', \quad (3.3.6)$$

La variance optimale est $\lambda_{\varpi^*}^2 = \frac{1}{\iota' \Lambda^{-1} \iota}$. Cependant, cette solution n'est pas applicable en pratique, Il n'est pas possible d'estimer Λ dès que les proxies $H^{(i)}(\Psi_n)$ sont non observables. Afin de surmonter cette difficulté, il suffit d'estimer la matrice de variance covariance du logarithme de la proxy géométrique $H_n^{(\varpi)}$ *i.e.*,

$$\Lambda_n = \text{var} \left((\log(H_n^{(1)}), \dots, (\log H_n^{(d)}))' \right) = \text{var} (\log(\sigma_n)) \iota \iota' + \Lambda, \quad (3.3.7)$$

pour obtenir le vecteur des coefficients optimaux ϖ^* , comme le montre le théorème suivant.

Théorème 3.3.2. (de Vilder et Visser, 2008)

Soit le modèle d'échelle donné en (3.2.2). Supposons que $(\lambda^{(i)})^2 = \text{var}(\log H^{(i)}(\Psi)) < \infty$, pour $i = 1, \dots, d$ et $\text{var}(\log(\sigma_n)) < \infty$. Soient Λ et Λ_n les matrices de variance covariance données en (3.3.5) et (3.3.7) respectivement. Le vecteur des coefficients optimaux ϖ^* donné en (3.3.6) ne dépend pas du processus (σ_n) et peut être exprimé sous la forme $\varpi^* = \frac{\Lambda_n^{-1}\iota}{\iota'\Lambda_n^{-1}\iota}$. La variance de logarithme de la proxy $H_n^{(\varpi^*)}$ est donnée par $\text{var}(\log(H_n^{(\varpi^*)})) = \text{var}(\log(\sigma_n)) + \lambda_{\varpi^*}^2$, où $\lambda_{\varpi^*}^2 = \frac{1}{\iota'\Lambda^{-1}\iota}$.

Démonstration (Théorème 3.3.2)

Le vecteur des coefficients optimaux ϖ^* ne dépend pas du modèle (σ_n) , d'après (3.3.4), il s'en suit :

$$\begin{aligned} \arg \min_{\varpi} \text{var}(\log(H_n^{(\varpi)})) &= \arg \min_{\varpi} \lambda_{\varpi}^2, \\ &= \arg \min_{\varpi} \text{var}(\log H^{(\varpi)}(\Psi_n)). \end{aligned}$$

Alors,

$$\arg \min_{\varpi} \varpi' \Lambda_n \varpi = \arg \min_{\varpi} \varpi' \Lambda \varpi. \quad (3.3.8)$$

En appliquant le multiplicateur de Lagrange l sous contrainte $\sum_{i=1}^d \varpi_i = 1$, on obtient l'équation lagrangienne suivante $\varpi' \Lambda_n \varpi + l(1 - \varpi' \iota)$. Par suite, en dérivant par rapport à ϖ , on aura : $2\Lambda_n \varpi - l\iota = 0$, d'où l'on tire $\varpi = \frac{1}{2}\Lambda_n^{-1}l\iota$ et en remplaçant cette dernière dans la contrainte $\iota'\varpi = 1$, on obtient $l = 2/\iota'\Lambda_n^{-1}\iota$. Alors, $\varpi = \frac{\Lambda_n^{-1}\iota}{\iota'\Lambda_n^{-1}\iota}$.

En utilisant (3.3.8), on obtient les égalités $\varpi^* = \frac{\Lambda_n^{-1}\iota}{\iota'\Lambda_n^{-1}\iota} = \frac{\Lambda^{-1}\iota}{\iota'\Lambda^{-1}\iota}$, ce qui nous permet d'écrire $\text{var}(\log(H_n^{\varpi^*})) = \text{var}(\log(\sigma_n)) + \lambda_{\varpi^*}^2$.

3.4 Variance réalisée

En finance et en économétrie financière, l'analyse des variations du prix d'un actif et de son instabilité est fondamentale. Il est bien connu que les rendements journaliers des actifs financiers sont difficiles, voire impossible à prédire, bien que leur volatilité semble relativement facile à prédire. De ce fait, l'estimation de la volatilité joue un rôle clé dans l'évaluation des produits dérivés, dans l'allocation des actifs ou dans la gestion du risque. Comment faire de l'inférence sur la volatilité à partir des rendements, constitue un thème de recherche central en économétrie financière.

Plusieurs mesures de volatilité ont été proposées dans la littérature et un nombre assez important de tentatives ont été faites, afin d'améliorer l'estimateur classique de volatilité, à savoir l'écart-type empirique des rendements, en passant des mesures basées sur des rendements de basse fréquence, vers celles basées sur des rendements de haute fréquence.

Parmi ces estimateurs, nous citons ceux développés par Parkinson (1980), Garman et Klass (1980), Rogers et Satchell (1991), Yang et Zhang (2000) et Alizahdeh, Brandt et Diebold (2002) (V_P , V_{GK} , V_{RS} , V_{YZ} et V_{ABD} respectivement)². Ces auteurs exploitent l'information disponible sur l'étendue de la journée ouvrable.

Théoriquement, sous l'hypothèse de continuité du processus sous-jacent des rendements, supposé suivre un mouvement brownien géométrique sans dérive, l'estimateur V_P est 5 fois plus efficace, V_{GK} , et V_{YZ} sont 7 fois plus efficaces, comparativement à l'estimateur classique, au sens du critère du rapport des variances. La réalité empirique (la non nullité de la dérive, la discontinuité du processus sous jacent et la non constance de la volatilité) révèle des insuffisances de ces estimateurs.

Dans le sillage de Merton (1980), compte tenu du caractère aléatoire de la volatilité, les chercheurs se sont intéressés, non plus directement à la volatilité instantanée, mais à ce que l'on désigne sous le nom de variance *intégrée* ou encore de variation quadratique et qui représente la somme (l'intégrale en raisonnant en temps continu) des carrés de la volatilité instantanée. La variance intégrée correspond ainsi, par exemple, à la variation d'un prix au cours d'une journée, obtenue grâce à des informations collectées à une fréquence intra-journalière (*i.e.*, à haute fréquence, 5, 10 ou 30 minutes par exemple). Plusieurs approches

²Pour plus de détails, voir l'annexe

ont été développées pour déterminer un estimateur fiable et robuste de la variance intégrée.

Les approches paramétriques sont fondées sur l'utilisation de modèles décrivant la dynamique de la variance des rendements. Le modèle *GARCH*, qui consiste à estimer conjointement des équations du rendement et de la volatilité conditionnelle d'un actif ainsi que le modèle exponentiel *ARCH* de Nelson (1992), qui permet un passage du temps continu au temps discret, sont des candidats pour la mesure de la variance intégrée. Ces modèles connaissent depuis plusieurs années déjà, tout particulièrement dans les milieux académiques, un développement extraordinaire (Avouyi-Dauvis et Jardet (2006)). Néanmoins, la plupart de ces modèles échouent à capter les faits stylisés observés dans les séries financières. Ce constat a été fait par Bollerslev (1987), Malmsten et Teräsvirta (2004), et Carnero *et al.* (2004). D'où, la nécessité de rechercher un cadre plus adéquat pour l'estimation et la prédiction de la variance conditionnelle des rendements des actifs financiers. En exploitant la disponibilité des données de haute fréquence, un nouveau concept est introduit et, il a permis de résoudre le problème. C'est celui de la variance réalisée.

La variance réalisée du prix d'un actif à une date n (un jour par exemple) est définie comme la somme des carrés de ses rendements observés à un pas de temps donné. La variance réalisée relève d'une approche non paramétrique. C'est un estimateur usuel de la variance intégrée, introduit par Merton en 1980 et généralisé dans plusieurs travaux (Andersen et Bollerslev (1998), Andersen, Bollerslev, Diebold et Labys, (2001) et (2003) et Barndorff-Nielsen et Shephard, (2002a), (2002b), (2005), et Meddahi (2002), ou encore Comte et Renault, (2001) qui ont étudié les propriétés statistiques de la variance réalisé.

En théorie, plus le pas de temps est fin, plus la variance réalisée converge vers la variance intégrée. Malheureusement, les prix d'actifs financiers, plus particulièrement à haute fréquence, subissent une très grande variété de frictions que l'on peut interpréter comme des imperfections du processus de transaction ou des effets dits de microstructure (différences entre les volumes de transaction, information contenue dans les variations de prix, composantes stratégiques des flux d'ordre, effets de contrôle des stocks, discrétisation des données, etc.). La présence des effets de microstructure fait que le prix efficient n'est pas observé.

En fait, le prix observé est la somme du prix efficient non observé plus une variable aléatoire qui représente les effets de microstructure. De ce fait, l'estimateur usuel de la

variance réalisée est biaisé et non convergent. De nombreux chercheurs se sont donc intéressés à trouver un estimateur convergent de la volatilité intégrée à partir des données entachées de bruit. (Aït-Sahalia *et al.* (2005), Bandi et Russell (2005, 2006), Zhang *et al.* (2005), Hansen et Lunde (2006), et d'autres auteurs).

3.4.1 Modèle en temps discret

Dans un premier temps, nous présentons un modèle en temps discret pour motiver la notion de la variance réalisée. Pour ce faire, nous considérons le modèle suivant, à temps discret pour les rendements journaliers *close-to-close*

$$r_n = \sigma_n Z_n,$$

où $\{Z_n\}_{n=1}^N$, est une suite de variables aléatoires gaussiennes indépendantes et identiquement distribuées (*N.i.i.d.*) de moyenne nulle et de variance égale à l'unité *i.e.*, $Z_n \sim N.i.i.d. (0, 1)$.

de Vilder et Visser (2008), supposent que la volatilité σ_n est constante au cours de la journée ouvrable (en anglais *business ou trading day*). Néanmoins, d'autres auteurs, en particulier McAleer et Medeiros (2008) définissent la volatilité journalière comme la moyenne arithmétique des volatilités *intraday*.

Le jour ouvrable n est normalisé en un intervalle unité $[0, 1]$. Nous nous intéressons aux rendements intra-journaliers échantillonnés en haute fréquence (à un pas Δ par exemple, $0 < \Delta < 1$), autrement dit, nous considérons une subdivision de cet intervalle en $1/\Delta$ sous-intervalles de même longueur. Ainsi, le rendement intra-journalier (relativement à la valeur précédente) est défini comme suit : $r_{n,k} = R_n(k\Delta) - R_n((k-1)\Delta)$, $k = 1, \dots, \frac{1}{\Delta}$ où le processus $R_n(\cdot)$ est défini en (3.2.2). C'est-à-dire que

$$r_{n,k} = \sigma_n(\Psi_n(k\Delta) - \Psi_n((k-1)\Delta))$$

Par exemple $\Psi_n(\vartheta) = \vartheta Z_n$. Dans ce cas,

$$r_{n,k} = \sigma_n Z_{n,\Delta}^{(k)},$$

où $Z_{n,\Delta}^{(k)}$ sont *N.i.i.d.* $(0, \Delta^2)$, et le rendement *close-to-close* est $r_n = \sum_{k=1}^{1/\Delta} r_{n,k}$, sous l'hypothèse que $R_n(0) = 0$.

Il s'ensuit que $\mathbb{E}(r_n^2 | \mathcal{G}_{n-1}) = \sigma_n^2$ et, comme on peut le voir à l'aide d'un calcul simple, $\text{var}(r_n^2 | \mathcal{G}_{n-1}) = 2\sigma_n^4$.

La variance réalisée définie comme la somme du carré des rendements intra-journaliers échantillonné en haute fréquence (à un pas Δ par exemple, $0 < \Delta < 1$), est donnée par

$$RV_n^{(2)} = \sum_{k=1}^{1/\Delta} r_{n,k}^2. \quad (3.4.1)$$

Le carré des rendements journaliers *close-to-close* peut s'écrire

$$r_n^2 = \left(\sum_{k=1}^{1/\Delta} r_{n,k} \right)^2 = \sum_{k=1}^{1/\Delta} r_{n,k}^2 + \sum_{k \neq j}^{1/\Delta} r_{n,k} r_{n,j},$$

tel que

$$\mathbb{E}(r_n^2 | \mathcal{G}_{n-1}) = \mathbb{E}(RV_n^{(2)} | \mathcal{G}_{n-1}) + \mathbb{E}\left(\sum_{k \neq j}^{1/\Delta} r_{n,k} r_{n,j} | \mathcal{G}_{n-1}\right).$$

Si les rendements intra-journaliers sont non corrélés, on obtient

$$\mathbb{E}(r_n^2 | \mathcal{G}_{n-1}) = \mathbb{E}(RV_n^{(2)} | \mathcal{G}_{n-1}) = \sigma_n^2.$$

Cependant, il est facile de montrer que

$$\text{var}(RV_n^{(2)} | \mathcal{G}_{n-1}) = (2\Delta) \sigma_n^4 < 2\sigma_n^4 = \text{var}(r_n^2 | \mathcal{G}_{n-1}).$$

En d'autres termes le carré de la volatilité peut être estimée de façon plus précise, en additionnant le carré des rendements intra-journaliers qu'en additionnant les rendements journaliers *close-to-close* au carré.

3.4.2 Propriétés statistiques de la variance réalisée en absence des effets de microstructure

Les résultats décrivant les propriétés statistiques de la variance réalisée sont établis dans le cadre d'un modèle en temps continu pour le processus des prix (en logarithme) en se basant

sur la théorie de la variation quadratique. Dans un premier temps ces propriétés statistiques sont décrites en absence des effets de microstructure. Ils seront pris en compte par la suite.

Modèle en temps continu en absence de bruit de microstructure

La théorie de la volatilité réalisée est étroitement liée à la disponibilité des données de haute fréquence. Ainsi, il est naturel de considérer ce problème de mesure de volatilité dans une classe de modèles de volatilité stochastique, à temps continu.

Supposons que les prix (en logarithme) d'un actif donné suivent le processus de diffusion en temps continu³ suivant

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t, \quad 0 \leq t \leq T$$

Les processus stochastiques $\mu(t)$ et $\sigma(t)$ (moyenne et volatilité conditionnelle instantanée respectivement) sont prédictibles, où $\mu(t)$ est à variations finies, tandis que $\sigma(t)$ est strictement positif et de carré intégrable. *i.e.*, $\mathbb{E} \left(\int_0^t \sigma_s^2 ds \right) < \infty$. La variation quadratique du processus des prix logarithmiques sur l'intervalle unité $[0, 1]$ est donnée par

$$\langle \log(S) \rangle_n = p \lim_{\Delta \rightarrow 0} \sum_{k=1}^{1/\Delta} (\log(S_{n,k}) - \log(S_{n,k-1}))^2. \quad (3.4.2)$$

Quand le pas d'échantillonnage Δ tend vers 0, on obtient

$$\langle \log(S) \rangle_n = \int_0^1 \sigma^2(s) ds. \quad (3.4.3)$$

Cette dernière quantité est connue sous le nom de *variance intégrée* (IV_n). Les équations (3.4.2), (3.4.3) montrent que, dans ce contexte⁴, la variance intégrée et la variation quadratique coïncident.

³Pour éliminer la présence des opportunités d'arbitrage, le processus S_t doit constituer une semimartingale.

⁴Le processus des prix logarithmiques est une semimartingale

Distribution asymptotique de la variance réalisée

La variance réalisée $RV_n^{(2)}$ calculée en additionnant les carrés des rendements à haute fréquence, trouve sa justification théorique dans la théorie de la variation quadratique. Notons $r_{n,k}$ le $k^{\text{ème}}$ rendement intra-journalier de la $n^{\text{ème}}$ journée ouvrable normalisée à l'intervalle $[0, 1]$ tel que :

$$r_{n,k} = \log(S_{n,k}) - \log(S_{n,k-1}), \quad k = 1, \dots, \frac{1}{\Delta},$$

la variance réalisée est définie comme suit

$$RV_n^{(2)} = \sum_{k=1}^{1/\Delta} r_{n,k}^2. \quad (3.4.4)$$

Comme nous l'avons déjà signalé, la théorie de la variation quadratique des processus de semimartingales assure la convergence uniforme en probabilité de la variance réalisée vers la variance intégrée, quand la fréquence d'échantillonnage $1/\Delta$ tend vers l'infini. Ainsi, sous l'hypothèse d'absence d'effets de microstructure, Andersen et *al* (2003) ont montré que la variance réalisée $RV_n^{(2)}$, donnée en (3.4.1) est un estimateur consistant de la variance intégrée IV_n , tel que

$$RV_n^{(2)} \xrightarrow{P} \int_0^1 \sigma_s^2 ds, \quad 1/\Delta \rightarrow \infty.$$

En s'inspirant des résultats de Jacod et Protter (1998), Barndorff-Nielsen et Shephard (2002a, b, 2004, 2005) ont obtenu la distribution asymptotique suivante de la variance réalisée

$$\frac{\sqrt{1/\Delta}}{\sqrt{2IQ_t}} \left(RV_n^{(2)} - \int_0^1 \sigma_s^2 ds \right) \xrightarrow{\mathcal{L}} N(0, 1),$$

où IQ_t est la quarticité intégrée définie par

$$IQ_t = \int_0^1 \sigma_s^4 ds,$$

où $\xrightarrow{\mathcal{L}}$ désigne la convergence en loi.

Comme la quarticité intégrée est inconnue, le calcul de la distribution asymptotique est infaisable. Barndorff-Nielsen et Shephard (2002) ont montré que la quarticité intégrée est estimée de façon consistante par la quarticité réalisée RQ_t , définie par

$$RQ_t = \frac{1/\Delta}{3} \sum_{k=1}^{1/\Delta} r_{n,k}^4,$$

et par suite la distribution asymptotique de la variance réalisée $RV_n^{(2)}$, peut être approximée par

$$\frac{\sqrt{1/\Delta}}{\sqrt{\frac{2}{3}RQ_t}} (RV_n^{(2)} - IV_n) \xrightarrow{\mathcal{L}} N(0, 1),$$

quand $\Delta \rightarrow 0$.

Barndorff-Nielsen et Shephard (2005b), Meddahi (2002), Gonçalves and Meddahi (2005), and Nielsen and Frederiksen (2006) ont étudié les propriétés à échantillon fini du comportement asymptotique de la variance réalisée et ils ont abouti au résultat suivant

$$\frac{\sqrt{1/\Delta}}{\sqrt{\frac{2}{3} \frac{2RQ_t}{(RV_n^{(2)})^{1/2}}}} (\log(RV_n^{(2)}) - \log(IV_n)) \xrightarrow{\mathcal{L}} N(0, 1),$$

quand $\Delta \rightarrow 0$.

L'introduction de l'opérateur logarithmique sur la variance réalisée, lui réduit considérablement l'erreur standard.

Rendements intra-journaliers et leurs différents schémas d'échantillonnage

En pratique, les prix sont observés à des intervalles de temps discrets et irréguliers. Supposons que pour une journée ouvrable "n", nous divisons l'intervalle $[0, 1]$ en d sous intervalles tel que $0 = \tau_{1,d} < \tau_{2,d} < \dots < \tau_{d,d} = 1$. La longueur du $k^{\text{ème}}$ sous intervalle est donnée par $\Delta_{k,d} = \tau_{k+1,d} - \tau_{k,d}$ et la variance intégrée qui lui est associée est définie comme $IV_{k,d} = \int_{\tau_{k,d}}^{\tau_{k+1,d}} \sigma^2(s) ds$, pour $k = 1, \dots, d$. Plusieurs plans d'échantillonnage ont été proposés dans la littérature, nous citons :

- (i) Le *calendar time sampling* (CTS). C'est le plan d'échantillonnage le plus utilisé, représentant le cas particulier où les instants $\tau_{k,d}$ sont équidistants, autrement dit, $\Delta = \frac{1}{d}$, pour chaque sous intervalle i . Par exemple, les prix peuvent être échantillonnés chaque 5 ou 15 minutes.
- (ii) Un autre plan d'échantillonnage est le plan *transaction time sampling* (TrTS), où les prix sont enregistrés pour chaque transaction ($\tau_{k,d}$ est l'instant d'une transaction).
- (iii) Le troisième schéma d'échantillonnage est connu sous le nom *tick time sampling* (TiTS) où $\tau_{i,d}$ correspond à l'instant de révision des quotations.
- (iv) Le dernier plan d'échantillonnage est celui de *business time sampling* (BTS). Les instants d'échantillonnage $\tau_{1,d}, \tau_{2,d}, \dots, \tau_{d,d}$ sont choisis de sorte que $\sigma_{k,d} = \frac{IV}{d}, \forall k = 1, \dots, d$.

À noter que ce dernier plan d'échantillonnage (BTS), les instants $\tau_{1,d}, \tau_{2,d}, \dots, \tau_{d,d}$ sont non observables, contrairement aux trois autres premiers.

3.4.3 Propriétés statistiques de la variance réalisée en présence des effets de microstructure

L'équation (3.4.4) énonce qu'au fur et à mesure que la fréquence d'échantillonnage $1/\Delta$ augmente, la variance réalisée approxime bien la variation quadratique (variance intégrée). Néanmoins, en pratique, ce résultat théorique rencontre deux limitations. Premièrement, la non disponibilité de quotation continue des prix même pour les actifs les plus liquides. Cette limitation introduit des erreurs de discrétisation lors de calcul de la variance réalisée. Deuxièmement, la présence des effets de microstructure tels que : discrétisation et arrondi des prix, les erreurs commises lors de l'enregistrement,...etc., induisent de fallacieuses autocorrélations dans les rendements intra-journaliers.

Dans cette sous-section, nous étudions les propriétés statistiques de la variance réalisée en présence des effets de microstructure, et les solutions proposées pour contourner ce problème.

Supposons que les prix (en logarithme) sont observés avec un bruit, alors

$$p_{n,k} = p'_{n,k} + \varepsilon_{n,k},$$

où $p'_{n,k}$ est le prix efficient (non observable) et $\varepsilon_{n,k}$ est le bruit de microstructure. Il s'ensuit que

$$r_{n,k} = r'_{n,k} + \varepsilon_{n,k} - \varepsilon_{n,k-1} = r'_{n,k} + v_{n,k},$$

où $r'_{n,k} = p'_{n,k} - p'_{n,k-1}$ est le rendement efficient. Il est clair que $r_{n,k}$ est un processus autocorrélé. Alors, $RV_n^{(2)}$ donnée en (3.4.4) définie comme étant la somme du carré des rendements intra-journalier, est un estimateur biaisé de la volatilité journalière, non observable. Ainsi,

$$RV_n^{(2)} = \sum_{k=1}^{1/\Delta} (r'_{n,k})^2 + 2 \sum_{i=1}^{1/\Delta} r'_{n,k} v_{n,k} + \sum_{i=1}^{1/\Delta} v_{n,k}^2,$$

où $1/\Delta$ est le nombre total des observations disponibles dans la $n^{\text{ème}}$ journée ouvrable.

Sous les hypothèse suivantes, Bandi et Russell (2005), ont montré que :

$$RV_n^{(2)} \xrightarrow{p.s} \infty, \text{ quand } 1/\Delta \rightarrow \infty.$$

Hypothèse 1

- (a) Le bruit de microstructure $\varepsilon_{n,k}$, est stationnaire au second ordre, de moyenne nulle.
- (b) La variance de $v_{n,k} = \varepsilon_{n,k} - \varepsilon_{n,k-1}$, est d'ordre $O(1)$.

À leur tour, Zhang et al (2005), considèrent l'hypothèse suivante :

Hypothèse 2

- (a) Le bruit de microstructure $\varepsilon_{n,k}$, est *i.i.d.*, de moyenne nulle.
- (b) Le bruit de microstructure $\varepsilon_{n,k}$ est indépendant du processus des prix.
- (c) La variance de $v_{n,k} = \varepsilon_{n,k} - \varepsilon_{n,k-1}$, est d'ordre $O(1)$.

Sous cette dernière hypothèse, ils ont montré que :

$$(1/\Delta)^{-\frac{1}{2}} [RV_n^{(2)} - IV_n - 2(1/\Delta) \mathbb{E}(\varepsilon_{n,k}^2)] \xrightarrow{\mathcal{L}} 2 [E(\varepsilon_{n,k}^4)]^{\frac{1}{2}} N(0, 1).$$

En effet, nous avons

$$RV_n^{(2)} \stackrel{L}{\approx} IV_n + \underbrace{2(1/\Delta) E(\varepsilon_{n,k}^2)}_{\text{biais dû au bruit}} + \underbrace{\left[\underbrace{4(1/\Delta) \mathbb{E}(\varepsilon_{n,k}^4)}_{\text{dû au bruit}} + \underbrace{\frac{2}{(1/\Delta)} \int_0^1 \sigma_t^4 dt}_{\text{dû à la discrétisation}} \right]}_{\text{variance totale}}^{\frac{1}{2}} N(0, 1),$$

où $\stackrel{L}{\approx}$ signifie la convergence en loi, à une constante près.

Zhang (2006), Ait Sahalia et al (2006) ont considéré le cas où le processus $\varepsilon_{n,k}$ ne constitue pas une séquence de v.a *i.i.d.*, comme le montre l'hypothèse suivante.

Hypothèse 3

- (a) Le processus, $\varepsilon_{t,i}$ est stationnaire, de moyenne nulle. De plus, $\mathbb{E}(\varepsilon_{n,k})^{4+\kappa} < \infty$, pour $\kappa > 0$.
- (b) Le bruit de microstructure $\varepsilon_{n,k}$ est indépendant du processus des prix.
- (c) La variance de $v_{n,k} = \varepsilon_{n,k} - \varepsilon_{n,ik-1}$, est d'ordre $O(1)$.

Sous cette hypothèse, ces trois auteurs, ont montré que

$$RV_n^{(2)} \stackrel{L}{\approx} IV_n + \underbrace{2(1/\Delta) \mathbb{E}(\varepsilon_{n,k}^2)}_{\text{biais dû au bruit}} + \underbrace{\left[\underbrace{4(1/\Delta) \Omega}_{\text{dû au bruit}} + \underbrace{\frac{2}{(1/\Delta)} \int_0^1 \sigma_t^4 dt}_{\text{dû à la discrétisation}} \right]}_{\text{variance totale}}^{\frac{1}{2}} N(0, 1),$$

où $\Omega = \text{var}(v_{n,1})^2 + 2 \sum_{k=1}^{\infty} \text{cov}(v_{n,1}^2, v_{n,k+1}^2)$.

Il est clair que RV_n diverge, quand $(1/\Delta)$ augmente. Bandi et Russell (2005), Zhang et al (2005), ont montré que la variance réalisée multipliée par $(2(1/\Delta))^{-1}$ est un estimateur consistant de la variance de bruit de microstructure, autrement dit,

$$\frac{1}{2(1/\Delta)} RV_n^{(2)} \xrightarrow{P} \text{var}(\varepsilon_{n,k}).$$

Pour éliminer le biais dû aux effets de microstructure, plusieurs solutions ont été proposées.

- 1) Andersen et al (2000, 2001, 2003), ont proposé d'échantillonner les rendements intrajournaliers à une fréquence basse (modérée). par exemple 5 ou 15 minutes.
- 2) Une méthode alternative est de filtrer les rendements intra-journaliers, soit selon un modèle autorégressif (*AR*) (Bollen et Inder (2002)), ou selon un modèle moyenne mobile (*MA*) (Ebens (1999), Maheu et McCurdy (2002), et Andersen et al (2001)).

3.4.4 Quelques propriétés des volatilités réalisées et de leur logarithme de l'indice S&P 500 *Futures*

Les propriétés de la distribution des volatilités réalisées et leurs logarithmes

Tout au long de reste de ce mémoire, nous utilisons les notations $RV5_n^{(2)}$, $RV5_n = \sqrt{RV5_n^{(2)}}$ et $\log(RV5_n)$ pour désigner la variance réalisée, la volatilité réalisée et le logarithme de la volatilité réalisée respectivement dans le cas "applications" et $RV_n^{2(m)}$, $RV_n^{(m)} = \sqrt{RV_n^{2(m)}}$ et $\log(RV_n^{(m)})$ dans le cas "simulations".

Dans ce qui suit, nous donnons quelques statistiques descriptives de la distribution empirique des volatilités réalisées ainsi que celle des volatilités réalisées logarithmiques de l'indice S&P 500 *futures*⁵.

	$RV5_n^{(2)}$	$RV5_n$	$\log(RV5_n)$
moyenne	4.30×10^{-5}	0.0061	-5.16
écart-type	4.35×10^{-5}	0.0024	0.33
skewness	4.45	2.10	0.67
kurtosis	34.26	9.97	3.8

Table 3.4.1 : Les propriétés de la distribution des séries $RV5_n^{(2)}$, $RV5_n$ et $\log(RV5_n)$

D'après les résultats de tableau (3.4.1), il est clair que la distribution des séries $RV5_n^{(2)}$ et $RV5_n$ sont asymétriques et leptokurtiques. À noter que le fait de transformer les variances

⁵Pour plus de détails sur la construction des $RV5_n$, voir la section (4) ci-dessous.

réalisées $RV5_n^{(2)}$ en volatilités réalisées $RV5_n$ a baissé la valeur de skewness de 4.45 à 2.10 et celle de kurtosis de 34.26 pour atteindre la valeur 9.97.

De son côté, la distribution de la série $\log(RV5_n)$ est presque symétrique et approximativement gaussienne. La valeur de skewness et de kurtosis sont 0.67 et 3.8 respectivement. Néanmoins, ces moments empiriques sont significativement différents de 0 et 3 (le test d'asymétrie $\left(\left| \left((996/6)^{1/2} 0.67 \right) \right| = |8.632| > 1.96 \right)$ et d'aplatissement normal $\left(\left| (996/24)^{1/2} (3.8 - 3) \right| = |5.15| > 1.96 \right)$, au seuil de 5%) d'une distribution gaussienne standard.

Les valeurs de skewness et de la kurtosis de la série $\log(RV5_n)$ que nous avons obtenues peuvent être comparées à celles obtenues par Ebens (1999) sur l'indice *DJIA* (sur la période 1993 – 1998), qui sont égales à 0.75 et 3.78 (respectivement) et à celles obtenues par Areal et Taylor (2000), 0.44 et 3.71 (respectivement) sur l'indice FTSE 100 *futures*. Sur le marché de taux de change Andersen et *al* (2000) ont obtenu les valeurs 0.35 et 3.27 (respectivement) pour le deutsche mark allemand / au dollar américain et les valeurs 0.28 et 3.47 (respectivement) pour yen japonais / au dollar américain.

En plus du cas pratique, nous avons procédé par simulations, où la simulation des volatilités réalisées $RV_n^{(m)}$ ($n = 5000, m = 81$) est faite en suivant les mêmes étapes décrites dans la section (4), chapitre 4.

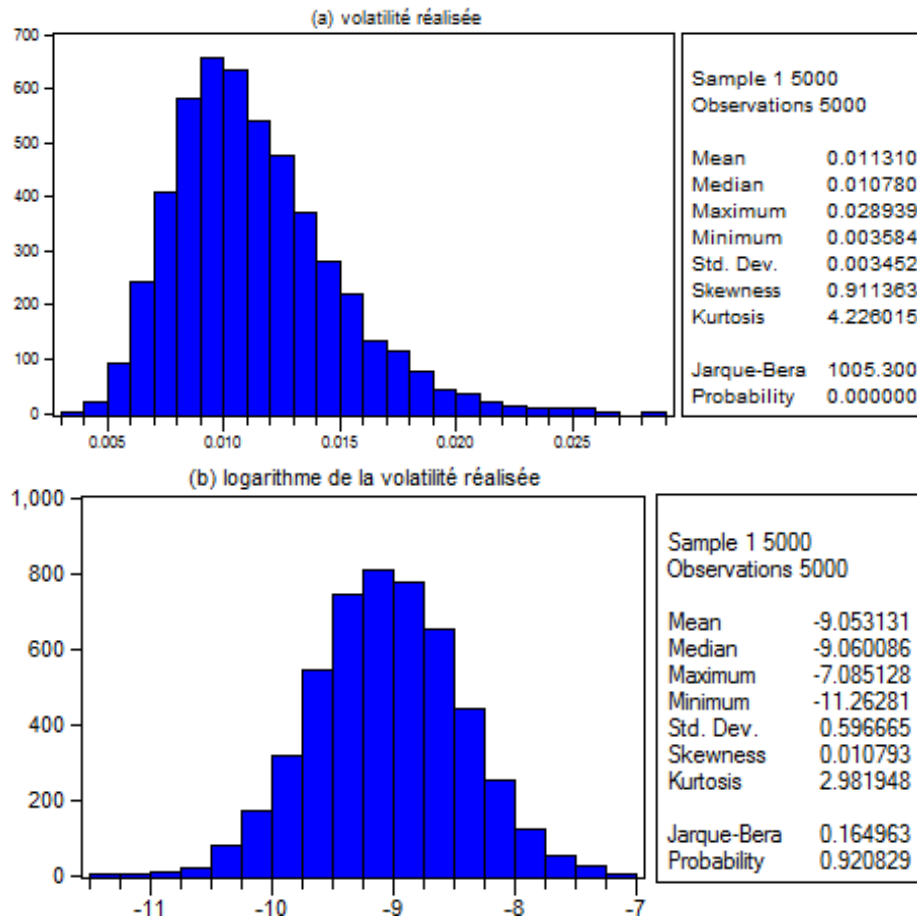


Figure 3.4.1 : Histogrammes et statistiques descriptives de la distribution des séries (a) $RV_n^{(m)}$ et (b) $\log\left(RV_n^{(m)}\right)$ simulées

Du graphe de l'histogramme de la volatilité réalisée $RV_n^{(m)}$ ⁶, nous constatons que cette dernière a une distribution leptokurtique (queues épaisses) et asymétriques. Le fait de faire une transformation, en prenant le logarithme de la volatilité réalisée, conduit à une réduction considérable dans la valeur des coefficients d'aplatissement et d'asymétrie. En effet, le test de Jacque-Bera accepte l'hypothèse nulle de normalité de la série $\log\left(RV_n^{(m)}\right)$.

Persistance de volatilité

Comme il a été signalé pour la première fois par Mandelbrot (1963) et Fama (1965), le regroupement des volatilités est l'une des régularités empiriques caractérisant les séries

⁶La simulation de la variance réalisée a été faite en suivant les mêmes étapes décrites dans la section (4), chapitre 4.

temporelles financières.

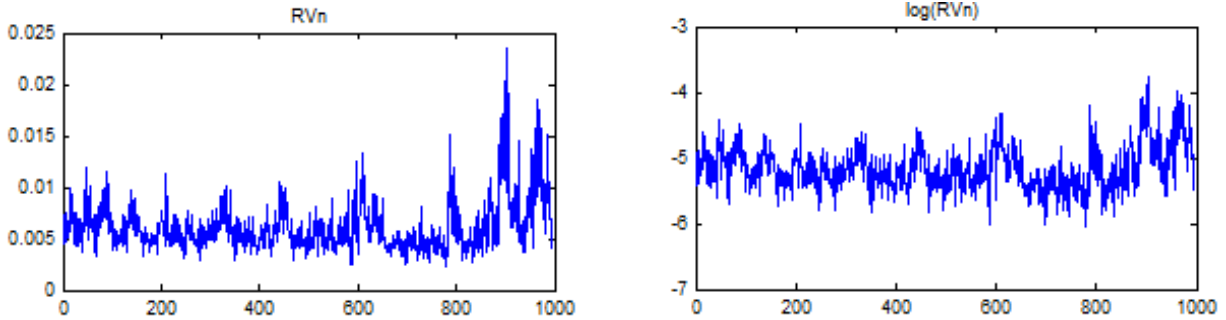


Figure 3.4.2 : Trajectoires des $RV5_n$ et $\log(RV5_n)$ de l'indice S&P 500 futures.

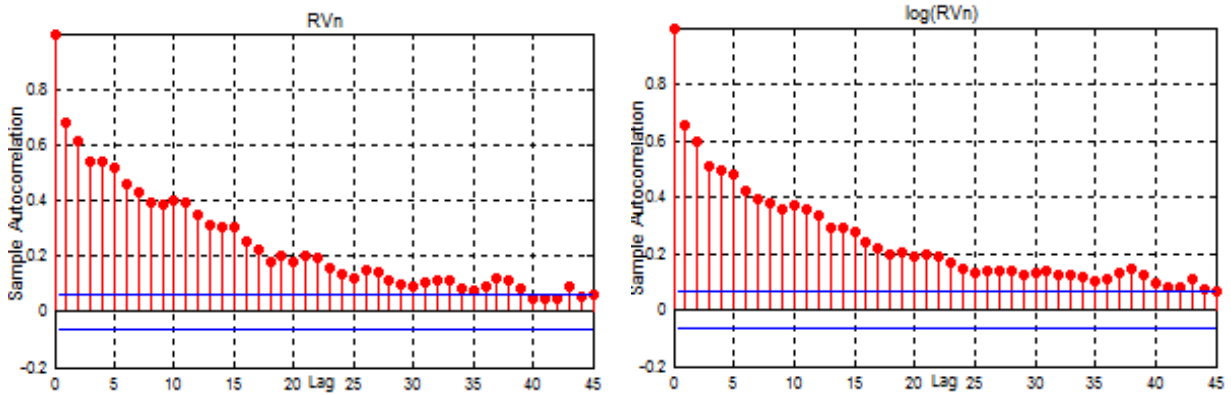


Figure 3.4.3 : Autocorrélations empiriques des séries $RV5_n$ et $\log(RV5_n)$ de l'indice S&P 500 futures.

Ce phénomène est illustré à travers les figures (3.4.2 *a* et *b*) et renforcé par les figures (3.4.3 *a* et *b*), où sont portés les autocorrélations empiriques des volatilités réalisés $RV5_n$ et leurs logarithmes $\log(RV5_n)$, qui décroissent lentement (décroissance hyperbolique).

Néanmoins ceci n'implique pas la non stationnarité des séries $RV5_n$ et $\log(RV5_n)$ de l'indice S&P 500, comme le montre le test d'*DF* suivant.

1) *Test de stationnarité (second ordre) des volatilités réalisées $RV5_n$.*

Considérons les trois modèles suivants :

$$\text{modèle 1 : } \Delta RV5_n = \phi RV5_{n-1} + \varepsilon_n,$$

$$\text{modèle 2 : } \Delta RV5_n = \phi RV5_{n-1} + c + \varepsilon_n,$$

$$\text{modèle 3 : } \Delta RV5_n = \phi RV5_{n-1} + c + \beta n + \varepsilon_n.$$

Nous commençons par estimer le modèle 3, incluant une constante et une tendance. Nous testons la présence d'une racine unitaire dans le processus en testant la nullité du paramètre ϕ . La valeur estimée de la statistique $DF = -13.97 < -3.41$ (lue à partir de la table de Dickey-Fuller pour un seuil de 5%). Alors, nous rejetons l'hypothèse nulle de racine unitaire ($\phi = 0$). Par suite, nous vérifions que le modèle 3 est le bon modèle. Cela revient à tester la nullité du paramètre β , en effectuant un test de Student, $t_{\hat{\beta}}$, où $\hat{\beta}$ désigne l'estimateur des MCO. La réalisation $|t_{\hat{\beta}}| = |2.34| > 1.96$, au seuil de 5%. Ainsi, nous rejetons l'hypothèse nulle de la non significativité de la tendance et par conséquent, le modèle 3 est le bon modèle. la série RV_n est TS (*Trend Stationary*), du fait de la présence de la tendance.

2) *Test de stationnarité (second ordre) du logarithme des volatilités réalisées* $\log(RV_n)$.

Considérons les trois modèles suivants :

$$\text{modèle 1 : } \Delta \log RV5_n = \phi \log RV5_{n-1} + \varepsilon_n,$$

$$\text{modèle 2 : } \Delta \log RV5_n = \phi \log RV5_{n-1} + c + \varepsilon_n,$$

$$\text{modèle 3 : } \Delta \log RV5_n = \phi \log RV5_{n-1} + c + \beta n + \varepsilon_n.$$

Nous commençons par estimer le modèle 3, incluant une constante et une tendance. Nous testons la présence d'une racine unitaire dans le processus en testant la nullité du paramètre ϕ . La valeur estimée de la statistique $DF = -14.41 < -3.41$ (lue à partir de la table de Dickey-Fuller pour un seuil de 5%). Alors, nous rejetons l'hypothèse nulle de racine unitaire ($\phi = 0$). Par suite, nous vérifions que le modèle 3 est le bon modèle. Cela revient à tester la nullité du paramètre β , en effectuant un test de Student, $t_{\hat{\beta}}$, où $\hat{\beta}$ désigne l'estimateur des MCO. La réalisation $|t_{\hat{\beta}}| = |1.30| < 1.96$, au seuil de 5%. Ainsi, nous acceptons l'hypothèse

nulle de la non significativité de la tendance et par conséquent, le modèle 3 est remis en cause.

Nous passons à l'estimation du modèle 2 et nous testons la présence d'une racine unitaire. $DF_{calculée} = -14.34 < DF_{tabulée} = -2.86$, nous rejetons l'hypothèse de la présence d'une racine unitaire dans le modèle 2. Il faut à nouveau évaluer la validité de notre diagnostic en vérifiant que le modèle 2 est le bon modèle. Pour cela, nous testons la nullité de la constante par le test statistique de Student. La valeur estimée de cette dernière $|t_{\hat{c}}| = |-14.31| > 1.96$, au seuil de 5%, où \hat{c} désigne l'estimateur des MCO. Donc, nous rejetons l'hypothèse nulle de la nullité la constante \hat{c} et le modèle 2 est le bon modèle.

Les volatilités réalisées $RV5_n$ et leurs logarithmes $\log(RV5_n)$ sont stationnaires au second ordre, mais caractérisées par une faible décroissance dans leurs fonctions d'autocorrélations. C'est une indication de présence de mémoire longue, autrement dit, de corrélations à longue portée.

3.5 Test de l'hypothèse de semimartingale pour l'indice S&P 500 *Futures*

Il est bien connu que la distribution des rendements logarithmiques des actifs financiers est asymétrique et exhibe des queues épaisses. Ainsi, ces caractéristiques sont en contradiction avec l'hypothèse classique que le modèle sous-jacent est un mouvement brownien géométrique.

Dans le cadre général d'un modèle de volatilité stochastique, le processus des prix S est décrit par une semimartingale continue donné par l'équation suivante :

$$\frac{dS(t)}{S(t)} = dA(t) + dM(t), \quad (3.5.1)$$

où A est une dérive (drift) et M est une martingale locale continue.

Dans cette section, nous testons l'hypothèse que le processus associé à l'indice S&P 500 *futures* est une semimartingale continue contre l'alternative que l'indice S&P 500 *futures* n'est pas une semimartingale continue. Ce test est effectué sous l'hypothèse que le processus

associé à l'indice SP 500 *futures* est à trajectoires continues.

Pour ce faire, nous considérons le résultat bien connu dans la théorie des martingales stipulant que chaque martingale locale continue M commençant à l'origine admet la représentation

$$M(t) = W(Q(t)), \quad (3.5.2)$$

où W est mouvement brownien standard et Q est la variation quadratique de M , *i.e.*, $Q = \langle M, M \rangle$ (voir théorème (1.4.1) chapitre 1). Autrement dit, il suffit de montrer que le processus des prix corrigés du terme drift A est un mouvement brownien subordonné à Q , *i.e.*, $M = W \circ Q$. Premièrement, nous devons estimer la dérive A , ensuite calculer les réalisations de la variation quadratique Q et enfin, tester le mouvement brownien.

3.5.1 Cadre théorique de l'approche utilisée

Selon les deux auteurs Peters et de Vilder (2006), le processus Q détermine la volatilité. Elle est obtenue moyennant la variation quadratique de la trajectoire du processus sous-jacent (M). Ainsi, ils définissent une nouvelle échelle du temps dit temps financier τ comme étant la valeur de Q en temps physique t . Ainsi, aux instants t_1, t_2, \dots, t_b , en temps physique correspondront les points de temps financiers $\tau_i = Q(t_i)$, $i = 1, 2, \dots, b$. Le temps financier τ est la valeur de Q qualifié de temps de changement à l'instant t .

La variation quadratique $Q = \langle M, M \rangle$ d'une martingale locale continue M , sur l'intervalle de temps $[0, c[$ peut être approximée par les processus Q_n , dit processus de sauts, commençant à l'origine $t_0 = 0$, avec les incréments $(M(t_i) - M(t_{i-1}))^2$, $\forall 0 = t_0 < t_1 < \dots < t_n = c$ qui s'ajoutent. Il en est de même si nous sommes en présence d'une semimartingale $X = A + M$ où A est un processus continu, à variations finies et M est une martingale locale continue tel que : $A(0) = M(0) = 0$. Autrement dit, $\langle X, X \rangle = \langle M, M \rangle$.

Considérons une grille de points des temps financiers équidistants. Si $M = W(Q)$ est un mouvement brownien subordonné à Q , alors les incréments définis par

$$\widetilde{W}_i = \frac{M(Q^{-1}(\tau_i)) - M(Q^{-1}(\tau_{i-1}))}{\sqrt{\Delta\tau}} = \frac{W(\tau_i) - W(\tau_{i-1})}{\sqrt{\tau_i - \tau_{i-1}}}, \quad (3.5.3)$$

sont des variables aléatoires *i.i.d.* $N(0, 1)$.

Où $Q^{-1}(\tau) = \inf_{t>0} \{Q(t) > \tau\}$.

Si, nous approximations la variation quadratique Q par la somme des incréments au carré du processus M , les incréments normalisés donnés en (3.5.3), resteront toujours indépendants et symétriques mais accompagnés d'une kurtosis qui peut être inférieure à 3 (mais jamais supérieure à 3), comme le montre la proposition suivante.

Proposition 3.5.1 (Peters et de Vilder, 2006)

Considérons b points dans le temps financier et supposons que le temps de changement Q est linéaire sur l'intervalle de temps $[Q^{-1}(\tau_i), Q^{-1}(\tau_{i-1})]$. Soient $t_j^{(i)}, j = 1, \dots, n_i$ des points équidistants dans cet intervalle avec $Q^{-1}(\tau_{i-1}) = t_0^{(i)} < \dots < t_{n_i}^{(i)} = Q^{-1}(\tau_i)$. La quantité W_i définie par $\frac{W(\tau_i) - W(\tau_{i-1})}{\sqrt{\sum_{j=1}^{n_i} (M(t_j^{(i)}) - M(t_{j-1}^{(i)}))^2}}$ a pour densité

$$f_{W_i}(w) = \frac{\Gamma(n_i/2)}{\sqrt{\pi n_i} \Gamma((n_i - 1)/2)} \left(1 - \frac{w^2}{n_i}\right)^{(n_i-3)/2} \mathbf{1}_{(-\sqrt{n_i}, \sqrt{n_i})}(w), w \in \mathbb{R}, \quad (3.5.4)$$

où $\Gamma(\cdot)$ désigne la fonction de Gamma.

Il est facile de montrer que le moment d'ordre $2k$ est à

$$\mathbb{E}(W_i^{2K}) = \left\{ n_i^k \left(k - \frac{1}{2}\right) \dots \left(\frac{1}{2}\right) \right\} / \left\{ \left(\frac{1}{2}n_i + k - 1\right) \dots \frac{1}{2}n_i \right\}, \quad (3.5.5)$$

En particulier, en remplaçant dans (3.5.5) k par 1 et 2, nous obtenons $\mathbb{E}(W_i^2) = 1$ et $\mathbb{E}(W_i^4) = 3n_i / (n_i + 2)$ respectivement.

Nous remarquons que si $n_i \rightarrow \infty$, la variable aléatoire W_i converge vers une variable aléatoire gaussienne standard. Néanmoins, en pratique, on est confronté à un nombre fini d'observations. Ce problème se manifeste clairement pour les intervalles de temps où la volatilité est très élevée, car dans ce cas, le temps financier court vite, ce qui fait que le nombre d'observations n_i disponibles dans l'intervalle de temps physique correspondant n'est pas grand.

Remarque :

- W_i est défini comme \widetilde{W}_i où $\Delta\tau$ est remplacé par la somme approximante

$$\sum_{j=1}^{n_i} \left(M \left(t_j^{(i)} \right) - M \left(t_{j-1}^{(i)} \right) \right)^2, \quad Q^{-1}(\tau_{i-1}) = t_0^{(i)} < \dots < t_{n_i}^{(i)} = Q^{-1}(\tau_i), \quad i = 1, 2, \dots, b.$$

3.5.2 Mise en oeuvre du test**Construction des incréments normalisés w_i**

Nous utilisons les notations en majuscule pour les variables aléatoires et en minuscules pour leurs réalisations correspondantes. Pour chaque jour ouvrable n que nous normalisons à $[0, 1]$, nous avons $\frac{1}{\Delta} + 1 = 82$ observations $s_{n,0}, s_{n,1}, \dots, s_{n,\frac{1}{\Delta}}$, et par suite $\frac{1}{\Delta}$ intervalles (81 intervalles de 5 minutes, pour les données *intraday* de haute fréquence dont nous disposons). Les incréments observés $r_{n,k}^*$ (rendements centrés) du processus M au $k^{\text{ème}}$ intervalle du jour n sont définis comme suit :

$$r_{n,k}^* = \log \left(\frac{s_{n,k}}{s_{n,k-1}} \right) - \Delta \bar{\mu},$$

où $\bar{\mu}$ désigne le rendement moyen calculé sur les 996 jours⁷, *i.e.*,

$$\bar{\mu} = \frac{1}{996} \sum_{n=1}^{996} \sum_{k=1}^{1/\Delta} \log \left(\frac{s_{n,k}}{s_{n,k-1}} \right).$$

Nous définissons

$$\begin{aligned} r_n^* &= r_{n,1}^* + \dots + r_{n,\frac{1}{\Delta}}^*, \\ rv5_n^{(2)} &= r_{n,1}^{*2} + \dots + r_{n,\frac{1}{\Delta}}^{*2}, \end{aligned}$$

comme étant l'accroissement journalier d'une réalisation du processus M et l'estimateur de l'accroissement journalier d'une réalisation du temps financier respectivement. Alors, la

⁷Pour la description des données voir le chapitre 4

trajectoire du processus M et l'estimateur \hat{q} du temps de changement Q associé à la trajectoire m du processus M , sur l'intervalle du temps $t \in [n - 1 + k\Delta, n - 1 + (k + 1)\Delta[$ sont donnés par

$$m(t) = r_1^* + \dots + r_{n-1}^* + r_{n,1}^* + \dots + r_{n,k}^*,$$

$$\hat{q}(t) = rv_1^{(2)} + \dots + rv_{n-1}^{(2)} + r_{n,1}^{*2} + \dots + r_{n,k}^{*2},$$

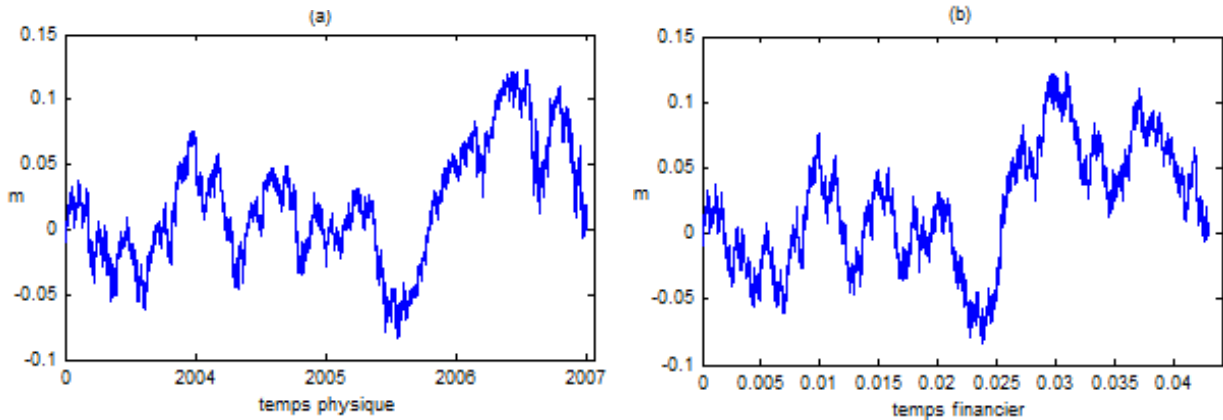


Figure 3.5.1 : (a) Rendements logarithmiques intra-journaliers centrés, en temps physique et (b) rendements logarithmiques intra-journaliers centrés dans le temps financier de l'indice SP 500.

D'après le tracé de la figure (3.5.1), il est clairement visible que la forme du tracé (a) est reproduite dans le tracé (b). Donc, les caractéristiques du processus associé à l'indice SP 500 sont conservées dans le temps financier autant que possible, sous une forme compressée.

Afin d'évaluer l'effet du terme drift $A(t)$ inconnu donnée en (3.5.1) sur l'estimateur \hat{q} , nous calculons la différence entre la variation quadratique totale de $m(t)$ et celle de $m(t) + \bar{\mu}t$. Cette différence est égale à $996 \times 81 \times (\bar{\mu}/81)^2 \simeq 2.95 \times 10^{-7}$ où $\bar{\mu} = -1.55 \times 10^{-4}$. Ainsi, l'effet de la fonction drift $A(t)$ est de l'ordre 10^{-7} . Donc, elle est d'une influence négligeable sur la variation quadratique.

Passons maintenant à la construction des incréments normalisés w_i . Les points de temps t_i où le processus M est évalué sont choisis selon la règle suivante (Peters et de Vilder (2006))

- Fixons un intervalle de temps d'une longueur $\Delta\tau$.
- Nous choisissons la suite de points t_1, \dots, t_b tel que $\hat{q}(t_i) - \hat{q}(t_{i-1}) \geq \Delta\tau$ et $\hat{q}(t'_i) - \hat{q}(t_{i-1}) < \Delta\tau$, pour tout $t'_i < t_i, i = 1, 2, \dots, b$.

Remarque :

Le choix de $\Delta\tau$ est très important. Une petite valeur de $\Delta\tau$ conduit à un biais dans l'observation des W_i donnés en (3.5.3), pendant les périodes caractérisées par des volatilités élevées, d'une part (comme nous l'avons déjà expliqué). D'autre part, une valeur élevée de $\Delta\tau$ diminue le nombre d'observations (*i.e.*, b), et par suite a un effet négatif sur la puissance du test utilisé (ici).

Les incréments normalisés w_i sont donnés par

$$w_i = \frac{m(t_i) - m(t_{i-1})}{\sqrt{\hat{\tau}_{i-1} - \hat{\tau}_i}} \quad (3.5.6)$$

où $\hat{\tau}_{i-1} - \hat{\tau}_i = \hat{q}(t_i) - \hat{q}(t_{i-1}), i = 1, 2, \dots, b$.

Nous avons essayé plusieurs valeurs de $\Delta\tau$ ($0.05ft, 0.06ft, 0.07ft, 0.08ft, 0.09ft, 0.1ft, 0.2ft, 0.3ft, 0.4ft$ et $0.5ft$), où $ft = 10^{-3}$ unités de temps financier. Tenant compte des deux contraintes discutées ci-dessus, nous sommes parvenues au choix $\Delta\tau = 0.00009$. Cela nous donne 463 points de temps financier.

Le test

Testons maintenant l'hypothèse que le processus associé à l'indice S&P 500 *futures* est une semimartingale continue contre l'alternative n'est pas une semimartingale continue, autrement dit, cela revient à effectuer le test suivant :

$$\begin{cases} H_0 : \text{le processus associé à l'indice S\&P 500 futures est une semimartingale continue} \\ H_1 : \text{le processus associé à S\&P 500 futures n'est pas une semimartingale continue} \end{cases}$$

Pour ce faire, il suffit de vérifier que les $w_i, i = 1, 2, \dots, b$, donnés en (3.5.6) sont des réalisations issues de variables aléatoires (v.a) *i.i.d.* gaussiennes standards.

Nous effectuons 5 tests. Les deux premiers pour la normalité, le troisième pour tester

l'absence des queues épaisses, et les deux derniers pour tester l'indépendance.

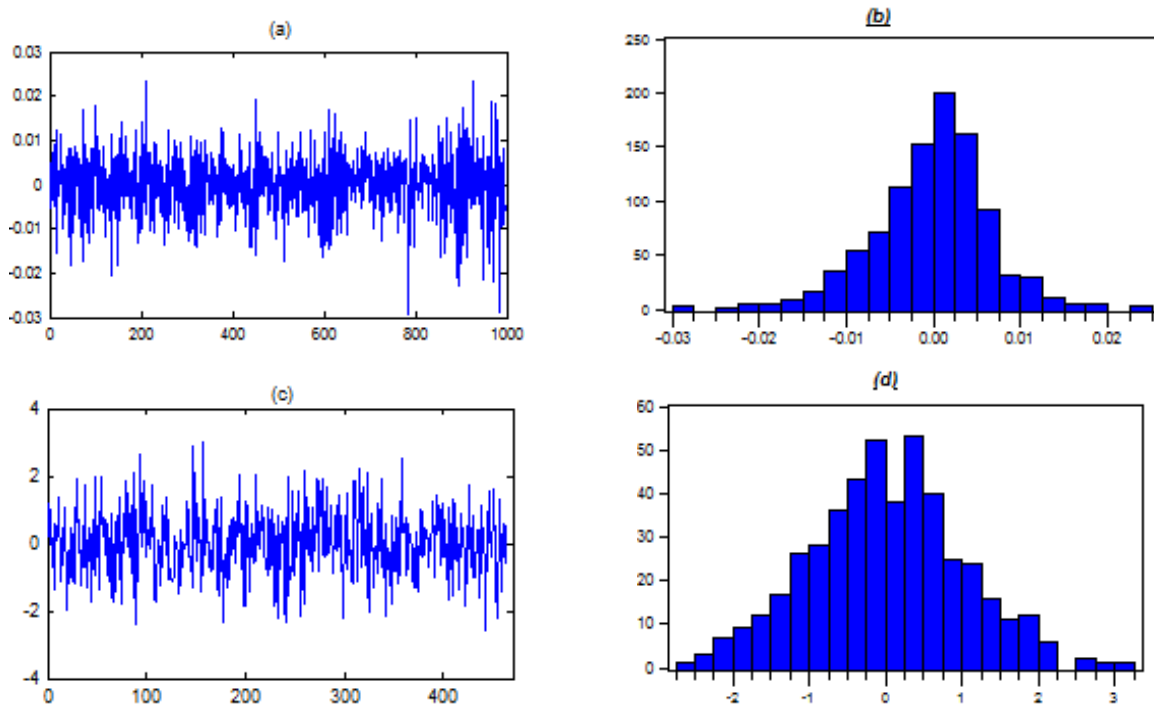


Figure 3.5.2 : Séries temporelles des rendements logarithmiques journaliers r_n^* et les incréments normalisés w_i ((a) et (c) respectivement) et leur histogramme correspondant ((b) et (d))

Test de Kolmogorov-Smirnov

Le premier test que nous utilisons est celui de Kolmogorov-Smirnov. Ce dernier est fondé sur la comparaison entre les fonctions de répartition empirique et théorique (G_X et F_X) respectivement, de la variable aléatoire X sous considération, dans notre cas, il s'agit d'une variable aléatoire gaussienne standard. La statistique du test est donnée par $KS = \max_x |G_X(x) - F_X(x)|$ (voir par, exemple, Shorack et Wellner (1986)). Les résultats de ce test sont présentés dans le tableau (3.5.1). Ce test est appliqué aussi sur les rendements logarithmiques $r_n = r_n^* + \bar{\mu}$. Comme on peut le voir, à partir des sorties de ce tableau, la normalité standard est acceptée pour la série des w_i , mais, elle ne l'est pas pour les rendements logarithmique journaliers r_n , au seuil de significativité $\alpha = 5\%$.

Test de Jarque-Bera

Les résultats de ce test se trouvent toujours dans le même tableau. Pour tenir compte de l'éloignement de la valeur espérée de la kurtosis $Kr = Kr^* < 3$, Peters et de Vilder

(2006), en se basant sur une analyse heuristique, proposent une version modifiée de la statistique de Jarque-Bera classique, à savoir $jb^* = (b/6) \left(\widehat{sw}^2 + \frac{1}{4} \left(\widehat{kr} - Kr^* \right)^2 \right)$, où $kr^* = b^{-1} \sum_{i=1}^b \frac{3n_i}{n_i+2} = 2.9328$.

Dans notre cas, $jb^* = 0.3962 < 5.99$, et par suite, nous acceptons la normalité des w_i . À noter que $jb^* = 0.3963 < jb = 0.5295$. Par contre la normalité est toujours rejetée pour les rendements r_n (au seuil $\alpha = 5\%$).

	ks	P-value	moyenne	écart-type	\widehat{sw}	\widehat{kr}	jb	P-value
r_n	0.49	0.00	-1.55×10^{-4}	0.0065	-0.370	4.394	103.546	0.00
w_i	0.0265	0.893	1.99×10^{-3}	0.99	0.0711	2.915	0.5295	0.767

Table. 3.5.1 : Test de normalité sur les deux séries r_n et w_i

Le test d'asymétrie $\left(\left| \left((b/6)^{1/2} \widehat{sw} \right) \right| = |0.624| < 1.96 \right)$ et d'aplatissement normal $\left(\left| (b/24)^{1/2} \left(\widehat{kr} - Kr^* \right) \right| = |-0.078| < 1.96 \right)$ confirment la normalité standard, au seuil $\alpha = 5\%$, où $b = 463$.

Test Binomial

Quand on a affaire à des données financières, les valeurs extrêmes sont d'une importance majeure. Pour vérifier la présence ou l'absence de queues épaisses dans la distribution des incréments normalisés w_i , nous comptons le nombre d'observations n_c dont la valeur dépasse $c = 2, 3, 3.5$, respectivement. Soit $p_c = P(|U| > c)$ où U est une variable aléatoire gaussienne standard. Pour b réalisations indépendantes de $|U|$, le nombre $N_c = \sum_{i=1}^b \mathbf{1}_{\{|U_i|>c\}}$, dépassant le niveau c , suit une loi binomiale de paramètres (b, p_c) de moyenne $\mu_c = bp_c$ et de variance $\sigma_c^2 = bp_c(1 - p_c)$. Les résultats du test sont donnés dans le tableau (3.5.2).

c	p_c	n_c	μ_c	σ_c	$P(N_c > n_c)$
2	0.0456	21	45.4176	6.5838	0.99
3	0.0027	1	2.6892	1.6376	0.50
3.5	4.6×10^{-4}	0	.4581	0.6767	0.077

Table. 3.5.2 : Test binomial

Les sorties de ce tableau valident l'hypothèse que les queues de la série w_i , se comportent comme celles d'une loi normale standard, au seuil $\alpha = 5\%$.

Test BDS

Ce test d'indépendance a été introduit dans la littérature financière par Brock, Dechert, Scheinkman et Lebaron (1987). La statistique BDS a été conçue pour tester l'hypothèse nulle que la série observée $\{x_1, \dots, x_N\}$, est une suite de variables aléatoires *i.i.d.* Elle mesure la signification statistique des intégrales de corrélations. En définissant une longueur e , l'intégrale de corrélation est la probabilité que deux points quelconques sont séparés d'une longueur e dans l'espace M_h de dimension h défini par :

$$M_1 : \underline{y}_j^{(1)} = x_j, M_2 : \underline{y}_j^{(2)} = (x_{j-1}, x_j)', \dots, M_h : \underline{y}_j^{(h)} = (x_{j-h+1}, \dots, x_j)'$$

Considérons $\underline{y}_j^{(h)}$ et $\underline{y}_{j'}^{(h)}$ deux vecteurs arbitraires tels que $|j - j'| \geq h$ et $x_k, x_{k'}$ deux observations arbitraires telles que $k \neq k'$. Si $(x_j)_{j=1}^N$ est une suite *i.i.d.*, alors,

$$P\left(\left|\underline{Y}_j^{(h)} - \underline{Y}_{j'}^{(h)}\right| < e\right) = (P(|X_k - X_{k'}| < e))^h,$$

$\underline{Y}_j^{(h)}, \underline{Y}_{j'}^{(h)}, X_k$ et $X_{k'}$ sont, respectivement, les vecteurs et variables aléatoires associés.

$|\cdot|$ désigne la norme $|\underline{Y}| = \sup_{i=1, \dots, h} |Y_i|$ et $e > 0$. Le test BDS utilise la différence $\hat{P}\left(\left|\underline{Y}_j^{(h)} - \underline{Y}_{j'}^{(h)}\right| < e\right) - \left(\hat{P}(|X_k - X_{k'}| < e)\right)^h$. Les estimateurs $\hat{P}\left(\left|\underline{Y}_j^{(h)} - \underline{Y}_{j'}^{(h)}\right| < e\right)$ et $\hat{P}(|X_k - X_{k'}| < e)$ sont obtenus en comptant le nombre de vecteurs $\underline{y}_j^{(h)}$ et le nombre d'observations x_j distants de moins de e et en prenant la moyenne arithmétique, *i.i.e*

$$\hat{P}\left(\left|\underline{Y}_j^{(h)} - \underline{Y}_{j'}^{(h)}\right| < e\right) = \frac{1}{n(n-1)} \sum_{j \neq j'} \mathbb{1}_{\left[\left|\underline{Y}_j^{(h)} - \underline{Y}_{j'}^{(h)}\right| < e\right]}$$

Sous l'hypothèse nulle, cette différence est asymptotiquement distribuée suivant une loi normale (Brock *et al*, 1987).

c	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$
$e = .25$	3.79×10^{-5}	1.24×10^{-4}	-2.32×10^{-5}	-2.64×10^{-5}	-7.83×10^{-6}
p	0.904	0.5290	0.8030	0.3650	0.5780
$e = .5$	-1.367×10^{-3}	-9.82×10^{-4}	-3.54×10^{-4}	-7.12×10^{-5}	1.47×10^{-6}
p	0.3130	0.2530	0.4780	0.7790	0.8990
$e = .75$	-1.426×10^{-3}	-4.206×10^{-3}	-4.111×10^{-3}	-3.995×10^{-3}	-5.009×10^{-3}
p	0.6610	0.4270	0.5690	0.6370	0.5410
$e = 1$	-2.132×10^{-3}	-3.498×10^{-3}	-3.151×10^{-3}	-2.394×10^{-3}	-1.471×10^{-3}
p	0.5140	0.3430	0.3360	0.3600	0.4570

Table. 3.5.3 : Test BDS sur la série w_i pour les différentes valeurs de h et e .

Le test a été effectué pour les différentes valeurs de dimensions d’empilement et de distance (h, e) respectivement. Les lignes donnant la p -value montrent que l’hypothèse d’indépendance est toujours acceptée.

Test d’absence d’autocorrélation (partielle)

Ce dernier test peut être considéré comme une vérification de l’indépendance (étant donné que l’hypothèse de normalité a été acceptée) des observations de la série w_i . Si la série w_i est *i.i.d.*, on s’attend à ce que les coefficients d’autocorrélations soient significativement nuls.

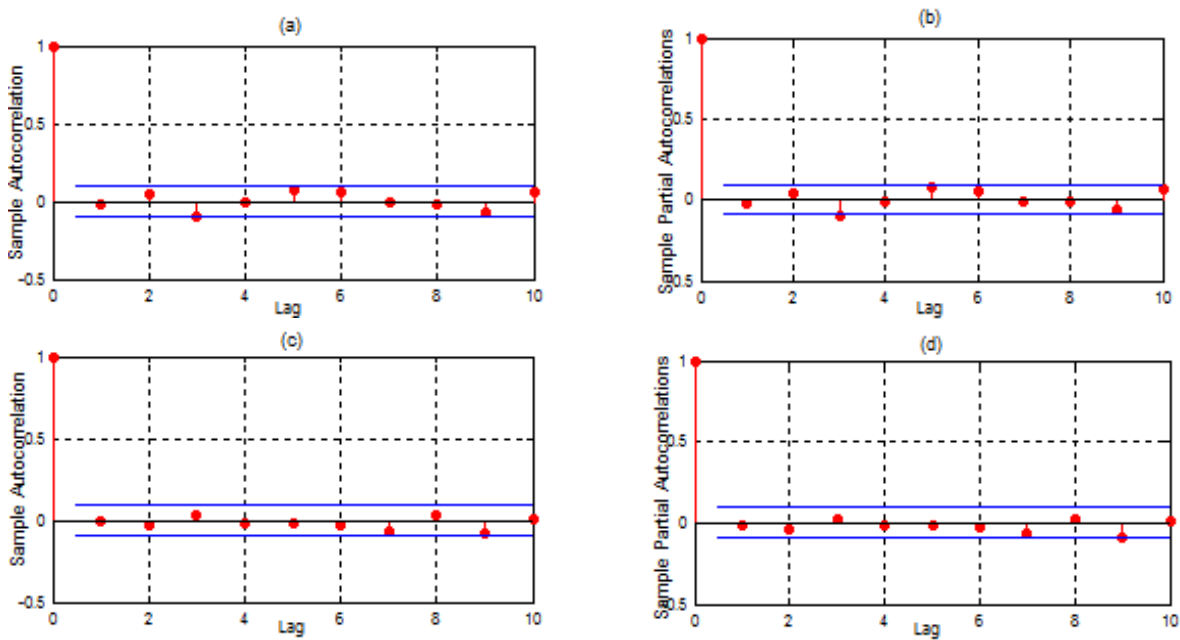


Figure 3.5.3 : Fonctions d’autocorrélations (empiriques) de la série w_i et de son carré w_i^2 (a et c) respectivement, ainsi que leurs fonctions d’autocorrélations partielles (b et d) respectivement

Les tracés de la figure (3.5.3) (fonction d'autocorrélation (partielle) empirique) pour les deux séries w_i (graphes du haut) et w_i^2 (graphes du bas) montrent que toutes les autocorrélations empiriques $\hat{\gamma}_l$ ($l = 1, \dots, 10$) de l'indice S&P 500 et de son carré se trouvent à l'intérieur de l'intervalle de confiance (les traits continus), au niveau $\alpha = 5\%$.

Au moyen du test de Ljung Box, nous testons quant à la significativité ou non des coefficients d'autocorrélation. Ce test est basé sur la statistique de Ljung Box : $R = b(b+1) \sum_{i=1}^l \hat{\gamma}_i / (b-k)$ où $\hat{\gamma}_i$ est l'estimateur de la $i^{\text{ème}}$ autocorrélation, l est le nombre de retards et b ($b = 463$) désigne la taille de l'échantillon. Sous l'hypothèse nulle d'absence d'autocorrélation, la statistique R suit asymptotiquement une loi de Khi-deux à l degrés de liberté.

le retard (l)	w_i			w_i^2		
	R	valeur critique	p -value	R	valeur critique	p -value
1	0.1799	2.7055	0.6715	0.0620	2.7055	0.8034
2	1.2057	4.6052	0.5472	0.7109	4.6052	0.7009
3	5.6975	6.2514	0.1273	1.0623	6.2514	0.7862
4	5.6985	7.7794	0.2228	1.2107	7.7794	0.8763
5	8.0416	9.2364	0.1540	1.2987	9.2364	0.9351
6	9.5370	10.6446	0.1456	1.6470	10.6446	0.9491
7	9.5400	12.0170	0.2162	3.5784	12.0170	0.8269
8	9.7126	13.3616	0.2858	4.0518	13.3616	0.8524
9	11.8944	14.6837	0.2193	7.6109	14.6837	0.5738
10	13.7623	15.9872	0.1841	7.6132	15.9872	0.6666

Table. 3.5.4 : Test d'absence d'autocorrélation pour les séries w_i et w_i^2 .

D'après les sorties de ce tableau, nous concluons que, pour les deux série w_i et w_i^2 , les coefficients d'autocorrélation sont non significatifs, car la p -value est supérieure à 10%. Donc, nous acceptons l'hypothèse d'absence d'autocorrélation, au seuil $\alpha = 10\%$.

Comme les incréments normalisés sont gaussiens et non autocorrélés, alors, ceci confirme leur indépendance.

En outre, des tests statistiques que nous avons présentés précédemment, nous décrivons quelques visualisations graphiques pour la normalité.

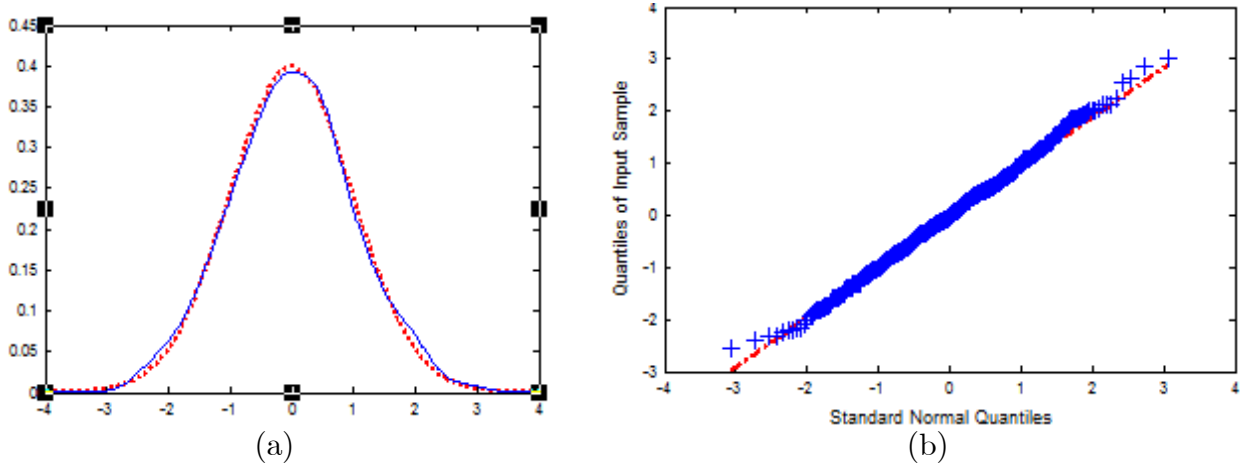


Figure 3.5.4 : (a) Estimateur de la densité des w_i (par la méthode du noyau) (trait plein) et densité normale centrée réduite (trait en pointillé) ainsi que (b) le QQ plot correspondant.

Nous constatons que les deux tracés de la figure (3.5.4a) ont un grand degré de ressemblance (la presque parfaite superposition)

À son tour, la figure (3.5.4b) présente le QQ plot qui est caractérisé par une quasitotalité de localisation des quantiles empiriques-quantiles théoriques ; sur la première bissectrice. Les points déviant reflètent la légère platikurticité de la distribution empirique des w_i ($k^* = 2.9328$).

Les tests statistiques, que nous avons effectués, nous montrent que les réalisations du processus M dans le temps financier sont des réalisations d'un mouvement brownien standard. Autrement dit, le processus M des prix S est un mouvement brownien subordonné au temps de changement Q . Sous l'hypothèse de continuité, nous ne pouvons pas rejeter l'hypothèse que le processus est une semimartingale, donné en (3.5.1), pour l'indice S&P 500 *futures*.

Distribution des rendements logarithmiques journaliers normalisés par les volatilités réalisées

Les rendements logarithmiques journaliers normalisés par leurs volatilités réalisées, *i.e.*,

$$\hat{u}_n^{(RV)} = \frac{r_n^*}{(rv_5^{(2)})^{\frac{1}{2}}} \text{ constituent un bruit blanc fort gaussien, comme le montre les figures (3.5.5),}$$

(3.5.6) et le tableau (3.5.5) ci-dessous.

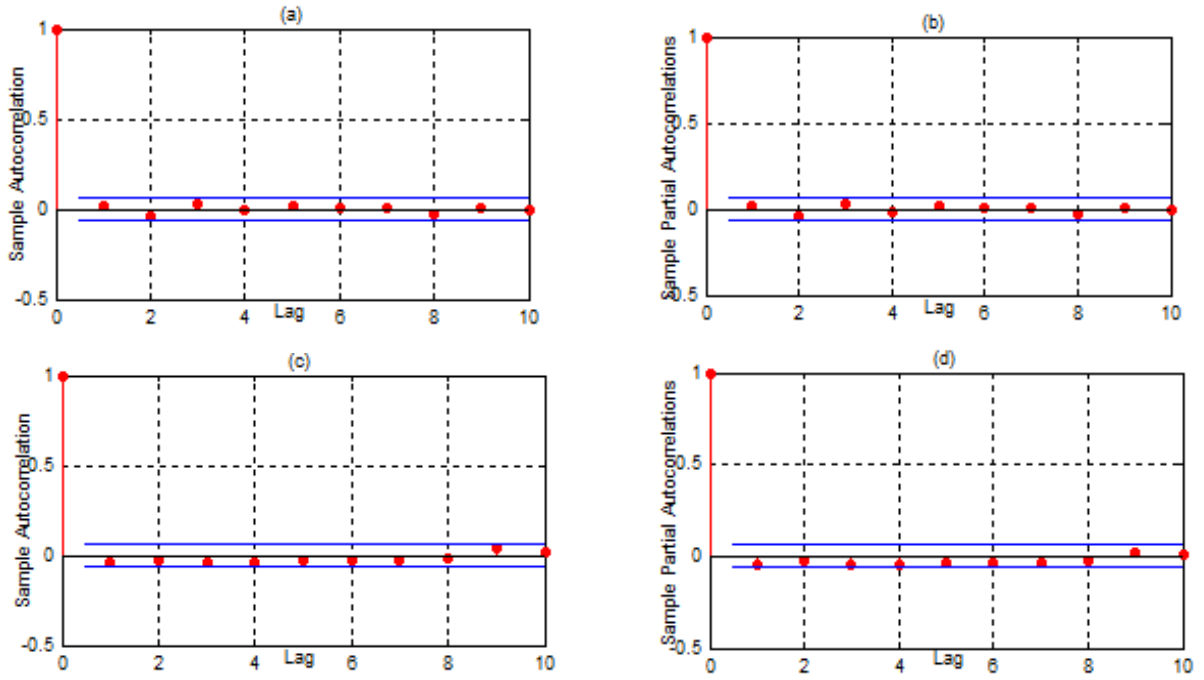


Figure 3.5.5 : Fonction d'autocorrélation (empirique) de la série $\hat{u}_n^{(RV)}$ et de son carré $(\hat{u}_n^{(RV)})^2$ (a et c) respectivement, ainsi que leur fonction d'autocorrélation partielle (b et d) respectivement.

le retard (l)	$\hat{u}_n^{(RV)}$			$(\hat{u}_n^{(RV)})^2$		
	R	valeur critique	p -value	R	valeur critique	p -value
1	0.1754	2.7055	0.6754	2.0143	2.7055	0.1558
2	1.8821	4.6052	0.3902	2.7661	4.6052	0.2508
3	2.4794	6.2514	0.4790	4.6443	6.2514	0.1998
4	2.5938	7.7794	0.6279	6.5267	7.7794	0.1631
5	2.8889	9.2364	0.7171	7.4897	9.2364	0.1867
6	2.9416	10.6446	0.8161	8.2223	10.6446	0.2223
7	2.9772	12.0170	0.8871	8.9593	12.0170	0.2556
8	3.7149	13.3616	0.8819	9.2981	13.3616	0.3178
9	3.7309	14.6837	0.9282	10.5456	14.6837	0.3081
10	3.7355	15.9872	0.9585	10.8766	15.9872	0.3672

Table 3.5.5 : Test d'absence d'autocorrélation pour les séries $\hat{u}_n^{(RV)}$ et $(\hat{u}_n^{(RV)})^2$.

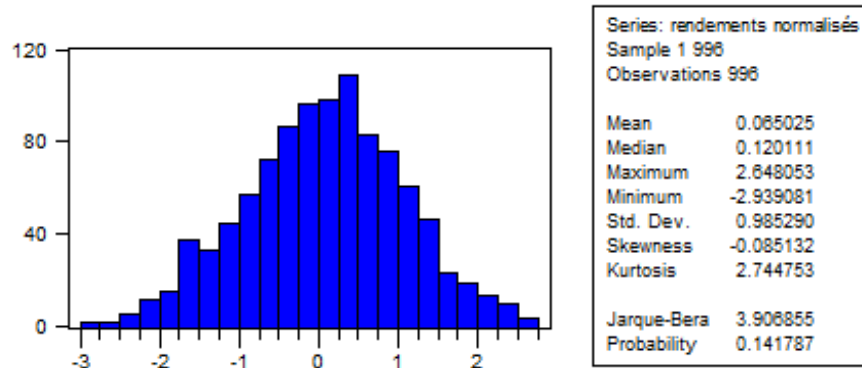


Figure 3.5.6 : Histogramme des rendements logarithmiques journaliers normalisés par leurs volatilités réalisées accompagné d'une table comportant les statistiques descriptives (moyenne, écart-type, skewness, kurtosis, la statistique de Jarque-bera,....,etc)

Remarques :

- 1) Nous concluons que les rendements logarithmiques journaliers normalisés par leurs volatilités réalisées sont *i.i.d.* gaussiens. Cette conclusion est en contradiction avec celle de Peters et de Vilder (2006), qui ont étudié la même série, mais pour la période allant du 1 janvier 1988 au 1 septembre 2001 qui est différente de la nôtre. Néanmoins, ils acceptent cette hypothèse pour l'indice boursier AEX (Amsterdam Exchange) (2002). Nous pensons que cela est peut être dû à la taille de l'échantillon (996 jours contre 3452) et/ou au pas d'échantillonnage (5 minutes contre 2 minutes).
- 2) D'autres auteurs, par exemple Andersen et *al* (2000) sur le taux de change du yen japonais et deutch mark allemand par rapport au dollar américain , Ané et Geman (2000) sur les actions technologiques Intel et Cisco systems , ont accepté l'hypothèse de normalité des rendements normalisés par leurs volatilités réalisées. Areal et Taylor (2000) tirent la même conclusion sur l'indice FTSE 100 *futures*.

Modélisation de la série des rendements journaliers r_n^*

Maintenant, nous nous intéressons à la modélisation de la série des rendements logarithmiques journaliers r_n^* de l'indice S&P 500 *futures* du 02/01/2004 au 31/12/2007 par la classe des modèles *GARCH*. Le processus des rendements des actifs financiers est de la classe des processus stochastiques, du second ordre, faiblement stationnaires, comme le confirme le test de *DF* (Dickey-Fuller) suivant.

Considérons les trois modèles suivants :

$$\begin{aligned} \text{modèle 1 : } \Delta r_n^* &= \phi r_{n-1}^* + \varepsilon_n, \\ \text{modèle 2 : } \Delta r_n^* &= \phi r_{n-1}^* + c + \varepsilon_n, \\ \text{modèle 3 : } \Delta r_n^* &= \phi r_{n-1}^* + c + \beta n + \varepsilon_n. \end{aligned}$$

Nous commençons par estimer le modèle 3, incluant une constante et une tendance. Nous testons la présence d'une racine unitaire dans le processus en testant la nullité du paramètre ϕ . Les résultats du test sont reproduits dans la figure (3.5.7).

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-31.50348	0.0000
Test critical values:		
1% level	-3.967298	
5% level	-3.414336	
10% level	-3.129291	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(RN*)
 Method: Least Squares
 Date: 11/03/11 Time: 08:35
 Sample (adjusted): 2 996
 Included observations: 995 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RN*(-1)	-1.000120	0.031746	-31.50348	0.0000
C	0.000135	0.000414	0.325538	0.7448
@TREND(1)	-2.60E-07	7.19E-07	-0.361605	0.7177

Figure 3.5.7 : Test de racine unitaire sur la série des rendements logarithmiques r_n^* : modèle 3.

La valeur estimée de la statistique $DF = -31.50 < -3.41$ (lue à partir de la table de Dickey-Fuller). Ainsi, dans ce cas pour un seuil de 5%, nous rejetons l'hypothèse nulle de racine unitaire ($\phi = 0$). Par suite, nous vérifions que le modèle à partir duquel nous avons fait le test (modèle 3) est le bon modèle. Cela revient à tester la nullité du paramètre β , en effectuant un test de Student, $t_{\hat{\beta}}$, où $\hat{\beta}$ désigne l'estimateur des MCO. La réalisation

$|t_{\hat{\beta}}| = |-0.36| < 1.96$, au seuil de 5%. Donc, nous acceptons l'hypothèse nulle de la non significativité de la tendance et par conséquent, le modèle 3 est remis en cause.

Nous passons à l'estimation du modèle 2 et nous testons la présence d'une racine unitaire. $DF_{calculée} = -31.51 < DF_{tabulée} = -2.86$, nous rejetons l'hypothèse de la présence d'une racine unitaire dans le modèle 2. Il faut à nouveau évaluer la validité de notre diagnostic en vérifiant que le modèle 2 est le bon modèle. Pour cela, nous testons la nullité de la constante par le test statistique de Student. La valeur estimée de cette dernière $|t_{\hat{c}}| = 0.024 < 1.96$, au seuil de 5%, où \hat{c} désigne l'estimateur des MCO. Donc, nous acceptons l'hypothèse nulle de la nullité la constante \hat{c} et le modèle 2 est remis en cause, (figure(3.5.8)).

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-31.51530	0.0000
Test critical values:		
1% level	-3.436703	
5% level	-2.864233	
10% level	-2.568256	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(RN*)
 Method: Least Squares
 Date: 11/03/11 Time: 08:42
 Sample (adjusted): 2 996
 Included observations: 995 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RN*(-1)	-1.000017	0.031731	-31.51530	0.0000
C	5.09E-06	0.000207	0.024623	0.9804

Figure 3.5.8 : Test de racine unitaire sur la série des rendements logarithmiques r_n^* : modèle 2.

Nous estimons le modèle 1. La valeur estimée de la statistique DF ; $DF_{calculée} = -31.53 < DF_{tabulée} = -1.94$ (les résultats du test sont reportés dans la figure (3.5.9) ci-dessous). Donc, nous rejetons l'hypothèse nulle de la racine unitaire.

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-31.53114	0.0000
Test critical values:		
1% level	-2.567288	
5% level	-1.941142	
10% level	-1.616485	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(RN*)
 Method: Least Squares
 Date: 11/03/11 Time: 08:44
 Sample (adjusted): 2 996
 Included observations: 995 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RN*(-1)	-1.000017	0.031715	-31.53114	0.0000

Figure 3.5.9 : Test de racine unitaire sur la série des rendements logarithmiques r_n^* : modèle 1.

Nous concluons que le processus des rendements $\{r_n^*, n \in \mathbb{Z}\}$, associé à l'indice boursier S&P 500 *futures* est stationnaire au second ordre.

L'observation du corrélogramme, figure (3.5.10), ci-dessous, de la série des rendements S&P 500 *futures* (r_n^*), permet de constater l'absence de corrélation de cette série. Ce constat est confirmé par le test de Ljung-Box, où les autocorrélations sont significativement nulles (la p -value est supérieure à 5%).

Autocorrelation	Partial Correlation	AC	PAC	R-Stat	Prob	
		1	-0.000	-0.000	3.E-07	1.000
		2	-0.058	-0.058	3.4099	0.182
		3	0.044	0.044	5.3484	0.148
		4	-0.008	-0.012	5.4135	0.247
		5	0.007	0.012	5.4634	0.362
		6	0.016	0.013	5.7103	0.456
		7	-0.006	-0.004	5.7463	0.570
		8	-0.069	-0.068	10.481	0.233
		9	-0.006	-0.007	10.516	0.310
		10	0.022	0.015	10.993	0.358
		11	0.015	0.020	11.231	0.424
		12	-0.024	-0.023	11.833	0.459
		13	0.042	0.045	13.635	0.400
		14	-0.008	-0.011	13.708	0.472
		15	0.004	0.011	13.723	0.547
		16	0.054	0.043	16.688	0.406
		17	0.028	0.030	17.475	0.423
		18	0.001	0.008	17.476	0.491
		19	-0.024	-0.023	18.056	0.519
		20	0.010	0.007	18.164	0.577

Figure 3.5.10 : Autocorrélations des rendements de l'indice S&P 500 sur la période 2004-2007

Par contre les carrés des rendements sont fortement autocorrélés. Le graphe de la fonction des autocorrélations de la série r_n^{*2} donné par la figure (3.5.11), reflète cette constatation qui est confirmée par le test de corrélation de Ljung-Box.

Autocorrelation	Partial Correlation	AC	PAC	R-Stat	Prob	
		1	0.081	0.081	6.5063	0.011
		2	0.086	0.080	13.883	0.001
		3	0.050	0.038	16.432	0.001
		4	0.038	0.026	17.913	0.001
		5	0.051	0.040	20.481	0.001
		6	0.045	0.032	22.471	0.001
		7	0.046	0.032	24.611	0.001
		8	0.058	0.043	27.955	0.000
		9	0.116	0.100	41.503	0.000
		10	0.085	0.059	48.706	0.000
		11	0.093	0.063	57.411	0.000
		12	0.075	0.044	63.068	0.000
		13	0.053	0.022	65.870	0.000
		14	0.042	0.012	67.636	0.000
		15	0.006	-0.022	67.676	0.000
		16	0.035	0.011	68.889	0.000
		17	0.051	0.027	71.520	0.000
		18	0.038	0.006	73.007	0.000
		19	-0.005	-0.041	73.038	0.000
		20	0.019	-0.011	73.410	0.000

Figure 3.5.11 : Autocorrélations des carrés des rendements de l'indice S&P 500 sur la période 2004 – 2007

L'absence de corrélation des rendements et la corrélation des carrés de ces rendements dans la série S&P 500, nous incitent à adopter une modélisation de type *ARCH* / *GARCH*. Ainsi, nous trouvons que les modèles candidats pour cette série sont *ARCH* (1), *ARCH* (2), *ARCH* (3), et *GARCH* (1, 1). Néanmoins, l'utilisation du critère d'Akaike et Schwarz

((-7.236, -7.227), (-7.242, -7.227), (-7.243, -7.223) et (-7.267, -7.252) respectivement), nous permet de retenir le modèle *GARCH* (1, 1), où l'estimation de sa moyenne conditionnelle n'est pas significative, par contre celle de sa variance conditionnelle est donnée par

$$\hat{\sigma}_n^2 = \underset{(2.789053)}{2.04 \times 10^{-6}} + \underset{(3.588302)}{0.051215} r_{n-1}^{*2} + \underset{(31.75049)}{0.901326} \hat{\sigma}_{n-1}^2, \quad (3.5.7)$$

où la significativité des paramètres est vérifiée au seuil $\alpha = 5\%$. (les valeurs des *t*-statistiques données entre parenthèses pour l'ensemble des paramètres dans l'équation ci-dessus sont supérieures à 1.96, en valeur absolue).

Autocorrelation	Partial Correlation	AC	PAC	R-Stat	Prob	
		1	0.005	0.005	0.0215	0.883
		2	-0.037	-0.037	1.3997	0.497
		3	0.032	0.032	2.4038	0.493
		4	-0.005	-0.007	2.4335	0.657
		5	0.012	0.015	2.5888	0.763
		6	0.006	0.004	2.6242	0.854
		7	-0.007	-0.006	2.6776	0.913
		8	-0.050	-0.050	5.1779	0.738
		9	-0.004	-0.005	5.1977	0.817
		10	0.010	0.007	5.3002	0.870
		11	0.011	0.014	5.4280	0.909
		12	-0.019	-0.019	5.8025	0.926
		13	0.041	0.043	7.4747	0.876
		14	-0.003	-0.006	7.4868	0.914
		15	0.008	0.012	7.5587	0.940
		16	0.051	0.045	10.233	0.854
		17	0.014	0.015	10.425	0.885
		18	-0.001	0.002	10.426	0.917
		19	-0.016	-0.017	10.691	0.934
		20	0.010	0.009	10.796	0.951

Figure 3.5.12 : Autocorrélations des rendements standardisés $\hat{u}_n^{(GARCH)}$ de l'indice S&P 500 futures sur la période 2004-2007.

Considérons la série des rendements standardisés estimés $\hat{u}_n^{(GARCH)} = \frac{r_n^*}{\hat{\sigma}_n}$. Cette dernière série forme un bruit blanc. Le graphe des autocorrélations de la série $\hat{u}_n^{(GARCH)}$ reflète cette constatation qui est confirmée par le test de Ljung-Box donné sur la même figure (3.5.12).

La même constatation est faite sur le carré de la série des rendements standardisés $\left(\hat{u}_n^{(GARCH)}\right)^2$, donnée dans la figure (3.5.13) ci-dessous.

Autocorrelation	Partial Correlation	AC	PAC	R-Stat	Prob	
		1	0.005	0.005	0.0215	0.883
		2	-0.037	-0.037	1.3997	0.497
		3	0.032	0.032	2.4038	0.493
		4	-0.005	-0.007	2.4335	0.657
		5	0.012	0.015	2.5888	0.763
		6	0.006	0.004	2.6242	0.854
		7	-0.007	-0.006	2.6776	0.913
		8	-0.050	-0.050	5.1779	0.738
		9	-0.004	-0.005	5.1977	0.817
		10	0.010	0.007	5.3002	0.870
		11	0.011	0.014	5.4280	0.909
		12	-0.019	-0.019	5.8025	0.926
		13	0.041	0.043	7.4747	0.876
		14	-0.003	-0.006	7.4868	0.914
		15	0.008	0.012	7.5587	0.940
		16	0.051	0.045	10.233	0.854
		17	0.014	0.015	10.425	0.885
		18	-0.001	0.002	10.426	0.917
		19	-0.016	-0.017	10.691	0.934
		20	0.010	0.009	10.796	0.951

Figure 3.5.13 : Autocorrélations des carrés des rendements standardisés $\left(\hat{u}_n^{(GARCH)}\right)^2$ de l'indice S&P 500 *futures* sur la période 2004-2007.

Comparaison entre les estimateurs utilisant respectivement σ_n selon un modèle paramétrique *GARCH*, et celui, non paramétrique, utilisant *RV5_n*

Considérons les rendements standardisés par les volatilités réalisées $\hat{u}_n^{(RV)}$, et les rendements standardisés par la racine carrée de la variance conditionnelle du modèle *GARCH* (1, 1), $\hat{u}_n^{(GARCH)}$, les résultats sont reportés dans le tableau (3.5.6) ci-dessous.

	$\hat{u}_n^{(RV)}$	$\hat{u}_n^{(GARCH)}$
moyenne	0.065	0.00087
écart-type	0.9859	1.00057
skewness	-0.0851	-0.437
kurtosis	2.744	4.418
jarque-bera	3.90	115.25
p-value	0.1417	0.00

Table 3.5.6 : Les propriétés de la distribution des rendements standardisés $\hat{u}_n^{(RV)}$ et $\hat{u}_n^{(GARCH)}$.

Contrairement aux rendements standardisés $\hat{u}_n^{(RV)}$, les rendements standardisés $\hat{u}_n^{(GARCH)}$, ne sont pas gaussiens. Le test de Jarque-Bera, confirme ce constat. La distribution de ces derniers est caractérisée par une skewness négative et des queues épaisses. Néanmoins, les $\hat{u}_n^{(GARCH)}$ ont une moyenne nulle et une variance égale à l'unité.

Ces résultats coïncident avec ceux obtenus par les auteurs Andersen et *al* (2000) sur le taux de change du yen japonais et deutsche mark allemand par rapport au dollar américain.

D'après cette étude nous concluons quant à la performance de l'estimateur $RV_n^{(2)}$, en utilisant des rendements intra-journaliers à haute fréquence. Ainsi, l'ajustement par un modèle *GARCH* est insuffisant pour éliminer l'excès de kurtosis, alors que la variance réalisée est capable d'accomplir cette tâche.

3.6 Conclusion

La volatilité des prix en finance est un phénomène non observable, pour contourner ce problème, plusieurs proxies de volatilité ont été introduites dans la littérature financière, allant des proxies basées sur des données de basse fréquence (journalière par exemple), vers celles basées sur les données de haute fréquence. La volatilité réalisée exploitant l'information intra-journalière (à haute fréquence) joue un rôle primordial en finance.

Le fait de supposer que le processus des prix est une semimartingale (ce qui élimine la présence des opportunités d'arbitrage); la variation quadratique, qui est la limite (en probabilité) de la variance réalisée en absence des effets de microstructure (néanmoins, leurs présences sont résolues en pratique par le choix d'un pas d'échantillonnage Δ modéré, de 5 minutes ou de 15 minutes, par exemple) coïncide avec la volatilité (intégrée).

Afin d'illustrer ces constats, nous avons fait une application sur la série de données réelles de l'indice S&P 500 *futures*, sur la période 2004 – 2007. Nous avons trouvé que le processus des prix qui lui est associé est une semimartingale. De plus, nous avons décrit les propriétés des volatilités réalisées de cet indice boursier qui sont en adéquation avec les résultats trouvés par plusieurs auteurs.

4

Estimation de modèles GARCH en présence de données de haute fréquence.

4.1 Présentation du modèle

Ce dernier chapitre consiste à montrer à quel point l'estimation des paramètres du modèle *GARCH* par la méthode du QMV sera améliorée, si une bonne proxy de volatilité est utilisée en présence de données de haute fréquence. Les propriétés asymptotiques des estimateurs, à savoir la consistance et la normalité asymptotique sont conservées *i.e.*, comme dans le cas classique pour les rendements *close-to-close*. Un gain en efficacité des estimateurs est obtenu si l'on travaille sur le logarithme des proxies H_n au lieu des proxies H_n . Lors de l'étude des propriétés statistiques, nous utilisons la théorie développée par Visser (2008) fondée sur la méthode du QMV. L'intuition sous-jacente est de construire la fonction de vraisemblance en se basant sur des proxies H_n , qui exploitent l'information contenue dans le processus des rendements intra-journaliers $R_n(\cdot)$. Une étude de simulation et une application sur la série de données de S&P 500 *futures* ont été faites. Elles confortent les résultats théoriques.

Pour modéliser la volatilité journalière, la plupart des travaux font appel aux rendements journaliers *close-to-close* r_n . Ainsi, nous supposons que la suite des rendements logarithmiques journaliers $(r_n)_{n \geq 1}$, est issue d'un modèle *GARCH* (1, 1), stationnaire où nous

utilisons la représentation suivante donnée par Drost et Klaassen (1997) :

$$r_n = v_n \kappa Z_n, \quad (4.1.1a)$$

$$v_n^2 = 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2, \quad (4.1.1b)$$

où les innovations Z_n sont *i.i.d.* de moyenne nulle et de variance égale à 1. Le paramètre κ est dit paramètre de normalisation, γ et β sont des paramètres à estimer. Le système (4.1.1) est équivalent aux équations d'un modèle *GARCH* classique donné en chapitre 2, dans le cas où $p = q = 1$, en posant

$$\sigma_n = v_n \kappa, \quad \omega = \kappa^2, \quad \alpha = \gamma \kappa^2.$$

Afin d'exploiter l'information contenue dans les données de haute fréquence, plus précisément des données intra-journalières, Visser (2008) a introduit un modèle non paramétrique pour le processus des prix intra-journaliers dans le système d'équations (4.1.1) où, pour chaque jour ouvrable n , il a considéré un processus en temps continu pour les rendements logarithmiques; noté $R_n(\cdot)$. Pour la simplicité des notations, nous réduisons la journée ouvrable à l'intervalle de temps unité $[0, 1]$. Par suite le modèle *GARCH* (1, 1) donnée en (4.1.1) devient

$$R_n(u) = v_n \kappa \Psi_n(u), \quad 0 \leq u \leq 1, \quad (4.1.2a)$$

$$v_n^2 = 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2, \quad (4.1.2b)$$

où l'équation (4.1.2a) est dite modèle d'échelle (en anglais : scaling model).

La suite de processus $(\Psi_n(\cdot))_n$ est supposée *i.i.d.*, et à trajectoires *càdlàg*, tel que $\mathbb{E}(\Psi_n(1)^2) = 1$. En fait, $R_n(0)$ et $R_n(1)$ donnent le rendement *overnight* et le rendement *close-to-close* r_n respectivement. À noter que le facteur d'échelle $v_n \kappa$ est le même que celui donné en équation (4.1.1a), dans le cas discret; de plus, il est supposé constant au cours de la journée ouvrable. L'hypothèse que la suite de processus $(\Psi_n)_{n \geq 0}$ est *i.i.d.* garantit que la suite de variables aléatoires $(Z_n)_{n \geq 1}$ l'est aussi. Ainsi, en posant $Z_n \equiv \Psi_n(1)$, nous pouvons retrouver les rendements *close-to-close* r_n , autrement dit, $r_n \equiv R_n(1)$.

Après avoir incorporé le processus des rendements logarithmiques intra-journaliers dans le modèle $GARCH(1,1)$, nous pouvons utiliser l'information contenue dans ces données intra-journalières et construire des proxies H_n pour les facteurs d'échelle.

Pour le modèle $GARCH(1,1)$, Visser (2008) a introduit des *proxies* pour la volatilité $v_n \kappa$. Rappelons d'abord qu'une variable aléatoire $H_n = H(R_n)$ est dite une proxy, si la fonctionnelle H est positive et de plus, possède la propriété d'homogénéité positive en R_n . Nous supposons que la variable aléatoire $H(\Psi)$ n'est pas identiquement nulle; autrement dit $\mu_2^H = \sqrt{\mathbb{E}(H^2(\Psi))} > 0$. Nous notons Z_H , les innovations normalisées données par $Z_H \equiv H(\Psi) / \mu_2^H < \infty$. Ainsi, $\mathbb{E}(Z_H^2) = 1$.

Par homogénéité on a

$$H_n = H(R_n) = v_n \kappa H(\Psi_n). \quad (4.1.3)$$

Par suite, en remplaçant $H(\Psi_n) = Z_H \mu_2^H$ dans (4.1.3), nous obtenons $H_n = v_n \kappa_H Z_{H,n}$, où $\kappa_H = \kappa \mu_2^H$, et les innovations $Z_{H,n} > 0$ sont *i.i.d.* et telles que $\mathbb{E}(Z_{H,n}^2) = 1$.

Alors, (4.1.2) devient

$$H_n = v_n \kappa_H Z_{H,n}, \quad (4.1.4a)$$

$$v_n^2 = 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2, \quad (4.1.4b)$$

et $\mathbb{E}(H_n^2 | \mathcal{F}_{n-1}) = v_n^2 \kappa_H^2$, où \mathcal{F}_{n-1} est l'information dont on dispose jusqu'à l'instant $n-1$.

Remarques :

- 1) Les paramètres γ et β donnés en équation (4.1.1b) sont identiques à ceux donnés en (4.1.4b). Ainsi, la *proxy* H_n et les rendements logarithmiques journaliers r_n se partagent le même facteur v_n et, par conséquent, leurs moments d'ordre deux conditionnels ont la même dynamique; autrement dit, des valeurs élevées du facteur d'échelle n'implique pas seulement des valeurs élevées pour les rendements r_n (en valeur absolue), mais aussi des valeurs élevées pour n'importe qu'elle proxy de volatilité H_n .
- 2) Comme *proxy* de volatilité, nous pouvons citer la volatilité réalisée journalière qui est

largement appliquée comme proxy et occupe une place primordiale en pratique, étant données ses propriétés statistiques. Néanmoins, elle souffre d'inconvénients, comme les effets de microstructure (voir, par exemple Barndorff-Nielsen et Shephard (2002) et Andersen, Bollerslev, Diebold, et Labys (2003)). La volatilité journalière réalisée est la racine carrée de la variance réalisée, qui est la somme des incréments des rendements intra-journaliers élevés au carré. D'autres *proxies* peuvent être utilisées.

4.2 Propriétés asymptotiques des estimateurs du quasi maximum de vraisemblance basés sur des proxies de volatilité

Cette section est consacrée à l'étude des propriétés statistiques asymptotiques des estimateurs du QMV gaussien (resp log-gaussien) du modèle *GARCH* (1,1) basés sur les *proxies* H_n (resp $\log(H_n)$). Elle consiste à généraliser le cas classique, basé sur les rendements journaliers *close-to-close* r_n , au cas des données de haute fréquence basé sur les proxies de volatilité H_n ou $\log(H_n)$ selon le cas.

4.2.1 Consistance et normalité asymptotique des estimateurs du QMV gaussien

Quasi maximum de vraisemblance gaussien

Étant donné les observations (y_1, y_2, \dots, y_N) issues d'un modèle *GARCH* (1,1), adapté à la filtration \mathcal{F}_{n-1}

$$\begin{aligned} y_n &= v_n \kappa Z_n, \\ v_n^2 &= 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2 \end{aligned} \tag{4.2.1}$$

Nous définissons les fonctions de moyenne et de variance conditionnelle $\mu_n(\theta)$ et $h_n(\theta)$ respectivement où ces dernières sont paramétrisées par un vecteur de dimension finie $\theta = (\kappa, \gamma, \beta)$. La vraie valeur (inconnue) du paramètre est $\theta^0 = (\kappa^0, \gamma^0, \beta^0) \in \Theta$. $\mu_n(\theta) = \mathbb{E}(y_n | \mathcal{F}_{n-1})$ et $h_n(\theta) = \text{var}(y_n | \mathcal{F}_{n-1})$.

La fonction de log-vraisemblance gaussienne $L_N(\theta; y_1, y_2, \dots, y_N)$ (à une constante près) est donnée par

$$\begin{aligned} L_N(\theta; y_1, y_2, \dots, y_N) &= \sum_{n=1}^N l_N(\theta; y_1, y_2, \dots, y_N), \\ &= -\frac{1}{2} \sum_{n=1}^N \left(\log(h_n(\theta)) + \frac{y_n^2}{h_n(\theta)} \right). \end{aligned}$$

Dans ce qui suit, nous adoptons le même principe pour établir les propriétés asymptotiques de l'estimateur du QMV $\hat{\theta}_N = \arg \max_{\theta \in \Theta} L_N(\theta)$, mais cette fois-ci en se basant sur les proxies de volatilité H_n , autrement dit, $y_n = H_n$. Nous traitons la proxy H_n comme si elle est issue d'une loi normale de moyenne μ_n égale à 0 et de variance $h_n = (v_n \kappa_H)^2$ (Comme le supposent Bollerslev et Wooldridge (1992), pour le *GARCH* classique). Alors, nous obtenons une fonction log-vraisemblance $L_{H,N}(\theta) = \sum_{n=1}^N l_{H,N}(\theta)$ à une constante près, tel que

$$\begin{aligned} L_{H,N}(\theta) &= -\frac{1}{2} \sum_{n=1}^N \left(\log(h_n(\theta)) + \frac{H_n^2}{h_n(\theta)} \right), \\ &= -\frac{1}{2} \sum_{n=1}^N \left(\log(v_n^2(\gamma, \beta) \kappa_H^2) + \frac{v_n^2(\gamma^0, \beta^0) (\kappa_H^0)^2 Z_{n,H}^2}{v_n^2(\gamma, \beta) \kappa_H^2} \right), \end{aligned}$$

où les innovations $Z_{H,n}$ sont indépendantes de $h_n(\theta)$ et de v_n , de plus elles satisfont $\mathbb{E}(Z_{H,n}^2) = 1$, comme nous l'avons déjà signalé.

Dans ce qui suit, $v_n \kappa_H$ sera noté par $\sigma_{H,n}$,

Pour étudier la consistance et la normalité asymptotique des estimateur du QMV $\hat{\theta}_N$, Visser (2008) considère les hypothèses suivantes.

B1 : Z_n est une séquence de variables aléatoires *i.i.d.* avec $\mathbb{E}(Z^2) = 1$, (l'indice n pouvant être omis).

B2 : Θ est un sous-espace compact de l'espace donné par $\kappa_H > 0$, $\gamma > 0$ et $\beta \in [0, 1[$, $\overset{0}{\Theta}$ est l'intérieur de Θ .

B3 : $\mathbb{E} \left(\log \left(\gamma^0 (\kappa^0)^2 Z^2 + \beta^0 \right) \right) < 0$,

B4 : Z^2 est non dégénérée,

B5 : $\mathbb{E}(Z_H^4) < \infty$,

Théorème 4.2.1. (Visser (2008))

Soit $\theta^0 = (\kappa_H^0, \gamma^0, \beta^0)$ et $\kappa_H^0 = \kappa^0 \mu_2^H$. Sous les hypothèses **B1-B5**, l'estimateur du QMV $\hat{\theta}_N$, basé sur les proxies H_n est asymptotiquement normal :

$\sqrt{N}(\hat{\theta}_N - \theta^0) \xrightarrow{\mathcal{L}} N(0, V_0)$, $N \rightarrow \infty$, avec $V_0 = \text{var}(Z_H^2) G_H^{-1}(\theta^0)$, où le terme général de la matrice G_H est donné par

$$G_H(\theta^0)_{i,j} = \mathbb{E}_{\theta^0} \left[\frac{1}{\sigma_H^4(\theta^0)} \left(\frac{\partial \sigma_{H,0}^2(\theta^0)}{\partial \theta_i} \right) \left(\frac{\partial \sigma_{H,0}^2(\theta^0)}{\partial \theta_j} \right) \right], \sigma_{H,n} = v_n \kappa_H.$$

Remarques :

- 1) La condition **(B1)** n'implique pas que $\mathbb{E}(Z) = 0$, car les rendements journaliers r_n peuvent ne pas avoir une moyenne nulle. Le fait que $\mathbb{E}(H_n^2 | \mathcal{F}_{n-1}) = v_n^2 \kappa_H^2$ assurent que la suite $(Z_{H,n})$ est *i.i.d.*, avec $\mathbb{E}(Z_{H,n}^2) = 1$.
- 2) La deuxième condition **(B2)** veut dire que le vrai vecteur des paramètres $\theta^0 \in \overset{0}{\Theta}$. Si θ^0 se trouve sur les bords de Θ , alors l'estimateur du QMV est toujours consistant, mais généralement n'est pas asymptotiquement gaussien (Francq et Zakoian (2008)). Puisque $\kappa_H = \kappa \mu_2^H$ et comme l'optimisation se fait sur la fonction $L_{H,N}$, il est raisonnable de poser la condition de positivité sur κ_H et non sur κ . Ainsi, dire κ_H est strictement positif implique la positivité stricte de κ .
- 3) La condition **(B3)** est la condition usuelle de la stationnarité stricte et d'ergodicité de processus (v_n) (l'indice n dans le terme général de la matrice G_H est remplacé par 0). À noter que si $\gamma^0 (\kappa^0)^2 Z^2 + \beta < 1$, la condition **(B3)** sera vérifiée en utilisant l'inégalité de Jensen, de plus le processus est faiblement stationnaire.
- 4) **(B4)** est une condition qui sert à l'identification de processus (v_n) .
- 5) La condition **(B5)** est nécessaire pour établir la normalité asymptotique des estimateur du QMV.

Le lemme suivant établi par Visser (2008), nous permet de comparer entre les matrices

de variance covariance V_0 des estimateurs de vecteur des paramètres (γ, β) , basé sur deux proxies H différentes.

Lemme 4.2.1. (Visser, 2008)

Le block des paramètres (γ, β) dans la matrice $G_H^{-1}(\theta^0)$, ne dépend pas de la proxy H .

D'après le théorème (4.2.1) et le lemme (4.2.1), il s'ensuit que pour comparer entre deux estimateurs du QMV, il suffit de comparer la variance de leurs innovations Z_H , sans tenir compte des valeurs du vecteur des paramètres à estimer.

Corollaire 4.2.1. (Visser, 2008)

Soit l'estimateur du QMV du vecteur des paramètres $\theta = (\gamma, \beta)$, fondés sur les proxies H_n et H'_n respectivement. Alors, l'efficacité relative asymptotique (EAR) de l'estimateur fondé sur H' relativement à celui fondé sur H est

$$EAR_{Gaussien}(H, H') = \frac{\text{var}(Z_H^2)}{\text{var}(Z_{H'}^2)}.$$

4.3 Estimation du modèle log-GARCH

4.3.1 Consistance et normalité asymptotique des estimateurs du QMV log-gaussien

Il est possible d'estimer les paramètres du modèle $GARCH(1,1)$ donné en (4.1), en utilisant le logarithme des proxies *i.e.*, $\log(H_n)$. Dans ce cas, on est en présence des estimateurs dits QMV log-gaussien, nous procédons de la même manière que précédemment. En introduisant l'opérateur logarithme sur les deux membres de l'équation (4.1.4a), on obtient $\log(H_n) = \log(v_n) + \log(\kappa_H) + \log(Z_{H,n})$. Définissons $\tilde{\kappa}_H = \kappa_H \exp\{\mathbb{E}(\log(Z_{H,n}))\}$, et

$$U_{H,n} = \frac{\log(Z_{H,n}) - \mathbb{E}(\log(Z_{H,n}))}{\sqrt{\text{var}(\log(Z_{H,n}))}}.$$

Par suite, nous obtenons la forme additive suivante :

$$\log(H_n) = \log(v_n) + \log(\tilde{\kappa}_H) + \lambda U_{H,n},$$

où les erreurs $U_{H,n}$ sont *i.i.d.* de moyenne nulle et de variance égale à l'unité.

La moyenne et variance conditionnelles de ces proxies logarithmiques sont

$$\begin{aligned}\mathbb{E}(\log(H_n) \mid \mathcal{F}_{n-1}) &= \log(v_n) + \log(\tilde{\kappa}_H), \\ \text{var}(\log(H_n) \mid \mathcal{F}_{n-1}) &= \lambda^2.\end{aligned}$$

respectivement.

Soit $\tilde{\theta} = (\tilde{\kappa}_H, \gamma, \beta)$ et $\eta = (\tilde{\theta}, \lambda)$. Nous traitons la proxy $\log(H_n)$ comme si elle était issue d'une loi normale de moyenne $\mu_n(\eta) = \log(\sigma_{H,n}(\tilde{\theta}))$ et de variance $h_n(\eta) = \lambda^2$.

Supposons que les conditions **B1-B5** sont vérifiées où la condition **B5** est remplacée par **B5'** : $\mathbb{E}(\log(Z_H^4)) < \infty$, alors

$$\sqrt{n}(\tilde{\eta}_n - \eta^0) \xrightarrow{\mathcal{L}} N(0, V_0), \quad n \rightarrow \infty,$$

$$\text{avec } V_0 = 4(\lambda^0)^2 \begin{pmatrix} G_H^{-1}(\tilde{\theta}^0) & \frac{1}{2}\mathbb{E}(U_H^3)\mathbb{E}\left(\frac{\partial \mu_n}{\partial \theta}(\tilde{\theta}^0)\right) \\ \frac{1}{2}\mathbb{E}(U_H^3)\mathbb{E}\left(\frac{\partial \mu_n}{\partial \theta}(\tilde{\theta}^0)\right) & \frac{1}{16}\text{var}(U_H^2) \end{pmatrix}, \text{ tel que } (\lambda^0)^2 =$$

$\text{var}(\log(Z_H))$.

L'efficacité relative asymptotique de ces estimateurs log-gaussiens est

$$EAR_{\log\text{-Gaussien}}(H, H') = \frac{\text{var}(\log(Z_H^2))}{\text{var}(\log(Z_{H'}^2))}$$

4.3.2 Gain en efficacité du QMV log-gaussien relativement au QMV gaussien

La comparaison de l'efficacité asymptotique des estimateurs $\hat{\gamma}$ et $\hat{\beta}$ du QMV gaussien relativement au log-gaussien, basé sur la même proxy H_n est donnée par

$$EAR_{(\log\text{-Gaussien}, \text{Gaussien})}(H, H) = \frac{\text{var}(\log(Z_H^2))}{\text{var}(Z_H^2)}.$$

4.4 Simulation et application

4.4.1 Simulation

Dans les sections précédentes, nous avons vu que l'estimateur du QMV des paramètres d'un processus *GARCH* sera considérablement amélioré une fois qu'une proxy de volatilité adéquate est utilisée en présence de données de haute fréquence.

Nous allons, dans cette section, illustrer ce résultat théorique, à l'aide d'une étude de simulation intensive. Pour ce faire, nous avons exploré les propriétés, en échantillon fini, de l'estimateur du QMV en faisant des simulations de Monte-Carlo. Ces dernières ont, pour objectif, de montrer le gain en consistance de l'estimateur du QMV selon le type de proxy de volatilité qui est mis en oeuvre. Dans ce travail, nous ne considérons que deux types de proxy de volatilité, à savoir, les rendements logarithmiques journaliers (*close-to-close*) $H_n = |r_n|$ (représentant le cas classique), et la volatilité réalisée, $H_n = RV_n^{(m)}$, basée sur m intervalles –comme dans le cas de données intra-journalières, où la journée ouvrable est divisée en m intervalles de temps–, comme proxy de volatilité (c'est le cas de données de haute fréquence).

Dans un premier temps, nous avons simulé des réalisations à partir du modèle *GARCH* (1, 1) donné en (4.1.2b) où la variable aléatoire Z_n suit une loi de Student à 5 degrés de liberté pour les cinq cas de figures suivants : $\theta^0 = (\gamma^0, \beta^0) = (.05, .9)$, $(.15, .8)$, $(.35, .6)$, $(.25, .6)$, et enfin $\theta^0 = (.05, .8)$ (voir la première colonne du tableau ci-dessous). Le paramètre de normalisation κ a été fixé à une valeur qui est égale à l'unité ($\kappa = 1$). Les estimations du vecteur des

paramètres $\theta = (\gamma, \beta)$ par QMV ont été obtenues en utilisant les quatre tailles d'échantillons $N = 250, 500, 1000, 2500$. Le nombre d'exécutions de ces simulations est 2000. Afin d'évaluer, en échantillon fini, la qualité de l'approximation asymptotique de la variance de l'estimateur nous avons effectué l'expérience de Monte-Carlo suivante. Pour la vraie valeur du vecteur des paramètres θ^0 et pour une taille N donnée, d échantillons sont simulés ($d = 2000$), conduisant à d estimations de $\hat{\theta}_N^{(i)}$ de θ , $i = 1, \dots, 2000$. On note $\bar{\theta}_N = (\bar{\gamma}_N, \bar{\beta}_N)'$ leur moyenne empirique.

L'écart-type empirique des estimations de θ est noté $RMSE(\theta) = \left\{ \frac{1}{d} \sum_{i=1}^d \left(\hat{\theta}_N^{(i)} - \bar{\theta}_N \right)^2 \right\}^{\frac{1}{2}}$.

Passons maintenant à l'estimation des paramètres γ et β du modèle *GARCH* (1, 1), mais cette fois-ci en présence de données de haute fréquence. Nous avons commencé d'abord par la génération des réalisations à partir du processus des rendements intra-journaliers $R_n(u) = v_n \kappa \psi_n(u)$, $u \in [0, 1]$. Pour cela nous devons générer $(m + 1)$ réalisations équidistantes dans l'intervalle $[0, 1]$ du processus $\Psi_n(\cdot)$ pour chaque jour n . Pour ce faire, nous avons considéré un processus de diffusion intra-journalier, avec un processus d'Orstein-Uhlenbeck pour le log du coefficient diffusion : $d\Psi = \exp(Y(u)) dW^{(1)}(u)$, $u \in [0, 1]$ où $Y(u)$ est un processus d'Orstein-Uhlenbeck $dY(u) = -(Y(u) - \mu)du + \sigma_Y dW^{(2)}(u)$. À noter que les deux mouvements browniens $W^{(1)}$ et $W^{(2)}$ sont non corrélés. Nous avons choisi comme valeurs initiales $\Psi(0) = 0$ et $Y(0) = y_{(0)}$ où $y_{(0)}$ a été généré à partir de sa distribution stationnaire gaussienne de moyenne μ et de variance $(\sigma_Y^2/2\delta)$. Pour $\mu = -\sigma_Y^2/(2\delta)$, la variance réalisée $RV^{2(m)}(\Psi)$ pour chaque nombre d'intervalles m , ainsi que la variation quadratique sur un intervalle unitaire a une espérance qui est égale à l'unité (pour notre cas on prend $m = 81$). Nous avons choisi comme valeur du vecteur des paramètres (δ, σ_Y, μ) les valeurs $(\frac{1}{2}, \frac{1}{4}, -\frac{1}{16})$ respectivement. Les estimations du vecteur des paramètres $\theta = (\gamma, \beta)$ par QMV pour des tailles d'échantillon (250, 500, 1000, 2500) sont portées sur la deuxième colonne du tableau ci-dessous, à partir de 2000 répliques ainsi que leur *RMSE* correspondant.

les vraies des valeurs paramètres	H_n =			$ r_n $	H_n =			$RV_n^{(81)}$
	$N = 250$	$N = 500$	$N = 1000$		$N = 250$	$N = 500$	$N = 1000$	
γ^0 RMSE	0.064 (0.0474)	0.0769 (0.0169)	0.0749 (0.0251)	0.0711 (0.0343)	0.0315 (0.0239)	0.054 (0.0104)	0.0508 (0.0067)	0.0491 (0.0041)
	0.924 (0.0424)	0.9189 (0.0298)	0.9163 (0.02)	0.9151 (0.0138)	0.702 (0.0804)	0.8959 (0.013)	0.8994 (0.0088)	0.9013 (0.0054)
β^0 RMSE	0.1722 (0.0595)	0.1783 (0.0341)	0.1766 (0.0338)	0.1759 (0.0239)	0.161 (0.0398)	0.1624 (0.034)	0.1624 (0.0297)	0.162 (0.0255)
	0.8266 (0.0583)	0.8241 (0.034)	0.8233 (0.0338)	0.8217 (0.0239)	0.798 (0.0368)	0.7952 (0.0304)	0.7953 (0.0269)	0.7968 (0.0228)
γ^0 RMSE	0.334 (0.0741)	0.3274 (0.0525)	0.3297 (0.0269)	0.3331 (0.04)	0.3698 (0.0644)	0.3845 (0.0474)	0.3822 (0.0351)	0.3769 (0.0227)
	0.6656 (0.0735)	0.6726 (0.0525)	0.6703 (0.04)	0.6668 (0.0269)	0.6195 (0.0565)	0.6165 (0.416)	0.615 (0.0324)	0.6148 (0.0219)
β^0 RMSE	0.308 (0.0923)	0.3268 (0.0428)	0.3261 (0.0419)	0.3241 (0.0251)	0.2356 (0.0601)	0.2354 (0.0331)	0.2361 (0.0282)	0.2376 (0.0202)
	0.6762 (0.077)	0.6726 (0.0388)	0.6724 (0.037)	0.6706 (0.0243)	0.6459 (0.048)	0.6464 (0.0259)	0.6453 (0.0217)	0.6448 (0.0163)
γ^0 RMSE	0.0634 (0.064)	0.0669 (0.0483)	0.0668 (0.0341)	0.0658 (0.021)	0.0479 (0.0171)	0.0428 (0.0048)	0.0433 (0.0077)	0.0448 (0.011)
	0.864 (0.055)	0.8632 (0.04)	0.8629 (0.0288)	0.8625 (0.0182)	0.8045 (0.0246)	0.8099 (0.0164)	0.8091 (0.0116)	0.8078 (0.0073)

Table 4.4.1 : Résultats de l'estimation des paramètres du modèle $GARCH(1, 1)$ par la méthode du QMV gaussien, en utilisant les proxies $|r_n|$ et $RV_n^{(81)}$.

Dans le tableau (4.4.1) nous avons présenté les résultats de l'estimation des paramètres du modèle *GARCH* (1,1) par la méthode du QMV gaussien, en utilisant les proxies $|r_n|$ et $RV_n^{(81)}$ simulées. Nous pouvons remarquer que l'estimation des paramètres est améliorée en utilisant la proxy $RV_n^{(81)}$ au fur et à mesure que la taille de l'échantillon augmente (en contrepartie le RMSE correspondant diminue), malgré le nombre d'itérations qui n'est pas assez important ($d = 2000$).

4.4.2 Description des données

Le *S&P 500* (Standard & Poor's 500) est un indice boursier basé sur 500 valeurs américaines, choisies selon des critères de taille et de liquidité, cotées sur les trois bourses américaines suivantes : la bourse Nasdaq, le New York Stock Exchange (NYSE), et l'American Stock Exchange. L'indice *S&P 500* a été créé en 1920. Il a détrôné le Dow Jones Industrials Average (DJIA) comme l'indice le plus représentatif du marché boursier américain, parce qu'il est composé d'un plus grand nombre de compagnies et que sa valeur tient compte de la capitalisation boursière¹ des compagnies contenues dans l'indice.

La disponibilité récente de données financières à haute fréquence offre une opportunité pour étudier les propriétés statistiques des marchés de façon très fine.

Un *futures* ou contrat à terme est un contrat standardisé négocié sur un marché organisé permettant de s'assurer ou de s'engager sur un prix convenu pour une quantité déterminée d'un produit donné (le sous-jacent) à une date future. La particularité des *futures* est d'aboutir à la livraison physique du produit à l'échéance. Le Chicago Board of Trade (CBOT) et le Chicago Mercantile Exchange (CME) sont les deux marchés les plus significatifs aux États-Unis. Ils ont fusionné en 2007 et donné naissance au CME Group, le plus grand marché de contrats à termes et d'options du monde. En 1982, a été introduit un contrat futures sur l'indice Standard & Poor's (S&P) 500 stock index.

Dans ce travail, nous utilisons les prix intra-journaliers échantillonnés à 5 minutes de 8 : 30 AM jusqu'à 15 : 15 PM sur la période de 2004 à 2007 des *futures* (contrats à termes) de l'indice boursier *S&P 500*. Au total, nous avons 996 jours ouvrables, nous avons éliminé

¹Nombre d'actions dans le capital \times Cours = valeur boursière de la société

les valeurs correspondant à des jours de fermeture à 12 : 15 PM.

Nous avons choisi cette période (2004 – 2007) par ce qu'elle est stable et le niveau de volatilité de l'indice S&P 500 *futures* sur cette période ne contient pas de sauts (elle évolue de façon homogène), comme le montre la figure (4.4.1) ci-dessous.

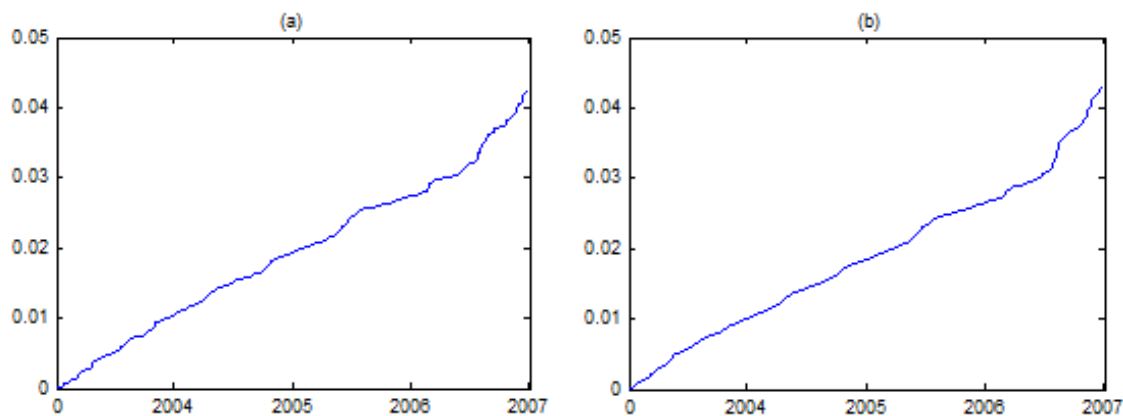


Figure 4.4.1 : (a) Cumul du carré des rendements journaliers et (b) les variances réalisées cumulées.

4.4.3 Mise en oeuvre du modèle

Dans cette sous-section, nous examinons l'efficacité des estimateurs dans le modèle *GARCH* (1, 1) donné en (4.1.1b), en utilisant les proxies de volatilité suivantes : les rendements *close-to-close* $|r_n^*|$, l'étendu des rendements max-min (hl_n) (nous utilisons la proxy de volatility de Parkinson), et enfin la volatilité réalisée basée sur 81 intervalles de 5 minutes. Cela est fait sans et avec le logarithme de la proxy H_n . Autrement dit, en utilisant la méthode du QMV gaussienne et celle log-gaussienne. Ces proxies sont construites à partir des données intra-journalières de l'indice boursier *S&P 500 futures* (voir la description des données ci-dessus)

L'efficacité des estimateurs du QMV gaussien (resp log-gaussien) est jugée à partir de la variance estimée des Z_H^2 (resp variance estimée du $\log(Z_H^2)$). Les résultats sont donnés dans le tableau (4.4.2) ci-dessous.

H_n	Gaussien $\widehat{var}(Z_H^2)$	Gain en efficacité	log-Gaussien $\widehat{var}(\log Z_H^2)$	Gain en efficacité
$ r_n^* $	3.435	1	—	—
hl_n	1.331	2.58	0.6	5.72
$RV5_n$	0.56	6.13	0.31	11.08

Table 4.4.2 : Variance estimée de Z_H^2 et $\log(Z_H^2)$.

Le gain en efficacité de l'estimateur du QMV gaussien (log-gaussien resp.) du modèle $GARCH(1,1)$ basé sur les trois proxies $|r_n^*|$, hl_n et $RV5_n$ ($\log|hl_n|$, $\log(RV5_n)$ resp.) est calculé par rapport à l'estimateur du QMV gaussien du modèle $GARCH(1,1)$ basé sur $|r_n^*|$, par exemple, $2.58 = 3.435/1.331$.

Les résultats de Tableau (4.4.2), nous montrent que l'estimation des paramètres du modèle $GARCH(1,1)$ en utilisant la proxy $RV5_n$, nous a permis de réaliser un gain en efficacité qui est égale à 6.13. Ce gain est plus important si nous utilisons la proxy logarithmique ($\log(RV5_n)$) (même constatation pour la proxy hl_n). Ainsi, l'estimateur du QMV log-gaussien est performant par rapport à celui du QMV gaussien. Cela peut être expliqué par la distribution des logarithmes des proxies (proxies logarithmiques) qui est approximativement gaussienne (Andersen et al (2001))

La figure (4.4.2) donne les courbes des proxies H_n (standardisées) que nous avons utilisées dans l'estimation du modèle $GARCH(1,1)$, ci-dessus. Il est remarquable qu'au fur et à mesure que l'on passe du graphique (a) au graphique (b), les courbes des proxies deviennent de moins en moins bruitées.

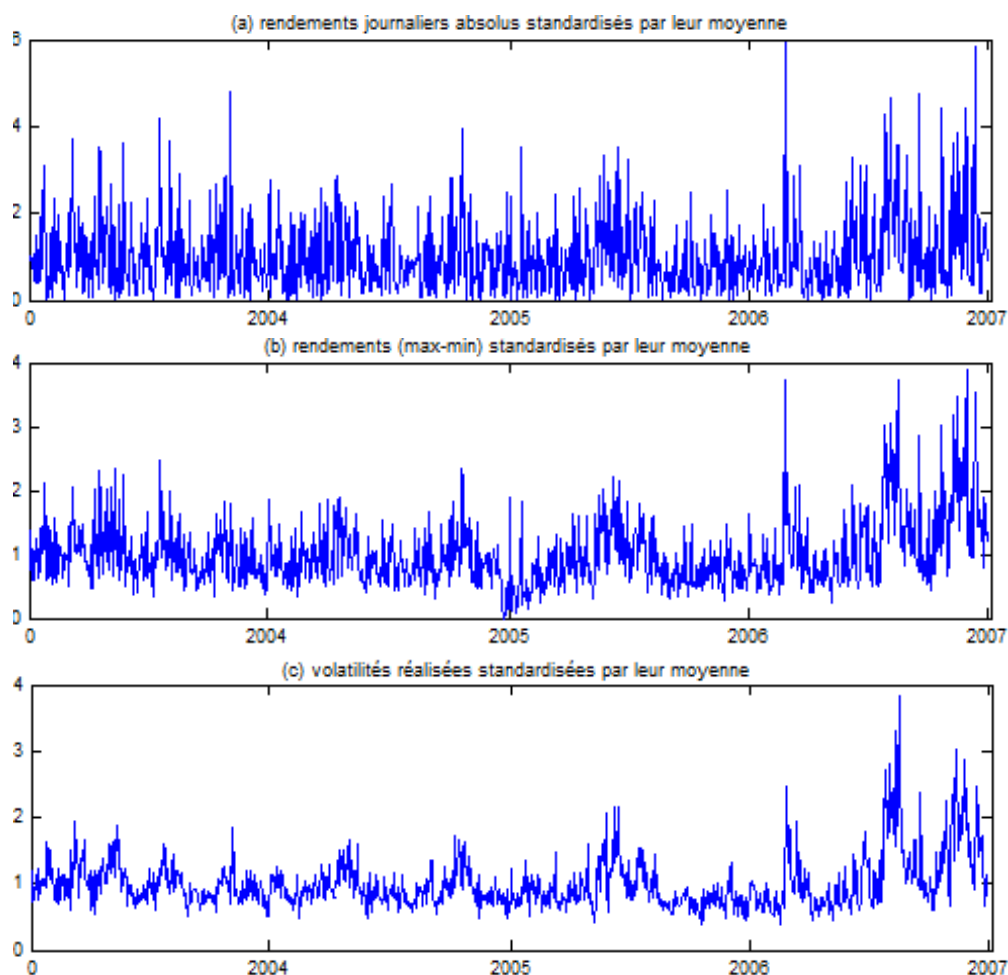


Figure 4.4.2 : Les trajectoires des proxies H_n ($|r_n^*$, hl_n et $RV5_n$ respectivement) standardisées par leur moyenne \overline{H}_n .

4.5 Semimartingales, variation quadratique et volatilité réalisée

Ces dernières décennies, la modélisation de la volatilité en utilisant les données de haute fréquence a connu un grand développement en finance. L'approche la plus utilisée est de supposer que le processus des prix logarithmiques est une semimartingale. Dans la théorie des semimartingales, la variation quadratique joue un rôle central. Elle est interprétée comme le carré de la volatilité. Comme nous l'avons vu dans le chapitre 3, la variation quadratique est la limite de la somme des carrés des incréments des rendements logarithmiques intrajournaliers, au fur et à mesure que le pas d'échantillonnage tend vers zéro.

Le modèle d'échelle donnée en (4.1.2a) est une semimartingale si le processus $\Psi(\cdot)$ l'est aussi. Par exemple, si le processus $\Psi(\cdot)$ est un mouvement brownien standard, alors sa variation quadratique sera égale à l'unité, *i.e.*, $\langle \Psi, \Psi \rangle = 1$, et par suite $\langle R_n, R_n \rangle = v_n^2 \kappa^2$. Néanmoins, nous n'avons pas toujours cette relation exacte.

La variation quadratique et le facteur d'échelle σ_n mesurent tous les deux les fluctuations de processus des prix logarithmiques intra-journaliers, mais sont deux concepts différents. Le facteur d'échelle est une variable latente qui détermine les fluctuations espérées de processus des prix, au cours de la journée ouvrable, tandis que la variation quadratique est liée aux fluctuations des trajectoires de ces processus, au cours de la journée ouvrable.

La variation quadratique est un estimateur non biaisé du carré du facteur d'échelle σ_n

$$\mathbb{E}(\langle R_n, R_n \rangle \mid \mathcal{F}_{n-1}) = \text{var}(r_n \mid \mathcal{F}_{n-1}) = v_n^2 \kappa^2.$$

Il est à noter que les données de haute fréquence ne permettent pas de calculer la variation quadratique à n'importe quel degré de précision. En pratique, on est confronté aux effets de microstructures de manière que l'utilisation des pas d'échantillonnage au dessous de certaines valeurs (d'ordre de 5 minutes, pour les marchés les plus liquides) n'améliore pas les estimations.

Pour comparer entre l'évolution de la volatilité σ_n (le facteur d'échelle) issue d'un processus *GARCH* et celle mesurée en utilisant la volatilité réalisée $RV5_n$ basée sur les rendements intra-journaliers de 5 minutes, nous avons calculé le cumul des carrés des $\hat{\sigma}_n$ (issus du modèle *GARCH*(1, 1)) et des $RV5_n$ à partir des données journalières et intra-journalières de l'indice S&P 500 *futures*, dans un premier temps (figure (4.5.1a)) et par simulation dans un deuxième temps (figure (4.5.1b)).

Pour les simulations, nous avons procédé comme suit. Nous avons simulé 996 réalisations issues d'un processus *GARCH*(1, 1), pour les valeurs suivantes des paramètres : $w = 2.04 \times 10^{-6}$, $\alpha = 0.051215$ et $\beta = 0.901326$, dont l'écriture est donnée en (3.5.7). La simulation de la variance réalisée a été faite en suivant les mêmes étapes décrites dans la section (4).

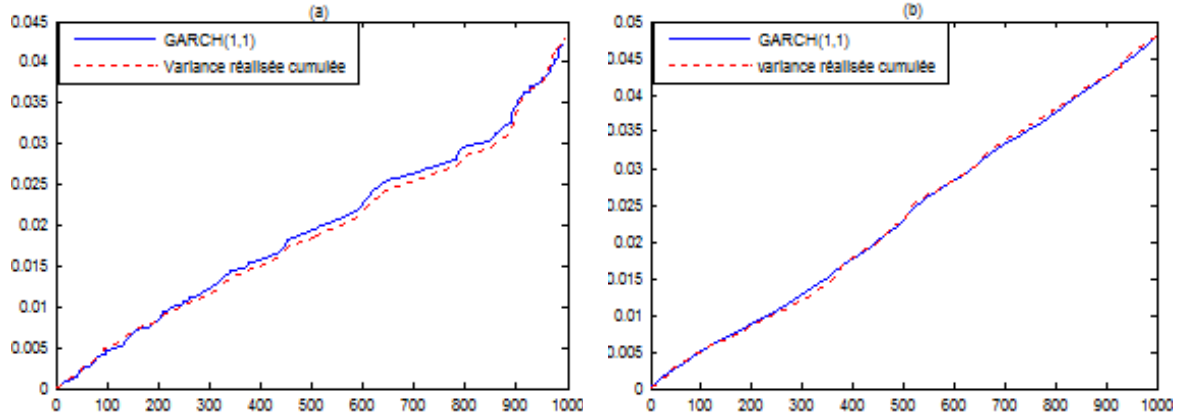


Figure 4.5.1 : Évolution cumulative de carré de facteur d'échelle $\hat{\sigma}_n$ et de la variance réalisée $RV5_n^{(2)}$ appliquée sur l'indice S&P 500 (a) et simulés (b).

Les graphes de la figure (4.5.1) ci-dessus, montrent une certaine ressemblance dans leur évolution. Cela se voit clairement, où les deux séries $\hat{\sigma}_n^2$ et $RV5_n^2$ évoluent de la même manière.

Volatilité (réalisée) et périodicité intra-journalières de l'indice S&P 500 futures

Nous terminons cette section par quelques visualisations graphiques des courbes de volatilité et de périodicité intra-journalières de l'indice S&P 500 futures pour l'année 2004 (cette dernière contient 248 jours ouvrables).

Les rendements intra-journaliers absolus moyens ($\frac{1}{248} \sum_{n=1}^{248} |r_{n,k}|$, $k = \overline{1, 81}$) ainsi que les variances réalisées intra-journalières moyennes sont caractérisés par un caractère régulier remarquable. C'est la forme *U* : la volatilité est plus prononcée à l'ouverture et à l'approche de fermeture des marchés financiers et se stabilise progressivement au cours de la journée ouvrable.

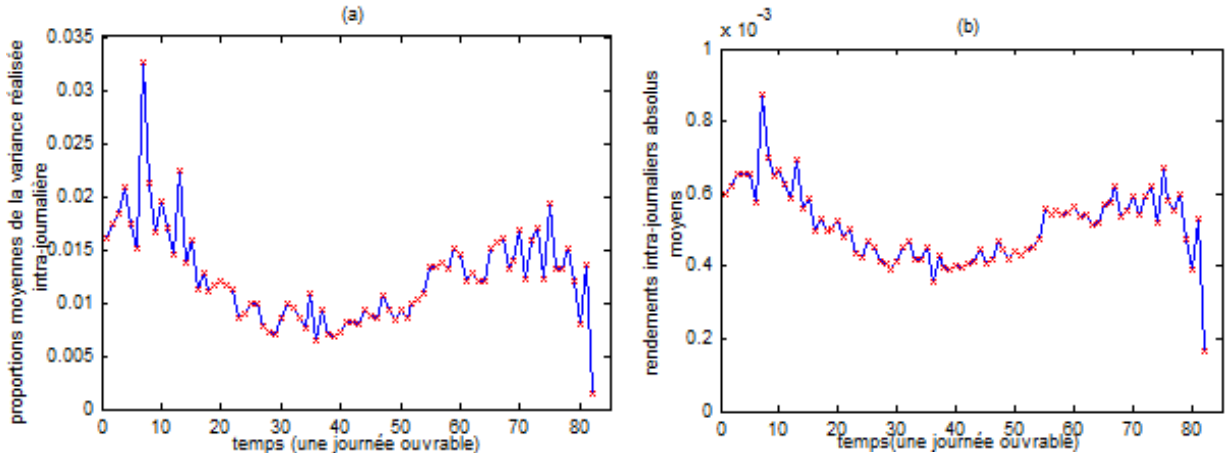


Figure 4.5.2 : La forme de la volatilité intra-journalière de l'indice S&P 500 futures sur l'année 2004.

Pour construire la variance réalisée intra-journalière moyenne, nous avons procédé comme suit :

Pour chaque jour ouvrable n ($n = 1, \dots, 248$), nous calculons :

1- Les rendements intra-journaliers de 5 minutes : $r_{n,k} = R_n(k\Delta) - R_n((k-1)\Delta)$, où Δ désigne le pas d'échantillonnage égale à 5 minutes.

2- Les variances réalisées : $RV_n^{(2)} = \sum_{k=1}^{81} r_{n,k}^2$.

3- Les proportions $\frac{r_{n,k}^2}{RV_n^{(2)}}$, $k = 1, \dots, 81$.

4- La proportion moyenne de la variance réalisée attribuée à chaque intervalle de 5 minutes. Autrement dit, nous calculons la quantité : $\frac{1}{248} \sum_{n=1}^{248} \frac{r_{n,k}^2}{RV_n^{(2)}}$, $\forall k = 1, \dots, 81$.

La courbe de la variance réalisée intra-journalière moyenne cumulée tracée sur la figure (4.5.3), ci-dessous, nous renseigne sur la manière d'évolution de la volatilité sur une journée ouvrable donnée. La croissance lente au niveau de la partie moyenne de cette courbe reflète les valeurs modérées dans le graphique de la figure (4.5.2a) entre le 18^{ème} intervalle (10 : 00 A.M) et 66^{ème} intervalle (2 : 00 P.M).

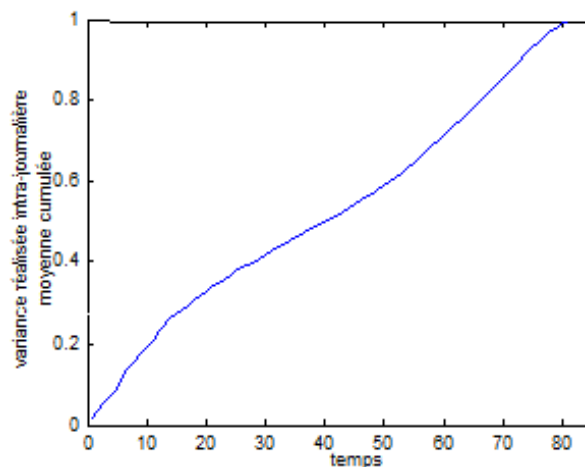


Figure 4.5.3 : La variance réalisée intra-journalière moyenne cumulée

Maintenant, retournons à la dépendance temporelle des rendements intra-journaliers (de 5 minutes) de l'indice S&P 500 *futures*. Le tracé de la figure (4.5.4) montre le corrélogramme (autocorrélations empiriques) des rendements absolus de 5 minutes, pour un retard $l = 810$ (10 jours ouvrables).

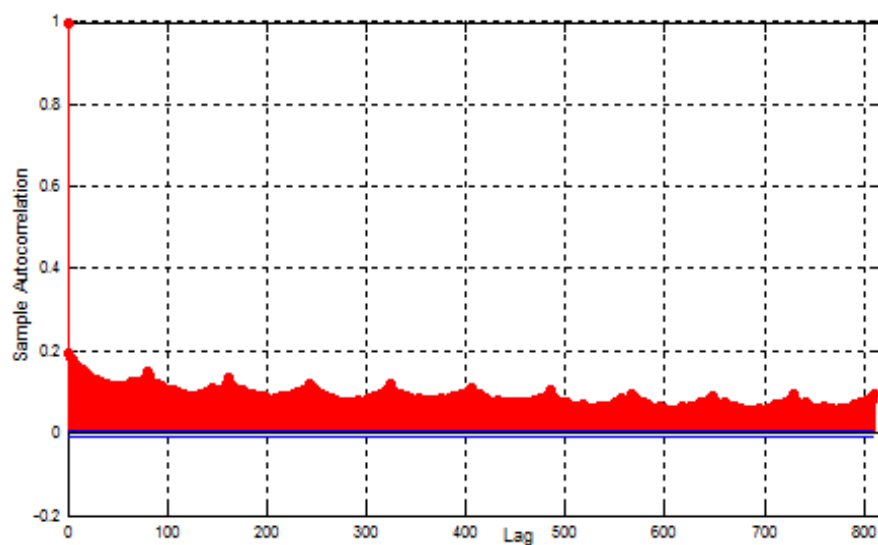


Figure 4.5.4 : Périodicité intra-journalière de l'indice S&P 500 *futures* des dix premiers jours ouvrables

Chaque jour ouvrable est caractérisée par une périodicité très saillante, où la forme U est reproduite. Ainsi, les modèles *GARCH* et les modèles de volatilité stochastique ne sont pas directement applicables, (Andersen et Bollerslev (1997)).

4.6 Conclusion

Dans ce chapitre nous avons étudié les propriétés statistiques asymptotiques de l'estimateur du modèle $GARCH(1, 1)$ par la méthode du QMV gaussien (log-gaussien resp.), basé sur les proxies H_n ($\log(H_n)$ resp.). Une étude de simulation a été faite afin de mettre en évidence les résultats théoriques.

Pour voir le gain en efficacité de l'estimateur du QMV (gaussien et log gaussien) du modèle $GARCH(1, 1)$, si une proxy adéquate (H_n et $\log(H_n)$) est utilisée (en exploitant les données journalières et intra-journalières (à haute fréquence)), nous avons appliqué cette méthode aux données financières, à savoir la série S&P 500 *futures*.

Conclusion générale

*il n'y a pas des problèmes qu'on se pose,
il y a des problèmes qui se posent, il n'y a pas de problèmes résolus,
il y a des problèmes plus ou moins résolus*

** Henri Poincaré***

L a modélisation des faits stylisés des séries financières (regroupements des volatilités, non corrélation des rendements, distributions à queues épaisses, etc.) nécessite l'utilisation de modèles relativement complexes.

L'apparition des modèles *ARCH/GARCH* (Engle (1982) et par la suite généralisé par Bollerslev en 1986) au cours de la décennie quatre-vingts, traduit le fait que l'on s'est rendu compte que les moments du second ordre (conditionnels) des séries temporelles issues des marchés financiers dépendent du temps et les chercheurs en finance ont donc commencé à modéliser cette dépendance.

Dans ce travail, nous nous sommes intéressées à la modélisation de la volatilité journalière des rendements des actifs financiers par le modèle GARCH en présence de données de haute fréquence. Ainsi, l'introduction d'un modèle en temps continu semble naturelle pour tenir compte des mouvements intra-journaliers des processus des prix. Ce modèle est dit modèle d'échelle. Comme le facteur d'échelle σ_n (la racine carré de la variance conditionnelle des rendements journaliers) est latent, une classe de statistiques basées sur les données intra-journalières est introduite (variance réalisée, par exemple).

La variance réalisée trouve sa justification théorique dans la théorie de la variation quadratique, où cette dernière suggère que le processus sous considération est une semimartingale.

De ce fait, nous avons testé l'hypothèse de semimartingale du processus des prix de l'indice S&P 500 *futures* où cette hypothèse est acceptée.

Dans le cadre de l'étude des données intra-journalières, nous avons montré que l'estimateur de la méthode du quasi-maximum de vraisemblance de modèle *GARCH* (1, 1) qui utilise la proxy de volatilité $RV5_n$ qui est construite en additionnant les rendements intra-journaliers de 5 minutes, nous a permis de réaliser un gain en efficacité et cela se manifeste plus clairement en l'appliquant sur les proxies logarithmiques et la même constatation est confirmée par une étude de simulation intensive.

Glossaire

Allocation d'actifs : Répartition d'un portefeuille en fonction des actifs le composant. Le gestionnaire divise son portefeuille en actions, obligations, produits dérivés, etc.

Arbitrage : Stratégie profitant d'incohérences momentanées entre des prix de plusieurs actifs ou contrats.

bid-ask : Terme anglo-saxon désignant la fourchette de cotations d'un titre entre le meilleur vendeur et le meilleur acheteur. Le bid désigne l'offre, à savoir la meilleure position de vente. L'ask désigne la demande à savoir le meilleur prix d'achat. Moins un titre est liquide, plus la fourchette bid-ask est large est moins le prix reflète la réalité du marché.

Call : Contract donnant le droit d'acheter un actif à une date future donnée (l'échéance) et à un prix convenu (prix d'exercice).

Carnet d'ordre : Carnet où sont regroupés l'ensemble des ordres d'achat et de vente concernant une valeur, ainsi que les derniers ordres exécutés. Les investisseurs ne peuvent avoir accès qu'aux 5 meilleures offres (ventes) et aux 5 meilleures demandes (achats), mais le carnet d'ordre d'une valeur peut comprendre plusieurs milliers d'ordres.

Contract *forward* : est comme un contract futures, un engagement ferme à acheter ou à vendre un actif (sous-jacent) à une date future donnée pour un prix convenu. Il est échangé le plus souvent entre deux établissements financiers ou entre un établissement et un client. À la différence des contracts futures, les contracts forward sont négociés sur des marchés de gré à gré (Over The Counter (OTC)), où les échanges sont conclus par téléphone ou par l'intermédiaire de réseaux informatiques.

Contract *futures* : est un accord entre deux parties pour acheter ou vendre un actif donné à une date future pour un prix convenu dit "prix *futures*".

Cotation : Fixation d'un cours d'une valeur suite à l'échange de titres. Pour qu'il y ait cotation, un acheteur doit trouver un vendeur, c'est-à-dire une contrepartie. Une fois que les deux protagonistes se sont mis d'accord sur le prix de transaction, la valeur est échangée. Ce prix de transaction est donc la cotation de la valeur à l'instant t .

Cours : Terme désignant la valeur d'un titre à un instant t . Le cours peut évoluer en permanence sur les marchés en continu. Le cours d'un titre s'explique par la rencontre entre

l'offre (les vendeur) et la demande (les acheteurs). L'équilibre entre l'offre et la demande vient de leur accord sur un prix de transaction. Ce prix d'équilibre désigne le cours.

Échéance : Fin de vie d'un contrat.

Fourchette : L'écart entre le prix demandé le plus élevé et le prix offert le plus bas. Plus cette fourchette est importante, plus on peut qualifier le marché d'étroit. Ainsi, sur un marché très liquide, cette fourchette sera extrêmement mince, alors sur un marché peu liquide, l'offre et la demande auront du mal à se rencontrer et donc à conclure une transaction.

Intraday : Achat et vente réalisée sur une même journée.

Liquide : Se dit d'un marché ou d'un titre où les échanges sont nombreux, la fourchette de cotations faible où il est facile d'acheter et de vendre ses titres. Plus un marché est liquide, plus la valorisation de ces actifs est conforme à la réalité. En effet, à chaque instant, le nombre d'acheteurs et de vendeur étant élevé. Le prix d'équilibre se rapprochera du prix théorique.

Marché de gré à gré : Marché où les transactions s'effectuent de gré à gré, en interaction bilatérale, c'est-à-dire directement entre le vendeur et l'acheteur, à l'opposé du marché organisé.

Marché organisé : Marché où les transactions sont organisées en interaction multilatérale, à l'aide d'un carnet d'ordre, et où les contrats sont standardisés.

Market-maker (ou teneur de marché) : Opérateur de marché qui fournit à tout instant des prix à l'achat et à la vente des actifs.

Option : Est un produit dérivé qui établit un contrat entre un acheteur et un vendeur, donnant le droit, et non l'obligation, d'acheter ou de vendre un actif dit sous-jacent (action, indice, etc.) à un cours déterminé à l'avance (strike ou prix d'exercice), pendant un temps donné ou à une date fixée (date d'échéance).

Prix ask : Prix offert le plus élevé auquel il est possible de vendre un actif sur le marché.

Prix bid : Prix demandé le plus faible auquel il est possible d'acheter un actif sur le marché.

Prix d'exercice : Prix auquel on peut acheter ou vendre l'actif sous-jacent en exerçant une option.

Prix spot : Prix pour une livraison immédiate.

Produit dérivé : Instrument financier dont le prix est dérivé du prix d'un autre actif. Un produit dérivé donne un droit d'agir sur un sous-jacent.

Put : Contract donnant le droit de vendre un actif à une date future donnée (l'échéance) et à un prix convenu (prix d'exercice).

Séance : Désigne la période entre l'ouverture et la clôture de la journée boursière. Ainsi, lorsque l'on dit que le titre a gagné 3% sur la séance, on compare le cours de clôture de la séance du jour par rapport au cours de clôture de la séance du veille. Pour les valeurs cotées en continu, un titre peut être échangé durant toute la séance.

Smile : variation de la volatilité implicite en fonction du prix d'exercice.

Sous-jacent : Le sous-jacent correspond à l'action, l'indice, devise, matière première, etc. sur lequel est basé l'option par exemple (le produit dérivé)..

Tick-by-tick : Se dit de données intraday, sur une seule séance. L'investisseur peut ainsi établir des graphiques en temps réel. Chaque transaction sur le marché, se traduit par un nouvel point dans le graphique et ainsi l'investisseur est auprès de l'évolution des marchés. Ce suivi en temps réel des marchés est très chronophage et n'est destiné qu'à certains investisseurs.

Titre : Terme désignant une valeur, une action ou toute autre valeur mobilière.

Valeurs : Terme générique dont les définitions sont nombreuses. La valeur peut ainsi désigner la valorisation d'une entreprise, ou encore être synonyme d'un titre ou d'une action.

Volatilité implicite : volatilité déduite de la valeur d'une option en utilisant le modèle de Black Scholes ou un modèle semblable.

Trading day : Expression anglo-saxonne désignant le temps pendant lequel un marché est ouvert aux transactions. Cette expression est donc parfaitement contraire à l'*overnight*. Durant le trading day, les investisseurs peuvent se positionner sur le marché ou bien déboucler leur position en fonction des conditions du marché. Certes ils sont exposés aux risques du marché mais ils sont en mesure de s'en dégager rapidement en débouclant leur position.

Structure d'autocovariance les carrés du processus GARCH

Soit h l'ordre de retard.

- Pour $h = 0$, nous avons

$$\begin{aligned}
 \mathbb{E}(X_n^{\otimes 2}) &= \mathbb{E}\{(B_n + A_n X_{n-1}) \otimes (B_n + A_n X_{n-1})\}, \\
 &= \mathbb{E}\{(B_n \otimes B_n) + (B_n \otimes A_n X_{n-1}) + (A_n X_{n-1} \otimes B_n) + (A_n X_{n-1} \otimes A_n X_{n-1})\}, \\
 &= \{\mathbb{E}(B_n^{\otimes 2}) + \mathbb{E}(B_n \otimes A_n X_{n-1}) + \mathbb{E}(A_n \otimes B_n X_{n-1}) + \mathbb{E}(A_n^{\otimes 2} X_{n-1}^{\otimes 2})\}, \\
 &= \mathbb{E}(B_n^{\otimes 2}) + \mathbb{E}(B_n \otimes A_n) \mathbb{E}(X_{n-1}) + \mathbb{E}(A_n \otimes B_n) \mathbb{E}(X_{n-1}) + \mathbb{E}(A_n^{\otimes 2} X_{n-1}^{\otimes 2}).
 \end{aligned}$$

Nous considérons les notations suivantes :

$$A^{(m)} = \mathbb{E}(A_n^{\otimes m}), \quad X^{(m)} = \mathbb{E}(X_n^{\otimes m}) \text{ et } B^{(m)} = \mathbb{E}(B_n^{\otimes m}).$$

D'où

$$\mathbb{E}(X_n^{\otimes 2}) = B^{(2)} + (\mathbb{E}(B_n \otimes A_n) + \mathbb{E}(A_n \otimes B_n)) X^{(1)} + A^{(2)} \mathbb{E}(X_n^{\otimes 2}),$$

et nous obtenons

$$\mathbb{E}(X_n^{\otimes 2}) = \left(I_{(p+q)^2} - A^{(2)} \right)^{-1} \{ B^{(2)} + (\mathbb{E}(B_n \otimes A_n) + \mathbb{E}(A_n \otimes B_n)) X^{(1)} \}. \quad (2.3.16)$$

Dans ce qui suit, nous montrons comment calculer chaque terme du membre de droite de (2.3.16).

Pour le calcul de $A^{(m)}$, nous pouvons utiliser la décomposition $A_n = Z_n^2 C + D$, où C et D sont des matrices déterministes de même dimension données par

$$C = \left(\begin{array}{ccc|ccc} \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ \hline & O_{(p+q-1) \times (q)} & & & O_{(p+q-1) \times (p)} & \end{array} \right),$$

$$D = \left(\begin{array}{ccc|ccc} & O_{1 \times (q)} & & & O_{1 \times (p)} & \\ & I_{(q-1)} & O_{(q-1) \times 1} & & O_{(q-1) \times p} & \\ \hline \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ & O_{(p-1) \times q} & & & I_{(p-1)} & O_{(p-1) \times 1} \end{array} \right)$$

de plus, on pose $\mu_m = \mathbb{E}(Z_n^m)$, nous obtenons

$$A^{(2)} = \mathbb{E}(Z_n^2 C + D)^{\otimes 2} = \mu_4 C^{\otimes 2} + C \otimes D + D \otimes C + D^{\otimes 2}.$$

Nous procédons de la même manière pour le calcul de $B^{(m)}$, *i.e.*, nous posons $B_n = Z_n^2 F + G$, où F et G sont des vecteurs déterministes (toujours de même dimension) donnés par

$$F = (\omega \ 0 \ \dots \ 0)', \quad G = (0 \ \dots \ \omega \ 0 \ \dots \ 0)'$$

Ainsi, on obtient pour $m = 2$

$$B^{(2)} = \mathbb{E}(Z_n^2 F + G)^{\otimes 2} = \mu_4 F^{\otimes 2} + F \otimes G + G \otimes F + G^{\otimes 2}.$$

Nous obtenons de même $\mathbb{E}(B_n \otimes A_n)$ et $\mathbb{E}(A_n \otimes B_n)$.

Pour $X^{(1)}$, on remarque que toutes ses composantes valent $\omega / \left(1 - \sum_{i=1}^{\max(p+q)} (\alpha_i + \beta_i)\right)$.

- Pour $h \geq 1$, nous aurons

$$\begin{aligned} \mathbb{E}(X_n \otimes X_{n-h}) &= \mathbb{E}\{(B_n + A_n X_{n-1}) \otimes X_{n-h}\}, \\ &= \mathbb{E}\{(B_{n+1} \otimes X_{n-h+1}) + (A_{n+1} X_n \otimes I_{p+q} X_{n-h+1})\}, \\ &= (B^{(1)} \otimes X^{(1)}) + \mathbb{E}(A_{n+1} \otimes I_{p+q}) \mathbb{E}(X_n \otimes X_{n-h+1}), \\ &= (B^{(1)} \otimes X^{(1)}) + (A^{(1)} \otimes I_{p+q}) \mathbb{E}(X_n \otimes X_{n-h+1}). \\ &= (B^{(1)} \otimes X^{(1)}) + (A^{(1)} \otimes I_{p+q}) \mathbb{E}(X_n \otimes X_{n-h+1}). \end{aligned} \tag{2.3.17}$$

Françq et Zakoïan ont proposé l'algorithme suivant pour le calcul des autocovariances γ_{ε^2} d'ordre h du processus (ε_n^2) (Françq et Zakoïan (2009) p 65).

Soit \underline{e}_1 le premier vecteur de la base canonique de $\mathbb{R}^{(p+q)^2}$.

- Définir les vecteurs $X^{(1)}$, $B^{(1)}$, $B^{(2)}$ et les matrices $A^{(1)}$, $A^{(2)}$, $\mathbb{E}(B_n \otimes A_n)$, $\mathbb{E}(A_n \otimes B_t)$ en fonction des α_i , β_i et de ω , μ_4 ;
- Calculer $(\mathbb{E}(X_n^{\otimes 2}))$ à l'aide de (2.3.16);
- Pour $h = 1, 2, \dots$, on calcule $\mathbb{E}(X_n \otimes X_{n-h})$ en utilisant (2.3.17);
- Pour $h = 0, 1, \dots$, on obtient $\gamma_{\varepsilon^2}(h) = \underline{e}_1 \mathbb{E}(X_n \otimes X_{n-h}) - (\underline{e}_1 X^{(1)})^2$.

Exemple

Dans ce qui suit, nous illustrons l'algorithme précédent pour le cas d'un $GARCH(1, 1)$. Pour ce faire, nous considérons l'écriture vectorielle donnée en (2.3.2).

Nous avons

$$X^{(1)} = \frac{\omega}{1-\alpha-\beta} \begin{pmatrix} 1 & 1 \end{pmatrix}', B^{(1)} = \begin{pmatrix} \omega & \omega \end{pmatrix}' \text{ et } A^{(1)} = \begin{pmatrix} \alpha & \beta \\ \alpha & \beta \end{pmatrix},$$

$$\mathbb{E}(B_n \otimes A_n) = \mathbb{E}(A_n \otimes B_n) = \omega \begin{pmatrix} \mu_4 \alpha & \alpha & \alpha & \alpha \\ \mu_4 \beta & \beta & \beta & \beta \end{pmatrix}',$$

$$A^{(2)} = \begin{pmatrix} \mu_4 \alpha^2 & \mu_4 \alpha \beta & \mu_4 \alpha \beta & \mu_4 \beta^2 \\ \alpha^2 & \alpha \beta & \alpha \beta & \beta^2 \\ \alpha^2 & \alpha \beta & \alpha \beta & \beta^2 \\ \alpha^2 & \alpha \beta & \alpha \beta & \beta^2 \end{pmatrix}, \text{ avec } C = \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix} \text{ et } D = \begin{pmatrix} 0 & 0 \\ \alpha & \beta \end{pmatrix},$$

$$B^{(2)} = \omega^2 \begin{pmatrix} \mu_4 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \text{ avec } F = \begin{pmatrix} \omega \\ 0 \end{pmatrix} \text{ et } G = \begin{pmatrix} 0 \\ \omega \end{pmatrix}.$$

Nous avons

$$B^{(2)} + (\mathbb{E}(B_n \otimes A_n) + \mathbb{E}(A_n \otimes B_n)) X^{(1)} = \omega^2 \frac{1+\alpha+\beta}{1-\alpha-\beta} \begin{pmatrix} \mu_4 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

La solution du système (2.3.16) est

$$X^{(2)} = \frac{\omega^2(1+\alpha+\beta)}{(1-\alpha-\beta)(1-3\alpha^2-2\alpha\beta-\beta^2)} \begin{pmatrix} \mu_4 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ainsi, pour $h = 0$ la fonction d'autocovariance $\gamma_{\varepsilon^2}(0) = \frac{\omega^2(1+\alpha+\beta)}{(1-\alpha-\beta)(1-3\alpha^2-2\alpha\beta-\beta^2)} \mu_4 - \left(\frac{\omega}{1-\alpha-\beta}\right)^2$.

Quand l'ordre h est supérieur ou égal à l'unité l'équation (2.3.17) s'écrit

$$\mathbb{E}(X_n \otimes X_{n-h}) = \frac{\omega^2}{1-\alpha-\beta} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \alpha & 0 & \beta & 0 \\ 0 & \alpha & 0 & \beta \\ \alpha & 0 & \beta & 0 \\ 0 & \alpha & 0 & \beta \end{pmatrix} \mathbb{E}(X_n \otimes X_{n-h+1}),$$

et cette dernière écriture nous permet de calculer la fonction d'autocovariance $\gamma_{\varepsilon^2}(\cdot)$.

Espace de Skorohod

Soit un ensemble non vide; on appelle topologie sur X une famille \mathcal{C} de parties de X , appelées ouverts telles que

- toute réunion d'ouverts est un ouvert
- toute intersection finie d'ouverts est un ouvert
- \emptyset et X sont des ouverts

Un exemple simple est celui où $\mathcal{C} = \{\emptyset, X\}$ mais cette topologie, appelée d'ailleurs topologie *grossière*, n'a aucun intérêt car elle a trop peu d'ouverts.

Définition : 1

\mathcal{C} est dite séparée si

$$a \neq b \implies \exists U, V \in \mathcal{C}; a \in U, b \in V, U \cap V = \emptyset.$$

- Une topologie séparée est donc une topologie qui suffisamment d'ouverts pour distinguer les points de X .
- Si \mathcal{C} est une topologie sur X , le couple (X, \mathcal{C}) s'appelle un espace topologique. Souvent, on sous-entend \mathcal{C} et on parle de l'espace topologique X .
- Un espace topologique (X, \mathcal{C}) est dit compact si tout recouvrement ouvert de X contient un sous-recouvrement fini. De façon symbolique

$$\forall (\omega_i) \in \mathcal{C}, X = \bigcup_I \omega_i \implies \exists J \subset I, J \text{ fini}; X = \bigcup_J \omega_i$$

- Un espace X est dit séparable s'il contient une partie dénombrable dense. L'ensemble \mathbb{R} des réels avec la topologie des réunions d'intervalles ouverts, est séparable puisque l'ensemble \mathbb{Q} des rationnels est dénombrable et dense dans \mathbb{R} .

Soit (Ω, \mathcal{F}, P) un espace probabilisé et E muni d'une distance. Si (Y^n) et Y sont des variables aléatoires à valeurs dans E , Il existe différentes modes de convergence de la suite (Y^n) vers la variable Y tels que la convergence presque sûre, la convergence en probabilité et la convergence en loi.

Si maintenant (Y^n) et (Y) sont des processus, on peut évidemment parler de la convergence dans l'un des sens précédemment cité de la suite de variables (Y_t^n) vers la variable Y_t pour un temps t fixé. Mais cette convergence ponctuelle ne rend pas compte du comportement asymptotique de la trajectoire globale des suites (Y^n) . Pour contourner ce problème,

nous avons besoin de définir :

- un espace de fonctions : nous prenons l'espace des processus càdlàg, à valeurs dans E . nous nommons cet espace $\mathbb{D}(E)$
- d'une distance dans cet espace, donc $\mathbb{D}(E)$ sera muni de sa tribu borélienne.

Ainsi, nous pouvons, parler de convergence de la suite de processus (Y^n) vers le processus (Y) .

D'abord rappelons ce que un espace séparable.

Définition : 2

Soit E un espace metrisable et séparable. Nous disons que E est un espace polonais si sa topologie peut être définie par une distance qui en fait un espace complet.

Soit E un espace polonais. L'espace des fonctions càdlàg de $\mathbb{R}^+ \rightarrow E$ est appelé espace de Skorohod. Nous le notons par $\mathbb{D}(E)$.

Les proxies de volatilité : $V_P, V_{GK}, V_{RS}, V_{ABD}$, et V_{YZ} .

Considérons les notations suivantes :

O_i : prix d'ouverture de la $i^{\text{ème}}$ journée ouvrable,

C_i : prix de clôture de la $i^{\text{ème}}$ journée ouvrable,

H_i : prix le plus haut de la $i^{\text{ème}}$ journée ouvrable,

L_i : prix le plus bas de la $i^{\text{ème}}$ journée ouvrable,

$$o_i = \ln (O_i / C_{i-1}),$$

$$u_i = \ln (H_i / O_i),$$

$$d_i = \ln (L_i / O_i),$$

$$c_i = \ln (C_i / O_i).$$

Nous définissons les proxies de volatilité (le carré de volatilité) suivantes, à savoir, V_P , V_{GK} , V_{RS} , V_{ABD} , et V_{YZ} .

$$V_P = \frac{1}{n} \sum_{i=1}^n \frac{1}{4 \ln(2)} [\ln(H_i) - \ln(L_i)]^2,$$

$$V_{GK} = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{2} \ln(H_i / L_i)^2 - (2 \ln(2) - 1) (\ln(C_i / O_i))^2 \right],$$

$$V_{RS} = \frac{1}{n} \sum_{i=1}^n [\ln(H_i / O_i) \ln(H_i / C_i) + \ln(L_i / O_i) \ln(L_i / C_i)],$$

$$V_{ABD} = \frac{1}{n} \sum_{i=1}^n [\ln(H_i) - \ln(L_i)],$$

$$V_{YZ} = V_o + k V_C + (1 - k) V_{RS},$$

où $V_o = \frac{1}{n-1} \sum_{i=1}^n (o_i - \bar{o})^2$, $V_c = \frac{1}{n-1} \sum_{i=1}^n (c_i - \bar{c})^2$, avec $\bar{o} = \frac{1}{n-1} \sum_{i=1}^n o_i$ et $\bar{c} = \frac{1}{n-1} \sum_{i=1}^n c_i$. La valeur du paramètre k , pour laquelle la proxy de volatilité V_{YZ} est optimisée (à une variance minimale) est $k_0 = \frac{0.34}{1.34 + (n+1)/(n-1)}$.

n : désigne la taille de l'échantillon.

Bibliographie

- [1] Aït-Sahalia, Y, Mykland, P.A et Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise”, *Review of Financial Studies*. 18, p. 351-416
- [2] Aït-Sahalia, Y., Mykland, P. A., Zhang, L. (2006). Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise. *Working Paper*, w11380, NBER.
- [3] Alizadeh ,S., Brandt, W. M., et Diebold, X.F. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance* 57 : 1047-1091
- [4] Andersen, T.G., Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, Vol. 4, Issues 2 – 3, pp 69 – 293.
- [5] Andersen, T.G. et Bollerslev, T. (1998) Answering the skeptics : yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885 – 906.
- [6] Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2009). Parametric and nonparametric volatility measurement. *In Handbook of Financial Econometrics* (eds Y. Aït-Sahalia et L. Hansen), Handbooks in Finance, volume 1, chapter 2. Forthcoming.
- [7] Andersen, T.G., Bollerslev, T., Diebold, F.X., et Labys, P. (2000). Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian. *Multinational Finance Journal*, 4, 159 – 179.
- [8] Andersen, T. G., Bollerslev, T., Diebold, F. X., Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics* 61 :43–76.
- [9] Andersen, T.G., Bollerslev, T., Diebold, F.X. et Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71, number 2, 579 – 625.
- [10] Andrews, D.W.K. (1999). Estimation when a parameter is on a boundary. *Econometrica*, 67, 1341–1384.

-
- [11] Ané, T., et Geman, H. (2000). Stochastic Subordination. *Risk*, 9, 146 – 149.
- [12] Areal N.M.P.C., et S.J.Taylor. (2000). The realized volatility of FTSE-100 futures prices, *Working paper*. Lancaster University, England LA1 4YX.
- [13] Avouyé-Dovi, S et Jardet, C. (2006). Microstructure des marchés monétaires et financiers. *Revue de la stabilité financière*, N° 9, Décembre 2006.
- [14] Bachelier, L. (1900). Théorie de la spéculation. *Annales de l'École Normale Supérieure*,3.
- [15] Bandi, F.M et Russel, J.R. (2003). Microstructure noise, realized volatility and optimal sampling. *Technical Report*. University of Chicago, Graduate School of Business.
- [16] Bandi, F. M., Russell, J. R. (2005). Microstructure noise, realized volatility, and optimal sampling. Unpublished paper. Graduate School of Business, University of Chicago.
- [17] Barndorff-Nielsen, O.E., et Shephard, N. (2002a). Econometric analysis of realized volatility and its use in stochastic volatility models. *Journal of the Royal Statistical Society B* 64, 253 – 280.
- [18] Bandi, F. M., Russell, J. R. (2006). Separating market microstructure noise from volatility. *Journal of Financial Economics* 79, 655 – 692.
- [19] Barndorff-Nielsen, O.E. et Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17, number 5, 457 – 477.
- [20] Barndorff-Nielsen, O. E., Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1 – 48.
- [21] Barndorff-Nielsen, O. E., et Shephard, N. (2005). How accurate is the asymptotic approximation to the distribution of realised volatility ? In : Andrews, D. W. K., Stock, J. H., eds Identification and Inference for Econometric Models. *A Festschrift for Tom Rothenberg*. Cambridge : Cambridge University Press, pp. 306 – 331.
- [22] Berkes, I. et Horvath, L. (2003a). Asymptotic results for long memory LARCH sequences. *Annals of Applied Probability* 13, 641 – 668.
- [23] Berkes, I. et Horvath, L. (2003b). The rate of consistency of the quasi-maximum likelihood estimator. *Statistics and Probability Letters* 61, 133 – 143.

- [24] Berkes, I., Horv ath, L. et Kokoszka, P. (2003a). Asymptotics for GARCH squared residual correlations. *Econometric Theory* 19, 515 – 540.
- [25] Berkes, I., Horv ath, L. et Kokoszka, P. (2003b). GARCH processes : structure and estimation. *Bernoulli* 9, 201 – 227.
- [26] Billingsley, P. (1995). *Probability and Measure (3 ed)*. John Wiley and Sons.
- [27] Black, F. (1976). Studies of stock price volatility changes. *Proceedings from the American Statistical Association, Business and Economic Statistics*. Section 177 – 181.
- [28] Bodie, Z. et Merton R.(2001). *Finance*. Pearson education.
- [29] Bollen, B., Inder, B. (2002). Estimating daily volatility in financial markets utilizing intraday data. *Journal of Empirical Finance* 9, 551 – 562.
- [30] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307 – 327.
- [31] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69, number 3, 542 – 547.
- [32] Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis* 9, 121 – 131.
- [33] Bollerslev, T. et Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11, number 2, 143–172.
- [34] Bougerol, P., et Picard, N. (1992a). Strict stationarity of generalised autoregressive processes. *Annals of Probability*, 20, 1714–1730.
- [35] Bougerol, P., et Picard, N. (1992b). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115–27.
- [36] Box, G.E.P. et Jenkins, G.M. (1970). *Time Series Analysis : Forecasting and Control*. San Francisco : Holden-Day.
- [37] Brandt, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advanced in Applied Probability*, 18, 211–220.

- [38] Brandt, M.W. et Jones, C.S. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business & Economic Statistics*, 24, number 4, 470 – 486.
- [39] Brockwell, P.J., et Davis, R.A. (1991). *Time Series : Theory and Methods*, 2nd edition. New York : Springer.
- [40] Brock, W., Dechert, W., Scheinkman, J., et LeBaron, B. (1987). “A Test for Independence Based on the Correlation Dimension,”. *Econometric Review*, 3, 197 – 235.
- [41] Brown, R. (1927). A brief account of microscopical observation made in the months of june, july and august, 1827, on the particules contains in the pollen of plants ; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical magazine* 4, pp. 161 – 173.
- [42] Carnero, M. A., Peña, D., Ruiz, E. (2004). Persistence and kurtosis in GARCH and stochastic volatility models. *Journal of Financial Econometrics* 2, 319 – 342.
- [43] Comte, F et Renault E. (2001). Long memory in continuous time stochastic volatility models. *Mathematical Finance*. 8, p. 291-323.
- [44] Christie, A.C. (1982). The stochastic behavior of common stock variances–value, leverage and interest rate effects. *Journal of Financial Economics*, 10, number 4, 407–432.
- [45] Dacorogna, M.M., Gençay, R., Müller, U., Olsen, R.B. and Pictet, O.V. (2001). *An Introduction to High-Frequency Finance*. London : Academic Press.
- [46] Delahaye, J. et Delahaye, F. (2007). *Finance d’entreprise : manuel et application*. Dunod.
- [47] de Vilder, de R.G. et Visser, M.P. (2008). Ranking and combining volatility proxies for GARCH and stochastic volatility models. *MPRA paper no.* 11001.
- [48] Doob, J.L. (1953). *Stochastic Processes*. John Wiley.
- [49] Drost, F.C. et Klaassen, C.A.J. (1997). Efficient estimation in semiparametric GARCH models. *Journal of Econometrics*, 81, 193 – 221.
- [50] Drost, F.C. et Nijman, T.E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, 61, number 4, 909 – 927.
- [51] Ebens, H. (1999). Realized Stock Index Volatility. *Working Paper* No. 420, Department of Economics. Johns Hopkins University, Baltimore.

- [52] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with estimates of variance of U.K. Inflation *Econometrica*, 50, 987–1008.
- [53] Engle, R.F. et Patton, A.J. (2001). What good is a volatility model? *Quantitative Finance*, 1, number 2, 237–245.
- [54] Engle, R.F., et Gallo, G.M. (2006). A multiple indicators model for volatility using intradaily data. *Journal of Econometrics*, 131, number 1 – 2, 2 – 27.
- [55] Fama, E. (1965). The Behavior of Stock Market Prices. *Journal of Business*, 38, 34 – 105.
- [56] Francq, C., et Zakoïan, J-M. (2004). Maximum likelihood estimation of pure *GARCH* and *ARMA-GARCH* processes. *Bernoulli*, 10, number 4, 605–637.
- [57] Francq, C. et Zakoïan, J. M. (2009). *Modèles GARCH : Structure, Inférence Statistique et Applications Financières*. Economica.
- [58] French, K.R., G.W. Schwert., R.F. Stambaugh. (1987). Expected Stock Returns and Volatility, *Journal of Financial Economics*, Vol. 19, pp.3 – 29.
- [59] Garman, M. B., et Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business* 53 :67–78.
- [60] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2006). Predicting volatility : getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131, number 1 – 2, 59 – 95.
- [61] Goldberg, R.R. (1976). *Methods of real analysis*, second edition, Wiley.
- [62] Gonçalves, S., et Meddahi, N. (2005). Bootstrapping Realized Volatility. *Working Paper*, Université de Montréal.
- [63] Gouriéroux, C. Monfort, A. et Trognon, A. (1984). Pseudo Maximum Likelihood Methods : Theory. *Econometrica* 52, 681 – 700.
- [64] Hall, P., et Heyde, C. C. (1980). *Martingale limit theory and its application..* Academic press.
- [65] Hansen, P. R et Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, vol. 24, p. 127-218.
- [66] Hurst, H. E. (1965). Long-term storage : An experimental study. *Constable*. London.

- [67] Jacod, J., et Protter, P. (1998), Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability* 26, 267 – 307.
- [68] Kallsen, J. et Vesenmayer, B. (2009). COGARCH as a continuous-time limit of GARCH(1,1). *Stochastic Processes and Their Applications*, 119, number 1, 74 – 98.
- [69] Karlin, S. et Taylor, H.M. (1975). *A first course in stochastic processes*. Academic press, USA.
- [70] Karatzas, I. et Shreve, S.E. (1991). *Brownian Motion and stochastic calculus*. Springer.
- [71] Kingman, J. C. (1986) : The ergodic theory of subadditive stochastic processes. *Journal of the Royal Statistical Society*, 30, 499–510.
- [72] Klüppelberg, C., Lindner, A. et Maller, R. (2004). A continuous-time GARCH process driven by a Lévy process : Stationarity and second-order behaviour. *Journal of Applied Probability*, 41, number 3, 601 – 622.
- [73] Lee, S.W. et Hansen, B.E. (1994). Asymptotic theory for the GARCH(1, 1) quasi-maximum likelihood estimator. *Econometric Theory*. 10, 29 – 52.
- [74] Lildholdt, P.M. (2002). Estimation of GARCH models based on open, close, high, and low prices. *Centre for Analytical Finance*. University of Aarhus. Working paper series N°. 128.
- [75] Ling, S. et McAleer, M. (2002). Necessary and sufficient moment conditions for the GARCH(r, s) and asymmetric power GARCH(r, s) models. *Econometric Theory* 18, 722 – 729.
- [76] Lumsdaine, R.L. (1995). Finite-sample properties of the maximum likelihood estimator in *Garch*(1,1) and *IGarch*(1,1) models : A monte carlo investigation. *Journal of Business & Economic Statistics*, 13, number 1, 1–10.
- [77] Lumsdaine, R.L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in *IGARCH*(1,1) and covariance stationary *GARCH*(1,1) models. *Econometrica*, 64, number 3, 575–596.
- [78] Maheu, J. M., McCurdy, T. H. (2002). Nonlinear features of realized FX volatility. *Review of Economics and Statistics* 84, 668 – 681.

- [79] Malmsten, H., Teräsvirta, T. (2004). Stylized Facts of Financial Time Series and Three Popular Models of Volatility. Working Paper, *Series in Economics and Finance* 563, *Stockholm School of Economics*.
- [80] Mandelbrot, B.B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394 – 419.
- [81] McAleer, M et Marcelo C. Medeiros. (2008). Realized volatility : a review. *Econometric Reviews*, 27 : 1, 10 — 45.
- [82] Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*. 17, 479 – 508.
- [83] Merton, R.C. (1980). On estimating the expected return on the market : an exploratory investigation. *Journal of Financial Economics*, 8, p. 323-361.
- [84] Nelson, D.B. (1990a). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, 6, number 3, 318 – 334.
- [85] Nelson, D.B. (1990b). ARCH models as diffusion approximations. *Journal of Econometrics*, 45, number 1 – 2, 7 – 38.
- [86] Nelson, D.B. (1992). Filtering and forecasting with misspecified ARCH models I : Getting the right variance with the wrong model. *Journal of Econometrics*, 52,.61 – 90.
- [87] Nielsen, M.Ø., Frederiksen, P. H. (2006). Finite Sample Accuracy and Choice of Sampling Frequency in Integrated Volatility Estimation. *Working Paper*, Cornell University.
- [88] Oksendal, B. (2003). *Stochastic Differential Equations*. Springer.
- [89] Osborne, M.F.M. (1959). Brownian motion in the stock market. *Operations Research*, 7, number 2, 145 – 173.
- [90] Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53 :61–65.
- [91] Peters, R.T. and de Vilder, R.G. (2006). Testing the continuous semimartingale hypothesis for the S&P 500. *Journal of Business & Economic Statistics*, 24, number 4, 444 – 454.

- [92] Portait, R. et Poncet, P. (2009). *Finance de marché : instruments de base, produits dérivés, portefeuilles et risques*. Dalloz.
- [93] Protter, P.E. (2005). *Stochastic integration and differential equations*. Applications of Mathematics, 2nd. edn. Berlin : Springer.
- [94] Rogers, L. C. G., et Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability* 1 :504–12.
- [95] Samuelson, P.A. (1965). Rational theory of warrant pricing. *Industrial Management Review*, 6, 13 – 39.
- [96] Schmalensee, R. and R. Trippi. (1978). Common stock volatility expectations implied by option premia. *Journal of Finance* 33 : 129 – 47.
- [97] Shorack, G., et Wellner, J. (1986). *Empirical Processes with Applications to Statistics*. Chichester, U.K. : Wiley.
- [98] Straumann, D., et T., Mikosch. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series : a stochastic recurrence equations approach. *The Annals of Statistics*, 34, number 5, 2449–2495.
- [99] Taylor, S. J., et Xu, X. (1997). The Incremental Volatility Information in One Million Foreign Exchange Quotations. *Journal of Empirical Finance*, 4, 317 – 340.
- [100] Visser, M.P. (2008). GARCH parameter estimation using high-frequency data. *MPRA paper n°* . 9076.
- [101] Visser, M.P. (2009). *Volatility Proxies and GARCH Models*. PhD thesis, Korteweg-de Vries Institute for Mathematics, Amsterdam.
- [102] Weiss, A. A. (1986). Asymptotic theory for ARCH models : estimation and testing. *Econometric Theory*, 2, 107–131.
- [103] Yang, D. et Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business* 73 :477-491.
- [104] Zhang, L., Mykland, P. A., Aït-Sahalia, Y. (2005). A tale of two time scales : determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association* 100, 1394 – 1411.

-
- [105] Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations : A multi-scale approach. *Bernoulli* 12, 1019 – 1043.